

Universidade Federal do Rio Grande do Sul

Programa de Pós-Graduação em Informática na Educação

Adriana Justin Cerveira Kampff

**Mineração de Dados Educacionais para
Geração de Alertas em Ambientes Virtuais de
Aprendizagem como Apoio à Prática Docente**

Porto Alegre

2009

Universidade Federal do Rio Grande do Sul

Programa de Pós-Graduação em Informática na Educação

Adriana Justin Cerveira Kampff

**Mineração de Dados Educacionais para
Geração de Alertas em Ambientes Virtuais de
Aprendizagem como Apoio à Prática Docente**

Tese apresentada como requisito parcial para
obtenção de grau de Doutor em Informática na
Educação.

Orientador: Prof. Dr. José Valdeni de Lima

Coorientador: Prof. Dr. Eliseo Berni Reategui

Porto Alegre

2009

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Kampff, Adriana Justin Cerveira

Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem como Apoio à Prática Docente. Adriana Justin Cerveira Kampff. – Porto Alegre: PPGIE/UFRGS, 2009.

186 f.:il.

Tese (doutorado) – Universidade Federal do Rio Grande do Sul. Centro Interdisciplinar de Novas Tecnologias na Educação. Programa de Pós-Graduação em Informática na Educação. Porto Alegre, RS - Brasil, 2009. Orientador: José Valdeni de Lima. Coorientador: Eliseo Berni Reategui.

1. Educação a Distância. 2. Mineração de Dados. 3. Sistemas de Alertas. I. Lima, José Valdeni de. II. Reategui, Eliseo Berni. III. Título.



SERVIÇO PÚBLICO FEDERAL
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO

**ATA DA SESSÃO DE DEFESA DE TESE DE DOUTORADO DE
ADRIANA JUSTIN CERVEIRA KAMPPF**

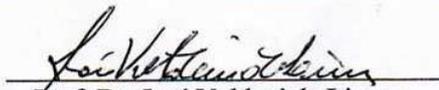
*Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem
como Apoio à Prática Docente*

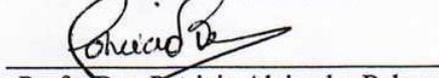
Às nove horas do dia vinte e um de dezembro de dois mil e nove, na sala 323 do PPGIE/CINTED, nesta Universidade, reuniu-se a Comissão de Avaliação, composta pelos Professores Doutores: Patricia Alejandra Behar, Julio Alberto Nitzke e Lucia Maria Martins Giraffa para a análise da Defesa de Tese intitulada "*Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem como Apoio à Prática Docente*" da doutoranda do Programa de Pós-Graduação em Informática na Educação Adriana Justin Cerveira Kampff, sob a orientação do Prof. Dr. José Valdeni de Lima e coorientação do Prof. Dr. Eliseo Berni Reategui.

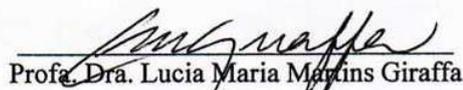
A Banca, reunida, após a apresentação e arguição, emite o parecer abaixo assinalado.

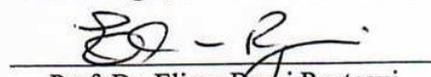
Considera a Tese aprovada.

Porto Alegre, 21 de dezembro de 2009


Prof. Dr. José Valdeni de Lima
Presidente e Orientador


Profa. Dra. Patricia Alejandra Behar
PGIE/UFRGS


Profa. Dra. Lucia Maria Martins Giraffa
PUCRS


Prof. Dr. Eliseo Berni Reategui
Coorientador


Prof. Dr. Julio Alberto Nitzke
UFRGS/ICTA

Agradecimentos

Quero agradecer às pessoas especiais que estão na minha vida, pela compreensão em todos os momentos nos quais foram privadas da minha presença: Darci, meu pai, Zulina, minha mãe, Fabiano, meu irmão, Hélio, meu esposo, Luiza, minha filha, e Vitor, meu filho. Cada um é um pedaço de mim. Obrigada pela paciência e pelo encorajamento. Aos meus grandes e bons amigos, quero dizer que o apoio de cada um, expresso em palavras e atitudes de incentivo, foi imprediscindível.

Não poderia deixar de destacar aqui um amigo especial, Irmão Firmino Caetano Biazus, que permitiu o meu licenciamento do Colégio Marista Rosário, durante nove meses, para gestar a tese, período esse dedicado à construção da proposta de tese, da sua fundamentação até a execução dos primeiros experimentos. Na pessoa dele, estendo o agradecimento aos demais colegas da instituição, que souberam compreender o meu afastamento e fizeram o possível para suprir todas as necessidades que se apresentaram.

À ULBRA quero agradecer o espaço de fomento à busca de alternativas para uma Educação a Distância de qualidade, permitindo o campo de reflexão e de experimentação. A idéia inicial nasceu da minha prática docente na instituição, especialmente na EAD. Esse esforço é reconhecido pela instituição, que apoiou a pesquisa e usufrui de seus resultados. Agradeço por tornar essa pesquisa possível. Quero registrar ainda o meu reconhecimento e o meu respeito a todos os alunos que buscam, por meio do investimento em Educação, construir um mundo melhor, destacando aqueles que, direta ou indiretamente, fizeram parte da pesquisa.

À UFRGS, universidade pública e de qualidade, devo grande parte do conhecimento construído ao longo dos anos de tese, das disciplinas cursadas no PGIE à orientação. Ao meu orientador, José Valdeni de Lima, por apontar caminhos, pela sua experiência e apoio, meu muito obrigada. Ao meu coorientador, Eliseo Berni Reategui, verdadeiro parceiro, por toda a atenção e dedicação ao meu trabalho, toda a minha gratidão.

Sumário

Lista de Abreviaturas	9
Lista de Figuras	10
Lista de Quadros	11
Lista de Tabelas	12
Resumo	13
Abstract.....	14
1. Introdução.....	15
1.1 <i>Problema de Pesquisa</i>	17
1.2 <i>Hipóteses</i>	17
1.3 <i>Objetivos</i>	18
1.4 <i>Estrutura do Documento</i>	19
2. Educação a Distância.....	21
2.1 <i>Referenciais para a EAD</i>	21
2.2 <i>Distância Transacional</i>	26
2.3 <i>Papéis Docentes em EAD</i>	28
2.3.1 <i>Autor</i>	29
2.3.2 <i>Especialista</i>	30
2.3.3 <i>Tutor</i>	31
2.3.4 <i>Acompanhamento de Cursos de Formação Docente para a EAD</i>	31
2.3.5 <i>Considerações</i>	32
2.4 <i>Mediação</i>	33
2.5 <i>Acompanhamento</i>	38
2.6 <i>Evasão e Desempenho</i>	39
2.6.1 <i>A Pesquisa sobre Evasão na Open University</i>	41
2.6.2 <i>Questões Externas e Internas à Instituição</i>	44
2.7 <i>Ambientes Virtuais de Aprendizagem</i>	45
3. Mineração de Dados.....	53
3.1 <i>Etapas da Mineração de Dados</i>	55
3.2 <i>Tipos de Tarefas de Mineração de Dados</i>	59
3.3 <i>Técnicas de Mineração de Dados</i>	60
3.3.1 <i>Árvore de Decisão</i>	60
3.3.2 <i>Regras de Classificação</i>	64
3.3.3 <i>Regras de Associação</i>	66
3.3.4 <i>Clusterização</i>	69
3.4 <i>Ferramentas de Mineração</i>	73
3.4.1 <i>RapidMiner</i>	74
3.4.2 <i>WizRule</i>	76

3.5 Aplicações de Mineração de Dados Educacionais	78
3.6 Sistema de Alertas.....	84
3.6.1 Mecanismos de Percepção (Awareness).....	86
3.6.2 Mecanismos de Alertas.....	88
4. Sistema de Alertas: Apoio à Prática Docente	92
4.1 Alertas como Apoio ao Acompanhamento e Mediação Docente	93
4.2 Arquitetura para Sistemas de Alertas em AVA	94
4.2.1 Interface do Sistema de Alertas	95
4.2.2 Gerenciador de Alertas Fixos.....	96
4.2.3 Gerenciador de Alertas de Padrões.....	100
4.2.4 Considerações sobre a Arquitetura.....	103
4.3 Contextualizando o Espaço de Realização dos Experimentos	103
4.4 Experimento Preliminar	106
4.5 Processo de Descoberta de Conhecimento.....	109
4.5.1 Definição da Amostra.....	110
4.5.2 Organização dos Dados	111
4.5.3 Mineração de Dados.....	117
4.6 Geração de Alertas.....	119
4.6.1 Turmas Acompanhadas.....	119
4.6.2 Alertas Gerados e Exemplos de Intervenções	120
4.7 Resultados.....	128
4.7.1 Comparação entre os Índices de Aprovação	129
4.7.2 Comparação entre os Índices de Reprovação	129
4.7.3 Comparação entre os Índices de Evasão	130
4.8 Pesquisa com os Alunos	130
5. Considerações Finais	138
6. Bibliografia.....	143
ANEXO 1	152
ANEXO 2	153
ANEXO 3	154
ANEXO 4	155
ANEXO 5	162
ANEXO 6	164
ANEXO 7	166
ANEXO 8	168
ANEXO 9	171
ANEXO 10.....	174

ANEXO 11..... 176
ANEXO 12..... 178
ANEXO 13..... 179
ANEXO 14..... 181
ANEXO 15..... 183
ANEXO 16..... 184

Lista de Abreviaturas

ARFF	<i>Attribute-Relation File Format</i>
AVA	Ambiente Virtual de Aprendizagem
BD	Banco de Dados
CSV	<i>Comma-Separated Values</i>
CVA	Comunidade Virtual de Aprendizagem
DT	Distância Transacional
DW	<i>Data Warehouse</i>
EAD	Educação a Distância
ETL	<i>Extract Transform Load</i>
GPL	<i>General Public License</i>
HTML	<i>Hypertext Markup Language</i>
IC	Instrumentalização Científica
IA	Inteligência Artificial
KDD	<i>Knowledge Discoverey in Databases</i>
LMS	<i>Learning Management Systems</i>
MD	Mineração de Dados
MEC	Ministério da Educação
OLAP	<i>On-line Analytical Process</i>
SAD	Sistema de Apoio à Decisão
SQL	<i>Structured Query Language</i>
SMA	Sistema Multiagente
TI	Tecnologia da Informação
TIC	Tecnologias da Informação e da Comunicação

Lista de Figuras

Figura 3.1 – Mineração de Dados: intersecção de várias áreas (adaptada de Silva, 2004) .	54
Figura 3.2 – Etapas para Descoberta de Conhecimento (adaptada de Fayyad et al., 1996)	56
Figura 3.3 – Árvore de Decisão gerada a partir dos dados da tabela 3.1	63
Figura 3.4 – Clusters gerados a partir dos dados da tabela 3.1	70
Figura 3.5 – Grafo de distância nominal gerado a partir dos dados da tabela 3.1	72
Figura 3.6 – RapidMiner: exemplo de montagem da árvore de operadores para MD	75
Figura 3.7 – RapidMiner: exemplo de visualização de resultados da MD	75
Figura 3.8 – WizRule: configuração dos parâmetros para geração das regras	76
Figura 3.9 – WizRule: exemplo de apresentação das regras geradas	77
Figura 4.1 – Arquitetura para Sistemas de Alertas em AVA	95
Figura 4.2 – Interface do Sistema Gerenciador de Alertas: recebimento dos alertas	96
Figura 4.3 – Exemplo de Configuração dos Alertas: Biblioteca Virtual	97
Figura 4.4 – Exemplo de Configuração dos Alertas: Fórum	98
Figura 4.5 – Exemplo de Configuração dos Alertas: Avaliação	99
Figura 4.6 – Exemplo de Configuração dos Alertas: Desempenho	99
Figura 4.7 – Sistema de Alertas: classificação	101
Figura 4.8 – Sistema de Alertas: envio de <i>e-mail</i>	102
Figura 4.9 – Sistema de Alertas: visualização de regras	102
Figura 4.10 – RapidMiner: visualização dos metadados	107
Figura 4.11 – Visualização Gráfica da Árvore de Decisão Gerada	108
Figura 4.12 – Visualização Textual da Árvore de Decisão Gerada.....	108
Figura 4.13 – SEstatNet: cálculo do tamanho da amostra	111
Figura 4.14 – Gráfico: questões relacionadas à dimensão estrutura	131
Figura 4.15 – Gráfico: Alunos que liam as mensagens enviadas pelo professor	133
Figura 4.16 – Gráfico: questões relacionadas aos alertas	134
Figura 4.17 – Gráfico: Alunos que cursariam outras disciplinas ou cursos a distância	137

Lista de Quadros

Quadro 3.1 – Dados em formato CSV gerados a partir da tabela 3.1	62
Quadro 3.2 – Regras de Decisão obtidas a partir da Árvore de Decisão da figura 3.3	64
Quadro 3.3 – Regras de Classificação geradas a partir dos dados da tabela 3.1	65
Quadro 4.1 – Valores do Campo Curso, com a respectiva quantidade de alunos	117
Quadro 4.2 – Exemplos de Regras Mês 1	123
Quadro 4.3 – Exemplo de mensagem enviada aos alunos após mês 2	125
Quadro 4.4 – Exemplo de mensagem enviada aos alunos após mês 3	126
Quadro 4.5 – Exemplo de mensagem enviada aos alunos após mês 4	127
Quadro 4.6 – Teste de Hipóteses: comparação entre os percentuais de aprovação	129
Quadro 4.7 – Teste de Hipóteses: comparação entre os percentuais de reprovação	129
Quadro 4.8 – Teste de Hipóteses: comparação entre os percentuais de evasão	130

Lista de Tabelas

Tabela 2.1 – Distância Transacional (adaptada a partir de Peters, 2001, p. 64)	27
Tabela 3.1 – Exemplo de Dados para Mineração	62
Tabela 3.2 – Tabela de distância nominal gerada a partir dos dados da tabela 3.1	72
Tabela 4.1 – Dados preparados para a Mineração	115
Tabela 4.2 – Mineração de Dados: atributos considerados e regras geradas	118
Tabela 4.3 – População 1 (Dados Históricos) x População 2 (Dados de Acompanhamento)	120
Tabela 4.4 – Alertas fixos: atividades avaliativas pendentes	122
Tabela 4.5 – Classificação Mês 2	125
Tabela 4.6 – Classificação Mês 3	126
Tabela 4.7 – Classificação Mês 4	127
Tabela 4.8 – População 1 (Dados Históricos) x População 2 (Dados Acompanhados)	128
Tabela 4.9 – Pesquisa com os Alunos	132

Resumo

A Educação a Distância (EAD) no país apresenta-se em franca expansão. Cresce, a cada dia, o número de alunos que estudam por meio dessa modalidade. Gerenciar seus processos de aprendizagem em Ambientes Virtuais de Aprendizagem (AVA), com qualidade de interação e de acompanhamento, exige cada vez mais do professor. Esta tese descreve uma pesquisa que identifica, por meio de Mineração de Dados (MD) gerados pela interação em AVA, comportamentos e características de alunos com risco de evasão ou reprovação e, então, alerta o professor. Tais alertas são gerados a partir de agrupamentos de alunos com características similares, para que o professor possa estabelecer comunicação personalizada e contextualizada com esses sujeitos. O trabalho propõe uma arquitetura para sistemas de alertas em AVA, com alertas pré-definidos e outros gerados a partir da mineração de dados. Para validação da arquitetura, foram utilizados dados de 1564 alunos, de edições anteriores de uma mesma disciplina a distância, para mineração e extração de regras de classificação. As regras foram aplicadas para gerar os alertas durante o acompanhamento de 230 alunos em turmas em andamento. De cada aluno, cerca de 230 atributos foram analisados. Ao final, foi possível comprovar que as intervenções realizadas pelo professor, a partir dos alertas, direcionadas a grupos que compartilhavam necessidades específicas, contribuíram para a melhoria dos índices de aprovação e para redução dos índices de evasão dos alunos na disciplina acompanhada.

Abstract

Distance Education is increasing very fast in Brazil. It grows every day the number of students engaged in this educacional mode. The professor must manage the teaching and the learning processes in Learning Management System (LMS), with quality of monitoring and interaction. This thesis describes a study that identifies, by Data Mining (DM) of educacional data generated by the interaction at LMS, profiles of students in risk of dropout or fail, and then alerts the professor. These alerts are based on groups of students with similar characteristics, so the professor can communicate with them in a personalized and contextualized way. The research proposes an architecture for warning systems in LMS, with pre-defined alerts and other generated from data mining. To validate the architecture, it was used data from 1564 students from previous editions of the same distance course, for data mining and extraction of classification rules. The rules were applied to generate alerts during the monitoring of 230 students in classes in progress. From each student, about 230 attributes were analyzed. In the end, it was possible to demonstrate that the interventions by the professor, based on the alerts, to groups that shared needs, contribute to higher rates of sucess and to reduce the dropout rates of students in the course.

1. Introdução

A Educação a Distância (EAD) no Brasil apresenta-se em franca expansão. Em 2007, mais de 2,2 milhões de brasileiros realizaram algum curso na modalidade a distância, de acordo com o Anuário Brasileiro Estatístico de Educação Aberta e a Distância (ABED, 2008). Em 2008, esse número cresceu para mais 2,5 milhões de alunos matriculados em 1.181 cursos, disponíveis em 257 instituições credenciadas (ABED, 2009).

No Ensino Superior, cursos de extensão, de graduação e de pós-graduação *lato sensu* são ofertados de forma integral na modalidade EAD. Hoje, há mais de 800 mil alunos brasileiros cursando graduação nesta modalidade (NISKIER, 2009). Também em cursos presenciais de graduação até 20% da carga horária total pode ser realizada a distância, conforme legislação vigente (MEC/Portaria n. 4.059, de 10 de dezembro de 2004.).

Segundo Carmo (1997), a significativa ampliação do número de alunos que buscam a EAD deve-se à combinação de três principais fatores: *“aumento demográfico da população jovem, especialmente nos países menos desenvolvidos; aumento das necessidades de formação continuada da população adulta; e crescente consciência da importância do nível de educação da população para o desenvolvimento econômico e social”*. Infelizmente, os sistemas educacionais não estão preparados para atender qualificadamente a toda essa demanda quantitativa e diversificada.

Ou seja, ao mesmo tempo em que a demanda é crescente por cursos em EAD, gerada pela necessidade de formação e democratização da educação, os índices de evasão e reprovação nesta modalidade de Educação mostram-se, ainda, elevados, de acordo com as pesquisas discutidas na seção 2.6. O termo evasão diz respeito aos alunos desistentes, que abandonam o curso antes de sua conclusão, não completando o percurso didático estabelecido. Já a reprovação refere-se aos alunos que não obtêm aproveitamento no curso, apesar de realizá-lo até o final, de acordo com os critérios definidos pela instituição de ensino à qual estão vinculados. Embora a idéia de que cursar uma graduação a distância, total ou parcialmente, traz

flexibilidade de tempo e local de estudo para o aluno, estudos evidenciam a dificuldade do aluno em manter uma rotina de estudos que possibilite seus avanços.

Com o crescimento do número de alunos, gerenciar seus processos de aprendizagem em Ambientes Virtuais de Aprendizagem (AVA), com qualidade de interação e de acompanhamento, visando à permanência e ao êxito desses sujeitos nos cursos que realizam, exige cada vez mais dos professores. Os dados gerados nas interações entre professores e alunos, dos alunos entre si e deles com os recursos disponibilizados são volumosos e pouco explorados: reuni-los e interpretá-los é uma atividade complexa e exaustiva.

Para acompanhar um número cada vez maior de alunos de forma adequada, é preciso lidar com um volume enorme de informações (de acesso ao ambiente, de desempenho e de comportamento) disponíveis em *logs* de sistema e relatórios variados. Analisar os diversos dados, gerados pelos alunos, costuma ser um processo demorado e complexo. A área de Mineração de Dados pode contribuir na descoberta de conhecimento potencialmente útil, por meio de algoritmos e ferramentas que processam esses dados de forma a buscar correlações importantes entre eles (FAYYAD et al., 1996).

Frente a este cenário, esta pesquisa busca, por meio de Mineração de Dados educacionais, identificar perfis de alunos com risco de evasão ou reprovação, visando à geração de alertas para a conscientização do professor. Tais alertas são gerados a partir de relações identificadas pelos processos de mineração de dados e são direcionados a um ou mais alunos com necessidades similares, para que o professor possa dirigir uma comunicação ou atividade personalizada e contextualizada para esses sujeitos, tornando sua mediação mais eficaz e colaborando para reduzir os índices de insucesso em cursos pela Web.

O foco desse trabalho é, portanto, a geração de alertas que auxiliem o professor na interação e no acompanhamento dos seus alunos por meio de AVA. Neste contexto, é preciso capacitar o professor para atuar nesta modalidade, pesquisar práticas pedagógicas adequadas para a EAD, analisar perfis de alunos e modelar ferramentas computacionais que permitam dar suporte a uma atuação

docente qualificada, tornando mais eficaz a comunicação com os estudantes e, conseqüentemente, potencializando o engajamento e a qualidade das aprendizagens desses sujeitos.

O mapeamento de práticas docentes eficazes em EAD, a realização de agrupamentos flexíveis e dinâmicos de alunos que compartilham necessidades específicas para intervenção do professor e a sensibilização deste por meio de alertas busca a diminuição da sobrecarga dos processos comunicativos e a ampliação da satisfação do professor com o acompanhamento que realiza.

1.1 Problema de Pesquisa

Considerando demandas cada vez maiores de Educação a Distância, que elevam o número de alunos por turma, e mantendo a perspectiva de mediação de qualidade, que contribua para o engajamento e a aprendizagem dos sujeitos, coloca-se o problema que motiva esta tese:

- Como gerenciar alertas, baseados em Mineração de Dados, em Ambientes Virtuais de Aprendizagem, para apoiar práticas docentes em Educação a Distância, de forma a tornar mais eficiente o acompanhamento que o professor realiza e sua comunicação com os alunos, contribuindo para a redução dos índices de evasão e reprovação?

Assim, busca-se propor mecanismos que possam dar suporte ao professor: na identificação de situações das quais deve estar ciente e/ou nas quais deve intervir, e na execução de ações pertinentes.

1.2 Hipóteses

Hoje, muitos professores, especialmente no Ensino Superior, não são adequadamente preparados para a atividade docente (não possuem formação pedagógica) e também não possuem conhecimento suficiente das ferramentas utilizadas para mediar atividades em EAD. Além disso, mesmo os professores com bons conhecimentos didáticos e capacitados nas ferramentas tecnológicas encontram dificuldades operacionais para acompanhar e mediar os processos de

aprendizagem dos alunos nos ambientes atuais, o que acaba por sobrecarregá-los com análises contínuas de dados, em um ir-e-vir entre ferramentas diferentes, que poderiam estar integradas e gerar alertas automatizados para a intervenção ou conscientização dos professores.

Esses alertas facilitariam o trabalho de acompanhamento realizado pelos professores, gerando comunicações mais voltadas às necessidades dos alunos. Por meio dos alertas seria ampliada também a satisfação dos alunos em relação ao acompanhamento que recebem de seus professores.

A partir disto, apresentam-se as hipóteses:

- *A mineração de dados educacionais contribui para a descoberta de características e comportamentos dos alunos que podem servir de base para a mediação do professor.*
- *A geração de alertas, com base em processos de mineração de dados, permite notificar o professor sobre grupos de alunos com tendência à reprovação e à evasão em EAD.*
- *A intervenção do professor, a partir dos alertas gerados na mineração de dados, contribui para a redução dos índices de evasão e reprovação e, conseqüentemente, para a ampliação do índice de aprovação.*

1.3 Objetivos

O objetivo geral da tese é propor uma arquitetura para sistemas de alertas em AVA, baseado em informações extraídas por processos de mineração de dados, buscando identificar estudantes com características e comportamentos que podem levar à evasão ou à reprovação. Um sistema de alertas em AVA potencializa a prática docente, em busca do engajamento do aluno. Para alcançar esse objetivo, apresentam-se os seguintes objetivos específicos:

- Estudar e aplicar técnicas de Mineração de Dados que identifiquem perfis de alunos com características de evasão ou reprovação, para intervenção preventiva do professor;

- Definir uma arquitetura para a construção de sistemas de alertas que auxiliem o professor na mediação em AVA com número expressivo de alunos – esses alertas devem agrupar alunos que compartilham necessidades similares;
- Integrar a mineração de dados com a geração dos alertas, aplicando e avaliando os resultados.

1.4 Estrutura do Documento

A tese apresenta, no capítulo 2, o referencial educacional, destacando a importância da mediação em EAD, de forma a estabelecer uma *proximidade comunicativa* entre os sujeitos envolvidos em cursos via Web. Faz-se necessário estabelecer uma *distância transacional* adequada, a partir do equilíbrio entre as dimensões de *estrutura*, *diálogo* e *autonomia*. Para tanto, é fundamental conhecer e dinamizar, de modo coerente com esta concepção, as diversas ferramentas que compõem um AVA. Além disso, descreve estudos sobre evasão e reprovação em EAD, apontando elementos que contribuem para a retenção do aluno que estuda nesta modalidade.

O capítulo 3 apresenta o referencial de Mineração de Dados (MD), destacando a motivação para seu uso no contexto desta pesquisa. As fases do processo de descoberta de conhecimento, da escolha dos dados à análise dos resultados, são apresentadas, destacando seu processo iterativo, que propõe retornar a fases anteriores em busca de melhores resultados. Técnicas para MD e trabalhos na área educacional contribuem para as definições do trabalho. Além disso, o capítulo traz a descrição de sistemas similares, que utilizam em algum grau mineração, mecanismos de percepção ou de alertas na área da educação.

No capítulo 4, descreve-se o trabalho de pesquisa realizado. Parte-se da proposição de uma arquitetura para geração dos alertas. Apresenta-se o contexto em que o trabalho é desenvolvido, estabelecendo as características da instituição e da disciplina em que o estudo é realizado. São apresentadas as escolhas das técnicas e ferramentas de Mineração de Dados empregadas. Além do experimento

inicial, criado para a validação da proposta, o estudo realizado para a tese é detalhado e os resultados são analisados.

Por fim, estão colocadas as considerações finais sobre a pesquisa realizada, as contribuições da tese e as indicações de trabalhos futuros.

2. Educação a Distância

Este capítulo, inicialmente, busca conceituar Educação a Distância (EAD), apresentando suas principais características e referenciais de qualidade. Os diversos papéis docentes em EAD, da constituição do material didático à mediação, também são apresentados, buscando discutir ações docentes que possam contribuir para estabelecer uma *distância transacional* adequada.

O conceito de *distância transacional* (PETERS, 2001) trata da distância comunicativa ou psicológica estabelecida na relação pedagógica, transcendendo a questão da distância geográfica posta em EAD. A *distância transacional* é analisada a partir da combinação das dimensões *estrutura* (do curso), *diálogo* (entre os participantes) e *autonomia* (do aluno). A combinação dessas variáveis depende de cada curso, de seus objetivos e público-alvo. A mediação e o acompanhamento dos alunos são questões determinantes para reduzir a evasão e a reprovação.

Em cursos a distância, especificamente via Web, as diversas ferramentas que compõe um AVA são fundamentais, como espaços pedagógicos. Uma análise criteriosa dos dados gerados nas interações realizadas nestes espaços virtuais pode contribuir para a que o professor acompanhe e gerencie seus alunos, com o intuito de engajá-los em suas atividades de aprendizagem e de que obtenham êxito nos estudos realizados.

2.1 Referenciais para a EAD

A Educação a Distância (EAD) refere-se à modalidade de ensino e aprendizagem em que os sujeitos envolvidos (professores e alunos), na maior parte do tempo, não estão reunidos em um mesmo local ou em um mesmo horário. Segundo Relatório do MEC (2002), são considerados cursos a distância aqueles em que “*pelo menos, 80% da carga horária correspondente às disciplinas curriculares não seja integralmente ofertada em atividades com frequência obrigatória de professores e alunos*”. Também em cursos presenciais de graduação 20% de suas disciplinas podem ser realizadas a distância (MEC/Portaria n. 4.059, de 10 de

dezembro de 2004.). Além da distância física e/ou temporal entre professores e alunos, a EAD caracteriza-se por ser mediada por recursos tecnológicos.

Educação a Distância é uma forma de ensino que possibilita a auto-aprendizagem, com a mediação de recursos didáticos sistematicamente organizados, apresentados em diferentes suportes de informação, utilizados isoladamente ou combinados, e veiculados pelos diversos meios de comunicação. (MEC/Decreto n. 2.494, de 10 de fevereiro de 1998.)

Entre as características das atividades educativas em EAD estão (KEEGAN, 1996, p. 50): a separação entre professores e alunos durante o processo de ensino e aprendizagem, constante ou na maior parte do tempo; a existência de uma organização educacional, tanto no planejamento e elaboração dos materiais quanto nos serviços de suporte aos estudantes, que diferencia a EAD de um estudo privado ou autodidata; o uso de várias mídias para unir professores e alunos e disponibilizar os conteúdos educacionais; a possibilidade de comunicação de via dupla entre professores e alunos, de modo que o estudante possa participar ou mesmo iniciar o diálogo; a possibilidade de encontros ocasionais, presenciais ou por meios eletrônicos, para propósitos didáticos ou de socialização.

O princípio fundamental que tem motivado a Educação a Distância tem sido o de ampliar as oportunidades educativas aos indivíduos, independentemente de qualquer limitação geográfica ou socioeconômica (UNESCO, 2002). A EAD – por apresentar-se como modalidade que permite levar a educação a locais distantes, flexibilizar os horários de estudos e envolver um número maior de sujeitos – é uma realidade que tem sido aceita e é estimulada por órgãos governamentais (COSTA, 2007), além de amplamente explorada pela iniciativa privada.

Moore & Kearsley (2007, p. 8) apresentam, de forma não exaustiva, uma lista de necessidades consideradas por responsáveis por políticas institucionais e governamentais ao ofertar cursos em EAD:

- acesso crescente a oportunidades de aprendizado e treinamento;
- proporcionar oportunidades para *atualizar aptidões*;
- melhorar a *redução de custos* dos recursos educacionais;
- apoiar a *qualidade* das estruturas educacionais existentes;
- melhorar a *capacitação* do sistema educacional;
- *nivelar desigualdades* entre grupos etários;
- direcionar campanhas educacionais para *públicos-alvo* específicos;
- proporcionar treinamento de emergência para *grupos-alvo importantes*;

- aumentar as aptidões para a educação em *novas áreas de conhecimento*;
- oferecer uma combinação de educação com *trabalho e vida familiar*;
- *agregar uma dimensão internacional* à experiência educacional.

Diante de tantas motivações, é preciso criar meios para desenvolver cursos em EAD que cumpram o objetivo de democratizar a educação. As estratégias e metodologias de ensino e aprendizagem empregadas dependem, em parte, do tipo de programa e da formação que se pretende realizar e, por outro lado, das características e potenciais das tecnologias utilizadas.

Moore & Kearsley (2007), ressaltam que, em EAD, o ensino e a aprendizagem são intencionais, ou seja, o professor cria meios para promover e apoiar a aprendizagem de quem se propõe a aprender. Destacando a “*natureza multidimensional dessa área*”, os autores adotam a seguinte definição:

Educação a distância é o aprendizado planejado que ocorre normalmente em um lugar diferente do local do ensino, exigindo técnicas especiais de criação de cursos e de instrução, comunicação por meio de várias tecnologias e disposições organizacionais e administrativas especiais. (MOORE & KEARSLEY, 2007, p. 2)

As primeiras experiências com EAD remontam ao século XIX, com sistemas de educação por correspondência, que ainda hoje são formas utilizadas em países menos desenvolvidos (UNESCO, 2002). A educação por correspondência também era chamada de *estudo em casa* pelas primeiras escolas com fins lucrativos ou de *estudo independente* pelas universidades (MOORE & KEARSLEY, 2007). Tais sistemas baseavam-se, em grande parte, em textos impressos, que serviam de guias de estudos. A interação, realizada por meio de documentos enviados pelo correio, era bastante limitada e a auto-aprendizagem era, portanto, a base do aprendizado (PALHARES, 2009).

Com o advento de mídias como o rádio e a televisão, surgiram programas educativos que transmitiam aulas gravadas ou ao vivo, destinadas a estudantes individuais ou a grupos que contavam com algum apoio presencial (TORRES, 2003). As possibilidades de interação durante as transmissões eram muito limitadas, necessitando do suporte de outros meios, tais como telefone, carta ou fax. O êxito na realização de cursos assim ofertados estava essencialmente nas mãos dos

sujeitos que os realizavam, em geral sem um acompanhamento maior das instituições geradoras dos mesmos. No Brasil, esses meios vêm deixando de ser utilizados para transmitir conteúdos ligados à educação formal, focando na disseminação de temas transversais, tais como cidadania e meio ambiente (BIANCO, 2009).

Com a popularização das Tecnologias da Informação e da Comunicação (TIC), surgem os cursos a distância por meio da Web, em diversos níveis de ensino, oportunizados por uma infra-estrutura com custos decrescentes, enquanto cresce potencialmente a qualidade e a abrangência geográfica, atendendo a diversos segmentos sociais e regiões antes excluídas. Além da convergência das mídias utilizadas até então, amplia-se a possibilidade de interação. Com as novas ferramentas da Internet, um número maior de indivíduos pode compartilhar suas aprendizagens em tempo real, bem como é possível a cada um manter uma interação pessoal com o professor ou com outro estudante, independentemente do local onde esteja (UNESCO, 2002).

É importante ressaltar, acompanhando a evolução dos recursos tecnológicos, que, ao desenvolver cursos a distância, *“o valor da proposta, mesmo quando adota os últimos desenvolvimentos da tecnologia, continua residindo, como em qualquer projeto educacional, na qualidade dos conteúdos e em suas propostas para o ensino”* (LITWIN, 2001, pág. 10).

Costa (2007) destaca que a *“educação a distância de qualidade tem uma complexidade que não deve ser subestimada. Muitas variáveis estão envolvidas e daí a necessidade de uma abordagem sistêmica”*. Para atender as crescentes demandas sem deixar de lado a qualidade, o MEC (2007) criou referenciais para a qualidade do ensino superior a distância, que englobam, fundamentalmente, questões pedagógicas, de infra-estrutura e de pessoal capacitado.

Os elementos que o MEC (2007) aponta devem ser considerados e estar descritos no Projeto Político Pedagógico de um curso a distância. São eles:

- *Concepção de educação e currículo no processo de ensino e aprendizagem*: apresentar uma definição clara da opção epistemológica, que deverá nortear o desenvolvimento e execução do curso;
- *Sistemas de Comunicação*: garantir a interação dos sujeitos nos processos de ensino e aprendizagem, com fins de informação, orientação, resolução de dúvidas e desenvolvimento de projetos compartilhados;
- *Material didático*: deve ser elaborado em conformidade com as necessidades do curso e público-alvo, recomendando-se a integração de diferentes mídias, produzidas por equipes interdisciplinares;
- *Avaliação*: engloba a avaliação da aprendizagem – processo contínuo de acompanhamento, com atividades a distância e presenciais – e a avaliação institucional – abrangendo a organização didático-pedagógica, instalações físicas e os corpos docente, de tutores, técnico-administrativo e discente;
- *Equipe multidisciplinar*: na EAD, a equipe multidisciplinar é necessária para planejar, implementar e gerenciar os cursos, composta por docentes, tutores presenciais e/ou a distância e técnicos-administrativos, que devem buscar atualização constante, de acordo com as demandas da modalidade;
- *Infra-estrutura de apoio*: diz respeito tanto à infra-estrutura na sede da Universidade, com espaços e recursos adequados ao trabalho das diversas equipes, quanto à infra-estrutura dos pólos distantes, com salas de trabalho, laboratórios, biblioteca, apoio de secretaria, entre outros;
- *Gestão acadêmico-administrativa*: o aluno distante deve ter as mesmas condições de acesso aos processos de inscrição, matrícula, protocolos e tesouraria, por exemplo, que o aluno presencial. Além disso, questões como a logística de produção e distribuição de material didático, bem como o acompanhamento e avaliação do aluno devem ser bem gerenciadas;
- *Sustentabilidade financeira*: na EAD, há investimentos iniciais elevados, referentes à capacitação das equipes para a modalidade, à produção de materiais didáticos, à implantação de pólos, aos recursos computacionais necessários, entre outros – é preciso garantir esse investimento inicial,

apontando metas de ingresso de alunos, para retorno dos investimentos em médio prazo.

Essa complexidade toda exige que se considere a proposta de organizações curriculares inovadoras, tais como o trabalho por eixos e temas integradores, gerenciados por equipes interdisciplinares (NEVADO et al., 2007); a produção de material didático interativo em diversas mídias, conforme as necessidades do curso e do público-alvo (BELISÁRIO, 2003); e, principalmente, a definição de uma infraestrutura de comunicação e mediação pedagógica adequada à *distância transacional* pretendida (PETERS, 2001).

Infelizmente, muitas instituições, a fim de abarcar este mercado educacional em expansão, lançam-se imaturamente na adoção de ferramentas para produção e gestão de cursos, passando a disponibilizar conteúdos e supondo que com o simples acesso a tais materiais o aluno estaria capacitado nos assuntos colocados. Se em situações de educação presencial o repasse de informações num modelo tradicional, centrado no ensino, na fala e no material do professor (LÉVY, 1999), já se mostra insuficiente para atender às necessidades atuais de formação, na EAD a situação é ainda mais complexa. A utilização de AVA como recurso tecnológico para a EAD via Web oportuniza que se rompam paradigmas tradicionais transmissivos (comunicação unidirecional), destacando o potencial interacionista (comunicação multidirecional) das novas tecnologias (SILVA, 2001).

Não basta democratizar o acesso à educação: é preciso garantir a qualidade da educação ofertada. Professores capacitados, planejamentos cuidadosos, materiais fartos e interativos, infra-estrutura acadêmica e administrativa bem gerenciada, alunos autônomos e motivados são elementos importantes para o sucesso na modalidade de EAD.

2.2 Distância Transacional

Para promover espaços de aprendizagem a distância, é preciso conceber propostas que propiciem uma adequada *distância transacional* (DT). O conceito da *distância transacional*, discutido por Moore (1993, p. 22), busca diferenciar a

distância física de alunos e professores (geográfica) da *distância comunicativa* (pedagógica / psicológica).

A *distância transacional* é dada pela combinação de três variáveis: a *autonomia* do aluno em buscar informações e estudar, estando motivado para tal; a *estrutura* do curso que orienta o percurso didático e oferece material adequado; e a possibilidade de *diálogo*, que aproxima, estabelece empatia, permite esclarecer dúvidas em prazos curtos de tempo, reorienta e cria novas possibilidades que podem impactar na reestruturação do curso e na construção da autonomia do aluno.

Em EAD, a distância física está posta, mas a distância comunicativa costuma variar muito, sendo baixa quando há interação intensa entre professores e alunos e alta quando os alunos desenvolvem estudos individuais sem oportunidade de esclarecer suas dúvidas. A tabela 2.1, adaptada a partir de Peters (2001, p. 64), traz representações extremas que expressam o foco em uma maior ou menor distância transacional.

Tabela 2.1 – Distância Transacional (adaptada a partir de Peters, 2001, p. 64)

Distância Transacional	Tipo	Exemplo
maior	Programa de ensino sem diálogo e sem estrutura	Estudo independente com base em leitura própria
	Programa de ensino sem diálogo, mas com estrutura	Programas didáticos no rádio e na televisão
	Programa de ensino com diálogo e com estrutura	Curso de estudo a distância típico
menor	Programa de ensino com diálogo, mas sem estrutura	Assistência tutorial

A combinação ideal das variáveis *autonomia*, *estrutura* e *diálogo* depende da natureza da atividade educativa proposta, como, por exemplo, um treinamento pontual ou uma formação mais ampla. O equilíbrio entre as variáveis *estrutura* e *diálogo*, no contexto de um dado curso, deve ser objetivo de todo o planejamento e acompanhamento, bem como a observação do grau de *autonomia* desejado nos alunos.

(...) distância transacional é uma função de três grandezas, que mudam de uma situação para outra, em parte inclusive são antagônicas ou até mesmo excludentes. Conforme as características das pessoas participantes, os objetivos e conteúdos do ensino, seu nível de exigências, os métodos necessários, os meios disponíveis e a cultura de estudo tradicional, é preciso encontrar a melhor relação possível dos três grupos de variáveis entre si e determinar, assim, uma distância transacional que se adapte exatamente a essa situação. Por conseguinte, a didática do ensino a distância consistiria daquelas estratégias e técnicas com que é possível estabelecer a distância transacional de acordo com as exigências de cada uma das situações educativas e de aprendizagem. (PETERS, 2001, p. 65)

A busca por uma *distância transacional* adequada passa, portanto, pela definição da estrutura do curso – recursos tecnológicos utilizados, produção de materiais interativos e desenvolvimento de formas de avaliação diversificadas –, pelas formas de interação previstas e pelo acompanhamento dos alunos, ajudando-os a construir suas autonomias. Os cursos devem ser organizados a partir dos conteúdos e das intervenções didáticas, que, em conjunto, apoiem a aprendizagem dos alunos.

Uma *distância transacional* inadequada ao tipo de curso ou público-alvo envolvido pode contribuir para um crescimento dos índices de evasão ou reprovação. Professores que atuam em EAD costumam acompanhar grupos grandes de alunos e precisam, mantendo a qualidade da mediação preconizada, de instrumentos que os auxiliem na gestão adequada desses grupos, possibilitando intervenções eficazes.

2.3 Papéis Docentes em EAD

A EAD, compreendida em uma visão sistêmica, da administração e produção dos cursos ao acompanhamento dos alunos, requer uma equipe de profissionais cada vez mais preparados para exercer as funções necessárias. Diante das inovações tecnológicas incorporadas na EAD, de novas demandas sociais e das exigências de alunos mais autônomos, uma das peças-chave do processo é o professor, “*chamado a desempenhar múltiplas funções, para muitas das quais não se sente, e não foi, preparado*” (BELLONI, 2008, p. 79).

Moran (2003, p. 41) acrescenta que os “*papéis do professor se multiplicam, diferenciam-se e complementam-se, exigindo uma grande capacidade de adaptação*

e criatividade diante de novas situações, propostas, atividades". Maggio (2001, p. 99) destaca que *"um bom docente cria propostas de atividades para a reflexão, apóia sua resolução, sugere fontes de informações alternativas, oferece explicações, facilita os processos de compreensão; isto é, guia, orienta, apóia, e nisso consiste seu ensino"*.

A adequada elaboração de materiais didáticos e de planos de ensino, bem como a apropriada orientação e acompanhamento dos alunos durante seus percursos na EAD, deve estar em consonância com as concepções pedagógicas do curso, os recursos disponíveis e o público-alvo. Considerando todas essas dimensões, a formação de docentes para atuar em EAD exige que se diferenciem diversos papéis docentes existentes nessa modalidade, descritos a seguir, e que se prepare os professores para exercê-los.

2.3.1 Autor

O professor autor é aquele responsável pela elaboração (conteúdo e metodologia) dos materiais didáticos, tanto escritos quanto audiovisuais, de um determinado curso ou disciplina. De acordo com os formatos dos materiais, podem ser assessorados por equipes interdisciplinares que orientam a produção, incluem recursos e organizam as versões finais dos materiais.

Para a produção escrita, o professor deve elaborar textos dialógicos, observando a utilização de recursos de diagramação e de estrutura do texto que levem os alunos a uma leitura envolvente e que tenham potencial para ativar suas redes de significações (SOLETIC, 2001). O material escrito é apresentado em várias formas, tais como artigos, livros didáticos, manuais, anotações de aula, ou mesmo hipertextos (MOORE & KEARSLEY, 2007).

Do ponto de vista do material audiovisual, o professor deve conhecer o processo de produção, executado por equipe interdisciplinar, que passa pela elaboração de pré-roteiros para as aulas, roteirização, produção dos recursos utilizados (entrevistas, enquetes, encenações), gravação do professor em estúdio e a edição (montagem) final (MORAN, 1995). A elaboração de material audiovisual de qualidade requer criatividade e conhecimento profissional especializado, o que

costuma ser bem mais oneroso do que a criação de materiais impressos (MOORE & KEARSLEY, 2007).

2.3.2 Especialista

O professor especialista é aquele que efetivamente interage com alunos e tutores, por meio de salas virtuais no AVA. O professor deve ser instrumentalizado na utilização de um AVA e vivenciar estudos como aluno em tais ambientes. A partir de materiais e dinâmicas interativas, o professor é levado a romper o paradigma de que não é possível estabelecer comunicação eficaz e afetividade a distância (FRANÇA et al., 2008).

Posteriormente, o professor é orientado a planejar sua ação docente por meio da construção de planos de aula. Esses planos devem evidenciar a utilização das ferramentas disponíveis (de compartilhamento de materiais, de comunicação e de gestão) para mediar o acesso dos alunos aos conteúdos e para oferecer apoio à construção de seus conhecimentos (CASTRO et al., 2002). O professor deve destacar ainda as formas como pretende manter os alunos informados sobre a organização da disciplina, como os motiva, engaja e resgata e como acompanha seus estudos e aprendizagens, estabelecendo ações que demonstrem a avaliação formativa.

As pesquisas descritas por Berge (1995) destacam que há quatro principais grupos de funções exercidas pelo professor *on-line*:

- *Pedagógica*: promover a aprendizagem, por meio de atos pedagógicos, tais como provendo informações, orientações, opiniões, questões, *feedback*, resumos e referências a fontes externas;
- *Gerencial*: gerenciar as ações dos alunos e organização do curso, tais como coordenação das tarefas da disciplina, coordenação das discussões estabelecidas e a coordenação da disciplina como um todo;
- *Social*: criar um ambiente de comunicação confortável, minimizando o isolamento, buscando estabelecer empatia, criando vínculos interpessoais, com certa medida de humor;

- *Suporte técnico*: apoio na utilização dos ambientes e recursos virtuais empregados, orientando o aluno com relação a como obter ajuda técnica, se necessário.

2.3.3 Tutor

O tutor atua acompanhando os alunos, a distância (por meio do AVA) ou presencialmente nos pólos (local externo à universidade que oferece estrutura física de salas de aula, laboratórios, bibliotecas e secretaria), orientando os alunos e dinamizando as aulas em consonância com a proposta e acompanhamento do professor especialista (MAGGIO, 2001).

2.3.4 Acompanhamento de Cursos de Formação Docente para a EAD

Durante os processos de formação docente acompanhados diretamente pela autora da tese, na instituição de Ensino Superior que provê os dados para esta pesquisa, foram inúmeros os depoimentos de surpresa (positiva) em relação ao cuidado e a qualidade do material didático produzido e sobre as possibilidades de acompanhamento e intervenção do professor no AVA. Ao todo, foram acompanhadas sete edições do curso de formação docente para a EAD, entre 2007 e 2009, com aproximadamente 330 professores capacitados, com carga horária de 60h/a em cada edição, sendo 30h/a presenciais e 30h/a distância.

Ao final da formação, os professores costumavam abandonar a idéia de que “*é fácil e menos trabalhoso*” atuar na EAD e passavam a externar a concepção de que a EAD é muito mais exigente em termos de tempo e atenção do professor para uma adequada dinamização das aulas, entendidas como organização de percursos de aprendizagem e suporte aos alunos em suas caminhadas. É preciso planejar a *distância transacional* desejada, combinando as dimensões de *estrutura* do curso, *diálogo* entre os atores educacionais e *autonomia* esperada dos alunos, auxiliando na construção desta.

Particularmente à pesquisa realizada na tese interessam os professores especialistas e tutores, que mediam as atividades junto aos alunos por meio de AVA.

Além de definir materiais e atividades (avaliativas ou não), organizando-os em uma estrutura definida e compartilhada com os alunos, é preciso orientá-los sobre formas de trabalho e organização, bem como criar espaços para o diálogo e fomentá-lo. O professor deve estar atento e intervir durante o curso, chamando os alunos à participação e oferecendo *feedback* sobre as atividades que desenvolvem e as dúvidas que apresentam. Moran (2007) afirma que “*na EAD, na maior parte do tempo, o professor não ‘leciona’, mas acompanha, gerencia, supervisiona, avalia o que está acontecendo. Seu papel de professor muda claramente: orienta, mais do que explica*”.

Durante os cursos de formação docente para a EAD acompanhados, surgiram discussões de elementos que os professores gostariam de ter – informações mais facilmente disponíveis e interligadas – para acompanhar os alunos no AVA, tendo em vista construir uma comunicação mais eficaz e uma intervenção do momento mais apropriado. A necessidade identificada também contribuiu para a pesquisa realizada na tese.

2.3.5 Considerações

Muito já se avançou na qualidade de produção de material didático para a EAD: livros impressos, hipertextos, videoaulas, recursos interativos na forma de jogos ou simuladores. Esses materiais, elaborados por professores conteudistas especialistas em suas áreas de conhecimento e por especialistas nas mídias utilizadas, mesmo que em muitos casos tenham excelente produção, tanto teórica quanto gráfica, não garantem por si só a qualidade da aprendizagem do aluno que com eles interage (MEC/SETEC/SEED, 2007).

O professor que acompanha uma turma de alunos em EAD deve transcender o papel de conteudista, ressaltando que, muitas vezes, o material didático já foi elaborado por outro professor. O que fará, então, esse professor? Ficará a disposição apenas para esclarecer dúvidas surgidas nos estudos dos alunos?

É preciso romper com muitos paradigmas vigentes no Ensino Superior (que têm o professor como transmissor absoluto de conteúdos especializados e o aluno

como receptor) para viabilizar EAD de qualidade e, conseqüentemente, com ganhos pedagógicos significativos, por meio de uma educação interativa (SILVA, 2001), dialógica (FREIRE, 2005). Os alunos, mesmo aqueles dotados de autonomia, beneficiam-se, inegavelmente, de professores mediadores durante seus percursos de estudo (PETERS, 2001) – que os acompanham, sugerem caminhos, incentivam o estabelecimento de relações entre os conteúdos apresentados, contextualizam as propostas –, bem como ampliam seus conhecimentos a partir de trocas com outros colegas, que compartilham suas dúvidas e seus achados.

Definitivamente, apenas reagir à demanda dos alunos, apresentando respostas às suas indagações, empobrece o papel do professor (ALVES & NOVA, 2003). O professor, com postura pró-ativa e mediadora, deverá gerir sua sala de aula virtual de forma a integrar seus alunos em uma comunidade virtual de aprendizagem (aproximando os alunos entre si, embora geograficamente distantes, e eles dos objetos de estudo), que estabeleça trocas significativas e onde cada um esteja engajado nos estudos (RAMAL, 2003).

2.4 Mediação

O professor pode, por meio de uma *estrutura* de curso bem planejada e de ferramentas de *diálogo* para promover trocas coletivas e atendimentos individuais, contribuir para o desenvolvimento da *autonomia* dos alunos. Conforme Litwin (2001, p. 14) é preciso não confundir autonomia com autodidatismo, “*pois um autodidata é o estudante que seleciona os conteúdos e não conta com uma proposta pedagógica e didática para o estudo*”.

Embora a modalidade a distância permita uma organização autônoma dos estudantes, não se deve esquecer que nela selecionam-se os conteúdos, orienta-se o prosseguimento dos estudos e propõem-se atividades para que os estudantes resolvam os mais complexos ou os mais interessantes problemas. Os programas de educação a distância contêm uma clara proposta didática, talvez, e em contradição com as crenças mais comuns, com maior conteúdo didático que as situações presenciais. (LITWIN, 2001, p. 14/15)

Litwin (2001, p. 14) salienta, ainda, que se criou a idéia de que para a EAD era necessário elaborar materiais auto-suficientes para gerar uma proposta de aprendizagem, mas “*se os materiais substituem as aulas convencionais e estas*

nunca são suficientes para assegurar o êxito da aprendizagem, é difícil que um material o assegure”.

Frente às novas necessidades educacionais, decorrentes de um modelo de EAD interacionista, professores e alunos precisam reaprender a exercer seus papéis. Os professores devem transcender à idéia de transmissão de conhecimento, já que os conteúdos são postos para o aluno em um momento definido, e devem passar a planejar suas aulas, ou melhor, possíveis percursos didáticos e interações, como mediação dos alunos na construção do conhecimento. Os alunos, por sua vez, precisam desenvolver novas atitudes de busca e acompanhamento das aulas.

Mediação e autonomia estão intimamente interligadas e ambas são importantes na aprendizagem do aluno. A apropriação (aprendizagem) é promovida a partir de um processo de mediação intencionalmente organizado e planejado pelo professor que cria as situações e espaços de problema assim como promove as trocas e as reflexões culminando numa aprendizagem autônoma. Este processo é permeado de momentos intensos de interação e mecanismos específicos para mediação que tem sido pouco aprofundado nos estudos específicos de EAD (...). (PASSERINO et al., 2007)

Belloni (2008, p. 103) acrescenta que estruturas adequadas de mediação são *“especialmente importantes em um país como o Brasil, onde os níveis de cultura geral e de escolaridade são, de modo geral, pouco elevados e onde a escola não instrumentaliza os jovens para o exercício da auto-aprendizagem”.*

“Um processo de ensino e aprendizagem centrado no estudante será então fundamental como princípio orientador de ações de EAD. Isto significa não apenas conhecer o melhor possível suas características socioculturais, seus conhecimentos e experiências, e suas demandas e expectativas, como integrá-las realmente na concepção de metodologias, estratégias e materiais de ensino, de modo a criar através deles as condições de auto-aprendizagem.” (BELLONI, 2008, p.31)

A mediação de grupos de aprendizagem em EAD, portanto, requer do professor que esteja atento às falas e aos comportamentos dos alunos. Costumam ser variados os espaços de compartilhamento de materiais e de interação. A comunicação, seja para manter o diálogo aberto (estabelecer empatia), para incentivar a participação (promover engajamento) ou para orientar/reorientar o percurso do aluno (visando à aprendizagem), deve ser adequada às suas necessidades específicas, nos contextos comunicativos adequados, apoiando e não

gerando sobrecarga nos canais comunicativos. O anexo 1 contém uma tabela, adaptada a partir de Palloff & Pratt (2004), com técnicas de mediação que podem auxiliar na organização das ações do professor. Mercer & Estepa (2001, p. 25) ressaltam que “*o bom ensino é um processo orientado, social e comunicativo*”.

... os alunos assumem papéis mais ou menos ativos durante o processo de aprendizagem. Podem, por exemplo, cumprir a função de meros ouvintes ou apenas responder às questões levantadas pelos professores, expressar de modo franco suas confusões ou fazer perguntas ao professor para que os ajude a aprofundar sua compreensão a respeito de um determinado assunto, entre outras coisas. Uma boa proposta pedagógica oferecerá aos estudantes as oportunidades e possibilidades de converter-se ativamente em protagonistas de seus próprios processos de aprendizagem. (MERCER & ESTEPA, 2001, p. 25)

Um professor atuante na orientação/reorientação dos alunos, promovendo debates problematizadores e dinamizadores das trocas sociais, influencia muito na percepção dos alunos de *proximidade comunicativa*. Belloni (2008, p. 54) coloca que “*o contato regular e eficiente, que facilita uma interação satisfatória e propiciadora de segurança psicológica entre os estudantes e a instituição ‘ensinante’, é crucial para a motivação do aluno, condição indispensável para a aprendizagem autônoma*”. Belloni (2008) afirma também que os problemas gerados pela distância geográfica podem ser mais facilmente superados com a utilização de sistemas eficientes de comunicação entre estudantes, tutores e professores.

Uma maior proximidade comunicativa cria condições para a formação de uma verdadeira comunidade de aprendizagem, onde, com respeito mútuo e liberdade, podem expressar suas hipóteses e discuti-las, lançando-se em busca de novos argumentos, confirmando-as ou refutando-as. O processo educacional exige diálogo aberto, entre pessoas que interagem e se afetam mutuamente (FREIRE, 2005).

Ao organizar as suas ações pedagógicas, o professor deve pensar como acompanhará e dinamizará as trocas entre ele e seus alunos (e desses entre si), estabelecendo vínculos e canais de comunicação que permitam trocas significativas. Bouchard (2002, p. 76) afirma que a “*aprendizagem se faz progressivamente por meio de questionamentos, de perguntas e de reformulações que exigem uma mediação pelo diálogo*”.

Belloni (2008, p. 27) afirma que *“a separação no tempo – comunicação diferida – talvez seja mais importante no processo de ensino e aprendizagem a distância do que a não contigüidade espacial”*. Toda comunicação se dá em contexto e o *“sentido emerge e se constrói no contexto, é sempre local, datado, transitório”* (LÉVY, 1999). O ato de comunicar, num primeiro momento, destina-se a transmissão de informações, mas o contexto em que ocorre confere significado a esse ato. No contexto, emergem novos comentários, novas interpretações, que podem alterar o sentido da mensagem anterior.

O jogo da comunicação consiste em, através de mensagens, precisar, ajustar, transformar o contexto compartilhado pelos parceiros. Ao dizer que o sentido de uma mensagem é uma “função” do contexto, não se define nada, já que o contexto, longe de ser um dado estável, é algo que está em jogo, um objeto perpetuamente reconstruído e negociado. Palavras, frases, letras, sinais ou caretas interpretam, cada um à sua maneira, a rede das mensagens anteriores e tentam influir sobre o significado das mensagens futuras. (LÉVY, 1999, p. 22)

A linguagem, independentemente da modalidade de educação, constitui-se na principal ferramenta de que dispõem professores e alunos para a construção do conhecimento. Na EAD, a linguagem escrita é um dos meios mais importantes de comunicação entre os professores e alunos de um curso (MERCER & ESTEPA, 2001). Assim, não apenas os materiais didáticos são, boa parte das vezes, apresentados em suporte escrito, como também é por meio de trabalhos escritos que os alunos habitualmente demonstram o que compreenderam sobre os estudos realizados. O retorno sobre os trabalhos dos alunos também costuma ser realizado pelos professores na forma de comunicações escritas.

Quando os sujeitos compartilham de um contexto comum, a mensagem emitida tende a ser interpretada pelos seus receptores de forma mais próxima a perspectiva do emissor, ativando redes de significações que permitam a adequada compreensão da mensagem. A interação entre os sujeitos é processo essencial para a aprendizagem que, para não ser inibida, exige diálogos humanizados, afetuosos e respeitosos (FREIRE, 2005).

Moore & Kearsley (2007) destacam que os mediadores em EAD devem utilizar técnicas de interação para:

- *Humanização*: criar um ambiente que enfatize a importância do sujeito e gere a sensação de relacionamento com o grupo. Isto pode ser feito utilizando os nomes dos alunos nas comunicações e solicitando que expressem opiniões e experiências pessoais.
- *Participação*: garantir que haja um bom nível de interação, questão possível por meio de técnicas de formular perguntas e de propor atividades em grupo para resolução de problemas.
- *Estilo da mensagem*: utilizar boas técnicas de comunicação, variando as formas de apresentar as informações, de promover integração dos assuntos abordados e de orientar os estudos.
- *Feedback*: fazer uso de tarefas e pesquisas para obter informações a respeito do progresso dos alunos em suas aprendizagens.

É preciso acompanhar de perto os processos de aprendizagem, oferecendo caminhos para construção individual e coletiva de significados, ajudando os alunos a estabelecer rotinas de estudos e posturas de buscas, oportunizando avaliações formativas com *feedback* e estudos adicionais.

Os professores precisam encontrar um equilíbrio na quantidade e na frequência do *feedback* que dão aos alunos, de maneira que estes saibam que os professores estão presentes e atentos, mas não de uma forma que sobrecarregue o grupo ou que domine a discussão. O papel do professor-facilitador nas discussões *on-line* é ajudar os alunos a sintetizarem o material estudado, auxiliando-os, assim, a irem em frente. Quando o professor intervém em demasia, há, de fato, uma interrupção do diálogo, pois ele se torna o centro das atenções. Por outro lado, a ausência de intervenção do professor pode também interromper a discussão, pois os alunos talvez se sintam confusos e incertos sobre o que se espera deles. Atingir o equilíbrio é, então, fundamental. (PALLOFF & PRATT, 2004, p. 155)

Araújo (2001) realizou uma pesquisa buscando comparar os objetivos pretendidos pelos cursos de formação docente para a EAD e as expectativas dos alunos em relação aos professores quando cursam algo nesta modalidade. Em geral, os alunos de cursos a distância esperam que seus professores façam-se “presentes” na sala de aula virtual, por meio de *feedback* constante e individualizado. De acordo com os resultados de tal pesquisa, para os alunos, a

“presença” do professor é mais importante do que o material didático ou a tecnologia empregada nos cursos.

Palloff & Pratt (2004) sintetizam as necessidades do aluno a distância, que devem pautar a mediação do professor:

Em resumo, o que o aluno virtual quer e precisa é algo muito claro: comunicação e *feedback*, interatividade e sentido de comunidade, direção adequada e capacitação para executar as tarefas exigidas. Se copiarmos o que acontece na sala de aula presencial, não atenderemos essas necessidades, causando angústia e frustração aos alunos. A solução é concentrar-se nas práticas que são centradas no aluno e que buscam atender às suas necessidades. (PALLOFF & PRATT, 2004, p. 153)

2.5 Acompanhamento

Na EAD, em função da distância física dos alunos em relação aos professores e suas instituições, o êxito dos cursos depende de sistemas eficazes de acompanhamento e avaliação. Apenas com bons mecanismos de monitoramento e avaliação é que os professores poderão identificar dificuldades específicas dos alunos e, assim, oferecer o tipo adequado de ajuda. Além disso, atrasos nas entregas de trabalhos ou “quebra” na comunicação também devem ser percebidos a tempo de uma intervenção corretiva.

O monitoramento eficaz exige uma rede de indicadores que disponibilizem os dados necessários sobre o desempenho do aluno e do professor; isso precisa ser feito *frequente e rotineiramente*, e os dados precisam ser transferidos com uma rotina similar para um centro de controle onde possam ser avaliados. Avaliação nesse contexto é o processo de análise dos dados de *feedback* colhidos pelo sistema de monitoramento, revisão e tomada de decisões a respeito de grau de adequação com que o sistema de educação a distância e suas várias partes estão operando; de que modo alunos, instrutores, profissionais de criação, administradores de recursos de comunicação operam juntos para atingir as metas de curto e longo prazo. Para um sistema educacional, as metas mais importantes são os resultados do aprendizado; no entanto, outras metas são legítimas e podem ser monitoradas e avaliadas (por exemplo, manter o custo compatível ou retificar desequilíbrios demográficos na população de alunos). (MOORE & KEARSLEY, 2007, p. 130)

Conforme Moore & Kearsley (2007, p. 130), *“uma das poucas generalizações que se pode fazer a respeito de qualquer programa de educação a distância – seja qual for a tecnologia de comunicação usada e o nível de conteúdo – é que um bom*

sistema de monitoramento e avaliação pode conduzir a um programa bem-sucedido, ao passo que um sistema ruim quase certamente levará ao fracasso”.

Para a qualidade da mediação, é fundamental acompanhar cada aluno, identificando suas necessidades e oferecendo auxílio personalizado. Da mesma forma que na sala de aula presencial o professor pode identificar sinais de que o aluno está enfrentando dificuldades, é preciso acompanhar indicadores para identificar que o aluno pode estar encontrando problemas durante o curso em EAD.

Esse acompanhamento individualizado, no entanto, torna-se um problema para o professor quando o número de alunos em uma turma é grande, já que na maioria das vezes as informações disponíveis para esse acompanhamento são escassas ou estão disponíveis em relatórios diversos não consolidados. É preciso desenvolver sistemas que apóiem o monitoramento e a avaliação – estando atento aos indicadores e reconhecendo indícios de falha – para a adequada tomada de decisão.

2.6 Evasão e Desempenho

A evasão e os problemas de desempenho são preocupações presentes na Educação, de forma geral, seja presencial ou a distância. Especificamente na EAD, devido às características inerentes da modalidade, tais como a distância física entre alunos e professores e a comunicação midiaticizada por recursos tecnológicos, os índices têm sido acompanhados e mostram-se altos, comprovados em diversas pesquisas sobre o tema, descritas nesta seção.

A retenção de alunos nos cursos *on-line* passou a ser uma questão significativa para os administradores, que descobriram que não é barato manter e criar tais cursos e programas. O *hardware* e o *software* são caros, bem como o treinamento, o suporte técnico e o pagamento dos professores pelo desenvolvimento e pela realização dos cursos. Os professores, também, se preocupam com o índice de retenção. Um professor universitário com quem falamos disse que, com frequência, aceitava um número de inscritos em seus cursos *on-line* que superava em um terço o número máximo de alunos, prevendo o abandono futuro e não querendo que os administradores de sua instituição ficassem indignados com seu possível quadro de evasão. O professor também observou que isso acabava por aumentar sua carga de trabalho, pois ele tinha de identificar os alunos que demonstravam dificuldades em começar o curso e aqueles que o abandonariam mais cedo ou mais tarde. (PALLOFF & PRATT, 2004, p. 24)

Carr (2000) apontava, no ano de publicação da sua pesquisa, que o abandono de cursos *on-line* girava em torno de 50 por cento dos alunos matriculados nos Estados Unidos. Segundo Longo (2009, p. 219), em publicação atualizada, “*considerando parâmetros internacionais, a evasão na EAD é altíssima, chegando, em alguns casos, a 65 por cento*”. Maia et al. (2004), em sua pesquisa, coletou dados de 22 instituições brasileiras de ensino superior, com 51 mil alunos matriculados em EAD, sendo 55% públicas e 45% privadas. Os índices de evasão em curso de graduação a distância foram em torno de 30%. Almeida & Ildete (2008), acompanhando turmas do curso livre “Aprender a Empreender pela Internet”, do SEBRAE – Serviço Brasileiro de Apoio às Micro e Pequenas Empresas, no período de fevereiro a maio de 2006, constataram uma taxa de evasão de 41,3% (totalizando 1021 alunos evadidos).

Palloff & Pratt (2004) citam que vários críticos da EAD apontam os elevados índices de desistência em cursos dessa modalidade como indicativo de má qualidade. Diaz (2002), porém, ressalta que os índices de abandono não têm relação direta com má qualidade ou o fracasso acadêmico, já que a sua pesquisa permitiu observar que os alunos *on-line* freqüentemente tinham desempenho (conceito final) melhor e mais satisfação do que os alunos presenciais em cursos análogos.

Diante desse quadro, Palloff & Pratt (2004, p. 135), colocam a seguinte questão: “*Então por que os alunos on-line desistem de suas aulas em maior número e o que pode ser feito para impedir isso?*” Diversas pesquisas buscam responder a essa questão e, na seqüência, algumas referências importantes são apresentadas.

Carr (2000) afirma que estudos têm demonstrado que os mesmos motivos que levam o aluno a optar pela modalidade a distância são os que interferem na sua permanência, tais como a possibilidade de conciliar seus estudos com uma longa jornada de trabalho ou com o atendimento da família.

Gibson (1998) estabeleceu três categorias de fatores que podem explicar os motivos de abandono do curso:

- Fatores do estudante, incluindo a preparação educacional anterior, atributos de motivação e persistência, bem como a sua autoconfiança acadêmica;
- Fatores situacionais, como apoio da família e do empregador, além de mudanças em circunstâncias da vida pessoal;
- Fatores do sistema educacional, englobando tanto a qualidade e as dificuldades com o material instrucional, como com o suporte oferecido pela instituição.

Os motivos para evasão são bastante complexos e abrangem questões internas e externas às instituições educacionais. Longo (2009, p. 219) destaca, de forma mais detalhada, mas não exaustiva, como causas mais freqüentes para a evasão: falta de tempo; questões profissionais; insatisfação com o conteúdo; estranhamento gerado pela mudança de paradigma, de presencial para a distância; falta de traquejo com as ferramentas; insatisfação com a turma; problemas pessoais; escolha inapropriada do curso; insatisfação com a tutoria ou mesmo ausência de tutoria; problemas técnicos e falta de suporte.

2.6.1 A Pesquisa sobre Evasão na Open University

Um estudo realizado com os alunos da *Open University* (Inglaterra), descrito por Tresman (2002), mostrou que os motivos que afetam o progresso dos estudantes apontam diferenças entre os cursos pesquisados e as regiões geográficas nas quais os alunos estudam. Em relação aos cursos, há uma considerável variação entre os motivos indicados para o abandono: o número de estudantes que indicaram como motivo as pressões em casa e no trabalho para sua desistência variaram entre 32 por cento (Curso de Saúde e Assistência Social) e 69 por cento (Curso de Artes).

O fato de estar em uma cidade cosmopolita como Londres, por exemplo, com muitas pessoas migrando o tempo todo, refletia esta característica nas altas taxas de abandono. Também mostrou-se relevante se o estudante era novo na *Open University*, indicativo de que era sua primeira experiência com EAD. Altas taxas de

evasão eram registradas para estudantes novos e estudantes com características demográficas de risco, ligadas a questões de gênero, idade, etnia, qualificações educacionais prévias, etc.

A qualidade da tutoria foi também um item significativo. No ano de 2000, 18 por cento dos estudantes descreveram-se como “muito insatisfeitos” com as formas de tutorias, a terceira razão mais citada pelos estudantes para justificar o abandono, seguida do tempo e pressão no trabalho. Seis por cento declararam que deixaram o curso especificamente por não estarem satisfeitos com seus tutores. Outros fatores citados pelos estudantes incluíram informações incompletas dos cursos e a insatisfação com os materiais. Esses fatores revelaram aspectos importantes sobre as expectativas dos alunos e indicam processos a serem melhorados na educação a distância.

Além dos motivos alegados, há também evidências na pesquisa, e em outros trabalhos realizados como parte do Programa de Retenção, de que a Universidade pode estar causando, ou pelo menos contribuindo, para a sensação de sobrecarga dos estudantes. Muitos estudantes reportaram que os cursos nos quais estavam envolvidos tomavam mais tempo, ou possuíam ritmos mais intensos, do que eles esperavam. Esses dados também revelaram expectativas irreais da Universidade em relação aos estudantes, em termos daquilo que eles poderiam realmente realizar no tempo disponível para completar o curso, apontando a necessidade de revisar os currículos sem prejudicar os objetivos de aprendizagem.

A partir do desenvolvimento de ações para integrar os estudantes no ambiente acadêmico, as instituições de ensino superior podem influenciar positivamente o comportamento dos alunos. Esta é a base do modelo de retenção de Tinto (2000), que argumenta que se as instituições se concentrarem em melhorar as experiências de aprendizagem dos alunos, crescerá a probabilidade deles persistirem em seus estudos.

Em linhas gerais, Tinto (2000) afirma que os alunos têm mais chances de aprender e persistir quando se encontram em ambientes que:

- Possuem altas *expectativas* para a sua aprendizagem e apresentam isto de forma clara e consistente: os alunos devem ser informados sobre o que se espera deles e o que é necessário para completar as disciplinas e o curso com sucesso.
- Fornecem *apoio* acadêmico e social para suas necessidades, essenciais para a promoção da retenção e da aprendizagem: em alguns casos, os estudantes chegam à universidade insuficientemente preparados para o rigor do estudo universitário, sendo necessário apoio ao desenvolvimento de habilidades básicas (por meio de cursos de nivelamento, monitorias ou atividades adicionais, por exemplo) e de hábitos de estudo.
- Provêem *feedback* freqüente sobre a sua aprendizagem, referindo-se não só à avaliação da aprendizagem e às ações que devem decorrer dessa avaliação, mas também diz respeito a sistemas de alertas precoces que avisem as instituições sobre estudantes que precisam de assistência, de forma que essa assistência possa ser fornecida com antecedência suficiente para fazer a diferença. Um atributo importante desses sistemas é que eles devem estar ligados à aprendizagem em curso, oferecendo elementos para a atuação institucional. Conforme os estudos de retenção, quanto mais se espera para intervir, mais difícil é reverter um processo de desgaste do aluno em seus estudos.
- Oportunizem o *envolvimento* com outros alunos e professores em aprendizagens significativas, em comunidades de aprendizagem que favoreçam o sentimento de pertença e engajamento: entre os alunos que persistem, estudantes mais ativamente envolvidos com os processos de aprendizagem, especialmente com os outros, aprendem mais e demonstram níveis melhores de desenvolvimento intelectual.
- Promovam *aprendizagem relevante*: quanto mais valor os estudantes encontrarem em suas aprendizagens, quanto mais elas estiverem conectadas aos seus interesses, mais estarão envolvidos em suas aprendizagens e, conseqüentemente, mais aprenderão e estarão estimulados a persistir.

Ao garantir que as informações sobre o curso refletem com maior precisão os cursos oferecidos, bem como a modalidade a distância de estudo praticada, os alunos têm mais probabilidade de ter sucesso em seus estudos, porque as suas expectativas estarão mais sintonizadas com a realidade do curso. Além disso, para melhorar os índices de retenção, é importante destacar que os tutores devem não apenas supervisionar o progresso do estudante, mas também monitorar e contatar aqueles que apresentam riscos de evasão.

2.6.2 Questões Externas e Internas à Instituição

Tresman (2002), em seu estudo sobre o retenção de estudantes da *Open University*, na Inglaterra, além dos estudos de Biazus (2004), com alunos do curso de Ciências Contábeis da UFSC, e de Pacheco et al. (2008), com alunos do curso de Administração também da UFSC, apresentam como questões externas principais de evasão às ligadas a carga horária de trabalho do aluno adulto e sua dificuldade de priorizar os horários para estudo, a não adaptação à modalidade, a incerteza da vocação (curso escolhido) e problemas de ordem financeira.

Tresman (2002) destaca que, em alguns casos, a evasão ocorre quando os alunos não estabeleceram adequadamente a integração dos seus estudos com seus ritmos de vida e, então, quando eles abandonam os cursos, costumam dizer: “o volume de trabalho é muito grande”, “o curso estava muito difícil”, “as mensalidades eram muito caras”.

Como questões internas aos cursos em EAD que influenciam na permanência dos alunos, estão a relevância do conteúdo para seus interesses, a qualidade do material, a usabilidade das tecnologias utilizadas, a metodologia empregada, a forma de avaliação, a interação com professores e colegas e a quantidade e natureza do *feedback* recebido com relação às tarefas realizadas e ao progresso no curso (MOORE & KEARSLEY, 2007).

Tresman (2002), Maia et al. (2004) e Favero & Franco (2006) destacam que o bom planejamento do curso, que leve em consideração as questões acima, impacta positivamente na permanência dos alunos, especialmente quando eles se sentem

apoiados. Lago (2003, p. 80) reforça que “*incentivos e a preocupação (e acompanhamento) por parte do curso para com o aluno (...) afastam a sensação de abandono, muitas vezes interpretada pela questão da distância*”, contribuindo para a permanência do aluno. É preciso apoiar o aluno a vencer barreiras tecnológicas, a criar hábitos de estudo e a organizar seu tempo para manter as atividades dentro dos prazos.

Apesar de todos os esforços para auxiliar os alunos a identificar o nível certo do curso, a fim de facilitar seu acesso ao aprendizado a distância, alguns alunos encontrarão problemas inesperados de trabalho, família ou saúde que ameaçam seu avanço acadêmico. Um serviço de apoio ao aluno tem de ser pró-ativo e reativo. Se ele reagir apenas aos alunos que pedirem ajuda, muitos desistirão. Métodos precisam ser desenvolvidos para identificar problemas logo no início e intervir para oferecer ajuda, muito embora o aluno possa não se apresentar para solicitá-la. O método principal consiste em monitorar cuidadosamente a produtividade das tarefas escolares. Se um aluno que normalmente produz boas tarefas escolares começa a piorar ou não produzir pontualmente, um *sinal vermelho* deve alertar o pessoal de apoio aos alunos para um possível problema potencial que poderá exigir pelo menos uma mensagem por e-mail oferecendo ajuda. Deixar de seguir tais passos pode significar que os problemas não-acadêmicos exigirão a atenção integral do aluno e é bem possível que o aluno desistirá do curso. (MOORE & KEARSLEY, 2007, p. 194)

Diferentes estratégias podem ser empregadas para incrementar os índices de retenção dos estudantes em EAD. À pesquisa desta tese, especialmente, tocam aspectos que podem ser trabalhados a partir do acompanhamento e da intervenção do professor.

2.7 Ambientes Virtuais de Aprendizagem

Os Ambientes Virtuais de Aprendizagem (AVA) surgem como soluções que integram diferentes recursos computacionais para permitir compartilhamento de materiais, em diferentes formatos, e variados tipos de comunicação, síncronas e assíncronas, em um mesmo espaço virtual. Estes sistemas derivam-se de estudos na área de Suporte ao Trabalho em Grupo (STG), que sustenta a produção de ferramentas que se utilizam de tecnologia da informação para que as pessoas possam trabalhar juntas mais eficientemente (ELLIS et al., 1991; HILLS, 1997).

Um sistema de STG tem por objetivo apresentar-se como um recurso de apoio às atividades de grupo, que auxilie no processo de colaboração dos mesmos,

visando a um ganho de desempenho. Para tanto, as ferramentas devem fornecer facilidades que possibilitem o compartilhamento de informação, a coordenação e o controle das atividades, bem como a comunicação entre os seus membros (MESQUITA et al., 1994).

Com base nestas considerações, para que ocorra interação de qualidade, segundo Borges et al. (1995), torna-se necessário:

- *Comunicação*: primordial para que as trocas se estabeleçam, mediada por computador e apoiada pelas telecomunicações e pela tecnologia de processamento computacional;
- *Negociação*: forma mais sofisticada de comunicação, exige alguma estruturação;
- *Coordenação*: processo fundamental para que o trabalho seja bem sucedido. Deve prover mecanismos de controle sobre as comunicações e negociações e de monitoramento sobre os resultados gerados.

Os itens acima se apresentam como questões de fundamental importância para a realização de tarefas entre indivíduos pertencentes a um grupo com objetivos comuns. Existem vários tipos de aplicações de sistemas de STG, tais como (ELLIS et al., 1991; BORGES et al., 1995): correio eletrônico, listas de discussão, quadros de avisos, fóruns, comunicadores instantâneos, videoconferência, editores cooperativos, sistemas de apoio à decisão e gerenciadores de fluxo de informações – *workflow*.

Na área educacional, desenvolve-se uma sub-área de STG, com a finalidade de investigar e gerar tecnologias para apoiar os processos de comunicação, negociação e coordenação no âmbito educativo: são os Sistemas de Gestão de Aprendizagem (*Learning Management Systems* – LMS), também denominados Ambientes Virtuais de Aprendizagem (AVA). Na EAD, os AVA são sistemas empregados para dar suporte a atividades educacionais, apoiando a relação pedagógica entre alunos e professores que estudam na modalidade a distância, ou mesmo como complemento de aulas presenciais.

Alguns dos AVA amplamente utilizados no Brasil são o *AulaNet*¹, o *Teleduc*² e o *Moodle*³, todos esses gratuitos. O *Teleduc* e o *Moodle* têm código aberto, favorecendo sua disseminação e que diversos grupos acadêmicos pesquisem e desenvolvam novas funcionalidades, visando a um melhor acompanhamento e dinamização das atividades didático-pedagógicas. Na pesquisa desta tese, foi utilizado o AVA *NetAula*⁴, de propriedade da instituição onde a pesquisa foi realizada.

O desenvolvimento do AVA *NetAula* é realizado pela Fábrica de *Software* do setor de Tecnologia da Informação da instituição geradora. Esse ambiente virtual de Educação a Distância foi desenvolvido para atender às necessidades da instituição, estando totalmente integrado com o sistema acadêmico. Isto significa que uma turma a distância criada, de acordo com o calendário acadêmico de início e fim do período letivo, automaticamente leva à geração da respectiva sala virtual. Na medida em que ocorre o lançamento da turma na carga horária do professor e que as matrículas dos alunos vão sendo realizadas, todos já estarão integrados à sala de aula virtual.

O ambiente *NetAula* comporta ferramentas para disponibilização de aulas, biblioteca virtual, quadro de avisos, correio eletrônico, fórum de discussão, sala de bate-papo, espaço para realização de avaliação, informações sobre os participantes da turma, além de relatórios e acompanhamento de desempenho. Através de um processo iterativo com os professores e demais profissionais que utilizam o ambiente, a equipe de desenvolvimento da plataforma está sempre buscando disponibilizar novas funcionalidades e melhorar as já existentes. Um dos principais pontos de aperfeiçoamento discutido é a melhora no processo de acompanhamento dos alunos, visto que um professor é responsável por muitas turmas e estas possuem uma grande quantidade de alunos.

Há diferentes perfis de usuários no sistema, com responsabilidades específicas. Basicamente, em uma sala de aula virtual, o professor trabalha com a

¹ Disponível em: http://www.eduweb.com.br/aulanet_o_aulanet.html. Acessado em: 30/06/2008.

² Disponível em: <http://www.teleduc.org.br/>. Acessado em: 30/06/2008.

³ Disponível em: <http://moodle.org/>. Acessado em: 30/06/2008.

⁴ Disponível em: <http://www.ulbra.br/ead>. Acessado em: 30/08/2009.

visão de administrador da sala, tendo acesso tanto a informações detalhadas dos alunos, como a funções específicas de organização do ambiente e correção das atividades (anexo 2). Com o perfil de aluno, os estudantes têm acesso aos materiais do curso, às ferramentas de comunicação, às atividades propostas e podem acompanhar seus desempenhos, interagindo com professores e colegas (anexos 3 e 4).

Os AVA são utilizados como locais para compartilhamento de materiais, realização de tarefas, criação e execução de projetos, postagem de dúvidas e busca de respostas e interações efetivas entre sujeitos autônomos, que proporcione a construção de conhecimento, superando os materiais e propostas de atividades iniciais e favorecendo as aprendizagens individuais e coletivas (FAGUNDES et al., 2006). Em um AVA, se formam comunidades virtuais de aprendizagem (CVA), que precisam estabelecer comunicações produtoras de sentido, em grupos que compartilham objetivos comuns. Um AVA, embora virtual, permite *territorializar* (LÉVY, 2000) o espaço compartilhado da comunidade de aprendizagem. É “lá” que a comunidade educativa se encontra.

Neste sentido, o professor deve empregar estratégias que contribuam para o envolvimento dos alunos dos estudos propostos (OKADA, 2003). Não basta disponibilizar conteúdos e abrir espaços de discussão, é preciso que o professor tenha uma postura de mediação, incentivando os alunos a participarem, acompanhando seus desempenhos, dando retorno e orientações durante toda a caminhada.

Os alunos, por sua vez, devem manter regularidade de acesso, com atenção às orientações do professor mediador, realizando as atividades nos prazos estabelecidos. Estudar nesses ambientes exige uma postura autônoma do aluno, pois há flexibilidade de horários e de formas de organização dos estudos, mas é preciso ter muito cuidado: flexibilidade nos estudos não quer dizer ausência de tarefas e prazos. O aluno deve acompanhar o curso de maneira efetiva e constante, participando das discussões da turma, sem acumular dúvidas e trabalhos avaliativos (OKADA, 2003; PAZ et al., 2003).

É essencial, ao iniciar uma disciplina ou curso em EAD, portanto, que professores e alunos tenham clareza da *estrutura* do curso, que estejam apresentados o objetivo da disciplina, a metodologia, as datas importantes de liberação e entrega de atividades, bem como a demanda de estudo e a importância de uma participação ativa nas discussões da turma (OKADA, 2003; PAZ et al., 2003). Tal *estrutura*, de acordo com a evolução do grupo, pode ser reavaliada no decorrer do tempo e visa a diminuir a *distância transacional*. É fundamental que cada aluno se sinta integrante do grupo e motivado para interagir, o que contribui para o seu engajamento e para um melhor desempenho acadêmico.

Dentre as ferramentas mais comuns que se encontram disponíveis em AVA, são destacadas:

- *Dados Pessoais / Perfil*: ferramenta que contém dados de cada sujeito, tais como nome, foto, apresentação pessoal e, em alguns casos, contatos externos (como telefone, e-mail externo e endereço) e que podem ser expostos a toda a turma ou apenas ao professor.
- *Materiais / Biblioteca Virtual*: há várias metáforas implementadas para o compartilhamento de materiais nos diversos AVA, incluindo ferramentas de autoria de páginas HTML ou o compartilhamento de arquivos e *links*.
- *Plano de Aula / Cronograma*: recursos deste tipo permitem organizar as seqüências didáticas, indicando a ordem e os períodos em que determinados materiais serão trabalhados e atividades serão realizadas – dá visibilidade à *estrutura* do curso. Com as atividades pedagógicas organizadas e apresentadas, os alunos podem organizar seus estudos.
- *Avisos / Mural*: trata-se de um recurso destinado a chamar a atenção do aluno, destacando novidades, lembrando prazos ou motivando para participação em atividades especiais. Muitas vezes, essa é a primeira ferramenta que o aluno visualiza ao ingressar no AVA, na página principal ou em uma janela *pop-up*.
- *Correio Eletrônico*: possibilita o envio de mensagens eletrônicas internas no ambiente, que podem ser endereçadas a toda a turma ou apenas a

sujeitos específicos. Muitas vezes é possível configurar para que se receba cópia das mensagens também no *e-mail* externo.

- *Fórum de Discussão*: o fórum possibilita a interação assíncrona e o registro das discussões em um espaço compartilhado pela turma. Os sujeitos podem postar suas contribuições no fórum, cada um a seu tempo, e voltar a ele, em qualquer momento, para ler o que os demais participantes postaram. É possível ler, reler, pesquisar e só após contribuir, o que acaba por qualificar as discussões, já que há mais tempo para amadurecer o que se pretende compartilhar com o grupo. Diferentes estratégias podem ser empregadas para dinamizar as discussões, definindo, por exemplo, o número mínimo de postagens qualificadas por aluno e o espaço de tempo para que ocorram, indicando a necessidade de comentar as postagens de outros colegas, de dar continuidade ou de sintetizá-las. É desejável que o fórum tenha espaço para a possibilidade de avaliação das postagens, pelo professor e pelos colegas.
- *Comunicadores Instantâneos*: há vários ambientes que contam com recursos para visualização de usuários *on-line* e a possibilidade de chamá-los para conversação em tempo real. Em muitos casos, o sujeito pode alterar o próprio *status*, de forma que os demais saibam se ele está disponível para conversação, ocupado ou ausente.
- *Sala de Bate-Papo*: permite interagir em tempo real, esclarecendo dúvidas e trocando referências, independentemente da localização geográfica dos sujeitos. Uma sala de bate-papo, para uma troca efetiva, não deve ter muitos participantes, para que todos possam contribuir, lendo e postando idéias. Cabe destacar que, em discussões educacionais, é importante existir o papel do mediador (exercido pelo professor ou por um aluno designado), que retoma o foco da discussão quando outros se distanciam dele e que apresenta questões instigadoras quando o grupo está apático.
- *Blog*: é uma ferramenta que facilita a publicação de informações. As informações postadas em um *blog* podem ser comentadas por quem o visita. Os comentários podem ser moderados ou não pelos proprietários do *blog*, cabendo a eles decidirem sobre a liberação ou exclusão desses.

Este recurso tem substituído, gradualmente, nos AVA, as ferramentas de *diário de bordo* e de *portifólio*. Um *blog* coletivo pode conter os registros de todas as etapas de construção do trabalho de um grupo e contar com os comentários dos colegas e professores.

- *Avaliação*: ferramentas de avaliação, na maioria dos AVA, permitem a oferta de três tipos de atividades – questões discursivas, questões objetivas e envio de arquivos. Quando uma avaliação é disponibilizada, o aluno passa a visualizar o conjunto de questões que a compõe e o seu prazo de entrega. Após a correção e a liberação do professor, o aluno passa a visualizar a nota obtida na atividade, o gabarito das questões objetivas e os comentários que podem ser acrescentados pelo professor em qualquer tipo de avaliação. Quando a atividade refere-se ao envio de arquivos, o professor pode disponibilizar outro arquivo com correções, mantendo o original do aluno para que ele possa comparar e realizar suas análises.
- *Relatórios de Desempenho*: apresentam, basicamente, dados de acesso ao ambiente (quantidades, ferramentas, datas e horários) e de aproveitamento (notas) das atividades realizadas.

Além dessas ferramentas, alguns AVA oferecem mecanismos para a realização de enquetes, manutenção de glossários, organização de FAQ (*Frequently Asked Questions* – Perguntas Frequentes), edição cooperativa de textos (ou *Wikis*) e gestão de grupos.

Para a mineração de dados e a geração posterior dos alertas, serão considerados: dados pessoais disponíveis, como sexo, idade, curso e pólo presencial; os acessos dos alunos ao AVA, tais como quantidade e frequência; as avaliações realizadas e seus respectivos desempenhos; o acesso aos materiais disponibilizados em áreas de conteúdo ou biblioteca virtual; as participações nos fóruns; e as trocas de *e-mail* realizadas.

A mineração dos dados provenientes dessas ferramentas facilitará a transposição dos conhecimentos gerados pela tese para AVA diversos, uma vez que

são as mais amplamente utilizadas na dinamização de cursos em EAD. Assim, pretende-se extrair conhecimento dos dados gerados pela interação dos alunos no ambiente e as relações com o êxito na conclusão do curso/disciplina, de forma a conscientizar o professor de situações em que intervenções são necessárias, contribuindo para a mediação docente.

3. Mineração de Dados

Com a ampla difusão do uso de sistemas informatizados, em diversas áreas, cresce a cada dia o volume de dados gerados e armazenados em bases digitais. Bases nas áreas de comércio, finanças, saúde, educação e ciência, por exemplo, têm sido muito valorizadas e, conseqüentemente, esforços vêm sendo empregados para analisá-las, uma vez que muitos padrões novos e relevantes, a partir dos dados disponíveis, podem ser identificados (WITTEN & FRANK, 2005).

Empresas que comercializam produtos pela Internet, por exemplo, necessitam analisar os dados gerados pela interação dos usuários, de forma a identificar perfis de consumo dos clientes e, analisando esses perfis, oferecer recomendações personalizadas a eles, tais como produtos similares aos já comprados, complementares, ou adquiridos por usuários com preferências semelhantes. Na área de crédito financeiro, é importante detectar perfis de bons e maus pagadores, bem como estimar o limite de crédito apropriado. A relação entre vários hábitos físicos e sociais com fatores hereditários pode ser relevante para definir a probabilidade de certas doenças se manifestarem em determinadas pessoas.

Para a capacidade humana, no entanto, é difícil analisar conjuntos enormes de dados e encontrar relações significativas entre eles, que não eram previamente conhecidas, embora sejam potencialmente úteis. Por isso, “*minerar dados*”, nas palavras de Deogun et al. (1997), pode ajudar a encontrar informações que valem “*ouro*”.

A Mineração de Dados (MD) é, portanto, uma forma de explorar e analisar bancos de dados, buscando identificar regras, padrões ou desvios. LIN & CERCONE (1997) definem a MD como um processo para a extração de informações implícitas, previamente desconhecidas e potencialmente úteis, a partir de grandes bases de dados, por meio de técnicas eficientes para a descoberta de conhecimento (KDD – *Knowledge Discovey in Databases*). Os mesmos autores ressaltam que a MD tem sido objeto de estudos interdisciplinares, envolvendo, principalmente, as áreas de Estatística, Inteligência Artificial (IA) e Banco de Dados (BD), representado na figura 3.1.

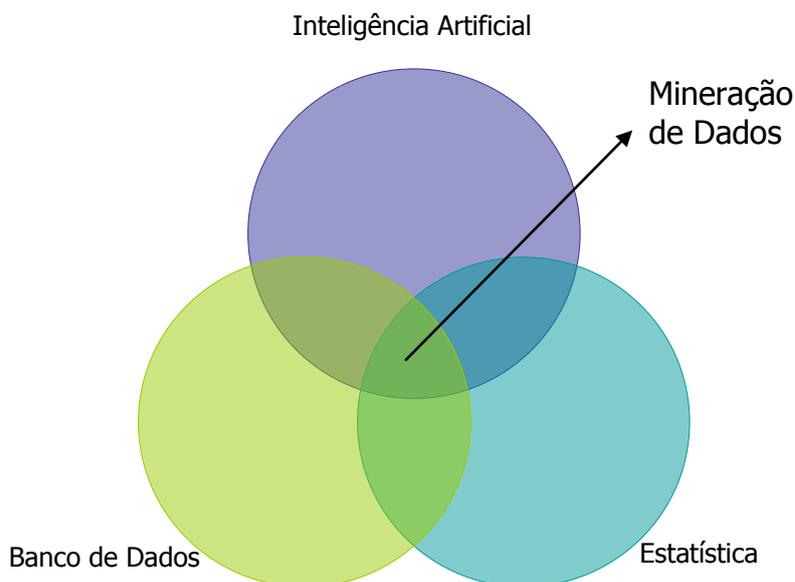


Figura 3.1 – Mineração de Dados: intersecção de várias áreas (adaptada de Silva, 2004)

A respeito das três áreas que contribuem de forma mais significativa para a MD, Zhou (2003, p. 143) reúne definições das diferentes áreas, que explicitam suas perspectivas, retomadas aqui a partir da bibliografia indicada e atualizada. Para a área de BD, MD é o “*processo de descoberta de conhecimento interessante a partir de grandes quantidades de dados armazenados em Bancos de Dados, Data Warehouses ou outros repositórios*” (HAN & KAMBER, 2006, p. 7). Na perspectiva da área de Inteligência Artificial, trata-se da “*extração de informação implícita, previamente desconhecida, e potencialmente útil a partir de dados*” (WITTEN & FRANK, 2005, p. xxiii). Para a estatística, MD é “*a análise de (freqüentemente grande) conjuntos de dados para encontrar relacionamentos inesperados e para resumir os dados em novas formas que sejam compreensíveis e úteis para o proprietário dos dados*” (HAND et al., 2001, p.1).

Minerar grandes bases de dados pode ajudar a localizar informações relevantes sobre características preferenciais de compra – tais como forma de pagamento, interesses por lançamentos ou incidência de vendas casadas –, sobre o que influencia na aquisição ou não de um serviço, ou também, como é objeto deste estudo, a respeito de quais características e/ou comportamentos contribuem para o êxito (ou não) na realização de um curso via Web.

A navegação em um ambiente, como um site de vendas ou um portal de conteúdos, também pode ser customizada a partir da MD. Alguns exemplos de aplicações de MD são descritos por Cunico (2004), nas áreas de comércio, finanças, seguros, medicina, governo, marketing, educação e telecomunicações.

Nesse capítulo, discutem-se as etapas necessárias à descoberta de conhecimento, as principais técnicas de MD utilizadas e aplicações em ambientes educacionais. Além disso, relaciona-se o potencial de agregar os resultados de MD aos sistemas de alertas.

3.1 Etapas da Mineração de Dados

Para minerar dados, é preciso, inicialmente, identificar os problemas que se pretende resolver e, então, definir o objetivo da mineração. Além disso, é importante verificar os dados que estão disponíveis – ou a possibilidade de obter aqueles que são necessários, em tempo adequado – para alcançar o objetivo estabelecido.

Rezende (2005) sintetiza que, durante a identificação dos problemas, é relevante: definir as metas do processo; estabelecer os critérios de desempenho; estabelecer a relação entre simplicidade e precisão do conhecimento extraído; e definir se um conhecimento resultante deve ser compreensível a seres humanos ou se um modelo do tipo caixa-preta é suficiente. Esses pontos interferem nas escolhas posteriores, nas diversas etapas do processo de descoberta de conhecimento.

As etapas para descoberta de conhecimento potencialmente útil estão representadas na figura 3.2, adaptada de Fayyad et al. (1996).

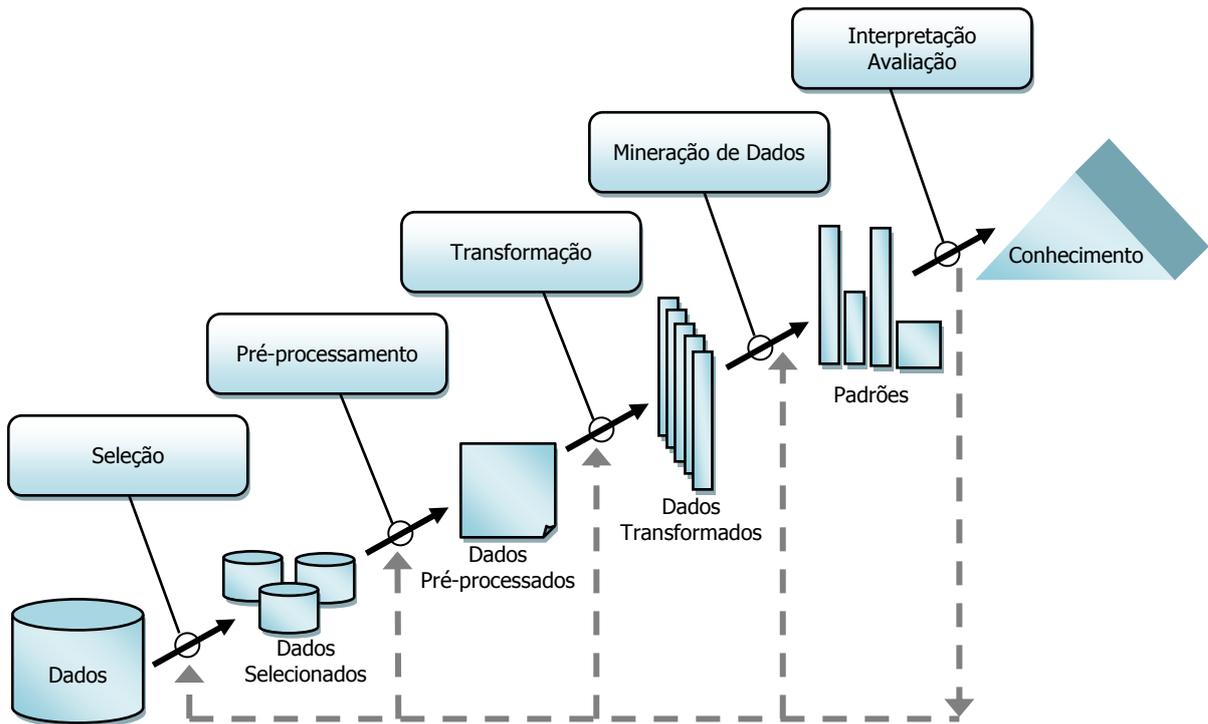


Figura 3.2 – Etapas para Descoberta de Conhecimento (adaptada de Fayyad et al., 1996)

A etapa de *seleção* está relacionada ao estudo dos dados que se tem disponível na base e na relevância desses para a busca de soluções para os problemas identificados. A seleção criteriosa de dados é a base de toda a descoberta de conhecimento.

Após a seleção dos dados, faz-se necessário prepará-los (*pré-processamento*). Essa etapa envolve diferentes fases, em um processo conhecido como qualificação de dados:

- *Padronização*: ajustar os dados a normas pré-definidas, uniformizando a forma de representação dos mesmos, como no caso de caracteres e endereços;
- *Verificação*: corrigir ou retirar dados inconsistentes, tais como cidade e estado que não são compatíveis em um endereço;
- *Enriquecimento*: acrescentar dados a registros incompletos da base de dados, como o DDD a partir da cidade e estado ou mesmo gerar novos campos a partir dos já existentes;

- *Deduplicação*: retirar dados duplicados, tais como mais de um cadastro para uma mesma pessoa;
- *Unificação*: agrupar dados que, originalmente, poderiam estar distribuídos, obtendo uma visão única.

Nesta fase também, dados com valores contínuos podem ser discretizados, ou seja, substituídos por valores nominais que caracterizem faixas de valores. Exemplificando, a partir da observação de faixas de valores de renda de um trabalhador, os valores deste campo podem ser substituídos por baixa, média ou alta; notas de alunos de 0 a 10 podem ser substituídas por intervalos ou conceitos.

Além disso, pode ser desejável reduzir o número de campos de dados a serem minerados. Nesse caso, é possível criar um novo campo que represente as informações de um conjunto, como, por exemplo, um campo que represente a frequência semanal de um usuário em um ambiente virtual ao invés da informação diária armazenada originalmente.

Para estas primeiras fases do processo de descoberta do conhecimento, é comum que se construa um *Data Warehouse* (DW). Um DW é um banco de dados voltado para consultas, armazenando dados históricos consolidados, oriundos de diversas fontes e não operacionais (INMON, 2002). Ferramentas de ETL (*Extract Transform Load*) auxiliam na extração de dados de diversas bases existentes em uma empresa, na transformação dos dados conforme as regras do negócio e na carga dos dados no DW.

A partir do DW, trabalha-se com um repositório de dados único, com todas as informações necessárias para a exploração, já devidamente organizadas. Almeida (2004) ressalta que é importante não confundir MD com ferramentas de consulta OLAP (*On-line Analytical Process*), que facilitam e flexibilizam as consultas a DW multidimensionais, mas não possuem a característica de descobrir informações ocultas no mesmo.

Com os dados pré-processados, é preciso transformá-los (*transformação*) nos formatos adequados aos algoritmos e ferramentas de mineração utilizadas, tais como em arquivos CSV (*Comma-Separated Values*), ARFF (*Attribute-Relation File Format*) ou em outros formatos especificados.

Depois da transformação dos dados, eles estão prontos para serem submetidos às técnicas de mineração propriamente ditas (seção 3.3). As primeiras etapas do processo de KDD, apresentadas até este ponto, representam em torno de 80% do esforço para a descoberta de conhecimento (ZHANG et al., 2003).

A fase de *mineração de dados* constitui-se na principal etapa do processo de descoberta de conhecimento, pois é nela que informações implícitas e potencialmente úteis serão extraídas. A escolha da técnica de MD mais adequada à descoberta de conhecimento depende de aspectos como a área de aplicação e os tipos de dados disponíveis. Conforme a técnica, então, vários parâmetros ainda precisam ser ajustados para obter resultados satisfatórios.

A avaliação dos resultados (*interpretação/avaliação*) envolve analisar o conhecimento gerado e utilizá-lo para auxiliar na solução dos problemas que motivaram a mineração. Portanto, os resultados gerados na mineração poderão ser incorporados por outros sistemas ou disponibilizados diretamente a determinadas pessoas, com o propósito simplesmente de ampliar a base de conhecimento ou, de forma mais importante, de servir de apoio para a tomada de decisão.

Assim, se a origem era descobrir o perfil de bons e maus pagadores em uma empresa, espera-se que o conhecimento gerado seja considerado no momento de liberar o crédito; se o problema era ampliar a venda de determinados produtos, que o resultado indique os fatores que influenciam em suas aquisições, para um trabalho de marketing focado em potenciais clientes; se o problema, em um curso a distância, é a reprovação, que relações entre comportamentos não exitosos possam ser descobertas e, assim, se possa traçar um plano de ação preventivo, que agregue valor à ação pedagógica e diminua os riscos de reprovação.

Cabe destacar que esse processo, sempre que necessário, retorna a fases anteriores, visando a melhorar os resultados (MANNILA, 1996; FAYYAD et al., 1996; BRACHMAN & ANAND, 1996). Portanto, o processo de KDD não é compreendido como linear e único, em que uma vez gerados os resultados estaria concluído de forma irrevogável. O processo de KDD é, de fato, um processo contínuo, cíclico, que possibilita retomar etapas anteriores, ajustá-las e, então, obter novos resultados.

Assim, com frequência, o processo envolve vários momentos de seleção de dados e sua preparação, aplicações de diversas técnicas de mineração e/ou algoritmos, adequando os parâmetros pertinentes, e novas análises dos resultados, tudo com o objetivo de refinar os conhecimentos extraídos.

3.2 Tipos de Tarefas de Mineração de Dados

As tarefas de MD podem ser divididas em dois grandes grupos: a descoberta direta de conhecimento e a indireta (CIOS et al., 1998).

A descoberta direta do conhecimento refere-se às tarefas em que se define um foco, um campo de dado, e se espera que a mineração explique como os demais se relacionam a ele. Busca explicar, por exemplo, como o comportamento de um aluno em um AVA relaciona-se com o seu desempenho em uma dada disciplina; ou como dados de escolaridade, idade, sexo e estado civil relacionam-se com a renda de um sujeito. A mineração tem um objetivo definido.

Entre as principais tarefas de mineração de dados com descoberta direta de conhecimento estão (BERRY & LINOFF, 1997):

- *Classificação*: consiste em, a partir de classes determinadas, associar cada registro de dados minerados a uma delas. Em relação à renda, por exemplo, classificar o sujeito em baixa, média ou alta; quanto ao desempenho, classificar o aluno em aprovado, reprovado ou evadido;
- *Estimativa*: diferentemente da classificação, que trabalha com classes pré-definidas (valores discretos), a estimativa lida com valores numéricos contínuos. A partir da análise dos dados, portanto, busca estimar um valor, como por exemplo, o valor do seguro de um veículo, para um dado cliente,

com certo tipo de carro, em uma dada região; na educação, poderia se buscar estimar a nota final do aluno a partir do desempenho em um grupo de tarefas já entregues;

- *Descrição*: a proposta é encontrar explicações para determinados fatos como, por exemplo, o fato de mulheres mais escolarizadas, em geral, terem menos filhos do que as de escolaridade mais baixa; ou as mulheres terem uma expectativa de vida mais alta do que os homens.

Já a descoberta indireta do conhecimento busca encontrar relações a partir do cruzamento entre todos os dados disponíveis, no intuito de localizar padrões. O propósito é encontrar relações novas e úteis, muitas vezes não previstas. Para tanto, por exemplo, realiza-se a *Clusterização*, que busca distribuir os variados dados analisados (heterogêneos) em grupos com características similares (mais homogêneos), denominados *clusters* (HAN & KAMBER, 2006). É importante destacar que na *Classificação*, as classes são pré-estabelecidas, enquanto na *Clusterização* os grupos são formados de acordo com a similaridade de características, analisadas na MD.

Para fins do presente estudo, considera-se a descoberta direta de conhecimento, com tarefas de classificação dos alunos em classes, conforme seus desempenhos em cursos realizados por meio da Web.

3.3 Técnicas de Mineração de Dados

Nesta seção, são apresentadas algumas das técnicas de MD mais utilizadas, estudadas para a análise de aplicações educacionais de MD e para a realização de experimentos de interesse da tese, a saber: *Árvores de Decisão*, *Regras de Classificação*, *Regras de Associação* e *Clustering*.

3.3.1 Árvore de Decisão

A técnica de *árvore de decisão* adapta-se bem a tarefas de *classificação*, com boa visualização das características que impactam em cada classe. Na fase de aprendizagem, um campo de dados do conjunto é identificado como central (classe)

e todos os demais são analisados sob o prisma de como influenciam na sua classificação (no valor desse campo).

A partir do campo escolhido como dado de saída (classe), que será apresentado nas *folhas* da árvore, o algoritmo localiza outro campo, mais fortemente relacionado à classe, e o coloca na *raiz* da árvore (topo). A partir da raiz, vão sendo colocados novos nós, conectando a raiz às folhas (formando ramos), com os próximos campos mais relacionados, de forma encadeada. Cada nó interno representa um teste sobre um determinado campo, com duas ou mais sub-árvores representando cada resultado possível do teste. Campos não determinantes são desprezados na construção da árvore (BERSON et al., 1999).

O algoritmo para a construção de uma árvore de decisão age recursivamente na fase de aprendizagem, subdividindo os dados até que as folhas sejam classes puras ou que exista um critério de parada especificado, como o número de casos enquadrados. A representação da árvore de decisão pode ser gráfica ou textual, podendo ser traduzida em regras do tipo SE <condição> ENTÃO <classificação>.

Para classificar um determinado registro, depois da árvore montada, parte-se do topo da árvore (raiz), respondendo as perguntas no caminho até chegar a uma folha, onde está a classificação.

3.3.1.1 Exemplo de Árvore de Decisão

Para melhor compreensão, criou-se um exemplo, composto por 13 conjuntos de dados fictícios, que representam dados simplificados de um determinado curso acadêmico (tabela 3.1). Cada conjunto de dados traz o nome de um aluno, o sexo deste, o número de atividades entregues por ele durante o período letivo, o desempenho médio alcançado nas atividades e o seu resultado final.

A tabela 3.1 foi convertida em formato CSV para que os dados pudessem ser submetidos aos algoritmos de mineração que cada técnica explora. O quadro 3.1 apresenta o conteúdo do arquivo gerado.

Tabela 3.1 – Exemplo de Dados para Mineração

Nome	Sexo	Atividades Entregues	Desempenho Médio	Resultado
Ana	F	3	Alto	Aprovado
André	M	1	Alto	Reprovado
Gabriel	M	3	Médio	Aprovado
Helena	F	2	Médio	Reprovado
Isabel	F	2	Baixo	Reprovado
Joana	F	1	Alto	Reprovado
João	M	2	Médio	Reprovado
José	M	3	Baixo	Reprovado
Lucas	M	2	Alto	Aprovado
Maria	F	3	Médio	Aprovado
Matheus	M	3	Alto	Aprovado
Pedro	M	1	Médio	Reprovado
Vera	F	3	Baixo	Reprovado

Quadro 3.1 – Dados em formato CSV gerados a partir da tabela 3.1

```

Nome;Atividades_Entregues;Desempenho_Medio;Sexo;Resultado
Ana;3;Alto;F;Aprovado
André;1;Alto;M;Reprovado
Gabriel;3;Medio;M;Aprovado
Helena;2;Medio;F;Reprovado
Isabel;2;Baixo;F;Reprovado
Joana;1;Alto;F;Reprovado
João;2;Medio;M;Reprovado
José;3;Baixo;M;Reprovado
Lucas;2;Alto;M;Aprovado
Maria;3;Medio;F;Aprovado
Matheus;3;Alto;M;Aprovado
Pedro;1;Medio;M;Reprovado
Vera;3;Baixo;F;Reprovado

```

Para gerar a árvore, a partir dos dados do arquivo CSV (quadro 3.1), utilizou-se o algoritmo *DecisionTree*, implementado na ferramenta de MD *RapidMiner* (seção 3.4.1), similar ao C4.5 (QUINLAN, 1992). O algoritmo ID3, que só trabalhava com valores nominais (discretos), serviu de base para o desenvolvimento do algoritmo C4.5, que trabalha também com dados numéricos (contínuos).

O algoritmo escolhido analisa os diversos campos, de forma iterativa, buscando identificar aquele que mais influencia nos valores das classes, que serão as *folhas* da árvore, valores esses presentes em um campo definido que deve ser informado previamente. No exemplo, as classes são *aprovado* e *reprovado* do campo *resultado*. O campo de maior influência para o valor do *resultado* é colocado no topo da árvore (*raiz*) e, então, o algoritmo segue buscando novos campos significativos para colocar nos *nós* abaixo, até que seja possível determinar as classes.

A partir da aplicação da técnica de árvore de decisão, por meio do algoritmo *DecisionTree*, foi gerada a árvore apresentada na figura 3.3. É possível observar que o fator mais importante para definir o sucesso do aluno no curso foi o número de atividades entregues e, na seqüência, o desempenho médio nas mesmas. O sexo não foi determinante e, por isso, foi desconsiderado pelo algoritmo na geração da árvore.

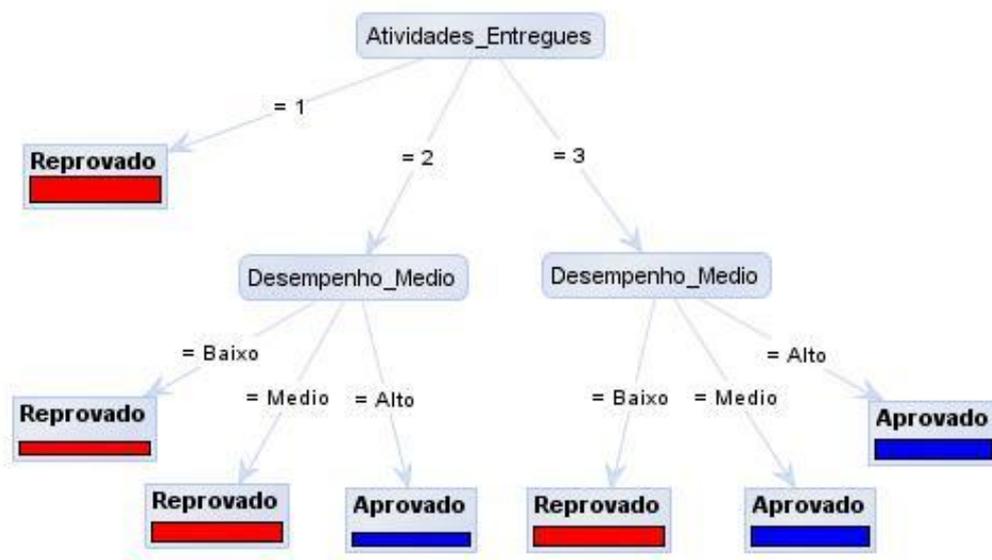


Figura 3.3 – Árvore de Decisão gerada a partir dos dados da tabela 3.1

A partir da árvore, é possível utilizar o conhecimento gerado para classificar outros alunos em um curso com as mesmas características. O quadro 3.2 mostra a árvore convertida em regras de decisão, também geradas pelo algoritmo *DecisionTree*, adaptadas e reordenadas aqui para facilitar a compreensão do exemplo.

Quadro 3.2 – Regras de Decisão obtidas a partir da Árvore de Decisão da figura 3.3

SE Atividades_Entregues = 1 ENTÃO Reprovado
SE Atividades_Entregues = 2 ENTÃO
SE Desempenho_Medio = Baixo ENTÃO Reprovado
SE Desempenho_Medio = Medio ENTÃO Reprovado
SE Desempenho_Medio = Alto ENTÃO Aprovado
SE Atividades_Entregues = 3 ENTÃO
SE Desempenho_Medio = Baixo ENTÃO Reprovado
SE Desempenho_Medio = Medio ENTÃO Aprovado
SE Desempenho_Medio = Alto ENTÃO Aprovado

3.3.2 Regras de Classificação

Um algoritmo de classificação tem por objetivo encontrar algum relacionamento entre os atributos e uma classe, gerando uma regra. Assim, o processo de classificação pode, posteriormente, utilizar essa regra para prever a classe de um novo registro.

Uma *regra de classificação* inferida de um *conjunto de treinamento* c_t — $\{atributo_{c_t}\} \cup atributo_de_classificac\tilde{a}o_o_c_t = \emptyset$ — é representada na forma *SE <condição> ENTÃO <classificação>*. Uma <condição> é um conjunto de termos <atributo_{c_t} θ valor>, $\theta = \{ '=', '<', '>', '\leq', '\geq', '\neq' \}$, e <classificação> é um termo <atributo_{de_classificac\tilde{a}o_c_t} θ valor>. Um valor do atributo de classificação é chamado de *classe* e <condição> é também chamada de *antecedente* da regra.

Uma diferença entre a geração de regras por algoritmos de *regras de classificação* e a geração por algoritmos de *árvores de decisão* é a restrição de que toda regra obtida a partir de uma árvore tenha o atributo raiz da árvore em seu

antecedente. Outra característica importante é que os algoritmos para geração de regras de classificação, para evitar conflito na aplicação das regras de classificação geradas, devem apresentá-las priorizando a qualidade das mesmas. A qualidade da regra pode levar em consideração (HAN & KAMBER, 2006): a *precisão*, dada pela relação entre o número de casos classificados corretamente sobre o número total de casos em que as condições apresentadas no antecedente da regra são verdadeiras; a *cobertura*, que é o número de casos classificados corretamente sobre o número total de casos analisados; e seu *comprimento*, que diz respeito à quantidade de atributos testados no antecedente da regra.

A ordem de apresentação das regras estabelece uma lista de decisão, a ser aplicada em seqüência. A regra que aparece primeiro na lista tem maior prioridade para predizer a classe. Quando um registro é classificado, nenhuma outra regra posterior de classificação será aplicada sobre ele.

3.3.2.1 Exemplo de Regras de Classificação

A técnica de regras de classificação foi aplicada aos dados da tabela 3.1, no formato CSV (quadro 3.1). Foi utilizado o algoritmo *RuleLearner*, implementado na ferramenta de MD *RapidMiner* (seção 3.4.1), que funciona de forma similar ao algoritmo RIPPER (*Repeated Incremental Pruning to Produce Error Reduction* – COHEN, 1995), tendo como *critério* a *accuracy* (precisão) para a geração das regras. O resultado é apresentado no quadro 3.3, de forma adaptada para facilitar a compreensão do exemplo.

Quadro 3.3 – Regras de Classificação geradas a partir dos dados da tabela 3.1

SE Atividades_Entregues = 1 ENTÃO Reprovado (0 / 3)
SE Desempenho_Medio = Alto ENTÃO Aprovado (3 / 0)
SE Atividades_Entregues = 2 ENTÃO Reprovado (0 / 3)
SE Desempenho_Medio = Medio ENTÃO Aprovado (2 / 0)
SENÃO Reprovado (0 / 2)
Correto: 13 dos 13 exemplos de treinamento.

O modelo de conhecimento gerado por meio de regras de classificação deve ser aplicado na seqüência em que se apresenta, com testes encadeados, para a classificação de novos casos. Quando um novo caso se enquadra na primeira regra, conforme o exemplo (quadro 3.3), ele é classificado como Reprovado. Caso não se encaixe na primeira regra, as condições das próximas regras são testadas para buscar enquadrá-lo em um dos padrões identificados. Senão atender a nenhuma das condições, pelo modelo, estará Reprovado.

Ao lado direito de cada regra há dois parâmetros: o primeiro refere-se ao número de casos do conjunto exemplo que se enquadram na regra e estão aprovados; o segundo indica ao número de casos que estão reprovados. No exemplo, o conjunto de regras geradas permite classificar corretamente todos os exemplos utilizados, o que em bases de dados volumosos nem sempre é possível.

3.3.3 Regras de Associação

Algoritmos para a geração de *regras de associação* têm por objetivo encontrar relações entre dados que ocorrem com determinada freqüência e que possam, então, ser utilizados para identificar padrões de comportamento. Uma das aplicações mais comuns desta técnica está na área do comércio, procurando identificar padrões de consumo que possam orientar ações de vendas, como, por exemplo, {estojo} \Rightarrow {lápis}. Essa regra quer dizer que quem compra estojo, em uma livraria, tende a comprar também lápis.

Agrawal et al. (1993) representa as *regras de associação* a partir do seguinte formalismo: $\Gamma = \{ i_1; i_2; i_m \}$ é um conjunto de m itens distintos e D uma base de dados formada por um conjunto de transações, onde cada transação T é composta por um conjunto de itens (*itemset*), tal que $T \subseteq \Gamma$. Uma regra de associação é uma expressão na forma $A \Rightarrow B$, onde $A \subset \Gamma$, $B \subset \Gamma$ e $A \cap B = \phi$.

A é denominado *antecedente* e B denominado *conseqüente* da regra. Tanto o *antecedente*, quanto o *conseqüente* de uma regra de associação podem ser formados por conjuntos contendo um ou mais itens. A quantidade de itens pertencentes a um conjunto de itens é chamada de *comprimento* do conjunto.

Cada regra de associação gerada por esta técnica ($A \Rightarrow B$) possui dois parâmetros:

- *Confiança* que é resultado obtido pelo número de vezes em que a A e B aparecem em uma mesma transação em relação ao número de vezes em que A aparece no conjunto de transações;
- *Suporte* indica o número de ocorrências da regra ($A \Rightarrow B$) no conjunto das transações.

Esses parâmetros interferem diretamente na quantidade e na qualidade das regras geradas e, portanto, é fundamental ajustá-los, no processo de MD, para a geração de regras de associação significativas para a análise. Por exemplo, ao minerar um grande conjunto de dados, poderão surgir muitas regras e pouco será útil uma regra com 100% de *confiança*, mas apenas 1 caso de *suporte*, o que pode indicar uma situação muito particular ou mesmo única, que possivelmente não possa ser generalizada. Em grandes volumes de dados, é preciso definir o número mínimo de casos em que as regras se aplicam (*suporte*), além de um valor mínimo de *confiança*.

3.3.3.1 Exemplo de Regras de Associação

Como exemplo, a tabela 3.1, foi minerada com o algoritmo de geração de regras de associação implementado na ferramenta *WizRule* (seção 3.4.2). Como a base exemplo contém apenas 13 casos, foi definido um mínimo de 2 casos de *suporte*, com 95% de *confiança*. O resultado é apresentado no quadro 3.4, visualmente organizado, de forma a facilitar a compreensão do exemplo. Para cada regra gerada, o *WizRule* apresenta valores como a *confiança* (1 equivalente a 100%) e o número de casos em que a regra se aplica (*suporte*).

Quadro 3.4 – Regras de Associação geradas a partir dos dados da tabela 3.1

- 1) **If Atividades_Entregues is 3,00 and Resultado is Reprovado**
Then Desempenho_Medio is Baixo
Rule's probability: 1,000
The rule exists in 2 records.
- 2) **If Desempenho_Medio is Alto and Resultado is Reprovado**
Then Atividades_Entregues is 1,00
Rule's probability: 1,000
The rule exists in 2 records.
- 3) **If Atividades_Entregues is 3,00 and Desempenho_Medio is Alto**
Then Resultado is Aprovado
Rule's probability: 1,000
The rule exists in 2 records.
- 4) **If Atividades_Entregues is 3,00 and Desempenho_Medio is Medio**
Then Resultado is Aprovado
Rule's probability: 1,000
The rule exists in 2 records.
- 5) **If Atividades_Entregues is 1,00**
Then Resultado is Reprovado
Rule's probability: 1,000
The rule exists in 3 records.
- 6) **If Desempenho_Medio is Baixo**
Then Resultado is Reprovado
Rule's probability: 1,000
The rule exists in 3 records.
- 7) **If Desempenho_Medio is Medio and Resultado is Aprovado**
Then Atividades_Entregues is 3,00
Rule's probability: 1,000
The rule exists in 2 records.
- 8) **If Sexo is F and Resultado is Aprovado**
Then Atividades_Entregues is 3,00
Rule's probability: 1,000
The rule exists in 2 records.
- 9) **If Atividades_Entregues is 2,00 and Desempenho_Medio is Medio**
Then Resultado is Reprovado
Rule's probability: 1,000
The rule exists in 2 records.
- 10) **If Atividades_Entregues is 2,00 and Sexo is F**
Then Resultado is Reprovado
Rule's probability: 1,000
The rule exists in 2 records.

Foram geradas 10 regras de associação, em que cada uma é independente da outra. As regras de associação, diferentemente das regras de classificação, não tem uma saída padrão, mas buscam descrever relações. A regra um, por exemplo, permite interpretar que se o número de atividades foi 3 e o resultado foi reprovado, então foi porque o desempenho foi baixo. As regras 8 e 10 trazem o campo sexo com um de seus componentes, conforme *confiança* e *suporte* estipulados, o que nas regras de classificação sempre foi desconsiderado.

É importante ressaltar que as regras de associação são mais comumente utilizadas para identificar padrões de ocorrência entre itens homogêneos, como, por exemplo, padrões de consumo de livros: as pessoas que compram determinado livro também costumam comprar um outro título específico. Em dados educacionais, utilizar técnicas de MD para a obtenção de regras de associação pode ser interessante para identificar padrões de navegação – quem acessa tais materiais ou recursos costuma acessar também tais outros – ou de êxito na realização de atividades – quem tem bom desempenho em determinadas tarefas, também costuma ter bom desempenho em certas outras / quem erra certas questões, também costuma errar tais outras.

3.3.4 Clusterização

A *clusterização* busca descobrir conhecimento de forma indireta, a partir da identificação de grupos de dados com características semelhantes. O objetivo desta técnica consiste em identificar agrupamentos de dados que podem ser classificados em uma classe comum, descoberta no processo de *clusterização*. Para a formação dos *clusters*, podem ser utilizados diversos algoritmos, com base estatística, procurando estabelecer os elementos *centróides* de cada *cluster* (BERRY & LINOFF, 1997).

Para realizar a *clusterização*, é preciso estabelecer critérios apropriados para a classificação de um conjunto de dados em subconjuntos, sendo os mais comuns os de *homogeneidade* e *separação*. A *homogeneidade* é medida no interior de um mesmo *cluster*, verificando o grau de similaridade; já a *separação* refere-se a

elementos de *clusters* diferentes, medindo o quão diversos são entre si (MARAVALLE et al., 1997).

3.3.4.1 Exemplo de Clusterização

Com o objetivo de exemplificar a criação automática de agrupamentos (*clusters*), utilizou-se o algoritmo *KMedoids*, implementado na ferramenta *RapidMiner* (seção 3.4.1), em que se define o número de *clusters* (parâmetro *K*) desejado e a medida de cálculo de distância entre os elementos a ser utilizada (parâmetro *measure*) para a geração dos agrupamentos.

A figura 3.4 mostra os agrupamentos gerados a partir dos dados da tabela 3.1, convertidos em CSV (quadro 3.1), considerando *K* com valores 2, 3 e 4, respectivamente, e *measure* como *NominalDistance*, que toma como base o cálculo da distância nominal (seção 3.3.4.2). Algoritmos de clusterização podem ser utilizados em aplicações educacionais, por exemplo, para formação de grupos de trabalho.

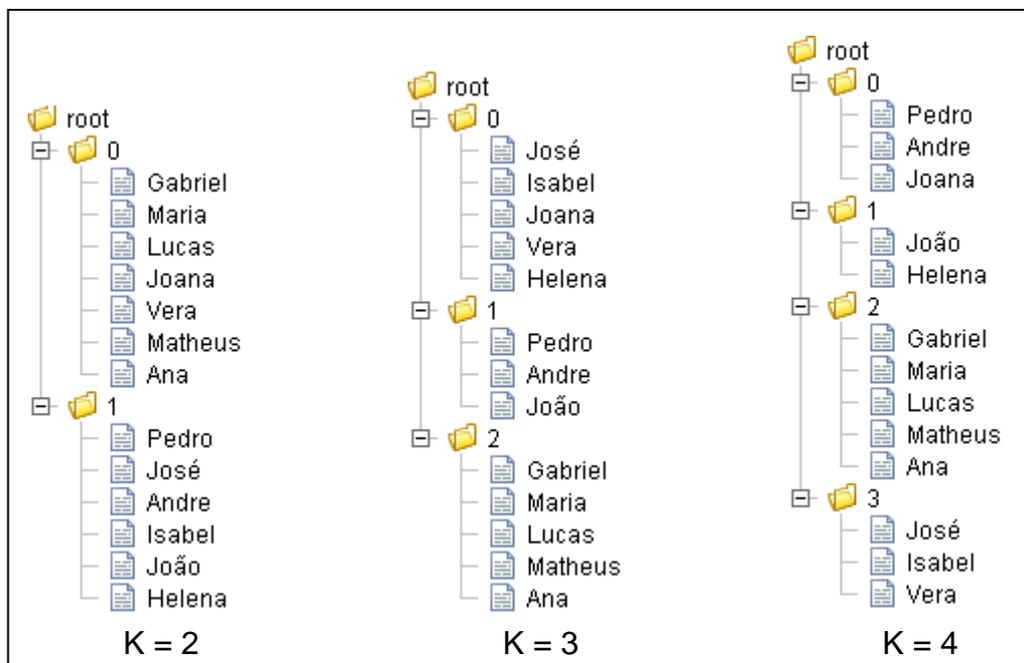


Figura 3.4 – Clusters gerados a partir dos dados da tabela 3.1

Como a *clusterização* é uma técnica de descoberta indireta de conhecimento, nem sempre os agrupamentos gerados são compreensíveis ao olhar humano, já que são realizados a partir da análise estatística dos diversos atributos e não de um atributo específico (BERRY & LINOFF, 1997). Na medida em que cresce o volume de dados e o número de atributos, a análise fica ainda mais complexa.

Quando se trabalha com valores contínuos nos atributos, uma medida de distância bastante difundida baseia-se no cálculo de distância euclidiana (distância entre dois pontos, dada pela aplicação da fórmula de Pitágoras). Para a geração dos exemplos, com atributos de valores discretos, utilizou-se o cálculo da distância nominal.

3.3.4.2 Distância Nominal

Para aplicar o cálculo da distância nominal, cada valor possível de um atributo, ou seja, cada *categoria*, torna-se um número. Como exemplo, considerando o atributo *Desempenho Médio* da tabela 3.1, observa-se que os valores possíveis são as categorias *Baixo*, *Médio* e *Alto*, portanto três. Assim, o atributo *Desempenho Médio* será representado, respectivamente, como 1, 2 ou 3. A aluna Ana (conforme a tabela 3.1) tem *Alto* desempenho médio e, então, o valor correspondente é 3.

A distância nominal final entre os registros (no exemplo, entre dois alunos) é uma função (soma) entre os resultados obtidos pela comparação dos valores de cada atributo. A comparação de cada atributo resulta em zero quando os valores são iguais e 1 quando são diferentes.

Há algoritmos que geram tabelas e grafos mapeando o grau de similaridade ou distância de cada nodo em relação aos demais. Para exemplificar como se dá a visualização dos resultados, utilizou-se o operador *ExampleSet2Similarity*, implementado na ferramenta *RapidMiner* (seção 3.4.1). Para gerar a distância de cada nodo em relação aos demais, é preciso definir o tipo de medida que se pretende utilizar (parâmetro *measure*). Para o exemplo realizado, utilizou-se, então, a medida *NominalDistance*, explicada nesta seção. Parte da tabela gerada é apresentada na tabela 3.2.

Tabela 3.2 – Tabela de distância nominal gerada a partir dos dados da tabela 3.1 (aluna Ana)

First	Second	Distance
Ana	Andre	3
Ana	Gabriel	2
Ana	Helena	3
Ana	Isabel	3
Ana	Joana	2
Ana	José	3
Ana	João	4
Ana	Lucas	2
Ana	Maria	1
Ana	Matheus	1
Ana	Pedro	4
Ana	Vera	2

A partir da análise da tabela 3.2, que apresenta os grau de distância entre os nodos (alunos), considerando os dados referentes à aluna Ana, é possível verificar que a aluna em questão tem como colegas mais próximos a Maria e o Matheus, com grau 1 de distância, sendo que todos entregaram 3 atividades e foram aprovados, e, além disso, Maria é do sexo feminino, como Ana, e Matheus tem desempenho Alto, como Ana. Maria difere de Ana pelo desempenho médio (Ana tem alto desempenho), já Matheus, pelo fato de ser do sexo masculino. Como colegas mais distantes, Ana tem o Pedro e o João, com grau 4, ambos homens, reprovados, com menos de três atividades entregues e desempenho médio nas mesmas. A representação visual do grafo completo é apresentada na figura 3.5.

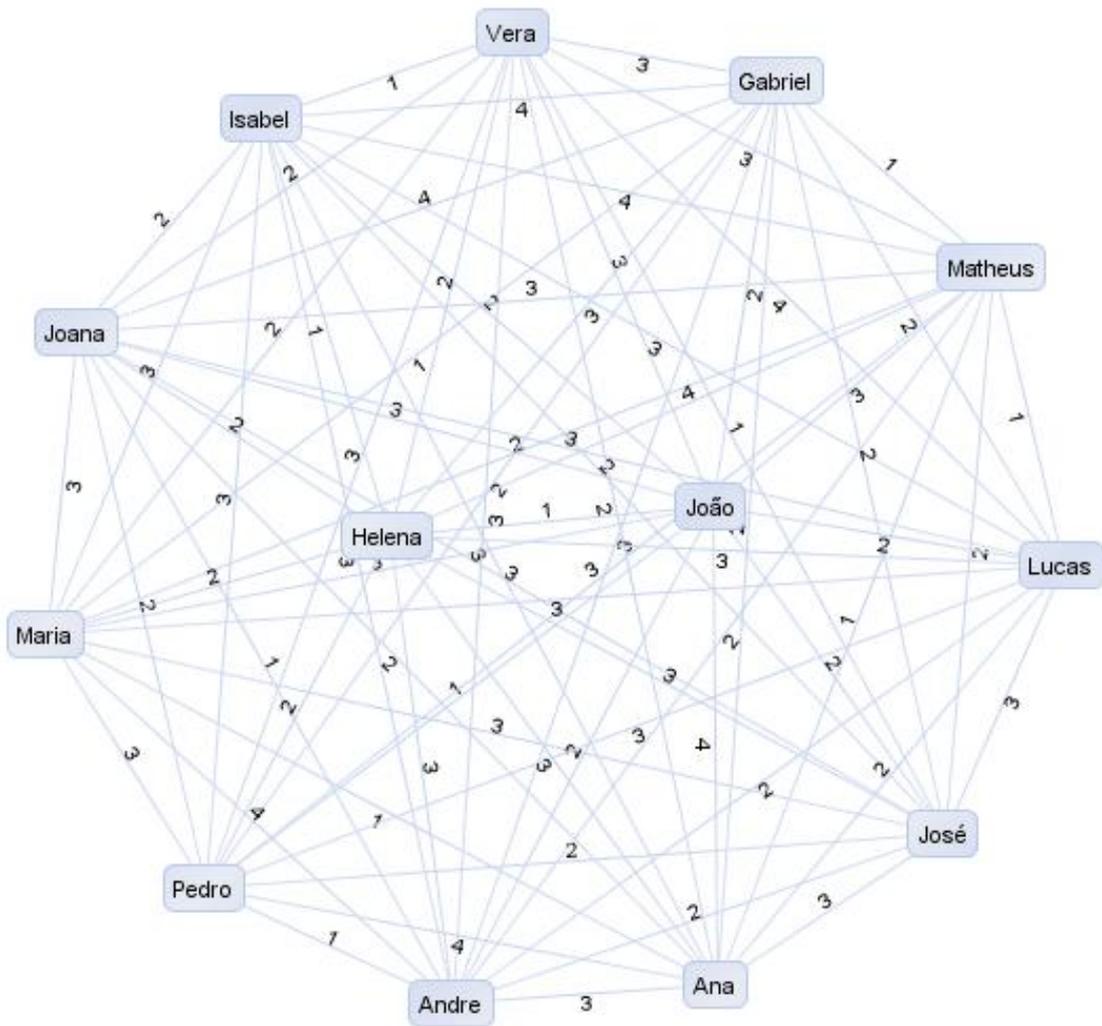


Figura 3.5 – Grafo de distância nominal gerado a partir dos dados da tabela 3.1

3.4 Ferramentas de Mineração

Com o crescente número de informações digitais disponíveis, também cresce o interesse na descoberta de conhecimento implícito no cruzamento dessas informações. Para tanto, multiplicam-se ferramentas que auxiliam nas diferentes etapas do processo de KDD e, em especial, na etapa de MD.

Tais ferramentas surgem tanto como aplicações comerciais – como o Oracle Data Mining⁵, o IBM DB2 Intelligent Miner⁶ e o WizRule⁷ – quanto com código aberto

⁵ Disponível em: http://www.oracle.com/global/br/solutions/business_intelligence/data-mining.html. Acessado em 30/10/2009.

⁶ Disponível em: <http://www-01.ibm.com/software/data/iminer/>. Acessado em 30/10/2009.

⁷ Disponível em: <http://www.wizsoft.com/default.asp?win=8&winsub=8>. Acessado em 30/10/2009.

– como o WEKA⁸ e o RapidMiner⁹ –, desenvolvidos a partir de pesquisa, explorando diversas tarefas, técnicas e algoritmos de mineração. Espera-se que o desenvolvimento de ferramentas de MD leve em conta não apenas os algoritmos implementados e os ajustes de parâmetros necessários, mas também que melhore os tempos de processamento e apresente os resultados de maneira a facilitar a interpretação do usuário.

Para o desenvolvimento das pesquisas realizadas para a presente tese, foram utilizadas duas ferramentas, cujas características são apresentadas nesta seção.

3.4.1 *RapidMiner*

A ferramenta *RapidMiner* é utilizada em mais de 30 países, tanto por companhias comerciais quando em atividades acadêmicas de pesquisa e ensino (RAPIDMINER, 2009). Cobrindo uma ampla faixa de tarefas de mineração, oferece diversos algoritmos para as técnicas estudadas durante o desenvolvimento da tese.

O ambiente *RapidMiner* foi desenvolvido pela Unidade de Inteligência Artificial da Universidade de *Dortmund* (Alemanhã), que lançou sua primeira versão em 2002. Desenvolvido em Java, em 2004, uma versão de código livre e aberto, sob a Licença Pública Geral (*General Public License* – GPL) foi disponibilizada no *SourceForge* (2009). A partir de então, muitas extensões, produzidas por desenvolvedores externos, tem sido incorporadas à ferramenta. (RAPIDMINER, 2009).

O *RapidMiner* trabalha a partir da combinação de operadores, demandando que o seu utilizador inclua aqueles necessários à tarefa que pretende realizar. A figura 3.6 apresenta uma tela do *RapidMiner*. À esquerda, na raiz do processo (*Root Process*), é possível observar que foi incluído um operador de entrada (*Input*), com os dados a serem minerados – para este exemplo, utilizou-se o arquivo CSV apresentado no quadro 3.1.

Na figura 3.6, à esquerda, há também um operador que implementa o algoritmo *DecisionTree* – para selecioná-lo, ao centro da tela, observa-se o caminho

⁸ Disponível em <http://www.cs.waikato.ac.nz/ml/weka/>. Acessado em 30/10/2009.

⁹ Disponível em: <http://rapid-i.com/content/view/123/141/lang,en/>. Acessado em: 30/10/2009.

percorrido: novo operador (*New Operator*), de aprendizagem (*Learner*), supervisionada (*Supervised*), árvore (*Tree*), e, entre os nove algoritmos ali disponíveis, na direita, em cinza mais escuro, está selecionado o *DecisionTree*.

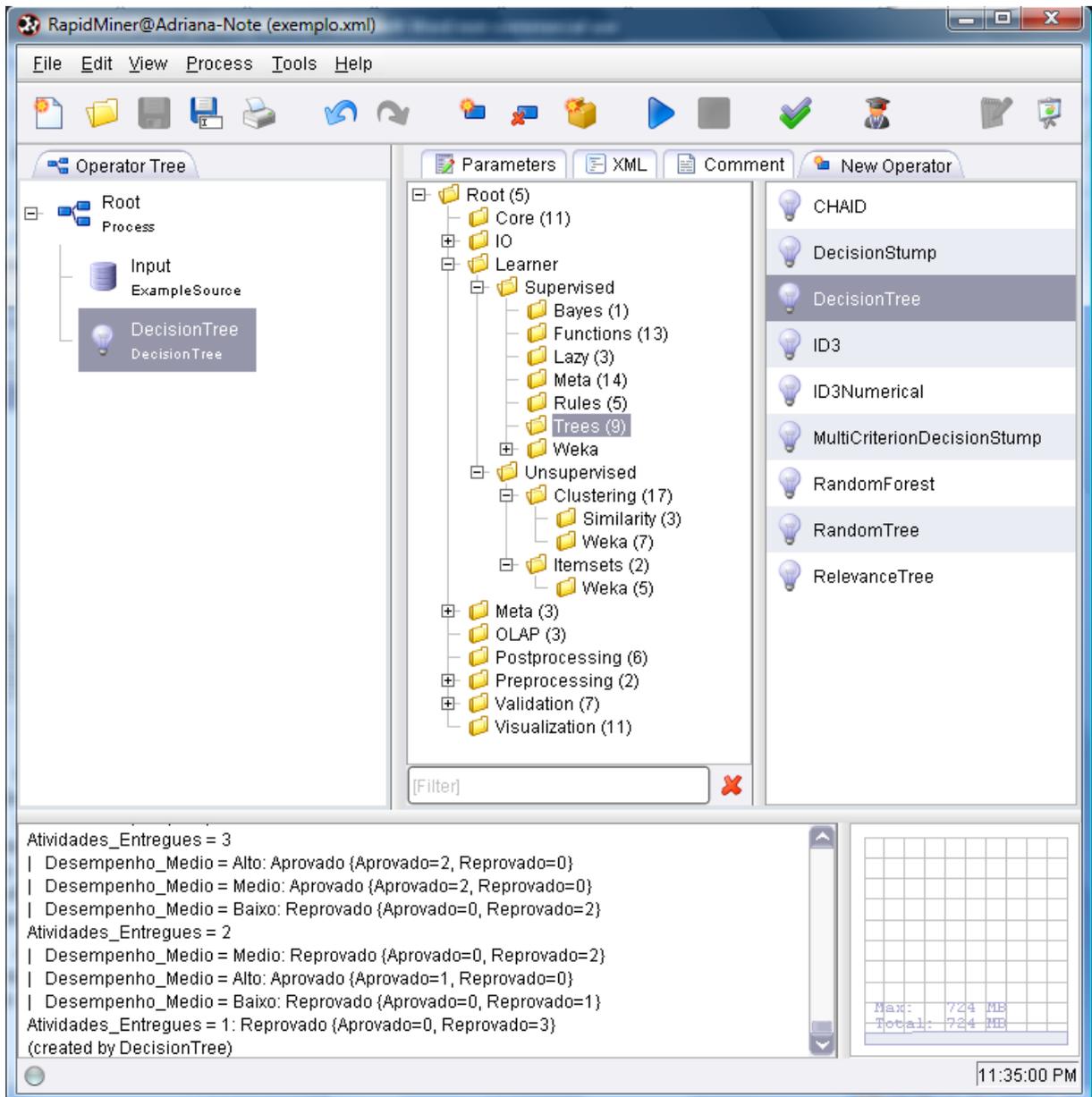


Figura 3.6 – RapidMiner: exemplo de montagem da árvore de operadores para MD

Cabe ressaltar que para cada operador incluído em um processo, há parâmetros pertinentes que devem ser ajustados. Após a finalização da escolha dos operadores e ajustes dos parâmetros, roda-se o processo (triângulo azul da barra de ferramentas) e, então, é possível visualizar os resultados da MD, no caso as regras geradas e a representação gráfica da árvore de decisão (figura 3.7).

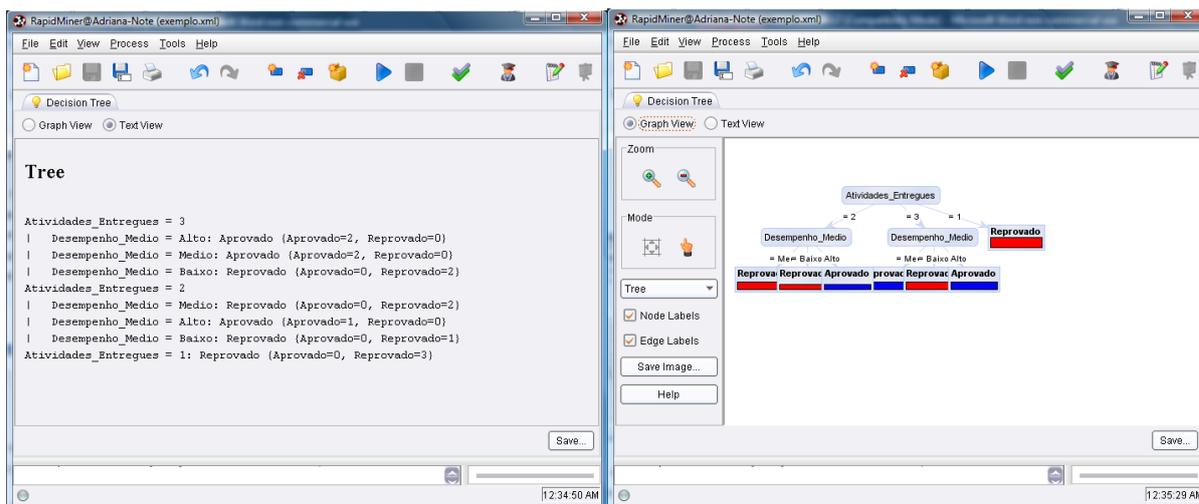


Figura 3.7 – RapidMiner: exemplo de visualização de resultados da MD

A árvore gerada neste processo em específico pode ser reorganizada para melhorar sua visualização, como se encontra na figura 3.3. As regras geradas podem ser visualizadas também na parte inferior da tela da escolha dos operadores (figura 3.6).

Com um grande número de operadores (mais de 400) referentes às diversas etapas de KDD (entrada e saída de dados e aprendizagem de máquina, entre outros), o *RapidMiner* permite combiná-los para a realização de experimentos em MD (MIERSWA et al., 2006). A combinação desses operadores flexibiliza o processo de escolha das tarefas, técnicas e algoritmos, bem como o ajuste de parâmetros e a visualização dos resultados gerados.

Como conseqüência dessa forma de manipulação do ambiente, quem o utiliza pode, com mais agilidade, gerar diversos experimentos e analisar quais algoritmos e parâmetros são mais adequados para o problema que pretende tratar. O *RapidMiner* foi utilizado, durante os estudos da tese, para a MD referente às técnicas de árvore de decisão, regras de classificação e clusterização.

3.4.2 *WizRule*

O *WizRule* é uma *software* para MD que gera regras do tipo SE <antecedente> ENTÃO <conseqüente>, com base em métodos estatísticos. Durante a realização dos experimentos da tese, mais especificamente na geração de regras

de associação, foi utilizada uma versão de demonstração, que trabalha com até mil registros (WIZRULE, 2009).

Como entrada de dados, é possível selecionar um arquivo no formato ASCII, dBase, MSAccess, ODBC ou OLE DB. Após, devem ser realizados os ajustes nos parâmetros necessários (como a *confiança* e o *suporte* da regra, entre outros), conforme itens apresentados na figura 3.8.

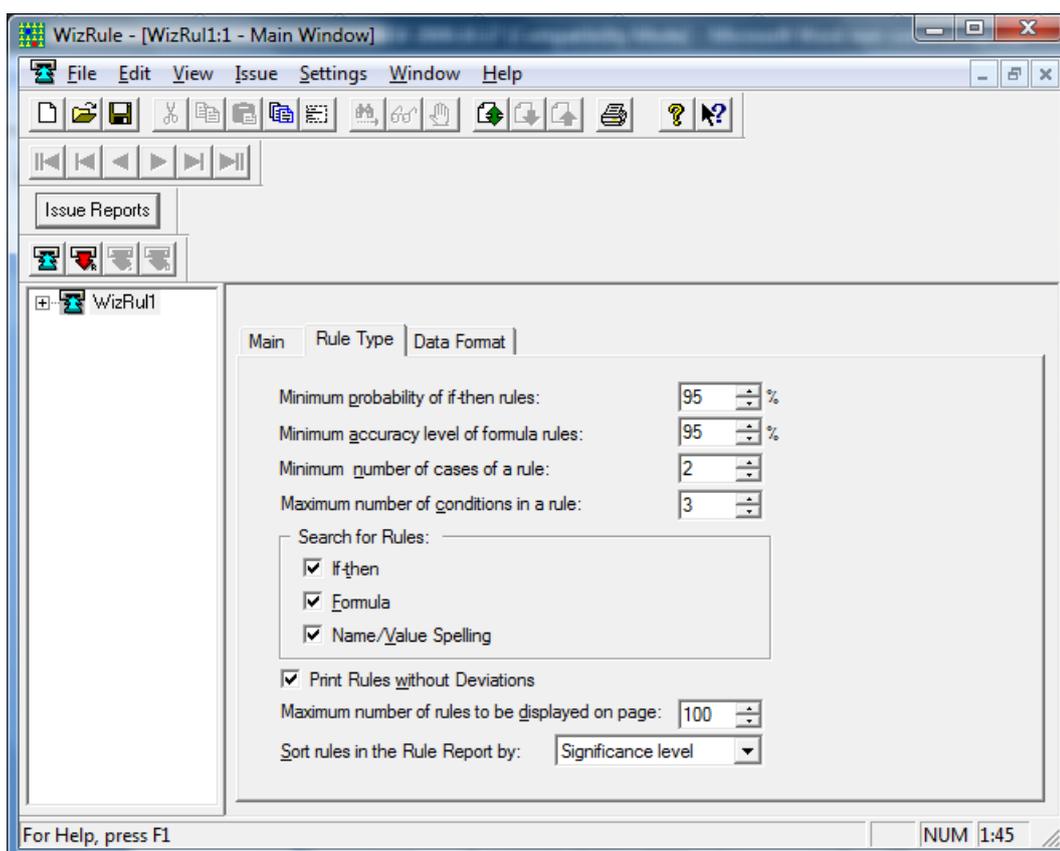


Figura 3.8 – WizRule: configuração dos parâmetros para geração das regras

Os resultados gerados são regras e informações complementares, tais como o *suporte* da regra específica no conjunto de registros, visualizados na figura 3.9.

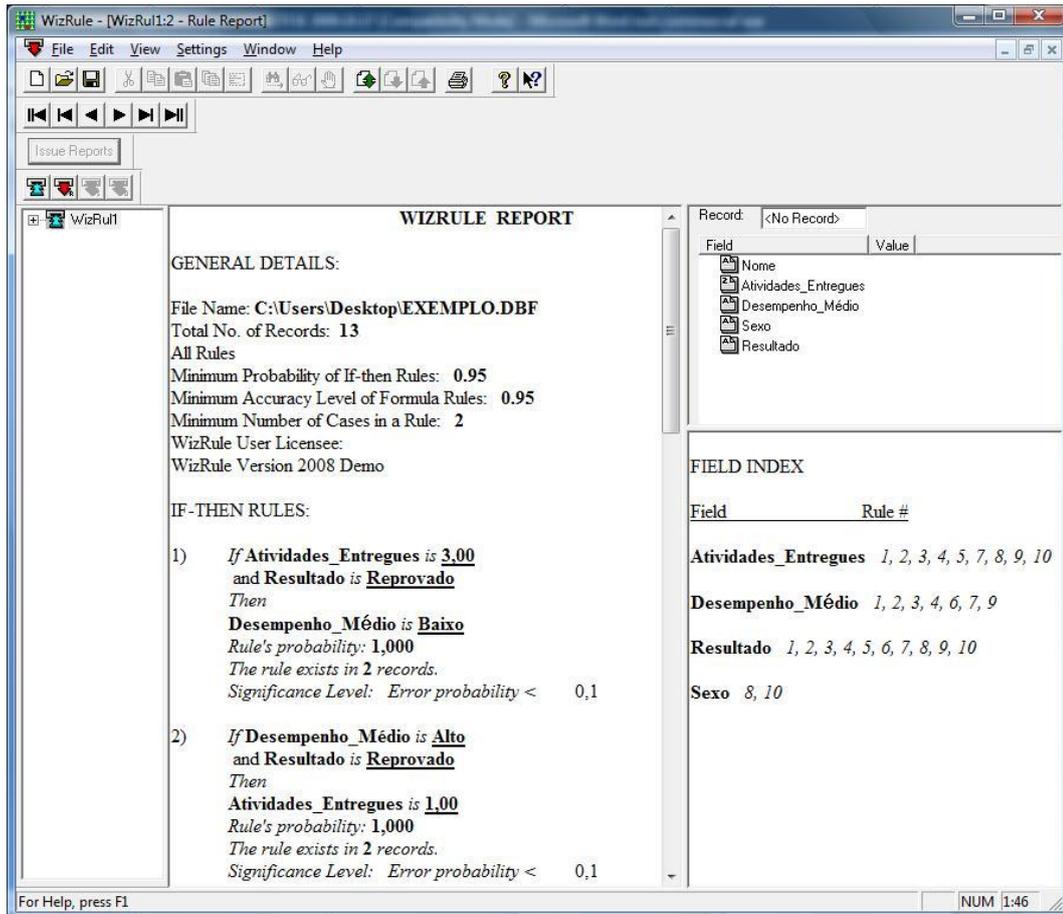


Figura 3.9 – WizRule: exemplo de apresentação das regras geradas

3.5 Aplicações de Mineração de Dados Educacionais

A Educação a Distância (EAD) tem um número cada vez maior de alunos. Imensos volumes de dados são gerados pela interação de usuários (professores e alunos) em Ambientes Virtuais de Aprendizagem (AVA), sendo sistematicamente armazenados e, em alguns casos, recuperados em forma de relatórios. Mas, freqüentemente, dados de *logs*, detalhados e extensos, são difíceis de interpretar e acabam sendo deixados de lado no cotidiano pedagógico. Portanto, torna-se importante utilizar mecanismos capazes de auxiliar na análise desses dados de maneira automatizada, buscando descobrir conhecimento que possa ajudar na resolução de problemas educacionais.

Técnicas de Mineração de Dados podem ser aplicadas a dados gerados em ambientes virtuais, para encontrar relações entre os dados disponíveis, visando a, por exemplo (GARCIA et al., 2007): otimizar os conteúdos em um portal educacional por meio da descoberta dos conteúdos que mais interessam aos usuários; descobrir relações interessantes entre a forma como os estudantes utilizam as informações disponíveis, de forma a prover *feedback* para o autor do curso; extrair padrões úteis para ajudar educadores e desenvolvedores de materiais a avaliar e interpretar as atividades de um curso *on-line*, as formas como são executadas e seus resultados; localizar erros estudantis que ocorrem freqüentemente juntos; encontrar atributos que permitam identificar padrões de diferenças de desempenho entre vários grupos de estudantes; guiar automaticamente as atividades dos alunos, gerando e recomendando materiais; construir agentes de recomendação para atividades de aprendizagem *on-line*; personalizar o ensino virtual com base na agregação de perfis de usuários e ontologias de domínio.

Processos de MD podem ser empregados, também, para descobrir características e comportamentos em alunos que indiquem risco de evasão ou reprovação e, então, essa descoberta pode contribuir para a atuação docente, de forma a evitar esses resultados indesejados. Utilizar técnicas de MD, portanto, possibilita identificar padrões de acesso, de realização de atividades e de interação dos alunos que os levam a obter êxito (ou não) e, dessa forma, oferecer embasamento para a construção de ferramentas que auxiliem na prática docente, buscando a redução dos índices de evasão e reprovação.

Na mineração de dados educacionais, a busca de padrões é realizada de forma que não são identificados os alunos, mas são identificadas suas características. A mineração pode ser feita sobre toda uma base de alunos matriculados nas disciplinas a distância de um dado curso ou apenas de uma disciplina específica, por exemplo. Para estabelecer os dados a serem minerados, é preciso tomar como base o problema que se pretende resolver, além dos dados que se tem à disposição. De forma iterativa, é preciso realizar as etapas de KDD até que os resultados possam auxiliar no tratamento dos problemas que originaram a busca de conhecimento. Esta seção contribui nesta direção, apresentando trabalhos na área de MD Educacionais.

Machado & Becker (2002) realizaram um estudo de caso com a utilização de técnicas de MD aplicadas à web (*Web Mining*), analisando comportamentos de usuários em AVA (suas interações com o ambiente) para localizar padrões em suas navegações. O objetivo era gerar um modelo de descoberta de padrões de navegação *web* em ambientes virtuais, para que, a partir da análise desses padrões, professores e desenvolvedores de sites educacionais pudessem verificar se a estrutura do curso está adequada ou modificações na forma de apresentação dos recursos oferecidos seriam necessárias. A MD, realizada após o curso, gera resultados que podem ser utilizados na reestruturação do mesmo curso para uma nova turma, não gerando intervenções em um curso em andamento.

Com o mesmo objetivo do trabalho descrito acima, Romero et al. (2008) desenvolveu uma ferramenta para minerar e apresentar visualmente dados de navegação de usuários. A partir dos dados previamente preparados, a MD é utilizada para descobrir padrões de seqüências de trilhas visitadas. Romero et al. (2008) coloca que o conceito de trilha está relacionado ao planejamento do percurso de aprendizagem, sendo uma rota para os alunos seguirem um dado elemento do currículo, como um assunto. Como resultado da MD, é apresentado um grafo em que as páginas *web* são os nodos e as linhas (com seta de direção) indicam o sentido do acesso (navegação) entre um nodo (página *web*) e outro, com números indicando a probabilidade de cada ligação. A geração do grafo tem como propósito auxiliar os professores e os desenvolvedores de materiais educacionais digitais na interpretação dos dados. Assim, os resultados focam as trilhas mais visitadas, para tomar decisão sobre a organização do espaço virtual de aprendizagem, bem como para o professor recomendar caminhos e atalhos para os estudantes de turmas posteriores, sem relação com questões como desempenho acadêmico.

Oliveira & Garcia (2004) realizaram um estudo sobre o perfil dos candidatos que se inscreveram em um processo seletivo do Centro Universitário em que atuam (UNIFOR). Para tal, aplicaram um questionário sócio-econômico-cultural, cujos dados foram minerados em dois experimentos distintos, com a técnica de regras de associação, considerando um total de 1121 alunos. No primeiro experimento, tomaram por base a cidade de domicílio dos alunos, a renda da família, o fator que os motivou a escolher a instituição e como obtiveram informações sobre o processo

seletivo. As regras geradas permitiram identificar diferentes públicos, relacionando regiões de origem, poder aquisitivo, meios de comunicação mais eficazes e motivos mais significativos. O segundo experimento considerou a idade, sexo, ano de término do Ensino Médio, tipo de escola em que o realizou (particular ou pública), turno em que estudava (diurno ou noturno), renda familiar, se possuía computador em casa e quantas vezes já havia realizado processo seletivo na instituição, possibilitando conhecer melhor aqueles que a procuram.

As autoras do estudo descrito acreditam que os resultados podem ser utilizados para melhorar as ações de marketing voltadas à divulgação da instituição, especialmente do processo seletivo. O estudo não faz referência à utilização desses dados, por exemplo, complementados por dados de desempenho no vestibular e curso escolhido, o que poderia trazer informações importantes para o acompanhamento acadêmico. Tais informações poderiam apoiar o desenvolvimento de ações de nivelamento e atividades extensionistas, entre outras, que tendem a impactar positivamente no desempenho dos alunos e nos índices de retenção nos cursos ofertados.

Considerando variados dados relacionados à vida acadêmico-administrativa de alunos universitários, o artigo de Amorim et al. (2008) descreve uma pesquisa que busca prever a evasão em cursos ofertados na modalidade a distância. Para tanto, utilizou dados de alunos de cinco cursos diferentes e 10 semestres de dados históricos. Foram consideradas informações acadêmicas gerais (matrícula; ano e semestre do ingresso do aluno; quantidade de disciplinas cursadas e percentual de aprovação antes do abandono ao curso; coeficiente de rendimento escolar – média de todo o curso; quantidade de disciplinas já concluídas); e financeiras dos alunos (quantidade de prestações em atrasos; percentual de desconto no semestre anterior ao abandono do curso; trancamentos, cancelamentos, transferências e finalização do curso). O autor utilizou parte dos dados para treinar o algoritmo e, submetendo depois os demais dados, obteve quase 90% de acerto.

Os dados abordados na pesquisa de Amorim et al. (2008) têm um cunho mais administrativo. O estudo está focado em descobrir padrões de risco de abandono nos cursos de EAD, de forma mais geral, mas não considera as variáveis

pedagógicas do cotidiano dos cursos a distância, ou seja, não trabalha com dados pedagógicos produzidos nos AVA. Tais dados contêm indicadores importantes com relação à permanência e ao êxito (aprovação) dos alunos, que são trabalhados em outras pesquisas, apresentadas na seqüência.

Dias et al. (2008) realizou um estudo tomando por base os dados de 272 alunos, agrupados a partir de sete turmas na área de Banco de Dados (três de especialização e quatro de graduação), de dois semestres letivos, da Universidade Federal do Pará. As turmas foram trabalhadas na modalidade semipresencial, com aulas presenciais e atividades a distância. Os dados coletados foram gerados a partir da utilização de um ambiente virtual de aprendizagem para estudo de SQL (*Structured Query Language*), chamado LabSQL.

De cada aluno, o experimento realizado por Dias et al (2008) considerou 18 atributos: sexo, código do curso, código do tipo de curso, código da disciplina, código da turma, código do coordenador da turma, o tempo que o usuário levou para se inscrever na turma após o início do período de inscrição, se trabalhou em equipe, se usou agenda de anotações do sistema, total de pontos obtidos pelo usuário na resolução dos exercícios, total de problemas resolvidos, média de pontos dos problemas resolvidos, média do nível de dificuldades dos problemas resolvidos, quantidade de acessos às páginas do ambiente LabSQL, se ficou acima da média de acessos de todas as turmas, se ficou acima da média de acessos de sua turma, se ficou acima da média de pontos de todas as turmas e se ficou acima da média de pontos da sua turma.

Após a mineração, foi possível verificar que a demora para inscrição influenciou diretamente na média da quantidade de acessos, na utilização ou não da agenda, no número de problemas respondidos e no total de pontos obtidos pelos usuários. O estudo, então, afirma a necessidade de envolver os alunos desde cedo no ambiente virtual e de buscar formas de apoiar os alunos atrasados. Com poucos atributos, muitos deles de caráter administrativo, os resultados foram limitados.

Lopes & Schiel (2004) modelaram um sistema para apoio à decisão (SAD) para auxiliar professores a identificar perfis de estudantes em EAD. Na etapa de

seleção de dados, definiram um conjunto de análise denominado MAD – Modelo Analítico de Dados, que engloba número de e-mails enviados, quantidade de acesso aos recursos didáticos, nota em cada item de avaliação, estado (UF), sexo, idade e situação (cursando, trancado ou desistente). Para a mineração, foi utilizando o *framework* WEKA, desenvolvido por Witten & Frank (2005). O sistema definido minerou dados fictícios, sem uma validação com dados reais.

Romero et al. (2008) realizou um experimento para comparar diferentes técnicas de MD (e vários algoritmos relacionados) para classificar estudantes (reprovado / FAIL, aprovado / PASS, aprovado com bom resultado / GOOD, aprovado com resultado excelente / EXCELLENT) com base em informações de uso do AVA e suas notas finais. Para tanto, considerou os dados de 438 alunos, provenientes de 7 cursos diferentes, da Universidade de Córdoba (Espanha). O propósito do experimento era identificar técnicas com melhores resultados, tanto do ponto de vista da precisão quanto da facilidade de interpretação, o que apontou para a geração de árvores de decisão, com uma precisão aproximada de 65% – o valor não é maior, segundo o pesquisador, pelo fato de que alguns alunos não realizaram todas as tarefas propostas no ambiente ou apresentaram poucas ações comunicativas, mas obtiveram êxito dos exames presenciais, cujos pesos eram significativos para a composição dos conceitos finais.

Além disso, a pesquisa de Romero et. al (2008) resultou em uma ferramenta de MD desenvolvida e incorporada ao AVA *Moodle*, para que o próprio professor pudesse selecionar os dados (tabelas específicas ou combinadas), configurar os parâmetros da mineração e executá-la, obtendo resultados para sua interpretação. A seleção dos dados a minerar e os ajustes de variáveis, no entanto, não são triviais para o professor leigo (compreendido aqui como sem formação na área de Computação ou correlatas), nem os resultados gerados são fáceis de interpretar para tais professores. Novamente, o professor tem um relatório para analisar, em linguagem mais difícil do que os demais, o que é um aspecto dificultador para que o resultado da mineração se torne conhecimento útil nas mãos do professor.

3.6 Sistema de Alertas

A MD Educacionais nesta tese constitui-se em um diferencial para alimentar um sistema de alertas, buscando conscientizar os professores sobre eventos que acontecem no AVA e alunos com tendência à evasão ou à reprovação, oferecendo suporte e recomendação. Receber informação relevante, em quantidade ideal e no momento certo, pode ampliar a satisfação do professor com o ambiente virtual utilizado e com a sua atuação no acompanhamento dos alunos.

O trabalho desenvolvido na tese busca identificar e propor um conjunto híbrido e configurável de alertas em AVA, com alertas pré-definidos e outros gerados a partir da mineração de dados. Esses alertas devem dar suporte à atuação pedagógica do professor no acompanhamento das situações de aprendizagem, com o monitoramento de perfis de alunos e de eventos no ambiente, notificando o professor sobre grupos de alunos que compartilham necessidades específicas. Cada alerta, portanto, notifica o professor sobre a situação de um grupo de alunos, formado por um ou mais sujeitos, para que o professor possa dirigir comunicações ou atividades específicas. Os grupos identificados são dinâmicos, sendo gerados no momento oportuno para conscientizar o professor.

Um professor reativo, basicamente, responde ao aluno esclarecendo dúvidas quando solicitado. Um professor pró-ativo e mediador buscará ampliar as possibilidades de aprendizagem do aluno:

- Apresentando formas de organização e estudo;
- Incentivando a busca de ajuda e colocando-se à disposição;
- Antecipando situações diversas e realizando orientações específicas;
- Indicando bibliografia complementar e favorecendo a troca de materiais localizados ou produzidos pelos alunos sempre que oportuno;
- Estimulando a participação dos fóruns de discussão;
- Incentivando a interação entre os alunos;
- Propondo a reestruturação no material didático aos administradores do curso;

- Estabelecendo uma relação de proximidade (com linguagem afetiva) junto aos alunos;
- Estando atento às dúvidas da turma e orientando os alunos do que pode ser uma dúvida coletiva – coletando e organizando as dúvidas mais frequentes para consulta;
- Chamando à participação alunos que estão distantes ou pouco envolvidos, bem como orientando aqueles que apresentam uma participação de baixa qualidade;
- Lembrando prazos de entrega de atividades para os alunos que ainda não o fizeram, estimulando esses alunos a cumprir as atividades;
- Acompanhando o desempenho dos alunos e oferecendo oportunidades adicionais de recuperação para aqueles que necessitarem, em uma perspectiva de avaliação formativa.

Muitas dessas práticas, para engajar os alunos e diminuir a *distância transacional*, podem se valer de alertas para apoiar a atuação do professor. Ferramentas, agentes ou assistentes virtuais podem auxiliar professores e alunos em tarefas educacionais (KAMPFF, 1999; REATEGUI et al., 2008). Alertas em AVA podem apresentar-se por meio de mecanismos de percepção (alterações de interface) e de notificações explícitas (mensagens síncronas ou assíncronas), desencadeados por eventos específicos, com sugestões de ações (como interação com os alunos ou alterações na estrutura do curso), conforme a análise dos contextos a partir dos quais foram gerados.

Os aspectos referenciados nesta seção, que devem ser considerados em um modelo consistente para a geração de alertas para apoio à prática docente, acrescido dos resultados da mineração de dados históricos do curso ou disciplina específica, diferenciam a pesquisa da tese realizada de estudos e sistemas similares existentes, como os apresentados na seqüência.

3.6.1 Mecanismos de Percepção (*Awareness*)

O trabalho realizado por Otsuka (2006), em sua tese de doutorado, focou o aprimoramento de ferramentas do AVA Teleduc e a modelagem de um sistema multiagente (SMA) para suporte à avaliação formativa, por meio da exploração da possibilidade de registro e posterior análise das interações que ocorrem nos cursos a distância. Para o desenvolvimento da solução, baseou-se na metodologia de Hadji de avaliação formativa, em que o avaliador deve realizar quatro grandes tarefas: *desencadear* (tarefas/comportamentos a serem observados/interpretados); *observar/interpretar* (esses comportamentos); *comunicar* (os resultados de sua análise e sua apreciação final); *remediar* (as dificuldades analisadas).

O modelo de avaliação formativa abrange duas atividades complementares: o apoio ao planejamento detalhado de atividades de aprendizagem a serem avaliadas, bem como a regulação das participações nessas atividades planejadas (por meio de notas e comentários); e a redução da quantidade de informação a ser analisada, ajudando o professor na recuperação e análise de dados relevantes para a mediação das participações nas atividades planejadas, de acordo com os critérios de avaliação definidos no planejamento de cada atividade de aprendizagem. Os mecanismos de mediação desenvolvidos visavam à diminuição da sobrecarga dos professores na recuperação, acompanhamento, análise e organização do volume de informações geradas a partir das interações desencadeadas no processo avaliativo. Isso foi feito incorporando recursos de percepção (*awareness*) no ambiente que indicavam situações mediadas e não-mediadas (sem *feedback*), por exemplo, e relatórios de atividades centradas no aluno, permitindo uma análise global do mesmo.

Para auxiliar os usuários de AVA, Gerosa et al. (2003) e Fuks et al. (2004) descrevem a inserção de mecanismos de percepção na interface do AVA Aulanet, que destacam onde há alterações, sejam novidades ou pendências, bem como identifica os usuários que estão trabalhando em uma mesma ferramenta, além de mecanismos que notificam externamente os usuários sobre a inserção de materiais ou desenvolvimento de discussões. O objetivo de utilizar mecanismos de ampliação

da percepção do que ocorre no ambiente é contribuir para que os usuários comuniquem-se e colaborem mais.

O ambiente AulaNet também oferece relatórios de acompanhamento da participação (GEROSA et al., 2003), que propiciam a percepção do grupo sobre as tarefas realizadas por seus membros e fornece subsídios ao professor. Os relatórios oferecem o conceito médio dos participantes por serviço, a porcentagem média de contribuições efetivas, a frequência de participações nos debates, o número de contribuições dos participantes por serviço e relatórios detalhados de cada serviço do curso. Estes relatórios possibilitam ao aluno conferir seu desempenho e compará-lo com o dos colegas, através de informações atualizadas continuamente. Também fornecem elementos para que o professor, atuando como mediador, organize, motive e avalie os alunos. Com base nos vários relatórios, cabe ao professor analisá-los para buscar interpretações mais globalizadas, definindo suas intervenções.

Pinheiro et al. (2003), em sua pesquisa, propõe um *framework* para suporte à *awareness* (percepção) em sistemas de *groupware* (trabalho em grupo). No desenvolvimento de projetos compartilhados, é importante perceber a contribuição dos demais participantes do projeto, de forma contextualizada, para não sobrepor o desenvolvimento de atividades e avançar com segurança. Portanto, além das ferramentas de comunicação, coordenação e memória de grupo (conhecimentos comuns ao grupo, tais como as interações entre os participantes e os produtos desenvolvidos por eles), normalmente presentes em ferramentas de *groupware*, o suporte à *awareness* pode auxiliar significativamente na fluência do trabalho, por exemplo, resumindo para cada participante as últimas alterações do projeto que o grupo desenvolve em conjunto e em qual atividade cada componente do grupo está trabalhando no momento.

Os estudos de mecanismos de percepção contribuem para o desenvolvimento também de sistemas de alertas, que devem ser pensados de forma a ajudar os usuários em atividades importantes, sem sobrecarregá-los com informações excessivas.

3.6.2 Mecanismos de Alertas

Martins et al. (2003) descreve um sistema de chat, denominado *jXChat*, desenvolvido para auxiliar na mediação de comunicações síncronas em ambientes de aprendizagem. Por meio de um agente tutor inteligente, monitora o ambiente e as trocas de mensagens entre os usuários, realizando notificações. No momento de um bate-papo agendado, o agente alerta por *e-mail* os membros ausentes. Caso um determinado aluno esteja na sala, mas não participando do debate, o sistema dispara uma mensagem *on-line* a ele, por meio de uma janela *pop-up*, com o intuito de alertá-lo e convidá-lo a participar. Além disso, o agente compara o texto digitado por cada aluno com as palavras-chave armazenadas na base de dados, visando a identificar se a participação de cada um refere-se ao assunto proposto, enviando uma notificação ao professor com os nomes dos alunos que não estão trabalhando na temática definida. O professor, então, decidirá se deve intervir ou não a fim de resgatar o aluno ao debate.

Ao final de uma seção de debate, o sistema *jXChat* fornece ao professor, por meio de relatórios: uma lista com os nomes dos alunos participantes e outra com os ausentes; hora de entrada, de saída e tempo de permanência de cada aluno; detalhes sobre a participação de cada aluno, como número de interações e quantidade de palavras-chave digitadas em cada interação; além de um arquivo com todo o conteúdo do que foi debatido, buscando assim auxiliar na avaliação do processo de ensino e aprendizagem em ambientes para EAD.

Em torno de setenta por cento dos 20 alunos que utilizaram o *jXChat* acreditam que o sistema pode melhorar seus desempenhos na disciplina e que é válida a utilização de tal ferramenta como apoio à metodologia, sentindo-se motivados com as mensagens que receberam. O *jXChat* monitora um conjunto fixo de eventos, que ocorrem no momento da discussão síncrona, não considerando questões relacionadas ao perfil dos alunos, mas apenas o comportamento na seção observada.

Caravoli & Coello (2004) propõem o *Alerts Forum*, sistema que permite que as mensagens publicadas pelos alunos em um fórum de discussão sejam avaliadas

pelos seus pares e essas avaliações, bem como estatísticas de participação, sejam utilizadas para emitir alertas que estimulem os alunos a participar. Os alertas descritos apenas valorizam o mérito do aluno com boa avaliação e exigem postagens daqueles que não atingiram o mínimo de participações previsto.

Musa et al. (2000) apresenta um gerenciador de alertas, integrado a um AVA chamado *Claroline*, para acompanhamento das atividades realizadas pelos alunos em cursos a distância, monitorando situações previamente definidas e enviando mensagens para os usuários. Um alerta é composto por: um evento de disparo, uma ou mais condições que devem ser verificadas e a ação que deve ser realizada. Os eventos podem ser de dois tipos: eventos do ambiente, gerados pelo professor ou aluno durante a sua interação com o ambiente de ensino, como acessar o ambiente, terminar algum exercício ou acessar uma determinada página; ou eventos temporais, com situações que devem ser checadas em um determinado período de tempo. Na implementação realizada, o professor não pode incluir um novo evento, ele apenas escolhe os eventos de uma lista já pré-definida pelo administrador do sistema de alertas.

Musa et al. (2001) detalha a arquitetura de um agente notificador, baseado no conceito de sistema de alertas inteligentes, que consiste na monitoração de um banco de dados para a detecção de determinadas condições e tomada de decisão. A situação de cada aluno é acompanhada individualmente, gerando alertas para cada uma das situações, para cada aluno. Para o professor, é possível gerar um alerta referente a cada aluno, para cada situação prevista, não realizando um agrupamento que facilitasse o gerenciamento ao professor por meio do recebimento de um único alerta sobre um conjunto de alunos com mesmas dificuldades.

Reategui et al. (2008) apresenta um sistema multiagente para controlar um agente personificado que visa ao apoio ao aluno, com base no seu perfil na interação com o AVA. Cada agente do sistema monitora uma funcionalidade específica e o agente mediador controla qual deve entrar em ação. O sistema multiagente que compõe o Assistente Pessoal foi organizado em diferentes módulos, cada um destinado a uma função específica, a saber: comunicação, *feedback*, recomendação, monitoramento e interação. Alguns destes agentes são simples,

necessitando apenas de um conjunto de consultas a determinadas tabelas de um banco de dados para aportarem seus resultados, tais como os de monitoramento e de *feedback*. Já o agente de comunicação em linguagem natural e o agente de recomendação colaborativa necessitam de mecanismos mais complexos, para a interpretação de linguagem natural e para a utilização adequada de técnicas de recomendação.

Campana et al. (2008) apresenta uma arquitetura multiagente para ambientes virtuais de aprendizagem implementada em um AVA, denominado AVAUFES, da Universidade Federal do Espírito Santo. O sistema visa ao monitoramento e ao suporte aos alunos e professores que utilizam o ambiente. Campana et al. (2008) afirma que características como o auxílio personalizado, o suporte à cooperação e o suporte às atividades docentes estão entre os problemas que ainda necessitam de melhores soluções, tanto do ponto de vista pedagógico quanto tecnológico, principalmente quando os cursos permitem um grande volume de participantes, tornando mais difícil a mediação e o acompanhamento das tarefas dos alunos.

Os requisitos especificados no trabalho de Campana et al. (2008) permitem solucionar diversos problemas observados nos AVA, automatizando as notificações esperadas, tais como: manter o aluno informado sobre as tarefas pendentes e seus prazos; receber informações sobre novos materiais postados no ambiente pelos colegas, seu grupo ou professor; ser avisado sobre mensagens enviadas pelo professor, podendo especificar que a notificação ocorra a partir de certa quantidade de mensagens recebidas; e receber confirmação de entrega de trabalhos. Segundo Campana et al. (2008), a camada multiagente projetada e implementada se mostra flexível e adaptável às necessidades e características dos alunos e professores, de modo a permitir que seus usuários possam se sentir mais informados e integrados ao ambiente virtual, auxiliando-os no controle de suas atividades. O conjunto de notificações é fixo, determinado na implementação do sistema.

Kosba (2004) projetou, em sua tese de doutorado, um “conselheiro artificial” que monitora o que está acontecendo em um curso de EAD, infere os problemas que podem ocorrer e recomenda possíveis intervenções. Em um sistema de acompanhamento específico, os alunos são monitorados individualmente. Seria

importante poder agrupá-los em grupos com características comuns, permitindo que o professor direcionasse suas comunicações a grupos que compartilham mesmas necessidades, não a cada aluno individualmente.

O que fica evidente em todas as pesquisas relatadas é que quando se utiliza um AVA, não basta disponibilizar conteúdos e abrir espaços de discussão, é preciso acompanhar o percurso do aluno, mediando o processo de aprendizagem. O foco do trabalho desta tese é gerar alertas que auxiliem a mediação do professor. Um aspecto fundamental a destacar é a MD, neste contexto, que permite descobrir conhecimento útil, encontrando relações entre as variáveis analisadas que podem contribuir na ação do professor para engajar e apoiar seus alunos.

4. Sistema de Alertas: Apoio à Prática Docente

A Educação a Distância (EAD) está em franca expansão no Brasil, sendo modalidade de oferta de um número crescente de cursos, tanto por instituições públicas quanto privadas. Aliado à ampla oferta é preciso garantir a eficácia dessa educação. Considerando o grande volume de alunos, é preciso pesquisar alternativas tecnológicas que auxiliem o professor na identificação de padrões de comportamento dos estudantes, na tomada de decisão e na interação com os alunos que necessitam de determinadas intervenções, estabelecendo comunicações mais eficazes.

A pesquisa realizada utilizou um conjunto híbrido de alertas, composto por alertas fixos e outros variáveis, gerados automaticamente a partir de técnicas de mineração de dados oriundos de ambientes virtuais de aprendizagem. Tais alertas visaram a apoiar a mediação docente durante o acompanhamento de turmas com grande número de alunos, ofertadas na modalidade a distância. Os alertas gerados, com base nos resultados da MD, no formato de regras de classificação, auxiliaram o professor na identificação de alunos, em turmas correntes, com tendência à reprovação ou à evasão. O objetivo dos alertas foi, neste contexto, conscientizar o professor sobre os alunos em risco, oportunizando que ele realizasse intervenções preventivas junto aos mesmos, contribuindo para reverter estes quadros antes da sua confirmação.

Entende-se que muitas das práticas docentes de resgate, engajamento e recuperação de estudos podem ser preventivas e dirigidas a alunos que, identificados por seus comportamentos, demandam uma atenção maior. De maneira mais eficaz, os professores podem ser auxiliados, então, por alertas, disponibilizados no Ambiente Virtual de Aprendizagem (AVA) utilizado, constituindo-se em um diferencial para a tomada de decisão e intervenção junto aos alunos, potencializando sua ação pedagógica. O papel do professor a distância está relacionado ao acompanhamento e orientação de seus alunos, devendo estar atento a uma ampla rede de indicadores para garantir o sucesso dos sujeitos que realizam os cursos ofertados.

4.1 Alertas como Apoio ao Acompanhamento e Mediação Docente

Os motivos que levam à evasão na EAD são complexos, abrangendo questões internas e externas às instituições, amplamente discutidos na seção 2.6. Ao presente estudo importam às questões relacionadas ao apoio ao aluno, envolvendo ações para seu engajamento e promoção de aprendizagem, com efetivo acompanhamento do professor.

Araújo (2001) e Palloff & Pratt (2004) apresentam como resultado de suas pesquisas o fato de que os alunos a distância, mais do que material de qualidade e boas tecnologias de suporte, esperam que seus professores façam-se “presentes” na sala de aula virtual, acompanhando-os de acordo com suas necessidades. Além disso, Moore & Kearsley (2007), afirmam que um bom sistema de monitoramento e avaliação é essencial para o sucesso no desenvolvimento de cursos a distância.

Todo curso deve tomar como ponto de partida o planejamento cuidadoso das ações didáticas para estabelecer uma *distância transacional* (comunicativa / pedagógica) adequada (PETERS, 2001) passando pela definição da uma boa estrutura para o curso, pelo diálogo constante e pelo apoio aos alunos, especialmente aqueles não habituados à modalidade, auxiliando-os na construção de suas autonomias. Para o bom acompanhamento docente, faz-se necessário ter à disposição indicadores que permitam identificar alunos em dificuldade e, então, decidir sobre as intervenções didáticas apropriadas.

O desenvolvimento de sistemas de alertas a partir de indicadores definidos automatiza parte do processo de gestão do professor, desonerando-o da análise de uma série de relatórios dispersos. Mesmo assim, na arquitetura proposta na tese, cabe ao professor escolher a ação que desenvolverá a partir do alerta, conforme sua concepção epistemológica e a conjectura do trabalho em desenvolvimento, tais como as atividades em andamento, a produtividade da turma ou especificidades dos alunos em evidência.

A comunicação estabelecida com os alunos, quando oportuna, deve ser adequada ao contexto a que diz respeito, utilizando-se técnicas de interação que

chamem os alunos à participação, provendo *feedback*, promovendo integração entre os assuntos abordados e as orientações de estudo, buscando enfatizar a importância de cada sujeito no processo de aprendizagem (MOORE & KEARSLEY, 2007).

Conforme Berge (1995) há quatro principais grupos de funções exercidas pelo professor *on-line*, discutidas mais detalhadamente da seção 2.3, passando por questões *pedagógicas*, relacionadas à promoção da aprendizagem, *gerenciais*, vinculadas à organização das atividades, *sociais*, visando a estabelecer empatia e reduzir a sensação de isolamento, e *suporte técnico*, oferecendo apoio na utilização dos ambientes e recursos virtuais empregados. Merce & Estepa (2001), Pallof & Pratt (2004), Moore & Kearsley (2007) e Belloni (2008), nos estudos apresentados na seção 2.4, reafirmam a importância de uma mediação centrada nas necessidades dos alunos.

Situações como baixa frequência ou ausência no ambiente virtual por um longo período, falta de participação ou contato com professores e colegas, pouco acesso aos materiais disponibilizados, perda de prazo para entrega de atividades ou desempenho insatisfatório nas tarefas são alguns elementos que podem ser monitorados com periodicidade. Além disso, combinações entre diversas características demográficas e comportamentos podem ser analisados em conjunto, por meio de mineração de dados, para descobrir padrões em que intervir pode auxiliar no resgate ou engajamento dos alunos.

4.2 Arquitetura para Sistemas de Alertas em AVA

Com base nos estudos realizados, propõe-se uma arquitetura híbrida para sistemas de alertas em AVA, baseada em alertas configurados pelo professor, a partir de indicadores do ambiente virtual, e alertas gerados a partir da mineração de dados educacionais. Os alertas são gerados sobre grupos de alunos para a intervenção do professor. A figura 4.1 apresenta a arquitetura proposta.

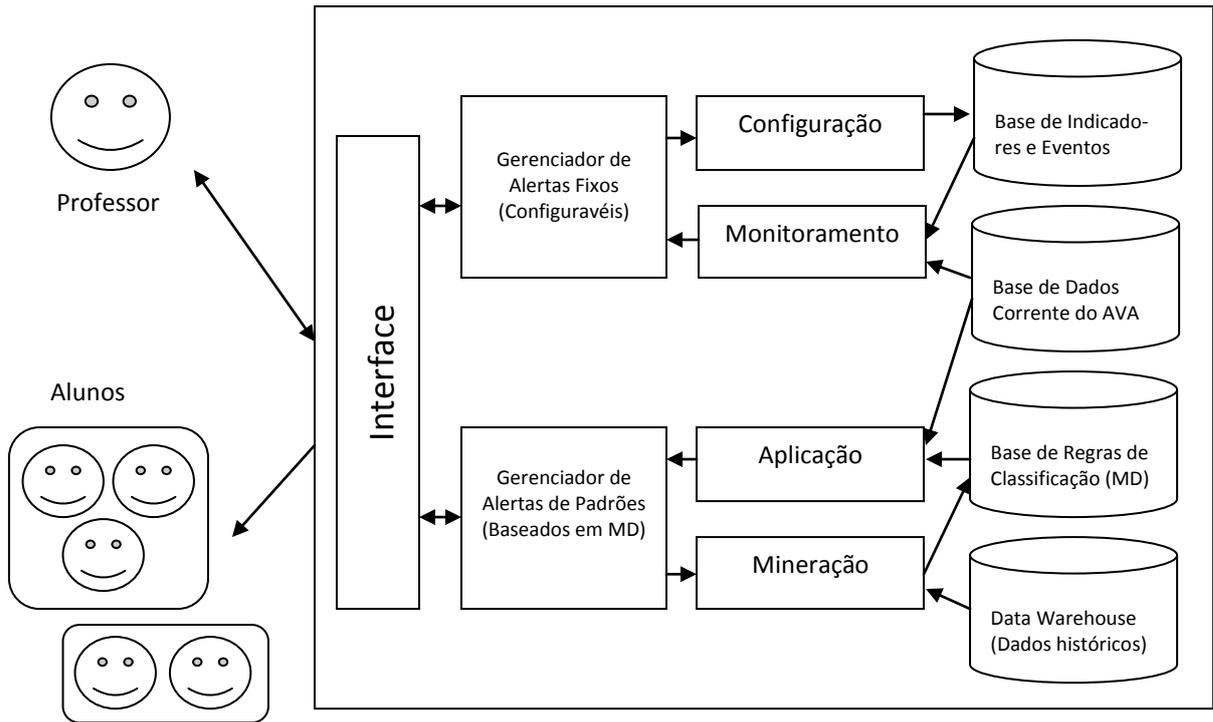


Figura 4.1 – Arquitetura para Sistemas de Alertas em AVA

4.2.1 Interface do Sistema de Alertas

É por meio do módulo de interface com o sistema de alertas que o professor configura os alertas fixos, com base em uma série de indicadores que podem ser monitorados no AVA, bem como recebe os alertas, tanto pelo gerenciador de alertas fixos (seção 4.2.2) quanto pelo gerenciador de alertas baseado em padrões obtidos por meio de mineração de dados (seção 4.2.3).

A figura 4.2 ilustra a interface de recebimento dos alertas por parte do professor, indicando se o alerta vem do monitoramento do gerenciador de alertas fixos configurado pelo professor (De: Configuração) ou se é oriundo da aplicação das regras de classificação baseadas na MD (De: Mineração). O sistema alerta o professor sobre os alunos que se enquadram nos parâmetros monitorados. No exemplo em evidência, o professor é alertado da ausência de acesso a certo material até a data estipulada, e tem a possibilidade de enviar uma mensagem diretamente aos alunos identificados ou de ser lembrado mais tarde sobre a ocorrência.

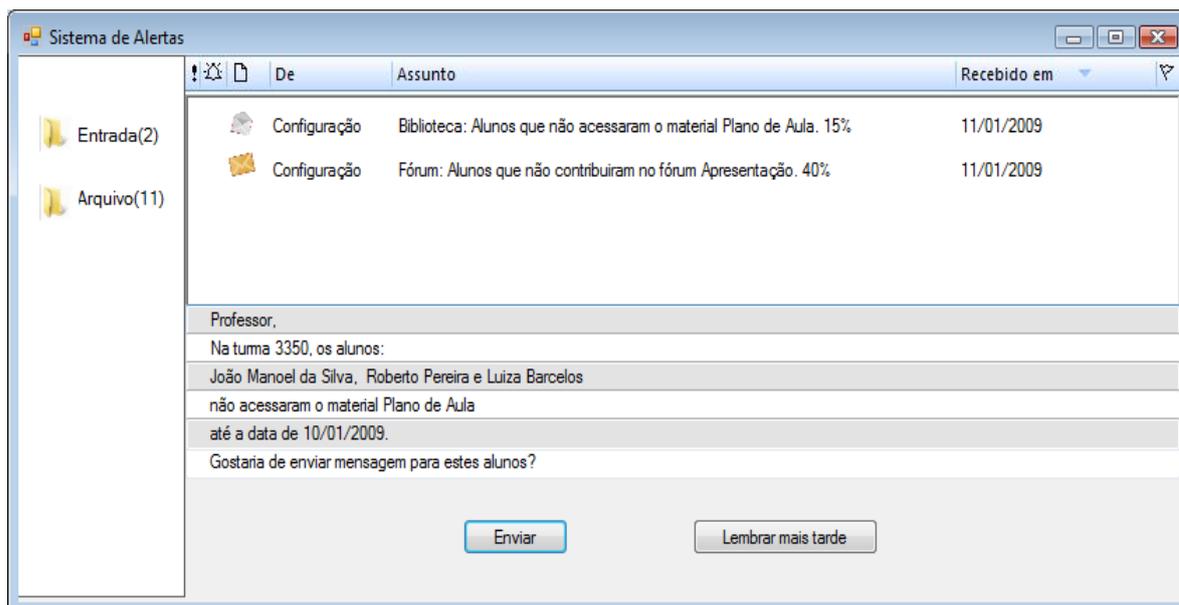


Figura 4.2 – Interface do Sistema Gerenciador de Alertas: recebimento dos alertas

4.2.2 Gerenciador de Alertas Fixos

Os alertas fixos estão relacionados a questões que o professor deseja acompanhar explicitamente, tais como acesso a materiais e realização de tarefas.

4.2.2.1 Configuração dos Alertas Fixos

A configuração dos aspectos que serão monitorados é armazenada em uma base de indicadores e eventos.

A figura 4.3 exemplifica possibilidades de configuração de alertas relativos aos materiais disponíveis na ferramenta Biblioteca Virtual. Para cada material disponibilizado, o professor pode indicar se o acesso é obrigatório, bem como a data limite preferencial.

Além disso, o professor deve indicar em que período deseja ser notificado sobre os alunos que não acessaram o material. Assim, o sistema gerará um alerta informando quais alunos não acessaram o recurso, por exemplo, um dia após a data limite indicada.

Esse processo pode ser repetido em vários períodos de tempo, indicando os alunos que ainda não acessaram o material monitorado. Os alunos também poderiam ser notificados diretamente pelo sistema, se o professor assim configurasse, sempre com cópia ao docente, para que esteja ciente dos alunos alertados.

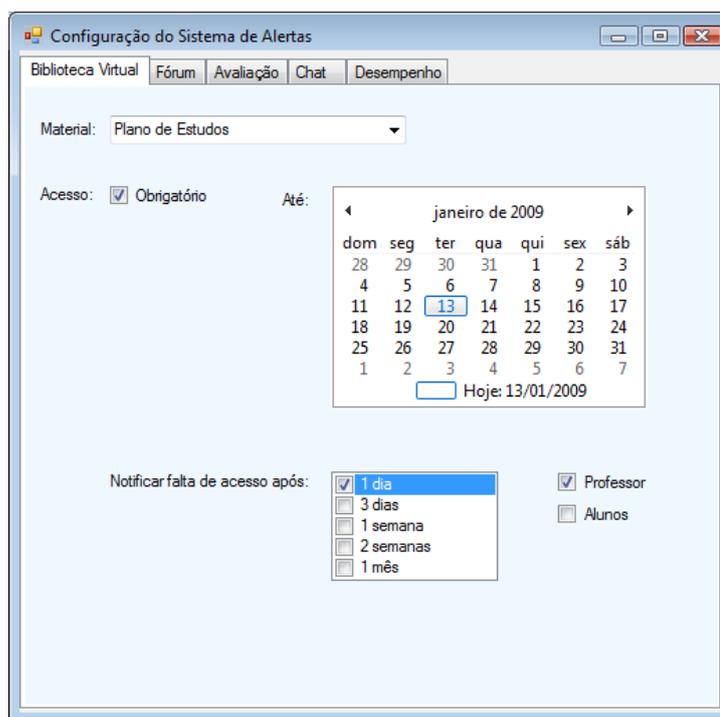


Figura 4.3 – Exemplo de Configuração dos Alertas: Biblioteca Virtual

Em relação à ferramenta Fórum, para cada discussão criada, é possível definir, para acompanhamento do sistema, se a postagem dos alunos é obrigatória e a data limite para realizá-la. É recomendado notificar a ausência de participação em um período que antecede o encerramento do prazo, para que o professor possa tomar consciência dos alunos que ainda não contribuíram e, então, intervir junto a eles.

Além disso, o professor pode desejar ser notificado sobre alunos com número de postagens inferior a um número estabelecido ou, em fóruns avaliativos, sobre notas abaixo de um limiar estipulado (figura 4.4).

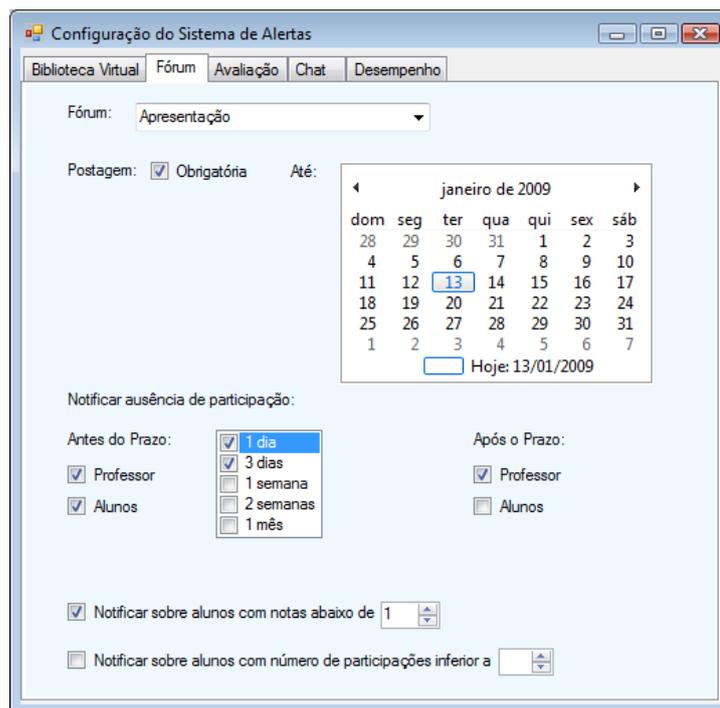


Figura 4.4 – Exemplo de Configuração dos Alertas: Fórum

Em relação à ferramenta avaliação, para cada tarefa, o professor pode configurar as notificações e avisá-lo com antecedência sobre alunos que não entregaram a atividade, para que possa contatá-los, com orientações para que ainda possam realizá-la no tempo disponível. É importante acompanhar, também, os alunos que não entregaram a atividade após o prazo expirado, saber quais são os alunos que ficaram abaixo de uma nota estabelecida (útil com questões de correção automática), bem como saber quais questões foram mais erradas pelo conjunto de alunos. A figura 4.5 exemplifica a configuração de alertas sobre as questões abordadas.

Considerando a ferramenta de *chat*, o professor pode indicar quem são os alunos que devem participar e configurar o sistema para alertá-lo sobre alunos ausentes em uma determinada seção, bem como sobre os alunos que estão inativos a certo tempo.

Há, ainda, questões gerais de desempenho dos alunos que o professor pode desejar acompanhar, tais como: ausência da sala de aula virtual por um determinado período de tempo, número de acessos ao ambiente ou à biblioteca inferior a um patamar estipulado, bem como baixa participação nos fóruns propostos. O professor

pode definir a periodicidade com que deseja receber tais alertas: diário, a cada três dias, semanal, quinzenal ou mensal. A figura 4.6 ilustra o processo de configuração de tais alertas.

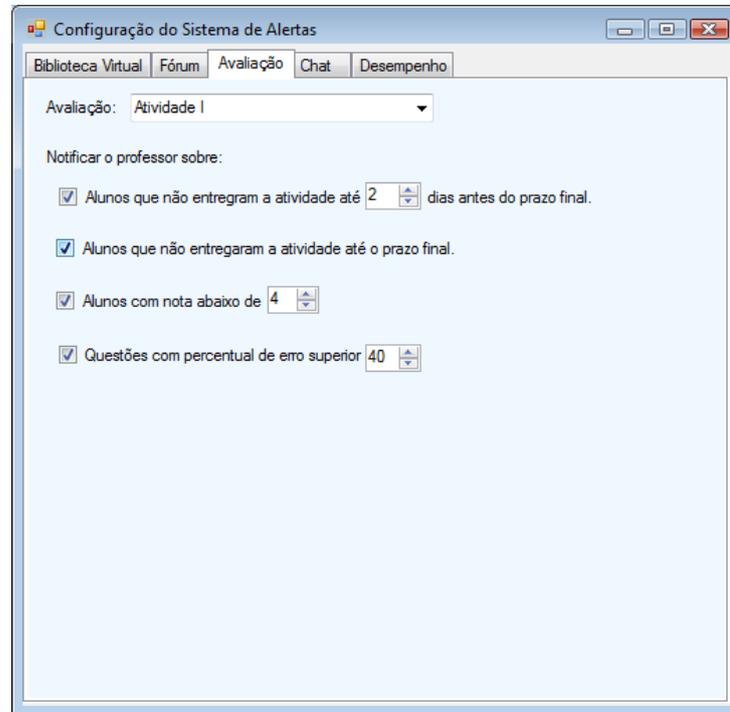


Figura 4.5 – Exemplo de Configuração dos Alertas: Avaliação

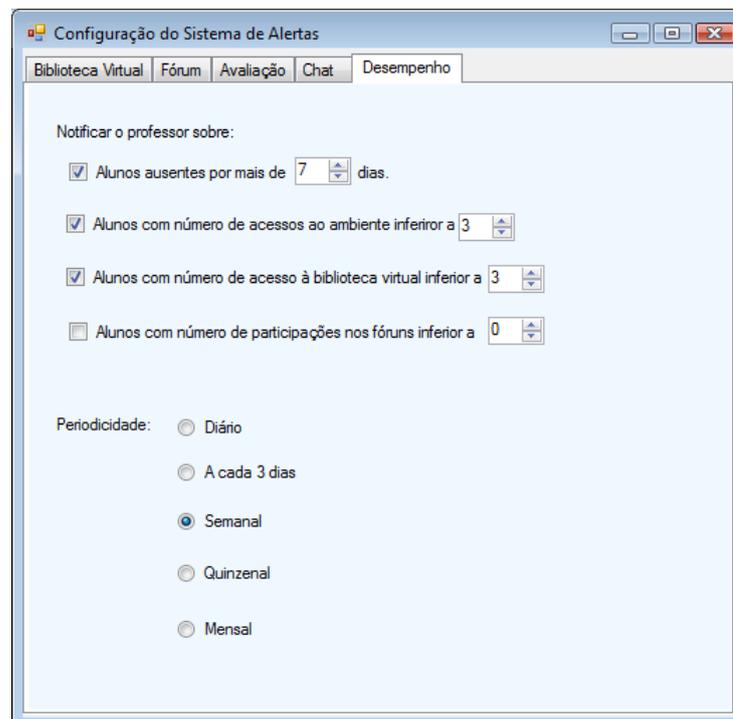


Figura 4.6 – Exemplo de Configuração dos Alertas: Desempenho

4.2.2.2 Monitoramento dos Alertas Fixos

Por meio de consulta à base de dados dos indicadores configurados e à base de dados do AVA, com as informações correntes dos alunos, o sistema gera os alertas necessários ao professor, visualizados por meio do módulo de interface.

4.2.3 Gerenciador de Alertas de Padrões

Os alertas de padrões, obtidos por meio de processos de descoberta de conhecimento, devem ser automaticamente gerenciados pelo sistema.

4.2.3.1 Mineração

O módulo de mineração consulta o Data Warehouse com dados históricos das disciplinas ou cursos correspondentes, previamente organizado, e gera as regras de classificação do período correspondente, armazenando-as na base de regras de classificação. O Data Warehouse deve ser atualizado a cada nova edição de um curso, o que levará a necessidade de atualização / geração de novas regras para aplicação em turmas subseqüentes.

4.2.3.2 Aplicação

O módulo aplicação do gerenciador de alerta de padrões busca, nos períodos de tempo definidos, por exemplo mensalmente, as regras de classificação pertinentes e as aplica aos dados correntes dos alunos de turmas em andamento, buscando classificá-los para notificar o professor sobre alunos em risco. A partir da classificação, alertada ao professor pelo módulo de interface, o professor poderá tomar as ações pertinentes, como enviar uma comunicação, reestruturar um material ou ampliar um prazo de entrega de atividades.

Implementação

O módulo aplicação, integrado ao gerenciador de alertas de padrões e a interface do sistema de alertas, foi implementado e validado para geração de alertas

ao professor, com base em MD, podendo salvar os nomes dos alunos em risco ou enviar uma mensagem para eles. O módulo recebe como entrada o arquivo de regras (obtidos no processo de mineração dos dados) e a planilha com os dados preparados das turmas atuais.

O módulo de aplicação do sistema de alertas, então, aplica as regras sobre os dados dos alunos e os classifica como evadido, aprovado, reprovado ou sem acesso. A figura 4.7 mostra o ambiente, com regras e dados extraídos do experimento realizado para a tese, para exemplificar seu funcionamento.

Arquivo Regras: C:\Users\Desktop\Mineracao_Mes11.log

Arquivo Excel: C:\Users\Desktop\Tabela_2009_1_Final_Ate_S4.xls

Total de Registros: 230

TENDE A SER EVADIDO = 27

ID_ALUNO	TURMA	UNIDADE	CURSO	NOME	EMAIL
395586	2918	SaoJeronimo	DIREITO		
428075	2918	SaoJeronimo	BIOLOGIA		
301933	2918	SaoJeronimo	BIOLOGIA		
393145	2918	SaoJeronimo	SISTEMASDEIN...		

TENDE A SER REPROVADO = 3

ID_ALUNO	TURMA	UNIDADE	CURSO	NOME	EMAIL
352429	2918	SaoJeronimo	ADMINISTRACAO		
395057	2918	SaoJeronimo	MATEMATICA		
429396	2986	Canoas	FISICA		

TENDE A SER APROVADO = 200

ID_ALUNO	TURMA	UNIDADE	CURSO	NOME	EMAIL
404919	2918	SantaMaria	DIREITO		
426552	2918	SantaMaria	DIREITO		
41451	2918	SantaMaria	PSICOLOGIA		
432496	2918	SantaMaria	EDUCACAOFISICA		

SEM ACESSO = 0

ID_ALUNO	TURMA	UNIDADE	CURSO	NOME	EMAIL
----------	-------	---------	-------	------	-------

Figura 4.7 – Sistema de Alertas: classificação

Para cada classificação, o professor pode salvar as informações dos alunos, clicando sobre o ícone salvar, ou enviar e-mail para todos relacionados a certa classificação, como ilustra a figura 4.8.

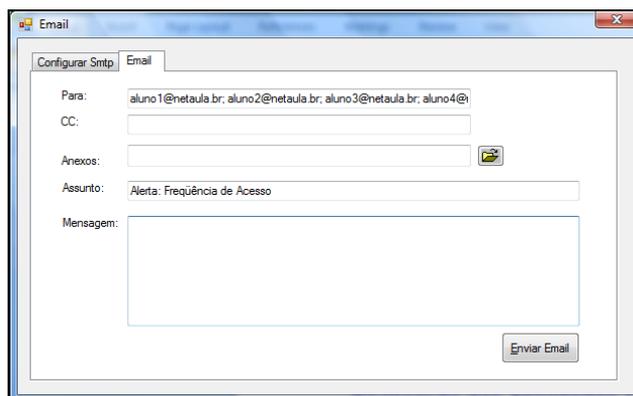


Figura 4.8 – Sistema de Alertas: envio de e-mail

A aba Visualização de Regras, por sua vez, permite que o professor veja, para cada regra específica, quais são os alunos classificados por ela. A figura 4.9 exemplifica esta funcionalidade, com os identificadores dos alunos, ao invés dos nomes, para preservar suas identidades.

Semestre	Turma	Id_Us	Unidade	Curso	Área
2009/1	2918	404919	Santa Maria	DIREITO	Cienc
2009/1	2918	395586	Sao Jeronimo	DIREITO	Cienc
2009/1	2918	428075	Sao Jeronimo	BIOLOGIA	Cienc

Total de Registros: 230

Visualização Regras

- # AF_TM1 > 6.500 and AG_TM1 > 25.500 and AF_TM1 <= 53 then Aprovado (66 / 0 / 0 / 0)
 - 432496
 - 425784
 - 424472
 - 423954
 - 42459
 - 394154
- # AF_TM1 > 6.500 and Unidade = Gravatai and Sexo = Masculino then Aprovado (26 / 0 / 0 / 0)
- # AF_TM1 > 6.500 and Unidade = Canoas and Area = SaudeeBem-EstarSocial then Aprovado (110 / 2 / 3 / 0)
 - 40935
 - 43002
 - 353096
 - 393617
 - 352892
 - 13358
- # FREQ_M1 > 1.500 and AB_TM1 > 17.500 and AF_TM1 > 26.500 and AB_TM1 <= 93.500 and AF_TM1 <= 32.500 then Aprovado (25 / 0 / 0 / 0)
 - 19576
- # FREQ_M1 > 1.500 and Area = TecnologiaeComputacao and FREQ_M1 > 3.500 then Aprovado (73 / 3 / 1 / 0)
 - 397875
 - 395000
 - 422790

Figura 4.9 – Sistema de Alertas: visualização de regras

O professor pode clicar sobre o identificador do aluno para retomar os dados completos do mesmo, o que permite ao docente analisar mais detalhadamente as informações de um aluno em particular, quando houver necessidade. Este ambiente foi desenvolvido como parte da pesquisa e servirá de base para integração de

alertas baseados em mineração de dados no AVA NetAula, com a incorporação automática de regras e dados dos alunos.

4.2.4 Considerações sobre a Arquitetura

A arquitetura apresentada para um sistema de alertas híbrido, a ser integrado a um AVA, apresenta os pontos essenciais para organização do sistema. Os dados a serem acompanhados e minerados poderão variar, conforme os recursos a serem monitorados em um AVA específico ou de forma a enfatizar mais determinados conjuntos de indicadores.

Para conceber o conjunto de alertas (notificações automáticas) validado na tese, para conscientização do professor e auxílio na sua ação mediadora, foram considerados os principais recursos computacionais disponíveis nos AVA. Além disso, técnicas de MD foram exploradas, para auxiliar na localização de características e comportamentos comuns aos alunos que evadem ou reprovam. Foram considerados dados demográficos e dados gerados nos AVA, em seus diversos espaços de informação e comunicação, detalhados no decorrer deste capítulo.

A partir dos estudos realizados, nas áreas de acompanhamento docente e retenção de alunos e de mineração de dados e sistemas de alertas, e da arquitetura proposta para sistemas de alertas em AVA, realizaram-se os experimentos da tese, descritos na seqüência, cujos resultados são analisados.

4.3 Contextualizando o Espaço de Realização dos Experimentos

Os experimentos foram realizados na Universidade Luterana do Brasil (ULBRA), instituição privada de abrangência nacional, que atua no Ensino Superior desde o ano de 1972. Além do campus Central, em Canoas, a ULBRA atua no Rio Grande do Sul (RS) com unidades de graduação, pós-graduação, extensão e pesquisa, em Cachoeira do Sul, Carazinho, Gravataí, Guaíba, Santa Maria, São Jerônimo e Torres. Está presente com unidades também em Itumbiara (GO), Porto

Velho e Ji-Paraná (RO), Manaus (AM), Santarém (PA) e Palmas (TO) (ULBRA, 2009).

Atenta ao desenvolvimento das novas tecnologias da informação e da comunicação e suas possibilidades de democratização da educação, em 2000, a ULBRA estabeleceu suas metas para a EAD, buscando criar condições para a oferta de cursos nesta modalidade e também para a ampliação da utilização de tecnologias na Educação Presencial. A partir de então, passou a planejar modelos de EAD viáveis, buscando desenvolver currículos diferenciados e materiais em linguagem multimídia, incorporar recursos tecnológicos de mediação e estabelecer programas de capacitação dos corpos docente e técnico-administrativo, constituindo também uma série de serviços para suporte a professores e alunos.

No primeiro semestre letivo de 2003, iniciou a oferta de disciplinas a distância para alunos da graduação presencial, alcançando o número de 1096 alunos matriculados. Em 17 de março de 2004, com a publicação no Diário Oficial da União da Portaria MEC nº 650/04, a ULBRA obteve seu credenciamento para a oferta do Programa Especial de Formação Pedagógica de Docentes das Disciplinas dos Currículos do Ensino Fundamental, do Ensino Médio e da Educação Profissional em nível Médio, nas áreas de Letras-Português e Matemática. Posteriormente, com a Portaria MEC nº 839, de 3 de abril de 2006, que ampliou a abrangência de sua atuação para todo o território nacional, novos cursos passaram a ser desenvolvidos (ULBRA/EAD, 2009). Atualmente, a ULBRA, por meio de Pólos de EAD, está presente em todos os estados brasileiros.

A ULBRA é hoje uma das maiores universidades do país, contando com mais de 150.000 alunos, distribuídos em 81 cursos de graduação, além dos diversos cursos de pós-graduação. Desses alunos, cerca de 80.000 estudam em cursos ofertados na modalidade a distância, distribuídos em 12 cursos de graduação e diversos de pós-graduação (ULBRA, 2009).

A cada semestre tem crescido, também, o número de alunos da graduação presencial que optam por realizar algumas disciplinas a distância, conforme possibilita o MEC (MEC/Portaria n. 4.059, de 10 de dezembro de 2004), compondo

até 20% da carga horária cursada pelo aluno. No RS, no segundo semestre de 2009, constam em torno de 5.000 matrículas, distribuídas em pouco mais de 20 disciplinas diferentes. As disciplinas ofertadas a distância para a graduação presencial atendem aos referenciais de qualidade para o EAD do MEC (abordados na seção 2.1), com materiais em mídias diversificadas elaborados por equipes interdisciplinares, professores capacitados, interação no AVA NetAula, atividades avaliativas a distância e presenciais e infra-estrutura de apoio nas diversas unidades da instituição.

Com professores sediados em Canoas, essas disciplinas a distância da graduação presencial são cursadas por alunos das diversas unidades do RS (Cachoeira do Sul, Canoas, Carazinho, Gravataí, Guaíba, Santa Maria, São Jerônimo e Torres). Cada turma acompanhada por um professor é composta por cerca de 80 alunos. Com calendário comum, as atividades de abertura e encerramento das salas virtuais, bem como as datas das provas presenciais, são compartilhadas e compatibilizadas com o calendário acadêmico das graduações presenciais. Para as avaliações presenciais obrigatórias, os alunos comparecem à unidade de origem à qual estão vinculados e as provas ocorrem simultaneamente nas diversas unidades.

Quatro disciplinas, que compõem o Ciclo de Formação Básica da Universidade, costumam contar com um número elevado de matrículas, tanto nas turmas presenciais quanto nas turmas a distância – são elas: Comunicação e Expressão, Sociedade e Contemporaneidade, Cultura Religiosa e Instrumentalização Científica. Para validar a geração de alertas a partir de dados históricos volumosos e, com base neles, o acompanhamento e intervenção do professor em turmas correntes, os experimentos da tese foram realizados a partir dos dados gerados por turmas ofertadas a distância da disciplina de Instrumentalização Científica (IC), que também contava com uma estrutura já bem estabelecida, da organização de materiais e atividades ao apoio e diálogo com os alunos.

A apresentação da pesquisa e solicitação de liberação de dados à instituição foi feita a partir de uma carta que explicava, em linhas gerais, o projeto e as necessidades envolvidas, disponível no anexo 5.

4.4 Experimento Preliminar

O experimento preliminar realizado para a proposta da tese tomou por base a mineração de dados reais de 161 estudantes de duas turmas de IC ofertadas na modalidade EAD, para alunos de cursos de graduação presencial. A disciplina minerada é semestral e os dados coletados referem-se aos dois primeiros meses de estudo, ou seja, até o fechamento da primeira nota utilizada para compor a nota final na disciplina – a nota final dos alunos é a média ponderada entre a primeira nota com peso um e a segunda nota com peso dois.

No experimento realizado, além de dados gerados em diversas ferramentas disponíveis em AVA (datas e quantidades de acessos aos conteúdos e arquivos nas bibliotecas virtuais; entrega das atividades de avaliações e notas; participações em fóruns, com postagens), pretendeu-se avaliar as relações com o sexo, a idade, a unidade presencial (pólo) e o curso de origem do aluno, analisando dados de disciplinas ofertadas totalmente a distância para alunos de cursos variados de graduação presencial. A partir desses dados, também foram consideradas as notas desses alunos, para verificar a existência de relações entre os comportamentos e os desempenhos. Os dados utilizados no experimento preliminar ora descrito estavam à disposição, seja no AVA ou no sistema acadêmico, por meio de relatórios.

Para realizar a mineração de dados, então, foram selecionados e preparados os dados relevantes para a investigação, excluindo-se dados redundantes e uniformizando a forma de representação dos mesmos que, originalmente, estavam em diferentes repositórios de dados, com formas de representação também diferenciadas. Após a preparação dos dados, para gerar as regras de classificação, foram utilizados os algoritmos *RuleLearner* e *DecisionTree* (apresentados nas seções 3.3.1 e 3.3.2, respectivamente) e a ferramenta de mineração *RapidMiner* (seção 3.4.1).

Os alunos pertenciam a três pólos diferentes (unidades presenciais onde ocorrem a aula introdutória e as provas finais, na proporção de 139/11/11, respectivamente de Canoas, Santa Maria e Gravataí) e a cinco áreas do conhecimento diferentes (conforme agrupamento da Universidade, 55/51/29/20/5). Entre os alunos da amostra, 78 eram masculinos e 83 femininos. 92 alunos tinham idade entre 16 e 25 anos, 52 entre 26 e 35 anos, 11 entre 36 e 45 anos e 6 possuíam 46 anos ou mais. Considerou-se a época do primeiro acesso ao AVA (primeira, segunda ou terceira quinzena da disciplina), a época do último acesso (semana corrente, semana passada, há mais de quinze dias ou há mais de um mês), o número de acessos ao AVA (que variou de zero a 80), o número de acessos à biblioteca virtual (que variou de zero a 50), o número de postagens em fóruns (de zero a 12) e o número de avaliações entregues (de zero a três).

Buscou-se relacionar estes dados com as notas obtidas pelos alunos durante os dois meses de acompanhamento da disciplina. Os dados preparados foram convertidos para o formato CSV e submetidos à ferramenta *RapidMiner*, que na figura 4.10 apresenta a visualização dos metadados.

Type	Name	Value Type	Statistics	Range
label	Faixa	nominal	mode = Nota: [7..9] (79)	Nota: [0..5] (41), Nota: [9..10] (12), Nota: [7..9] (79), Nota: [5..7] (29)
regular	Sexo	nominal	mode = F (83)	M (78), F (83)
regular	Pólo	nominal	mode = Canoas (139)	Canoas (139), SantaMaria (11), Gravataí (11)
regular	Área	nominal	mode = Educação,CiênciasEArtes (55)	SaúdeEBemEstarSocial (51), TecnologiaEComputação (20), Educação,CiênciasEArtes (55), CiênciasSociaisAplicadas (29), CiênciasAgrárias (6)
regular	FaixaEtária	nominal	mode = Idade: [16..25] (92)	Idade: [36..45] (11), Idade: [26..35] (52), Idade: [16..25] (92), Idade > 45 (6)
regular	PriAc1	nominal	mode = Quinzena1 (113)	Quinzena1 (113), Quinzena4 (1), NC (4), Quinzena2 (32), Quinzena3 (10)
regular	ÚltAc1	nominal	mode = EssaSemana (114)	MaisDeUmMês (5), EssaSemana (114), NC (4), De7a15DiasAtrás (23), De15a30DiasAtrás (15)
regular	NroAces	integer	avg = 20.969 +/- 15.140	[0.000 ; 83.000]
regular	T.Fór	integer	avg = 2.950 +/- 2.793	[0.000 ; 12.000]
regular	T.Bib	integer	avg = 12.224 +/- 8.946	[0.000 ; 50.000]
regular	T.Ava	nominal	mode = 3 (121)	0 (12), 3 (121), 1 (5), 2 (23)

Figura 4.10 – RapidMiner: visualização dos metadados

A partir dos dados tratados, foi utilizado o algoritmo *RuleLearner* para a descoberta direta de conhecimento (aprendizagem supervisionada), gerando regras no formato *Se <condição> Então <classificação>*. O foco neste experimento preliminar era descobrir a relação entre comportamentos dos alunos – como acesso ao ambiente e às leituras sugeridas, entrega das atividades, postagens em fóruns – e características demográficas – como faixa etária, sexo, curso e pólo de origem – e suas possíveis repercussões nos desempenhos dos alunos.

Aplicou-se, também, o algoritmo *DecisionTree* para a geração de *árvores de decisão*, alterando-se os dados de desempenho (ao invés de faixas de valores, trabalhou-se com os conceitos de acima ou abaixo na média) que apresenta os resultados de maneira gráfica e textual, como apresentado nas figuras 4.11 e 4.12.

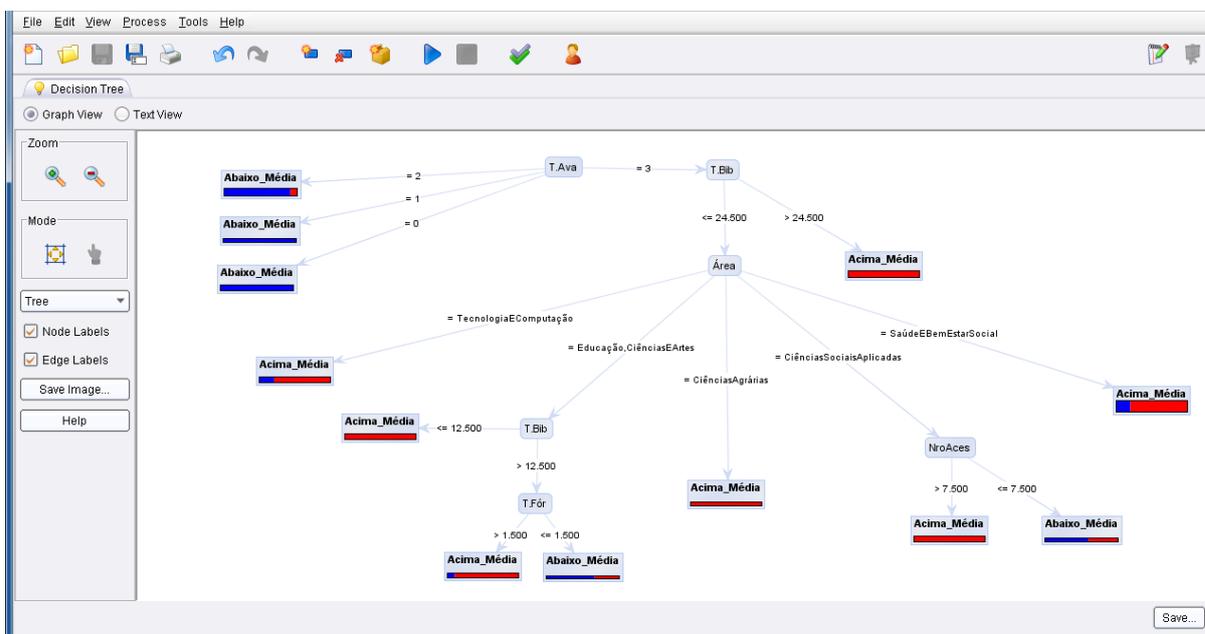


Figura 4.11 – Visualização Gráfica da Árvore de Decisão Gerada

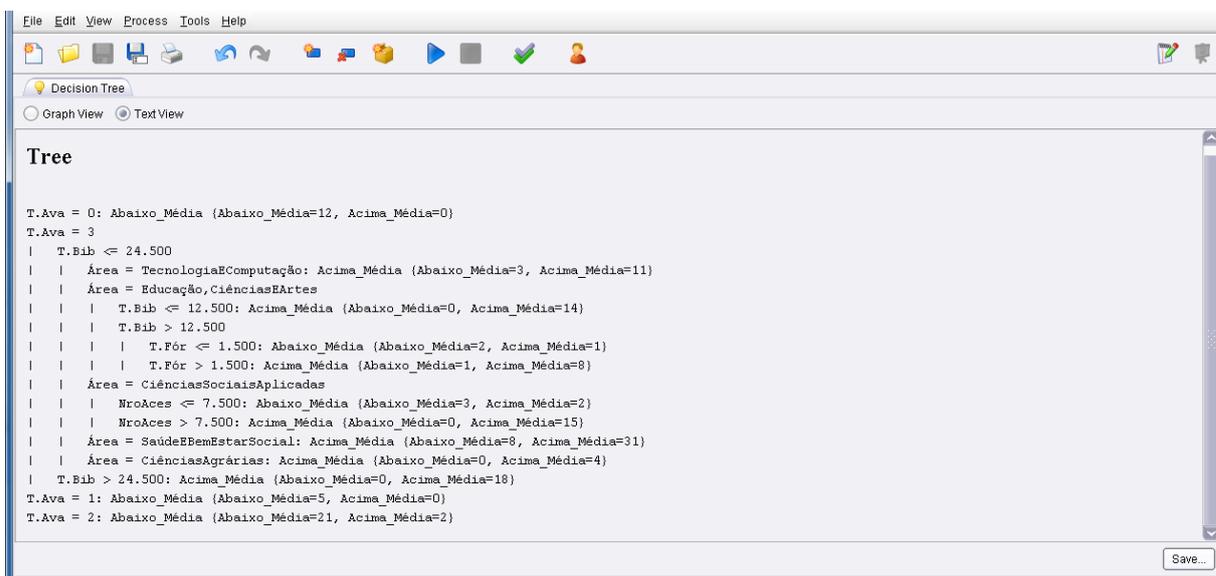


Figura 4.12 – Visualização Textual da Árvore de Decisão Gerada

Ao interpretar a *Árvore de Decisão* gerada, observou-se que dificilmente os alunos que não realizaram todas as atividades avaliativas propostas obtiveram nota

acima da média – reforçou-se a importância de engajar e envolver os alunos na realização das atividades. A análise mais relevante referiu-se, portanto, aos alunos que realizaram todas as tarefas propostas (T.Ava=3). Desses, todos os alunos com mais de 24 acessos aos materiais de estudo, periodicamente postados na Biblioteca Virtual do AVA utilizado, obtiveram bons resultados (Acima_Média).

Entre os alunos da Área de Educação, Ciências e Artes, por exemplo, foi possível identificar que, para aqueles que tiveram mais de 12 e menos de 25 acessos à Biblioteca Virtual, o que distinguiu os alunos acima ou abaixo da média foi o volume de postagens realizadas nos fóruns de discussão: os alunos acima da média tiveram, no mínimo, duas postagens nos fóruns. As informações sobre idade, sexo e unidade presencial não foram relevantes para a classificação dos sujeitos.

Observou-se, a partir da mineração realizada no experimento preliminar, o potencial de analisar os dados gerados em AVA para descoberta de conhecimento, resultando em regras que podem ser aplicadas em novas turmas para classificar os alunos em função de suas características e comportamentos.

Para resultados mais efetivos, a amostra foi ampliada para a realização do experimento final da tese, bem como foram acrescentados dados que permitem acompanhar a periodicidade de acesso às diversas ferramentas existentes no AVA utilizado pelos alunos, ampliando a seleção de dados a outras ferramentas de comunicação não consideradas neste experimento preliminar. É preciso dispor de um volume de dados maior do que o inicialmente utilizado para a efetiva localização de padrões que possam subsidiar a atuação do professor.

4.5 Processo de Descoberta de Conhecimento

A mineração de dados educacionais disponíveis nos AVA – partindo de dados demográficos (sexo, idade, pólo), comportamentais (quantidade e frequência de acessos aos recursos de informação e comunicação disponíveis e ações realizadas), de desempenho (entrega e notas das tarefas) e suas relações com o resultado final (evasão, reprovação e aprovação) – possibilita a identificação de alunos com

características comuns que demandam intervenções similares por parte dos professores.

Com o objetivo, então, de descobrir padrões que permitam identificar alunos com tendência à evasão e à reprovação, além de apontar também aqueles que tendem à aprovação, para alertar o professor e apoiar sua ação docente, dá-se início ao processo de descoberta de conhecimento (KDD), conforme as etapas descritas por Fayyad et al. (1996), explicadas na seção 3.1.

As primeiras etapas do processo (seleção dos dados, pré-processamento e transformação) demandam um grande esforço referente à aquisição e ajuste dos dados. A seção 4.5.1 apresenta a definição da amostra utilizada e a seção 4.5.2 detalha a organização dos dados.

A seção 4.5.3 destina-se a explicar as minerações realizadas, cujos resultados (regras) são utilizados na geração dos alertas. O foco da presente pesquisa é a descoberta direta de conhecimento, com tarefas de classificação (seção 3.2).

4.5.1 Definição da Amostra

A disciplina de Instrumentalização Científica (IC) foi escolhida para a MD em função do volume de dados disponíveis. No período 2006/1 a 2008/2, a disciplina já contava com seis semestres de ocorrência, com cerca de 4.800 alunos tendo cursado IC no formato de EAD.

Para este estudo, foram considerados os dados históricos de alunos, distribuídos em 20 turmas, em dois semestres diferentes (5 turmas em 2006/2 e 15 turmas em 2008/2). As turmas eram constituídas por alunos provenientes de diferentes cursos e de 8 unidades presenciais distintas.

A amostra de dados minerados, de 1564 alunos, é considerada adequada para representar a população de alunos que cursou IC em EAD, sendo superior a amostra necessária para um nível de confiança de 99.9% (883), conforme cálculo de

tamanho da amostra obtido por meio do ambiente estatístico SEstatNet (NASSAR et al., 2009), apresentado na figura 4.13.

Tamanho da Amostra	
Estimação de Percentual	
Tamanho da População	4800
Precisão da Estimativa	50 ± 5%
Nível de Confiança	95%
Tamanho da Amostra	356
Perda Amostral	Nenhuma
Para outros Níveis de Confiança	
Nível de Confiança	Tamanho da Amostra
99.9%	883
99%	583
90%	256

Figura 4.13 – SEstatNet: cálculo do tamanho da amostra

A partir de 2006/2, a estrutura da disciplina estabilizou-se, contando com um número determinado de tarefas e pesos, tarefas essas disparadas e recebidas em semanas específicas, conforme plano de ensino e aprendizagem da disciplina. Cada uma das tarefas avaliativas oportunizadas aos alunos, embora renovada a cada semestre, busca avaliar aspectos definidos da disciplina. Por exemplo, a Atividade1 da disciplina, em qualquer um dos semestres considerados, vale 2,5 pontos para compor a Nota de Grau1 e diz respeito à avaliação dos conteúdos e competências trabalhados nos dois primeiros capítulos disponibilizados.

Assim, apresentando uma estrutura de trabalho com poucas variações ao longo dos semestres, a disciplina de IC apresentou-se adequada para a busca de padrões de características e comportamentos dos alunos que a cursam, de forma a utilizar esses padrões para classificar, de forma antecipatória, as tendências de resultados dos próximos alunos a cursá-la. O objetivo de antecipar um possível resultado, a partir dos padrões identificados, serve de base para alertar o professor, que deve atuar de forma a prevenir (ou reverter) prognósticos ruins.

4.5.2 Organização dos Dados

De cada aluno, foram obtidos os seguintes dados:

- Identificação do semestre;

- Identificação da turma;
- Identificação do aluno;
- Unidade presencial;
- Curso de graduação em que está matriculado;
- Sexo do aluno;
- Idade do aluno;
- Nota de cada avaliação da disciplina, data final de entrega estipulada pelo professor e data de entrega efetivada pelo aluno, perfazendo um total de cinco atividades;
- Notas das avaliações: grau1, grau2, substituição de grau e média final;
- Datas do primeiro e do último acesso à sala virtual da disciplina;
- Para cada uma das 19 semanas do semestre letivo:
 - Número de acessos ao ambiente virtual;
 - Número de acessos à ferramenta aula (conteúdo *on-line*);
 - Número de acessos à biblioteca virtual (materiais complementares);
 - Número de acessos ao fórum;
 - Número de postagens no fórum;
 - Número de e-mails enviados;
 - Número de e-mails recebidos.

Os dados históricos foram solicitados à TI da Universidade, conforme especificação apresentada no anexo 6. Muitos dos dados, como as quantidades de acesso às ferramentas especificadas, não estavam armazenados de forma organizada, sendo necessário analisar os *logs* da época de ocorrência das turmas para obter as informações.

Embora todas as informações solicitadas tenham sido entregues, os dados vieram bastante desorganizados, em muitos casos com distribuição diversa da

solicitada, sendo necessário remontá-los em uma tabela única que permitisse uma visão integrada dos mesmos, em que cada linha representasse a totalidade das informações de um aluno e as colunas apresentassem cada um de seus atributos.

Ao reunir, inicialmente, os dados obtidos, cada aluno foi representado por 158 atributos. Todos os atributos foram validados, possibilitando o trabalho com uma tabela bastante completa. Na fase de preparação de dados, novos atributos foram criados a partir dos existentes, de forma a tornar as informações mais ricas, contextualizadas ou resumidas.

Como os cursos são muitos, criou-se o atributo Área. Na Universidade, os cursos são agrupados em seis grandes áreas (Ciências Agrárias e Ambientais; Ciências Jurídicas; Ciências Sociais Aplicadas; Educação, Ciências e Artes; Saúde e Bem-Estar Social; e Tecnologia e Computação) o que pode contribuir para uma visão sistêmica.

As datas de primeiro e último acesso ao AVA foram transformadas em um número (de 1 a 19) que indicava a semana do curso em que tais acessos ocorreram, comparando com as datas do calendário do curso no semestre específico da turma. Vale destacar que todas as datas foram substituídas, na etapa de preparação dos dados, pelo número da semana a que correspondiam, conforme o calendário a que pertenciam. As datas variam ao longo dos semestres e esta informação direta não permitia, na forma apresentada, buscar padrões ou relações entre os semestres minerados e os futuros. Para cada tarefa avaliativa entregue, comparou-se a data de entrega do aluno com a data de entrega prevista no sistema, verificando se a atividade foi entregue no prazo, com atraso ou não foi entregue.

As contagens semanais (acessos ao ambiente, às aulas, à biblioteca, ao fórum, postagens em fóruns, e-mails recebidos e enviados), foram trabalhadas em intervalos mensais, agrupadas, em geral, de quatro em quatro semanas, transformadas em total mensal (soma dos acessos semanais no mês indicado) e frequência mensal (0 – sem acesso; 1 – acessos em apenas uma das semanas; 2 – acessos em apenas duas das semanas; 3 – acessos em três das semanas; 4 – acessos em todas as semanas do mês analisado).

As notas de G1 foram validadas a partir da soma das três primeiras atividades avaliativas. A nota da prova presencial de G2 foi obtida a partir da nota final de G2 disponível, subtraindo-se as notas das atividades 4 e 5 que constavam nas avaliações realizadas no AVA para a composição do grau. A partir dos acessos dos alunos, das entregas das atividades e da média final, criou-se o campo Resultado, contendo as classes Aprovado, Reprovado, Evadido ou Sem Acesso.

A tabela final, enriquecida, alcançou o número de 230 atributos. Para a MD, foram selecionados 87 atributos, já que alguns foram resumidos ou agrupados em novos atributos gerados durante a fase de preparação dos dados. Os 87 atributos finais minerados estão disponíveis da tabela 4.1. A tabela traz as seguintes informações:

- Tipo de dado para a mineração: *id* (campo de identificação do registro, no caso o identificador do aluno), *label* (atributo que contém as classes que devem constar no conseqüente da regra de classificação, no caso o resultado do aluno), *regular* (demais atributos considerados, que serão avaliados para a composição do antecedente da regra).
- Nome dos campos de dados utilizados (a descrição dos campos está disponível no anexo 8).
- Tipo do valor: nominal (categorias) e numéricos (inteiro e real).
- Dados estatísticos: para os campos do tipo nominal, mostra a moda (mode), isto é, o valor (categoria) mais freqüente, seguido da quantidade de vezes em que aparece no conjunto total de registros; para os atributos de tipos numéricos, apresenta a média (avg) e o desvio padrão.
- Faixa de valores: para os campos do tipo nominal, apresenta cada valor possível e a quantidade de vezes em que se faz presente no grupo de dados analisados; para os atributos numéricos, apresenta a faixa de valores que o atributo pode assumir (a partir dos valores mínimo e máximo identificados nos dados trabalhados).

Tabela 4.1 – Dados preparados para a Mineração

	Tipo MD	Nome do Campo	Tipo do Valor	Dados Estatísticos	Faixa de Valores
1	id	Id_Us	nominal	mode = 42117 (1)	
2	label	Resultado	nominal	mode = Aprovado (1286)	Aprovado (1286), Evadido (212), Reprovado (52), Sem_Acesso (14)
3	regular	Unidade	nominal	mode = Canoas (892)	Canoas (892), SantaMaria (69), CachoeiradoSul (84), Gravataí (162), Carazinho (102), Torres (120), SãoJerônimo (56), Guaíba (79)
4	regular	Curso	nominal	mode = EDUCACAOFISICA (207)	Vide Quadro 4.1.
5	regular	Área	nominal	mode = EducaçãoCiênciaseArtes (473)	CiênciasSociaisAplicadas (260), TecnologiaeComputação (222), CiênciasJurídicas (207), EducaçãoCiênciaseArtes (473), SaúdeeBem-EstarSocial (363), CiênciasAgrárias (39)
6	regular	Sexo	nominal	mode = Feminino (876)	Feminino (876), Masculino (688)
7	regular	Idade	integer	avg = 26.264 +/- 7.480	[17.000 ; 58.000]
8	regular	Nota1	real	avg = 1.939 +/- 0.783	[0.000 ; 2.500]
9	regular	Nota2	real	avg = 1.557 +/- 0.733	[0.000 ; 2.500]
10	regular	Nota3	real	avg = 3.233 +/- 1.554	[0.000 ; 5.000]
11	regular	NAt_G1	integer	avg = 2.618 +/- 0.836	[0.000 ; 3.000]
12	regular	Entrega1	nominal	mode = Sim (1399)	Sim (1399), Nao (155), Atraso (10)
13	regular	Entrega2	nominal	mode = Sim (1351)	Sim (1351), Nao (194), Atraso (19)
14	regular	Entrega3	nominal	mode = Sim (1324)	Sim (1324), Nao (221), Atraso (19)
15	regular	Nota4	real	avg = 0.654 +/- 0.339	[0.000 ; 1.000]
16	regular	Nota5	real	avg = 1.257 +/- 0.682	[0.000 ; 2.000]
17	regular	Nota6	real	avg = 3.900 +/- 1.794	[0.000 ; 7.000]
18	regular	NAt_G2	integer	avg = 2.543 +/- 0.978	[0.000 ; 3.000]
19	regular	Entrega4	nominal	mode = Sim (1287)	Sim (1287), Nao (261), Atraso (16)
20	regular	Entrega5	nominal	mode = Sim (1272)	Sim (1272), Nao (269), Atraso (23)
21	regular	Pri_Ac	integer	avg = 1.936 +/- 1.428	[0.000 ; 11.000]
22	regular	Ult_Ac	integer	avg = 17.357 +/- 2.919	[0.000 ; 19.000]
23	regular	AG_TM1	integer	avg = 9.836 +/- 9.306	[0.000 ; 94.000]
24	regular	FREQ_M1	integer	avg = 2.691 +/- 1.278	[0.000 ; 4.000]
25	regular	AG_TM2	integer	avg = 13.102 +/- 9.945	[0.000 ; 109.000]
26	regular	FREQ_M2	integer	avg = 3.230 +/- 1.024	[0.000 ; 4.000]
27	regular	AG_TM3	integer	avg = 8.700 +/- 7.658	[0.000 ; 83.000]
28	regular	FREQ_M3	integer	avg = 2.319 +/- 0.883	[0.000 ; 3.000]
29	regular	AG_TM4	integer	avg = 8.419 +/- 8.093	[0.000 ; 102.000]
30	regular	FREQ_M4	integer	avg = 2.684 +/- 1.269	[0.000 ; 4.000]
31	regular	AG_TM5	integer	avg = 10.500 +/- 9.332	[0.000 ; 89.000]
32	regular	FREQ_M5	integer	avg = 2.666 +/- 1.146	[0.000 ; 4.000]
33	regular	AB_TM1	integer	avg = 18.895 +/- 23.754	[0.000 ; 245.000]
34	regular	AB_FM1	integer	avg = 2.396 +/- 1.341	[0.000 ; 4.000]
35	regular	AB_TM2	integer	avg = 24.322 +/- 31.932	[0.000 ; 418.000]
36	regular	AB_FM2	integer	avg = 2.960 +/- 1.207	[0.000 ; 4.000]
37	regular	AB_TM3	integer	avg = 14.236 +/- 18.106	[0.000 ; 284.000]
38	regular	AB_FM3	integer	avg = 2.051 +/- 1.024	[0.000 ; 3.000]
39	regular	AB_TM4	integer	avg = 15.893 +/- 21.557	[0.000 ; 260.000]
40	regular	AB_FM4	integer	avg = 2.510 +/- 1.347	[0.000 ; 4.000]

(continua na próxima página)

	Tipo MD	Nome do Campo	Tipo do Valor	Dados Estatísticos	Faixa de Valores
41	regular	AB_TM5	integer	avg = 15.754 +/- 18.608	[0.000 ; 236.000]
42	regular	AB_FM5	integer	avg = 2.407 +/- 1.196	[0.000 ; 4.000]
43	regular	AF_TM1	integer	avg = 11.419 +/- 15.635	[0.000 ; 185.000]
44	regular	AF_FM1	integer	avg = 2.044 +/- 1.409	[0.000 ; 4.000]
45	regular	AF_TM2	integer	avg = 11.029 +/- 16.253	[0.000 ; 243.000]
46	regular	AF_FM2	integer	avg = 2.148 +/- 1.375	[0.000 ; 4.000]
47	regular	AF_TM3	integer	avg = 6.159 +/- 10.530	[0.000 ; 132.000]
48	regular	AF_FM3	integer	avg = 1.380 +/- 1.127	[0.000 ; 3.000]
49	regular	AF_TM4	integer	avg = 7.453 +/- 11.732	[0.000 ; 164.000]
50	regular	AF_FM4	integer	avg = 1.709 +/- 1.419	[0.000 ; 4.000]
51	regular	AF_TM5	integer	avg = 3.437 +/- 6.328	[0.000 ; 84.000]
52	regular	AF_FM5	integer	avg = 1.128 +/- 1.130	[0.000 ; 4.000]
53	regular	PF_TM1	integer	avg = 1.356 +/- 1.710	[0.000 ; 14.000]
54	regular	PF_FM1	integer	avg = 0.932 +/- 0.954	[0.000 ; 4.000]
55	regular	PF_TM2	integer	avg = 1.094 +/- 1.941	[0.000 ; 33.000]
56	regular	PF_FM2	integer	avg = 0.671 +/- 0.854	[0.000 ; 4.000]
57	regular	PF_TM3	integer	avg = 0.561 +/- 1.122	[0.000 ; 12.000]
58	regular	PF_FM3	integer	avg = 0.503 +/- 0.742	[0.000 ; 3.000]
59	regular	PF_TM4	integer	avg = 0.659 +/- 1.076	[0.000 ; 9.000]
60	regular	PF_FM4	integer	avg = 0.497 +/- 0.704	[0.000 ; 4.000]
61	regular	PF_TM5	integer	avg = 0.189 +/- 0.664	[0.000 ; 8.000]
62	regular	PF_FM5	integer	avg = 0.127 +/- 0.391	[0.000 ; 3.000]
63	regular	EE_TM1	integer	avg = 0.053 +/- 0.296	[0.000 ; 5.000]
64	regular	EE_FM1	integer	avg = 0.043 +/- 0.219	[0.000 ; 2.000]
65	regular	EE_TM2	integer	avg = 0.124 +/- 0.610	[0.000 ; 11.000]
66	regular	EE_FM2	integer	avg = 0.087 +/- 0.345	[0.000 ; 3.000]
67	regular	EE_TM3	integer	avg = 0.127 +/- 0.580	[0.000 ; 9.000]
68	regular	EE_FM3	integer	avg = 0.091 +/- 0.354	[0.000 ; 3.000]
69	regular	EE_TM4	integer	avg = 0.047 +/- 0.284	[0.000 ; 4.000]
70	regular	EE_FM4	integer	avg = 0.040 +/- 0.227	[0.000 ; 3.000]
71	regular	EE_TM5	integer	avg = 0.120 +/- 0.532	[0.000 ; 9.000]
72	regular	EE_FM5	integer	avg = 0.091 +/- 0.355	[0.000 ; 3.000]
73	regular	ER_TM1	integer	avg = 6.003 +/- 5.242	[0.000 ; 19.000]
74	regular	ER_FM1	integer	avg = 2.256 +/- 1.664	[0.000 ; 4.000]
75	regular	ER_TM2	integer	avg = 6.491 +/- 5.722	[0.000 ; 32.000]
76	regular	ER_FM2	integer	avg = 2.972 +/- 1.688	[0.000 ; 4.000]
77	regular	ER_TM3	integer	avg = 4.714 +/- 3.744	[0.000 ; 26.000]
78	regular	ER_FM3	integer	avg = 2.067 +/- 1.219	[0.000 ; 3.000]
79	regular	ER_TM4	integer	avg = 3.170 +/- 2.056	[0.000 ; 9.000]
80	regular	ER_FM4	integer	avg = 2.098 +/- 1.276	[0.000 ; 4.000]
81	regular	ER_TM5	integer	avg = 7.556 +/- 4.913	[0.000 ; 20.000]
82	regular	ER_FM5	integer	avg = 3.002 +/- 1.688	[0.000 ; 4.000]
83	regular	G1	real	avg = 6.750 +/- 2.554	[0.000 ; 10.000]
84	regular	G2	real	avg = 5.811 +/- 2.477	[0.000 ; 10.000]
85	regular	S1	real	avg = 0.261 +/- 1.254	[0.000 ; 9.000]
86	regular	S2	real	avg = 0.552 +/- 1.799	[0.000 ; 9.500]
87	regular	MF	real	avg = 6.325 +/- 2.194	[0.000 ; 9.900]

Quadro 4.1 – Valores do Campo Curso, com a respectiva quantidade de alunos

ADMINISTRACAO (141)	ENFERMAGEM (80)	MATEMATICA (27)
AGRONOMIA (17)	ENGENHARIAAGRICOLA (3)	MATEMATICAAPLICINFORMATICA (5)
ANALISEEDESDESISTEMAS (2)	ENGENHARIAAMBIENTAL (11)	MEDICINA (56)
ARQUITETURAEURBANISMO (13)	ENGENHARIAACIVIL (10)	MEDICINAVETERINÁRIA (18)
ARTESVISUAIS (7)	ENGENHARIADEPETROLEO (5)	ODONTOLOGIA (50)
AUTOMACAOINDUSTRIAL (1)	ENGENHARIADEPLASTICOS (8)	OPTOMETRIA (5)
BIOLOGIA (44)	ENGENHARIAELETRICA (18)	PEDAGOGIA (73)
BIOMEDICINA (33)	ENGENHARIA MECANICA (18)	PSICOLOGIA (63)
CIENCIA D A COMPUTACAO (35)	ENGENHARIA MEC AUTOMOTIVA (12)	QUIMICA (9)
CIENCIA POLITICA (6)	ENGENHARIA QUIMICA (10)	QUIMICA INDUSTRIAL (34)
CIENCIAS CONTÁBEIS (19)	ESTETICA ECOSMETICA (1)	SECRETARIADO EXECUTIVO (11)
CIENCIAS SOCIAIS (1)	FARMACIA (19)	SEGURANCANOTRABALHO (1)
COMUNICACAO SOCIAL-JORN (16)	FISICA (8)	SERVICOSOCIAL (37)
COMUNICACAO SOCIAL-PEP (12)	FISIOTERAPIA (39)	SISTEMAS DE INFORMACAO (51)
COMUNICACAO SOCIAL-RP (5)	FONO AUDIOLOGIA (17)	SUPERIORDETECNOLOGIA EM AGRON
DANCA (1)	GEOGRAFIA (11)	EGOCIO (1)
DESIGN (27)	HISTORIA (19)	TEOLOGIA (12)
DIREITO (207)	LETRAS-INGLES (8)	
EDUCACAO FISICA (207)	LETRAS-PORTUGUES (20)	

Uma análise inicial da tabela permite observar que o resultado mais freqüente é a aprovação, com nota final (MF) média de 6.325 e desvio padrão de +/- 2.194. A maioria dos estudantes é do sexo feminino, cursa Educação Física, graduação vinculada à área de Educação, Ciências e Artes. É possível, também, observar que as quantidades de acesso ao ambiente, aos materiais e aos fóruns têm médias e limites bem maiores do que as quantidades de envio de *e-mails* ou postagens nos fóruns, indicativo de que mesmo sem uma atitude explícita de comunicação no ambiente, os alunos freqüentam ativamente os diversos espaços da sala de aula virtual.

4.5.3 Mineração de Dados

A descoberta de conhecimento por meio da MD, para a geração de alertas, foi organizada em cinco momentos determinados, para geração de regras para cada um dos cinco meses do semestre letivo. Os dados demográficos juntamente com os dados referentes ao primeiro mês de interação no ambiente (mês um, englobando dados da semana 1 até a semana 4) e o resultado final alcançado (conhecido nos dados históricos) foram minerados para observar como as características dos alunos e seus comportamentos no ambiente, já no primeiro mês, se relacionavam (influenciavam) seus resultados finais, ou seja, como a relação entre esses dados permitia classificar os alunos. As regras geradas servem de base para, em edições posteriores da disciplina, classificar / prever o resultado dos alunos com os dados já do primeiro mês.

O processo de mineração foi realizado também ao final do mês dois (semana 8), mês três (semana 11), mês quatro (semana 15) e mês cinco (semana 19), de forma análoga ao descrito para o mês um, sempre incorporando os dados gerados no espaço de tempo do início de curso até a semana indicada. As minerações foram realizadas no ambiente *RapidMiner* (apresentado na seção 3.4.1), com o algoritmo *RuleLeaner* para gerar as regras de classificação (técnica descrita na seção 3.3.2). A tabela 4.2 organiza os atributos utilizados em cada uma das minerações realizadas e indica o anexo em que estão as regras obtidas, que foram utilizadas para a geração de alertas durante o acompanhamento de novas turmas.

Tabela 4.2 – Mineração de Dados: atributos considerados e regras geradas

Mês	No. de Atributos	Atributos	Regras Geradas
Mês1	20	Id_Us, Resultado, Unidade, Curso, Área, Sexo, Idade, Pri_Ac AG_TM1, FREQ_M1 AB_TM1, AB_FM1 AF_TM1, AF_FM1 PF_TM1, PF_FM1 EE_TM1, EE_FM1 ER_TM1, ER_FM1	Anexo 8 (74 regras)
Mês2	36	Id_Us, Resultado, Unidade, Curso, Área, Sexo, Idade, Pri_Ac Nota1, Nota2, Entrega1, Entrega2 AG_TM1, FREQ_M1, AG_TM2, FREQ_M2 AB_TM1, AB_FM1, AB_TM2, AB_FM2 AF_TM1, AF_FM1, AF_TM2, AF_FM2 PF_TM1, PF_FM1, PF_TM2, PF_FM2 EE_TM1, EE_FM1, EE_TM2, EE_FM2 ER_TM1, ER_FM1, ER_TM2, ER_FM2	Anexo 9 (77 regras)
Mês3	52	Id_Us, Resultado, Unidade, Curso, Área, Sexo, Idade, Pri_Ac, G1 Nota1, Nota2, Nota3, Nat_G1, Entrega1, Entrega2, Entrega3 AG_TM1, FREQ_M1, AG_TM2, FREQ_M2, AG_TM3, FREQ_M3 AB_TM1, AB_FM1, AB_TM2, AB_FM2, AB_TM3, AB_FM3 AF_TM1, AF_FM1, AF_TM2, AF_FM2, AF_TM3, AF_FM3 PF_TM1, PF_FM1, PF_TM2, PF_FM2, PF_TM3, PF_FM3 EE_TM1, EE_FM1, EE_TM2, EE_FM2, EE_TM3, EE_FM3 ER_TM1, ER_FM1, ER_TM2, ER_FM2, ER_TM3, ER_FM3	Anexo 10 (58 regras)
Mês4	64	Id_Us, Resultado, Unidade, Curso, Área, Sexo, Idade, Pri_Ac, G1 Nota1, Nota2, Nota3, Nat_G1, Entrega1, Entrega2, Entrega3 AG_TM1, FREQ_M1, AG_TM2, FREQ_M2, AG_TM3, FREQ_M3, AG_TM4, FREQ_M4 AB_TM1, AB_FM1, AB_TM2, AB_FM2, AB_TM3, AB_FM3, AB_TM4, AB_FM4 AF_TM1, AF_FM1, AF_TM2, AF_FM2, AF_TM3, AF_FM3, AF_TM4, AF_FM4 PF_TM1, PF_FM1, PF_TM2, PF_FM2, PF_TM3, PF_FM3, PF_TM4, PF_FM4 EE_TM1, EE_FM1, EE_TM2, EE_FM2, EE_TM3, EE_FM3, EE_TM4, EE_FM4 ER_TM1, ER_FM1, ER_TM2, ER_FM2, ER_TM3, ER_FM3, ER_TM4, ER_FM4	Anexo 11 (56 regras)
Mês5	87	Id_Us, Resultado, Unidade, Curso, Área, Sexo, Idade, Pri_Ac, Ult_Ac, G1, G2, S1, S2, MF Nota1, Nota2, Nota3, Nat_G1, Entrega1, Entrega2, Entrega3 Nota4, Nota5, Nota6, Nat_G2, Entrega4, Entrega5 AG_TM1, FREQ_M1, AG_TM2, FREQ_M2, AG_TM3, FREQ_M3, AG_TM4, FREQ_M4, AG_TM5, FREQ_M5 AB_TM1, AB_FM1, AB_TM2, AB_FM2, AB_TM3, AB_FM3, AB_TM4, AB_FM4, AB_TM5, AB_FM5 AF_TM1, AF_FM1, AF_TM2, AF_FM2, AF_TM3, AF_FM3, AF_TM4, AF_FM4, AF_TM5, AF_FM5 PF_TM1, PF_FM1, PF_TM2, PF_FM2, PF_TM3, PF_FM3, PF_TM4, PF_FM4, PF_TM5, PF_FM5 EE_TM1, EE_FM1, EE_TM2, EE_FM2, EE_TM3, EE_FM3, EE_TM4, EE_FM4, EE_TM5, EE_FM5 ER_TM1, ER_FM1, ER_TM2, ER_FM2, ER_TM3, ER_FM3, ER_TM4, ER_FM4, ER_TM5, ER_FM5	Anexo 12 (9 regras)

4.6 Geração de Alertas

A última etapa do processo de KDD é a interpretação/avaliação dos resultados, buscando utilizá-los para auxiliar na solução dos problemas que motivaram a mineração. Na pesquisa realizada, as regras de classificação obtidas na etapa de MD foram utilizadas para auxiliar no acompanhamento de novas turmas, alertando o professor sobre alunos em risco e suas características, com o propósito de servir de apoio à ação docente.

4.6.1 Turmas Acompanhadas

A partir das regras de classificação obtidas e da definição de indicadores, acompanhou-se quatro turmas de IC durante o semestre de 2009/1, totalizando 230 alunos, gerando alertas e realizando intervenções junto a eles. Os dados das turmas de acompanhamento foram solicitados à TI da Universidade, a cada período de tempo estipulado (semana 4, 8, 11, 16 e 19), conforme especificação apresentada no anexo 13.

A instituição iniciou a disponibilização dos dados a partir da semana 8, quando, então, após a limpeza e preparação dos dados, de forma análoga ao realizado com os dados históricos (seção 4.5.2), submeteu-se os dados às regras de classificação geradas pela MD (seção 4.5.3).

Para comparar as populações e, posteriormente, os resultados, desconsiderou-se em ambas os alunos que nunca acessaram as aulas. Assim, na população 1 (dados históricos), reduziu-se a análise a 1550 alunos e na população 2 (dados em acompanhamento), utilizando-se o mesmo critério, trabalhou-se com 230 alunos.

Os dados obtidos mostram uma população a ser acompanhada com perfil similar ao da população utilizada para a geração das regras. Os dados históricos continham alunos das diversas unidades da ULBRA no RS (em um total de 8), já os dados das turmas de acompanhamento possuíam representantes de 7 das unidades, ambas com predominância de alunos de Canoas.

Nos dados históricos, a moda (medida estatística) do curso é Direito, com 205 alunos, seguida do curso de Educação Física, com 204. Nas turmas acompanhadas em 2009/1, a moda é Educação Física, com 35 alunos, seguida de Direito, com 26. Nas duas populações, a área de cursos mais freqüente é a de Educação, Ciências e Artes.

O sexo feminino é mais comum nos dois grupos. Em relação à idade, a faixa etária é a mesma, de 17 a 58 anos, com média de 26,2 anos e desvio padrão em torno de 7. A tabela 4.3 compara os dados entre as populações utilizadas na pesquisa.

Tabela 4.3 – População 1 (Dados Históricos) x População 2 (Dados de Acompanhamento)

Atributos			Históricos		Acompanhamento	
Tipo	Nome	Valor Tipo	Estatística	Faixa	Estatística	Faixa
regular	Unidade	nominal	mode = Canoas (886)	Canoas (886), SantaMaria (68), CachoeiradoSul (84), Gravataí (160), Carazinho(102), Torres (117), SãoJerônimo (56),Guaíba (77)	mode = Canoas (124)	SantaMaria (31), SãoJerônimo (37), Canoas (124), Torres (9), Carazinho (9), Gravataí (13), Guaíba (7)
regular	Curso	nominal	mode = DIREITO (205)	55 cursos. 2º curso em número de alunos: EDUCACAOFISICA (204)	mode = EDUCACAOFISICA (35)	45 cursos. 2º curso em número de alunos: DIREITO (26)
regular	Área	nominal	mode = Educação Ciências e Artes (467)	CiênciasSociaisAplicadas (260), TecnologiaeComputação (219), CiênciasJurídicas (205), EducaçãoCiênciaseArtes (467), SaúdeeBem-EstarSocial (360), CiênciasAgrárias (39)	mode = Educação Ciências e Artes (68)	CiênciasJurídicas (26), EducaçãoCiênciaseArtes (68), SaúdeeBem-EstarSocial (44), TecnologiaeComputação (44), CiênciasSociaisAplicadas (42), CiênciasAgrárias (6)
regular	Sexo	nominal	mode = Feminino (868)	Feminino (868), Masculino (682)	mode = Feminino (117)	Feminino (117), Masculino (113)
regular	Idade	Integer	avg = 26.210 +/- 7.415	[17.000 ; 58.000]	avg = 26.226 +/- 7.126	[17.000 ; 58.000]

4.6.2 Alertas Gerados e Exemplos de Intervenções

Para o experimento realizado, foram definidos como indicadores de acompanhamento:

- Entrega das atividades avaliativas – alerta fixo, a partir do prazo/evento definido pelo professor. Com dois dias de antecedência do prazo para

entrega de cada tarefa, o professor foi notificado sobre quem eram os alunos que ainda estavam com a tarefa pendente;

- Regras de classificação – alertas do sistema, a partir da aplicação das regras resultantes da MD. No final do mês 2, 3 e 4, o professor foi notificado sobre grupos de alunos com tendência à aprovação, à reprovação ou à evasão, indicando a regra da mineração de dados que permitiu classificar cada grupo de alunos.

Cada alerta, com base no evento ou regra específica, trouxe a regra e o grupo de alunos, notificando o professor e sugerindo o contato com os estudantes. A partir da análise do professor, outras ações puderam ser tomadas, como, por exemplo, a ampliação de prazos ou a oferta de materiais complementares.

4.6.2.1 Alertas sobre Atividades Avaliativas Pendentes

De forma fixa, para cada uma das atividades avaliativas, representadas de A1 a A5 na tabela 4.4, com dois dias de antecedência para o prazo final de entrega, o professor foi alertado sobre o grupo de alunos que ainda não havia enviado sua atividade. Com base no alerta, nas características da atividade e do grupo de alunos, o professor pode enviar uma comunicação mais adequada para estes alunos.

A tabela 4.4, para cada uma das atividades avaliativas da disciplina de IC, mostra o número de alunos que já havia entregue a tarefa com dois ou mais dias de antecedência, o número de alunos que recebeu o alerta sobre o prazo da atividade, a quantidade de alunos que realizou a tarefa neste prazo final, o total de alunos que entregou as atividades e a mensagem enviada.

O número de alunos que deixou para enviar as tarefas avaliativas no prazo final é bastante elevado. Não é possível, em um primeiro olhar, afirmar que os alertas lembrando prazos e atividades pendentes tenham levado o aluno a fazê-las ou que teriam deixado de realizá-las caso não fossem contatados. Mas, os índices de aprovação foram superiores e os índices de evasão menores em comparação àqueles obtidos nas turmas anteriores, sem os alertas, conforme teste estatístico

(seção 4.7). Além disso, os alunos, em pesquisa de satisfação com os alertas (seção 4.8), afirmaram que as mensagens os auxiliavam na lembrança de atividades pendentes, como a realização de atividades avaliativas – 85,45% concordaram plenamente com esta afirmação, seguido de 7,27% que também concordaram.

Tabela 4.4 – Alertas fixos: atividades avaliativas pendentes

	No. Entrega 2d. Antes	No. de Alertados	No. Entrega 2d. Finais	No. Entrega Total	Mensagem encaminhada aos alunos
A1	132	98	85	217	Prezados Alunos, Após o estudo dos capítulos 1 e 2, realize a Atividade 1 - postagem de síntese/reflexão em fórum, conforme orientação incluída no mesmo. Prazo final: 22/04. Pontuação: 2,5.
A2	127	103	82	209	Prezados Alunos, Após o estudo dos capítulos 3, 4 e 5, realize a Atividade 2 - ler o texto motivacional da atividade e realizar as pesquisas indicadas nas questões de avaliação. Atenção: é preciso dedicar certo tempo para realizá-las, não deixe para a última hora. Prazo: 22/04. Pontuação: 2,5.
A3	184	46	18	202	Prezado(a) Aluno(a), Para que possamos dedicar mais tempo para os estudos e as pesquisas relacionadas, decidimos ampliar a possibilidade de entrega da Atividade 3. Assim, você tem a oportunidade de fazê-la e enviá-la até o dia 10/05 (domingo próximo) às 22h. Cabe ressaltar que a atividade vale 5,0 pontos para a G1, o que representa uma parte significativa da nota. Aproveite a oportunidade e engaje-se nos estudos. Esperamos sua atividade! Logo após, iniciaremos os trabalhos de G2 e manter um ritmo de estudos é parte importante para o processo de aprendizagem! Atenciosamente, Profa. Adriana
A4	138	92	75	213	Prezado(a) Estudante! Após finalizar o estudo dos capítulos 11, 12 e 13, realize a Atividade 4, que corresponde a 2 pontos na nota da G2. A atividade proposta é interpretativa, buscando que você possa aplicar os conhecimentos na análise das situações apresentadas! A atividade está ativa até 26/6. Se você tiver alguma dificuldade, entre em contato! Bons estudos! Estamos aguardando sua entrega! Abraços!!! Profa. Adriana
A5	143	87	70	213	Prezado(a) Estudante! Após finalizar o estudo do capítulo 14, além de participar do Fórum sobre Direitos Autorais, realize a Atividade 5, que corresponde a 1 ponto na nota da G2. A atividade está ativa até 26/6. Se você tiver alguma dificuldade, entre em contato! Bons estudos! Estamos aguardando sua entrega! Abraços!!! Profa. Adriana

4.6.2.2 Alertas Gerados pela Mineração de Dados

Após o final dos meses 2, 3 e 4, o professor foi alertado sobre a tendência de resultado dos seus alunos (aprovação, reprovação ou evasão), com a identificação da regra e dados dos alunos, para decisão sobre o tipo de intervenção.

Observações sobre o Mês 1

Os dados do mês 1 das turmas de acompanhamento não foram submetidos às regras obtidas na mineração dos dados históricos, para a geração dos alertas, por terem sido liberados pela instituição após o prazo necessário para sua utilização. É importante ressaltar também que os dados do mês 1 tendem a ser menos precisos para definir a classificação (predizer o resultado final), já que o volume de informações disponíveis no primeiro mês é menor e neste período os alunos estão ingressando e se ambientando, tanto à disciplina quanto, em boa parte das vezes, à modalidade a distância.

De qualquer modo, ao analisar as regras geradas, é possível observar padrões interessantes a partir das características demográficas dos alunos e de seus primeiros comportamentos de acesso. Para exemplificar, destacam-se algumas regras, no quadro 4.2, retiradas do anexo 8. As regras são aplicadas em seqüência para a classificação dos alunos, o que significa dizer que, uma vez classificado, o aluno não é submetido às regras posteriores.

Quadro 4.2 – Exemplos de Regras Mês 1

1.	if AF_TM1 > 6.500 and AG_TM1 > 25.500 and AF_TM1 <= 53 then Aprovado
4.	if FREQ_M1 > 1.500 and AB_TM1 > 17.500 and AF_TM1 > 26.500 and AB_TM1 <= 93.500 and AF_TM1 <= 32.500 then Aprovado
6.	if FREQ_M1 > 1.500 and AF_TM1 > 13.500 and Idade > 26.500 and Idade <= 35.500 then Aprovado
7.	if FREQ_M1 > 1.500 and AF_TM1 > 13.500 and AF_FM1 > 2.500 and Curso = DIREITO then Aprovado
24.	if AG_TM1 > 4.500 and AG_TM1 <= 12.500 and Pri_Ac > 1.500 and PF_TM1 <= 1.500 and AB_TM1 > 5.500 and AB_TM1 <= 19.500 and ER_TM1 > 1.500 and ER_TM1 <= 14.500 and Área = EducaçãoCiênciaseArtes then Aprovado

A regra 1, obtida pela mineração dos dados históricos, indica que se o acesso total à sala virtual no primeiro mês foi superior a 25 (AG_TM1 > 25.500) e o acesso total aos fóruns esteve entre 7 (AF_TM1 > 6.500) e 53 (AF_TM1 <= 53) vezes, o

aluno tende a se enquadrar da classe Aprovado. A regra 4 evidencia questões de frequência ao ambiente em no mínimo 2 das quatro semanas ($FREQ_M1 > 1.500$), com acessos às aulas e à biblioteca entre 18 ($AB_TM1 > 17.500$) e 93 ($AB_TM1 \leq 93$) e acessos aos fóruns entre 27 ($AF_TM1 > 26.500$) e 32 ($AF_TM2 \leq 32.500$) vezes como padrão para enquadramento na classe Aprovado.

As regras 6, 7 e 24 trazem a combinação de mais elementos, como faixa etária, curso, área do curso do aluno, frequência de acesso a ferramentas diversificadas do AVA, semana em que foi realizado o primeiro acesso, quantidade de postagens em fóruns, quantidade de e-mails recebidos, para identificar alunos que tendem a aprovar. As regras baseiam-se em padrões identificados nas ocorrências anteriores da disciplina, relacionando os resultados obtidos pelos alunos aos dados disponíveis no período de tempo considerado.

Em função do grande número de regras aplicadas e alertas gerados ao final dos meses 2, 3 e 4, os dados foram agrupados para a análise e alguns exemplos das intervenções realizadas a partir dos alertas são apresentados a seguir.

Alertas do Mês 2 e Resultados

As regras aplicadas (anexo 9) analisam 34 atributos regulares de cada aluno das turmas acompanhadas. A tabela 4.5 apresenta o número de alunos classificados em cada classe (aprovação, evasão e reprovação), as regras que os classificaram e observações sobre as regras e sobre os resultados obtidos ao final do semestre, após a conclusão do experimento.

O professor, ao ser alertado sobre os grupos de alunos em cada uma das três classificações, pode, por exemplo, emitir uma mensagem única para o grupo desejado, ou analisar os subgrupos vinculados a cada regra de uma classificação específica, de forma a observar mais detalhadamente as especificidades de cada grupo. Dos 46 alunos acompanhados de maneira mais próxima (44 classificados como possíveis evasores e 2 como possíveis reprovados), 38 obtiveram aprovação ao final do curso.

Tabela 4.5 – Classificação Mês 2

Classificação: Tendência a	No. de Alunos classificados	Regras Empregadas	Observações
Aprovação	184	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17, 19, 20, 21, 22, 24, 27, 34, 36, 38, 41, 43, 44, 46, 48, 56, 60, 74.	Total de regras em que houve classificação: 31. Predomínio das regras de 1 a 17, classificando 144 dos 184 elementos.
Evasão	44	28, 33, 37, 40, 42, 45, 47, 49, 51, 59.	Total de regras em que houve classificação: 10. Total destes alunos que obtiveram aprovação ao final do curso: 36. Total destes alunos evadidos ao final do curso: 6. Total destes alunos reprovados ao final do curso: 2.
Reprovação	2	57, 61.	Total de regras em que houve classificação: 2. Total destes alunos que obtiveram aprovação ao final do curso: 2.

Basicamente, havia dois grandes indicadores nas regras que classificaram os alunos como possíveis evasores ou reprovados: baixa frequência de acesso aos fóruns e baixo rendimento nas primeiras atividades. Para envolvê-los em seus estudos, de forma a participarem das discussões da turma, bem como retomar os estudos relacionados, foram enviadas mensagens especificamente para esses alunos, colocando a importância das leituras complementares e da participação e leitura dos fóruns de discussão. O quadro 4.3 apresenta um exemplo de mensagem enviada.

Quadro 4.3 – Exemplo de mensagem enviada aos alunos após mês 2

Prezados Alunos,

Após o Estudo dos capítulos 6, 7 e 8, bem como da leitura do texto "O que é científico?" - participe no fórum sobre esses capítulos, que não vale nota diretamente, mas auxilia na sistematização do conhecimento e demonstra o engajamento do aluno no estudo da disciplina. Além disso, leia a contribuição dos colegas!

Esperamos que todos tenham êxito na construção dos conhecimentos trabalhados! Imprimam ritmo de estudo e organização, sabendo que os professores estão aqui para apoiá-los nos seus percursos de aprendizagem!

Abraços,

Profa. Adriana

Alertas do Mês 3 e Resultados

As regras aplicadas (anexo 10) analisam 50 atributos regulares de cada aluno das turmas acompanhadas. A tabela 4.6 apresenta o número de alunos classificados em cada classe (aprovação, evasão e reprovação), as regras que os classificaram e observações sobre as regras e sobre os resultados obtidos ao final do semestre, após a conclusão do experimento.

Tabela 4.6 – Classificação Mês 3

Classificação: Tendência a	No. de Alunos classificados	Regras Empregadas	Observações
Aprovação	197	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 19, 20, 22, 24, 26, 30, 31, 33, 35, 44, 48, 50	Total de regras em que houve classificação: 28. Predomínio das regras de 1 a 8, classificando 140 dos 197 elementos.
Evasão	25	15, 18, 21, 23, 28, 34, 36, 38, 41, 47	Total de regras em que houve classificação: 10. Total destes alunos que obtiveram aprovação ao final do curso: 17. Total destes alunos evadidos ao final do curso: 7. Total destes alunos reprovados ao final do curso: 1.
Reprovação	8	37, 40, 45, 49, 51, 52	Total de regras em que houve classificação: 6. Total destes alunos que obtiveram aprovação ao final do curso: 7. Total destes alunos evadidos ao final do curso: 1. Total destes alunos reprovados ao final do curso: 0.

As grandes questões ligadas a maioria das classificações dos alunos como evasores ou reprovados apresentava reflexo ainda do primeiro acesso ao ambiente, feito tardiamente. Além disso, baixo acesso geral ao ambiente, baixa nota de G1 e a ausência de postagens nos fóruns estavam entre os antecedentes de muitas das regras de mau prognóstico. Resgatar o aluno é essencial e uma das formas de fazer isso é envolvê-los em atividades diferenciadas. O quadro 4.4 apresenta um exemplo de mensagem enviada, incentivando os alunos a participar do estudo de caso proposto.

Quadro 4.4 – Exemplo de mensagem enviada aos alunos após mês 3

<p>Prezado(a) Estudante,</p> <p>A nossa discussão no fórum, sobre Direitos Autorais, é bastante relevante. Compreender a temática e conhecer a lei que trata do assunto em nosso país nos auxiliará em nossos trabalhos acadêmicos e mesmo em situações do nosso cotidiano, em que tantos não respeitam os direitos dos autores de produção intelectual.</p> <p>Estamos em fase de debates no fórum e gostaríamos de ouvir sua opinião. Participe!</p> <p>Dica importante: Leia as contribuições já postadas no fórum, pois as reflexões dos colegas podem auxiliar em uma compreensão mais ampla sobre os Direitos Autorais!</p> <p>Bons estudos e aguardo a sua contribuição!</p> <p>Abraços,</p> <p>Profa. Adriana</p>
--

Alertas do Mês 4 e Resultados

As regras aplicadas (anexo 11) analisam 62 atributos regulares de cada aluno das turmas acompanhadas. A tabela 4.7 apresenta o número de alunos classificados em cada classe (aprovação, evasão e reprovação), as regras que os classificaram e observações sobre as regras e sobre os resultados obtidos ao final do semestre, após a conclusão do experimento.

Tabela 4.7 – Classificação Mês 4

Classificação: Tendência a	No. de Alunos classificados	Regras Empregadas	Observações
Aprovação	170	1, 2, 3, 5, 8, 9, 10, 11, 12, 13, 14, 15, 18, 20, 21, 26, 31, 48, 52	Total de regras em que houve classificação: 19. Predomínio das regras de 1 a 11, classificando 125 dos 170 elementos.
Evasão	41	13, 16, 19, 25, 29, 30, 33, 42, 45	Total de regras em que houve classificação: 9. Total destes alunos que obtiveram aprovação ao final do curso: 30. Total destes alunos evadidos ao final do curso: 11. Total destes alunos reprovados ao final do curso: 0.
Reprovação	19	32, 35, 38, 40, 43, 47	Total de regras em que houve classificação: 6. Total destes alunos que obtiveram aprovação ao final do curso: 13. Total destes alunos evadidos ao final do curso: 3. Total destes alunos reprovados ao final do curso: 3.

Dos 60 alunos em evidência (41 classificados como possíveis evasores e 19 como reprovados), 43 obtiveram êxito na conclusão da disciplina. O mês 4 fechou na semana 16 e a prova presencial ocorreu na semana 17, sendo a substituição de grau, para os alunos que necessitaram, realizada na semana 19. O quadro 4.5 apresenta um exemplo de mensagem enviada, orientando os alunos a organizar seus estudos finais.

Quadro 4.5 – Exemplo de mensagem enviada aos alunos após mês 4

Prezado(a) Estudante,

Tudo bem?

No próximo sábado, dia 11/7, teremos a nossa **prova de substituição de grau de Instrumentalização Científica!** Está aí a oportunidade para, retomando os estudos, alcançar os objetivos propostos na disciplina e obter aprovação.

Nesta disciplina, desde o começo, combinados que o aluno deveria estabelecer seu ritmo de estudos, estando atento aos prazos. De qualquer modo, envio algumas dicas... Você pode organizar seus estudos em seis grandes grupos temáticos:

1. **Primeiro Grupo:** Capítulos 1 (*O Ato de Estudar*) e Capítulo 2 (*Conhecimento e Método*).
2. **Segundo Grupo:** Capítulo 3 (*Mecanismos de Busca, Diretórios e Banco de Dados*), Capítulo 4 (*Bibliotecas Virtuais, Enciclopédias e Portais*) e Capítulo 5 (*Periódicos e Portais*).
3. **Terceiro Grupo:** Capítulo 6 (*Projeto de Pesquisa*), Capítulo 7 (*A Pesquisa Científica*) e Capítulo 8 (*Tipos de Pesquisa*).
4. **Quarto Grupo:** Capítulo 9 (*O Qualis*) e Capítulo 10 (*A Plataforma Lattes*).
5. **Quinto Grupo:** Capítulo 11 (*Relatório de Pesquisa*) e Capítulo 12 (*Artigo e Linguagem Científica*).
6. **Sexto Grupo:** Capítulo 13 (*Aplicação das Normas da ABNT em um Editor de Textos*) e Capítulo 14 (*Ética e Aspectos Legais da Utilização da Informação*).

Minha recomendação, neste momento, é que você se organize para os estudos finais, rumo à substituição. Você pode estudar um grupo de conteúdos por dia, ou então dois ou três grupos por dia, dependendo da sua disponibilidade. É interessante, também, retomar os fóruns temáticos e as avaliações realizadas na NetAula.

Desejo um bom estudo e me coloco à disposição, mais uma vez, para auxiliar na resolução de dúvidas durante seus estudos.

Atenciosamente,

Prof. Adriana (Instrumentalização Científica)

Observações sobre o Mês 5

Para fins de fechamento do experimento, foram aplicadas as regras pós mês 5 (anexo 12) que analisam 85 atributos regulares de cada aluno das turmas acompanhadas. Todos os alunos foram classificados corretamente, ou seja, o resultado obtido por cada aluno acompanhado foi igual à classificação indicada pela aplicação das regras geradas na MD históricos.

4.7 Resultados

Para validar os resultados efetivos dos alertas, aplicou-se o teste de hipóteses para os percentuais de aprovação, de evasão e de reprovação dos alunos acompanhados, em relação aos dados históricos. A tabela 4.8 apresenta os números absolutos e percentuais de cada tipo de resultado alcançado pelos alunos, na população 1 (dados históricos dos semestres 2006/2 e 2008/2) e na população 2 (dados das turmas acompanhadas no semestre 2009/1).

Tabela 4.8 – População 1 (Dados Históricos) x População 2 (Dados Acompanhados)

	Histórico População 1		Acompanhamento População 2	
	Número de Alunos	Percentual	Número de Alunos	Percentual
Aprovado	1286	82,97	199	86,52
Reprovado	52	3,35	7	3,04
Evadido	212	13,68	24	10,43
Total	1550	100,00	230	100,00

Com amostras grandes de dados ($n_1 = 1550$ e $n_2 = 230$), considerou-se a Distribuição Normal para os testes. Tomando por base um nível de significância α de 0,1, o aumento ou a redução dos índices (teste unilateral) obtem-se o valor para Z_{TAB} de -1,28 (teste lateral à esquerda) e de 1,28 (teste lateral à direita). Aplica-se o teste estatístico apresentado na equação (1) para cálculo do Z_{CALC} . Na equação, p representa cada uma das proporções (da população 1 e da população 2, respectivamente, escrita em valores de 0 a 1); q é seu complemento, ou seja $1 - p$; n é o tamanho da amostra.

$$(1) \quad Z_{CALC} = \frac{(p_1 - p_2)}{\sqrt{\frac{p_1 \cdot q_1}{n_1} + \frac{p_2 \cdot q_2}{n_2}}}$$

Equação (1): teste cálculo do Z para comparação de proporções

4.7.1 Comparação entre os Índices de Aprovação

Quadro 4.6 – Teste de Hipóteses: comparação entre os percentuais de aprovação

<p>$H_0: p_1 = p_2$. Os percentuais de aprovação são iguais nas populações analisadas (dados históricos P1 e turmas com alertas P2).</p> <p>$H_1: p_1 < p_2$. O percentual de aprovados nas turmas sem alertas (dados históricos P1) foi inferior ao percentual de aprovados nas turmas com alertas P2.</p>
--

A partir do teste de hipóteses apresentado no quadro 4.6, buscou-se validar estatisticamente a hipótese H_1 . Em função do tipo de H_1 , é aplicado o teste unilateral à esquerda. O Z_{CALC} resultou no valor -1,45141, inferior ao Z_{TAB} de -1,28, com nível de significância $\alpha = 0,1$. Há evidências, portanto, para rejeitar H_0 e aceitar H_1 .

Ao nível de significância de 10% existem evidências para aceitar que o índice de aprovação nas turmas com os alertas foi significativamente superior ao índice de aprovação observado na amostra dos dados históricos.

4.7.2 Comparação entre os Índices de Reprovação

Quadro 4.7 – Teste de Hipóteses: comparação entre os percentuais de reprovação

<p>$H_0: p_1 = p_2$. Os percentuais de reprovação são iguais nas populações analisadas (dados históricos P1 e turmas com alertas P2).</p> <p>$H_1: p_1 > p_2$. O percentual de reprovados nas turmas sem alertas (dados históricos P1) foi superior ao percentual de reprovados nas turmas com alertas P2.</p>

A partir do teste de hipóteses apresentado no quadro 4.7, buscou-se validar estatisticamente a hipótese H_1 . Em função do tipo de H_1 , é aplicado o teste unilateral

à direita. O Z_{CALC} resultou no valor 0,253924, inferior ao Z_{TAB} de 1,28, com nível de significância $\alpha = 0,1$. Há evidências, portanto, para aceitar H_0 e rejeitar H_1 .

Ao nível de significância de 10% existem evidências para rejeitar que o índice de reprovação nas turmas com os alertas tenha sido significativamente inferior ao índice de reprovação observado na amostra dos dados históricos, aceitando a hipótese de igualdade estatística entre os índices obtidos nos dois grupos.

4.7.3 Comparação entre os Índices de Evasão

Quadro 4.8 – Teste de Hipóteses: comparação entre os percentuais de evasão

<p>$H_0: p_1 = p_2$. Os percentuais de evasão são iguais nas populações analisadas (dados históricos P1 e turmas com alertas P2).</p> <p>$H_1: p_1 > p_2$. O percentual de evadidos nas turmas sem alertas (dados históricos P1) foi superior ao percentual de evadidos nas turmas com alertas P2.</p>

A partir do teste de hipóteses apresentado no quadro 4.8, buscou-se validar estatisticamente a hipótese H_1 . Em função do tipo de H_1 , é aplicado o teste unilateral à direita. O Z_{CALC} resultou no valor 1,479774, superior ao Z_{TAB} de 1,28, com nível de significância $\alpha = 0,1$. Há evidências, portanto, para rejeitar H_0 e aceitar H_1 .

Ao nível de significância de 10% existem evidências para aceitar que o índice de evasão nas turmas com os alertas tenha sido significativamente inferior ao índice de evasão observado na amostra dos dados históricos.

4.8 Pesquisa com os Alunos

A pesquisa junto aos alunos, após a conclusão da disciplina, visava a evidenciar a percepção dos sujeitos sobre a vivência realizada em EAD, em múltiplos aspectos e, em especial, em relação às mensagens (alertas) recebidas. Sem caráter obrigatório e avaliativo, os alunos foram sensibilizados a responder o formulário enviado, que apresentava a motivação da pesquisa, conforme o anexo 15. Por se tratar de uma pesquisa com participação espontânea, portanto voluntária,

a análise foi realizada sobre o número de devoluções obtidas, totalizando 55 respondentes.

O questionário foi elaborado para avaliar as grandes dimensões que nortearam a proposta didática desenvolvida, ou seja, minimizar a *distância transacional*, de forma a ampliar o acompanhamento e a retenção dos alunos. Para tanto, buscou-se coletar as opiniões dos alunos em uma escala LiKert – expressa pelos valores Concordo Totalmente, Concordo, Indiferente, Discordo e Discordo Totalmente – a partir de questões que remetiam às dimensões *estrutura*, *diálogo* e *autonomia*, discutidas na seção 2.2. As questões e a tabulação dos resultados encontram-se na tabela 4.9.

As questões de 1 a 4 referiram-se a aspectos da estrutura da disciplina, à forma como as diversas etapas foram organizadas e colocadas em prática. A satisfação dos alunos com a estrutura que possibilita o acompanhamento das atividades é evidente, expressa pela soma dos índices de Concordo Totalmente e Concordo, variando de 94,55% a 100% nas questões pertinentes.

No gráfico (figura 4.14), as áreas azuis e vermelhas representam o percentual de alunos que avaliaram positivamente a estrutura da disciplina.

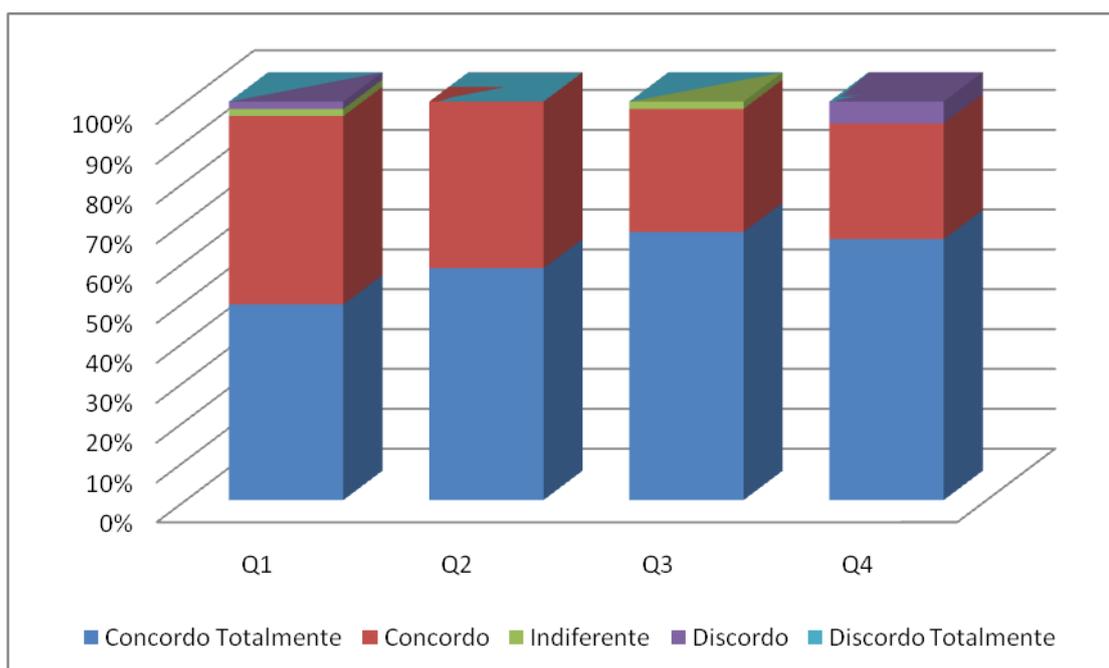


Figura 4.14 – Gráfico: questões relacionadas à dimensão estrutura

Tabela 4.9 – Pesquisa com os Alunos

		Concordo Totalmente	PCT (%)	Concordo	PC (%)	PCT+PC (%)	Indiferente	PI (%)	Discordo	PD (%)	Discordo Totalmente	PDT (%)	PD+PDT (%)	Respostas	Per. (%)
Q1	O Ambiente Virtual de Aprendizagem (NetAula) foi adequado para as demandas de estudar a distância.	27	49,09	26	47,27	96,36	1	1,82	1	1,82	0	0,00	1,82	55	100
Q2	O Plano de Ensino e Aprendizagem, juntamente com o Cronograma de Atividades do semestre, disponíveis na Biblioteca Virtual, traziam detalhes suficientes para o bom andamento da disciplina.	32	58,18	23	41,82	100,0	0	0,00	0	0,00	0	0,00	0,00	55	100
Q3	O Plano de Atividades, na parte central da sala virtual, facilitou a organização e o acesso aos materiais e tarefas, bem como a compreensão do encadeamento das atividades da disciplina.	37	67,27	17	30,91	98,18	1	1,82	0	0,00	0	0,00	0,00	55	100
Q4	Os critérios de avaliação e prazos de liberação, entrega e devolução de notas foram informados e cumpridos.	36	65,45	16	29,09	94,55	0	0,00	3	5,45	0	0,00	5,45	55	100
Q5	Os avisos gerais foram relevantes e atualizados.	39	70,91	13	23,64	94,55	2	3,64	1	1,82	0	0,00	1,82	55	100
Q6	Os professores utilizaram uma linguagem clara e objetiva ao comunicar-se.	43	78,18	10	18,18	96,36	1	1,82	1	1,82	0	0,00	1,82	55	100
Q7	As mensagens enviadas pelos professores eram lidas, seja no e-mail externo ou no e-mail interno da NetAula.	37	67,27	16	29,09	96,36	2	3,64	0	0,00	0	0,00	0,00	55	100
Q8	As mensagens enviadas por meio da NetAula eram recebidas no seu e-mail externo.		0,00	54	98,18	98,18		0,00	1	1,82		0,00	1,82	55	100
Q9	O recebimento de uma mensagem no e-mail externo o levava a acessar o ambiente.	34	61,82	14	25,45	87,27	5	9,09	2	3,64	0	0,00	3,64	55	100
Q10	As mensagens auxiliavam na lembrança de atividades pendentes, tais como leitura de materiais, realização de atividades avaliativas e participação em fóruns.	47	85,45	4	7,27	92,73	3	5,45	1	1,82	0	0,00	1,82	55	100
Q11	As mensagens foram importantes para a sua organização e engajamento na disciplina.	41	74,55	9	16,36	90,91	3	5,45	2	3,64	0	0,00	3,64	55	100
Q12	As mensagens auxiliaram no estabelecimento de um ritmo de estudos a distância.	35	63,64	13	23,64	87,27	5	9,09	2	3,64	0	0,00	3,64	55	100
Q13	As mensagens permitiram estabelecer um vínculo mais próximo com o professor, pois você percebeu que este estava atento às suas necessidades.	38	69,09	12	21,82	90,91	3	5,45	2	3,64	0	0,00	3,64	55	100
Q14	A quantidade e a frequência de mensagens foram adequadas ao momento da disciplina e à sua necessidade.	40	72,73	12	21,82	94,55	2	3,64	0	0,00	1	1,82	1,82	55	100
Q15	As mensagens auxiliaram na compreensão das relações entre os conteúdos e as atividades, bem como da importância de realizá-las para a aprendizagem dos assuntos trabalhados.	40	72,73	11	20,00	92,73	4	7,27	0	0,00	0	0,00	0,00	55	100
Q16	Houve interesse da sua parte para realizar as leituras dos conteúdos liberados e materiais complementares sugeridos.	32	58,18	13	23,64	81,82	4	7,27	6	10,91	0	0,00	10,91	55	100
Q17	Houve organização da sua parte para realizar os trabalhos solicitados e cumprir os prazos.	27	49,09	23	41,82	90,91	2	3,64	3	5,45	0	0,00	5,45	55	100
Q18	Você manteve uma postura ativa de aprendizagem, acessando regularmente o ambiente NetAula.	29	52,73	22	40,00	92,73	0	0,00	4	7,27	0	0,00	7,27	55	100
Q19	Você participou dos fóruns, visando aprofundamento de temas e esclarecimento de dúvidas.	31	56,36	19	34,55	90,91	3	5,45	2	3,64	0	0,00	3,64	55	100
Q20	Você leu as diversas postagens dos fóruns para aprender e contribuir com os colegas.	24	43,64	21	38,18	81,82	7	12,73	3	5,45	0	0,00	5,45	55	100
Q21	Você cursaria outras disciplinas ou cursos na modalidade a distância.	41	74,55	8	14,55	89,09	5	9,09	1	1,82	0	0,00	1,82	55	100

As questões de 5 a 15 referiam-se as formas de comunicação estabelecidas pelo professor, com diferentes objetivos e pretendiam coletar percepções de cada uma delas. O item 5 relacionava-se aos avisos gerais, disponíveis por meio do mural de recados do AVA, buscando saber se os alunos identificavam ali recados importantes e atualizados para o bom andamento da disciplina como um todo. 23,64% dos alunos concordaram e 70,91% concordaram totalmente em relação à relevância e atualização dos avisos. Para 3,64% esse item foi indiferente e apenas 1 aluno (1,82%) discordou.

Com professores capacitados e experientes em EAD, o item 6 mostra que 96,36% dos alunos acredita que os professores utilizaram uma linguagem clara e objetiva ao estabelecer comunicação, sendo que 78,18% concordam totalmente com esta afirmativa. Conforme respostas para a questão sete, 53 alunos (96,36%) liam sistematicamente as mensagens enviadas pelos professores, seja internamente no ambiente virtual de aprendizagem ou no seu *e-mail* pessoal externo (cópia), e somente para 2 alunos (3,64%) a leitura das mensagens era indiferente (figura 4.15). Apenas um dos respondentes não recebia cópia das mensagens em seu *e-mail* externo (item 8, com duas respostas possíveis: Concordo ou Discordo).

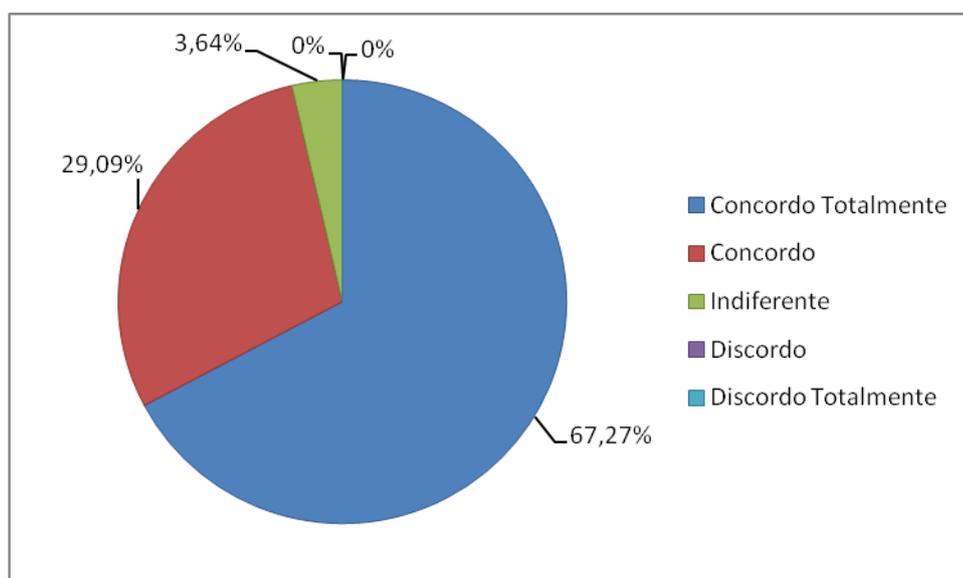


Figura 4.15 – Gráfico: Alunos que liam as mensagens enviadas pelo professor (Q7)

Os itens de 9 a 15 referem-se a percepção dos alertas no apoio às ações discentes. Para analisar as respostas, vale destacar que 100% dos alunos, conforme

resposta à questão dois, concordavam que o Plano de Ensino e Aprendizagem, juntamente com o Cronograma de Atividades do semestre, disponíveis na Biblioteca Virtual, traziam detalhes suficientes para o bom andamento da disciplina. Mesmo assim, os alertas tiveram um papel fundamental na organização e engajamento dos estudantes, como é possível observar ao analisar os itens a seguir.

O percentual de alunos que concordaram plenamente com as afirmações relacionadas aos alertas (mensagens) é bastante elevado, como é possível observar no gráfico de barras (figura 4.16).

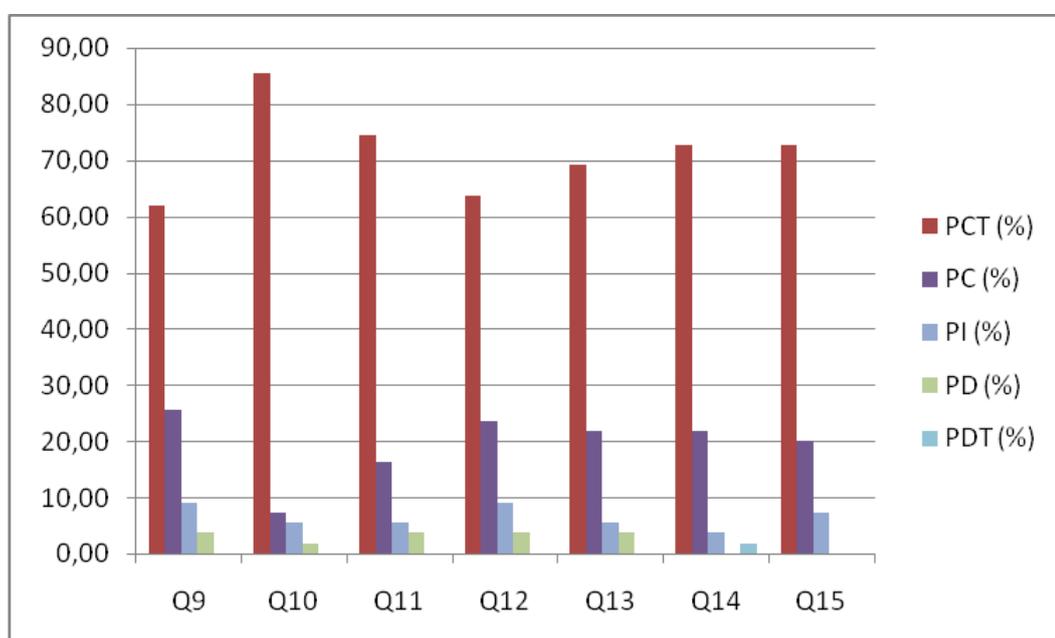


Figura 4.16 – Gráfico: questões relacionadas aos alertas

Para 87,27% dos alunos o recebimento de uma mensagem (alerta) os levava a acessar a sala de aula virtual (item 9) – 61,28% concordam totalmente e 25,45% concordam. Em 92,73% dos casos as mensagens auxiliavam na lembrança de atividades pendentes, tais como leitura de materiais, realização de atividades avaliativas e participação em fóruns, sendo que 85,45% dos respondentes concordaram totalmente com esta afirmação. Em relação a este item (número 10), apenas para 5,45% dos alunos essas mensagens eram indiferentes e 1,82% discordavam do auxílio das mensagens para chamar a atenção às pendências.

As mensagens foram importantes para a organização dos estudos e engajamento na disciplina, conforme retorno de 90,91% dos respondentes, sendo que 74,55% concordaram totalmente com a afirmativa apresentada no item 11. Em relação à questão 12, que aborda o estabelecimento de um ritmo adequado de estudos a distância (considerando-se a flexibilidade de tempo, embora sem ausência de prazos), 63,64% dos alunos concordaram totalmente que as mensagens foram importantes para tal e 23,64% também concordaram, totalizando 87,27% de avaliação positiva. 9,09% dos respondentes acreditam que as mensagens não tenham interferido no ritmo de estudos que estabeleceram e 3,64% discordam que as mensagens tenham auxiliado, não sendo sensibilizados por elas.

Em 90,91% dos casos (questão 13), os alunos concordaram que as mensagens recebidas permitiam estabelecer um vínculo mais próximo com o professor, ou seja, minimizar a distância transacional, reduzindo a sensação de isolamento, percebendo que o professor acompanhava e estava atento às suas necessidades. Para 5,45%, o estabelecimento de vínculo mais próximo com o professor por meio das mensagens foi indiferente e 3,64% discordaram que as mensagens pudessem evidenciar uma preocupação maior do professor em relação às suas necessidades.

94,55% dos alunos julgaram que a quantidade e a frequência de mensagens foram adequadas ao momento da disciplina e à sua necessidade (questão 14). 92,73% acreditam que as mensagens auxiliaram na compreensão das relações entre os conteúdos e as atividades, bem como da importância de realizá-las para a aprendizagem dos assuntos trabalhados, aspecto que foi indiferente para 7,27% dos alunos e do qual ninguém discordou.

O bloco de questões de 16 a 20 relacionava-se à dimensão *autonomia* do aluno. Sobre o interesse em realizar as leituras dos conteúdos liberados e materiais complementares sugeridos (questão 16), 58,18% concordaram plenamente que houve interesse pessoal, seguido de 23,64% que concordaram. 7,27% foram indiferentes a este item, o que indica que liam o necessário, enquanto 10,91% discordavam que havia um interesse pessoal genuíno para a realização das leituras.

Analisando os resultados já somados das respostas concordo totalmente e concordo das questões 17 a 20 (com pequena vantagem para o primeiro critério, comparativamente aos demais itens da pesquisa), 90,91% dos estudantes afirmam que houve organização da sua parte para realizar os trabalhos solicitados e cumprir os prazos, 92,73% afirmam que mantiveram uma postura ativa de aprendizagem, acessando regularmente o ambiente NetAula, 90,91% dizem ter participado dos fóruns com vista ao aprofundamento de temas e esclarecimento de dúvidas e 81,82% colocam que leram as diversas postagens dos fóruns para aprender e contribuir com os colegas.

As respostas obtidas às questões ligadas à autonomia do aluno (Q16 a Q20) são reflexo da boa estrutura da disciplina, comprovada pelos índices das questões de 1 a 4, e pela forma de diálogo estabelecida, avaliadas das questões de 5 a 15, com enfoque especial as mensagens (alertas), que visavam a apoiar à construção da autonomia do aluno. Os índices de acesso a fóruns, por exemplo, foram superiores a partir do terceiro mês, comparativamente ao mesmo período das edições de IC anteriores, em função dos alertas para acesso e participação nas discussões.

Com boa resposta dos alunos ao trabalho desenvolvido na disciplina e ao acompanhamento realizado pelos professores – o que refletiu do aumento dos índices de aprovação e redução da evasão (seção 4.7) – o trabalho desenvolvido na disciplina repercutiu também em uma boa auto-avaliação dos alunos sobre suas próprias atitudes de aprendizagem.

Quando indagados sobre se cursariam outras disciplinas ou cursos na modalidade a distância, 74,55% concordaram plenamente, seguidos de 14,55% que também concordaram (total de 89,09%), demonstrando uma avaliação positiva sobre o estudo nessa modalidade. 9,09% dos casos poderiam vir a cursar outra disciplina a distância, se fosse necessário. E 1,82% respondeu que não cursaria outra disciplina ou curso a distância, demonstrando a inadaptação ao trabalho realizado (figura 4.17).

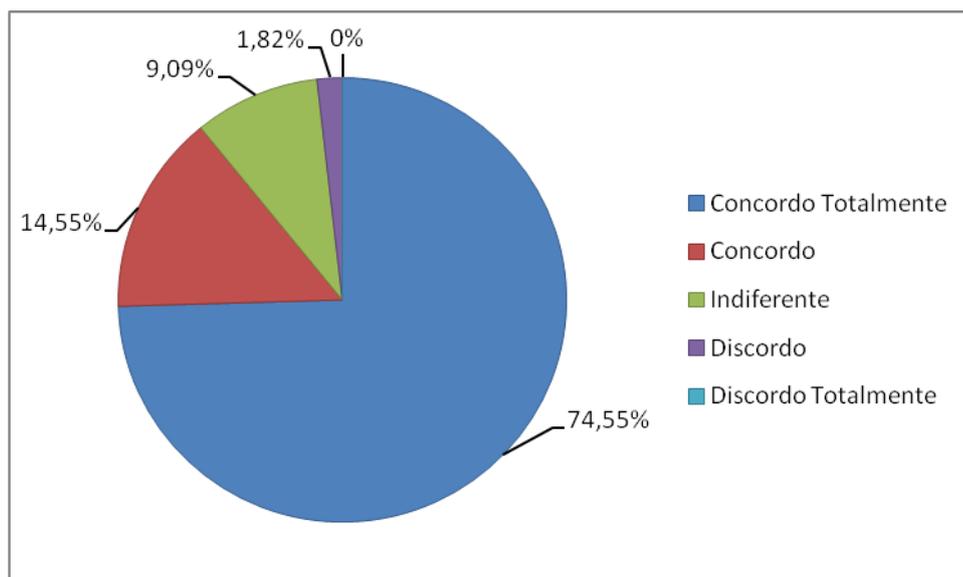


Figura 4.17 – Gráfico: Alunos que cursariam outras disciplinas ou cursos a distância (Q21)

5. Considerações Finais

Cresce, a cada dia, a quantidade de alunos matriculados em cursos realizados na modalidade a distância. Embora a modalidade esteja se tornando cada vez mais comum entre os estudantes, permanecem altos os índices de evasão, conforme as pesquisas analisadas na tese, com destaque para o trabalho de Tresman (2002), com amplo levantamento realizado na *Open University* (Inglaterra). Um bom acompanhamento docente, identificando e atuando junto a alunos em risco, pode auxiliar na retenção destes sujeitos em seus programas de formação.

Como Belloni (2008) afirma, muitos alunos que ingressam na modalidade a distância, não construíram autonomia para gerir seus processos de aprendizagem. Em cursos formais a distância, elaborados de forma intencional, com estruturas definidas e formas de diálogo com o professor, cabe ao aluno, em atitude ativa, organizar seus horários de estudo, de acordo com os prazos estabelecidos, para estudar os temas necessários, realizar as atividades propostas e interagir com professores e colegas. Neste contexto, com tantos apelos diários, é comum que os alunos priorizem outras atividades em detrimento dos estudos, que possuem horários mais flexíveis.

Assim, muitas vezes, as tarefas se acumulam e já não é mais possível dar conta do que ficou para trás. Moore & Kearsley (2007) ressaltam a importância de um papel pró-ativo do professor, acompanhando indicadores, observando alunos em risco de abandono e oferecendo ajuda. Estes autores sugerem, inclusive, que é preciso desenvolver “métodos” para identificar dificuldades dos alunos precocemente, gerando “alertas vermelhos” para o pessoal de apoio da instituição. Envolver os alunos, mantendo uma comunicação mais voltada às suas necessidades, contribuiu para que os alunos persistam em seus estudos.

O trabalho desenvolvido na tese buscou definir uma arquitetura para sistemas de alertas em AVA, com alertas pré-definidos e outros gerados a partir da mineração de dados. Esses alertas devem dar suporte à atuação pedagógica do professor no acompanhamento das situações de aprendizagem, com o monitoramento de perfis de alunos e de eventos no ambiente, notificando o professor sobre grupos de alunos

que compartilham necessidades específicas. Cada alerta, portanto, notifica o professor sobre a situação de um grupo de alunos, formado por um ou mais sujeitos, para que o professor possa dirigir comunicações ou atividades específicas. Os grupos identificados são dinâmicos, sendo gerados no momento oportuno para conscientizar o professor.

A arquitetura proposta diferencia-se de outros trabalhos similares, sobretudo por combinar uma proposta híbrida como fonte dos alertas gerados ao professor: indicadores fixos com parâmetros configuráveis mais a aplicação de regras geradas com base em mineração de dados. Sistemas de alertas como propostos por Musa et al. (2000) e Kosba (2004) ou de acompanhamento como Reategui et al. (2008) e Campana et al. (2008) apontam para o monitoramento de eventos fixos do ambiente, gerando interação individual com o aluno identificado. O professor pode ser avisado, em alguns casos, sobre o contato gerado pelo sistema com os alunos, mas recebe uma cópia de cada notificação e não o agrupamento de alunos com características comuns.

Ao invés de gerar alertas automáticos para os alunos, o sistema de alertas proposto na tese visa a conscientizar o professor para que este realize uma mediação efetiva, baseada na necessidade do grupo de alunos em evidência por um determinado alerta, bem como que esta comunicação seja contextualizada com o que ocorre na disciplina. Receber um contato do professor, que demonstra interesse e acompanha seus alunos, tende, pelo resultado da pesquisa com os alunos, a criar um ambiente comunicativo mais próximo e fluente, favorecendo o engajamento efetivo dos alunos em seus estudos.

A mineração de dados, por sua vez, traz um diferencial importante para a geração de alertas: mais do que eventos fixos, pré-programados, os alertas podem basear-se na descoberta de relações, de padrões de comportamento e características de alunos em edições anteriores do curso, analisando a combinação de vários aspectos para predizer seus resultados. Antecipar prognósticos pode ser útil para seguir apoiando alunos que estão engajados em seus estudos e, como foco principal do estudo realizado, para intervir junto a alunos em risco de reprovação ou evasão, contribuindo ativamente para a retenção e êxito dos alunos.

No trabalho realizado na tese foram considerados dados de 1564 alunos de edições anteriores do curso ministrado para a extração das regras por mineração, e mais 230 alunos acompanhados para a geração dos alertas, com cerca de 230 atributos analisados de cada aluno. A realização do experimento durante a tese permitiu comprovar que as intervenções realizadas pelo professor, a partir dos alertas, direcionadas a grupos que compartilhavam necessidades específicas, contribuiu para a melhoria dos índices de aprovação e para redução dos índices de evasão dos alunos na disciplina acompanhada.

Assim, apresentam-se como contribuições da tese:

- A definição de uma arquitetura híbrida para sistemas de alertas integrados em AVA, a partir de indicadores do ambiente configuráveis pelo professor e de técnicas automatizadas de mineração de dados.
- A utilização de técnicas de MD sobre dados educacionais oriundos de AVA, aliados a dados demográficos definidos, para a descoberta de características e comportamentos dos alunos que podem servir de base para a mediação do professor.
- A integração da mineração de dados históricos com a geração dos alertas em turmas correntes, para notificar o professor sobre grupos de alunos com tendência à reprovação e à evasão em EAD, aplicando e avaliando os resultados.
- A validação das hipóteses que relacionavam a intervenção do professor, realizada a partir dos alertas, e a contribuição para a redução da evasão e para a ampliação da aprovação de alunos na disciplina em que o experimento foi realizado.
- A realização de pesquisa com os alunos, para avaliar a importância de uma maior proximidade comunicativa a partir das mensagens enviadas pelo professor, com frequência e quantidades adequadas às necessidades de cada aluno.

Durante a pesquisa, observou-se que o sistema de alertas contribuiu de maneira importante na gestão do ambiente virtual. Muitas das práticas docentes de

engajamento dos alunos em seus estudos, bem como de resgate dos estudantes em processo de evasão, foram realizadas a partir do encaminhamento de mensagens aos alunos que, identificados por seus comportamentos, demandavam uma atenção maior, de forma que voltassem a participar ativamente das atividades da disciplina. Os alertas (automatizados pelo sistema), disponibilizados em AVA, constituíram-se em um diferencial para a tomada de decisão e intervenção junto aos alunos, visando a uma ação pedagógica de qualidade. Professores iniciantes também podem se beneficiar dos alertas, chamando a atenção a aspectos que demandam intervenções.

Como trabalhos futuros, apresentam-se alguns desafios:

- Realizar um estudo para identificar os dados mais relevantes para serem gerados e armazenados com o intuito de alimentar uma boa MD educacional.
- Acrescentar dados qualitativos à mineração, empregando técnicas de *Text Mining* para observar a pertinência das participações dos alunos aos temas de discussões estabelecidas em AVA.
- Estabelecer parâmetros para a priorização na apresentação/disposição dos alertas.
- Investigar quais dos alertas enviados ao professor, para sua intervenção, podem ser automaticamente enviados aos alunos, com *conscientização* do professor pelo sistema. Nesse caso, verificar se a mensagem deve ser previamente criada pelo professor, favorecendo a pessoalidade e a adequação da mesma aos possíveis alunos da(s) turma(s) que orienta ou se é possível automatizar o envio de mensagens baseado no comportamento (individual ou em grupo) histórico de sucesso já comprovado pelo sistema em edições anteriores. O sistema poderá utilizar uma marcação para o vocativo (nome do aluno), além de personalização, como em função de gênero (masculino/feminino).
- Realizar um estudo que priorize os alunos com bom potencial de aprendizagem, observando se seus desempenhos são melhorados a partir

da identificação desse potencial e da realização de intervenções específicas.

As áreas de mineração de dados educacionais e sistemas de alertas, integradas, têm muito a contribuir para a área da Educação e, em especial, para a EAD. Os desafios da EAD, para democratizar e qualificar os processos de ensino e aprendizagem, são grandes. Ferramentas que integrem conhecimento dessas áreas, potencialmente, contribuem para identificar padrões em grupos normalmente bastante heterogêneos e para apoiar professores, experientes ou não, a gerir seus espaços de aula.

6. Bibliografia

- ABED. **Associação Brasileira de Educação a Distância. Informações do Anuário Brasileiro Estatístico de Educação Aberta e a Distância (Abraead/2007).** Disponível em: http://www2.abed.org.br/noticia.asp?Noticia_ID=275. Acessado em: 30/11/2008.
- ABED. **Associação Brasileira de Educação a Distância. Informações do Anuário Brasileiro Estatístico de Educação Aberta e a Distância (Abraead/2008).** Disponível em: http://www2.abed.org.br/visualizaDocumento.asp?Documento_ID=261. Acessado em: 20/04/2009.
- AGRAWAL, R.; IMIELINSKI, T.; SRIKANT R. **Mining Association Rules between Sets of Items in Large Databases.** In: Proc. of the ACM SIGMOD Intl. Conf. on Management of Data, US, Washington, 1993, 207–216.
- ALMEIDA, Ivana Carneiro; ILDETE, Maria. **Educação a distância: Um estudo dos motivos de desistência de um curso a distância via Internet.** 14^o. Congresso Internacional da ABED de Educação a Distância, Santos/SP, 2008. Disponível em: <http://www.abed.org.br/congresso2008/tc/54200862040PM.pdf>. Acessado em: 25/07/2009.
- ALMEIDA, Manoel Roberto Aguirre de. **Sistema Híbrido Neuro-Fuzzy-Genético para Mineração Automática de Dados.** Dissertação de Mestrado. Rio de Janeiro: Departamento de Engenharia Elétrica/PUC-Rio, 2004.
- ALVES, Lynn; NOVA, Cristiane. **Educação a distância: Limites e Possibilidades.** Capítulo do Livro: ALVES, Lynn; NOVA, Cristiane. **Educação a Distância: Uma nova concepção de aprendizado e interatividade.** São Paulo: Futura, 2003.
- AMORIM, Maurício J.V.; BARONE, Dante; MANSUR, Anré Uebe. **Técnicas de Aprendizado de Máquina Aplicadas na Previsão de Evasão Acadêmica.** In: Anais do XIX Simpósio Brasileiro de Informática na Educação. SBC. Fortaleza, 2008.
- ARAÚJO, J. P. **O que os aprendizes esperam dos professores na Educação a Distância On-line?** 2001. Disponível em: www.comunicar.pro.br/artigos/expect.htm. Acessado em: 24/08/2009.
- BELISÁRIO, Aluizio. **O material didático na educação a distância e a constituição de propostas interativas.** Capítulo do Livro: SILVA, Marco. **Educação Online.** São Paulo: Loyola, 2003.
- BELLONI, Maria Luiza. **Educação a distância.** 5a. ed. Campinas: Autores Associados, 2008.
- BERGE, Zane L. **The Role of the Online Instructor/Facilitator.** 1995. Disponível em: http://www.cordonline.net/mntutorial2/module_2/Reading%20202-1%20instructor%20role.pdf. Acessado em: 20/05/2009.

- BERRY, Michael J. A.; LINOFF, Gordon. **Data Mining Techniques: For Marketing, Sales, and Customer Support**. US: John Wiley & Sons, 1997.
- BIAZUS, Cleber Augusto. **Sistema de fatores que influenciam o aluno a evadir-se dos cursos de graduação na UFSM e na UFSC**: um estudo no curso de Ciências Contábeis. Florianópolis, 2004. Tese. (Doutorado em Engenharia de Produção). Universidade Federal de Santa Catarina.
- BIANCO, Nelía R. Del. **Aprendizagem por rádio**. Capítulo do Livro: LITTO, Frederic Michael; FORMIGA, Manuel Marcos Maciel. **Educação a Distância: o estado da arte**. São Paulo: Pearson Educacion do Brasil, 2009.
- BORGES, Marcos Roberto da Silva; CAVALCANTI, Maria Cláudia Reis; CAMPOS, Maria Luiza Machado. **Suporte por Computador ao Trabalho Cooperativo**. In: Jornada de Atualização em Informática, 14., 1995, Canela. Anais... Porto Alegre: Instituto de Informática da UFRGS. 45p.
- BOUCHARD, Paul. **Autonomia e Distância Transacional na Formação a Distância**. Capítulo do Livro: ALAVA, Séraphin. **Ciberespaço e Formações Abertas: rumo a novas práticas educacionais?** Porto Alegre: Artmed, 2002.
- BRACHMAN, Ronald J. ANAND, Tej. **The Process of Knowledge Discovery in Databases**. In: FAYYAD, Usama M. et al. **Advances in Knowledge Discovery and Data Mining**. Menlo Park: AAAI Press, 1996. P. 37-57.
- CAMPANA, Vitor Faiçal et al. **Agentes para Apoiar o Acompanhamento das Atividades em Ambientes Virtuais de Aprendizagem**. In: Anais do XIX Simpósio Brasileiro de Informática na Educação. SBC. Fortaleza, 2008.
- CARAVOLI, Juliano Trevisan; COELLO, Juan Manuel Adán. **Alerts Forum: um Sistema Emissor de Alertas a partir da Avaliação de Mensagens Categorizadas em Fóruns de Discussão**. VII Congresso Iberoamericano de Informática Educativa. 2004. Disponível em: <http://www.niee.ufrgs.br/ribie2004/Trabalhos/Comunicacoes/com39-47.pdf>. Acessado em: 04/04/2008.
- CARMO, H. **Ensino Superior a Distância**. Lisboa: Universidade Aberta, 1997.
- CASTRO, Nivalde et al. **O Estudo a Distância com Apoio da Internet**. ABED, 10/09/2002. Disponível em: <http://www.abed.org.br/publique/cgi/cgilua.exe/sys/start.htm?inford=137&sid=116&tpl=printerview>. Acessado em: 04/10/2008.
- CARR, Sarah. **As Distance Education Comes of Age, the Challenge Is Keeping the Students**. Chronicle of Higher Education, Feb. 11, 2000. Disponível em: <http://chronicle.com/article/As-Distance-Education-Comes-of/14334>. Acessado em: 20/07/2009.
- CIOS, Krzysztof J.; PEDRYCZ, Witold; SWINIARSKI, Roman. **Data mining methods for knowledge discovery**. USA: Kluwer Academic Publisher, 1998.

- COSTA, José da. **Modelos de Educação Superior a Distância e Implementação da Universidade Aberta do Brasil**. Revista Brasileira de Informática na Educação – V. 15 N. 2 - Maio a Agosto de 2007.
- CUNICO, Luiz Homero Bastos. **Técnicas em Data Mining Aplicadas na Predição de Satisfação de Funcionários de uma Rede de Lojas do Comércio Varejista**. Dissertação de Mestrado. Curitiba: Pós-Graduação em Métodos Numéricos em Engenharia – Programação Matemática, Setores de Tecnologia e Ciências Exatas, Universidade Federal do Paraná, 2004.
- DEOGUN, Jitender S. et al. **Data Mining: Trend in Research and Development**. In: LIN, T. Y.; CERCONE, N. Rough Sets and Data Mining: Analysis for Imprecise Data. Springer, 1997.
- DIAS, Maxwell Macedo et al. **Aplicação de Técnicas de Mineração de Dados no Processo de Aprendizagem na Educação a Distância**. In: Anais do XIX Simpósio Brasileiro de Informática na Educação. SBC. Fortaleza, 2008.
- DIAZ, David P. **Online Drop Rates Revisited**. The Technology Source. May/June, 2002. Disponível em: http://technologysource.org/article/online_drop_rates_revisited/ Acessado em: 20/07/2009.
- ELLIS, Clarence A.; GIBBS, Simon J.; REIN, Gail L. **Groupware – Some Issues and Experiences**. Communications of the ACM, New York, v. 34, n. 1, p. 38-58, Jan. 1991.
- FAGUNDES, Léa da Cruz et al. **Projetos de Aprendizagem - Uma experiência mediada por ambientes telemáticos**. Revista Brasileira de Informática na Educação – V. 14 N. 1 - Janeiro a Abril de 2006.
- FAVERO, Rute Vera Maria; FRANCO, Sérgio Roberto Kieling. **Um estudo sobre a permanência e a evasão na Educação a Distância**. Revista Novas Tecnologias na Educação – V. 4 N.º 2 – Dezembro de 2006.
- FAYYAD, U. M.; PIATESKY-SHAPIRO, G.; SMYTH, P. **From Data Mining to Knowledge Discovery: An Overview**. Advances in Knowledge Discovery and Data Mining, AAAI Press, 1996.
- FRANÇA, George; SILVA, Liliam Maria da; SANTOS, Luciana Aparecida; VASCONCELOS, Paulo A. C. **Design Instrucional: Metodologias, Comunicação, Afetividade e Aprendizagem**. Revista Inter Science Place – V. 1 N. 2 – Novembro de 2008.
- FREIRE, Paulo. **Pedagogia do Oprimido**. 40ª. Ed. Rio de Janeiro: Paz e Terra, 2005.
- FUKS, Hugo et al. **O Modelo de Colaboração 3C no Ambiente AulaNet**. Revista Informática na Educação: Teoria e Prática, Vol 7, No. 1, 2004. Porto Alegre, UFRGS, pp. 25-48.

- GARCIA, Enrique et al. **Drawbacks and solutions of applying association rule mining in learning management systems**. International Workshop on Applying Data Mining in e-Learning (ADML'07), in Second European Conference on Technology Enhanced Learning (EC-TEL07). Crete, Greece, September, 2007. Disponível em: <http://www.win.tue.nl/~mpechen/conf/adml07/>. Acessado em: 20/05/2009
- GEROSA, Marco Aurélio; FUKS, Hugo; LUCENA, Carlos José Pereira de. **Suporte à Percepção em Ambientes de Aprendizagem Colaborativa**. Revista Brasileira de Informática na Educação – V.11 N. 2 – 2003.
- GIBSON, Chere Campbell. **The distance learner's academic self-concept**. Capítulo de Livro: GIBSON, Chere Campbell. Distance learners in higher education: Institutional responses for quality outcomes. Madison, WI: Atwood Publishing, 1998.
- HAN, Jiawei; KAMBER, Micheline. **Data Mining: concepts and techniques**. 2nd. ed. San Francisco: Morgan Kaufmann Publishers / Elsevier, 2006.
- HAND, David; MANNILA, Heikki; SMYTH, Padhraic. **Principles of data mining**. USA: Massachusetts Institute of Technology, 2001.
- HILLS, Mellanie. **Intranet as Groupware**. United States: John Wiley & Sons Inc., 1997.
- INMON, W. H. **Building the Data Warehouse**. United States: John Wiley & Sons Inc., 2002.
- KAMPFF, Adriana Justin Cerveira Kampff. **Utilização de Agentes para Suporte ao Trabalho em Grupo na Internet**. Dissertação de Mestrado. Porto Alegre: PGCC/UFRGS, 1999.
- KEEGAN, Desmond. **Foundations of distance education**. 3a. Edition. London: Routledge, 1996.
- KOSBA, Essam Mahmoud Abdel Monemng. **Generating Computer-Based Advice in Web-Based Distance Educacion Environments**. Tese de Doutorado. University of Leeds. School of Computing. December, 2004.
- LAGO, Andréa Ferreira. **Aluno: on-line; Senha: comunidade: considerações sobre EAD a partir de experiência como aluna on-line**. Capítulo do Livro: ALVES, Lynn; NOVA, Cristiane. **Educação a Distância: Uma nova concepção de aprendizado e interatividade**. São Paulo: Futura, 2003.
- LÉVY, Pierre. **As tecnologias da inteligência: o futuro do pensamento na era da Informática**. 8ª. Ed. Rio de Janeiro: Ed. 34, 1999.
- LÉVY, Pierre. **Cibercultura**. 2ª. Ed. São Paulo: Ed. 34, 2000.
- LIN, T. Y; CERCONE, N. **Rough Sets and Data Mining: Analysis for Imprecise Data**. Springer, 1997.

- LITWIN, Edith. **Das Tradições à Virtualidade**. Capítulo do Livro: LITWIN, Edith. **Educação a Distância: temas para o debate de uma nova agenda educativa**. Porto Alegre: Artmed, 2001.
- LONGO, Carlos Roberto Juliano. **A EAD na pós-graduação**. Capítulo do Livro: LITTO, Frederic Michael; FORMIGA, Manuel Marcos Maciel. **Educação a Distância: o estado da arte**. São Paulo: Pearson Educacion do Brasil, 2009.
- LOPES, Claudivan Cruz; SCHIEL, Ulrich. **Uma Estratégia para Aplicar Mineração de Dados no Acompanhamento do Aprendizado na EaD**. XIII Seminário de Computação. Blumenau, SC, Setembro, 2004. Disponível em: <http://www.inf.furb.br/seminco/2004/artigos/108-vf.pdf>. Acessado em: 20/11/2008.
- MACHADO, Leticia dos Santos; BECKER, Karin. **O uso da Mineração de Dados na Web aplicado a um Ambiente de Ensino a Distância**. I Workshop de Teses e Dissertações em Banco de Dados. Gramado, 2002.
- MAGGIO, Mariana. **O Tutor em Educação a Distância**. Capítulo do Livro: LITWIN, Edith. **Educação a Distância: temas para o debate de uma nova agenda educativa**. Porto Alegre: Artmed, 2001.
- MAIA, Marta de Campos; MEIRELLES, Fernando de Souza; PELA, Silvia Krueger. **Análise dos Índices de Evasão nos Cursos Superiores a Distância do Brasil**. 11º. Congresso Internacional de Educação a Distância – ABEC. Salvador. Abril de 2004. Disponível em: <http://www.abed.org.br/congresso2004/por/pdf/073-TC-C2.pdf>. Acessado em: 20/09/2008.
- MANNILA, H. **Data mining: Machine learning, statistic and databases**. In: Proceedings of the 8th International Conference on Scientific and Statistical Database Management, 1996, pp. 1–8.
- MARAVALLE, M.; SIMEONE, B.; NALDINI, R. **Clustering on trees**. Computational Statistics & Data Analysis, v. 24, n. 2, 1997, p. 217-234.
- MARTINS, Francisco José; FERRARI, Débora Nice; GEYER, Cláudio F. R. **jXChat – Um sistema de Comunicação Eletrônica Inteligente para apoio a Educação a Distância**. In: Anais do XIV Simpósio Brasileiro de Informática na Educação. SBC. NCE – IM/UFRJ, 2003.
- MEC/DECRETO. **N.º 2.494, de 10 de fevereiro de 1998**. Disponível em: <http://portal.mec.gov.br/seed/arquivos/pdf/tvescola/leis/D2494.pdf>. Acessado em: 24/08/2008.
- MEC - MINISTÉRIO DE EDUCAÇÃO DO BRASIL. **Relatório Final da Comissão Assessora para Educação Superior a Distância**. Agosto, 2002. Disponível em: http://www.nead.ufma.br/arquivos/RELATORIO_FINAL_DA_COMISSAO_DE_EAD_MEC.pdf. Acessado em: 30/08/2008.

- MEC/PORTARIA. **Nº 4.059, de 10 de dezembro de 2004.** Disponível em: http://portal.mec.gov.br/sesu/arquivos/pdf/nova/acs_portaria4059.pdf. Acessado em: 24/08/2008.
- MEC/SETEC/SEED. **Referenciais de Qualidade para Educação Superior a Distância.** Agosto, 2007. Disponível em: <http://portal.mec.gov.br/seed/arquivos/pdf/legislacao/refead1.pdf>. Acessado em: 01/07/2009.
- MEC/SETEC/SEED. **Referenciais para Elaboração de Material Didático para EAD no Ensino Profissional e Tecnológico.** Disponível em: http://www.etecbrasil.mec.gov.br/gCon/recursos/upload/file/ref_materialdidatico.pdf. Acessado em: 01/08/2008.
- MERCER, Neil; ESTEPA, Francisco Gonzáles. **A Educação a Distância, o Conhecimento Compartilhado e a Criação de uma Comunidade de Discurso Internacional.** Capítulo do Livro: LITWIN, Edith. **Educação a Distância: temas para o debate de uma nova agenda educativa.** Porto Alegre: Artmed, 2001.
- MESQUITA, Cláudia S. F.; SALGADO, Ana Carolina; IOCHPE, Cirano. **O Gerenciamento da Comunicação em um Ambiente de Projeto e Cooperação.** In: Conferencia Latinoamericana de Informática (CLEI), 20., 1994. Proceedings... [S.1.:s.n.], 1994. p. 925-933.
- MIERSWA, Ingo et al. **YALE: Rapid Prototyping for Complex Data Mining Tasks.** In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-06), 2006. p. 935-940.
- MOORE, Michael G. **Theory of transactional distance.** Capítulo do Livro: KEEGAN, Desmond. **Theoretical principles of distance education.** Oxon: Routledge, 1993.
- MOORE, Michael G.; KEARSLEY, Greg. **Educação a distância: uma visão integrada.** São Paulo: Thomson Learning, 2007.
- MORAN, José Manuel. **O vídeo na sala de aula.** Revista Comunicação & Educação. São Paulo, ECA-Ed. Moderna, [2]: 27 a 35, jan./abr. de 1995.
- MORAN, José Manuel. **Contribuições para uma pedagogia da educação online.** Capítulo do Livro: SILVA, Marco. **Educação online.** São Paulo: Edições Loyola, 2003.
- MORAN, José Manuel. **A educação que desejamos: Novos desafios e como chegar lá.** Revista Comunicação & Educação. Campinas: Papirus, 2007.
- MUSA, Daniela Leal; SISSON, Gustavo de Abreu; OLIVEIRA, José Palazzo M. de. **Experiências no Desenvolvimento e Integração de um Gerenciador de Alertas para Ambientes de Ensino a Distância na Internet.** Disponível em: [http://www.inf.ufrgs.br/~palazzo/OAI/04%20Ideas%20\(Dani,%20Gustavo\).pdf](http://www.inf.ufrgs.br/~palazzo/OAI/04%20Ideas%20(Dani,%20Gustavo).pdf). Acessado em: 20/03/2000.

- MUSA, Daniela Leal et al. **Agente para auxílio a avaliação de aprendizagem em ambientes de ensino na Web**. In: Anais do XI Simpósio Brasileiro de Informática na Educação. SBC, 2001.
- NASSAR, Silvia M.; WRONSCKI, Vilson R.; OHIRA, Masanao et al. **SEstatNet - Sistema Especialista para o Ensino de Estatística na Web**. Disponível em: <http://www.sestat.net>. Acessado em: 27/10/2009.
- NEVADO, Rosane Aragón de; CARVALHO, Marie Jane Soares de; MENEZES, Crediné Silva de. **Educação a distância mediada pela internet: uma abordagem interdisciplinar na formação de professores em serviço**. Capítulo do Livro: NEVADO, Rosane Aragón de; CARVALHO, Marie Jane Soares de; MENEZES, Crediné Silva de. **Aprendizagem em rede e educação a distância: estudos e recursos para formação de professores**. Porto Alegre: Ricardo Lenz, 2007.
- NISKIER, Arnaldo. **Os aspectos culturais e a EAD**. Capítulo do Livro: LITTO, Frederic Michael; FORMIGA, Manuel Marcos Maciel. **Educação a Distância: o estado da arte**. São Paulo: Pearson Educacion do Brasil, 2009.
- OKADA, Alexandra Lilavati Pereira. **Desafios para EAD: Como fazer emergir a colaboração e a cooperação em ambientes virtuais de aprendizagem?** Capítulo do Livro: SILVA, Marco. **Educação Online**. São Paulo: Loyola, 2003.
- OTSUKA, Joice Lee. **Modelo de suporte à avaliação formativa baseado em sistemas multiagentes para ambientes de EaD**. Tese de Doutorado. Instituto de Informática, Unicamp. Campinas, SP, 2006.
- OLIVEIRA, Aracele Garcia de; GARCIA, Denise Ferreira. **Mineração da Base de Dados de um Processo Seletivo Universitário**. Infocomp – Revista de Ciência da Computação, v. 3. n. 2, p. 38-43, 2004.
- PACHECO, Andressa S. V. et al. **Fatores Dificultadores à Permanência dos Alunos no Curso de Graduação em Administração a Distância da Universidade Federal de Santa Catarina**. RENOTE. Revista Novas Tecnologias na Educação, v. 6, p. 1, 2008.
- PALHARES, Roberto. **Aprendizagem por correspondência**. Capítulo do Livro: LITTO, Frederic Michael; FORMIGA, Manuel Marcos Maciel. **Educação a Distância: o estado da arte**. São Paulo: Pearson Educacion do Brasil, 2009.
- PALLOFF, Rena M. PRATT, Keith. **O aluno virtual: um guia para trabalhar com estudantes on-line**. Porto Alegre: Artmed, 2004.
- PASSERINO, Liliana M.; GLUZ, João Carlos; VICARI, Rosa Maria. **MEDIATEC – Mediação Tecnológica em Espaços Virtuais para Apoio ao Professor Online**. RENOTE. Revista Novas Tecnologias na Educação, v. 5, p. 1, 2007.

- PAZ, Carolina Rodrigues et al. **Monitoria online em educação a distância: o caso LED/UFSC**. Capítulo do Livro: SILVA, Marco. **Educação Online**. São Paulo: Loyola, 2003.
- PETERS, Otto. **Didática do Ensino a Distância**. São Leopoldo: Unisinos, 2001.
- PINHEIRO, Manuele Kirsch; LIMA, José Valdeni de; BORGES, Marcos R. S. **A framework for awareness support in groupware systems**. In: Computer in Industry, v. 52, 2003, p. 47-57.
- QUINLAN, Ross. **C4.5: Programs For Machine Learning**. Morgan Kaufmann Publishers, 1992.
- RAMAL, Andrea Cecilia. **Educação a distância: Entre Mitos e Desafios**. Capítulo do Livro: ALVES, Lynn; NOVA, Cristiane. **Educação a Distância: Uma nova concepção de aprendizado e interatividade**. São Paulo: Futura, 2003.
- RAPIDMINER. **RapidMiner-I: Report the Future**. Disponível em: <http://rapid-i.com/>. Acessado em: 30/09/2009.
- REZENDE, Solange Oliveira. **Mineração de Dados**. In: Mini-Cursos do V Encontro Nacional de Inteligência Artificial. SBC. São Leopoldo, 2005.
- REATEGUI, Eliseo Berni; RIBEIRO, Alexandre; BOFF, Elisa. **Um Sistema Multiagente para Controle de um Assistente Pessoal Aplicado a um Ambiente Virtual de Aprendizagem**. Revista Novas Tecnologias na Educação – V. 6 N.º 1 – Julho de 2008.
- ROMERO, Cristóbal; VENTURA, Sebastián; ESPEJO, Pedro G.; HERVÁS, César. **Data Mining Algorithms to Classify Students**. First International Conference on Educational Data Mining. Canadá. Junho, 2008. Disponível em: http://www.educationaldatamining.org/EDM2008/uploads/proc/1_Romero_3.pdf. Acessado em: 10/10/2008.
- SILVA, Marcelino P. S. **Mineração de Dados: Conceitos, Aplicações e Experimentação com Weka**. ERI RJ/ES IV. Vitória – ES / Rio das Ostras – RJ. Novembro, 2004. Disponível em: <http://www.sbc.org.br/bibliotecadigital/download.php?paper=35>. Acessado em: 30/06/2009.
- SILVA, Marco. **Sala de Aula Interativa**. Rio de Janeiro: Quartet, 2ª Ed., 2001.
- SOLETIC, Angeles. **A Produção de Materiais Escritos nos Programas de Educação a Distância: Problemas e Desafios**. Capítulo do Livro: LITWIN, Edith. **Educação a Distância: temas para o debate de uma nova agenda educativa**. Porto Alegre: Artmed, 2001.
- SOURCEFORGE. **Source Forge – Find and Develop Open Source Software**. Disponível em: <http://sourceforge.net/>. Acessado em: 30/09/2009.
- TINTO, Vincent. **Taking Student Retention Seriously: Rethinking the First Year of College**. NACADA Journal, 2000. Disponível em: <http://www.sdcity.edu/support/SCS/DrTinto/TSRSfirstyear.pdf>. Acessado em: 30/03/2009.

- TORRES, Patrícia Lupion. **Pioneirismo em Educação a Distância**. Natal: CEFET-RN, 2003.
- TRESMAN, S. **Towards a Strategy for Improved Studenty Retention in Programmes of Open, Distance Education: A Case Study from the Open University**. The International Review of Research in Open and Distance Learning, Vol 3, No 1 - April, 2002.
- ULBRA. **Sobre o Ensino a Distância**. Disponível em: <http://www.ulbra.br/sobreulbra/>. Acessado em: 01/08/2009.
- ULBRA/EAD. **Sobre o Ensino a Distância**. Disponível em: http://www.ulbra.br/ead/sobre_ead.htm. Acessado em: 01/08/2009.
- WITTEN, I.H.; FRANK, E. **Data Mining: Practical Machine Learning Tools and Techniques**. 2nd. ed. San Francisco: Morgan Kaufmann Publishers / Elsevier, 2005.
- WIZRULE. **WizRule**. Disponível em: <http://www.wizsoft.com/default.asp?Win=8&winsub=50>. Acessado em: 30/09/2009.
- ZHANG, Shichao; ZHANG, Chengqi; YANG, Qiang. **Data Preparation for Data Mining**. In: Applied Artificial Intelligence, 17, p. 375–381, 2003.
- ZHOU, Zhi-Hua. **Three perspectives of data mining**. Artificial Intelligence, Volume 143, Issue 1, p. 139-146. January, 2003.

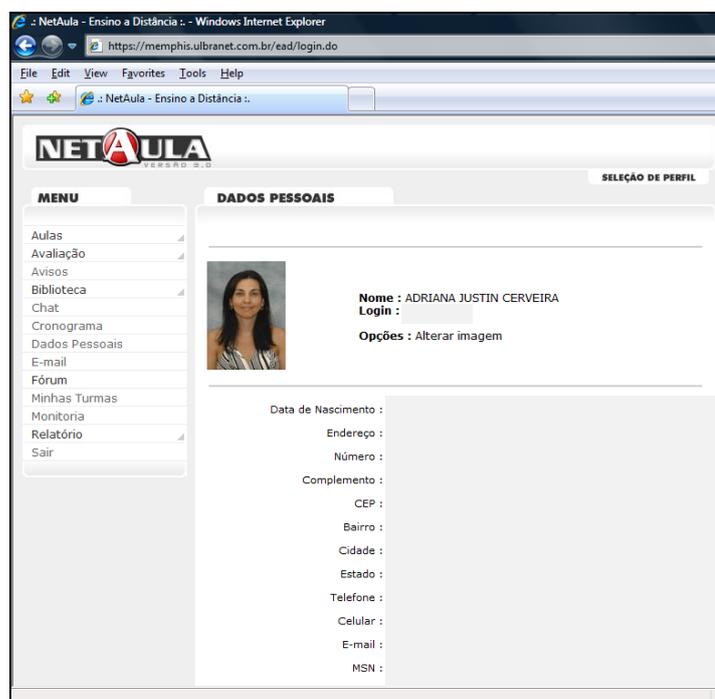
ANEXO 1

Técnicas centradas no aluno - Tabela adaptada a partir de PALLOFF & PRATT, 2004, p. 34/35

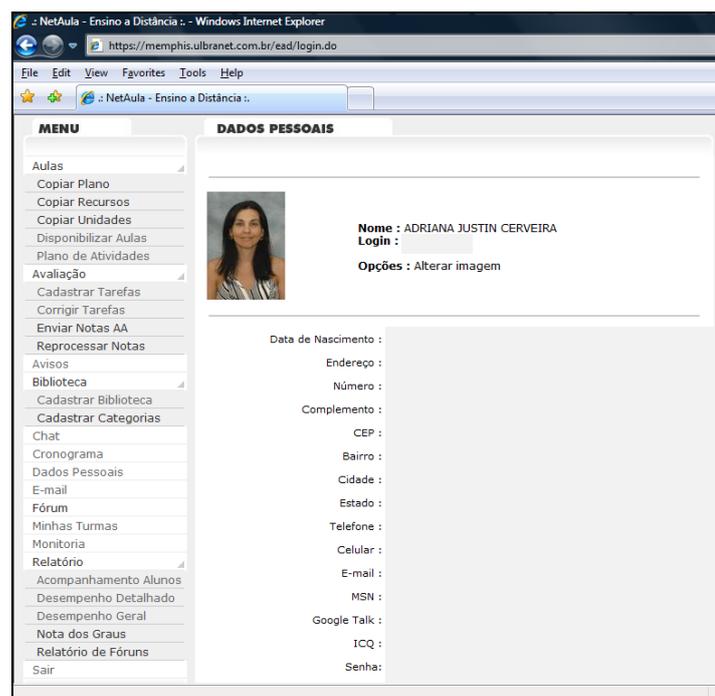
Características dos alunos <i>on-line</i>	Técnicas centradas no aluno
Acesso e habilidade	<ul style="list-style-type: none"> • Use somente tecnologia que sirva aos objetivos de aprendizagem. • Mantenha a tecnologia em um nível simples, a fim de que seja transparente para o aluno. • Certifique-se de que os alunos têm as habilidades necessárias para usar a tecnologia para o curso (...) pela realização de uma orientação tecnológica.
Abertura	<ul style="list-style-type: none"> • Comece o curso com apresentações, envio de biografia ou perfis. • Use exercícios que “quebrem o gelo” na aula para fazer com que os alunos se conheçam. • Use atividades de aprendizagem que levem em consideração a experiência e a resolução de problemas. • Crie uma área social ou “lounge” para os alunos.
Comunicação	<ul style="list-style-type: none"> • Envie suas diretrizes para comunicação, incluindo a <i>netiqueta</i>. • Exemplifique como realizar uma boa comunicação. • Explique o que constitui uma mensagem substancial para discussão. • Estimule a participação na criação das diretrizes do curso. • Acompanhe os alunos que não participam ou cujo nível de participação muda.
Comprometimento	<ul style="list-style-type: none"> • Envie uma mensagem com suas expectativas em relação à utilização do tempo. • Desenvolva e envie expectativas claras para a realização de trabalhos, prazos de entrega e meios pelos quais a avaliação será elaborada. • Seja claro quanto aos requisitos para publicação. Considere criar uma agenda de publicação para os alunos. • Apóie o desenvolvimento das boas habilidades de gerenciamento de tempo.
Colaboração	<ul style="list-style-type: none"> • Inclua estudos de caso; trabalhos em pequenos grupos, atividades “quebra-cabeça”, simulações e facilitação em rodízio para estimular o pensamento crítico. • Faça com que os alunos enviem seus trabalhos para o <i>site</i> do curso com a expectativa de que eles darão um <i>feedback</i> crítico a seus colegas. • Faça perguntas abertas para estimular a discussão e incentive os alunos a fazer o mesmo.
Reflexão	<ul style="list-style-type: none"> • Inclua uma área de reflexão no curso e estimule seu uso. • Faça perguntas abertas nas discussões e que estimulem a reflexão sobre o material utilizado, bem como sua integração.
Flexibilidade	<ul style="list-style-type: none"> • Varie as atividades do curso para atender a todos os estilos de aprendizagem e oferecer um interesse adicional e abordagens múltiplas aos tópicos trabalhados. • Negocie as diretrizes dos cursos com os alunos, de modo a promover seu engajamento. • Inclua a internet como uma ferramenta e um recurso de ensino e estimule os alunos a buscar referências que possam compartilhar.

ANEXO 2

Visualização do Ambiente Professor do AVA NetAula



Anexo 2 – Figura 1 – NetAula: Ferramentas do Ambiente Professor



Anexo 2 – Figura 2 – NetAula: Ferramentas Expandidas do Ambiente Professor

ANEXO 3

Visualização do Ambiente Aluno do AVA NetAula

NetAula - Ensino a Distância :: - Windows Internet Explorer

https://memphis.ulbranet.com.br/ead/login.do

File Edit View Favorites Tools Help

pdf Search PDF

NetAula - Ensino a Distância ::

NETAULA VERSÃO 3.0

PLANO DE ATIVIDADES

Procure seguir este Plano de Atividade, atentando aos prazos e acompanhando sempre as novidades. Clique nos links para ter acesso aos itens do plano ou acesse diretamente as ferramentas através dos ícones no menu barra superior. Tão logo esteja encerrada a atividade, clique na imagem à direita (X) para apontar que você entende que a concluiu (V)

Prezado(a) Aluno(a),

Este é o nosso ambiente de aula, nosso espaço de interatividade: participe!!!

Abraços,

Profa. Adriana e Prof. Laíno

Nosso Espaço

Fórum de Apresentação	(05/03/2009 - 21/04/2009)	X
Guia de Utilização da NetAula	(05/03/2009 - 31/07/2009)	X
Fórum de Dúvidas sobre a Utilização da NetAula	(05/03/2009 - 16/07/2009)	X
Plano de Ensino e Bibliografia	(05/03/2009 - 16/07/2009)	X
Cronograma de Atividades e Avaliação		X

CALENDÁRIO

Julho 2009

D	S	T	Q	Q	S	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

PARTICIPAÇÕES

- 2 postagen(s)
- 0 acesso(s)
- 2 enviado(s)
- 28 não lido(s)
- 14 acesso(s)
- 4 realizada(s)

Done

Anexo 3 – Figura 1 – NetAula: com Plano de Atividades (ao centro) e Contadores (à esquerda)

ANEXO 4

Ambiente Aluno do AVA NetAula

- Ferramentas disponíveis, no estágio atual, no AVA da instituição pesquisada. -

No Portal da instituição, mediante usuário e senha, se tem acesso à plataforma de ensino e aprendizagem utilizada. Ao acessar o ambiente, deve-se escolher a sala virtual de acesso, já que para organizar a interação há uma sala para cada módulo ou disciplina.

A partir da escolha da sala virtual, ocorre o acesso aos recursos da plataforma: informações gerais tais como avisos, cronogramas e acompanhamento de desempenho; acesso aos conteúdos e materiais complementares por meio do sumário e da biblioteca virtual; ferramentas internas de interação tais como correio eletrônico, sala de bate-papo e fórum; espaços para entrega e acompanhamento de resultados das avaliações; informações pessoais e acesso aos nomes e imagens (se disponibilizadas) dos colegas. A figura 1 apresenta a barra de ferramentas do ambiente e o quadro 1 detalha os recursos acessíveis por meio de cada ícone.



Anexo 4 – Figura 1 – NetAula: Ferramentas do Ambiente Aluno

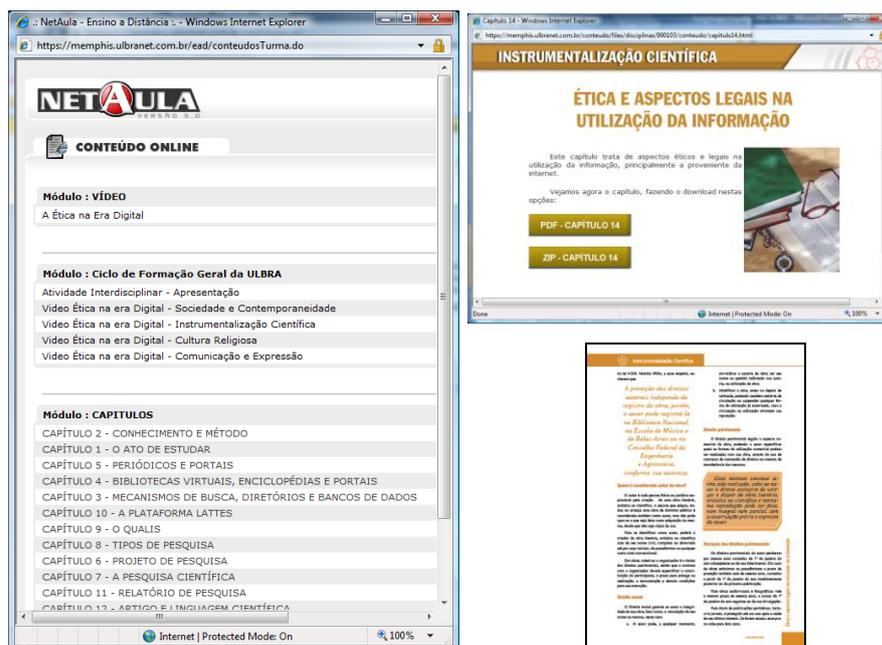
A parte central da plataforma apresenta um plano de atividades, conforme cronograma estabelecido pelo professor. É importante destacar que as datas de liberação de conteúdos e avaliações, além de avisos para o bom andamento das atividades da turma, disponíveis no ícone “informações gerais”, surgem em uma janela *pop-up* quando o aluno ingressa na sala virtual.

Anexo 4 – Quadro 1 – Ferramentas da NetAula

Ícones	Ferramentas
	Conteúdo <i>On-line</i>
	Informações Gerais
	E-mail
	Avaliação
	Dados pessoais
	Biblioteca
	Chat
	Fórum
	Minha turma
	Trocar de turma
	Sair

Conteúdo On-line

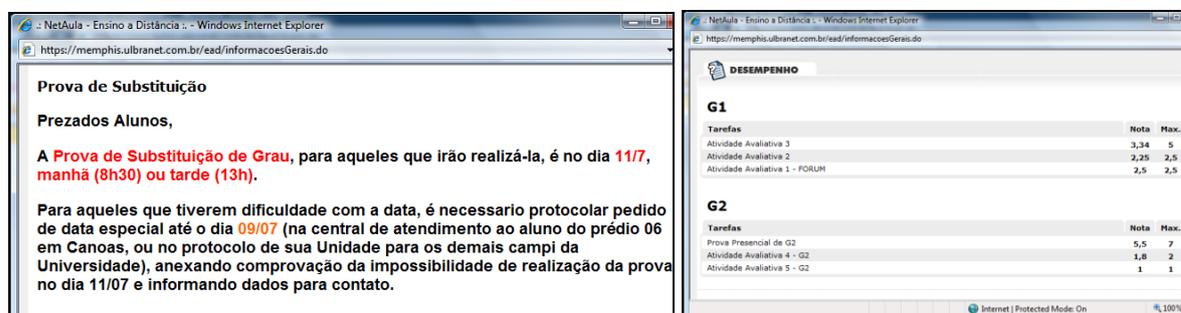
Na ferramenta Conteúdo *On-line* estão disponíveis os materiais base da disciplina ou módulo em estudo, normalmente diagramados nos formatos HTML ou PDF, além de vídeos produzidos sobre as temáticas abordadas. A figura 2 apresenta uma lista de materiais disponíveis nessa ferramenta, com exemplos abertos.



Anexo 4 – Figura 2 – NetAula: Conteúdo *On-line*

Informações Gerais

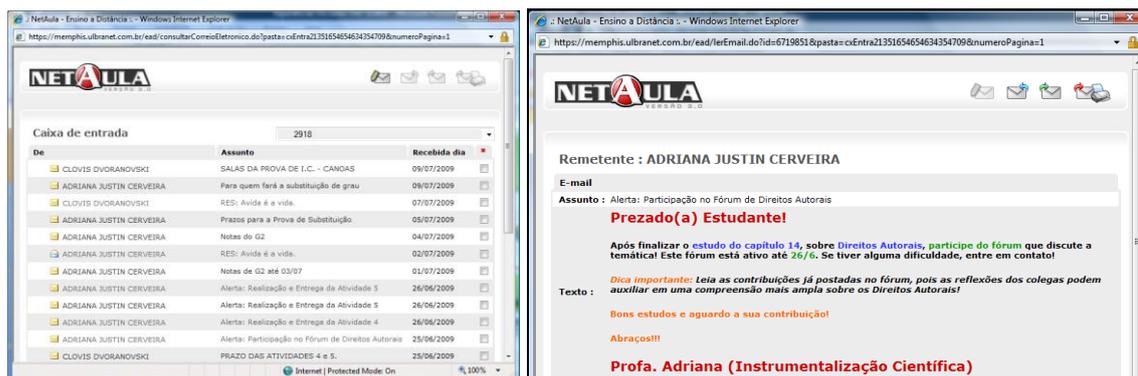
As informações gerais (acessíveis ao ingressar no ambiente por meio de uma janela *pop-up* ou por meio do ícone correspondente) apresentam aspectos gerais que visam a auxiliar o aluno em suas atividades (figura 3). Por meio do “mural de recados”, o professor pode chamar a atenção dos alunos para as novidades que estão surgindo no ambiente, tais como materiais complementares que disponibilizou ou discussões que estão interessantes. É sempre oportuno, também, retomar prazos e orientações de formas de estudo que auxiliem ao aluno na sua organização.



Anexo 4 – Figura 3 – NetAula: Informações Gerais

E-mail

A ferramenta “e-mail” implementa um correio eletrônico interno à sala virtual específica em que o aluno se encontra (figura 4). Assim, é possível escrever uma mensagem apenas para as pessoas daquela turma, professores ou alunos, selecionando um ou mais nomes na lista da turma.

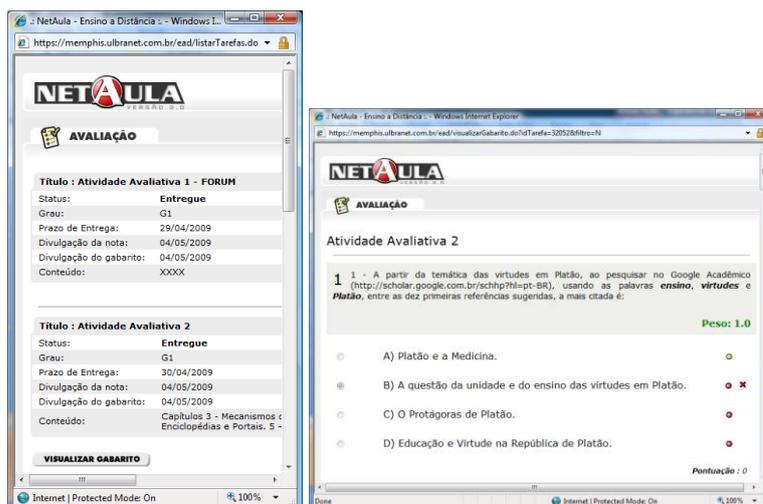


Anexo 4 – Figura 4 – NetAula: E-mail

De forma similar ao correio eletrônico convencional, é possível enviar arquivos anexados, responder, encaminhar ou imprimir as mensagens recebidas. Uma característica dessa ferramenta é que, ao enviar um e-mail para alguém, a pessoa recebe uma cópia da mensagem em seu endereço eletrônico externo (desde que esse esteja atualizado, conforme visualização nos “dados pessoais”).

Avaliação

A ferramenta “avaliação” permite a oferta de quatro tipos de atividades: questões discursivas; questões objetivas, de escolha simples ou múltipla; questões associativas; e envio de arquivos. Quando o professor disponibiliza uma avaliação, o aluno passa a visualizar o conjunto das questões colocadas e o prazo de entrega. No caso de questões discursivas, objetivas e associativas, o aluno pode fazer as questões (todas ou parte), salvá-las e retomá-las depois, enviando ao professor, conforme o prazo, quando finalizar. Ao enviar, o status da atividade muda de “não entregue” para “entregue”. A figura 5 apresenta a visualização de um conjunto de atividades avaliativas em andamento.



Anexo 4 – Figura 5 – NetAula: Avaliação

Após a correção e liberação do professor, o aluno passa a visualizar a nota obtida na atividade, o gabarito das questões objetivas e os comentários que podem ser acrescentados pelo professor em qualquer tipo de avaliação. Quando a atividade refere-se ao envio de arquivo, o professor pode disponibilizar um outro arquivo com correções, mantendo o original do aluno para que esse possa comparar e realizar suas análises.

Dados Pessoais

O recurso “dados pessoais” contém as informações de contato do próprio aluno. É por meio dessa ferramenta que o aluno pode enviar a sua foto e alterar seus dados pessoais, tais como telefone de contato e *e-mail* pessoal, facilitando que o professor possa localizá-lo em caso de necessidade. Manter essas informações atualizadas deve ser um compromisso do aluno. É importante destacar que os demais colegas não têm acesso a essas informações.

Biblioteca Virtual

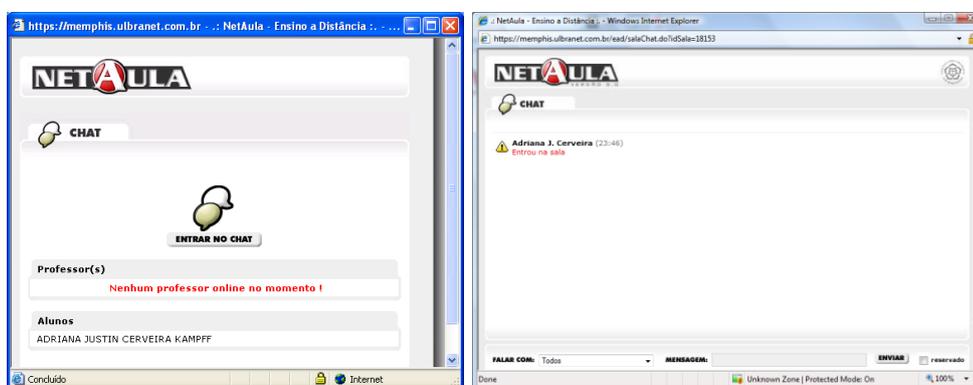
A ferramenta “biblioteca virtual” possibilita o acesso a diferentes materiais e indicações do professor (figura 6). Por meio dela, é possível ter acesso a arquivos em variados formatos e *links* para *sites* na Internet.



Anexo 4 – Figura 6 – NetAula: Biblioteca Virtual

Chat

Na ferramenta de *chat*, é possível visualizar quais professores e alunos estão conectados na plataforma naquele momento (figura 7). Clicando no botão “entrar no *chat*”, tem-se acesso à sala de conversação, que é única por turma. O professor pode agendar com os alunos um bate-papo (atividade síncrona) para a discussão de determinados temas; ou mesmo os próprios alunos podem combinar um horário para estudarem conjuntamente pontos do conteúdo ou organizarem trabalhos em grupo.

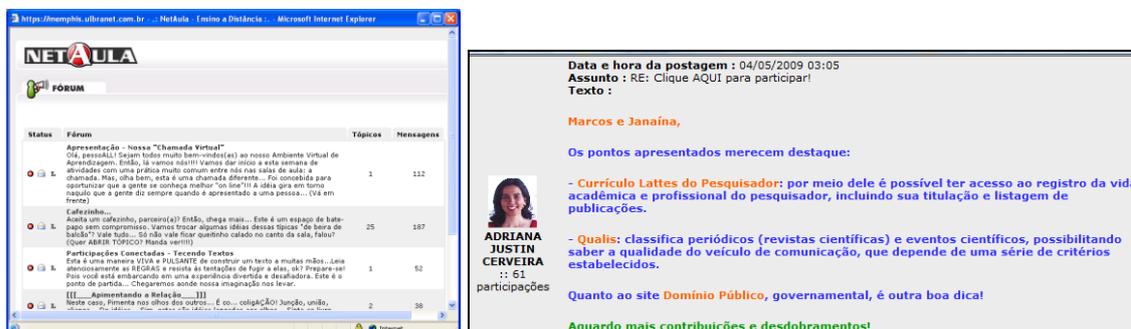


Anexo 4 – Figura 7 – NetAula: Chat

Fórum

A ferramenta fórum é uma das mais interessantes do ponto de vista pedagógico. Podem-se criar diversas discussões, categorizadas conforme a organização do professor, visando a esclarecer dúvidas, aprofundar temas ou mesmo estabelecer vínculos afetivos que favoreçam o engajamento dos alunos no ambiente de estudo a distância (figura 8). O que é discutido nos fóruns fica ali registrado durante todo o período letivo, podendo o aluno utilizar esse espaço como

fonte de estudo valiosa, onde várias questões foram apresentadas, respondidas, sintetizadas e aprofundadas, favorecendo a construção do conhecimento de modo coletivo.



The image shows a screenshot of the NetAula forum interface. On the left, there is a table with columns for 'Status', 'Fórum', 'Tópicos', and 'Mensagens'. The table lists several topics, including 'Apresentação - Nossa "Chamada Virtual"', 'Cafezinho', and 'Participações Conectadas - Tecendo Textos'. On the right, there is a detailed view of a post by Adriana Justin Cerveira. The post includes the date and time of posting (04/05/2009 03:05), the subject (RE: Clique AQUI para participar!), and the text of the post. The text discusses the importance of having access to the researcher's curriculum and professional information, and mentions the quality of the communication vehicle, which depends on a series of criteria established by the site. The post also mentions the site's public, governmental domain and expresses a desire for more contributions and unfoldings.

Status	Fórum	Tópicos	Mensagens
0	1	1	112
0	1	25	187
0	1	1	52
0	1	2	38

Data e hora da postagem : 04/05/2009 03:05
Assunto : RE: Clique AQUI para participar!
Texto :

Marcos e Janaína,

Os pontos apresentados merecem destaque:

- **Currículo Lattes do Pesquisador:** por meio dele é possível ter acesso ao registro da vida acadêmica e profissional do pesquisador, incluindo sua titulação e listagem de publicações.
- **Qualis:** classifica periódicos (revistas científicas) e eventos científicos, possibilitando saber a qualidade do veículo de comunicação, que depende de uma série de critérios estabelecidos.

Quanto ao site Domínio Público, governamental, é outra boa dica!

Aguardo mais contribuições e desdobramentos!

ADRIANA JUSTIN CERVEIRA
 :: 61 participações

Anexo 4 – Figura 8 – NetAula: Fórum

ANEXO 5

Apresentação do Projeto de Pesquisa à Instituição para fins de Liberação dos Dados

Mineração de Dados gerados na Interação com o Ambiente NetAula: Busca da Descoberta de Padrões de Acesso e Interação dos Alunos

A Educação a Distância (EAD) está em franca expansão no Brasil, sendo modalidade de oferta de um número crescente de cursos, tanto por instituições públicas quanto privadas. Nesse cenário, a Universidade Luterana do Brasil apresenta-se como referência, em função de sua abrangência no território nacional, número de alunos matriculados na modalidade e qualidade do material didático ofertado. Aliado à ampla oferta é preciso garantir a eficácia dessa educação. Formar professores para atuar nessa nova modalidade e prover tecnologias adequadas para o diálogo com o aluno, para a identificação das suas necessidades e orientação dos percursos didáticos oferecidos e trilhados tornam-se questões fundamentais. Considerando o grande volume de alunos, é preciso pesquisar e construir ferramentas que auxiliem ao professor na identificação de padrões de comportamento, na tomada de decisão e na interação com os alunos que necessitam de determinadas intervenções, estabelecendo comunicações mais eficazes.

A presente proposta de trabalho, neste contexto, busca – a partir do referencial teórico, da pesquisa com professores que atuam nesta modalidade e, especialmente, da mineração de dados gerados no Ambiente Virtual de Aprendizagem da Instituição (NetAula) – identificar padrões de acesso e interação de alunos que os levam a obter êxito (ou não) e, dessa forma, oferecer embasamento para a construção de ferramentas que auxiliem na prática docente, reduzindo índices de evasão e reprovação. A descoberta de padrões fornecerá indicadores para alertar os professores sobre possíveis comportamentos de alunos com propensão à evasão ou à reprovação, possibilitando uma intervenção e acompanhamento pró-ativo, de maneira a reverter este quadro antes da sua confirmação.

Entende-se que muitas das práticas docentes de resgate, engajamento e recuperação de estudos podem ser preventivas e dirigidas a alunos que, identificados por seus comportamentos, demandam uma atenção maior. De maneira mais eficaz, os professores podem ser auxiliados por alertas (automatizados pelo sistema) e históricos (gerados sob demanda do professor), disponibilizados no Ambiente Virtual de Aprendizagem (AVA) utilizado, constituindo-se em um diferencial para a tomada de decisão e intervenção junto aos alunos, visando a uma ação pedagógica de qualidade.

Para a busca de padrões, não serão identificados os alunos, mas agrupados por características. A mineração poderá ser feita sobre toda base de alunos matriculados nas disciplinas a distância da graduação presencial e nas disciplinas da graduação a distância. Como projeto inicial, são necessários os dados gerados na NetAula em 2006/2 e 2008/2 das turmas a distância da disciplina de Instrumentalização Científica, disciplina essa que compõe o Ciclo de Formação Geral da Universidade (juntamente com Comunicação e Expressão, Cultura Religiosa, e Sociedade e Contemporaneidade), com sede na Unidade Canoas e alunos das Unidades Sul.

Além das tabelas com os dados gerados nas diversas ferramentas (datas e quantidades – acessos aos conteúdos e arquivos na biblioteca; acessos à ferramenta de avaliações, ações, datas de entrega, acertos, notas; envio de e-mail; participações em fóruns...) e dos logs de acesso (datas e ferramentas), pretende-se avaliar as relações com o sexo, a idade, o curso e o unidade presencial do aluno. A partir desses dados, também são necessárias as notas finais de G1, G2 e S para buscar relações entre os comportamentos e os desempenhos. Esses dados podem ser fornecidos nas tabelas originais (convertidos em formato Access) ou em uma planilha única, linearizada, a ser estabelecida entre a TI e a pesquisadora, para que possam ser submetidos às ferramentas que implementam os recursos de mineração de dados.

ANEXO 6

Encaminhamento da Solicitação de Dados Históricos à TI

Projeto de Mineração de Dados Educacionais

1ª Etapa

Problema: Identificar características e comportamentos de alunos (em suas variações temporais) e suas relações com aprovação, reprovação ou evasão em disciplinas na modalidade a distância.

Dados Necessários:

- Semestre
- Turma
- Id Aluno
- Unidade Presencial
- Curso
- Sexo
- Idade
- Notas finais – G1, G2, S, MF, status (aprovado, reprovado, cancelado...)
- Quantidade de acessos ao Ambiente (Sala Virtual) por semana
- Quantidade de acessos aos Materiais de Aula por semana
- Quantidade de acessos à Biblioteca Virtual por semana
- Quantidade de acessos aos Fóruns por semana
- Quantidade de postagens nos Fóruns por semana
- Quantidade de e-mails recebidos por semana
- Quantidade de e-mails enviados por semana
- Avaliações entregues – nota, datas (liberação, prazo e entrega) e grau que compõe (de cada uma).

Formato - planilhas Excel, por turma:

Planilha1_TurmaX (Dados Gerais)

Semestre / Turma / IdAluno / UnidadePresencial / Curso / Sexo / Idade / G1 / G2 / S / MF / Status

Planilha2_TurmaX (Quantidade de Acesso à Sala Virtual por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha3_TurmaX (Quantidade de Acesso aos Materiais de Aula por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha4_TurmaX (Quantidade de Acesso à Biblioteca Virtual por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha5_TurmaX (Quantidade de Acesso aos Fóruns por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha6_TurmaX (Quantidade de Postagens nos Fóruns por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha7_TurmaX (Quantidade de E-mails recebidos por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha8_TurmaX (Quantidade de E-mails enviados por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha9_TurmaX (Detalhamento das Avaliações Entregues)

Semestre / Turma / IdAluno / Nota1 / DataEntrega1 / GrauQueCompõe1 ... NotaN / DataEntregaN / GrauQueCompõeN

Turmas Alvo (Disciplina de Instrumentalização Científica):

- **Semestre 2006/2**
 - 3128, 3129, 3130, 3132, 3215
- **Semestre 2008/2**
 - 3237, 3242, 3247, 3251, 3255, 3259, 3263, 3302, 3303, 3304, 3305, 3306, 3307, 3350, 3383

Intervalos de Tempo - Semanas:

Semanas / Semestre	2006.2	2008.2
S1	07/08/2006 a 13/08/2006	04/08/2008 a 10/08/2008
S2	14/08/2006 a 20/08/2006	11/08/2008 a 17/08/2008
S3	21/08/2006 a 27/08/2006	18/08/2008 a 24/08/2008
S4	28/08/2006 a 03/09/2006	25/08/2008 a 31/08/2008
S5	04/09/2006 a 10/09/2006	01/09/2008 a 07/09/2008
S6	11/09/2006 a 17/09/2006	08/09/2008 a 14/09/2008
S7	18/09/2006 a 24/09/2006	15/09/2008 a 21/09/2008
S8	25/09/2006 a 01/10/2006	22/09/2008 a 28/09/2008
S9	02/10/2006 a 08/10/2006	29/09/2008 a 05/10/2008
S10	09/10/2006 a 15/10/2006	06/10/2008 a 12/10/2008
S11	16/10/2006 a 22/10/2006	13/10/2008 a 19/10/2008
S12	23/10/2006 a 29/10/2006	20/10/2008 a 26/10/2008
S13	30/10/2006 a 05/11/2006	27/10/2008 a 02/11/2008
S14	06/11/2006 a 12/11/2006	03/11/2008 a 09/11/2008
S15	13/11/2006 a 19/11/2006	10/11/2008 a 16/11/2008
S16	20/11/2006 a 26/11/2006	17/11/2008 a 23/11/2008
S17	27/11/2006 a 03/12/2006	24/11/2008 a 30/11/2008
S18	04/12/2006 a 10/12/2006	01/12/2008 a 07/12/2008
S19	11/12/2006 a 17/12/2006	08/12/2008 a 14/12/2008

ANEXO 7

Descrição dos Dados para Mineração

	Tipo MD	Nome do Campo	Tipo do Valor	Descrição
1	id	Id_Us	nominal	Identificador do aluno.
2	label	Resultado	nominal	Resultado (Status Final): Aprovado, Reprovado, Evadido ou Sem Acesso.
3	regular	Unidade	nominal	Unidade (Polo) presencial em que o aluno está matriculado.
4	regular	Curso	nominal	Curso de graduação ao qual o aluno está vinculado.
5	regular	Área	nominal	Área a que pertence o curso que o aluno realiza, conforme agrupamento da Universidade.
6	regular	Sexo	nominal	Sexo do aluno.
7	regular	Idade	integer	Idade do aluno.
8	regular	Nota1	real	Nota obtida na primeira atividade, que compõe nota de grau 1.
9	regular	Nota2	real	Nota obtida na segunda atividade, que compõe nota de grau 1.
10	regular	Nota3	real	Nota obtida na terceira atividade, que compõe nota de grau 1.
11	regular	NAt_G1	integer	Número de atividades entregues para composição de grau 1.
12	regular	Entrega1	nominal	Status de entrega da primeira atividade, conforme prazo (Sim, Não, Atraso).
13	regular	Entrega2	nominal	Status de entrega da segunda atividade, conforme prazo (Sim, Não, Atraso).
14	regular	Entrega3	nominal	Status de entrega da terceira atividade, conforme prazo (Sim, Não, Atraso).
15	regular	Nota4	real	Nota obtida na quarta atividade, que compõe nota de grau 2.
16	regular	Nota5	real	Nota obtida na quinta atividade, que compõe nota de grau 2.
17	regular	Nota6	real	Nota obtida na sexta atividade (prova presencial), que compõe nota de grau 2.
18	regular	NAt_G2	integer	Número de atividades entregues para composição de grau 2.
19	regular	Entrega4	nominal	Status de entrega da quarta atividade, conforme prazo (Sim, Não, Atraso).
20	regular	Entrega5	nominal	Status de entrega da quinta atividade, conforme prazo (Sim, Não, Atraso).
21	regular	Pri_Ac	integer	Número da semana em que foi realizado o primeiro acesso ao ambiente.
22	regular	Ult_Ac	integer	Número da semana em que foi realizado o último acesso ao ambiente.
23	regular	AG_TM1	integer	Quantidade total de acessos à sala virtual no mês 1 (semanas de 1 a 4).
24	regular	FREQ_M1	integer	Frequência de acessos à sala virtual no mês 1.*
25	regular	AG_TM2	integer	Quantidade total de acessos à sala virtual no mês 2 (semanas de 5 a 8).
26	regular	FREQ_M2	integer	Frequência de acessos à sala virtual no mês 2.*
27	regular	AG_TM3	integer	Quantidade total de acessos à sala virtual no mês 3 (semanas de 9 a 11).
28	regular	FREQ_M3	integer	Frequência de acessos à sala virtual no mês 3.*
29	regular	AG_TM4	integer	Quantidade total de acessos à sala virtual no mês 4 (semanas de 12 a 15).
30	regular	FREQ_M4	integer	Frequência de acessos à sala virtual no mês 4.*
31	regular	AG_TM5	integer	Quantidade total de acessos à sala virtual no mês 5 (semanas de 16 a 19).
32	regular	FREQ_M5	integer	Frequência de acessos à sala virtual no mês 5.*
33	regular	AB_TM1	integer	Quantidade total de acessos aos conteúdos de aula e biblioteca virtual no mês 1.
34	regular	AB_FM1	integer	Frequência de acessos aos conteúdos de aula e biblioteca virtual no mês 1.*
35	regular	AB_TM2	integer	Quantidade total de acessos aos conteúdos de aula e biblioteca virtual no mês 2.
36	regular	AB_FM2	integer	Frequência de acessos aos conteúdos de aula e biblioteca virtual no mês 2.*
37	regular	AB_TM3	integer	Quantidade total de acessos aos conteúdos de aula e biblioteca virtual no mês 3.
38	regular	AB_FM3	integer	Frequência de acessos aos conteúdos de aula e biblioteca virtual no mês 3.*
39	regular	AB_TM4	integer	Quantidade total de acessos aos conteúdos de aula e biblioteca virtual no mês 4.
40	regular	AB_FM4	integer	Frequência de acessos aos conteúdos de aula e biblioteca virtual no mês 4.*
41	regular	AB_TM5	integer	Quantidade total de acessos aos conteúdos de aula e biblioteca virtual no mês 5.

	Tipo MD	Nome do Campo	Tipo do Valor	Descrição
42	regular	AB_FM5	integer	Frequência de acessos aos conteúdos de aula e biblioteca virtual no mês 5.*
43	regular	AF_TM1	integer	Quantidade total de acessos aos fóruns de discussão no mês 1.
44	regular	AF_FM1	integer	Frequência de acessos aos fóruns de discussão no mês 1.*
45	regular	AF_TM2	integer	Quantidade total de acessos aos fóruns de discussão no mês 2.
46	regular	AF_FM2	integer	Frequência de acessos aos fóruns de discussão no mês 2.*
47	regular	AF_TM3	integer	Quantidade total de acessos aos fóruns de discussão no mês 3.
48	regular	AF_FM3	integer	Frequência de acessos aos fóruns de discussão no mês 3.*
49	regular	AF_TM4	integer	Quantidade total de acessos aos fóruns de discussão no mês 4.
50	regular	AF_FM4	integer	Frequência de acessos aos fóruns de discussão no mês 4.*
51	regular	AF_TM5	integer	Quantidade total de acessos aos fóruns de discussão no mês 5.
52	regular	AF_FM5	integer	Frequência de acessos aos fóruns de discussão no mês 5.*
53	regular	PF_TM1	integer	Quantidade total de postagens nos fóruns de discussão no mês 1.
54	regular	PF_FM1	integer	Frequência de postagens aos fóruns de discussão no mês 1.*
55	regular	PF_TM2	integer	Quantidade total de postagens nos fóruns de discussão no mês 2.
56	regular	PF_FM2	integer	Frequência de postagens aos fóruns de discussão no mês 2.*
57	regular	PF_TM3	integer	Quantidade total de postagens nos fóruns de discussão no mês 3.
58	regular	PF_FM3	integer	Frequência de postagens aos fóruns de discussão no mês 3.*
59	regular	PF_TM4	integer	Quantidade total de postagens nos fóruns de discussão no mês 4.
60	regular	PF_FM4	integer	Frequência de postagens aos fóruns de discussão no mês 4.*
61	regular	PF_TM5	integer	Quantidade total de postagens nos fóruns de discussão no mês 5.
62	regular	PF_FM5	integer	Frequência de postagens aos fóruns de discussão no mês 5.*
63	regular	EE_TM1	integer	Quantidade total de e-mails enviados no mês 1.
64	regular	EE_FM1	integer	Frequência de envio de e-mails no mês 1.*
65	regular	EE_TM2	integer	Quantidade total de e-mails enviados no mês 2.
66	regular	EE_FM2	integer	Frequência de envio de e-mails no mês 2.*
67	regular	EE_TM3	integer	Quantidade total de e-mails enviados no mês 3.
68	regular	EE_FM3	integer	Frequência de envio de e-mails no mês 3.*
69	regular	EE_TM4	integer	Quantidade total de e-mails enviados no mês 4.
70	regular	EE_FM4	integer	Frequência de envio de e-mails no mês 4.*
71	regular	EE_TM5	integer	Quantidade total de e-mails enviados no mês 5.
72	regular	EE_FM5	integer	Frequência de envio de e-mails no mês 5.*
73	regular	ER_TM1	integer	Quantidade total de e-mails recebidos no mês 1.
74	regular	ER_FM1	integer	Frequência de recebimento de e-mails no mês 1.*
75	regular	ER_TM2	integer	Quantidade total de e-mails recebidos no mês 2.
76	regular	ER_FM2	integer	Frequência de recebimento de e-mails no mês 2.*
77	regular	ER_TM3	integer	Quantidade total de e-mails recebidos no mês 3.
78	regular	ER_FM3	integer	Frequência de recebimento de e-mails no mês 3.*
79	regular	ER_TM4	integer	Quantidade total de e-mails recebidos no mês 4.
80	regular	ER_FM4	integer	Frequência de recebimento de e-mails no mês 4.*
81	regular	ER_TM5	integer	Quantidade total de e-mails recebidos no mês 5.
82	regular	ER_FM5	integer	Frequência de recebimento de e-mails no mês 5.*
83	regular	G1	real	Nota final de Grau 1.
84	regular	G2	real	Nota final de Grau 2.
85	regular	S1	real	Nota de Substituição de Grau 1.
86	regular	S2	real	Nota de Substituição de Grau 2.
87	regular	MF	real	Média Final do aluno.

* Frequência: 0 – nenhum acesso no mês; 1 – acessos em uma das semanas do mês; 2 – acessos em duas das semanas do mês; 3 – acessos em três semanas do mês; 4 – acessos em quatro das semanas do mês.

ANEXO 8

Regras de Classificação – Dados Históricos Mês 1

Foram geradas 74 regras ao todo, utilizando o algoritmo *RuleLearner*, implementado na ferramenta *RapidMiner*, como resultado da MD históricos de 1564 alunos, com 20 atributos de cada aluno, referentes ao mês um.

RuleModel

1. if AF_TM1 > 6.500 and AG_TM1 > 25.500 and AF_TM1 <= 53 then Aprovado (66 / 0 / 0 / 0)
2. if AF_TM1 > 6.500 and Unidade = Gravataí and Sexo = Masculino then Aprovado (26 / 0 / 0 / 0)
3. if AF_TM1 > 6.500 and Unidade = Canoas and Área = SaúdeeBem-EstarSocial then Aprovado (110 / 2 / 3 / 0)
4. if FREQ_M1 > 1.500 and AB_TM1 > 17.500 and AF_TM1 > 26.500 and AB_TM1 <= 93.500 and AF_TM1 <= 32.500 then Aprovado (25 / 0 / 0 / 0)
5. if FREQ_M1 > 1.500 and Área = TecnologiaeComputação and FREQ_M1 > 3.500 then Aprovado (73 / 3 / 1 / 0)
6. if FREQ_M1 > 1.500 and AF_TM1 > 13.500 and Idade > 26.500 and Idade <= 35.500 then Aprovado (57 / 0 / 0 / 0)
7. if FREQ_M1 > 1.500 and AF_TM1 > 13.500 and AF_FM1 > 2.500 and Curso = DIREITO then Aprovado (30 / 0 / 0 / 0)
8. if FREQ_M1 > 1.500 and AF_TM1 > 13.500 and AF_FM1 > 2.500 and AB_TM1 <= 45.500 and AG_TM1 > 14.500 and AB_TM1 > 28.500 then Aprovado (26 / 0 / 0 / 0)
9. if FREQ_M1 > 1.500 and Idade <= 29.500 and AF_FM1 > 3.500 and AB_TM1 <= 21.500 and Área = EducaçãoCiênciaseArtes then Aprovado (19 / 0 / 0 / 0)
10. if FREQ_M1 > 1.500 and Unidade = Canoas and AF_TM1 > 0.500 then Aprovado (340 / 42 / 16 / 0)
11. if AG_TM1 > 1.500 and PF_TM1 > 1.500 and AF_TM1 <= 12.500 and AF_TM1 > 10.500 then Aprovado (26 / 0 / 0 / 0)
12. if AG_TM1 > 1.500 and Curso = DIREITO and Idade <= 27.500 then Aprovado (53 / 5 / 3 / 0)
13. if AF_TM1 > 0.500 and Pri_Ac <= 2.500 and AB_TM1 <= 29.500 and AF_TM1 > 1.500 and Sexo = Feminino and ER_TM1 <= 3.500 then Aprovado (42 / 2 / 0 / 0)
14. if AF_TM1 > 0.500 and PF_TM1 > 0.500 and AF_TM1 <= 12.500 and Área = SaúdeeBem-EstarSocial and EE_TM1 <= 0.500 then Aprovado (28 / 0 / 0 / 0)
15. if AF_TM1 > 0.500 and Idade > 25.500 and FREQ_M1 > 2.500 and AG_TM1 <= 9.500 then Aprovado (23 / 0 / 0 / 0)
16. if AF_TM1 > 0.500 and PF_TM1 > 0.500 and AB_TM1 <= 10.500 and AB_TM1 <= 2.500 and Pri_Ac > 1.500 then Aprovado (18 / 0 / 0 / 0)
17. if Idade <= 22.500 and ER_FM1 <= 2.500 and Curso = MEDICINA then Aprovado (13 / 0 / 0 / 0)
18. if AG_TM1 > 1.500 and AB_FM1 > 2.500 and Unidade = Gravataí and Área = CiênciasSociaisAplicadas then Aprovado (10 / 0 / 0 / 0)
19. if AG_TM1 > 1.500 and AF_TM1 > 5.500 and AB_TM1 <= 10.500 and AB_TM1 > 3.500 and AF_TM1 <= 15 then Aprovado (14 / 0 / 0 / 0)
20. if AG_TM1 > 1.500 and Unidade = Guaíba and PF_TM1 <= 0.500 then Aprovado (11 / 0 / 0 / 0)
21. if AG_TM1 > 1.500 and AB_TM1 > 44.500 and AB_TM1 <= 59 then Aprovado (12 / 0 / 0 / 0)
22. if Idade <= 22.500 and Sexo = Masculino and AG_TM1 <= 3.500 then Aprovado (41 / 7 / 3 / 2)
23. if AF_TM1 > 2.500 and Idade <= 43.500 and Idade > 33 and PF_TM1 > 0.500 then Aprovado (12 / 0 / 0 / 0)
24. if AG_TM1 > 4.500 and AG_TM1 <= 12.500 and Pri_Ac > 1.500 and PF_TM1 <= 1.500 and AB_TM1 > 5.500 and AB_TM1 <= 19.500 and ER_TM1 > 1.500 and ER_TM1 <= 14.500 and Área = EducaçãoCiênciaseArtes then Aprovado (7 / 0 / 0 / 0)
25. if Área = SaúdeeBem-EstarSocial and Curso = ODONTOLOGIA then Aprovado (15 / 3 / 1 / 0)
26. if Unidade = Gravataí and AG_TM1 > 2.500 then Aprovado (22 / 6 / 3 / 0)

27. if Pri_Ac > 2.500 and Idade > 29.500 and AF_TM1 <= 3.500 and ER_TM1 > 1.500 and Área = EducaçãoCiênciaseArtes then Aprovado (5 / 0 / 0 / 0)
28. if Pri_Ac > 1.500 and Área = SaúdeeBem-EstarSocial and AB_TM1 <= 3.500 then Aprovado (20 / 5 / 1 / 0)
29. if AB_TM1 > 15.500 and AG_TM1 <= 12.500 and PF_TM1 <= 1.500 and AB_TM1 <= 19.500 and Idade <= 24.500 then Aprovado (5 / 0 / 0 / 0)
30. if Idade <= 25.500 and AF_TM1 <= 3.500 and AF_TM1 > 0.500 and ER_TM1 <= 7.500 and Sexo = Masculino then Evadido (0 / 7 / 0 / 0)
31. if Sexo = Masculino and AF_TM1 > 2.500 then Aprovado (27 / 14 / 4 / 0)
32. if Idade <= 33.500 and AB_TM1 > 0.500 and AG_TM1 <= 6.500 and ER_TM1 > 10.500 and FREQ_M1 > 1.500 then Evadido (0 / 6 / 0 / 0)
33. if Idade <= 33.500 and Pri_Ac > 2.500 and Idade > 22.500 and Pri_Ac > 6.500 and Idade <= 30 then Evadido (0 / 7 / 0 / 0)
34. if Pri_Ac > 5.500 and Idade > 34 then Aprovado (4 / 0 / 0 / 0)
35. if AB_FM1 > 2.500 and Idade > 26.500 and Idade <= 35.500 then Aprovado (6 / 0 / 0 / 0)
36. if AG_TM1 > 0.500 and Idade <= 37.500 and AB_FM1 <= 1.500 and Curso = EDUCACAOFISICA and AB_TM1 > 2.500 then Evadido (0 / 10 / 0 / 0)
37. if Pri_Ac > 1.500 and Curso = ADMINISTRACAO and AG_TM1 <= 6.500 then Aprovado (14 / 4 / 1 / 0)
38. if AG_TM1 > 0.500 and AG_TM1 <= 1.500 and Idade <= 30.500 and Idade > 19.500 and Idade <= 24.500 and AB_TM1 > 0.500 then Evadido (0 / 6 / 0 / 0)
39. if Pri_Ac > 1.500 and AB_TM1 <= 12.500 and Área = EducaçãoCiênciaseArtes and Pri_Ac > 3.500 then Aprovado (19 / 7 / 2 / 0)
40. if Pri_Ac > 2.500 and Idade > 25.500 and Idade <= 31.500 and AB_TM1 <= 2.500 and Sexo = Feminino then Evadido (0 / 8 / 0 / 0)
41. if AG_TM1 > 0.500 and AB_TM1 <= 21.500 and AB_TM1 > 6.500 and AF_TM1 > 4.500 then Evadido (0 / 8 / 0 / 0)
42. if Pri_Ac > 0.500 and ER_TM1 > 8.500 and Área = TecnologiaeComputação and Idade <= 42.500 then Aprovado (7 / 0 / 0 / 0)
43. if AG_TM1 > 0.500 and AF_TM1 <= 4.500 and Idade <= 31.500 and FREQ_M1 <= 1.500 and Área = CiênciasSociaisAplicadas then Evadido (0 / 5 / 0 / 0)
44. if Pri_Ac > 1.500 and Idade > 30.500 and ER_TM1 <= 6 and AB_TM1 > 7.500 and Idade <= 57 then Aprovado (9 / 0 / 0 / 0)
45. if AG_TM1 > 0.500 and AF_TM1 <= 2.500 and AB_TM1 > 6.500 and Idade > 25.500 and Idade <= 43.500 and AB_TM1 <= 21 then Evadido (0 / 10 / 0 / 0)
46. if Pri_Ac > 0.500 and AG_TM1 <= 3.500 and Sexo = Masculino and ER_TM1 > 6 then Aprovado (6 / 0 / 0 / 0)
47. if AG_TM1 > 0.500 and AF_TM1 <= 2.500 and FREQ_M1 <= 1.500 and ER_TM1 > 1.500 and Idade > 21.500 and AG_TM1 <= 6.500 then Evadido (0 / 9 / 0 / 0)
48. if Pri_Ac > 2.500 and Idade <= 32.500 and Curso = DIREITO and ER_FM1 <= 2.500 then Aprovado (8 / 0 / 0 / 0)
49. if AG_TM1 > 1.500 and Idade <= 29.500 and AB_TM1 <= 33 and AF_TM1 > 1.500 and Idade > 23.500 then Aprovado (4 / 0 / 0 / 0)
50. if Pri_Ac > 0.500 and Idade > 27.500 and Curso = DIREITO and AB_TM1 <= 10.500 then Evadido (0 / 5 / 0 / 0)
51. if AG_TM1 > 1.500 and AB_TM1 <= 25 and ER_TM1 <= 7.500 and AF_TM1 <= 0.500 and Sexo = Feminino and AB_TM1 <= 9.500 then Aprovado (5 / 0 / 0 / 0)
52. if AG_TM1 > 2.500 and AG_TM1 <= 14.500 and AB_TM1 > 25 then Evadido (0 / 5 / 0 / 0)
53. if FREQ_M1 > 1.500 and AG_TM1 <= 3.500 and Idade > 21.500 and Idade <= 28 then Aprovado (6 / 0 / 0 / 0)
54. if Pri_Ac > 0.500 and AB_TM1 <= 10.500 and ER_TM1 <= 3.500 and Idade <= 28.500 and AG_TM1 > 1.500 then Evadido (0 / 6 / 0 / 0)
55. if AB_TM1 > 5.500 and Idade <= 22.500 and EE_TM1 <= 0.500 and Idade <= 20.500 then Aprovado (5 / 0 / 0 / 0)
56. if Idade <= 21.500 and Unidade = Canoas and Pri_Ac > 0.500 then Evadido (0 / 4 / 0 / 0)
57. if AG_TM1 > 0.500 and AB_TM1 <= 1.500 and Idade > 28.500 then Aprovado (4 / 0 / 0 / 0)
58. if Pri_Ac > 0.500 and AB_TM1 <= 8.500 and ER_FM1 <= 2.500 and AG_TM1 > 1.500 then Evadido (0 / 4 / 0 / 0)
59. if AF_TM1 > 3.500 and AB_TM1 <= 54 then Reprovado (0 / 0 / 5 / 0)
60. if AB_TM1 > 5.500 and ER_TM1 <= 7.500 and AB_TM1 <= 11.500 then Aprovado (4 / 0 / 0 / 0)
61. if Pri_Ac > 0.500 and AG_TM1 <= 7 and AB_TM1 > 3.500 and ER_TM1 > 6.500 and AB_TM1 > 6.500 then Evadido (0 / 3 / 0 / 0)
62. if Pri_Ac <= 0.500 then Sem_Acesso (0 / 0 / 0 / 11)

63. if AB_TM1 <= 0.500 and ER_TM1 <= 3 and Pri_Ac > 3.500 and Idade > 23 then Evadido (0 / 3 / 0 / 0)
64. if Idade > 23.500 and Idade <= 27.500 and Pri_Ac <= 4 then Reprovado (0 / 0 / 5 / 0)
65. if AF_TM1 > 0.500 and AG_TM1 <= 35 then Aprovado (5 / 0 / 0 / 0)
66. if AB_TM1 > 3.500 and AB_TM1 <= 5 then Evadido (0 / 2 / 0 / 0)
67. if ER_TM1 <= 3 and Unidade = Carazinho then Evadido (0 / 1 / 0 / 0)
68. if Idade <= 31 and ER_TM1 > 2 and Pri_Ac > 2 then Aprovado (3 / 0 / 0 / 0)
69. if Área = TecnologiaeComputação then Reprovado (0 / 0 / 3 / 0)
70. if Idade > 21.500 and Idade <= 30 then Evadido (0 / 2 / 0 / 0)
71. if Curso = CIENCIAPOLITICA then Evadido (0 / 1 / 0 / 0)
72. if Unidade = Gravataí then Sem_Acesso (0 / 0 / 0 / 1)
73. if Unidade = CachoeiradoSul then Aprovado (1 / 0 / 0 / 0)
74. else Reprovado (0 / 0 / 1 / 0)

correct: 1424 out of 1564 training examples.

ANEXO 9

Regras de Classificação – Dados Históricos Mês 2

Foram geradas 77 regras ao todo, utilizando o algoritmo *RuleLearner*, implementado na ferramenta *RapidMiner*, como resultado da MD históricos de 1564 alunos, com 36 atributos de cada aluno, referentes ao mês dois.

RuleModel

1. if Nota2 > 0.920 and AF_TM1 > 6.500 and Nota1 > 1.900 and AF_TM2 > 10.500 and AB_TM1 <= 85.500 and Área = SaúdeeBem-EstarSocial then Aprovado (70 / 0 / 0 / 0)
2. if Nota2 > 0.920 and AF_TM1 > 6.500 and AG_TM2 > 11.500 and FREQ_M1 > 2.500 and Curso = DIREITO then Aprovado (60 / 0 / 0 / 0)
3. if Nota2 > 0.920 and Nota1 > 2.225 and AF_TM1 > 6.500 and AG_TM2 > 11.500 and Área = EducaçãoCiênciaseArtes then Aprovado (94 / 0 / 0 / 0)
4. if Nota2 > 1.275 and Nota1 > 2.225 and FREQ_M2 > 2.500 and ER_TM2 <= 8.500 and Sexo = Masculino and Área = CiênciasSociaisAplicadas then Aprovado (46 / 0 / 0 / 0)
5. if Nota2 > 1.275 and Nota1 > 2.225 and AB_TM2 > 11.500 and ER_TM2 <= 8.500 and Área = TecnologiaeComputação then Aprovado (56 / 0 / 0 / 0)
6. if Nota2 > 1.275 and Nota1 > 2.225 and AB_TM2 > 11.500 and ER_TM2 <= 8.500 and AB_TM2 <= 34.500 and Área = EducaçãoCiênciaseArtes then Aprovado (56 / 0 / 0 / 0)
7. if Nota2 > 1.275 and Nota2 > 1.950 and AG_TM2 > 6.500 and ER_TM1 > 3.500 and Idade <= 20.500 then Aprovado (56 / 0 / 0 / 0)
8. if Nota2 > 1.275 and AF_TM2 > 4.500 and Nota2 > 1.950 and AF_TM1 > 1.500 and AB_TM1 > 20.500 and AG_TM1 > 5.500 then Aprovado (68 / 0 / 0 / 0)
9. if Nota2 > 0.920 and Nota1 > 1.900 and FREQ_M2 > 2.500 and AB_TM2 <= 9.500 and FREQ_M1 > 2.500 and ER_TM1 > 4.500 then Aprovado (74 / 0 / 0 / 0)
10. if Nota2 > 0.920 and Nota1 > 1.900 and FREQ_M2 > 2.500 and ER_TM2 <= 5.500 and Unidade = Canoas and PF_TM1 > 0.500 then Aprovado (48 / 0 / 0 / 0)
11. if Nota2 > 0.920 and Nota1 > 1.900 and AB_TM1 <= 9.500 and AF_TM1 > 1.500 and AB_TM2 > 7.500 and Nota2 > 1.655 then Aprovado (59 / 0 / 0 / 0)
12. if Nota2 > 0.920 and Nota1 > 1.900 and AG_TM2 > 22.500 and AF_TM2 <= 35 then Aprovado (33 / 0 / 0 / 0)
13. if Nota2 > 0.920 and Nota1 > 1.625 and AF_TM1 > 1.500 and AB_TM1 <= 6.500 and AB_TM2 <= 6.500 and Sexo = Masculino then Aprovado (26 / 0 / 0 / 0)
14. if Nota1 > 1.375 and Nota2 > 1.275 and AF_TM1 > 1.500 and ER_TM1 > 8.500 and AF_TM1 <= 17.500 and ER_TM2 <= 11.500 and AB_TM2 <= 6.500 then Aprovado (21 / 0 / 0 / 0)
15. if Nota1 > 1.375 and Nota2 > 1.275 and AG_TM2 > 14.500 and AF_TM2 <= 27 and AB_TM1 > 31.500 then Aprovado (19 / 0 / 0 / 0)
16. if Nota1 > 1.375 and Nota2 > 1.275 and PF_TM1 > 1.500 and AF_TM1 <= 12.500 and AB_FM2 > 1.500 and ER_TM2 <= 25.500 then Aprovado (51 / 0 / 0 / 0)
17. if Nota1 > 1.375 and Nota2 > 1.275 and AG_TM2 > 6.500 and AF_TM2 <= 6.500 and AB_TM1 <= 6.500 and ER_TM2 <= 7.500 then Aprovado (32 / 0 / 0 / 0)
18. if Nota1 > 1.625 and AF_TM2 > 1.500 and FREQ_M1 > 3.500 and Nota2 > 0.665 and ER_TM1 <= 6.500 then Aprovado (17 / 0 / 0 / 0)
19. if Nota1 > 1.625 and Idade <= 22.500 and ER_FM1 <= 2.500 and AG_TM2 > 6.500 and AB_TM2 <= 35.500 and Nota2 <= 2.220 then Aprovado (38 / 0 / 0 / 0)
20. if Nota2 > 0.920 and AF_TM2 > 1.500 and ER_TM1 > 9.500 and PF_TM2 <= 0.500 and AF_TM2 <= 13 then Aprovado (31 / 0 / 0 / 0)

21. if Nota2 > 0.920 and AF_TM2 > 13.500 and Idade > 24.500 and AF_TM2 <= 27 and AB_TM1 > 3 then Aprovado (22 / 0 / 0 / 0)
22. if Nota1 > 1.375 and ER_TM1 > 5.500 and AB_TM1 <= 1.500 and AF_TM2 <= 18 and ER_TM2 <= 11.500 then Aprovado (19 / 0 / 0 / 0)
23. if Nota1 > 1.375 and AB_FM1 > 2.500 and Unidade = Gravataí and FREQ_M1 <= 3.500 then Aprovado (12 / 0 / 0 / 0)
24. if Nota1 > 1.375 and Área = SaúdeeBem-EstarSocial and ER_TM1 <= 2.500 then Aprovado (27 / 2 / 3 / 0)
25. if Nota1 > 1.175 and AB_FM1 > 2.500 and AB_TM2 <= 28.500 and AB_TM2 > 16.500 and Unidade = Canoas and AG_TM1 <= 13.500 then Aprovado (17 / 0 / 0 / 0)
26. if Nota2 > 0.920 and AG_TM2 > 23.500 and ER_TM2 <= 8.500 then Aprovado (14 / 0 / 0 / 0)
27. if Nota1 > 1.175 and ER_TM1 > 7.500 and AG_TM1 > 5.500 and AF_TM1 <= 7.500 and AB_TM1 <= 14.500 and AG_TM2 <= 19.500 then Aprovado (16 / 0 / 0 / 0)
28. if AG_TM2 <= 4.500 and Nota1 <= 1.875 and AB_TM1 > 0.500 and AF_FM2 <= 0.500 and ER_FM2 > 3.500 and Sexo = Feminino then Evadido (0 / 20 / 0 / 0)
29. if Nota1 > 1.175 and AG_TM1 > 13.500 and Nota2 <= 1.625 and AG_TM2 <= 18 and AG_TM2 > 4.500 then Aprovado (17 / 0 / 0 / 0)
30. if Nota2 <= 0.920 and AF_TM2 <= 4.500 and Pri_Ac > 1.500 and Área = EducaçãoCiênciaseArtes and Nota1 <= 1.750 then Evadido (1 / 24 / 1 / 0)
31. if Nota2 > 0.920 and AG_TM1 <= 5.500 and ER_TM2 > 9.500 and AB_TM2 > 4.500 and AB_FM2 <= 3.500 then Aprovado (12 / 0 / 0 / 0)
32. if Nota2 > 0.920 and ER_TM2 <= 4.500 and AB_FM1 <= 1.500 and Idade <= 32.500 and Sexo = Feminino then Aprovado (9 / 0 / 0 / 0)
33. if Pri_Ac > 2.500 and Nota2 <= 1.655 and ER_TM1 <= 7.500 and Idade > 23.500 and AF_TM1 <= 0.500 and Idade <= 31 and AB_TM2 <= 17 and AG_TM2 > 0.500 then Evadido (0 / 19 / 0 / 0)
34. if Nota1 > 1.775 and AB_TM2 <= 6.500 and AF_TM2 > 3.500 and AB_TM2 > 3.500 then Aprovado (9 / 0 / 0 / 0)
35. if FREQ_M2 > 3.500 and Entrega2 = Nao and AB_TM1 <= 13 then Aprovado (10 / 1 / 0 / 0)
36. if Nota2 > 0.920 and AG_TM1 <= 3.500 and Idade > 27.500 and AF_TM2 <= 2 and ER_TM2 <= 18 then Aprovado (11 / 0 / 0 / 0)
37. if ER_FM1 <= 2.500 and Idade > 23.500 and AF_TM1 <= 2.500 and AG_TM1 > 1.500 and Nota2 <= 1.530 and AG_TM1 <= 10.500 then Evadido (0 / 16 / 0 / 0)
38. if Idade <= 25.500 and Unidade = Canoas and Nota1 <= 0.500 then Aprovado (23 / 6 / 3 / 2)
39. if ER_TM2 > 11.500 and Idade <= 23.500 and AB_TM1 > 9.500 and AF_FM2 <= 3.500 then Evadido (0 / 10 / 0 / 0)
40. if ER_FM1 <= 2.500 and AB_FM2 > 1.500 and Nota1 > 1.775 and Área = EducaçãoCiênciaseArtes and FREQ_M2 > 3.500 then Evadido (1 / 10 / 0 / 0)
41. if Nota1 > 1.175 and AB_FM2 > 3.500 and Nota2 > 1.625 and Nota1 > 2.125 and AB_TM2 > 15.500 then Aprovado (11 / 0 / 0 / 0)
42. if Área = CiênciasSociaisAplicadas and AF_TM2 > 1.500 then Evadido (5 / 16 / 2 / 0)
43. if Nota1 > 1.175 and AB_TM2 <= 15.500 and AB_TM2 > 7.500 and AG_TM1 <= 6.500 and AB_TM2 > 11.500 and Idade > 23.500 then Aprovado (6 / 0 / 0 / 0)
44. if AG_TM2 > 3.500 and AB_FM2 <= 1.500 and Idade > 20.500 and AF_TM2 <= 1.500 and AF_FM1 <= 2.500 then Aprovado (9 / 0 / 0 / 0)
45. if AB_FM2 <= 2.500 and AB_TM2 > 3.500 and Idade > 21.500 and Nota1 > 2.375 then Evadido (0 / 8 / 0 / 0)
46. if Nota1 > 1.900 and AB_TM2 <= 15.500 and Idade <= 21.500 and AB_TM1 > 2.500 and AG_TM1 > 1.500 and AG_TM1 <= 10.500 and AB_TM1 <= 7 then Aprovado (7 / 0 / 0 / 0)
47. if Pri_Ac > 2.500 and AB_TM2 <= 6.500 and AF_TM2 <= 3.500 and Nota1 <= 2.125 and Sexo = Feminino then Evadido (0 / 9 / 0 / 0)
48. if PF_TM2 > 0.500 and ER_TM2 > 7.500 and FREQ_M2 <= 3.500 and Nota1 > 1.875 then Aprovado (8 / 0 / 0 / 0)
49. if Nota2 > 2.125 and ER_TM1 > 1 and Pri_Ac <= 4.500 then Evadido (0 / 9 / 0 / 0)
50. if AB_TM2 > 4.500 and AB_TM1 <= 21.500 and ER_FM1 <= 2.500 and Pri_Ac <= 2.500 and Nota2 > 1.125 and Nota2 <= 2.250 then Aprovado (9 / 0 / 0 / 0)
51. if AG_TM2 <= 9.500 and AB_TM1 > 3.500 and Nota2 <= 1.655 and AF_TM1 <= 1.500 then Evadido (0 / 9 / 0 / 0)
52. if Idade <= 32.500 and AB_TM2 > 30.500 and AB_TM2 <= 79 and Nota1 <= 0.250 then Aprovado (6 / 0 / 0 / 0)
53. if ER_FM1 <= 2.500 and AG_TM1 > 5.500 and AB_TM2 > 26 and Nota2 > 0.565 then Evadido (0 / 7 / 0 / 0)
54. if AB_TM1 > 0.500 and AG_TM1 <= 5.500 and Nota2 <= 0.375 and AF_TM2 > 3.500 then Aprovado (7 / 0 / 0 / 0)
55. if AB_FM2 <= 2.500 and AF_TM2 > 0.500 and Nota1 <= 1.875 and AF_TM1 > 6.500 then Evadido (0 / 9 / 0 / 0)

56. if AB_TM1 > 0.500 and Pri_Ac > 2.500 and AF_FM2 <= 0.500 then Aprovado (5 / 0 / 0 / 0)
57. if FREQ_M2 > 2.500 and ER_TM1 <= 8.500 and Idade > 24.500 and AF_FM2 > 1.500 then Reprovado (0 / 0 / 11 / 0)
58. if AB_TM1 > 2.500 and AB_TM1 <= 20.500 and AB_FM2 > 3.500 and ER_TM1 <= 8.500 then Aprovado (6 / 0 / 0 / 0)
59. if Pri_Ac > 3.500 and AB_FM2 <= 2.500 and AG_TM2 > 5.500 then Evadido (0 / 6 / 0 / 0)
60. if AB_TM2 > 4.500 and Área = TecnologiaeComputação and Nota1 > 1.875 then Aprovado (5 / 0 / 0 / 0)
61. if AG_TM2 > 3.500 and ER_TM1 <= 8.500 and ER_TM1 > 0.500 and Nota1 <= 1.375 then Reprovado (0 / 0 / 6 / 0)
62. if Pri_Ac > 1.500 and PF_TM1 <= 0.500 and Idade > 25.500 and ER_FM1 <= 2.500 then Evadido (0 / 7 / 0 / 0)
63. if AF_TM2 > 0.500 and Nota2 <= 0.875 and Idade > 22.500 and AB_TM2 <= 11 then Aprovado (7 / 0 / 0 / 0)
64. if Entrega2 = Sim and Idade <= 22.500 then Reprovado (5 / 4 / 12 / 0)
65. if Idade <= 27.500 and Nota1 <= 1.875 and AB_TM1 > 5.500 and Idade <= 25.500 then Evadido (0 / 8 / 0 / 0)
66. if ER_TM1 <= 7.500 and AG_TM2 > 0.500 and AG_TM1 <= 7.500 and Entrega1 = Sim then Aprovado (7 / 0 / 0 / 0)
67. if AB_TM1 > 4.500 and Pri_Ac > 1.500 then Reprovado (0 / 0 / 4 / 0)
68. if Pri_Ac > 1.500 and AB_TM2 <= 7.500 then Evadido (0 / 6 / 0 / 0)
69. if Pri_Ac <= 0.500 then Sem_Acesso (0 / 0 / 0 / 11)
70. if Nota1 > 1.875 and AB_TM2 <= 18.500 and ER_TM2 <= 8.500 then Reprovado (0 / 0 / 6 / 0)
71. if AG_TM1 > 12 and Idade > 23.500 then Aprovado (4 / 0 / 0 / 0)
72. if Sexo = Feminino and Idade > 25.500 then Evadido (0 / 4 / 0 / 0)
73. if Curso = EDUCACAOFISICA then Reprovado (0 / 0 / 2 / 0)
74. if ER_FM2 > 3.500 and AB_TM1 <= 15 then Aprovado (4 / 0 / 0 / 0)
75. if PF_TM1 > 0.500 then Evadido (0 / 2 / 0 / 0)
76. if Sexo = Masculino then Reprovado (0 / 0 / 2 / 0)
77. else Sem_Acesso (0 / 0 / 0 / 1)

correct: 1528 out of 1564 training examples.

ANEXO 10

Regras de Classificação – Dados Históricos Mês 3

Foram geradas 58 regras ao todo, utilizando o algoritmo *RuleLearner*, implementado na ferramenta *RapidMiner*, como resultado da MD históricos de 1564 alunos, com 52 atributos de cada aluno, referentes ao mês três.

RuleModel

1. if G1 > 5.850 and AF_TM1 > 6.500 and AF_TM2 > 10.500 and Sexo = Masculino then Aprovado (141 / 0 / 0 / 0)
2. if G1 > 5.850 and G1 > 7.050 and AF_TM1 > 7.500 and Sexo = Feminino and AG_TM1 <= 44 then Aprovado (322 / 2 / 0 / 0)
3. if G1 > 5.550 and G1 > 7.350 and G1 > 8.050 and ER_TM3 <= 13 and AF_TM3 <= 15 and ER_TM1 <= 0.500 then Aprovado (73 / 0 / 0 / 0)
4. if G1 > 5.550 and G1 > 7.350 and ER_FM3 > 2.500 and AF_TM2 <= 17.500 and Sexo = Masculino and Área = EducaçãoCiênciaseArtes then Aprovado (31 / 0 / 0 / 0)
5. if G1 > 5.550 and G1 > 7.350 and PF_TM1 > 0.500 and AB_TM2 <= 34 and AB_TM3 <= 44 then Aprovado (145 / 0 / 0 / 0)
6. if G1 > 5.550 and Nota3 > 2.970 and FREQ_M2 > 3.500 and PF_FM3 > 0.500 then Aprovado (59 / 0 / 0 / 0)
7. if G1 > 5.550 and AF_TM3 <= 1.500 and Nota3 > 2.970 and AB_TM2 <= 82 and AG_TM1 > 7.500 then Aprovado (42 / 0 / 0 / 0)
8. if G1 > 5.550 and AB_FM1 <= 1.500 and AF_TM2 <= 12.500 and Nota3 > 3.025 then Aprovado (90 / 0 / 0 / 0)
9. if G1 > 3.950 and G1 > 5.550 and ER_TM3 <= 4.500 and AB_TM2 <= 82 and AB_FM1 > 2.500 and AB_TM2 <= 27.500 then Aprovado (35 / 0 / 0 / 0)
10. if G1 > 3.950 and G1 > 5.850 and AF_TM2 > 1.500 and PF_FM2 <= 0.500 and ER_FM2 > 3.500 and AF_TM2 <= 17.500 and AG_TM3 <= 23 then Aprovado (51 / 0 / 0 / 0)
11. if G1 > 3.750 and G1 > 5.550 and AB_TM3 > 10.500 and AF_TM1 > 1.500 and Idade > 24.500 and AG_TM2 > 4.500 then Aprovado (33 / 0 / 0 / 0)
12. if G1 > 3.750 and Nota3 > 3.080 and AB_TM2 <= 9.500 and ER_TM2 <= 11.500 and ER_TM1 > 5.500 then Aprovado (21 / 0 / 0 / 0)
13. if G1 > 3.750 and AB_TM1 <= 10.500 and AB_TM2 > 11.500 and AB_TM1 > 3.500 and ER_FM1 <= 2.500 and Idade > 17.500 then Aprovado (24 / 0 / 0 / 0)
14. if G1 > 3.750 and AG_TM3 > 4.500 and NA_t_G1 <= 2.500 and FREQ_M2 > 3.500 and AG_TM1 > 7.500 then Aprovado (15 / 0 / 0 / 0)
15. if G1 <= 3.150 and Pri_Ac > 1.500 and Sexo = Feminino and AG_TM1 <= 4.500 then Evadido (0 / 38 / 2 / 0)
16. if G1 > 3.750 and G1 > 5.550 and Nota3 <= 2.870 and Nota1 > 2.100 and AB_TM1 <= 8.500 and PF_TM3 <= 1.500 then Aprovado (24 / 0 / 0 / 0)
17. if G1 > 3.750 and Nota3 > 3.080 and AG_TM3 > 13.500 and PF_TM3 <= 3.500 then Aprovado (10 / 0 / 0 / 0)
18. if AG_TM3 <= 3.500 and AB_TM1 > 4.500 and Nota1 <= 1.625 and AG_TM3 <= 1.500 then Evadido (0 / 23 / 0 / 0)
19. if AF_TM2 > 2.500 and ER_TM3 > 5.500 and AB_TM2 <= 12.500 and ER_TM3 > 6.500 and EE_TM2 <= 0.500 then Aprovado (19 / 0 / 0 / 0)
20. if G1 > 3.750 and Idade <= 29.500 and ER_TM2 <= 4.500 and AF_FM3 > 1.500 and PF_TM2 <= 3.500 then Aprovado (13 / 0 / 0 / 0)
21. if PF_FM3 <= 0.500 and Pri_Ac > 1.500 and Área = EducaçãoCiênciaseArtes and AB_TM2 > 15 then Evadido (1 / 20 / 2 / 0)
22. if Nota1 > 1.175 and AB_TM2 > 13.500 and AG_TM1 <= 6.500 and Nota3 > 3.080 and AB_FM1 <= 3.500 then Aprovado (14 / 0 / 0 / 0)
23. if AB_TM3 <= 9.500 and Unidade = SantaMaria and ER_TM1 > 1 then Evadido (0 / 9 / 0 / 0)

24. if Nota1 > 1.175 and Idade <= 29.500 and Nota3 <= 0.250 and AB_TM2 > 9.500 and AF_TM3 <= 9 then Aprovado (17 / 0 / 0 / 0)
25. if AB_TM3 <= 9.500 and ER_TM2 > 10 and Entrega1 = Sim then Evadido (0 / 14 / 0 / 0)
26. if Nota3 > 1.535 and PF_FM3 > 0.500 and Idade <= 29 and AF_FM3 <= 2.500 and Nota2 > 0.815 and PF_TM3 <= 1.500 then Aprovado (13 / 0 / 0 / 0)
27. if ER_TM2 <= 4.500 and AB_TM2 > 9.500 and Nota3 <= 3.705 and FREQ_M3 > 1.500 and Idade > 21.500 then Evadido (0 / 14 / 0 / 0)
28. if ER_FM3 <= 2.500 and AB_TM2 > 3.500 and AB_TM2 <= 8.500 and Sexo = Masculino then Aprovado (14 / 0 / 0 / 0)
29. if FREQ_M2 <= 2.500 and Pri_Ac > 1.500 and AF_TM2 <= 1.500 and AG_TM3 <= 0.500 and Idade > 21.500 then Evadido (0 / 11 / 0 / 0)
30. if AF_TM3 <= 6.500 and AB_TM1 > 0.500 and FREQ_M2 <= 2.500 and AF_TM1 <= 1.500 and ER_FM1 > 3.500 and AB_FM1 <= 2.500 then Evadido (0 / 9 / 0 / 0)
31. if Nota3 > 1.535 and Entrega1 = Nao and ER_TM1 > 5.500 then Aprovado (7 / 1 / 0 / 0)
32. if G1 > 3.750 and AB_TM2 <= 6.500 and ER_TM1 > 11.500 then Aprovado (8 / 0 / 0 / 0)
33. if FREQ_M1 > 1.500 and Idade > 24.500 and AB_TM1 <= 15.500 and AB_TM2 > 9.500 and Sexo = Feminino then Evadido (0 / 10 / 0 / 0)
34. if Nota3 > 1.535 and ER_TM1 <= 3.500 and Nota2 <= 1.530 and G1 > 4.650 and AF_FM3 <= 2.500 then Aprovado (11 / 0 / 0 / 0)
35. if AB_TM3 <= 9.500 and AB_TM2 > 2.500 and Nota3 <= 2.100 and ER_TM2 > 3.500 and AB_TM3 > 3.500 and Idade <= 29.500 then Evadido (0 / 10 / 0 / 0)
36. if AB_FM3 > 1.500 and Sexo = Feminino and AG_TM3 > 7.500 then Aprovado (16 / 4 / 3 / 0)
37. if ER_FM1 <= 2.500 and Pri_Ac > 3.500 and Idade > 22.500 and Pri_Ac <= 5.500 then Evadido (0 / 8 / 0 / 0)
38. if AG_TM2 > 9.500 and AB_TM2 > 17.500 and PF_TM3 > 0.500 then Reprovado (0 / 0 / 9 / 0)
39. if AF_FM2 > 2.500 and Idade <= 23.500 then Evadido (0 / 8 / 0 / 0)
40. if AB_FM3 > 0.500 and AB_TM1 <= 18 and ER_TM1 > 8.500 and Área = TecnologiaeComputação and Nota3 <= 3.750 then Aprovado (7 / 0 / 0 / 0)
41. if AG_TM2 > 4.500 and AF_FM1 > 1.500 and AG_TM3 > 1.500 and AF_TM2 <= 3.500 then Reprovado (0 / 0 / 9 / 0)
42. if FREQ_M1 > 1.500 and Sexo = Masculino and AB_TM3 <= 7.500 then Evadido (5 / 17 / 3 / 0)
43. if AF_TM3 > 0.500 and ER_FM3 <= 2.500 and AF_TM3 <= 2.500 then Aprovado (8 / 0 / 0 / 0)
44. if AG_TM2 > 2.500 and Nota2 <= 1.815 and AB_TM2 > 17.500 then Reprovado (0 / 0 / 7 / 0)
45. if AB_TM3 > 4 and NAt_G1 > 2.500 and Idade > 24.500 then Aprovado (7 / 0 / 0 / 0)
46. if Nota1 > 0.125 and Nota3 <= 1.850 and AF_TM2 <= 7.500 and Pri_Ac <= 5.500 then Reprovado (0 / 0 / 8 / 0)
47. if Pri_Ac <= 0.500 then Sem_Acesso (0 / 0 / 0 / 13)
48. if AB_TM2 > 2.500 and Idade > 21.500 and AF_TM2 <= 12 then Evadido (0 / 9 / 0 / 0)
49. if AF_TM3 > 3.500 then Aprovado (4 / 0 / 0 / 0)
50. if AG_TM2 > 3.500 and Nota2 <= 0.750 then Reprovado (0 / 0 / 4 / 0)
51. if Idade > 22.500 and AF_TM2 > 1 then Aprovado (3 / 0 / 0 / 0)
52. if Curso = EDUCACAOFISICA then Reprovado (0 / 0 / 3 / 0)
53. if Entrega1 = Sim and Idade <= 23.500 then Evadido (0 / 3 / 0 / 0)
54. if Idade > 29.500 then Aprovado (2 / 0 / 0 / 0)
55. if Idade > 22.500 and Sexo = Masculino then Reprovado (0 / 0 / 2 / 0)
56. if Unidade = Canoas and AG_TM3 > 0.500 then Evadido (0 / 2 / 0 / 0)
57. if Unidade = Canoas then Aprovado (1 / 0 / 0 / 0)
58. else Sem_Acesso (0 / 0 / 0 / 1)

correct: 1541 out of 1564 training examples.

ANEXO 11

Regras de Classificação – Dados Históricos Mês 4

Foram geradas 56 regras ao todo, utilizando o algoritmo *RuleLearner*, implementado na ferramenta *RapidMiner*, como resultado da MD históricos de 1564 alunos, com 64 atributos de cada aluno, referentes ao mês quatro.

RuleModel

1. if G1 > 5.850 and AF_TM1 > 6.500 and AF_TM2 > 10.500 and Sexo = Masculino then Aprovado (141 / 0 / 0 / 0)
2. if G1 > 5.850 and G1 > 7.050 and AF_TM1 > 7.500 and Sexo = Feminino and AG_TM1 <= 44 then Aprovado (322 / 2 / 0 / 0)
3. if G1 > 5.550 and G1 > 7.350 and G1 > 8.050 and ER_TM3 <= 13 and AF_TM4 <= 14.500 and Sexo = Masculino then Aprovado (139 / 0 / 0 / 0)
4. if G1 > 5.550 and Nota3 > 2.970 and AB_TM2 <= 82.500 and PF_TM4 > 0.500 and ER_TM4 <= 4.500 then Aprovado (101 / 0 / 0 / 0)
5. if G1 > 5.550 and G1 > 7.350 and Nota1 <= 2.375 and AF_TM4 <= 6.500 and ER_FM3 > 2.500 and AB_TM2 <= 31.500 then Aprovado (97 / 0 / 0 / 0)
6. if G1 > 5.550 and ER_TM4 <= 3.500 and AB_TM3 > 4.500 and AB_TM2 <= 43.500 and Nota1 <= 2.375 then Aprovado (60 / 0 / 0 / 0)
7. if G1 > 5.550 and AF_TM1 > 2.500 and AB_TM3 > 12.500 and AF_TM4 <= 6 then Aprovado (42 / 0 / 0 / 0)
8. if G1 > 3.950 and G1 > 5.850 and AG_TM2 > 14.500 and AB_TM4 > 28.500 then Aprovado (31 / 0 / 0 / 0)
9. if G1 > 3.950 and G1 > 6.250 and AB_TM2 <= 6.500 and ER_TM2 <= 12.500 and Idade <= 28.500 then Aprovado (47 / 0 / 0 / 0)
10. if G1 > 3.750 and Nota3 > 3.030 and Unidade = Canoas and Nota3 <= 4.085 then Aprovado (58 / 5 / 0 / 0)
11. if G1 > 3.750 and AG_TM4 > 3.500 and AB_TM1 <= 26.500 and FREQ_M1 > 2.500 and AB_TM1 <= 12.500 and AB_TM4 <= 18 and AB_TM2 > 3.500 then Aprovado (30 / 0 / 0 / 0)
12. if G1 > 3.750 and AG_TM4 > 3.500 and AB_TM1 <= 26.500 and Nat_G1 <= 2.500 and AF_TM2 > 3.500 and AG_TM3 <= 11 then Aprovado (17 / 0 / 0 / 0)
13. if G1 <= 3.150 and AF_TM4 <= 0.500 and AB_TM1 > 0.500 and AB_TM1 > 4.500 and Nota3 <= 2.250 then Evadido (0 / 41 / 0 / 0)
14. if G1 > 4.150 and Nota1 > 2.400 and AB_FM2 > 2.500 and AF_TM2 <= 9.500 and Idade <= 30.500 and Nota2 <= 2.125 and ER_TM2 <= 25.500 then Aprovado (35 / 0 / 0 / 0)
15. if AG_TM4 > 4.500 and PF_TM4 > 0.500 and AG_TM3 > 5.500 and AF_TM3 <= 18.500 and AG_TM4 > 11 then Aprovado (20 / 0 / 0 / 0)
16. if AG_TM4 <= 4.500 and G1 <= 4.950 and Pri_Ac > 1.500 and FREQ_M4 <= 0.500 and AG_TM3 > 1.500 then Evadido (0 / 17 / 0 / 0)
17. if AG_TM4 <= 4.500 and G1 <= 3.850 and Pri_Ac > 1.500 and AG_TM3 <= 0.500 and Sexo = Feminino then Evadido (0 / 15 / 0 / 0)
18. if G1 > 3.250 and AB_TM1 <= 3.500 and ER_FM1 > 2.500 and Idade > 24 and ER_TM3 <= 11.500 then Aprovado (16 / 0 / 0 / 0)
19. if AG_TM4 <= 4.500 and AB_TM1 > 2.500 and ER_TM1 <= 2.500 and AB_TM1 > 8.500 then Evadido (0 / 18 / 0 / 0)
20. if Nota3 > 1.750 and Nota3 > 4.185 and G1 <= 6.250 then Aprovado (13 / 0 / 0 / 0)
21. if G1 > 5.550 and AF_TM2 > 1.500 and AG_TM1 > 6.500 and AB_FM4 <= 3.500 then Aprovado (17 / 0 / 0 / 0)
22. if FREQ_M4 <= 2.500 and Idade <= 31.500 and AG_TM2 <= 4.500 and Pri_Ac > 1.500 and AF_TM2 <= 1.500 and Nota2 <= 0.250 and AG_TM2 > 0.500 then Evadido (0 / 14 / 0 / 0)
23. if AB_TM4 > 0.500 and Nota2 > 1.555 and AB_TM1 <= 11.500 and ER_FM3 <= 2.500 and AB_FM1 <= 1.500 and

- AB_TM2 > 3.500 then Aprovado (16 / 0 / 0 / 0)
24. if Pri_Ac > 2.500 and Nota2 > 0.725 and Nota3 <= 2.110 and AG_TM3 <= 3.500 and Idade > 25.500 then Evadido (0 / 9 / 0 / 0)
 25. if FREQ_M4 <= 2.500 and AB_TM1 > 0.500 and AG_TM2 <= 2.500 and AF_TM1 <= 3.500 and ER_FM2 > 3.500 then Evadido (0 / 14 / 0 / 0)
 26. if AG_TM4 > 7.500 and G1 <= 3.950 and Unidade = Canoas then Aprovado (14 / 0 / 0 / 0)
 27. if G1 > 3.650 and AF_TM4 <= 1.500 and AG_TM1 > 4.500 and AB_FM4 <= 1.500 then Evadido (0 / 10 / 0 / 0)
 28. if Nota1 > 1.375 and AG_TM2 <= 10.500 and Sexo = Masculino and AG_TM2 > 7.500 then Aprovado (14 / 1 / 1 / 0)
 29. if G1 > 5.150 and AF_TM3 > 1.500 and Nota3 > 2.640 and Idade > 23.500 then Evadido (0 / 8 / 0 / 0)
 30. if Idade <= 25.500 and Nota3 > 2.870 and AF_TM2 <= 0.500 then Evadido (0 / 7 / 0 / 0)
 31. if ER_TM1 > 8.500 and Nota2 > 1.625 and Idade <= 22.500 and AF_TM3 <= 1.500 then Aprovado (9 / 0 / 0 / 0)
 32. if AG_TM1 > 8.500 and Nota1 > 2.125 and AF_FM3 <= 1.500 then Reprovado (0 / 0 / 8 / 0)
 33. if Idade <= 24.500 and Nota2 > 1.875 and ER_TM1 > 1 and AG_TM3 <= 21 then Evadido (0 / 10 / 0 / 0)
 34. if AB_TM3 > 5.500 and Idade <= 32.500 and Idade > 22.500 and AF_FM3 > 2.500 and AF_TM1 <= 31 then Aprovado (8 / 0 / 0 / 0)
 35. if AG_TM2 > 2.500 and ER_TM1 <= 8.500 and Nota1 <= 1.125 and Nota3 > 1.155 and Nota3 <= 3.945 then Reprovado (0 / 0 / 6 / 0)
 36. if AF_TM1 > 5.500 and AB_TM1 <= 13.500 and G1 > 3.650 then Evadido (0 / 7 / 0 / 0)
 37. if G1 > 0.900 and AG_TM2 <= 4.500 and ER_TM4 <= 3.500 then Aprovado (9 / 0 / 0 / 0)
 38. if Entrega2 = Sim and Curso = EDUCACAOFISICA then Reprovado (2 / 2 / 9 / 0)
 39. if ER_TM1 <= 3.500 and Nota3 <= 2.450 and AG_TM1 > 2 and AB_TM1 > 2.500 then Evadido (0 / 9 / 0 / 0)
 40. if AG_TM3 > 12.500 and AF_TM3 <= 11 then Reprovado (0 / 0 / 7 / 0)
 41. if AB_TM3 > 5.500 and Sexo = Feminino and AG_TM3 > 6.500 then Aprovado (10 / 0 / 0 / 0)
 42. if AB_TM2 > 2.500 and Nota3 <= 2.450 and AB_TM1 <= 0.500 and AG_TM4 <= 5.500 then Evadido (0 / 7 / 0 / 0)
 43. if ER_TM1 <= 8.500 and AG_TM2 > 1.500 and G1 <= 5.600 and AB_TM2 <= 11 then Reprovado (0 / 0 / 10 / 0)
 44. if AB_FM4 > 0.500 and AF_TM4 <= 2.500 and AB_TM3 > 5 then Aprovado (7 / 0 / 0 / 0)
 45. if FREQ_M1 > 2.500 and Idade <= 26 and AG_TM1 > 4 then Evadido (0 / 8 / 0 / 0)
 46. if Pri_Ac <= 0.500 then Sem_Acesso (0 / 0 / 0 / 13)
 47. if AB_TM2 > 21.500 and PF_TM4 <= 0.500 then Reprovado (0 / 0 / 5 / 0)
 48. if AB_FM4 > 0.500 and AB_TM3 <= 3.500 and Pri_Ac <= 3.500 then Aprovado (7 / 0 / 0 / 0)
 49. if Unidade = Carazinho then Evadido (0 / 3 / 0 / 0)
 50. if AF_FM1 > 1.500 and Nota2 > 0.625 then Reprovado (0 / 0 / 3 / 0)
 51. if Área = CiênciasSociaisAplicadas then Evadido (0 / 3 / 0 / 0)
 52. if AB_TM2 > 0.500 then Aprovado (4 / 0 / 0 / 0)
 53. if Unidade = Canoas and ER_TM2 > 5 then Reprovado (0 / 0 / 2 / 0)
 54. if Sexo = Masculino then Evadido (0 / 2 / 0 / 0)
 55. if Unidade = Canoas then Reprovado (0 / 0 / 1 / 0)
 56. else Sem_Acesso (0 / 0 / 0 / 1)

correct: 1551 out of 1564 training examples.

ANEXO 12

Regras de Classificação – Dados Históricos Mês 5

Foram geradas 9 regras ao todo, utilizando o algoritmo *RuleLearner*, implementado na ferramenta *RapidMiner*, como resultado da MD históricos de 1564 alunos, com 87 atributos de cada aluno, referentes ao mês cinco.

RuleModel

1. if MF > 5.800 then Aprovado (1286 / 0 / 0 / 0)
2. if MF <= 3.850 and Pri_Ac > 0.500 and S2 <= 3.650 and ER_TM5 > 8.500 then Evadido (0 / 94 / 0 / 0)
3. if S2 <= 3.100 and S1 <= 1.200 and Pri_Ac > 0.500 and Nota2 <= 2.375 and ER_TM5 <= 7.500 then Evadido (0 / 88 / 0 / 0)
4. if AG_TM5 > 6.500 and S1 > 1.200 then Reprovado (0 / 0 / 18 / 0)
5. if S2 > 1.600 and AB_TM4 <= 29 then Reprovado (0 / 0 / 27 / 0)
6. if AG_TM5 > 2.500 and S1 <= 2.150 and AG_TM5 <= 20 then Evadido (0 / 27 / 0 / 0)
7. if NA_t_G2 <= 0.500 then Sem_Acesso (0 / 0 / 0 / 14)
8. if Nota5 <= 1.065 then Reprovado (0 / 0 / 7 / 0)
9. else Evadido (0 / 3 / 0 / 0)

correct: 1564 out of 1564 training examples.

ANEXO 13

Encaminhamento da Solicitação de Dados Atualizados à TI

Projeto de Mineração de Dados Educacionais

2ª Etapa

Problema: A partir das regras obtidas pela mineração de dados históricos (já obtidos na 1ª etapa do projeto), na etapa atual serão analisadas as características e os comportamentos, em suas variações temporais, dos alunos de quatro turmas em andamento, buscando identificar tendências de evasão e reprovação para que o professor intervenha antes da confirmação dessas tendências, ou seja, que possa reverter esses quadros e contribuir para o engajamento e aprovação dos alunos.

Dados Necessários:

- Semestre
- Turma
- Id Aluno
- Nome
- Unidade Presencial
- Curso
- Sexo
- Idade
- Notas finais – G1, G2, S, MF, status (aprovado, reprovado, cancelado...)
- Quantidade de acessos ao Ambiente (Sala Virtual) por semana
- Quantidade de acessos aos Materiais de Aula por semana
- Quantidade de acessos à Biblioteca Virtual por semana
- Quantidade de acessos aos Fóruns por semana
- Quantidade de postagens nos Fóruns por semana
- Quantidade de e-mails recebidos por semana
- Quantidade de e-mails enviados por semana
- Avaliações entregues – nota, datas (liberação, prazo e entrega) e grau que compõe (de cada uma).

Formato - planilhas Excel, por turma, no mesmo formato da 1ª Etapa:

(ou em planilha única, conforme anexo, com as informações até o período de tempo do semestre indicado, destacado na tabela de intervalos de tempo por semana)

Planilha1_TurmaX (Dados Gerais)

Semestre / Turma / IdAluno / Nome / UnidadePresencial / Curso / Sexo / Idade / G1 / G2 / S / MF / Status

Planilha2_TurmaX (Quantidade de Acesso à Sala Virtual por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha3_TurmaX (Quantidade de Acesso aos Materiais de Aula por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha4_TurmaX (Quantidade de Acesso à Biblioteca Virtual por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha5_TurmaX (Quantidade de Acesso aos Fóruns por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha6_TurmaX (Quantidade de Postagens nos Fóruns por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha7_TurmaX (Quantidade de E-mails recebidos por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha8_TurmaX (Quantidade de E-mails enviados por semana)

Semestre / Turma / IdAluno / S1 / S2/ S3 / S4 / S5 / S6 / S7 / S8 / S9 / S10 / S11 / S12 / S13 / S14 / S15 / S16/ S17 / S18 / S19

Planilha9_TurmaX (Detalhamento das Avaliações Entregues)

Semestre / Turma / IdAluno / Nota1 / DataEntrega1 / GrauQueCompõe1 ... NotaN / DataEntregaN / GrauQueCompõeN

Turmas Alvo (Disciplina de Instrumentalização Científica):

- **Semestre 2009/1**
 - 2918, 2919, 2978, 2986

Intervalos de Tempo - Semanas:

Semanas / Semestre	2009.1
S1	04/03/2009 a 10/03/2009
S2	11/03/2009 a 17/03/2009
S3	18/03/2009 a 24/03/2009
S4	25/03/2009 a 31/03/2009
S5	01/04/2009 a 07/04/2009
S6	08/04/2009 a 14/04/2009
S7	15/04/2009 a 21/04/2009
S8	22/04/2009 a 28/04/2009
S9	29/04/2009 a 05/05/2009
S10	06/05/2009 a 12/05/2009
S11	13/05/2009 a 19/05/2009
S12	20/05/2009 a 26/05/2009
S13	27/05/2009 a 02/06/2009
S14	03/06/2009 a 09/06/2009
S15	10/06/2009 a 16/06/2009
S16	17/06/2009 a 23/06/2009
S17	24/06/2009 a 30/06/2009
S18	01/07/2009 a 07/07/2009
S19	08/07/2009 a 14/07/2009

Enviar os dados nas semanas 4, 8, 11, 15 e 19.

ANEXO 14

Estatística dos Dados Históricos x Dados de Acompanhamento

Anexo 14 – Tabela 1 – Estatística dos Dados Históricos x Dados de Acompanhamento

Atributos			Históricos		Acompanhamento	
Tipo	Nome	Valor Tipo	Estatística	Faixa	Estatística	Faixa
regular	Nota1	real	avg = 1.957 +/- 0.764	[0.000 ; 2.500]	avg = 2.148 +/- 0.607	[0.000 ; 2.500]
regular	Nota2	real	avg = 1.571 +/- 0.721	[0.000 ; 2.500]	avg = 1.922 +/- 0.653	[0.000 ; 2.500]
regular	Nota3	real	avg = 3.261 +/- 1.529	[0.000 ; 5.000]	avg = 3.123 +/- 1.414	[0.000 ; 5.000]
regular	NAt_G1	integer	avg = 2.641 +/- 0.802	[0.000 ; 3.000]	avg = 2.739 +/- 0.627	[0.000 ; 3.000]
regular	Entrega1	nominal	mode = Sim (1397)	Sim (1397), Nao (143), Atraso (10)	mode = Sim (217)	Sim (217), Nao (13)
regular	Entrega2	nominal	mode = Sim (1351)	Sim (1351), Nao (180), Atraso (19)	mode = Sim (209)	Sim (209), Nao (21)
regular	Entrega3	nominal	mode = Sim (1324)	Sim (1324), Nao (207), Atraso (19)	mode = Sim (188)	Sim (188), Nao (28), Atraso (14)
regular	Nota4	real	avg = 0.660 +/- 0.335	[0.000 ; 1.000]	avg = 0.812 +/- 0.291	[0.000 ; 1.000]
regular	Nota5	real	avg = 1.268 +/- 0.675	[0.000 ; 2.000]	avg = 1.309 +/- 0.543	[0.000 ; 2.000]
regular	Nota6	real	avg = 3.935 +/- 1.763	[0.000 ; 7.000]	avg = 3.663 +/- 1.388	[0.000 ; 7.000]
regular	NAt_G2	integer	avg = 2.566 +/- 0.952	[0.000 ; 3.000]	avg = 2.765 +/- 0.726	[0.000 ; 3.000]
regular	Entrega4	nominal	mode = Sim (1287)	Sim (1287), Nao (247), Atraso (16)	mode = Sim (208)	Sim (208), Nao (17), Atraso (5)
regular	Entrega5	nominal	mode = Sim (1272)	Sim (1272), Nao (255), Atraso (23)	mode = Sim (208)	Sim (208), Nao (17), Atraso (5)
regular	Pri_Ac	integer	avg = 1.953 +/- 1.423	[1.000 ; 11.000]	avg = 2.265 +/- 1.406	[1.000 ; 8.000]
regular	Ult_Ac	integer	avg = 17.508 +/- 2.453	[1.000 ; 19.000]	avg = 18.026 +/- 1.208	[11.000 ; 19.000]
regular	AG_TM1	integer	avg = 9.923 +/- 9.302	[0.000 ; 94.000]	avg = 8.013 +/- 7.722	[0.000 ; 52.000]
regular	FREQ_M1	integer	avg = 2.714 +/- 1.260	[0.000 ; 4.000]	avg = 2.396 +/- 1.210	[0.000 ; 4.000]
regular	AG_TM2	integer	avg = 13.220 +/- 9.911	[0.000 ; 109.000]	avg = 12.635 +/- 10.506	[0.000 ; 92.000]
regular	FREQ_M2	integer	avg = 3.259 +/- 0.981	[0.000 ; 4.000]	avg = 3.048 +/- 1.014	[0.000 ; 4.000]
regular	AG_TM3	integer	avg = 8.778 +/- 7.648	[0.000 ; 83.000]	avg = 10.565 +/- 9.326	[0.000 ; 78.000]
regular	FREQ_M3	integer	avg = 2.339 +/- 0.860	[0.000 ; 3.000]	avg = 2.530 +/- 0.677	[0.000 ; 3.000]
regular	AG_TM4	integer	avg = 8.495 +/- 8.090	[0.000 ; 102.000]	avg = 8.652 +/- 8.843	[0.000 ; 65.000]
regular	FREQ_M4	integer	avg = 2.708 +/- 1.249	[0.000 ; 4.000]	avg = 2.765 +/- 1.078	[0.000 ; 4.000]
regular	AG_TM5	integer	avg = 10.595 +/- 9.320	[0.000 ; 89.000]	avg = 12.487 +/- 10.620	[0.000 ; 84.000]
regular	FREQ_M5	integer	avg = 2.690 +/- 1.123	[0.000 ; 4.000]	avg = 2.830 +/- 0.943	[0.000 ; 4.000]
regular	AB_TM1	integer	avg = 19.064 +/- 23.794	[0.000 ; 245.000]	avg = 13.313 +/- 14.655	[0.000 ; 78.000]
regular	AB_FM1	integer	avg = 2.416 +/- 1.329	[0.000 ; 4.000]	avg = 1.791 +/- 1.187	[0.000 ; 4.000]
regular	AB_TM2	integer	avg = 24.542 +/- 31.992	[0.000 ; 418.000]	avg = 19.117 +/- 18.934	[0.000 ; 161.000]
regular	AB_FM2	integer	avg = 2.986 +/- 1.179	[0.000 ; 4.000]	avg = 2.413 +/- 1.034	[0.000 ; 4.000]
regular	AB_TM3	integer	avg = 14.363 +/- 18.137	[0.000 ; 284.000]	avg = 10.826 +/- 14.593	[0.000 ; 112.000]
regular	AB_FM3	integer	avg = 2.069 +/- 1.011	[0.000 ; 3.000]	avg = 1.404 +/- 0.922	[0.000 ; 3.000]
regular	AB_TM4	integer	avg = 16.036 +/- 21.601	[0.000 ; 260.000]	avg = 12.800 +/- 14.812	[0.000 ; 141.000]
regular	AB_FM4	integer	avg = 2.532 +/- 1.332	[0.000 ; 4.000]	avg = 1.791 +/- 1.237	[0.000 ; 4.000]
regular	AB_TM5	integer	avg = 15.897 +/- 18.632	[0.000 ; 236.000]	avg = 12.943 +/- 14.947	[0.000 ; 63.000]
regular	AB_FM5	integer	avg = 2.429 +/- 1.179	[0.000 ; 4.000]	avg = 1.387 +/- 0.924	[0.000 ; 4.000]

regular	AF_TM1	integer	avg = 11.522 +/- 15.668	[0.000 ; 185.000]	avg = 5.722 +/- 7.900	[0.000 ; 43.000]
regular	AF_FM1	integer	avg = 2.063 +/- 1.402	[0.000 ; 4.000]	avg = 1.161 +/- 1.105	[0.000 ; 4.000]
regular	AF_TM2	integer	avg = 11.128 +/- 16.292	[0.000 ; 243.000]	avg = 7.313 +/- 16.336	[0.000 ; 170.000]
regular	AF_FM2	integer	avg = 2.167 +/- 1.366	[0.000 ; 4.000]	avg = 1.265 +/- 1.210	[0.000 ; 4.000]
regular	AF_TM3	integer	avg = 6.214 +/- 10.561	[0.000 ; 132.000]	avg = 4.083 +/- 8.236	[0.000 ; 88.000]
regular	AF_FM3	integer	avg = 1.393 +/- 1.124	[0.000 ; 3.000]	avg = 0.804 +/- 0.938	[0.000 ; 3.000]
regular	AF_TM4	integer	avg = 7.521 +/- 11.764	[0.000 ; 164.000]	avg = 13.839 +/- 24.660	[0.000 ; 262.000]
regular	AF_FM4	integer	avg = 1.725 +/- 1.416	[0.000 ; 4.000]	avg = 1.713 +/- 1.144	[0.000 ; 4.000]
regular	AF_TM5	integer	avg = 3.468 +/- 6.348	[0.000 ; 84.000]	avg = 7.887 +/- 9.587	[0.000 ; 71.000]
regular	AF_FM5	integer	avg = 1.138 +/- 1.130	[0.000 ; 4.000]	avg = 1.370 +/- 0.941	[0.000 ; 4.000]
regular	PF_TM1	integer	avg = 1.368 +/- 1.713	[0.000 ; 14.000]	avg = 0.987 +/- 0.872	[0.000 ; 5.000]
regular	PF_FM1	integer	avg = 0.941 +/- 0.954	[0.000 ; 4.000]	avg = 0.843 +/- 0.641	[0.000 ; 3.000]
regular	PF_TM2	integer	avg = 1.104 +/- 1.947	[0.000 ; 33.000]	avg = 1.391 +/- 1.128	[0.000 ; 9.000]
regular	PF_FM2	integer	avg = 0.677 +/- 0.856	[0.000 ; 4.000]	avg = 1.157 +/- 0.723	[0.000 ; 3.000]
regular	PF_TM3	integer	avg = 0.566 +/- 1.126	[0.000 ; 12.000]	avg = 0.687 +/- 0.795	[0.000 ; 5.000]
regular	PF_FM3	integer	avg = 0.507 +/- 0.744	[0.000 ; 3.000]	avg = 0.609 +/- 0.621	[0.000 ; 3.000]
regular	PF_TM4	integer	avg = 0.665 +/- 1.079	[0.000 ; 9.000]	avg = 0.530 +/- 0.630	[0.000 ; 2.000]
regular	PF_FM4	integer	avg = 0.502 +/- 0.706	[0.000 ; 4.000]	avg = 0.509 +/- 0.595	[0.000 ; 2.000]
regular	PF_TM5	integer	avg = 0.190 +/- 0.667	[0.000 ; 8.000]	avg = 0.443 +/- 0.539	[0.000 ; 2.000]
regular	PF_FM5	integer	avg = 0.128 +/- 0.393	[0.000 ; 3.000]	avg = 0.422 +/- 0.494	[0.000 ; 1.000]
regular	EE_TM1	integer	avg = 0.054 +/- 0.297	[0.000 ; 5.000]	avg = 0.065 +/- 0.336	[0.000 ; 4.000]
regular	EE_FM1	integer	avg = 0.044 +/- 0.220	[0.000 ; 2.000]	avg = 0.061 +/- 0.289	[0.000 ; 3.000]
regular	EE_TM2	integer	avg = 0.125 +/- 0.613	[0.000 ; 11.000]	avg = 0.096 +/- 0.361	[0.000 ; 2.000]
regular	EE_FM2	integer	avg = 0.088 +/- 0.346	[0.000 ; 3.000]	avg = 0.074 +/- 0.262	[0.000 ; 1.000]
regular	EE_TM3	integer	avg = 0.128 +/- 0.582	[0.000 ; 9.000]	avg = 0.187 +/- 0.663	[0.000 ; 6.000]
regular	EE_FM3	integer	avg = 0.092 +/- 0.355	[0.000 ; 3.000]	avg = 0.130 +/- 0.396	[0.000 ; 2.000]
regular	EE_TM4	integer	avg = 0.048 +/- 0.286	[0.000 ; 4.000]	avg = 0.096 +/- 0.437	[0.000 ; 4.000]
regular	EE_FM4	integer	avg = 0.041 +/- 0.228	[0.000 ; 3.000]	avg = 0.074 +/- 0.307	[0.000 ; 2.000]
regular	EE_TM5	integer	avg = 0.121 +/- 0.535	[0.000 ; 9.000]	avg = 0.191 +/- 0.566	[0.000 ; 4.000]
regular	EE_FM5	integer	avg = 0.092 +/- 0.356	[0.000 ; 3.000]	avg = 0.174 +/- 0.480	[0.000 ; 3.000]
regular	ER_TM1	integer	avg = 5.995 +/- 5.239	[0.000 ; 19.000]	avg = 5.948 +/- 2.846	[0.000 ; 14.000]
regular	ER_FM1	integer	avg = 2.255 +/- 1.664	[0.000 ; 4.000]	avg = 2.335 +/- 0.883	[0.000 ; 4.000]
regular	ER_TM2	integer	avg = 6.488 +/- 5.715	[0.000 ; 32.000]	avg = 7.539 +/- 1.098	[0.000 ; 12.000]
regular	ER_FM2	integer	avg = 2.974 +/- 1.687	[0.000 ; 4.000]	avg = 3.974 +/- 0.279	[0.000 ; 4.000]
regular	ER_TM3	integer	avg = 4.704 +/- 3.738	[0.000 ; 26.000]	avg = 6.030 +/- 1.519	[4.000 ; 17.000]
regular	ER_FM3	integer	avg = 2.065 +/- 1.219	[0.000 ; 3.000]	avg = 3 +/- 0	[3.000 ; 3.000]
regular	ER_TM4	integer	avg = 3.168 +/- 2.056	[0.000 ; 9.000]	avg = 6.517 +/- 1.137	[5.000 ; 11.000]
regular	ER_FM4	integer	avg = 2.097 +/- 1.275	[0.000 ; 4.000]	avg = 3 +/- 0	[3.000 ; 3.000]
regular	ER_TM5	integer	avg = 7.554 +/- 4.916	[0.000 ; 20.000]	avg = 13.400 +/- 2.435	[10.000 ; 23.000]
regular	ER_FM5	integer	avg = 3.001 +/- 1.688	[0.000 ; 4.000]	avg = 4 +/- 0	[4.000 ; 4.000]
regular	G1	real	avg = 6.810 +/- 2.483	[0.000 ; 10.000]	avg = 7.206 +/- 2.081	[0.000 ; 10.000]
regular	G2	real	avg = 5.864 +/- 2.425	[0.000 ; 10.000]	avg = 5.783 +/- 1.961	[0.000 ; 10.000]
regular	S1	real	avg = 0.263 +/- 1.259	[0.000 ; 9.000]	avg = 0.383 +/- 1.529	[0.000 ; 8.400]
regular	S2	real	avg = 0.557 +/- 1.807	[0.000 ; 9.500]	avg = 0.503 +/- 1.709	[0.000 ; 8.300]
regular	MF	real	avg = 6.382 +/- 2.119	[0.000 ; 9.900]	avg = 6.461 +/- 1.571	[0.700 ; 9.700]

ANEXO 15

Formulário de Satisfação com os Alertas - Respondidos pelos Alunos

Termo de Consentimento

A professora Adriana Justin Cerveira Kampff, orientada pelos professores José Valdeni de Lima e Eliseo Berni Reategui, está realizando uma pesquisa, junto ao Ensino a Distância da Universidade Luterana do Brasil, sobre possibilidades de intervenção junto aos alunos, que possam ajudá-los a organizar seus estudos nesta modalidade. Para tanto, durante o semestre, os alunos foram acompanhados e receberam mensagens para auxiliá-los no engajamento e na realização das atividades propostas.

Neste momento, gostaríamos que respondessem as questões apresentadas neste formulário, se forma a colher sua percepção sobre o Ambiente Virtual utilizado, a organização das aulas da disciplina, a sua autonomia e o auxílio das mensagens recebidas.

Desde já agradecemos sua colaboração.

Nome:

Aceito participar da pesquisa: Sim / Não

Tela inicial do Formulário com Orientações



The screenshot shows a web browser window with the following content:

- Browser title: Pesquisa em EAD - Windows Internet Explorer
- Address bar: <http://spreadsheets.google.com/viewform?formkey=dHIRWDFidE4eFRr>
- Page title: Pesquisa em EAD
- Form content:
 - Pesquisa em EAD**
 - A professora Adriana Justin Cerveira Kampff, orientada pelos professores Jose Valdeni de Lima e Eliseo Berni Reategui, esta realizando uma pesquisa, junto ao Ensino a Distancia da Universidade Luterana do Brasil, sobre possibilidades de intervencao junto aos alunos, que possam ajuda-los a organizar seus estudos nesta modalidade.
 - Para tanto, durante o semestre, os alunos foram acompanhados, na disciplina de Instrumentalizacao Cientifica, e receberam mensagens para auxilia-los no engajamento e na realizacao das atividades propostas.
 - Neste momento, gostaríamos que respondessem as questoes apresentadas neste formulario, de forma a colher sua percepcao sobre o Ambiente Virtual utilizado, a organizacao das aulas da disciplina, a sua autonomia e o auxilio das mensagens recebidas, dedicando em torno de 5 a 10 minutos para essa importante tarefa.
 - Desde ja agradecemos sua colaboracao, fundamental para a validacao da pesquisa!
 - * Required
 - Nome Completo: *
(apenas para controle de respondentes, não será utilizado para análise da pesquisa)
 -

Anexo 15 – Figura 1 – Pesquisa junto aos alunos

ANEXO 16

Publicações durante o Doutorado

Artigos completos publicados em periódicos

1. LOH, Stanley; LICHTNOW, Daniel; KAMPFF, Adriana Justin Cerveira; OLIVEIRA, José Palazzo Moreira de. **Recommendation of Complementary Material during Chat Discussions.** Knowledge Management & E-Learning: An International Journal (KM&EL), 2009. *(Aceito para Publicação)*
2. JELINEK, K. R.; KAMPFF, Adriana Justin Cerveira. **A Geometria que existe além do olhar: levando a Geometria da natureza para dentro da escola.** Educação Matemática em Revista, 2009. *(Aceito para Publicação)*
3. PORTELA, Tanise; KAMPFF, Adriana Justin Cerveira. **Jogos Educacionais: Interação Apoiada por Agentes Afetivos.** RENOTE. Revista Novas Tecnologias na Educação, 2009. *(Aceito para Publicação)*
4. KAMPFF, Adriana Justin Cerveira; REATEGUI, Eliseo; LIMA, José Valdeni de. **Mineração de dados educacionais para a construção de alertas em ambientes virtuais de aprendizagem, como apoio a prática docente.** RENOTE. Revista Novas Tecnologias na Educação, v. 6, p. 1, 2008.
5. KAMPFF, Adriana Justin Cerveira. **Inclusão digital: iniciativas potencializadoras.** Caderno Marista de Educação, 2008.
6. KAMPFF, Adriana Justin Cerveira; OLIVEIRA, Andréia; MIRANDA, Roxane Leandra dos Santos; ROSO, Mafalda. **Aprendizagem Colaborativa Em Construção Coletiva Multimídia.** RENOTE. Revista Novas Tecnologias na Educação, v. 5, p. 1-9, 2007.
7. NICOLAO, M.; BERG, A.; CABRAL, Anderson Ricardo Yanzer; KAMPFF, Adriana Justin Cerveira; ZEVE, C. M. D. **Gestão do Conhecimento em Ambientes de Ensino a Distância nas Universidades.** Logos (Canoas), v. 1, p. 51-61, 2007.
8. LICHTNOW, Daniel; LOH, Stanley; KAMPFF, Adriana Justin Cerveira; PRIMO, Tiago; GARIN, Ramiro Saldaña; OLIVEIRA, José Palazzo Moreira de; LIMA, José Valdeni de. **O Uso de Técnicas de Recomendação em um Sistema para Apoio à Aprendizagem Colaborativa.** Revista Brasileira de Informática na Educação, v. 14, p. 49-59, 2006.
9. KONRATH, Mary Lúcia Pedroso; KAMPFF, Adriana Justin Cerveira; GOMES, Fábio de Jesus Lima; CARVALHO, Marie Jane Soares; NEVADO, Rosane Aragón de. **Nós no Mundo: Objeto de Aprendizagem voltado para o 1º Ciclo do Ensino Fundamental.** RENOTE. Revista Novas Tecnologias na Educação, Porto Alegre, v. 4, n. 1, p. 1-10, 2006.
10. KAMPFF, Adriana Justin Cerveira; FERREIRA, André Luis Andrejew; CARVALHO, Maria Helena Sório de; LIMA, José Valdeni de. **Um Hiperdocumento para Introdução à Geometria Plana.** RENOTE. Revista Novas Tecnologias na Educação, Porto Alegre, v. 3, n. 2, p. 1-10, 2005.
11. KAMPFF, Adriana Justin Cerveira; LIRA, Antônio da Fonseca de; REITZ, Doris Simone; GOMES, Fábio de Jesus Lima; FONSECA, Luís Carlos Costa; MACHADO, Nádie Christina Ferreira. **Relação entre o Perfil do Usuário e a Escolha do Perfil do Tutor.** RENOTE. Revista Novas Tecnologias na Educação, Porto Alegre, v. 3, n. 1, p. 1-10, 2005.
12. KAMPFF, Adriana Justin Cerveira; MACHADO, José Carlos; CAVEDINI, Patrícia. **Novas Tecnologias e Educação Matemática.** RENOTE. Revista Novas Tecnologias na Educação, Porto Alegre, v. 2, n. 2, p. 1-10, 2004.

Livros publicados/organizados ou edições

1. KAMPPFF, Adriana Justin Cerveira. **Tecnologia da Informática e Comunicação na Educação**. Curitiba: IESDE Brasil S.A., 2006.

Capítulos de livros publicados

1. KAMPPFF, Adriana Justin Cerveira. **Computadores e educação**. In: ULBRA. (Org.). Tecnologias da informação e da comunicação na educação. Curitiba: IBPEX, 2007, v., p. 31-45.
2. KAMPPFF, Adriana Justin Cerveira. **Softwares educacionais**. In: ULBRA. (Org.). Tecnologias da informação e da comunicação na educação. Curitiba: IBPEX, 2007, v., p. 47-62.
3. KAMPPFF, Adriana Justin Cerveira. **Internet**. In: ULBRA. (Org.). Tecnologias da informação e da comunicação na educação. Curitiba: IBPEX, 2007, v., p. 97-113.
4. KAMPPFF, Adriana Justin Cerveira. **Comunicação virtual**. In: ULBRA. (Org.). Tecnologias da informação e da comunicação na educação. Curitiba: IBPEX, 2007, v., p. 115-131.
5. KAMPPFF, Adriana Justin Cerveira. **Comunidades virtuais de aprendizagem**. In: ULBRA. (Org.). Tecnologias da informação e da comunicação na educação. Curitiba: ULBRA, 2007, v., p. 133-147.
6. KONRATH, Mary Lúcia Pedroso; KAMPPFF, Adriana Justin Cerveira; GOMES, Fábio de Jesus Lima; HAETINGER, D.; RELA, E.; GELATTI, L. S.; KIST, T.; CARVALHO, Marie Jane Soares; NEVADO, Rosane Aragón de; CARNEIRO, M. L. F.. **Objetos de Aprendizagem, Formação de Professores e Práticas Pedagógicas no Contexto Escolar das Séries Iniciais**. In: Rosane Aragón de Nevado; Marie Jane Soares de Carvalho; Crediné Silva de Menezes. (Org.). Aprendizagem em rede na educação a distância: estudos e recursos para a formação de professores. 1 ed. Porto Alegre: Ricardo Lenz, 2007, v. 1, p. 131-149.

Trabalhos completos publicados em anais de congressos

1. NICOLAO, M. ; CABRAL, Anderson Ricardo Yanzer ; ZEVE, C. M. D. ; KAMPPFF, Adriana Justin Cerveira ; BERG, A. . Persistência do Conhecimento em Ambientes de Ensino a Distância. In: VII SEMINFO - Seminário de Informática RS'2008, 2008, Torres. Anais do VII Seminário de Informática - RS. Torres : Ulbra, 2008.
2. JELINEK, K. R.; KAMPPFF, Adriana Justin Cerveira. **O Resgate da Geometria no Ensino Fundamental e a Valorização da Construção do Pensamento Geométrico**. In: IV Congresso Internacional de Ensino da Matemática, 2007, Canoas. Anais do IV Congresso Internacional de Ensino da Matemática, 2007.
3. KONRATH, Mary Lúcia Pedroso; KAMPPFF, Adriana Justin Cerveira; NEVADO, Rosane Aragón de; CARVALHO, Marie Jane Soares. **Nós no Mundo: Objeto de Aprendizagem Interativo**. In: Simpósio Brasileiro de Informática na Educação, 2006, Brasília. Anais do SBIE 2006. 2006. p. 1-10.
4. LOW, Stanley; LICHTNOW, Daniel; KAMPPFF, Adriana Justin Cerveira; OLIVEIRA, José Palazzo Moreira de. **Sistema de Recomendações em Discussões Eletrônicas - Apoio à Construção do Conhecimento em Comunidades Virtuais de Aprendizagem**. In: Congresso da Sociedade Brasileira de Computação - Workshop de Informática na Escola, 2005, São Leopoldo. Anais do XXV Congresso da Sociedade Brasileira de Computação, 2005. p. 2504-2513.
5. KAMPPFF, Adriana Justin Cerveira; DIAS, Márcia Gládis Cantelli. **A Construção do Conhecimento em Ambientes de Pesquisa e de Autoria Multimídia: Uma Tarefa Compartilhada por Alunos e Professores**. In: XXIV Congresso da Sociedade Brasileira de Computação - Workshop de Informática na Escola, 2004, Salvador. Anais do SBC 2004, 2004. p. 655-663.

6. KAMPPFF, Adriana Justin Cerveira; DIAS, Márcia Gládis Cantelli. **A Construção do Conhecimento em Ambientes de Pesquisa e de Autoria Multimídia: Uma Tarefa Compartilhada por Alunos e Professores.** In: Congresso da Rede Ibero-Americana de Informática Educativa, 2004, Monterrey. Anais do RBIE, 2004. p. 1-10.
7. KAMPPFF, Adriana Justin Cerveira; CAVEDINI, Patrícia. **Ambientes informatizados de aprendizagem Matemática: O Estudo da geometria no ensino fundamental.** In: Congresso da Rede Ibero-Americana de Informática na Educação, 2004, Monterrey. Anais do RIBIE, 2004. p. 1-10.
8. KAMPPFF, Adriana Justin Cerveira; CAVEDINI, Patrícia; MACHADO, José Carlos. **Novas Tecnologias e Educação Matemática.** In: XXIV Congresso da Sociedade Brasileira de Computação - Workshop de Informática na Escola, 2004, Salvador. Anais do SBC 2004, 2004. p. 555-563.

Resumos expandidos publicados em anais de congressos

1. KAMPPFF, Adriana Justin Cerveira; REATEGUI, Eliseo; LIMA, José Valdeni de. **Concepções para a construção de alertas em ambientes virtuais de aprendizagem, como apoio a prática docente, em busca de mediações mais efetivas.** In: Simpósio Brasileiro de Informática na Educação, 2008, Fortaleza. Anais do Simpósio Brasileiro de Informática na Educação, 2008.
2. KAMPPFF, Adriana Justin Cerveira; OLIVEIRA, Andréia; MIRANDA, Roxane Leandra dos Santos; ROSO, Mafalda. **Collective Development of Multimedia Books: Cooperation as an Educational Approach.** In: ED-MEDIA 2007-World Conference on Educational Multimedia, Hypermedia & Telecommunications, 2007, Vancouver. ED-MEDIA 2007-World Conference on Educational Multimedia, Hypermedia & Telecommunications. Vancouver: EdITLib - Digital Library for Information Technology and Education, 2007.
3. KAMPPFF, Adriana Justin Cerveira; MIRANDA, Roxane Leandra dos Santos; NAZARIO, Silvana Reali. **Construção Coletiva de Livros Multimídia: Uma Abordagem para Temas Transversais.** In: Simpósio Brasileiro de Informática na Educação, 2006, Brasília. Anais do SBIE 2006, 2006. p. 1-3.
4. KAMPPFF, Adriana Justin Cerveira; FERREIRA, André Luis Andrejew; COSTA, Íris Elisabeth Tempel; LIMA, José Valdeni de; BEHAR, Patrícia Alejandra. **Uma Oficina Virtual de Aprendizagem para Sensibilização de Professores de Matemática no Uso de Tecnologias Educacionais.** In: XVI Simpósio Brasileira de Informática na Educação, 2005, Juiz de Fora. Anais do XVI Simpósio Brasileiro de Informática na Educação, 2005.