

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE CIÊNCIAS ECONÔMICAS  
PROGRAMA DE PÓS-GRADUAÇÃO DE MESTRADO EM ECONOMIA**

**OLAVO SEVERO GUIMARÃES**

**PREVISÃO DE *DEFAULT* DA DÍVIDA PÚBLICA:  
UMA APLICAÇÃO DE *MACHINE LEARNING***

**PORTO ALEGRE**

**2021**

**OLAVO SEVERO GUIMARÃES**

**PREVISÃO DE *DEFAULT* DA DÍVIDA PÚBLICA:  
UMA APLICAÇÃO DE *MACHINE LEARNING***

Dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Ciências Econômicas da UFRGS como requisito parcial para a conclusão do curso de Mestrado Profissional em Economia

Orientador: Prof. Dr. Sergio Marley Modesto Monteiro

**PORTO ALEGRE**

**2021**

### CIP - Catalogação na Publicação

Guimarães, Olavo Severo  
Previsão de default da dívida pública: uma  
aplicação de machine learning / Olavo Severo  
Guimarães. -- 2021.  
88 f.  
Orientador: Sergio Marley Modesto Monteiro.

Dissertação (Mestrado) -- Universidade Federal do  
Rio Grande do Sul, Faculdade de Ciências Econômicas,  
Programa de Pós-Graduação em Economia, Porto Alegre,  
BR-RS, 2021.

1. Machine learning. 2. Árvores decisórias. 3.  
Sinais de crise. 4. Crises moratórias. I. Monteiro,  
Sergio Marley Modesto, orient. II. Título.

**OLAVO SEVERO GUIMARÃES**

**PREVISÃO DE *DEFAULT* DA DÍVIDA PÚBLICA:  
UMA APLICAÇÃO DE *MACHINE LEARNING***

Dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Ciências Econômicas da UFRGS como requisito parcial para a conclusão do curso de Mestrado Profissional em Economia

Orientador: Prof. Dr. Sergio Marley Modesto Monteiro

Aprovado em: Porto Alegre, \_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

---

Presidente: Prof. Dr. Sérgio Marley Monteiro – Orientador  
UFRGS

---

Prof. Dr. Carlos Eduardo Schönerwald da Silva  
UFRGS

---

Prof.<sup>a</sup> Dra. Luiza Peruffo  
UFRGS

## RESUMO

Este trabalho utilizou um método de *machine learning* para auxiliar a previsão de *defaults* da dívida pública de 66 países, detectando sinais prévios. Por conta de sua transparência, o algoritmo C4.5 de árvores de classificação foi adotado. Tal algoritmo obteve êxito na previsão das crises da década de 1980 e do início da década de 1990, registrando, a partir daí, perda na capacidade de aprendizado para prever as crises das duas décadas seguintes. Limitações do próprio algoritmo e mudanças estruturais da economia internacional na década de 1990, com diminuição do número de crises moratórias, podem ter contribuído para este resultado. As árvores decisórias obtidas com o método também foram exploradas a fim de se avaliar a própria lógica algorítmica, tentando traçar sua relação com a teoria econômica atinente, que também foi exposta neste trabalho, e com fatos da história econômica mundial. Em termos de variáveis preditivas, o algoritmo corroborou a existência de “*serial defaulters*”, posto que o histórico de *default* dos países quase sempre figurou como importante sugestivo de crises futuras. Outras variáveis que se mostraram importantes para prever crises foram aquelas relacionadas com as reservas internacionais – principalmente a variação das reservas e a taxa da dívida a curto prazo sobre as reservas. Isto demonstra o forte nexos existente entre crises moratórias e cambiais, sobretudo nos *defaults* da década de 1980, que o algoritmo conseguiu prever. Por fim, os juros globais e a variação do PIB também se mostraram variáveis relevantes. Conclui-se que, apesar de suas limitações, o algoritmo pode contribuir para o estudo de crises moratórias por conta de sua transparência e capacidade de estabelecer relações condicionais e não lineares.

**Palavras-chave:** Machine learning. Árvores decisórias. Sinais de crise. Crises moratórias.

## ABSTRACT

This thesis has used a machine learning method to help forecast 66 countries' sovereign debt defaults by detecting early warning signals. Due to its transparency, the C4.5 classification trees algorithm was chosen. This algorithm has succeeded while trying to forecast sovereign debt defaults in the 1980's decade and in the beginning of the 1990's decade, having decayed, after that, its learning capacity to forecast the two following decade's crises. The algorithm's own limitations and structural changes in the international economy in the 1990 decade, with debt crises reduction, may have contributed to this result. The decision trees obtained with the method were also explored to assess the own algorithm rationale, trying to trace its relation with the pertaining economic theory, which was also exposed in this thesis, and with the world economic history facts. Regarding predicting variables, the algorithm has corroborated the existence of "serial defaulters" since the countries' default history was always an important suggestive of future crises. Other important variables were those related to international reserves - mainly the reserves variation and the short-term debt ratio over reserves. This demonstrates the strong nexus among debt and currency crises, especially regarding 1980's decade defaults, which the algorithm has succeeded in forecasting. Finally, global rates and GDP variation were shown as relevant variables. It is concluded that, despite its limitations, the algorithm may contribute to the study of debt crises due to its transparency and capacity for establishing conditional and non-linear relations.

**Key-Words:** Machine learning. Decision trees. Early warning signals. Sovereign debt crises.

## LISTA DE FIGURAS

Figura 1 – Processo de descoberta de conhecimento em bases de dados .....	31
Figura 2 – Curva ROC e AUC.....	36
Figura 3 – Exemplo de árvore de classificação .....	39
Figura 4 – Gráfico – Relação entre probabilidade e entropia.....	40
Figura 5 – AUC x Ano – Modelo Global .....	48
Figura 6 – Árvore gerada para previsão de crises do ano de 1979 – Modelo Global .....	50
Figura 7 – Árvore gerada para previsão de crises do ano de 1979 – Modelo Global .....	52
Figura 8 – Árvore gerada para previsão de crises do ano de 1979 – Modelo Global .....	53
Figura 9 – Árvore gerada para previsão de crises do ano de 1982 – Modelo Global .....	56
Figura 10 – AUC x Ano – Modelo América Latina.....	59
Figura 11 – Árvore gerada para previsão de crises do ano de 1990 – Modelo Global ....	60
Figura 12 – Árvore gerada para previsão de crises do ano de 1987 – Modelo América Latina .....	61
Figura 13 – Árvore gerada para previsão de crises do ano de 2000 – Modelo América Latina.....	62
Figura 14 – Gráfico – Importância das variáveis para as previsões – Modelo América Latina.....	63
Figura 15 – AUC x Ano – Modelo África.....	65
Figura 16 – Árvore gerada para previsão de crises do ano de 1980 – Modelo África .....	66
Figura 17 – Árvore gerada para previsão de crises do ano de 1985 – Modelo África .....	67
Figura 18 – Distribuição dos registros de <i>default</i> na África de 1975 a 1984 e a relação dívida pública/PIB .....	68
Figura 19 – Gráfico – Importância das variáveis para as previsões – Modelo África.....	69
Figura 20 – AUC x Ano – Modelo Ásia .....	71
Figura 21 – Árvore gerada para previsão de crises do ano de 1989 – Modelo Ásia.....	72
Figura 22 – Árvore gerada para previsão de crises do ano de 1997 – Modelo Ásia .....	73
Figura 23 – Gráfico – Importância das variáveis para as previsões – Modelo Ásia.....	75

## LISTA DE TABELAS

<b>Tabela 1 – Relação de episódios de <i>default</i> da base de dados .....</b>	<b>43</b>
<b>Tabela 2 – Lista das variáveis da base de dados utilizada .....</b>	<b>45</b>
<b>Tabela 3 – Resultados do Modelo Global.....</b>	<b>47</b>
<b>Tabela 4 – Resultados do Modelo América Latina.....</b>	<b>57</b>
<b>Tabela 5 – Resultados do Modelo África.....</b>	<b>64</b>
<b>Tabela 6 – Resultados do Modelo Ásia.....</b>	<b>70</b>



## LISTA DE ABREVIATURAS E SIGLAS

EWS	<i>Early Warning Signal</i>
FMI	Fundo Monetário Internacional
PIB	Produto Interno Bruto
ROC Curve	<i>Receiver Operating Characteristic Curve</i>
AUC	<i>Area Under the Curve</i>

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>11</b>
<b>2 CRISES MORATÓRIAS: UMA REVISÃO DA LITERATURA</b> .....	<b>13</b>
2.1 REVISÃO DA LITERATURA TEÓRICA.....	13
2.1.1 Profecias autorrealizadas .....	13
2.1.2 Modelos sudden-stop .....	14
2.1.3 Excesso de endividamento.....	17
2.1.4 Crescimento econômico.....	17
2.1.5 O papel dos juros.....	18
2.1.6 Intolerância a dívidas e pecado original .....	18
2.1.7 Dívidas de curto prazo .....	19
2.1.8 Mudança de disposição .....	20
2.1.9 Considerações .....	21
2.2 INDICADORES DE CRISE - EWS: REVISÃO DA LITERATURA EMPÍRICA.....	21
2.2.1 modelos de limites.....	22
2.2.2 Modelos logit.....	24
2.2.3 Modelos classificadores de <i>machine learning</i> .....	26
2.2.4 Heterogeneidades temporais e espaciais .....	28
2.2.5 Considerações .....	29
<b>3 METODOLOGIA</b> .....	<b>31</b>
3.1 MINERAÇÃO DE DADOS E <i>MACHINE LEARNING</i> .....	31
3.1.1 Treinamento e teste do algoritmo .....	34
3.1.2 Métricas de avaliação de modelos de classificação.....	35
3.2 ALGORITMO E BASE DE DADOS UTILIZADOS .....	37
3.2.1 A plataforma weka .....	37
3.2.2 Balanceamento de classes.....	37
3.2.3 O algoritmo de árvores de classificação c4.5 .....	38
3.2.4 Descrição da base de dados .....	42
<b>4 ANÁLISE DOS RESULTADOS</b> .....	<b>46</b>
4.1 MODELO GLOBAL.....	46

4.2 MODELO AMÉRICA LATINA .....	57
4.3 MODELO ÁFRICA .....	63
4.4 MODELO ÁSIA .....	69
<b>5 CONCLUSÃO .....</b>	<b>75</b>
<b>REFERÊNCIAS .....</b>	<b>79</b>
<b>ANEXO A.....</b>	<b>85</b>
<b>ANEXO B.....</b>	<b>86</b>
<b>ANEXO C.....</b>	<b>87</b>
<b>ANEXO D.....</b>	<b>88</b>

## 1 INTRODUÇÃO

O excesso de endividamento público e as crises moratórias, embora sejam tão antigas quanto a própria existência dos empréstimos, passaram a ser temas de maior preocupação à medida que mercados financeiros se desenvolveram e se integraram internacionalmente. Neste sentido, o final da década de 70 foi especialmente traumático, pois marcou a transição de anos de crédito barato para o chamado “choque Volcker”, em que o banco central norte-americano pressionaria, por anos, as taxas de juros globais, com efeitos catastróficos para países com endividamento externo, os quais muitas vezes deixaram de cumprir seus compromissos financeiros. Nesta época se consolida o campo de estudo denominado de *country (sovereign) risk analysis* - risco-país em nossa tradução - sendo os preditores de crises moratórias usualmente chamados de *early warning signals* (EWS), os quais são objetos de estudo do presente trabalho.

Tal literatura acumula aplicações de métodos econométricos, em especial modelos logit, para a detecção de sinais de defaults. Posteriormente, os modelos de limites foram desenvolvidos no âmbito do Fundo Monetário Internacional (FMI), mas focaram-se, sobretudo, em crises cambiais. Embora tais modelos tenham méritos, métodos de *machine learning* se apresentam como novas alternativas para se tentar detectar os sinais mais importantes de defaults e aprimorar a previsão de tais crises.

Algoritmos de *machine learning* são programas que, a partir de regras estatísticas, criam seus próprios programas para a resolução de problemas. Sendo uma espécie de inteligência artificial, eles têm tido crescente aplicação em variadas indústrias e áreas de pesquisa, sendo problemas de classificação e previsão uma de suas áreas de aplicação mais proeminente.

Neste contexto, este trabalho busca entender como tais algoritmos podem auxiliar na previsão de defaults da dívida pública e seus sinais. Escolheu-se, para tal, o uso do algoritmo de árvores de classificação C4.5, em razão de sua transparência e facilidade de interpretação, que o distingue de outros métodos mais potentes, mas ao mesmo tempo mais opacos.

É premissa deste trabalho que algoritmos de *machine learning*, embora sejam potentes e promissoras ferramentas, não devem substituir o julgamento humano, sobretudo em questões complexas e estratégicas. Trata-se, sim, de um instrumento complementar. Assim, faz-se fundamental que o problema investigado seja também estudado à luz da literatura científica a ele atinente. Portanto, realizou-se neste trabalho, inicialmente, um apanhado do

que a literatura econômica teórica e empírica dizem sobre crises moratórias. Na sequência, apresenta-se a área de *machine learning*, detalhando-se o método adotado com a descrição do algoritmo e da base de dados utilizados. Por fim, os resultados obtidos com as árvores decisórias são apresentados. A partir de um método exploratório, aproveitou-se da transparência do algoritmo para se analisar o seu raciocínio à luz da literatura econômica estudada e dos fatos históricos. Por fim, as conclusões cabíveis são expostas.

## 2 CRISES MORATÓRIAS: UMA REVISÃO DA LITERATURA

Crises de endividamento nacional são tão antigas quanto os próprios empréstimos. Na antiguidade europeia, a maioria das crises fiscais foi resolvida por meio do aumento da base monetária. Episódios de inadimplência e renegociação das dívidas passam a aparecer mais na história ocidental no século XVI, havendo registros na França, Espanha e Portugal. Porém, é a partir do século XVIII que as crises da dívida, as renegociações e as inadimplências aumentaram tanto em termos quantitativos, quanto geográficos. Isto se deve principalmente ao desenvolvimento dos mercados financeiros, ao crescente fluxo de capitais internacionais e à emergência de novos Estados soberanos (Sturzenegger, Zettelmeyer, 2006).

Como afirma Dani Rodrik (2011, p.8), “estas crises ocorreram não porque elas eram imprevisíveis, mas porque elas não foram previstas”. De fato, diversos economistas, analistas e acadêmicos conseguiram prever que um país incorreria em crise moratória em algum momento, mas detectá-las com maior precisão segue sendo um desafio. Ademais, mesmo quem previu estas crises atribui diferentes causas imediatas e básicas a elas.

É oportuno observar assim que o estudo do *default* da dívida soberana abriga componentes de inferência causal e de predição. Enquanto a primeira tenta explicar o porquê das crises, a segunda se limita a prevê-las. Por certo, o estudo de um fornece argumentos e *insights* para o estudo do outro, mas é preciso diferenciá-los. Esta revisão de literatura está organizada de modo a cobrir primeiro uma vertente de trabalhos teóricos, com viés explicativo, para em seguida analisar uma vertente de trabalhos empíricos, com foco majoritariamente preditivo.

### 2.1 REVISÃO DA LITERATURA TEÓRICA

#### 2.1.1 Profecias autorrealizadas

Uma primeira e simples explicação para episódios de crises moratórias é a de uma profecia autorrealizada. Imagine-se que há uma dívida nacional de curto prazo. A moratória será decretada se o governo não conseguir rolar a dívida, isto é, pegar novos empréstimos para pagar os anteriores. Ocorre que, se um financiador prevê um episódio deste gênero, o melhor para ele é não emprestar. Se vários financiadores passam a concordar com isto, ocorre uma crise moratória, ainda que o julgamento deles esteja errado (SACHS, 1983; COLE, KEHOE, 2000).

Isto ajuda a explicar os perigos das dívidas de curto prazo, mais atreladas a países não desenvolvidos. Países desenvolvidos têm mais facilidade para emitir dívidas de longo prazo em sua própria moeda, o que rompe tal ciclo deletério se o país conseguir emitir uma dívida com prazo suficiente para superar depressões econômicas advindas da crise moratória (COLE, KEHOE, 2000; STURZENEGGER, ZETTELMEYER, 2006). Esta diferença na estrutura das dívidas de diferentes nações é uma das possíveis causas de *default* e será mais bem abordada em outra seção.

Uma restrição da teoria de profecia autorrealizada, contudo, é que há, mesmo em países emergentes, mecanismos contratuais e informais para solucionar o problema de coordenação descrito (STURZENEGGER, ZETTELMEYER, 2006), o que sugere que ela não é tão realista a ponto de explicar as referidas crises em toda sua complexidade.

### 2.1.2 Modelos sudden-stop

Um ditado de banqueiros diz “*It is not speed that kills, it is the sudden stop*”<sup>1</sup>. Isto nos leva a uma segunda explicação, que pode ser encontrada nos chamados modelos *sudden-stop* (parada súbita). Estes modelos, introduzidos nos anos 1990, tentam explicar os efeitos de reduções repentinas no influxo de capital estrangeiro nos países, o que tipicamente ocorre quando investidores estrangeiros param de investir nestes mercados (BOONMAN, 2015). Um dos possíveis efeitos destes eventos seriam crises moratórias.

Note-se que estes modelos, em geral, já tomam as paradas súbitas como dadas – isto é, exógenas ao modelo. Não obstante, exploraremos algumas explicações fornecidas pela literatura para estes eventos, a fim de ilustrar as causas das crises moratórias em diferentes profundidades – isto é, tanto causas últimas e aparentes, quanto causas primárias e fundamentais.

Em relação àquelas explicações apontadas para a redução do influxo de investimentos incluem aumentos repentinos nas taxas de juros internacionais, incerteza política e diminuição da produtividade (STURZENEGGER, ZETTELMEYER, 2006). Os aumentos nas taxas de juros em grandes centros financeiros alterariam a percepção de risco-retorno dos investidores, os quais, alterando suas carteiras, gerariam alterações nos fluxos de capitais.

De modo semelhante, aumentos da incerteza política e econômica tanto local quanto mundial podem ter o efeito de estimular investidores a buscarem países mais seguros para

---

<sup>1</sup> Não é a velocidade que mata, mas a parada súbita (tradução nossa). Citado originalmente por Dornbusch, Goldfajn e Valdés (1995, p.1).

investir, ensejando *sudden-stops*. Em termos de fatores econômicos, é importante ressaltar a associação existente entre tais eventos e recorrentes déficits em conta corrente, conforme demonstram Cavallo, Eichengreen e Panizza (2018)<sup>2</sup>.

Por fim, como contratos de empréstimo são firmados em unidades de bens transacionáveis, mas frequentemente dependem da renda de bens não transacionáveis, choques de produtividade podem gerar dificuldades creditícias e afugentar capitais (MENDONZA, 2002).

Há também versões de crises autorrealizáveis causadas por investidores internacionais, em que uma desaceleração exógena da entrada de capitais leva uma empresa a falir, destruindo canais de crédito e capital humano e gerando uma espiral depressiva, o que por fim valida previsões pessimistas e diminui mais ainda o influxo de investimento, podendo gerar até fugas de capitais (CALVO, 1998).

Por sua vez, explicações mais sistêmicas e profundas envolvem teorias sobre ciclos financeiros que enxergam o capitalismo como um processo instável de produção, a exemplo da tradição minskyana, que tem embasado abordagens recentes relacionando vulnerabilidades externas com tais ciclos<sup>3</sup>. Para esta vertente, as variações no fluxo de capitais são causadas pela própria natureza frágil do sistema capitalista atual, que não oferece robustez mesmo contra pequenos choques, gerando uma instabilidade permanente que penaliza, sobretudo, os países em desenvolvimento. Estes seriam desfavorecidos pelo atual sistema financeiro internacional, que não lhes permite endividamento em moeda própria, de modo a torná-los mais dependentes das flutuações cambiais.

Assim, para economistas minskyanos, as instabilidades nos fluxos de capitais (e suas consequências) não seriam causadas pelos (maus) fundamentos das economias subdesenvolvidas, como descontrole fiscal e histórico de inflação, mas sim por variações nos ciclos financeiros internacionais. Aumentos nas taxas de juros internacionais e da incerteza política seriam, por conseguinte, reações a sucessivas etapas do ciclo – este sim, a causa primeira das variações no nível de tais investimentos. Neste contexto, a solução para amenizar as causas imediatas seria uma maior regulação dos fluxos de capitais, em especial os de curta-duração, mais sujeitos a movimentos especulativos, de modo a atenuar os efeitos dos ciclos financeiros (Kaltenbrunner e Paineira, 2015).

---

<sup>2</sup> Nesse sentido, Reinhart e Rogoff (2009b) argumentam que prolongados déficits em conta corrente são importantes indicativos de crises financeiras.

<sup>3</sup> Ver, por exemplo, Paineira e Kaltenbrunner (2015).



A teoria de Minsky, segundo o próprio autor, segue a vertente da instabilidade econômica, na tradição de Marx, Keynes, Schumpeter e Kalecki; a qual contrasta com teorias como as de Friedman e Lucas, que argumentam a existência de um equilíbrio no crescimento econômico que eventualmente é perturbado por mudanças na oferta monetária (Ferri e Minsky, 1992). De maneira análoga, as soluções de controles de capitais são motivos de controvérsia, entre diferentes vertentes, quanto ao seu sucesso e efetividade<sup>4</sup>.

Conforme exposto, o consenso sobre o que gera as reduções de influxo ainda não se formou. Interessa-nos, ainda assim, as suas consequências. Duas delas têm sido amplamente documentadas: a redução no produto e a depreciação da moeda (Catão, 2006). Partindo delas, diversos desdobramentos são possíveis. Examinemos o processo mais detalhadamente.

As reduções de influxo de capitais estrangeiros gerariam uma redução na demanda agregada (o que reduziria também o produto e a arrecadação fiscal do governo), bem como uma depreciação da moeda. Se a dívida for constituída ao menos em parte em moeda estrangeira, a depreciação dificultará o pagamento dela, podendo levar à moratória em casos extremos. Ademais, a própria baixa da arrecadação também pode aumentar os juros que o mercado cobra, o que, por sua vez, também dificulta a rolagem da dívida.

Se a taxa de câmbio for fixa e o governo decidir mantê-la estável, haverá queda nas reservas internacionais, a qual, se severa, pode incentivar especulações ou gerar uma crise no balanço de pagamentos (BOONMAN, 2015). Alternativamente, o governo poderia aumentar os juros para atrair capitais; isto, contudo, tende a deprimir o produto e aumentar o serviço da dívida.

Como se pode notar, *sudden-stops* gerariam crises moratórias através de (ameaças de) crises cambiais e seus desdobramentos, estando também associadas a recorrentes déficits em conta corrente. De fato, as crises moratórias frequentemente estão relacionadas a outras crises, como bancárias ou cambiais (REINHART, 2002)<sup>5</sup>. Note-se que, mesmo sendo a dívida pública apenas parcialmente em moeda estrangeira, a desestabilização por conta de uma crise cambial pode vir a ocorrer: o *default* da dívida em moeda estrangeira pode dificultar o

---

<sup>4</sup> Em meio a tais discussões, cabe ressaltar a visão institucional do FMI (2012), no sentido de que não há uma solução única para a questão da liberalização e do controle de fluxo de capitais, sendo necessário sopesar diversos aspectos, sobretudo o estágio de desenvolvimento institucional e financeiro dos países destinatários dos fluxos. Para uma maior discussão, ver: KOKENYNE, Annamaria; BABA, Chikako. **Effectiveness of capital controls in selected emerging markets in the 2000s**. 2011; MAGUD, Nicolas E.; REINHART, Carmen M.; ROGOFF, Kenneth S. **Capital controls: myth and reality-a portfolio balance approach**. National Bureau of Economic Research, 2011; EDWARDS, Sebastian. Capital controls, sudden stops, and current account reversals. In: **capital controls and capital flows in Emerging Economies: policies, practices, and consequences**. University of Chicago Press, 2007. p. 73-120.

<sup>5</sup>Kindleberger e Aliber (2011) oferecem uma boa perspectiva histórica da relação entre crises moratórias, cambiais e *crashes* de forma geral.

rolamento da dívida em moeda nacional. Neste contexto, se o pagamento da dívida interna acarretar inflação, isto tende a depreciar ainda mais a moeda no mercado internacional (STURZENEGGER, ZETTELMAYER, 2006). Registre-se, enfim, que abordamos *sudden-stops* neste item, mas situações semelhantes e ainda mais graves são geradas pela fuga de capitais, em que os capitais não só param de entrar, como também passam a sair do país.

### 2.1.3 Excesso de endividamento

As situações acima descritas tratam de choques externos – por trás delas há a suposição de que, em tempos normais, a dívida poderia ser paga. Contudo, a causa da inadimplência pode ser simplesmente que os governos tomam mais dinheiro emprestado do que podem pagar. A explicação para este fenômeno costuma envolver risco moral, no sentido de que os responsáveis por aumentar a dívida não são os mesmos cidadãos que lidam com os custos dela (STURZENEGGER, ZETTELMAYER, 2006).

Sob a ótica do excesso de endividamento, as aparentes profecias autorrealizáveis poderiam ser, na verdade, consequências da distribuição das probabilidades de *default* em relação ao endividamento, conforme argumenta Bi (2012):

[...]o prêmio pelo risco de *default*, refletindo a probabilidade de *default* nacional, é uma função não linear do nível de endividamento do governo. O prêmio de risco começa a aumentar quando o nível de endividamento do governo se aproxima da extremidade inferior da distribuição dos limites fiscais[...] (Tradução nossa)<sup>6</sup>.

Assim, mesmo a altos níveis de endividamento, seria possível manter juros baixos, mas à medida que o limite fiscal vai se expandindo, haveria um momento em que os juros subiriam rapidamente, dificultando a rolagem da dívida e causando crises moratórias.

### 2.1.4 Crescimento econômico

De qualquer forma, é certo que o endividamento não pode ser analisado separadamente do crescimento econômico. De fato, hoje a principal relação abordada quando se fala em dívida nacional é a da relação desta com o Produto Interno Bruto (PIB). Esta é uma

---

<sup>6</sup> “(...) the default risk premium, reflecting the probability of sovereign default, is a nonlinear function of the level of the government debt. The risk premium begins to emerge as the level of government debt approaches the lower end of the distribution of fiscal limits (...)” (BI, 2012, p. 4).

medida que é usada na literatura tanto como explicação dos *defaults* por excesso de endividamento quanto como variável preditiva. Os escritos de Evsey Domar (1944) foram seminais no sentido de analisar o endividamento nacional sob tal ótica.

Tal autor ponderou, no contexto do pós-guerra, que a dívida em si não era um problema tão grave, desde que houvesse um crescimento suficiente da renda nacional. Para que a dívida fosse sustentável, seria necessário que a relação dívida/PIB se mantivesse constante no longo prazo. Assim, a própria dívida poderia crescer indefinidamente, desde que dentro dos limites impostos pelo crescimento da economia.

### **2.1.5 O papel dos juros**

Pasinetti (2005), por sua vez, revisitou os estudos de Domar atentando para o fato de que a relação dívida/PIB depende fundamentalmente da relação entre a taxa de crescimento econômico e a taxa de juros. Neste sentido, o autor aponta que houve uma mudança radical a partir da década de 1970, no contexto das crises do petróleo e dos aumentos de preços então observados: os juros passaram a ser tratados, pelos bancos centrais, como ferramenta de combate à inflação.

Deste momento em diante, endividamentos que, com taxas menores de juros, não causavam maiores problemas nas décadas de capitalismo keynesiano de 1950 e 1960, passaram a suscitar maiores preocupações fiscais. Assim, o referido autor considera que a mudança de paradigma da teoria econômica em relação aos juros é uma importante causa das crises moratórias que se observaram nas décadas de 1980 e 1990.

### **2.1.6 Intolerância a dívidas e pecado original**

Outra ferramenta teórica usada para explicar as crises moratórias reside no conceito de intolerância a dívidas (*debt intolerance*). A expressão foi cunhada por Reinhart, Rogoff e Savastano (2003) para designar a dificuldade que alguns países têm de administrar níveis de dívida que seriam manejáveis por países desenvolvidos. Ou seja, um endividamento ser excessivo ou não depende, além do nível da dívida em relação ao PIB, da condição do país endividado e sua intolerância à dívida, a qual, segundo os autores, depende da história da nação mutuária, em especial seu histórico de *defaults* e de inflação.

Outro conceito parecido, mas diferente, que ajuda a explicar as diferentes condições que os países apresentam para lidar com suas dívidas é o de pecado original (*original sin*). Este conceito, criado por Eichengreen e Hausmann (1999), refere-se à dificuldade que alguns países têm de contrair dívidas em sua própria moeda. Contudo, diferente da intolerância à dívida, o pecado original não foi atribuído a características nacionais, como boas instituições e responsabilidade fiscal. A única característica que se mostrou estatisticamente relevante nos estudos de Eichengreen, B., Hausmann, R. e Panizza, U.(2003) foi o tamanho do país – países menores têm mais dificuldade de emitir dívida em moeda própria.

Segundo estes defensores do pecado original, a raiz do problema está na própria estrutura dos portfólios globais e dos mercados financeiros internacionais (EICHENGREEN, HAUSMANN, 1999). Ao mesmo tempo em que diversificar o investimento em diversas moedas dilui o risco, em um mundo com custos de transação, há um nível ótimo de diversificação. Isto implica a existência de um limite no número de moedas a serem consideradas pelos investidores, gerando um cenário em que a adoção de uma moeda diminui as chances de adoção da outra. Neste contexto, países com economias maiores e com mais emissão de moeda seriam mais atraentes para terem suas moedas adotadas pelos portfólios globais.

### **2.1.7 Dívidas de curto prazo**

Conforme mencionado anteriormente, outra condição do endividamento que é agravante, segundo a literatura, é a de uma dívida excessivamente composta de títulos de curto prazo, pois ela poderia abrir as portas para crises autorrealizáveis decorrentes de mudanças de expectativas dos agentes (BRONER, LORENZONI, SCHMUKLER, 2013; COLE e KEHOE, 2000; STURZENEGGER e ZETTELMEYER, 2006). Por outro lado, uma vertente da literatura enxerga dívidas de curto prazo como dispositivos de compromisso prévio que ajudam a dividir o risco entre mutuantes e mutuários ou fornecem aos emprestadores maior controle sobre a ação dos devedores, reduzindo o risco de *default* (Rodrik, 1999). Uma explicação alternativa está atrelada ao custo da dívida: países tomam créditos de curto prazo porque os investidores exigem um prêmio maior pelo alto risco de deter dívidas de longo prazo, fazendo com que empréstimos de vencimento mais breve sejam mais baratos (BRONER, LORENZONI, SCHMUKLER, 2013).

Perfis de endividamento de menor maturidade costumam ser associados a países emergentes (COLE, KEHOE, 2000; STURZENEGGER, ZETTELMEYER, 2006). Contudo,

há controvérsia a respeito de tal suposição: um estudo de 2018 do FMI concluiu que não há diferença na maturidade da dívida soberana em relação a países de diferentes níveis de renda. O estudo, todavia, documenta que a maturidade média do débito corporativo é de fato menor nos países emergentes (FMI, 2019).

Outros achados relevantes do referido estudo em relação à duração da dívida é que este prazo tende a diminuir junto com a redução da aversão ao risco global ou dos juros norte-americanos, contudo estes fatores explicariam apenas uma pequena porcentagem da variância de maturidade (FMI, 2019). Por sua vez, Broner, Lorenzoni e Schmukler (2013) documentam que países emergentes emitem dívidas de prazo menor em momentos de crises financeiras, evidência que contribui para o entendimento de que a capacidade de emissão de títulos de longa duração depende da percepção de risco dos investidores, o que tende a penalizar mais países emergentes.

### **2.1.8 Mudança de disposição**

Por fim, alternativamente, mudanças políticas em um país (revoluções, trocas de governo etc.) podem diminuir a disposição de pagar, ainda que a capacidade permaneça a mesma (STURZENEGGER, ZETTELMEYER, 2006). Estas situações de *default* costumam ser referidas pela literatura especificamente como repúdio da dívida (*debt repudiation*) (SACHS, 1984). Por exemplo, na 1ª Guerra Mundial, a Turquia, a Bulgária e o Império Austro-Húngaro suspenderam o pagamento de suas dívidas para credores nacionais de países inimigos e a Itália, a Turquia e o Japão fizeram o mesmo na 2ª Guerra Mundial. Ademais, governos revolucionários repudiaram suas dívidas após a tomada do poder no México (1914), na Rússia (1917), na China (1949), na Tchecoslováquia (1952) e em Cuba (1960) (STURZENEGGER, ZETTELMEYER, 2006).

Note-se, contudo, que mudanças de disposição não foram analisadas apenas em decorrência de eventos políticos extraordinários: Eaton e Gersovitz (1981) sugeriram e encontraram evidências de que países mais dependentes de comércio exterior sofreriam mais com possíveis retaliações por um *default*, que envolvem custos de reputação e restrições ao acesso ao mercado de crédito internacional. Assim, características diferentes dos países podem estar atreladas a uma maior ou menor disposição de honrar com seus compromissos financeiros, posto que a relação custo-benefício da inadimplência não é a mesma para todas as nações.

### 2.1.9 Considerações

Conforme podemos ver, não há um consenso na literatura sobre o que causa as crises moratórias: há diferentes explicações que, em geral, não são excludentes entre si. Além disto, as causas para uma crise moratória podem variar e é frequente que mais de uma causa concorra para explicar o mesmo evento. Por exemplo, um país pode deixar de pagar sua dívida tanto porque ela é demasiadamente estruturada em títulos de curto prazo quanto pelo aumento dos juros internacionais e pela eleição de um governo que se declare contra os credores. A discussão, então, seria sobre qual fator contribuiu mais para o resultado.

Ademais, registre-se que mesmo cenários em que não chega a haver qualquer *default* ou renegociação da dívida podem causar suficientes transtornos ao país, até porque muitas medidas adotadas para evitar crises moratórias são recessivas, como o aumento dos juros, por exemplo. Neste sentido, existe uma literatura que se dedica a estudar as novas formas de volatilidade e vulnerabilidade perenes que a integração financeira trouxe<sup>7</sup>. Este estudo, contudo, irá se concentrar na detecção de eventos mais extremos de *default*. Portanto, no próximo item, passaremos a abordar os principais sinalizadores destas crises segundo a literatura empírica, a qual será dividida conforme seus métodos.

## 2.2 INDICADORES DE CRISE - EWS: REVISÃO DA LITERATURA EMPÍRICA

Há um espectro amplo de técnicas utilizadas para detectar *early warning signals* (EWS) de riscos sistêmicos, como crises bancárias, cambiais e moratórias. Adotaremos, assim, a divisão proposta pelo projeto de tomografia de riscos sistêmicos financiado pela União Europeia, o SYRTO, que considera haver três principais abordagens para o estudo de EWS: os modelos de limites, os modelos logit/probit e os métodos de mineração de dados (SIEGMANN, 2017).

Por sua vez, em relação às variáveis preditivas e explicativas de crises moratórias tipicamente utilizadas na literatura, Manasse e Roubini (2009) as separam em medidas de solvência, que captam a capacidade que um país tem de pagar, como a relação dívida/PIB; medidas de liquidez, que captam a capacidade de refinanciar sua dívida em um momento, com medidas como a relação entre a dívida de curto prazo e as reservas internacionais; medidas

---

<sup>7</sup> Ver, por exemplo, Kaltenbrunner e Paineira (2015).

políticas e institucionais<sup>8</sup>, que captam a disposição de pagar; e, por fim, outras variáveis macroeconômicas que captam tanto a capacidade quanto a disposição para pagar, como a taxa de câmbio.

Nos próximos itens, apresentaremos brevemente cada grupo de modelos e elencaremos alguns achados empíricos obtidos através deles, com ênfase para as variáveis mais significativas dos estudos e com o nível de precisão obtido nas previsões. Em relação a este, o leitor notará certa falta de uniformidade na métrica avaliativa: isto ocorre pois, de fato, as pesquisas nesta área têm utilizado diferentes métricas para avaliar e divulgar os resultados obtidos, a depender muito dos próprios tipos de modelos utilizados. Falaremos mais a respeito de diferentes métricas para avaliar problemas de classificação no capítulo 2 desta dissertação. Independentemente disso, os estudos também são feitos com diferentes variáveis, de países e de períodos diferentes, e o próprio tamanho da amostra de validação varia conforme cada estudo (e há estudos que só reportam os resultados dentro da amostra). Tudo isto impede a comparação direta entre eles. Dito isso, elencar-se-ão os principais achados de cada estudo, na ótica deste trabalho, a fim de se desenhar um panorama geral da literatura.

### 2.2.1 Modelos de limites

Estes modelos, não paramétricos, atribuem um “sinal de alerta” sempre que certas variáveis consideradas importantes para prever o evento ultrapassem um limiar histórico (por exemplo, a variável está em um valor correspondente a 95% do maior registro no passado). Tal método foi gestado no próprio FMI no final da década de 1990 para prever crises cambiais e é também chamado de “*Noise-to-Signal*” (SIEGMANN, 2017).

Modelos de limites usualmente são avaliados pela chamada relação sinal-ruído (*noise-to-signal ratio*). Esta relação captura a taxa entre alarmes falsos e eventos corretamente previstos, a partir da seguinte equação (KAMINSKY, 1999):

$$\frac{\text{Percentual de Falsos Negativos (Erros Tipo II)}}{1 - \text{Percentual de Falsos Positivos (Erros Tipo I)}}$$

---

<sup>8</sup> Manasse e Roubini (2009), por exemplo, incorporam em sua lista de candidatas a variáveis dados como índices de abertura econômica, *Freedom House Index* e anos até a próxima eleição (contudo, após uma triagem, apenas o número de anos até as eleições presidenciais permaneceu no modelo dos pesquisadores).

Reinhart (2002), tentando identificar crises cambiais e bancárias com a técnica, utilizando-se de uma média de 5 indicadores, obteve, respectivamente, 70 e 72% de previsões corretas, a uma relação sinal-ruído de 0.45 e 0.5. No mesmo estudo, Reinhart (2002) também concluiu que mais 84% das crises moratórias em países emergentes de sua amostra estão associadas a uma crise cambial (ocorrendo antes ou depois em um período de 24 meses) – o inverso, contudo, não é verdadeiro: um pouco menos da metade das crises cambiais estavam associadas com *defaults*. Por sua vez, para países desenvolvidos, nenhuma conexão entre essas crises foi detectada.

Também tentando prever crises cambiais e bancárias, Kaminsky (1999) obteve, analisando a taxa real de câmbio, relações sinal-ruído de 0,2 para crises cambiais e 0,3 para crises bancárias. Depois destas variáveis, a taxa real de juros globais mostrou-se a melhor para prever tais crises, exibindo um sinal-ruído de 0,4. Neste estudo, a variável *dummy* que captava desregulamentação nos mercados financeiros domésticos e mundiais não obteve taxas boas: 1 para crises cambiais (ou seja, a cada alarme verdadeiro, a *dummy* emitiu também um alarme falso) e 0,8 para crises bancárias.

O método *Noise-to-Signal* é bastante disseminado na previsão de outras crises, mas há poucas pesquisas que o tenham utilizado para detectar crises moratórias. Wijayanti e Rachmanira (2020) encontraram na relação entre dívida externa e renda nacional bruta o melhor predictor, detectando 72,7% das crises com um sinal ruído de 0,09, seguido da relação entre o pagamento de serviço da dívida sobre a renda nacional bruta (61% e 0,2) e da taxa cambial (60% e 0,06). Contudo, apenas resultados dentro da amostra são divulgados, e a literatura reporta uma baixa associação, no contexto de EWS para *defaults*, entre ajuste dentro da amostra e desempenho preditivo fora dela (FUERTES, KALOTYCHOU, 2006). Savona e Vezzoli (2015), também com o método *Noise-to-Signal*, obtiveram, fora da amostra, uma taxa AUC (*Area Under The Curve*)<sup>9</sup> de 0,66 e sensibilidade (taxa de verdadeiro-positivos) de 0,77.

Por fim, através de um modelo de limites, Dawood, Horsewood e Strobel (2017) buscaram variáveis para prever crises com dois anos de antecedência, encontrando como melhores sinais, segundo taxa sinal-ruído, a relação de crédito contraído junto ao FMI/PIB (0,21), os juros globais (0,33), o saldo na balança de transações correntes (0,33) e a relação reservas/PIB (0,36). Os autores encontraram algumas evidências de heterogeneidade regional

---

<sup>9</sup> Quanto mais perto de 1, melhor é a AUC. As diferentes métricas de avaliação de modelos utilizada na literatura de EWS serão avaliadas no capítulo 3 deste trabalho, junto à discussão dos resultados.



nos determinantes de default<sup>10</sup>. Note-se, contudo, que estes resultados foram obtidos dentro da amostra. Fora da amostra, apenas resultados com modelos regionais foram divulgados, com resultados bastante díspares: enquanto no grupo de países desenvolvidos e na América Latina o percentual de crises prevista foi bom (85% e 100%), no grupo de nações africanas e do Oriente Médio o percentual foi de 0%.

### 2.2.2 Modelos logit

O segundo método diz respeito aos modelos logit. Tais modelos estatísticos são especificados com variáveis dependentes binárias, aceitam diversas variáveis independentes e têm fácil interpretação, por isto encontram largo uso na previsão de eventos de crise. Em verdade, modelos logit e suas variações, como a multinomial, por exemplo, são nitidamente os mais usados pela literatura de EWS no âmbito de crises moratórias.

A variação multinomial destes modelos aceita mais de uma variável dependente, não sendo necessário restringir os resultados a duas categorias. No caso da previsão de default, faz-se possível, por exemplo, definir um espectro maior de cenários, como de normalidade, crise e de ajuste (pós-crise) – por exemplo, em Ciarlone e Trebeschi (2005). Alguns autores sugerem esta mudança, pois acreditam que os indicadores podem se comportar diferentemente em tempos normais e em tempos pós-crise, de modo que combinar as duas situações em uma só categoria poderia formar um “viés pós-crise”. Outros autores, para combater tal viés, eliminaram as observações pós-crise da amostra - por exemplo, Fuertes e Kalotychou (2007) e Savona e Vezzoli (2015). Ainda que isto implique em perda de informação, esta foi também a abordagem adotada na base de dados utilizada neste trabalho.

Mcfadden et al (1985) construíram modelos logit, obtendo como melhor resultado 82,5% de precisão dentro da amostra para o período de 1971 a 1982, com 14,5% de erros tipo I e 23,5% de erros tipo II. Nestes modelos, as variáveis que se mostraram mais importantes foram reservas/PIB, importações/PIB, crescimento real do PIB e uma variável *dummy* de histórico de crises moratórias.

Peter (2002), por sua vez, elaborou um modelo logit com dados de 1984 a 1997 e obteve 86% de precisão fora da amostra, com 16,5% de erros tipo I (alarme falso) e 12,5% de erros tipo II (crises não previstas). A variável mais importante foi o histórico recente de

---

<sup>10</sup> Por exemplo, a taxa de juros globais não previu tão bem as crises em países desenvolvidas (taxa sinal-ruído de 0,62). Por sua vez, a relação das reservas com o PIB obteve resultados muito piores no Sudeste Asiático (0,78) em relação à América Latina (0,25) e a África (0,10). Semelhante comportamento ocorreu em relação à abertura comercial (importações mais exportações/PIB): Sudeste Asiático (0,53), América Latina (0,33) e África (0,17).

pagamento dos países analisados. Outras variáveis importantes foram a variação do crescimento do produto nacional bruto per capita e um índice de risco político.

Manasse e Roubini (2009) incluíram na definição de crise casos próximos de default, evitados por vultosos empréstimos do FMI, o que foi seguido por vários trabalhos seguintes<sup>11</sup>, incluindo o presente. O período analisado foi de 1970 a 2002. O modelo logit estimado detectou corretamente 74% das crises dentro da amostra, a uma taxa de alarmes falsos de apenas 6%. Esta baixa taxa se manteve no experimento fora da amostra; o percentual de crises previstas, contudo, caiu para 45%.

Em relação aos determinantes de crise, o modelo logit de Manasse, Roubini e Schimmelpfennig (2003) revelou forte associação de crises moratórias com as relações dívida externa/PIB e dívida curto prazo/reservas: quanto maiores, mais chance de o país enfrentar dificuldades. Ademais, maiores juros norte-americanos aumentaram a chance de crise, ocorrendo o inverso com maiores saldos na balança de transações correntes. Ainda, alto crescimento real do PIB mostrou-se inversamente atrelado a crises, e períodos de alta inflação, ou alta volatilidade na inflação, revelaram-se associados a maior probabilidade de dificuldades moratórias. Os referidos autores usaram estes achados para criar um modelo CART de árvores de classificação, o qual será referido no próximo item.

Ciarlone e Trebeschi (2005) desenvolveram um modelo multinomial logit, em que há três estados possíveis em relação a crises moratórias (normalidade, crise e ajuste pós-crise), obtendo 76% de precisão fora da amostra, com capacidade de previsão de 80% das crises e uma taxa de 20% de alarmes falsos.

Fuertes e Kalotychou (2006), usando dados de 1983 a 2002, desenvolveram diversos modelos logit, encontrando, para seu melhor modelo, taxas de erro tipo I e II de 16% e 25%, respectivamente. Segundo o achado dos autores, serviram como importantes indicadores dos modelos: variáveis que sinalizam a estrutura da dívida, como dívida externa/PIB, dívida a agências bi- ou multilaterais/dívida total, dívida de curto prazo/dívida total, crédito do FMI/exportações; uma variável macroeconômica - produto nacional bruto per capita; uma variável de estabilidade macroeconômica - volatilidade do crescimento do produto nacional bruto; e, por fim, um sinal de abertura – comércio externo<sup>12</sup>/PIB. A intuição por trás desta última variável, com impacto negativo nas chances de *default*, seria que países mais globalmente integrados incorrem em um custo de oportunidade maior ao não honrarem suas dívidas, tendo, por conseguinte, mais disposição em pagá-las. Tal resultado está de acordo

---

<sup>11</sup> Por exemplo, Fioramanti (2008) e Dawood, Horsewood e Strobel (2017).

<sup>12</sup> Importações + Exportações

com os achados empíricos e teóricos de Eaton e Gersovitz (1981), já referidos. Outro resultado importante dos autores foi que, a despeito da literatura teórica ressaltar a vulnerabilidade dos países emergentes a mudanças de sentimento do mercado, *proxies* de aversão ao risco e volatilidade dos mercados não aumentaram o poder preditivo dos modelos.

Catão e Milesi-Ferretti (2014), usando dados de 1970 a 2011, especificaram diversos modelos probit para detectar eventos de crises e concluíram que déficits em conta corrente teriam um poder preditivo maior que qualquer outra variável na maioria das especificações.

Dawood, Horsewood e Strobel (2017) desenvolveram dois modelos logit regionais: um modelo binomial em que anos de crise e anos pós-crise são unificados como uma só observação e outro modelo multinomial com uma categoria de crise, outra de normalidade e uma terceira de ajuste (pós-crise). Os modelos tiveram bons resultados dentro da amostra, mas não mantiveram o desempenho fora dela – mesmo o modelo binomial, que se mostrou melhor em quase todos os experimentos, previu corretamente apenas 25% das crises da América Latina. Os autores apresentam uma variação do modelo com previsões recursivas, o qual consegue resultados ótimos (100% de crises detectadas na América Latina e 71,4% no grupo de países desenvolvidos). Contudo, a base de dados para teste fora da amostra era pequena e nela havia apenas um episódio de crise na América Latina.

### 2.2.3 Modelos Classificadores de *Machine Learning*

Por fim, o terceiro método foi denominado pelo SYRTO genericamente como “*data mining*”, ou mineração de dados, o que abrange o uso de modelos classificadores de *machine learning*, maneira pela qual nos iremos referir a eles. Estas são as técnicas mais novas: trata-se do uso de princípios algorítmicos utilizados no ramo de inteligência artificial, como reconhecimento de imagens e de vozes, para ajudar na compreensão e previsão de crises financeiras.

Tais técnicas, além de conseguir resultados por vezes superiores aos da estatística tradicional - ver, por exemplo, Fioramanti (2008) e Savona e Vezzoli (2015) - também possibilitam, por vezes, a melhor compreensão dos fenômenos. Este é especialmente o caso das árvores de regressão e classificação – dentro do universo de *machine learning*, estes algoritmos não são os mais potentes classificadores; todavia, eles são de fácil interpretação e oferecem esquemas que podem enriquecer a compreensão dos problemas, ainda que eles tenham viés preditivo, sem pretensões explicativas. Aprofundaremos a lógica destes

algoritmos específicos em seção mais apropriada – neste tópico, elencaremos alguns estudos na área com métodos de *machine learning* em geral.

Manasse, Roubini e Schimmelpfennig (2003) criaram modelos preditivos usando árvores de classificação (CART), a fim de comparar com o modelo logit. O CART obteve acurácia fora da amostra de 85% de previsões acertadas com taxa de alarmes falsos de 21%. Os autores concluíram que o modelo logit dispara menos alarmes falsos, mas as árvores detectam mais crises. Ademais, foi concluído que alto endividamento (maior que 49,7% do PIB) e alta inflação (maior que 10,5% anual) estão associados aos maiores riscos de *default*. Por sua vez, as condições menos associadas ao risco de *default*, segundo tal esquema, são as de baixo endividamento externo, baixa relação de dívida a curto prazo em relação às reservas e baixo endividamento externo em relação às receitas. Contudo, baixo endividamento externo não é uma condição necessária para evitar crises moratórias. Os países podem evitá-las, segundo o esquema de árvore gerado, através de baixa inflação, baixa relação de necessidade de financiamento externo com as reservas e baixa relação da dívida pública com as receitas. Este tipo na análise é um dos maiores atrativos das árvores de classificação e regressão: elas permitem análises condicionais de cenários, o que ajuda muito na formulação de políticas e previsão de cenários.

Neste sentido, Roubini e Manasse (2009) usaram o algoritmo CART de árvores de classificação para analisar diferentes tipos de crises moratórias. Eles argumentam que limites incondicionais, como a relação dívida/PIB, têm pouco valor para avaliarmos a probabilidade de *default*, posto que outras variáveis, como nível de endividamento a curto prazo, podem afetar tais limites. Os autores propuseram três categorias de risco: (a) riscos de solvência, (b) riscos de liquidez e (c) riscos macroeconômicos decorrentes da taxa de câmbio.

Os primeiros caracterizam-se por uma dívida externa acima de 49,7% do PIB, junto a um desbalanço fiscal e monetário e alta necessidade de financiamento externo. Os segundos seriam casos de níveis moderados de endividamento, mas com dívida a curto prazo a uma taxa superior a 130% das reservas e em um ambiente de incerteza global. Por fim, os riscos macroeconômicos decorreriam de baixo crescimento econômico junto a taxas relativamente fixas de câmbio. Roubini e Manasse (2009) conseguiram prever 93,9% dos *defaults* a uma taxa de alarme falso de 18,5%.

Fuertes e Kalotychou (2007) realizaram predições com o algoritmo de *K-Clustering* e combinações deste com o modelo logit. Os modelos concordaram em 70% das vezes e houve bastante convergência nos indicadores relevantes. Em relação a estes, destaca-se, quanto ao setor externo: volatilidade do crescimento das exportações e relação balanço comercial/PIB;

quanto à exposição ao crédito, dívida externa/PIB, dívida de curto prazo/dívida total e crédito junto ao FMI/exportações; quanto às condições domésticas, a relação crédito privado/PIB, crescimento do PIB, volatilidade do crescimento do PIB e taxa de câmbio real; por fim, ambos modelos utilizaram o nível de abertura comercial<sup>13</sup>. Com dados de 75 países do período de 1983 a 2000, os pesquisadores concluíram que, embora individualmente o algoritmo de K-Clustering tenha gerado as piores previsões, a combinação deste algoritmo com o modelo logit rende resultados melhores do que aqueles alcançados separadamente pelos métodos.

Fioramanti (2008) usou uma rede neural de duas camadas para prever *defaults* com dados do período de 1981 a 2004. Fora da amostra, o pesquisador obteve uma excelente taxa de acerto de 87,45% de não-cries e de 91,80% das crises. Contudo, conforme reconhece o próprio autor, a desvantagem do modelo que ele não oferece intuições imediatas ao *policy maker*, sendo uma ferramenta estritamente preditiva.

Savona e Vezzoli (2015), utilizando-se de árvores de decisão (CART), aprimoradas pelo algoritmo CRAGGING (*Cross-validation Aggregating*) e agregadas pelo algoritmo de florestas aleatórias (*Random Forest*) defenderam que esta combinação traria a melhor descrição da amostra mantendo-se a capacidade preditiva, prevendo, fora da amostra, 88% das crises de *default* e 64% das não-cries. Além disso, os autores encontraram que o histórico de *default* e uma alta relação de dívida curto prazo sobre as reservas seriam os principais determinantes de crises, seguidas da taxa de crescimento do PIB e juros norte-americanos.

Savona, Vezzoli e Ciavolino (2015), usando método semelhante, acharam indicadores diferentes para cada perfil de país. Para países emergentes, os melhores indicadores foram históricos de *default*, taxa de juros do tesouro americano, relação entre dívida de curto prazo e reservas internacionais e uma variável de contágio (número de países com *default* no mesmo ano). Por outro lado, para Portugal, Irlanda, Grécia e Espanha, as melhores variáveis para determinação de *default* foram as taxas de inflação, a relação dívida/PIB, as taxas de crescimento das exportações e do PIB.

## 2.2.4 Heterogeneidades temporais e espaciais

Mcfadden et al (1985) argumentaram a provável existência de alguma heterogeneidade temporal: a ligação entre atributos macroeconômicos e os episódios moratórias variariam em

---

<sup>13</sup> Importações + Exportações

diferentes épocas. Isto poderia afetar as suposições dos modelos. À heterogeneidade temporal, soma-se a heterogeneidade espacial: a literatura vem investigando se diferentes países e regiões podem ter diferentes determinantes de crises moratórias.

Fuertes e Kalotychou (2006) concluíram que, embora haja diferenças nacionais, regionais e temporais na economia (refletindo, por exemplo, diferentes fases de ciclos de negócios, mudanças nos mercados de capitais internacionais e mudanças na natureza econômica dos países emergentes), o impacto destas heterogeneidades nas probabilidades de *default* não chega a prejudicar modelos simples de EWS. De fato, analisando-se o desempenho fora da amostra, os modelos logit mais simples dos autores, que não controlavam heterogeneidades temporais e entre as nações, exibiram um desempenho preditivo melhor. Restringir a amostra dos países a regiões tampouco resultou em ganhos significativos de predição em relação a uma análise global. Contudo, quando analisados dentro da amostra, modelos com controles para heterogeneidades regionais descreveram muito bem os dados, mostrando que, mesmo se tal controle não tiver utilidade preditiva, ele é importante para a formulação de políticas e melhor compreensão dos fenômenos reportados.

Por sua vez, Darwood, Horsewood e Strobel (2017), conforme já referido, apresentaram diferentes determinantes para regiões diferentes, tanto em seus modelos logit, quanto de limites (embora as diferenças não tenham sido sempre as mesmas). Ademais, os autores concluíram que a heterogeneidade especial relevante é a regional, e não nacional, corroborando o concluído por Fuertes e Kalotychou (2006). Contudo, os estudos divergem quanto à utilidade da divisão por região para melhor as previsões: Darwood, Horsewood e Strobel (2017) concluem que os modelos regionais apresentam ganhos de predição.

### 2.2.5 Considerações

Conforme exposto, os estudos empíricos de previsão de *default* variam quanto aos sinais encontrados e aos pesos concedidos a cada variável; contudo, pode-se dizer, de maneira geral, que a literatura encontra importância preditiva em (i) indicadores de estrutura da dívida pública, em especial os que considerando sua maturidade; (ii) indicadores de endividamento externo, como dívida externa/PIB, ou de liquidez externa, em especial os que consideram o nível das reservas de alguma forma; (iii) histórico de *default*; (iv) indicadores de crescimento econômico e (v) taxa de juros internacionais. Pode-se notar, assim, que os estudos empíricos de predição de *default* dialogam bastante com várias das teorias explicativas desenvolvidas para estes eventos, descritas alhures.

Em relação aos métodos, nota-se que o uso da metodologia sinal-ruído é bem diminuto, sendo mais comum para predição de outras crises, como cambiais ou bancárias. Os modelos logit, por sua vez, configuraram, por bastante tempo, a principal metodologia para o estudo de eventos de *default*. Uma vantagem significativa destes modelos é que eles permitem testes de inferência causal. Os algoritmos de árvore, bem como diversos métodos em *machine learning*, não permitem que se estabeleça causalidade com rigor estatístico. Por sua vez, os algoritmos de árvore têm como grande atrativo o fato de explorarem e criarem interações entre as variáveis, de modo a tornar possível e fácil a análise condicionada destas. Isto é especialmente interessante pois, como indica a literatura tanto teórica quanto explicativa, crises moratórias são frequentemente eventos multifatoriais, de modo que uma análise completa sempre terá de considerar mais de uma variável. Outra característica dos algoritmos de árvore é o fato de, sendo um método não paramétrico, não dependerem de suposições quanto à distribuição dos dados. Isto, contudo, também limita a realização de testes estatísticos em seus resultados.

Por fim, nota-se que há pouca literatura sobre segmentação regional dos países para predição de *default* e os estudos que existem apresentam resultados conflitantes quanto à utilidade de tal divisão.

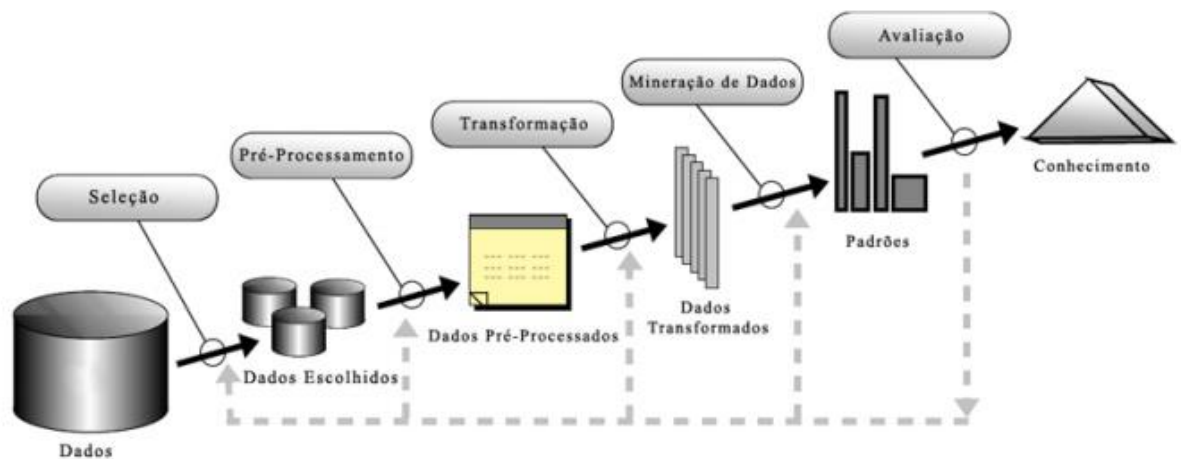
### 3 METODOLOGIA

Neste capítulo, introduzir-se-ão, inicialmente, alguns conceitos de mineração de dados e *machine learning* para, na sequência, adentrar-se na metodologia específica e base de dados utilizados no presente trabalho.

#### 3.1 MINERAÇÃO DE DADOS E *MACHINE LEARNING*

A mineração de dados é uma etapa do processo de descoberta de conhecimento em bases de dados (*knowledge discovery in databases - KDD*), o qual se encontra abaixo representado:

**Figura 1 – Processo de descoberta de conhecimento em bases de dados**



**Fonte:** CAMILO e SILVA, 2009, p.3.

Inicialmente, identificam-se os dados relevantes para o problema em questão. Seleciona-se então a base de dados a ser utilizada e é realizada a coleta. A próxima fase é o processo de limpeza dos dados, que geralmente envolve filtrar, combinar e preencher valores vazios (CAMILO, SILVA, 2009). Na limpeza de dados, valores absurdos também podem ser tratados. Com os dados pré-processados, explora-se então a base de dados com técnicas (algoritmos) de mineração, a fim de se encontrar padrões ou se realizar tarefas de predição, associação ou agrupamento – para estas tarefas, algoritmos de aprendizado de máquina (*machine learning*) podem ser utilizados. Com os resultados, realiza-se a avaliação deles e deriva-se algum conhecimento do experimento.



A mineração de dados lida com a análise dos dados em busca de padrões úteis ou informativos, o que, por si só, não é nada novo. Na mineração de dados, estes são armazenados eletronicamente, e a busca é automatizada – ou ao menos aumentada – por um computador, o que é relativamente recente. Em especial, a abundância de dados digitalizados e acessíveis da atualidade é um fato sem precedentes (dando origem à era do *Big Data*), tendo ocorrido também, recentemente, avanços no *design* e condições de processamento de algoritmos de mineração de dados, o que, conjuntamente, vêm impulsionando a mineração de dados.

Em relação aos algoritmos de *machine learning*, os primeiros datam de meados dos anos 1950, como o Perceptron de Rosenblatt<sup>14</sup>, o primeiro algoritmo de rede neural moderno. Tais modelos são algoritmos estatísticos treinados com uma base de dados a fim de obter resultados fora da amostra. O desenvolvimento destas técnicas veio como uma expansão tanto do conhecimento em estatística quanto de ciência da computação.

Em alguns casos, desenvolvimentos paralelos nas duas áreas convergiram para soluções semelhantes – este é o caso dos algoritmos de árvores de decisão, que serão usados neste trabalho. Quatro estatísticos de universidades californianas escreveram um livro sobre árvores de classificação e regressão em meados da década de 1980<sup>15</sup>. Ao mesmo tempo, durante os anos 1970 e início dos anos 1980, um pesquisador australiano de *machine learning*, J.Ross Quinlan, estava desenvolvendo um sistema de classificação em forma de árvores. Ambos desenvolveram algoritmos parecidos, mas tanto um criador quanto os outros, só souberam da coincidência muito tempo depois (Witten et al, 2016). Como se vê, a fronteira entre *machine learning* e estatística pode ser nebulosa.

*Machine learning* é também uma espécie de inteligência artificial. Com ela, o ser humano desiste de listar regras lógicas para o programa e deixa que algoritmos estatísticos criem suas próprias regras a partir do “estudo” da base de dados. Por exemplo, ao se criar um modelo para tradução entre duas línguas, ao invés de se tentar convocar linguistas especializados nos dois idiomas para elencar exaustivamente regras lógicas gramaticais a fim de orientar o algoritmo, alimenta-se o modelo com uma vasta base de textos e adotam-se algoritmos estatísticos que aprendem a probabilidade de as palavras aparecerem juntas. (Trata-se de um exemplo simplificado, e nada impede que os dois paradigmas sejam

---

<sup>14</sup> Ver *paper* seminal em: ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, v. 65, n. 6, p. 386, 1958. Disponível em: <https://www.ling.upenn.edu/courses/cogs501/Rosenblatt1958.pdf>. Acesso em 15 de maio de 2021.

<sup>15</sup> Classification and Regression Trees Leo Breiman, Jerome Friedman, Charles J. Stone, R.A. Olshen (1984)

combinados de alguma forma, mas serve aqui para demonstrar o fundamento estatístico das técnicas de *machine learning* em contraste com outras espécies de inteligência artificial).

Neste contexto, maiores e melhores bases de dados geralmente possibilitam um melhor treino do algoritmo, pois o “material de estudo” para o aprendizado de máquina é mais amplo. Contudo, estes algoritmos e as largas bases de dados requeridas têm um custo computacional tanto em capacidade de memória quanto em velocidade de processamento. Por esta e outras razões, a área não prosperou tanto em décadas passadas, até que, em meados de 2010, uma soma de fatores combinados reavivou o movimento. A capacidade de processamento havia obtido largos avanços e os dados foram se tornando abundantes na internet. Por fim, melhores algoritmos surgiram e foram sendo usados para reconhecimento de voz, de imagens e outras finalidades pelos principais *players* do mercado. Com o tempo, os algoritmos foram se popularizando e hoje estão sendo utilizados para os mais variados fins, da indústria à ciência. Por exemplo, recentemente cientistas divulgaram que a partir da análise de diversas gravações de tosses (forçadas), um algoritmo alcançou bons resultados na detecção de covid-19<sup>16</sup>.

Na ciência econômica, o uso de *machine learning* no meio acadêmico vem aos poucos ganhando corpo (VARIAN,2014; BAJARI et al, 2015; ATHEY,2019). Um dos entraves desta inserção é que boa parte dos métodos de *machine learning* não são apropriados para detectar inferência causal, e muito do trabalho empírico em economia tem como interesse primário estimar um efeito causal, como o efeito de um aumento de preços ou do salário mínimo (ATHEY, 2019). Todavia, Athey (2019) prevê uma literatura dedicada a aprimorar métodos de *machine learning* para o uso de inferência casual e opina que economistas terão um papel importante nestes avanços.

Ademais, trabalhos econômicos empíricos também são bastante focados na predição de crises ou do desempenho de variáveis econômicas. Para estes problemas, *machine learning* oferece soluções bastante imediatas. Por exemplo, Vasconcelos (2017) usou técnicas de *machine learning* para estimar o PIB de diferentes países, e Nyman e Ormerod (2017) usou os algoritmos para prever recessões.

Os algoritmos de *machine learning* podem ser divididos em duas categorias: aprendizado supervisionado ou não-supervisionado. Nos primeiros, o algoritmo é alimentado com observações rotuladas, a fim de que ele as classifique futuramente. Por sua vez, nos

---

<sup>16</sup> Algoritmo americano identifica “tosse de Covid” inaudível para humanos. Reportagem da Istoé. EDIÇÃO Nº 2692 20/08. Disponível em: <https://istoe.com.br/algoritmo-americano-identifica-tossida-de-covid-inaudivel-para-humanos/> Acesso em 10 de agosto de 2021.

segundos, os algoritmos acham padrões nos registros, geralmente a fim de agrupá-los (*clusterização*). Neste estudo, o algoritmo de *machine learning* utilizado será do tipo supervisionado, posto que ele receberá registros (variáveis) já rotulados de países que enfrentaram uma crise moratória ou não.

### 3.1.1 Treinamento e teste do algoritmo

De maneira simples, treinar o algoritmo é deixar que o programa aprenda relações existentes em uma base de dados e assim crie seus próprios parâmetros decisórios. Por sua vez, testar o algoritmo é avaliar seu desempenho em uma base de dados independente, que não foi usada para desenvolver o modelo. Esta separação é fundamental, pois é ela que garante que os resultados do modelo são generalizáveis. Por isso, nenhuma informação adicional da base de testes deve ser usada no desenvolvimento do algoritmo com a base de treino.

Em relação à divisão da base de dados para teste e treino, existem diferentes abordagens, sendo as duas principais a divisão teste-treino e o método de validação cruzada (*cross-validation*).

A primeira abordagem consiste simplesmente em dividir a base de dados: uma parte é exclusivamente dedicada para o treino, onde ocorre o aprendizado do algoritmo, e a outra parcela é exclusivamente voltada para teste e verificação de resultados. Esta divisão pode ocorrer em diferentes percentuais, sendo comuns as proporções de 70/30 e 80/20. Contudo, um dos problemas desta abordagem é que ela envolve desperdício de dados para treinar o algoritmo.

Na validação cruzada, ocorre uma divisão aleatória da base de treinos em um número  $k$  de subconjuntos e cada subconjunto é usado uma vez para teste e  $k-1$  vezes para treino. Ao final, o desempenho do algoritmo nos  $k$  subconjuntos é sumarizado com base na média entre os testes de previsão realizados em cada subconjunto. É bastante comum também que a validação cruzada seja usada não como substituição da avaliação com a base de testes, mas sim como etapa preliminar a ela, a fim de se estimar os resultados do algoritmo em bases de dados autônomas.

No caso do experimento realizado nesta pesquisa, não seria possível utilizar a técnica de validação cruzada, pois a base de dados envolve observações com relação temporal. Realizar sua divisão aleatória para fins de treino permitiria o uso de dados do futuro para prever o passado. O uso de *cross-validation* poderia permitir, por exemplo, que o algoritmo

usasse dados de 2000 para prever *defaults* de 1999 ou de 1993, o que não seria consistente com um exercício de previsão.

Por outro lado, conforme dito, a divisão da base de dados nas proporções usualmente adotadas no âmbito de *machine learning*, como 80/20 e 70/30, implicaria importantes perdas de informação para o treinamento do algoritmo, aspecto importante, sobretudo quando se lida com bases de dados pequenas, como a utilizada nesta pesquisa.

Por tais razões, a seguinte abordagem foi utilizada como metodologia de treinamento e avaliação do algoritmo: inicialmente, utilizaram-se dados de alguns poucos anos para o aprendizado inicial do algoritmo. Em seguida, com o modelo desenvolvido com estes dados, testou-se seu desempenho com as observações do ano seguinte ao último ano utilizado para treino do algoritmo. Por exemplo, treinou-se o algoritmo com os dados de 1975 a 1978 e o modelo foi testado nos dados de 1979. Ato contínuo, treinou-se o algoritmo com os dados de 1975 a 1979, realizando-se o teste nos dados de 1980, e assim sucessivamente.

Esta metodologia evita as contradições que ocorreriam com a validação cruzada e ao mesmo tempo aproveita ao máximo a informação contida na base de dados, sendo amplamente utilizada na literatura de sinalização de crises (ver, por exemplo: CHAMON, MANASSE, 2007; FUERTES, KALOTYCHOU, 2007; DUTTAGUPTA, CASHIN, 2011; DAWOOD, HORSEWOOD, 2017).

### **3.1.2 Métricas de avaliação de modelos de classificação**

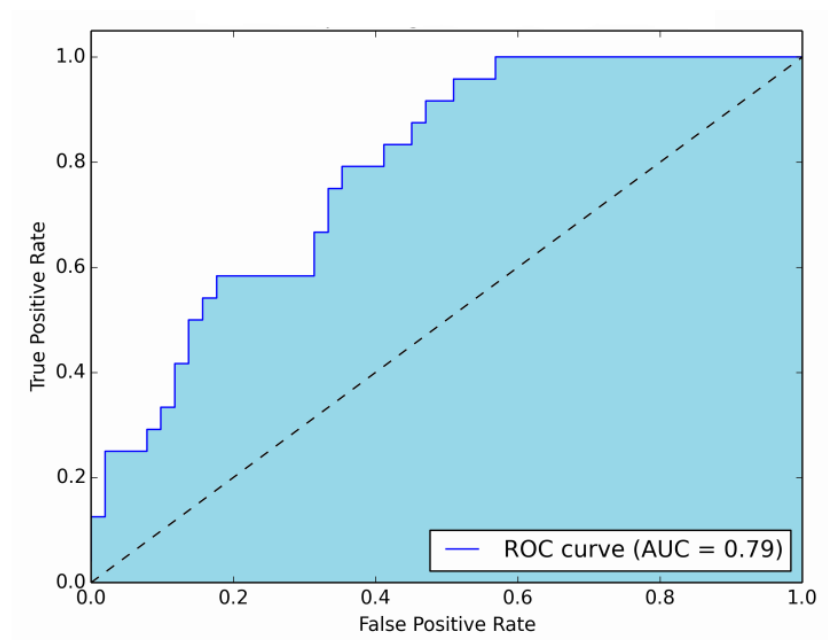
O experimento que executamos se trata de um problema de classificação, pois tentamos prever se nossas amostras se encaixam na classe “*default*” ou “*não default*”. Uma das principais questões que problemas de classificação trazem é qual será a métrica para julgar o desempenho do algoritmo.

É intuitivo usar o percentual de amostras que foram classificadas corretamente, a acurácia geral. Contudo, esta métrica é insuficiente no caso de bases de dados desbalanceadas – ou seja, em que uma categoria é muito mais prevalente que a outra. Por exemplo, se um algoritmo tenta detectar tumores em exames de imagem e apenas 0,5% das imagens contém algum tumor, o programa será capaz de classificar corretamente 99,5% das imagens prevendo sempre que não há tumor algum. Por óbvio, apesar do percentual de quase 100% de acurácia, este programa seria absolutamente inútil, se não contraprodutivo.

Tendo este problema em mente, outras métricas são utilizadas para avaliar o desempenho de classificadores, como o número de falso positivos, falso negativos e verdadeiros positivos, assim como a área ROC – *receiver operating characteristic*.

ROC significa *receiver operating characteristic*, um termo usado para caracterizar o *tradeoff* entre as taxas de falsos positivos (FPR)<sup>17</sup> e de verdadeiros positivos (TPR)<sup>18</sup>. É frequente que haja este dilema: modelos mais sensíveis acusarão mais eventos, detectando-os mais; contudo, muitas vezes isto ocorre ao custo de mais alarmes falsos. Assim, diferentes parâmetros dos modelos podem ser alterados para tornar o classificador mais ou menos sensível. A curva ROC, contudo, permite a comparação entre dois modelos mesmo considerando esta possível variação de cada modelo. Ela traça a FPR no eixo X e a TRP no eixo Y, de modo que uma linha perfeitamente diagonal representa um classificador aleatório. Quanto mais ao canto superior esquerdo estiver a curva, melhor é o classificador, conforme demonstra a imagem abaixo.

**Figura 2: Curva ROC e AUC**



**Fonte:** Statsexchange<sup>19</sup>

Uma consequência disto é que quanto maior a área abaixo da curva ROC, melhor pode ser considerado o classificador. Por isso, a área abaixo da curva, *Area Under the Curve*

<sup>17</sup> A taxa de falsos positivos é calculada pela seguinte fórmula: 
$$\frac{\text{Falsos Positivos}}{\text{Falsos Positivos} + \text{Verdadeiros Negativos}}$$

<sup>18</sup> A taxa de verdadeiros positivos é calculada pela seguinte fórmula: 
$$\frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}}$$

<sup>19</sup> DERNONCOURT, Frank. What does AUC stand for and what is it? Disponível em: <https://stats.stackexchange.com/questions/132777/what-does-auc-stand-for-and-what-is-it> Acesso em 12 de abril de 2021.

(AUC) é a medida mais usada em *machine learning* para avaliar classificadores em diversos contextos. Segundo Savona e Vezzoli (2015), o mesmo vale para modelos de sinais: usar a curva ROC fornece melhores informações sobre a confiabilidade na classificação de episódios de *default* e *não-default*. Assim, uma AUC de 0.5 significa que o modelo ou algoritmo não conseguiu aprender a ponto de discriminar bem as classes; por sua vez, considera-se em geral que, a partir de 0.7, o modelo tem um desempenho aceitável; por fim, a partir de 0.9, os resultados são considerados excelentes.

## 3.2 ALGORITMO E BASE DE DADOS UTILIZADOS

Introduzidos os principais conceitos de mineração de dados e *machine learning*, bem como as principais métricas para problemas de classificação, passa-se a expor, nos próximos itens, o método específico utilizado neste trabalho – o ambiente de implementação, o filtro, o algoritmo e a base de dados utilizados.

### 3.2.1 A plataforma Weka

Todas as etapas do experimento foram realizadas no ambiente de *software* WEKA - Waikato Environment for Knowledge Analysis<sup>20</sup>. Trata-se de um software de código aberto, em linguagem Java, desenvolvido na Universidade de Waikato (Nova Zelândia), que possibilita o tratamento de dados e a implementação dos principais algoritmos de *machine learning* do atual estado da arte, para tarefas como pré-processamento, classificação, agrupamento, regressão, visualização, entre outras. Note-se que a plataforma é interoperável com ambientes R e Python, havendo pacotes com os comandos do Weka nestas linguagens de programação, possibilitando assim a reprodução dos experimentos em diferentes ambientes.

### 3.2.2 Balanceamento de classes

Além das técnicas já abordadas de preparação de dados, que, no caso do experimento realizado, incluiu a remoção de valores absurdos e estruturação dos dados para leitura em ambiente Weka, este trabalho usou o filtro *ClassBalancer*, usado para problemas de classificação em que há uma classe minoritária de observações.

---

<sup>20</sup> Utilizou-se a versão 3.8.4 – a mais recente. Para mais informações a respeito do software, ver: <https://www.cs.waikato.ac.nz/ml/weka/index.html> Acesso em 28 de dezembro de 2020.

Algoritmos de classificação são planejados para minimizar o erro de classificação geral, independentemente da distribuição das classes de dados (ELREEDY, ATIYA, 2019). Por isto, eles acabam com suas previsões enviesadas para a classe majoritária. Reprisa-se aqui, conforme referido acima, que, se há duas classes de eventos, A e B, sendo que 95% das observações se encaixam na primeira categoria, o algoritmo pode obter 95% de acerto na classificação geral apenas prevendo sempre o evento do tipo A. Existem técnicas para corrigir isto, sendo o filtro *ClassBalancer* uma delas.

O método é simples e consiste em atribuir a cada observação um valor positivo como peso, de modo que a soma dos pesos das observações pertencentes a cada classe resulta no mesmo valor. No caso do experimento realizado, um peso muito maior foi dado para as observações de eventos de *default*, que são minoritários, de modo a contrabalancear a tendência do algoritmo de prever sempre a classe majoritária para minimizar o erro.

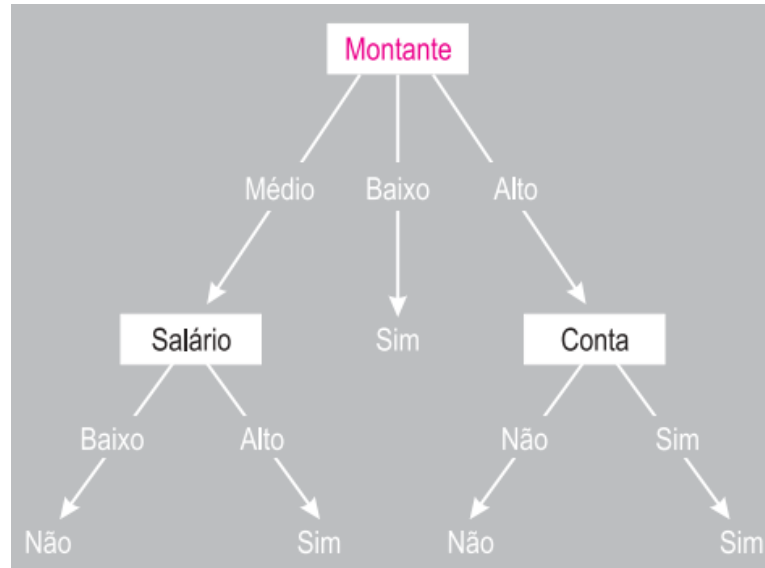
Para aplicar o filtro de balanceamento de classes, foi usada a ferramenta *FilteredClassifier* do Weka, de modo que o filtro foi aplicado apenas nas bases de treino, nos procedimentos de desenvolvimento dos modelos, não sendo aplicado nas bases usadas para teste.

### 3.2.3 O Algoritmo de árvores de classificação C4.5

Uma vez realizada a limpeza de dados e a correção do seu desbalanceamento com o já descrito filtro, o algoritmo C4.5, proposto por Quinlan (1993), foi implementado. No ambiente Weka, o algoritmo é apelidado de J48, por ser uma aplicação em Java. Trata-se de um método não paramétrico de partição recursiva que dá origem a uma árvore de decisão classificatória.

Esta árvore é representada de maneira invertida, ou seja, com as raízes em cima e as folhas (nós terminais) embaixo, de modo que a cada nível que se desce na árvore há uma progressão de informações requeridas que enfim levam a uma classificação ou previsão. Para uma melhor compreensão, colaciona-se um exemplo simplificado de árvore de classificação:

**Figura 3: Exemplo de árvore de classificação para classificar a concessão de empréstimos**



**Fonte:** LEMOS, STEINER e NIEVOLA (2005).

A árvore acima, designada para conceder ou não um empréstimo, tem como variáveis o montante a ser emprestado, o valor do salário do consumidor e a propriedade ou não de uma conta bancária. De acordo com as respostas de cada subproblema, chega-se a uma classificação final. Este tipo de esquema encontra aplicações nas mais diversas áreas, desde a classificação de solos até o diagnóstico de doenças.

Existem diferentes maneiras de se criar uma árvore de decisão, sendo o algoritmo C4.5 um dos principais algoritmos deste tipo. Visualmente as árvores de decisão se diferem principalmente quanto à ordem de aparição das variáveis dependentes (de cima para baixo, primeiro aparecem os atributos mais determinantes para a classificação) e na forma que cada atributo é usado para gerar nós inferiores (em direção às folhas).

Neste contexto, os algoritmos de árvore precisam de uma métrica para definir a importância de um atributo, sendo o índice de Gini e medidas de entropia as mais utilizadas. O algoritmo C4.5, em específico, utiliza a taxa de ganho de informação normalizada para decidir, a cada nó da árvore, qual atributo é o mais eficiente para realizar a próxima partição.

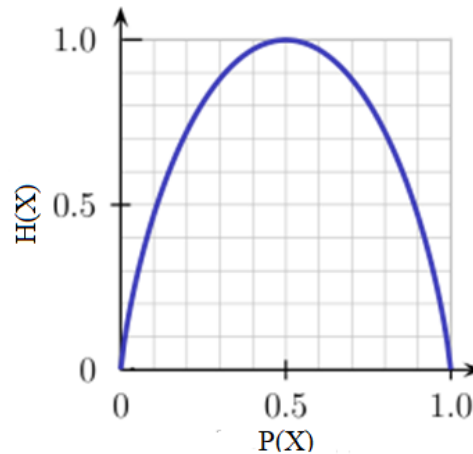
O ganho de informação é igual à diferença em entropia. Entropia (H), por sua vez, é uma medida de impureza definida por:

$$H = - \sum p(x) \log p(x)$$



Em que  $p(x)$  é a probabilidade de um resultado. Assim, conforme o gráfico abaixo ilustra, para uma situação em que há dois resultados possíveis, a entropia (impureza) é máxima quando uma observação extraída da amostra tiver 50% de chance de pertencer a cada resultado:

**Figura 4: Relação entre probabilidade e entropia**



**Fonte:** Elaboração do autor

O objetivo é escolher o atributo que separe melhor as classes, isto é, que gere menos impureza (entropia) nos subgrupos resultantes. A entropia também pode ser considerada como a média de informação necessária para identificar a classe de um caso  $C$  em um uma base de dados  $T$  – média aqui denominada  $info(T)$ .

Conforme dito, o algoritmo C4.5 utiliza a taxa de ganho de informação normalizada (“*gain ratio*”) para escolher um atributo  $X$  a ser utilizado para cada partição. Tal taxa é calculada conforme as equações abaixo (QUINLAN,1993):

$$info(T) = - \sum_{j=1}^k \frac{freq(C_j, T)}{|T|} \times \log_2 \left( \frac{freq(C_j, T)}{|T|} \right) \text{ bits} \quad (1)$$

$$info\ gain(X) = info(T) - \sum_{j=1}^p \left( \frac{|T_j|}{|T|} \times info(T_j) \right) \text{ bits} \quad (2)$$

$$split\ info(X) = - \sum_{j=1}^n \left| \frac{T_j}{T} \right| \times \log_2 \left( \frac{|T_j|}{|T|} \right) \quad (3)$$

$$gain\ ratio(X) = \frac{info\ gain(X)}{split\ info(X)} \quad (4)$$

Após a variável com o maior ganho de informação normalizada ser escolhida para tal partição, o algoritmo segue repetindo o procedimento com os subgrupos resultantes de cada divisão. Em pseudocódigo (simplificação), o algoritmo seria este:

1. Para cada atributo  $a$ , encontrar o ganho de informação normalizado pelo particionamento em  $a$
2. Seja  $a_{\text{maior}}$  o atributo com maior ganho da informação normalizado
3. Criar um nó de decisão que particiona o conjunto de dados em  $a_{\text{maior}}$
4. Repetir a operação nos subconjuntos obtidos através da divisão em  $a_{\text{maior}}$ , e adicionar aqueles nós como filhos de nó (Kotsiantis et al, 2007).

Note-se que o algoritmo C4.5 pode utilizar o mesmo atributo mais de uma vez numa mesma árvore, em locais diferentes do esquema. Ademais, há um mecanismo de poda (*pruning*) para impor um limite ao procedimento de partições, para tentar evitar que as árvores não fiquem tão complexas e por demais ajustadas às bases de treino<sup>21</sup>.

Árvores de classificação oferecem como vantagem o fato de não requererem nenhuma padronização ou normalização para atributos numéricos. Além disto, alguns algoritmos, como o C4.5, são capazes de combinar preditivos categóricos com numéricos para realizar classificações. Outra vantagem é que o “raciocínio” das árvores é mais facilmente interpretado do que o de outros algoritmos, oferecendo assim um componente explicativo interessante, além da função meramente preditiva. Este componente explicativo possibilita a análise condicional das variáveis (por exemplo, um nível  $i$  de inflação pode ser um bom sinal de *default* desde que a relação dívida/PIB esteja acima de determinado patamar), em contraste com uma análise isolada.

Em relação a desvantagens, registramos que árvores muito grandes podem perder a facilidade de interpretação. Ademais, há classificadores mais potentes no universo de *machine learning*, sendo que as árvores classificatórias facilmente sofrem de *overfitting* (adequação excessiva à base de treino, impedindo bons resultados em amostras externas).

---

<sup>21</sup> Uma explicação mais detalhada do algoritmo se encontra em Quinlan (1993).

Registra-se, ainda, que o algoritmo C4.5, como a maioria dos algoritmos de *machine learning*, possui hiperparâmetros, que são parâmetros externos ao modelo. Também chamados de *tuning parameters*, não há uma fórmula analítica disponível para calcular um valor apropriado para eles (KUHN, JOHNSON, 2013). Diferentes hiperparâmetros oferecem resultados diferentes, tanto na precisão do modelo quanto nas árvores geradas. Os hiperparâmetros ideais dependem da base de dados utilizada e das finalidades do experimento. Como a base de treinos utilizada neste trabalho variou a cada implementação do algoritmo, sendo incorporado um ano a mais de observações, conforme explicado alhures, optou-se por manter as configurações padrões do algoritmo no ambiente Weka, para fins de uniformidade e de melhor avaliação do algoritmo. A única modificação ocorrida foi a alteração do parâmetro que controla o número mínimo de observações nos nós terminais (*MinNumObj*), que foi reduzido de 2 (padrão) para 1 nos modelos de previsão por continente, por se tratar de bases de dados bem menores do que as usadas no modelo global. Ademais, para tornar as árvores mais simples, não foram permitidas repartições múltiplas (isto é, em vários subgrupos), apenas binárias.

### 3.2.4 Descrição da base de dados

A base de dados do experimento foi coletada e concedida pelo pesquisador Roberto Savona, que a usou nos artigos “A data-driven explanation of country risk: emerging markets vs. eurozone debt crises” e “Fitting and Forecasting Sovereign Defaults using Multiple Risk Signals”. Em verdade, a base de dados atualiza e estende outra base de dados coletada por Manasse e Roubini (2009). Uma descrição detalhada da base de dados é fornecida no apêndice do segundo artigo citado (ver SAVONA, VEZZOLI, 2015).

Trata-se de observações anuais de 66 países emergentes somados a Grécia, Irlanda, Portugal e Espanha. A base de dados utiliza a definição de *default* de Manasse, Roubini e Schimmelpfennig (2003), em que, além de episódios técnicos de inadimplência, inclui-se episódios em que o país teve acesso a largos empréstimos do tipo não-preferencial do FMI excedendo em 100% as suas cotas no fundo. Assim, episódios de *default* que só foram evitados por conta de auxílios robustos do FMI são inseridos na categoria. Abaixo podemos ver a lista de episódios de *default* da base de dados:

**Tabela 1: Relação de episódios de *default* da base de dados**

Ano	Nº de Crises	Países
1975	2	Quênia, Zimbábue
1976	1	Peru
1977	2	Jamaica, México
1978	4	Egito, Peru, Turquia, Zâmbia
1979	5	Honduras, Quênia, Malawi, Nicarágua, Maurícia
1980	8	Bangladesh, Bolívia, Costa Rica, Coréia, Madagascar, Marrocos, Paquistão e Filipinas
1981	10	República Dominicana, El Salvador, Etiópia, Honduras, Índia, Polônia, Jamaica, Romênia, Tailândia e Zâmbia
1982	10	Argentina, Equador, Haiti, Hungria, Quênia, Malawi, México, Nigéria, Peru e Turquia
1983	12	Brasil, Burkina Faso, Chile, Coréia, Maurícia, Nigéria, Filipinas, Serra Leoa, Uruguai, Venezuela, Zimbábue e Zâmbia
1984	1	Egito
1985	3	Camarões, África do Sul e Tailândia
1986	7	Bolívia, Gabão, Madagascar, Marrocos, Paraguai, Serra Leoa e Romênia
1987	2	Jamaica e Uruguai
1988	3	Malawi, Trinidad e Tobago e Tunísia
1989	2	Jordânia e África do Sul
1990	1	Uruguai
1991	3	Algéria, Etiópia e Hungria
1992	1	Zimbábue
1993	1	África do Sul
1994	4	Quênia, Lituânia, Filipinas e Hungria
1995	2	México e Venezuela
1996	3	Jordânia, Cazaquistão e Moldova
1997	5	Indonésia, Coréia, Serra Leoa, Sri Lanka e Tailândia
1998	6	Argentina, Brasil, Moldova, Paquistão, Filipinas e Ucrânia
1999	4	Equador, Gabão, México e Turquia
2000	3	Argentina, Uruguai e Zimbábue
2001	1	Brasil
2002	6	Gabão, Indonésia, Moldova, Paraguai, Turquia e Uruguai
2004	1	Camarões
2005	1	Venezuela

2008	5	Equador, Hungria, Letônia, Paquistão e Ucrânia
2010	3	Grécia, Irlanda e Jamaica

Tradução de tabela de Savona e Vezzoli (2015), apêndice A.

O período abrange desde 1975 até 2010, sendo a fonte dos dados os órgãos estatísticos do Banco Mundial, do FMI, do índice *Freedom House* e da agência de rating Standard & Poor's. Os valores foram transformados pelos referidos autores em bilhões de dólares de 2013.

As variáveis incluem medidas de liquidez, fatores macroeconômicos, históricos, políticos e econômicos e estão com uma defasagem de um ano em relação aos *defaults*, posto que o objetivo é prever estes com um ano de antecedência – estudos de crises moratórias no geral usam o horizonte de um ano, enquanto os de crises cambiais tipicamente usam um período de dois anos (FUERTES, KALOTYCHOU, 2007). Ademais, como o objetivo da base de dados é prever a entrada em crises moratórias e não a permanência em *default*, observações subsequentes a uma crise foram excluídas. Isto é, a condição para a inclusão de um episódio de *default* foi que no ano anterior o país não estivesse nesta condição.

A base de dados inclui ainda *dummies* para registrar se o petróleo é a principal exportação do país, se o país tem acesso aos mercados de capitais internacionais e se houve empréstimo do FMI. As duas primeiras *dummies* são as únicas que não contêm o *lag* de um ano referido. Ademais, o histórico de *default* é contado a partir de 1970.

Segue a lista de variáveis da base de dados utilizada:

**Tabela 2: Lista das variáveis da base de dados utilizada**

Variável	Mínimo	Máximo	Média <sup>22</sup>	Desvio Padrão
Região	NA	NA	NA	NA
Histórico de <i>default</i>	0	5	1	1,36
Dívida externa/PIB	0,14	223,21	49,01	32,34
Dívida externa/Exportações	2,62	3620,63	216,26	251,11
Serviço da dívida de longo prazo/Reservas	0,00	32,02	0,74	1,59
Dívida de curto prazo/PIB	0,00	58,07	6,03	6,21
Dívida de curto prazo/Reservas	0,00	3239,99	94,37	94,38
Dívida pública/PIB	1,03	249,36	51,97	35,87
Variação anual das reservas	-93,33	3024,06	22,24	111,50

<sup>22</sup> Em anexo, encontram-se gráficos com as distribuições.

(%)				
Abertura Comercial (Exportações + Importações/PIB)	10,98	220,41	71,55	35,25
Varição anual do PIB (%)	-42,45	39,49	3,86	5,42
Saldo da Conta Corrente/PIB	-39,67	39,58	-2,88	6,67
Exportações	0,03	525,29	22,89	47,56
FMI	NA	NA	NA	NA
Varição anual das Exportações (%)	-47,3	126,68	6,09	11,61
Varição anual das Importações (%)	-47,1	117,96	6,49	13,31
Varição anual do Investimento Estrangeiro Direto (IED) (%)	121956,72	49975,75	54,40	3288,92
Taxa anual de inflação	-100	24411,03	41,41	621,81
M2/Reservas	0,00	7427,44	40,27	415,95
Investimento Estrangeiro Direto (IED)/ PIB	-15,03	52,05	2,41	3,46
Freedom House Index	NA	NA	NA	NA
Varição cambial (taxa residual de câmbio sobre tendência linear)	-9937,77	7466,71	-22,81	416,64
Acesso ao mercado de capitais internacionais	NA	NA	NA	NA
Petróleo	NA	NA	NA	NA
Juros dos EUA	0,15	14,04	5,36	3,02

**Fonte:** Elaboração do autor a partir da base de dados.

## 4 ANÁLISE DOS RESULTADOS

Conforme exposto no capítulo anterior, nesta pesquisa, o algoritmo C4.5 foi treinado sempre com a base de dados contendo os anos anteriores ao ano cujas observações serviram de teste. Ressalta-se ainda que as informações de cada ano contam com um ano de atraso. Assim, as variáveis usadas para prever a ocorrência ou não de *default* no ano de 1979, por exemplo, são registros do ano de 1978.

Desta maneira, foram realizadas diversas previsões, uma para cada ano. A cada previsão anual, uma árvore foi produzida, de modo que não seria possível apresentar, neste trabalho, todas as árvores geradas com o algoritmo.

Além disto, houve uma tendência do algoritmo de gerar árvores cada vez mais complexas com o aumento da base de treino, sendo as últimas árvores das séries muitas vezes grandes demais para reprodução e interpretação neste trabalho. Realiza-se, portanto, uma análise exploratória de alguns esquemas gerados, comentando-se de forma geral as árvores geradas no experimento.

As previsões foram realizadas tanto com a base de dados inteira quanto com as observações divididas por região (América Latina, Ásia e África). O algoritmo apresentou aprendizado muito fraco para a Europa, provavelmente devido à baixíssima incidência de *defaults* naquele continente, ainda maior que nos outros - na base de dados trabalhada, havia apenas dez episódios de 1975 a 2010. Assim, os resultados não são reportados. Os demais resultados obtidos são a seguir descritos e analisados.

### 4.1 MODELO GLOBAL

Inicialmente, tentamos prever as crises moratórias treinando e testando o algoritmo com os dados de todos os países disponíveis. O algoritmo apresentou bom desempenho ao tentar prever os *defaults* de 1979 e da década de 1980, conforme demonstram os valores AUC (abaixo). Contudo, os resultados pioram a partir do ano de 1993. Deste ano em diante, poucos *defaults* são precisamente previstos, o que ocorre junto a uma taxa alta de alarmes falsos. Este bom desempenho na década de 1980 seguido de piora nas décadas seguintes foi uma tendência ocorrida também nos modelos continentais. As tabelas abaixo sumarizam os resultados encontrados no modelo treinado com a base de dados global:

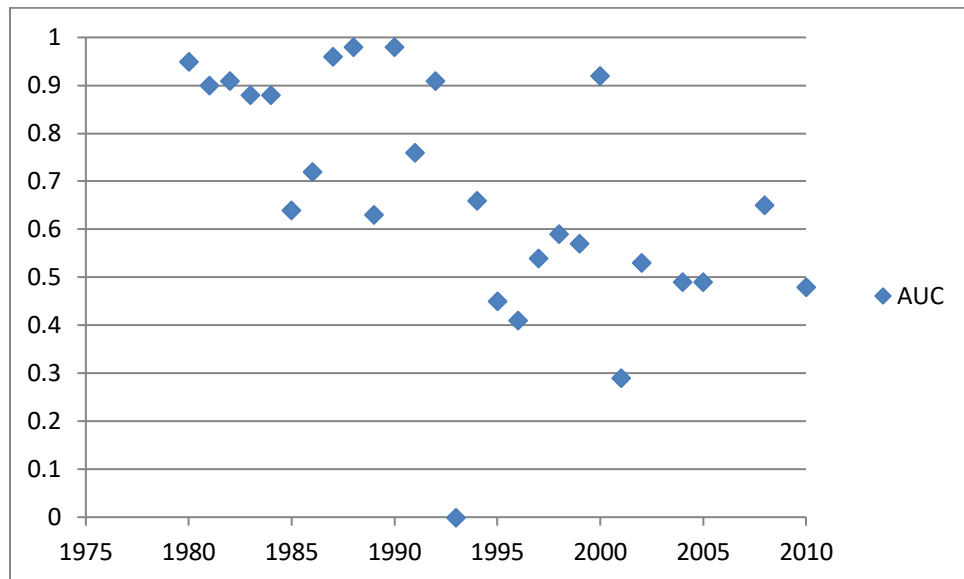
**Tabela 3 – Resultados do Modelo Global**

Ano	N	Número de defaults	Defaults corretamente previstos	Defaults incorretamente previstos	Acurácia Geral	AUC*
1979	57	5	5	4	92,98%	0,95
1980	55	8	8	6	89,09%	0,95
1981	55	10	10	9	83,63%	0,90
1982	51	10	10	6	88,23%	0,91
1983	43	12	11	7	85,79%	0,88
1984	38	1	1	11	71,05%	0,88
1985	41	3	2	15	60,97%	0,64
1986	41	7	6	14	63,41%	0,72
1987	36	2	2	4	88,88%	0,96
1988	36	3	3	8	77,77%	0,98
1989	36	2	1	6	80,55%	0,63
1990	40	1	1	5	87,5%	0,98
1991	46	3	2	5	86,95%	0,76
1992	46	1	1	6	86,95%	0,91
1993	49	1	0	9	79,59%	0,0
1994	52	4	2	7	82,69%	0,66
1995	57	2	0	6	85,96%	0,45
1996	58	3	0	8	81,03%	0,41
1997	61	5	1	7	81,96%	0,54
1998	61	6	2	7	81,96%	0,59
1999	60	4	1	7	83,33%	0,57
2000	61	3	3	8	86,88%	0,92
2001	62	1	0	14	77,41%	0,29
2002	61	6	1	1	90,16%	0,53
2003	66	0	-	2	96,96%	-
2004	67	1	0	1	97,01%	0,49
2005	68	1	0	1	97,05%	0,49
2006	70	0	-	2	97,14%	-
2007	70	0	-	5	92,85%	-
2008	70	5	1	4	85,71%	0,65
2009	65	0	-	7	89,23%	-
2010	70	3	0	2	92,86%	0,48



\* Média ponderada entre as duas classes. Fonte: Elaboração do autor

**Figura 5 – AUC x Ano – Modelo Global**



**Fonte:** Elaboração do autor

Como se nota, até o ano de 1992, o algoritmo foi capaz de detectar 63 de 68 (92,6%) crises moratórias contidas na base de dados. Isto foi feito a uma taxa razoável de alarmes falsos – foram contados 114 neste período – 18,3% das observações. Contudo, a partir do ano de 1992, os resultados do algoritmo de árvores pioram sensivelmente, de modo que, considerado o período todo, o esquema foi capaz de prever 68% dos *defaults*, emitindo alarmes falsos 34% das vezes. Porém, diversos destes alarmes não podem ser considerados propriamente falsos positivos tal como considerado em ciências biomédicas, conforme demonstraremos adiante.

Primeiramente, faz-se necessário notar que a comparação dos resultados com a literatura não pode ser feita de maneira direta, pois os estudos são feitos, em geral, com diferentes bases de dados, cada com uma com suas variáveis e sua cobertura em termos de países e de intervalo temporal (FUERTES e KALOTYCHOU, 2007). Além disso, os testes fora da amostra às vezes são de períodos históricos diferentes e variam em extensão, havendo muita variância nos resultados de bases de teste menores.

Feitas estas ressalvas, analisam-se outros resultados da literatura, atentando-se sempre para os resultados em bases independentes (“*out-of-sample*”).

Savona e Vezzoli (2015), com o método *Noise-to-Signal*, obtiveram, fora da amostra, uma AUC de 0,66 e Dawood, Horsewood e Strobel (2017) obteve 0,36 como melhor relação sinal-ruído, usando a relação reservas/PIB.

Ciarlone e Trebeschi (2005), com o modelo logit, previram 80% das crises de 2000 a 2002 emitindo alarmes falsos 20% das vezes. Fuertes e Kalotychou (2007), também com o modelo logit, previram 81,3% das crises de 1996 a 2002 a uma taxa de alarmes falsos de 22,6%.

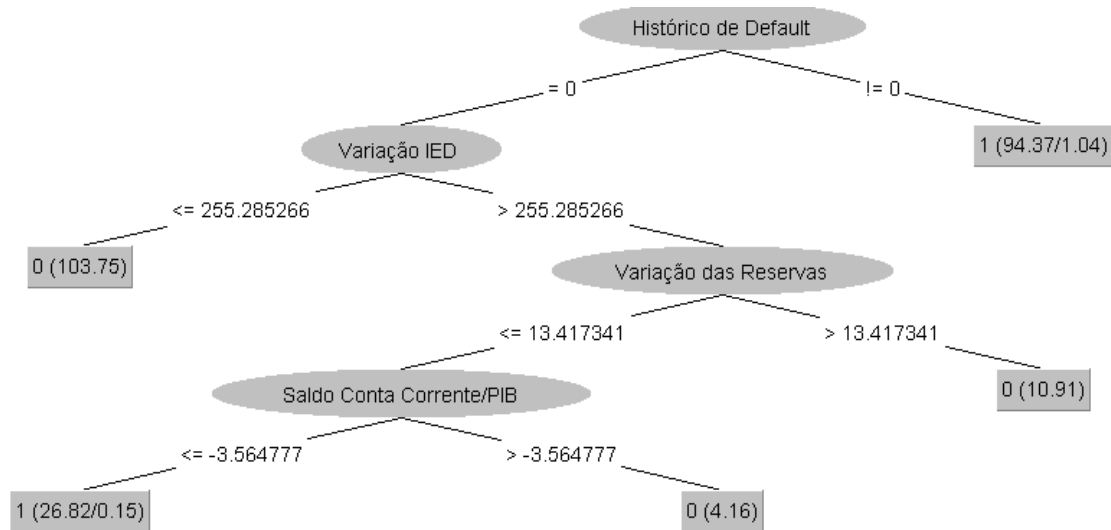
Manasse e Roubini (2009), com a mesma base de dados que foi estendida e usada neste trabalho, previram 57% das crises de 1990 a 2002 a uma taxa de alarmes falsos de 19%, com o algoritmo CART de árvores decisórias. Savona e Vezzoli (2015), por sua vez, também utilizando-se do algoritmo CART, aprimorado e agregado pelo algoritmo CRAGGING (*Cross-validation Aggregating*), conseguiram prever 88% das crises de 1991 a 2010 da base de dados deste trabalho, porém a uma especificidade de 64%, o que mostra que foram emitidos alarmes falsos para 36% das amostras de não *default*. Estes autores reportam ainda AUC de 0,83.

Manasse, Roubini e Schimmelpfennig (2003) previram, com modelo logit e com o algoritmo CART, respectivamente, 69% e 85% das crises 1990 a 2002 a uma taxa de alarmes falsos de 5% e 15%. Finalmente, Fioramanti (2008) utilizando um algoritmo de rede neurais previu com sucesso 88% das crises de 1990 a 2001 a uma taxa de alarmes falsos de 15%.

Conforme se pode ver, até o início da década de 90, o algoritmo C4.5 conseguiu resultados bons, comparados com a literatura; declinando, todavia, nas previsões seguintes. A literatura de *machine learning* registra que, por vezes, o algoritmo C4.5 não tem um método de poda das árvores tão eficaz, de modo que, quanto mais instâncias são adicionadas à base de treino, mais as árvores crescem, mesmo quando isto não aumenta a performance da árvore em amostras independentes (Witten et al, 2016). Isto foi precisamente o que ocorreu na presente pesquisa: as árvores foram ficando por demais complexas, perdendo parte da principal vantagem do método, que reside na facilidade de interpretação. No presente caso, o aumento de complexidade, em geral, não serviu para prever devidamente as crises posteriores à década de 1980.

Em relação às árvores geradas, abaixo é possível ver o primeiro produto do algoritmo, usado para prever crises do ano de 1979. As terminações sinalizadas com “0” significam que não há previsão de *default*, enquanto as sinalizadas com “1” indicam a previsão do evento.

**Figura 6: Árvore gerada para previsão de crises do ano de 1979 – Modelo Global**



**Fonte:** Elaboração do autor

Desde já é possível ver um padrão que se repetiu na maioria das árvores geradas: o histórico de *default* como primeiro atributo a dividir as observações. Nos anos de 1979, 1980 e 1981, o algoritmo previu *default* sempre que o histórico de *default* dos países era diferente de zero. Com o aumento da base de treino, a tendência do algoritmo foi gerar árvores cada vez mais complexas, de modo que, a partir da previsão para o ano de 1982, o histórico de *default* passou a ser conjugado com outras variáveis, conforme mostraremos abaixo.

Ademais, nota-se, para o esquema acima, a importância da balança de pagamentos como previsor de *defaults*, o que é indicado pela escolha das variáveis de variação anual do investimento estrangeiro direto (“Variação IED”, em percentual) e pelo saldo da conta corrente em relação ao PIB.

A respeito destes, cabe registrar que a variação de IED é usada pelo algoritmo de maneira contraintuitiva, considerada a teoria econômica, pois associa menores influxos de IED à não ocorrência de *default*. De fato, o algoritmo eventualmente toma decisões que, embora sejam transparentes, geram dúvidas quanto a sua lógica, o que pode sinalizar tanto um aprendizado fraco quanto um excesso de ajuste à base de treinos. Contudo o algoritmo também pode estar estabelecendo relações com alguma racionalidade econômica que não é usualmente considerada.

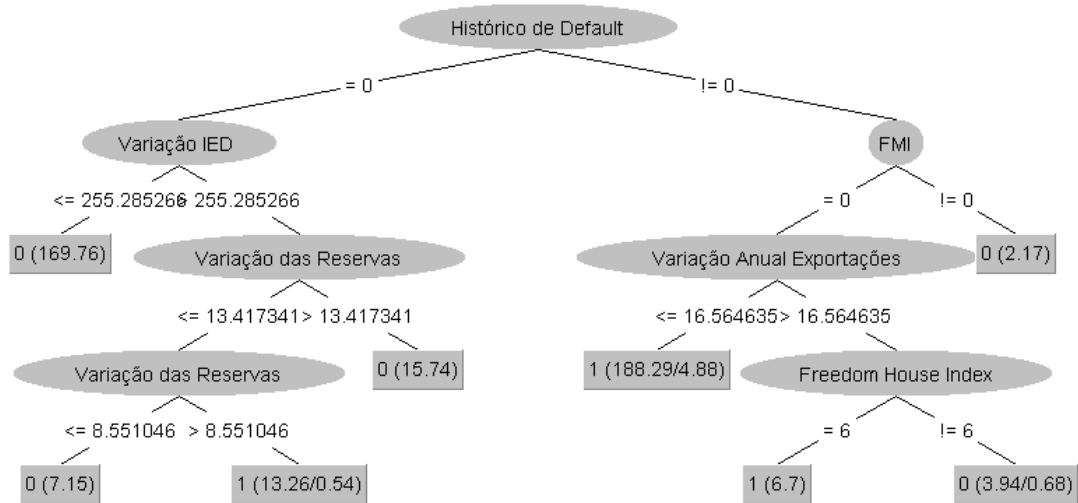
No caso em comento, nota-se que o limiar de variação de IED estabelecida pelo algoritmo é bastante alta (255%). Neste sentido, uma racionalidade possível seria a de que excessivos influxos de IED podem sinalizar alguma bolha ou apreciar excessivamente a moeda, de modo a causar problemas cambiais e moratórios quando o sentimento do mercado muda. Por exemplo, em 2007, a variação de IED da Hungria em relação ao ano anterior foi de 262,88%; em 2008, após a crise, o país sofreu forte desvalorização cambial e teve de ser socorrido pela União Europeia e pelo FMI com um pacote de 25 bilhões de dólares (o que foi previsto pelo algoritmo, que sinalizou *default*).

Voltando ao esquema gerado, nota-se que ele também escolhe o atributo de variação das reservas, o que indica, por sua vez, que a deterioração das reservas foi outro sinal considerado precedente de *defaults*, o que é condizente com a teoria econômica, em especial a que enfatiza o papel de *sudden stops* em crises moratórias, encontrando amparo também em pesquisas empíricas – neste sentido, ver, por exemplo, Dawood, Horsewood e Strobel (2017).

Em relação aos *defaults* previstos, a simples árvore gerada conseguiu prever com sucesso as crises moratórias de Honduras, Quênia, Malawi, Nicarágua e Ilhas Maurício. Foram previstas também crises no México, Egito, Serra Leoa e Zâmbia. Neste sentido, é de se registrar que de fato foram registrados *defaults* na Zâmbia, em 1981, e no México, em 1982. Por sua vez, crises moratórias ocorreram em Serra Leoa em 1983 e no Egito em 1984. Nota-se, assim, que se o horizonte de previsão fosse estendido por mais alguns anos, é possível que o algoritmo não tivesse incorrido em erro.

De fato, conforme argumentam Fuertes e Kalotychou (2007), escolher um horizonte de previsão envolve um *tradeoff*: quanto maior o horizonte, menores são os *defaults* perdidos, mas maiores são os falsos alarmes. Assim, o melhor horizonte depende das prioridades e finalidades de cada situação. Alarmes falsos podem significar oportunidades perdidas para investidores e políticas custosas e desnecessárias; por outro lado, algumas repercussões notáveis de crises moratórias não evitadas são prejuízos a investidores, declínio do produto, aumento do desemprego e das taxas de pobreza (FUERTES, KALOTYCHOU, 2007). Fuertes e Kalotychou (2007) ainda pontuam que estudos de crises cambiais costumam usar um horizonte de previsão de dois anos, enquanto os estudos de crises moratórias costumam usar o horizonte de um ano, a exemplo do que foi feito neste trabalho.

**Figura 7: Árvore gerada para previsão de crises do ano de 1982 – Modelo Global**



**Fonte:** Elaboração do autor

A figura acima exibe a árvore gerada para prever crises no ano de 1982. Conforme se nota, a partir deste ano, o simples fato de ter havido algum *default* anterior já não foi mais suficiente para o algoritmo prever a repetição do evento.

O algoritmo passa a conjugar o referido histórico com outras informações, como a existência de empréstimo junto ao FMI no anterior, que o algoritmo considera um sinal de incoerência de *default*<sup>23</sup>. O esquema gerado usa ainda, no lado direito da árvore, a variação das exportações, estabelecendo-se o patamar de crescimento de 16%. Países com variação abaixo deste limiar, no ano de 1981, que havia histórico de *default* e que não tomaram empréstimo junto ao FMI, foram previstos em moratória pelo algoritmo.

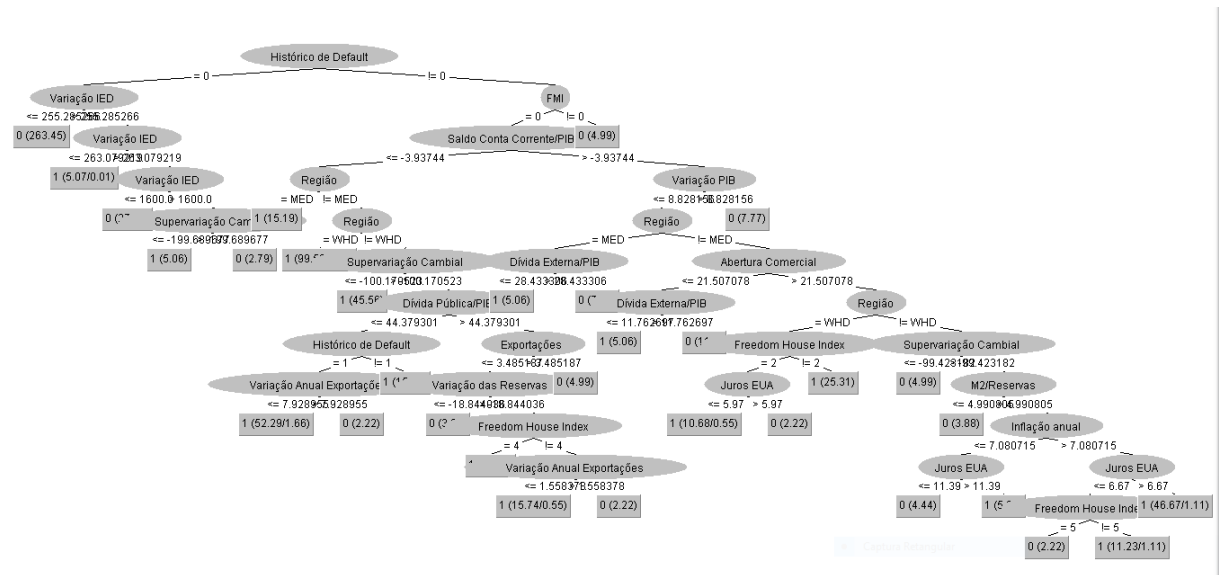
Por sua vez, o lado esquerdo da árvore demonstra outra característica do algoritmo, compartilhada por outros métodos de *machine learning*: o estabelecimento de relações não-lineares. Por exemplo, o esquema prevê, nos últimos dois nós, *default* para países com variação das reservas entre 8,55 e 13,42%, mas não para observações superiores e inferiores ao intervalo. Esta capacidade, se por um lado permite aos algoritmos que percebam nuances que a análise humana dificilmente captaria, por outro lado aumenta a chance de *overfitting*, que ocorre quando o modelo se ajusta tanto à base de dados com a qual ele foi criado que perde o poder de fazer previsões para observações independentes, pois cria regras que só

<sup>23</sup> Ressalta-se aqui, conforme dito, que na base de dados utilizada, considera-se que, se o país teve acesso a empréstimos do tipo não preferencial do FMI excedendo em 100% as suas cotas no fundo, houve *default*. Assim, os empréstimos da *dummy* FMI são aqueles inferiores a este patamar.

servem para a base de treinamento. É precisamente para tentar evitar este efeito que todos os resultados reportados precisam decorrer da base de testes, a exemplo do que se faz aqui.

Conforme dito, em todos os modelos, houve uma tendência de aumento da complexidade das árvores geradas conforme se aumentava a base de treinos. A árvore gerada para prever *defaults*, em países do mundo todo, em 1990, dificilmente cabe em uma figura (o esquema detalhado da árvore abaixo se encontra no anexo deste trabalho):

**Figura 8: Árvore gerada para previsão de crises do ano de 1990 – Modelo Global**



**Fonte:** Elaboração do autor

No ano de 1990, o algoritmo conseguiu prever a única crise existente na base de dados, qual seja, do Uruguai. Foram também previstas, incorretamente, crises na África do Sul, em Trinidad e Tobago, Tunísia, Hungria e Egito. Contudo, vale observar que em dois países destes de fato houve *default* na sequência – África do Sul, em 1993, e Hungria, em 1991.

As árvores geradas para os anos seguintes ao ano de 1990 são tão complexas quanto a figura acima ou ainda mais. As previsões foram se deteriorando, conforme demonstram os valores AUC da Tabela 3. No ano de crise de 2008, por exemplo, de 5 crises, apenas a crise na Hungria foi corretamente prevista.

Várias razões podem ter contribuído para a diminuição do desempenho do algoritmo. A um, conforme já dito, a literatura aponta que, por vezes, o algoritmo C4.5 aumenta demais a complexidade das árvores com o aumento da base de treinos, isto é, não tem um mecanismo eficiente de poda, o que pode levar a previsões piores. A dois, as próprias variáveis que serviram bem para a previsão de *defaults* na década de 80, como o histórico de *default* e os

juros norte-americanos, podem não ter servido mais tão bem para prever crises posteriores, ou os patamares decisórios apreendidos pelo algoritmo ficaram defasados.

De fato, as crises moratórias ficaram bem mais escassas a partir de meados da década de 1990, o que atrapalha o aprendizado do algoritmo, ainda que seja usado o filtro de balanceamento de classes. Historicamente, é oportuno ressaltar, como explicação para tal, dentre outros, a criação do Plano Brady, de reestruturação das dívidas de diversos países emergentes, com início em 1989 e implementação ao longo dos primeiros anos da década de 1990, beneficiando diversos “inadimplentes em série”: Argentina, Brasil, Bulgária, Costa Rica, República Dominicana, Equador, México, Marrocos, Nigéria, Filipinas, Polônia, Uruguai e Venezuela.

Diferentemente de planos anteriores, como o Plano Baker, os Estados Unidos entraram diretamente nas negociações, procurando evitar a ocorrência do *free rider* - isto é, de bancos que só se beneficiariam das concessões feitas por outros bancos, sem renegociar significativamente seus créditos (SILVA, CARVALHO, MEDEIROS, 2009). O foco do Plano Brady era reestruturar as dívidas para um horizonte de longo prazo, usando instrumentos de mercado para isto, de modo que soluções diferentes foram adotadas em cada país.

No Brasil, as principais características de sucesso do plano firmado em abril de 1994, conforme Silva, Carvalho e Medeiros (2009), foram a existência de garantias (a serem efetivadas pela compra de títulos do Tesouro dos Estados Unidos) para os novos títulos, e, principalmente a possibilidade de o Brasil fazer operações com os novos títulos emitidos. Esta última característica iniciou uma nova fase de endividamento externo de nosso país em que, a despeito das sucessivas crises internacionais vindouras, o país continuou acessando o mercado externo com flexibilidade de escolha, tanto em relação ao momento para efetivar a operação quanto ao próprio mercado para que seriam emitidos os títulos e os instrumentos utilizados (SILVA, CARVALHO, MEDEIROS, 2009). Foi após o Plano Brady que o país conseguiu, pela primeira vez, emitir títulos com prazo de 30 anos (SILVA, CARVALHO, MEDEIROS, 2009).

Assim, como se nota, no início da década de 90, as dívidas nacionais foram reestruturadas e alguns países obtiveram maior flexibilidade na emissão de títulos, o que, junto com reformas estruturais dos países emergentes e um ambiente global de juros menores, ajudou a reestabelecer o acesso destes países aos mercados de crédito para rolagem de suas dívidas (CLARK, 1993). Com isto, é possível conjecturar que as condições para crises moratórias tenham sido alteradas, prejudicando o algoritmo.

Em relação às variáveis mais importantes para as previsões do algoritmo, cabe pontuar que não há uma medida universalmente aceita de relevância das variáveis, ainda que várias diferentes tenham sido propostas (Witten et al, 2016). A explicabilidade (*explainability*) de *machine learning* é uma área em desenvolvimento e diferentes métodos são empregados atualmente.

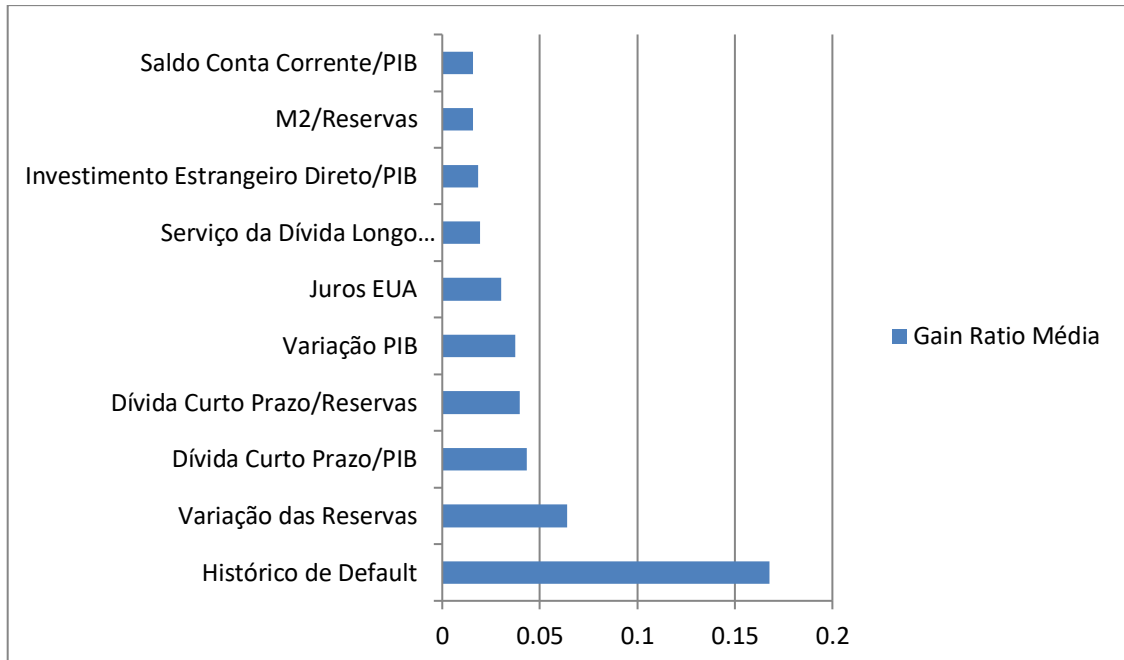
Em árvores decisórias, em geral, os atributos mais relevantes para a previsão encontram-se na parte de cima da árvore, sendo este o método usado para medir relevância, por exemplo, pelo documento de trabalho 12/22 do FMI (GHOSH et al, 2014). Todavia, a visualização da árvore nem sempre pode ser considerada conclusiva: pode haver variáveis que dividem o grupo de observações (ou algum subgrupo dele) de maneira semelhante. Neste caso, uma variável será omitida, mas sua ausência não significará falta de importância do atributo. Ademais, cabe registrar que as árvores de classificação são modelos instáveis, no sentido de que pequenas alterações nos hiperparâmetros podem gerar esquemas com variáveis diferentes. (Neste trabalho, contudo, os parâmetros foram mantidos constantes, conforme dito.)

A fim de se contornar estas questões, uma alternativa complementar, adotada por este trabalho, é a de se calcular, para todos os atributos, a principal métrica de partição do algoritmo C4.5, a fim de se ver, em potencial, a relevância de cada um. Trata-se da taxa de ganho de informação normalizada (“*gain ratio*”), mencionada na seção 2.2.3, que mede qual a proporção de informação gerada pela divisão através de um atributo  $X$  que tem utilidade para a tarefa de classificação almejada (isto é, que aparenta auxiliar na divisão entre *defaults* e não *defaults*). Diferentes bases de treino podem apresentar diferentes *gain ratio* para cada atributo. Portanto, como os dados de treino foram alterados a cada ano de previsão, a *gain ratio* de cada atributo foi recalculada a cada nova previsão, extraindo-se, ao final, uma média da taxa, a qual pode ser vista no gráfico abaixo.

Vale observar, contudo, que o ranking é feito com a base de treinos completa, mas não se avalia a taxa de ganho de informação normalizada (*gain ratio*) decorrente de partições dentro de subgrupos gerados pela árvore. Isto explica a razão pela qual mesmo as variáveis utilizadas pelas árvores não foram apontadas como relevantes na medida abaixo. É dizer, o ranking abaixo avalia a capacidade de discriminação das variáveis considerando a base de treinos como um todo, mas há variáveis que podem funcionar melhor em bases de treino já reduzidas por alguma divisão criada com outra variável. Em outras palavras, o ranking não é capaz de antecipar a capacidade preditiva condicional das variáveis, apenas a importância incondicional na base de treino completa. Portanto, ele serve de complemento às árvores.



**Figura 9: Gráfico – Importância das variáveis para as previsões – Modelo Global**



**Fonte:** Elaboração do autor

Conforme se pode ver, o histórico de *defaults* passados foi o atributo mais relevante para se separar os eventos de crise e de normalidade, podendo ser vista também como a informação mais importante para prevê-los, sob a ótica do algoritmo adotado. De fato, conforme exibido acima, o histórico de *default* figurou sempre no topo da árvore, iniciando as partições. Isto corrobora a existência de inadimplentes em série (“*serial defaulters*”), conforme Reinhart e Rogoff (2009a). Para estes autores, riscos de crédito e riscos políticos são os principais fatores para que isto ocorra: países com um histórico de inadimplência têm dificuldades de pegar empréstimos do resto do mundo e, portanto, devem respeitar limites mais estritos de endividamento para evitar crises, sobretudo no que tange à dívida externa.

A variação das reservas mostrou-se a segunda variável mais relevante. Tal fato, conjugado com a aparição da relação entre o saldo da conta corrente e o PIB em diversas árvores, demonstra a relação das crises moratórias com crises cambiais, *sudden stops* e déficits em conta corrente, dialogando com os achados de Cavallo, Eichengreen e Panizza (2018) e Catão e Milesi-Ferretti (2014).

Em seguida, praticamente ao mesmo nível de relevância, aparecem as variáveis de relação da dívida curto prazo com o PIB e com as reservas, bem como a variação do PIB. Por

fim, os juros dos Estados Unidos, que sinalizam os juros globais, também se mostraram uma variável importante, com *gain ratio* média ligeiramente inferior às demais.

Estes resultados são bastante parecidos com os encontrados por Savona e Vezzoli (2015), que usaram a mesma base de dados, mas com método diferente<sup>24</sup>. Estes autores concluíram que o histórico de *default*, a relação da dívida de curto prazo com as reservas, a variação do PIB e a taxa de juros dos Estados Unidos seriam os principais determinantes das crises moratórias analisadas.

Similarmente, através de um modelo de limites, Dawood, Horsewood e Strobel (2017) encontraram como melhores sinais para prever crises, com dois anos de antecedência, a relação reservas/PIB, os juros globais e o saldo na balança de transações correntes (0,33). Através de modelos logit, Mcfadden et al (1985) e Peter (2002) também encontraram o histórico de *default* dos países como uma das variáveis mais relevantes.

#### 4.2 MODELO AMÉRICA LATINA

A fim de se experimentar se a segregação da base de dados por continentes e regiões do globo poderia atingir melhores resultados, bem como verificar se a importância das variáveis seria a mesma, o algoritmo foi testado para gerar modelos regionais. A tabela abaixo reporta os modelos obtidos com dados dos países da América Latina:

**Tabela 4 – Resultados do Modelo América Latina**

Ano	N	Número de <i>defaults</i>	<i>Defaults</i> corretamente previstos	<i>Defaults</i> incorretamente previstos	Acurácia Geral	AUC*
1979	17	2	2	1	94,11%	0,97
1980	16	2	2	3	81,25%	0,89
1981	15	4	3	1	86,66%	0,83
1982	11	5	4	0	90,90%	0,90
1983	6	4	4	0	100%	1,0
1984	2	0	-	0	100%	-
1985	3	0	-	1	66,66%	-
1986	5	2	2	2	60%	0,83
1987	3	2	2	0	100%	1,0

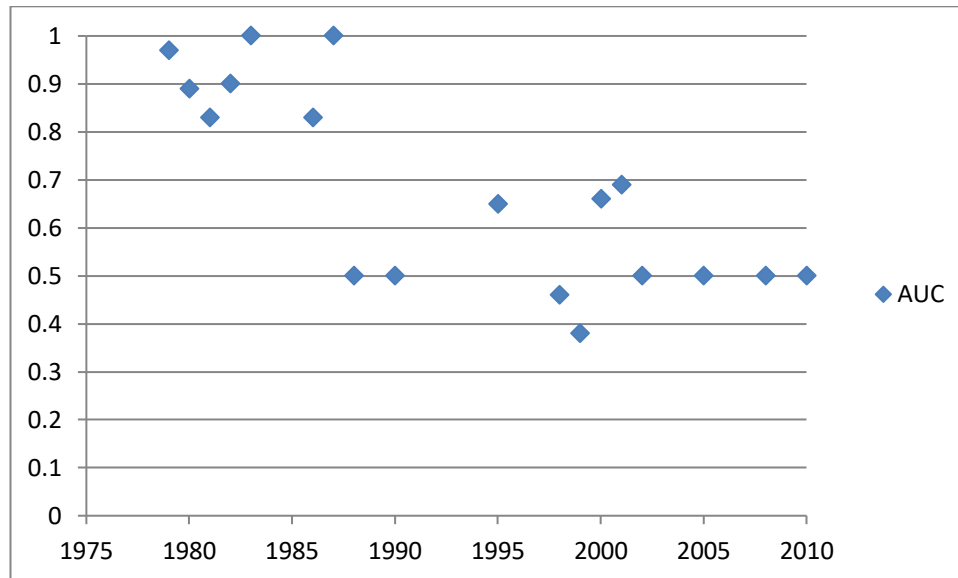
<sup>24</sup> Os referidos autores utilizaram a combinação do algoritmo CRAGGING com o Random Forest, sendo este baseado na agregação de diversas previsões de árvores de classificação que usam o índice de Gini para criar partições.

1988	2	1	1	1	50%	0,5
1989	1	0	-	1	0%	-
1990	2	1	1	1	50%	0,5
1991	5	0	-	4	20%	-
1992	6	0	-	3	50%	-
1993	7	0	-	3	57,41%	-
1994	10	0	-	1	90%	-
1995	12	2	2	4	66,66%	0,65
1996	12	0	-	2	83,33%	-
1997	13	0	-	1	92,31%	-
1998	15	2	0	1	80%	0,46
1999	15	2	0	3	66,66%	0,38
2000	14	2	1	2	78,57%	0,66
2001	14	1	1	11	21,42%	0,69
2002	13	2	0	0	84,61%	0,50
2003	14	0	-	3	78,57%	-
2004	15	0	-	0	100%	-
2005	16	1	0	0	93,75%	0,50
2006	18	0	-	0	100%	-
2007	18	0	-	1	94,44%	-
2008	18	1	0	0	94,44%	0,50
2009	17	0	-	0	100%	-
2010	18	1	0	0	94,44%	0,50

\* Média ponderada entre as duas classes

**Fonte:** Elaboração do autor

**Figura 10 – AUC x Ano – Modelo América Latina**



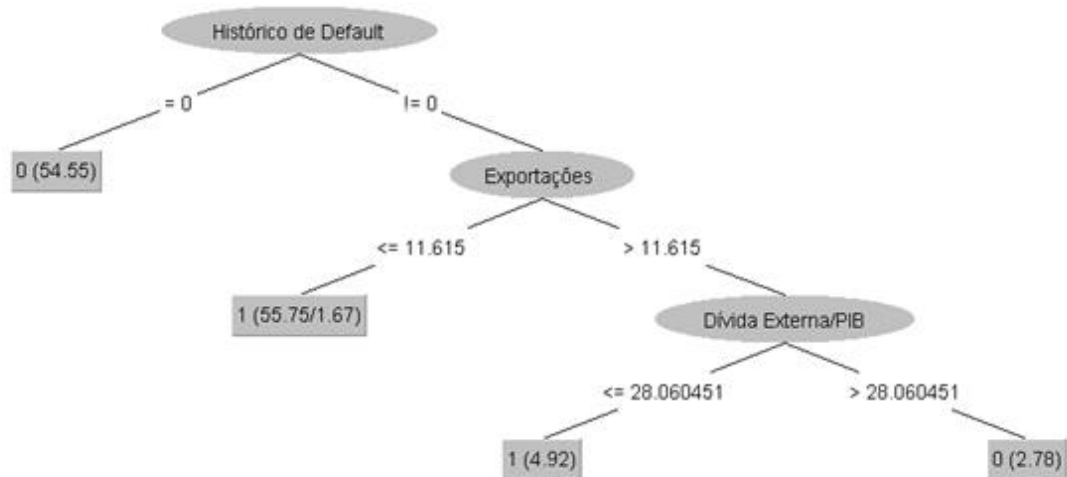
**Fonte:** Elaboração do autor

Um dos desafios de se testar o algoritmo em modelos regionais é que, com base de dados menores, a incidência de crises moratórias é ainda mais baixa, o que impacta no aprendizado do algoritmo e, também, dificulta sua testagem. Com amostras de teste muito pequenas, poucos erros ou acertos podem mudar completamente os resultados obtidos, gerando muita variância. Lembre-se aqui que, como o objetivo do modelo era prever a entrada em *default*, países que entraram em crise e nela permaneceram continuamente não tiveram as observações da continuidade incluídas na base de dados, o que também contribuiu para reduzir ainda mais as observações.

A despeito disto, os resultados acima reportam o desempenho obtido apenas com o treino e teste do algoritmo fazendo uso exclusivo da base de dados referente a países da América Latina. Nota-se a mesma tendência vista no modelo global, isto é, um bom desempenho nas crises da década de 1980, seguida de uma piora em que o algoritmo deixa de discriminar bem as classes de eventos a partir da década de 1990, incluindo os anos 2000.

De 20 *defaults* registrados na década de 1980, o algoritmo previu corretamente 18 (90%). A árvore abaixo retrata o esquema gerado pelo algoritmo, por exemplo, para prever crises no ano de 1982:

Figura 11: Árvore gerada para previsão de crises do ano de 1980 – Modelo América Latina



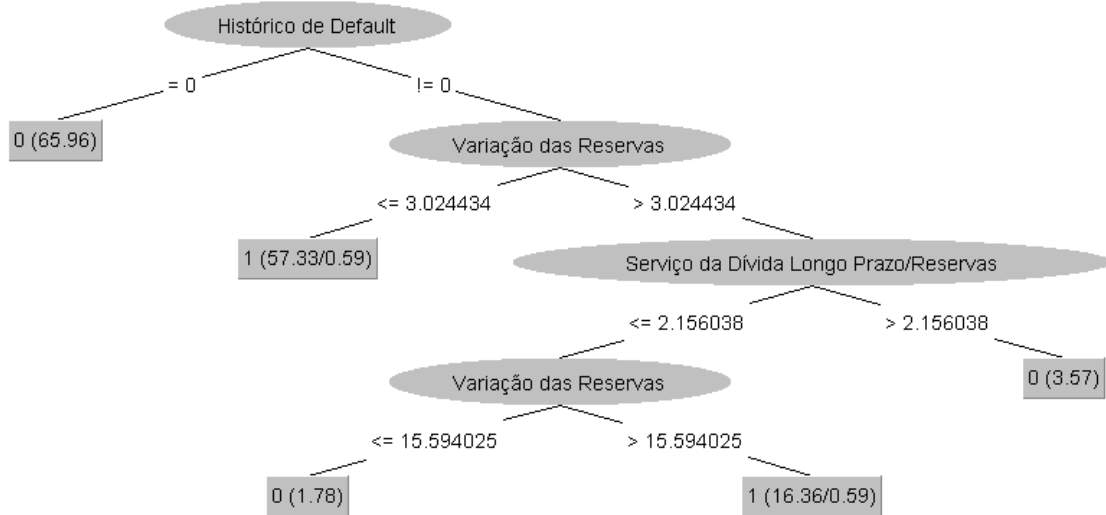
Fonte: Elaboração do autor

Neste ano, no Peru, por exemplo, que tinha histórico de *default*, foi previsto uma crise moratória porque registrou crescimento das exportações abaixo de 11 bilhões de dólares<sup>25</sup>. Por sua vez, o nó abaixo demonstra uma regra que aparenta não fazer sentido e que, de fato, lhe fez incorrer em erro ao não prever o *default* do México neste ano. O algoritmo associou maiores níveis de dívida externa à inoccorrência de *default*, e o México apresentou crescimento das exportações acima de 11,6 bilhões de dólares e relação dívida externa/PIB de 31,27%. Ainda assim, as regras bastante simples aprendidas pelo algoritmo conseguiram prever 4 das 5 crises daquele ano.

No ano de 1987, por sua vez, havia apenas três países na amostra, mas o algoritmo consegue detectar bem, com a seguinte árvore, os dois *defaults* (no Uruguai e na Jamaica):

<sup>25</sup> Valores de 2013.

**Figura 12: Árvore gerada para previsão de crises do ano de 1987 – Modelo América Latina**



**Fonte:** Elaboração do autor

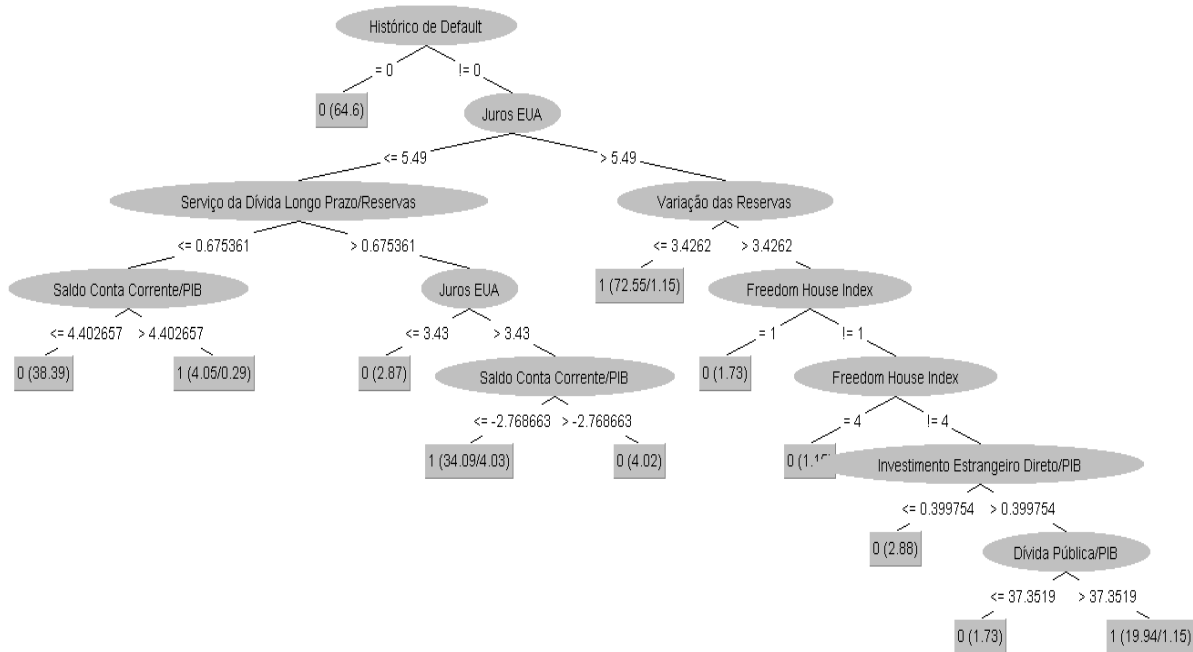
O *default* na Jamaica é previsto a partir do histórico de *default* e da variação das reservas, sendo intuitivo do ponto de vista da teoria econômica. O *default* na Jamaica, por sua vez, tem a compreensão mais difícil, posto que é usada também a taxa do serviço da dívida a longo prazo com as reservas. Números maiores poderiam significar tanto reservas menores quanto serviço da dívida a longo prazo maiores: ambos deveriam aumentar a probabilidade, segundo a racionalidade macroeconômica; contudo, o algoritmo usa a variável de maneira contraintuitiva. Percebe-se, assim, que o fato de o algoritmo ser transparente nem sempre significa que ele será totalmente compreensível. Contudo, algoritmos de árvore ao menos permitem que lógicas questionáveis sejam detectadas, o que não ocorre, por exemplo, com algoritmos de redes neurais.

Nos anos 1990, quase não foram registrados *default* no continente; apesar disto, o algoritmo disparou consideráveis alarmes, conforme demonstra a Tabela 4. Nesta época, conforme fora explanado, diversos países latinos se beneficiaram do chamado Plano Brady, de reestruturação da dívida pública, o que contribuiu para a redução de episódios de crises moratórias.

O esquema abaixo, gerada para prever crises no ano 2000, demonstra a tendência do algoritmo de, com bases de treino maiores, produzir árvores mais complexas (e, no nosso

caso, menos precisas). Conforme se pode ver, à medida que o algoritmo foi expandindo sua base de treinos da América Latina, também a variação das reservas e os juros americanos passam a ter relevância na previsão dos eventos moratórios.

**Figura 13: Árvore gerada para previsão de crises do ano de 2000 – Modelo América Latina**



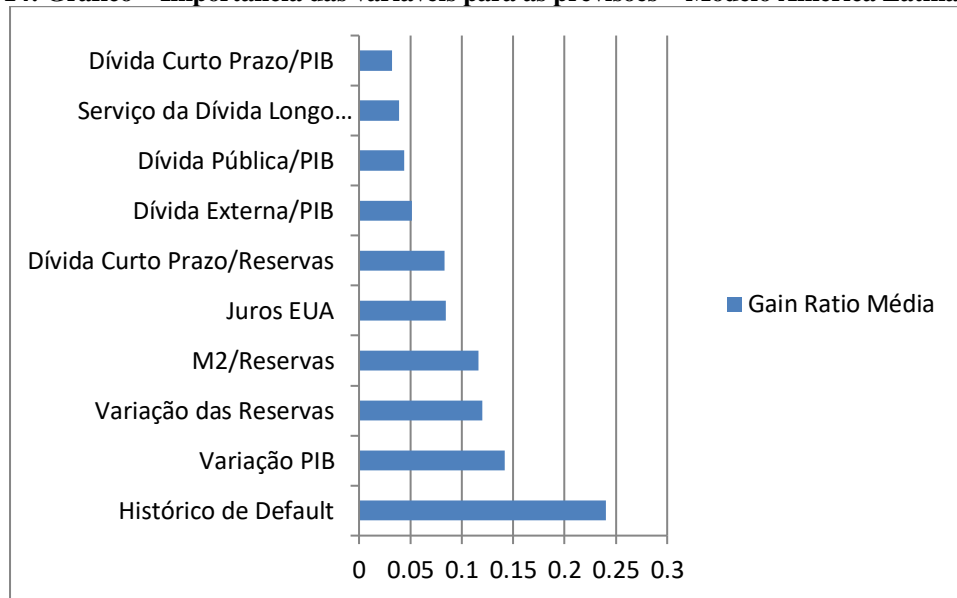
**Fonte:** Elaboração do autor

Em 2000, foram registrados *default* no Uruguai e na Argentina, mas o algoritmo só previu o evento neste segundo país. A árvore prevê ainda *defaults* no México e no Brasil, cabendo registrar que, para os critérios da base de dados, houve *default* no Brasil no ano seguinte, dado os vultosos empréstimos tomados junto ao FMI em 2001. Contudo, apesar de prever novamente, e desta vez acertadamente, a crise moratória brasileira de 2001, o algoritmo disparou 11 alarmes falsos. Observando-se a árvore gerada, percebe-se que a maior parte dos alarmes falsos foram gerados por conta de uma regra de que variações das reservas menores do que 3,4% gerariam *default*, e neste ano diversos países latinos experimentaram decréscimo em suas reservas, sem, contudo, incorrer em crises moratórias. É possível, assim, que tal variável já não seja mais tão preditiva quanto fora na década de 1980.

Para o ano de 2008, a árvore não previu a única crise moratória no Equador. Vale frisar, contudo, que há relativo consenso no sentido de que este foi um típico *default* que dependeu mais da disposição do que da capacidade do país de pagar, pois a dívida do país não era insustentável (BOHOSLAVSKY, RAFFER, 2017). À época, o então presidente de

esquerda Rafael Correa declarou ilegítima parte da dívida contraída. Isto nos lembra que, para além de qualquer modelo, haverá sempre um componente político que precisará ser analisado.

**Figura 14: Gráfico – Importância das variáveis para as previsões – Modelo América Latina**



**Fonte:** Elaboração do autor

Por fim, o gráfico acima demonstra as variáveis que, para o algoritmo, foram mais relevantes para suas previsões na América Latina. Apesar da alta semelhança com os atributos relevantes para o modelo global, nota-se que o histórico de *default*, a variação das reservas e a variação do PIB foram ainda mais importantes para as previsões feitas para a América Latina. Ademais, aparece com bastante importância a relação M2/Reservas, o que não é visto no modelo global, talvez pelo alto histórico de inflação desta região.

#### 4.3 MODELO ÁFRICA

A tabela abaixo reporta os resultados obtidos pelo algoritmo com a base de dados de países pertencentes ao continente africano:



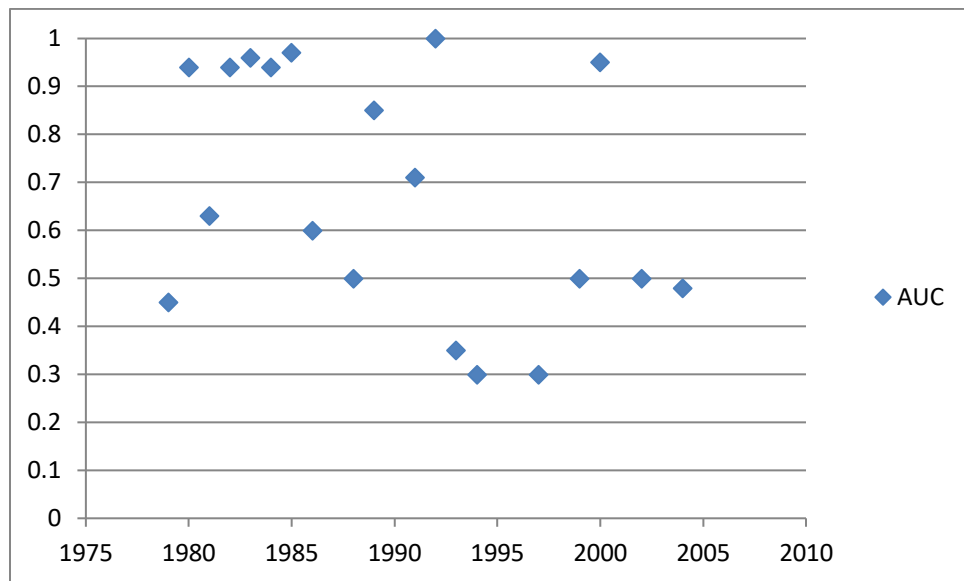
**Tabela 5 – Resultados do Modelo África**

Ano	N	Número de defaults	Defaults corretamente previstos	Defaults incorretamente previstos	Acurácia Geral	AUC*
1979	22	3	0	2	77,27%	0,45
1980	19	2	2	3	84,21%	0,94
1981	21	2	1	5	71,42%	0,63
1982	21	3	3	3	85,71%	0,94
1983	18	6	6	1	94,44%	0,96
1984	17	1	1	5	70,59%	0,94
1985	19	2	2	5	73,69%	0,97
1986	17	4	2	4	64,70%	0,60
1987	13	0	-	0	100%	-
1988	14	2	0	0	85,71%	0,50
1989	14	1	1	2	85,71%	0,85
1990	14	0	-	0	100%	-
1991	15	2	1	1	86,66%	0,71
1992	14	1	1	0	100%	1,0
1993	15	1	0	4	66,66%	0,35
1994	16	1	0	4	68,75%	0,30
1995	16	0	-	3	81,25%	-
1996	17	0	-	2	88,23%	-
1997	19	1	0	3	78,95%	0,30
1998	18	0	-	2	88,88%	-
1999	19	1	0	0	94,74%	0,50
2000	20	1	1	2	90%	0,95
2001	19	0	-	1	94,74%	-
2002	19	1	0	0	94,74%	0,50
2003	23	0	-	1	95,65%	-
2004	23	1	0	1	91,30%	0,48
2005	23	0	-	0	100%	-
2006	23	0	-	1	95,65%	-
2007	23	0	-	1	95,65%	-
2008	23	0	-	0	100%	-
2009	23	0	-	1	95,65%	-
2010	23	0	-	0	100%	-

\* Média ponderada entre as duas classes

**Fonte:** Elaboração do autor

Figura 15 – AUC x Ano – Modelo África

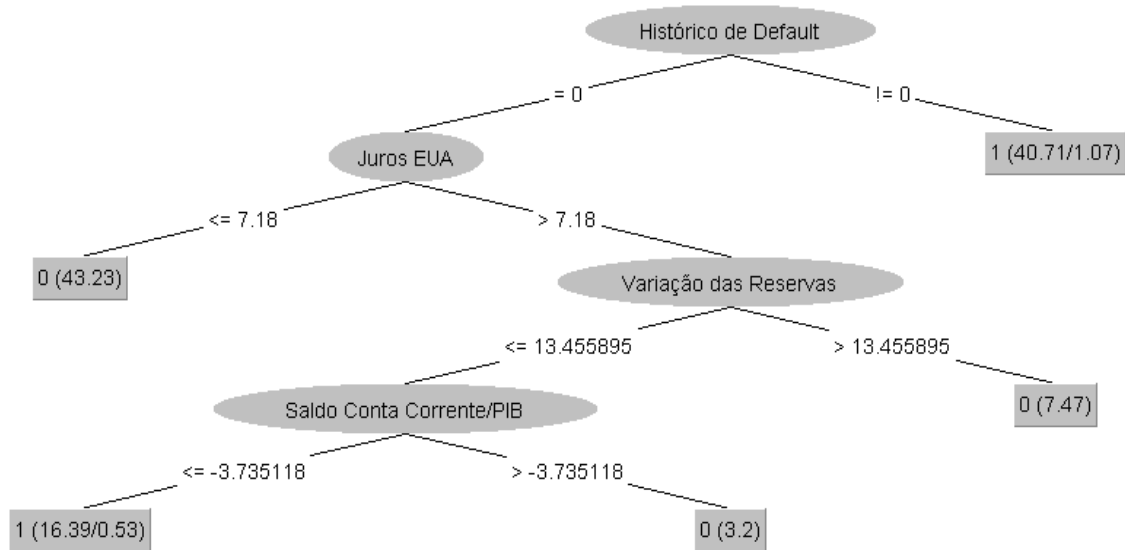


Fonte: Elaboração do autor

Como pode ser visto, mais uma vez, o algoritmo apresenta bom desempenho para crises da década de 1980, mas não consegue manter o aprendizado para as demais crises da década de 1990. Nos anos 2000, há pouquíssimas crises a testar as árvores (2 na África), que também não são previstas; contudo, poucos (6) alarmes falsos foram disparados.

De 23 *defaults* na década de 1980, o algoritmo previu corretamente 18 (78,2%). A árvore abaixo retrata o esquema gerado pelo algoritmo, por exemplo, para prever crises no ano de 1980:

**Figura 16: Árvore gerada para previsão de crises do ano de 1980 – Modelo África**

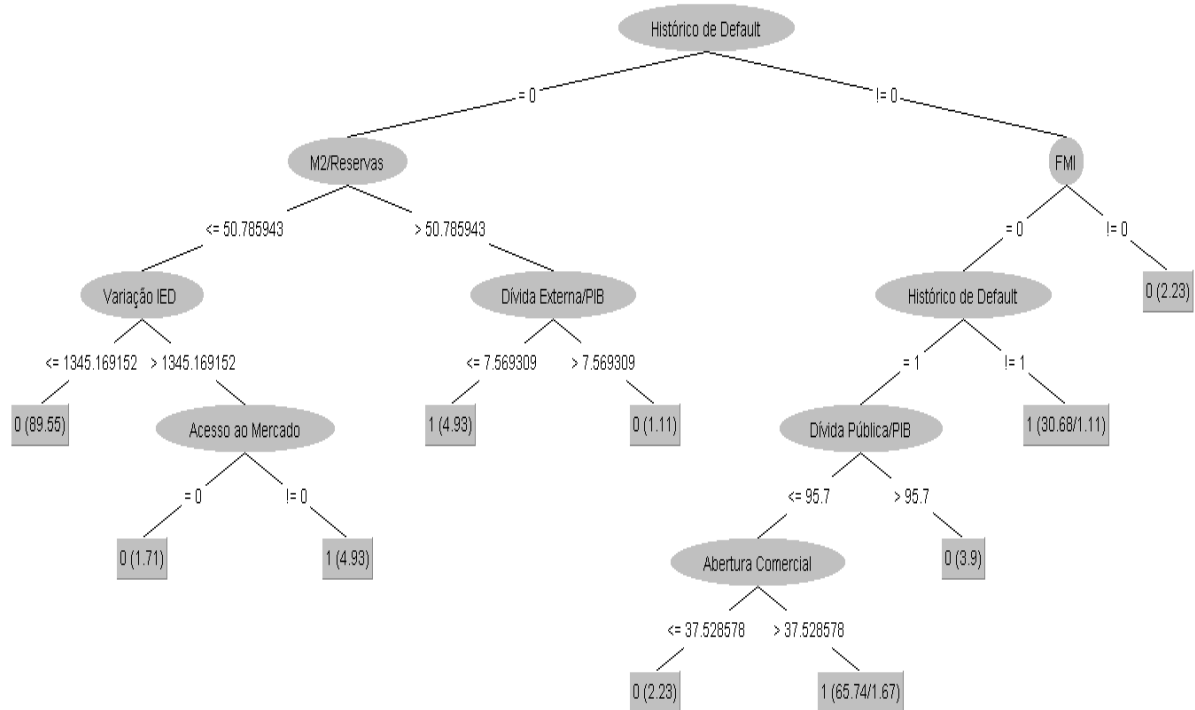


**Fonte:** Elaboração do autor

A árvore gerada elencou regras com sentido do ponto de vista da teoria econômica: juros maiores aumentam a probabilidade de *default*, conforme argumentou, por exemplo, Pasinetti (2005). Ademais, variação das reservas e saldo da conta corrente em relação ao PIB maiores sinalizam maior proteção contra crises cambiais. Contudo, os *defaults* previstos para este ano se basearam apenas no lado direito da árvore, de países com histórico de *default*, de modo que o esquema previu corretamente crises no Marrocos e em Madagascar e incorretamente no Egito e na Nigéria. Ou seja, a falta de sofisticação para países com histórico de *default* acabou levando o algoritmo a disparar alarmes falsos, mas foi o suficiente para prever algumas crises dada a existência de “inadimplentes em série” e a alta incidência de *defaults* na década de 1980.

Com o aumento da base de treinos, o algoritmo foi se tornando mais complexo, e, por vezes, criando regras que contrariam a intuição econômica, como demonstra a árvore feita para prever crises de 1985 (em anexo detalhada para melhor visualização):

**Figura 17: Árvore gerada para previsão de crises do ano de 1985 – Modelo África**

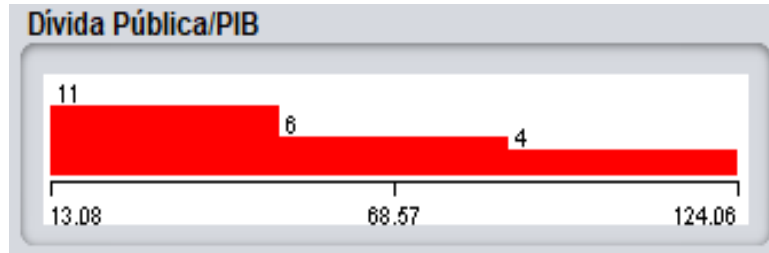


**Fonte:** Elaboração do autor

Neste ano, a árvore previu acertadamente duas crises (Camarões e África do Sul), mas emitiu cinco alarmes falsos (Egito, Quênia, Malawi, Ilhas Maurício e Zimbábwe). Nota-se, por exemplo, no lado direito da árvore, condicionado à existência de histórico de *default*, que os países são submetidos a um nó decisório envolvendo a relação dívida pública/PIB e outro, sucessivo, envolvendo a abertura comercial. Em confronto com a teoria econômica, uma dívida pública/PIB acima de 95,7% levou o algoritmo a prever a não ocorrência de *default*. Este nó, aparentemente sem sentido, sugere que tal variável tem pouca capacidade de sinalizar *defaults*, já que o algoritmo não conseguiu aprender dela.

Analisando-se a base de dados, nota-se que, de fato, nas 21 entradas de *default* ocorridas de 1975 até 1984, com que o algoritmo aprendeu para prever *defaults* no ano de 1985, não há relação aparente entre maiores taxas de dívida pública/PIB e a incidência de *defaults*, conforme demonstra a distribuição abaixo:

**Figura 18:** Distribuição dos registros de *default* na África de 1975 a 1984 e a relação dívida pública/PIB



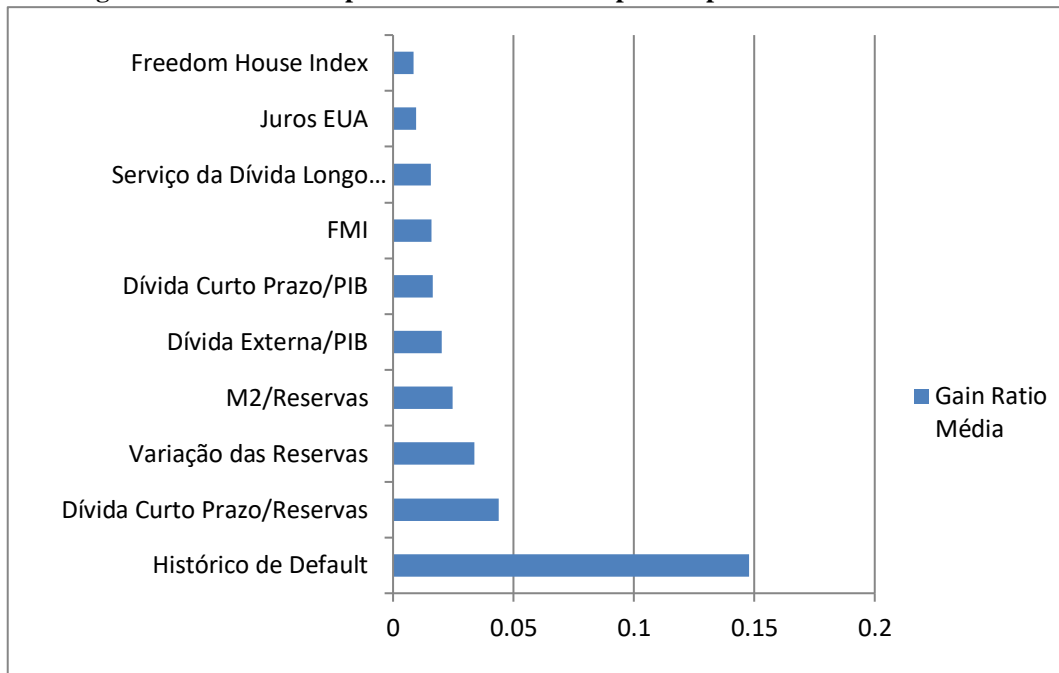
**Fonte:** Elaboração do autor

Assim, a regra criada pelo algoritmo neste nó, apesar de aparentemente desprovida de lógica, é compreensível a partir da análise da base de dados. Note-se que, considerando que o foco do modelo é preditivo, isto não quer dizer que aumentos do endividamento público não estejam na causa das crises moratórias e devam ser cuidados, mas sim que tal variável é importante à medida de sua interação com outras variáveis, as quais, por sua vez, podem servir como melhores sinalizadores de *default*.

Abaixo do referido nó, o algoritmo cria outra regra que contraria a teoria econômica. Como dito alhures, Eaton e Gersovitz (1981) sugeriram e encontraram evidências de que países mais dependentes de comércio exterior sofreriam mais com possíveis retaliações por um *default*, sendo assim menos dispostos a enfrentá-los. O algoritmo, todavia, associou maior abertura comercial (importações + exportações/PIB > 37,5) à incidência de *default* e assim conseguiu prever as crises na África do Sul e em Camarões. Uma possível racionalidade para isto é que países mais abertos sofram mais com variações cambiais e, conforme se nota, crises cambiais têm bastante relação com crises moratórias, ao menos na década de 1980.

A partir da previsão do ano de 1987, as árvores foram ficando cada vez mais complexas. Por exemplo, enquanto a árvore exibida acima, para a previsão de 1985, contém 10 “folhas”, a árvore para prever o ano de 1990 possui 19, quase o dobro.

Outra dificuldade que se nota, a partir dos anos 1990, é a dificuldade de aprendizado do algoritmo, pois ocorrem pouquíssimas crises moratórias. Igualmente, é difícil testar o aprendizado do modelo – por exemplo, a partir dos anos 2000, a acurácia geral é sempre superior a 90%, mas não se pode avaliar devidamente a capacidade de discriminação do algoritmo, pois a maioria dos anos não tem qualquer registro de crise moratória.

**Figura 19: Gráfico – Importância das variáveis para as previsões – Modelo África**

**Fonte:** Elaboração do autor

Por fim, a figura acima demonstra as variáveis consideradas mais relevantes, pelo algoritmo, para previsões dos eventos em continente africano. Como se nota, há preponderância das mesmas variáveis utilizadas na América Latina, mas chama atenção que os juros norte-americanos tiveram menos importância para as previsões. A variação do PIB e a relação dívida pública/PIB, consideradas por si só, sem condicionantes, não tiveram relevância alguma no aprendizado do algoritmo.

#### 4.4 MODELO ÁSIA

A tabela abaixo reporta os resultados obtidos pelo algoritmo com a base de dados de países pertencentes ao continente asiático:

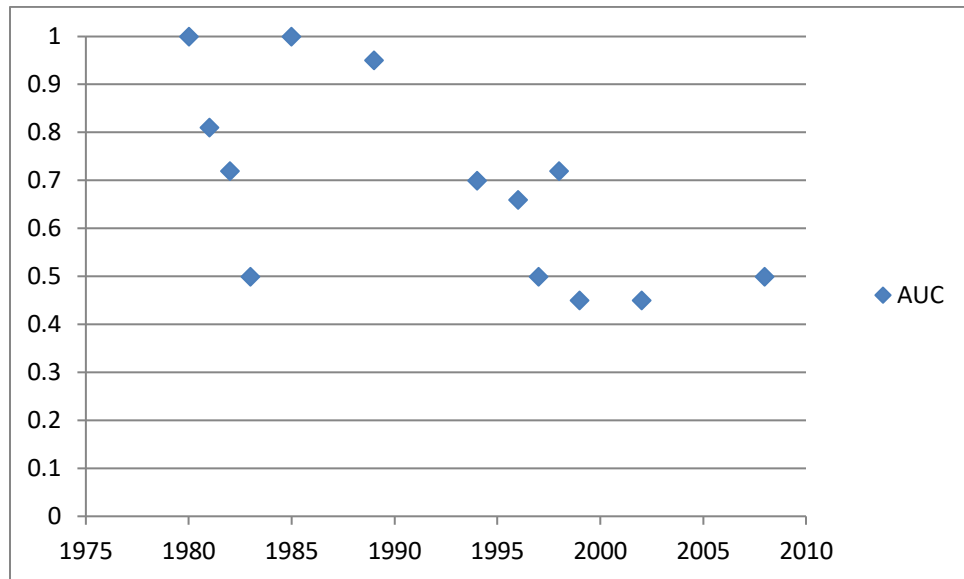
**Tabela 6 – Resultados do Modelo Ásia**

Ano	N	Número de defaults	Defaults corretamente previstos	Defaults incorretamente previstos	Acurácia Geral	AUC*
1979	11	0	-	0	100%	-
1980	11	4	4	0	100%	1,0
1981	10	2	2	3	70%	0,81
1982	12	1	1	6	50%	0,72
1983	12	2	0	0	83,33%	0,50
1984	11	0	-	5	54,54%	-
1985	11	1	1	2	81,82%	1,0
1986	11	0	-	2	81,82%	-
1987	11	0	-	1	90,91%	-
1988	11	0	-	0	100%	-
1989	11	1	1	1	90,91%	0,95
1990	10	0	-	3	70%	-
1991	11	0	-	0	100%	-
1992	11	0	-	0	100%	-
1993	12	0	-	0	100%	-
1994	10	2	1	1	83,33%	0,70
1995	13	0	-	2	84,61%	-
1996	13	2	1	2	76,92%	0,66
1997	13	4	0	0	69,23%	0,50
1998	12	2	1	0	91,66%	0,72
1999	11	1	0	0	90,90%	0,45
2000	11	0	-	1	91,66%	-
2001	13	0	-	2	84,61%	-
2002	13	2	0	1	76,92%	0,45
2003	13	0	-	1	92,30%	-
2004	13	0	-	0	100%	-
2005	13	0	-	0	100%	-
2006	13	0	-	1	92,30%	-
2007	13	0	-	3	76,92%	-
2008	13	1	0	0	92,30%	0,50
2009	12	0	-	2	83,33%	-
2010	13	0	-	0	100%	-

\* Média ponderada entre as duas classes

**Fonte:** Elaboração do autor

**Figura 20 – AUC x Ano – Modelo Ásia**

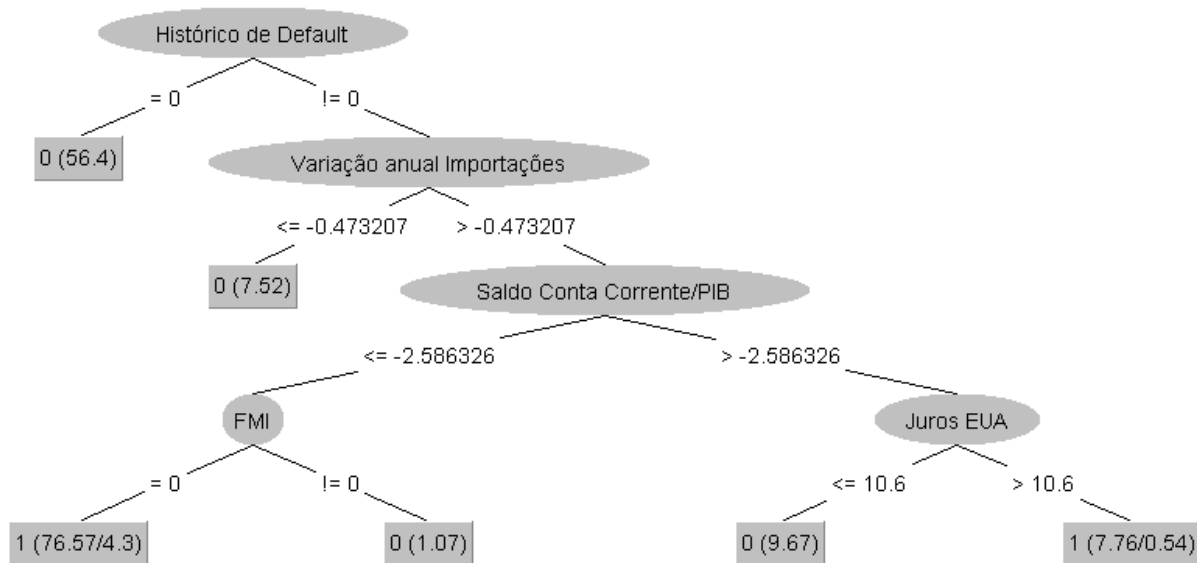


**Fonte:** Elaboração do autor

A baixa incidência de crises moratórias constantes na base de dados relativa ao continente asiático dificultou o treino, o teste e a devida avaliação do algoritmo. Nas primeiras quatro tentativas de previsão, o algoritmo simplesmente previu *default* para todos os países com algum histórico de *default*. A partir da previsão de 1983, o algoritmo criou árvores mais complexas, mas com baixa capacidade de predição, deixando de prever os dois *defaults* ocorridos naquele ano (Coreia do Sul e Filipinas). Para 1985, com uma base de treinos maior, o algoritmo conseguiu sofisticar mais seu esquema e obter um bom resultado, prevendo a única crise que houve naquele ano (Tailândia), ao custo de dois alarmes falsos (Índia e Turquia). Não há registro, na base de dados, de crises nos anos subsequentes, o que ocorre apenas em 1989. Deste ano, algoritmo conseguiu prever a crise na Jordânia, disparando apenas um alarme falso em relação à Tailândia. A árvore gerada para a previsão dos eventos de 1989 pode ser vista abaixo:



**Figura 21: Árvore gerada para previsão de crises do ano de 1989 – Modelo Ásia**



**Fonte:** Elaboração do autor

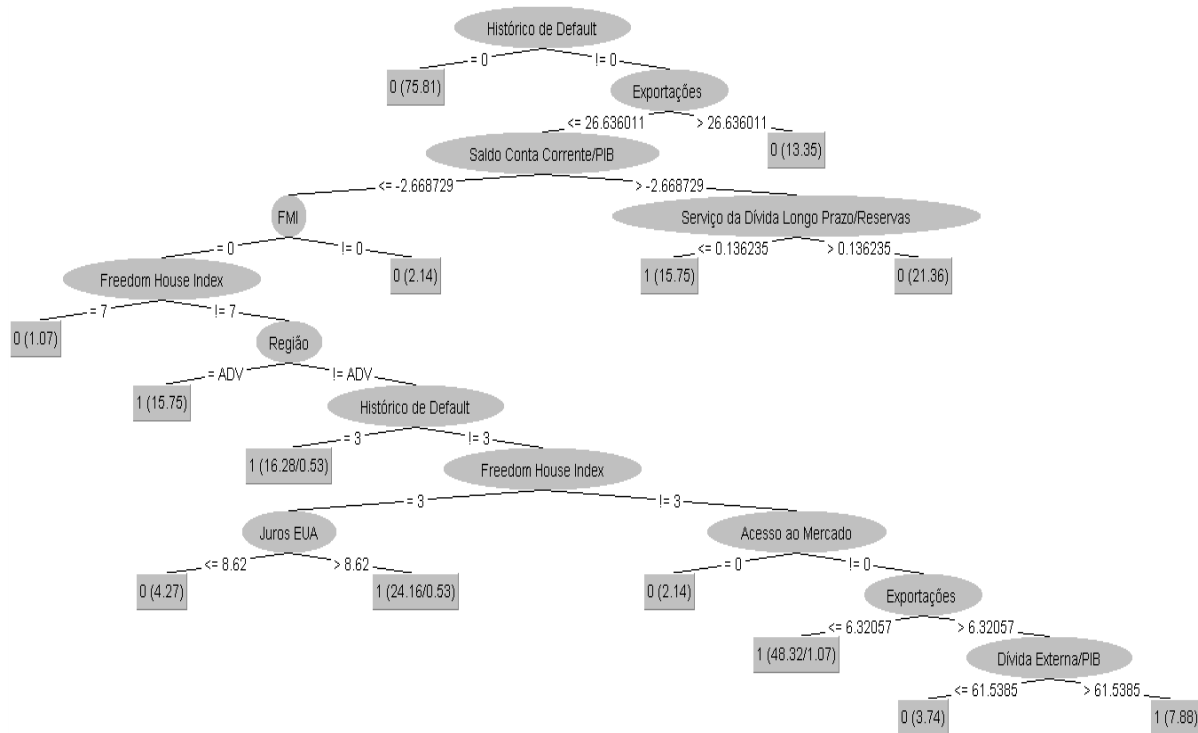
Nota-se que o produto do algoritmo enfatiza, além do histórico de *default*, a variação anual das importações, associando taxas de diminuição abaixo de -0,4% à ausência de crises. Considerando a natureza das crises moratórias da década de 1980, trata-se de um nó com amparo na teoria econômica, vez que aumentos nas importações podem estar associados a déficits na balança comercial e consequentes depreciações da moeda. No mesmo sentido, o nó seguinte da árvore usa a taxa saldo da conta corrente/PIB para separar as observações. Por fim, a exemplo de outras árvores, os juros globais são usados como divisor, bem como a contração de empréstimos junto ao FMI.

De 1989 em diante, poucas crises ocorreram no continente asiático. Para 1994, o algoritmo não previu a crise ocorrida na Turquia, mas conseguiu acusar a crise das Filipinas, prevendo também, incorretamente, crise no Paquistão (que, de fato, entrou em *default* quatro anos mais tarde). Em relação ao ano de 1996, o algoritmo prevê com precisão a crise na Jordânia, mas deixa de acusar a crise no Cazaquistão. Ainda para 1996, o algoritmo dispara dois falsos alarmes, para Paquistão e Filipinas – ambos os países entrariam em *default* dois anos mais tarde.

Contudo, os quatro *defaults* registrados na crise do Sudeste Asiático de 1997 (Indonésia, Sri Lanka, Turquia e Coreia do Sul), principal crise do continente da década de 90, não foram detectados pelo algoritmo, que não acusou crise alguma para aquele ano. Para 1998, o algoritmo previu novamente crise no Paquistão, mas desta vez corretamente; todavia,

a árvore não detectou a crise ocorrida em Filipinas. A árvore gerada para o ano de 1997 pode ser visualizada através do seguinte esquema:

**Figura 22: Árvore gerada para previsão de crises do ano de 1997 – Modelo Ásia**



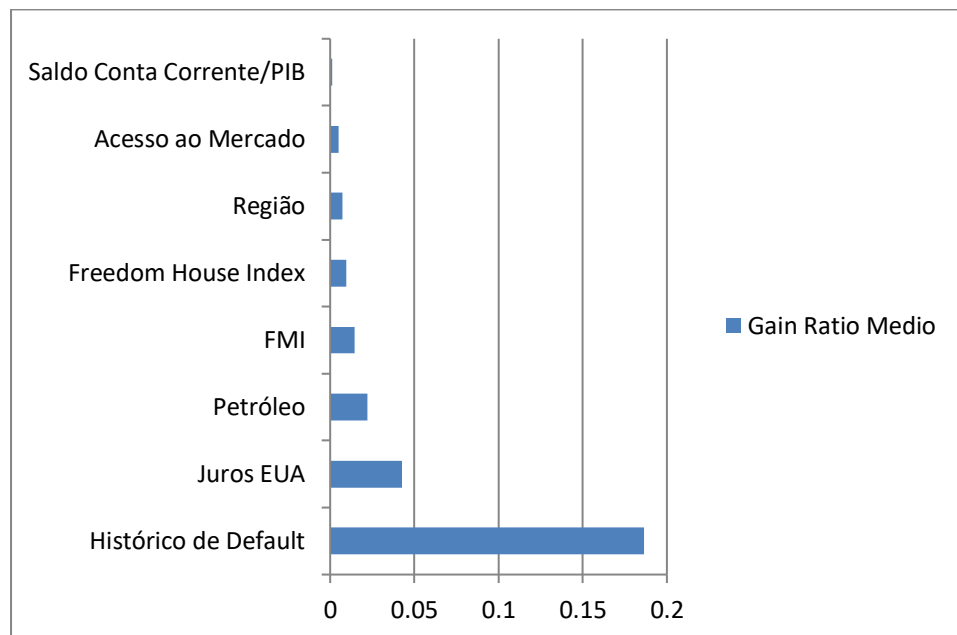
**Fonte:** Elaboração do autor

Nota-se que as crises na Indonésia, Coreia e Tailândia não foram previstas por conta do nó que estabelecia um limiar de 26 bilhões de dólares nas exportações – o algoritmo previa que, as exportações ultrapassando este valor, não ocorreriam crises, contudo não foi o que se observou nos referidos países. A crise no Sri Lanka, por sua vez, não foi prevista por conta do nó gerado para os juros americanos, que previa que sob as condições dos nós superiores, juros abaixo de 8,62% não gerariam *defaults*. Assim, aparentemente, o algoritmo se pautou por valores da década de 1980, o que não foi suficiente para o aprendizado das crises do final do milênio, ainda que estas tenham características em comum, como desbalanços cambiais.

Quanto aos anos do novo milênio, houve pouquíssimas crises moratórias (duas em 2002 e uma em 2008); neste contexto, embora o algoritmo tenha disparado poucos alarmes falsos, ele foi incapaz de detectar qualquer uma das três crises. Registra-se, porém, que o algoritmo previu crise em 2001 na Turquia, que ocorreu efetivamente em 2002, e em 2007 no Paquistão, que ocorreu de fato em 2008.

Por fim, a figura abaixo demonstra as variáveis consideradas mais relevantes, pelo algoritmo, para previsões dos eventos em continente asiático. Como se nota, há muito menos variáveis relevantes do que outros continentes, emergindo como principais indicadores apenas o histórico de *default* e a taxa de juros praticada nos Estados Unidos. Uma explicação para isto é que, como referido, a baixa incidência de crises moratórias no continente asiático constante na base de dados dificultou o aprendizado do algoritmo.

**Figura 23: Gráfico - Importância das variáveis para as previsões – Modelo Ásia**



**Fonte:** Elaboração do autor

Conforme se pôde ver, os modelos regionais não obtiveram resultados consistentemente melhores que o modelo global. Estes resultados oferecem indicativos a favor dos achados de Fuertes e Kalotychou (2006), para quem “restringir a amostra de países a regiões não oferece ganhos significantes de previsão sobre uma abordagem mais ampla global”. Por sua vez, também oferecem indicativos em contraste ao obtido por Darwood, Horsewood e Strobel (2017), que concluem que os modelos regionais apresentam ganhos de predição.

## 5 CONCLUSÃO

Este trabalho utilizou algoritmos de *machine learning* para auxiliar a previsão de *defaults* da dívida pública, detectando sinais prévios (“*early warning signals*”, conforme usualmente referido na literatura). Para isto, testou-se o algoritmo de árvores de classificação C4.5, que, com base em uma medida baseada na teoria da informação (*gain ratio*), cria uma série de regras que são interligadas de modo a criar árvores decisórias. Além dos resultados obtidos, analisou-se neste trabalho, a partir de um método exploratório, a própria lógica algorítmica, tentando traçar sua relação com a teoria econômica atinente, que também foi exposta neste trabalho, e com fatos da história econômica mundial.

Tal algoritmo obteve êxito na previsão das crises da década de 1980 e do início da década de 1990, registrando, a partir daí, perda na capacidade de aprendizado. Até o ano de 1992, o método foi capaz de detectar 63 de 68 (92,6%) crises moratórias contidas na base de dados. Isto foi feito a uma taxa de emissão de 18,3% de alarmes falsos, a qual é semelhante ou superior aos resultados obtidos com modelos de limites e logit, mas inferior a alguns resultados reportados com métodos de *machine learning*. Não obstante, cumpre ressaltar que os resultados não são diretamente comparáveis, por se referirem a períodos diferentes e, eventualmente, bases de dados diferentes. Ademais, alguns falsos positivos não podem ser considerados puramente incorretos, vez que renunciaram alguma crise que foi evitada por medidas cabíveis ou que ocorreu em um horizonte temporal maior que o de um ano, adotado neste trabalho.

Em termos de variáveis preditivas, o algoritmo corroborou a existência de “*serial defaulters*”, posto que o histórico de *default* dos países quase sempre figurou como importante sugestivo de crises futuras. Isto indica que se deve ter mais cautela com países com maiores históricos de *default* do que com outros - especialmente em relação a dívidas externas, conforme demonstram outras variáveis relevantes. Contudo, tal variável pode ter tido mais importância na década de 1980 do que nos anos seguintes, considerando que o algoritmo detectou poucas crises nas demais décadas.

Outras variáveis que se mostraram importantes para prever crises foram aquelas relacionadas com as reservas – principalmente a variação das reservas e a taxa da dívida a curto prazo sobre as reservas. Isto demonstra o forte nexos existente entre crises moratórias e cambiais, sobretudo nos *defaults* da década de 1980. Porém, o Plano Brady, iniciado em 1989, possibilitou a alguns países emergentes maior flexibilidade na negociação de seus títulos e assim na rolagem de seus compromissos, de modo que algumas nações, como o Brasil,

conseguiram reestruturar suas dívidas a depender menos das variações externas. Para o ano de 2001, por exemplo, o algoritmo previu diversos falsos alarmes baseado em um aprendizado que superestimava a importância da diminuição das reservas.

A variação do PIB também se mostrou uma variável relevante, para o algoritmo, tanto no modelo global quanto no dedicado à América Latina. A taxa de juros dos Estados Unidos, que sinalizam os juros globais, também se mostrou uma variável importante de discriminação, sobretudo sob certas condições prévias, o que sugere o perigo de momentos de iliquidez global para certas nações.

A relação da dívida pública com o PIB, frequentemente usada para sinalizar perigos de inadimplência, não foi, para o algoritmo, uma variável com alto poder de classificação. Aqui é preciso enfatizar uma das limitações do modelo, qual seja, sua incapacidade de estabelecer causalidades. Assim, o desuso de tal variável não significa ausência de causalidade entre maiores endividamentos e *defaults*, mas sim que, para este método, tal variável foi importante apenas à medida de sua interação com outras variáveis, as quais, por sua vez, podem servir como melhores sinalizadores de *default*. Conquanto seja possível que, com outros métodos preditivos, a referida variável tenha, por si só, maior capacidade preditiva, este trabalho sugere que não se deve exagerar na capacidade incondicional desta variável de prever crises moratórias.

Em relação ao decréscimo de desempenho observado a partir das previsões para o início da década de 1990, várias razões podem ter contribuído para tal resultado. A literatura de *machine learning* registra que, por vezes, o algoritmo C4.5 não tem um método de poda das árvores tão eficaz, de modo que, quanto mais instâncias são adicionadas à base de treino, mais as árvores crescem, mesmo quando isto não aumenta a performance da árvore em amostras independentes (Witten et al, 2016). Isto foi precisamente o que ocorreu na presente pesquisa: à medida que se aumentou a base de treinamento, as árvores ficaram por demais complexas, perdendo parte da principal vantagem do método, que reside na facilidade de interpretação. Além disso, no presente caso, o aumento de complexidade, em geral, não serviu para prever devidamente as crises posteriores à década de 1980. Neste sentido, conclui-se que o algoritmo CART, que usa o índice de Gini para gerar árvores de classificação e possui um mecanismo de poda que analisa os custos da complexidade, pode gerar resultados com melhor interpretação, conforme Manasse, Roubini e Schimmelpfennig (2003). Por sua vez, algoritmos *ensemble*, que agregam dezenas de esquemas decisórios, ainda que não apresentem um produto tão interpretável, podem ter resultados preditivos melhores, conforme demonstram Savona e Vezzoli (2015).

Este trabalho testou também a segmentação da base de dados em continentes para comparação de resultados. Em geral, êxitos e limitações parecidos com o modelo global foram observados, possivelmente porque tal divisão reduziu ainda mais a base de treino, diminuindo também o número de *defaults* para aprendizado (e, também, para avaliação) do modelo. Assim, apesar de os modelos regionais não apresentarem ganhos preditivos em relação ao modelo global, este trabalho não é conclusivo a respeito da inexistência de heterogeneidade regional em matéria de EWS, posto que o tamanho das bases de dados é sabidamente um fator que afeta o aprendizado de máquina.

Em relação a tal aprendizado, nota-se que a criação de regras pelo algoritmo, se por um lado ilumina relações e dependências já esperadas, também, por vezes, cria associações inéditas, que podem tanto significar aprendizado deficiente, base de dados de baixa qualidade ou sobreajuste à base de treinos, quanto relações efetivamente existentes e negligenciadas pelo pensamento humano. Em se tratando de métodos algorítmicos, nota-se que transparência nem sempre gera compreensão. Contudo, o fato de se poder visualizar e questionar a lógica do algoritmo ou mesmo aprender com ele já é uma das vantagens que algoritmos de árvores apresentam em relação a outros algoritmos do gênero “caixa preta”, como redes neurais.

Outra vantagem do algoritmo é a criação de regras não lineares e condicionais, que também, muitas vezes, passam ao largo da compreensão humana e de outros modelos econométricos. Contudo, a própria natureza do algoritmo, que recursivamente divide as amostras, faz com que cada vez mais as relações estabelecidas sejam mais “locais” – assim, o esquema estabelece ótimos locais e condicionais, mas não relações que se apliquem a toda a base de dados.

O fato de o algoritmo aceitar variáveis numéricas e categóricas, sem quaisquer assunções sobre distribuições, também é uma vantagem em relação a métodos econométricos. O algoritmo C4.5 também pode ser usado com variáveis colineares, apesar de ser necessário atentar, nestes casos, para não se confundir omissão de variáveis nas árvores com irrelevância.

Todavia, a referida ausência de assunções também limita a possibilidade de se conduzir testes estatísticos nos resultados ou de se calcular os efeitos marginais de cada variável na probabilidade de pertencimento a uma classe, como fazem, por exemplo, modelos logit/probit.

Conclui-se, por fim, que apesar de suas limitações, algoritmos de aprendizado de máquina podem ser ferramentas preciosas, se não para a previsão de crises, ao menos para a detecção de sinais de crise. Por conta da própria natureza opaca destes algoritmos, não se deve delegar a eles decisões estratégicas; contudo, como ferramentas acessórias à tomada de

decisão, tais métodos podem ajudar a estabelecer focos de atenção. Neste sentido, as árvores de decisão são ferramentas que podem ajudar por conta de sua facilidade de interpretação, sendo esquemas comunicáveis até a audiências não técnicas.

## REFERÊNCIAS

- ADHATRAO, Kalpesh et al. Predicting students' performance using ID3 and C4. 5 classification algorithms. **International Journal of Data Mining & Knowledge Management Process (IJDKP)**, [S.l.], v.3, n.5, 2013. Disponível em: <https://bit.ly/3jkmad0>. Acesso em: 20 out. 2020.
- ATHEY, Susan. The Impact of Machine Learning on Economics. **The economics of artificial intelligence**. University of Chicago Press, p. 507-552, 2019. Disponível em: <https://www.degruyter.com/document/doi/10.7208/9780226613475-023/html> . Acesso em 14 set. 2020.
- BAJARI, Patrick; NEKIPELOV, Denis STEPHEN P., Ryan; MIAOYU, Yang. Machine learning methods for demand estimation. **American Economic Review**, v.105, n.5, p.481–85, 2015. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/aer.p20151021>. Acesso em: 02 fev. 2021.
- BI, Huixin. Sovereign default risk premia, fiscal limits, and fiscal policy. **European Economic Review**, v. 56, n. 3, p. 389–410, 2012. Disponível em: <https://bit.ly/3sD2Pdi>. Acesso em: 20 out. 2020.
- BOHOSLAVSKY, Juan Pablo; RAFFER, Kunibert (ed.). **Sovereign debt crises: what have we learned?**. London: Cambridge University Press, 2017.
- BOONMAN, Tjeerd Menno. **Sovereign debt defaults and currency crises in Latin America**. Tese (Doutorado). University of Groningen. Groningen, 2015. Disponível em: <https://bit.ly/3dKhFay>. Acesso em: 20 out. 2020.
- BRONER, Fernando A.; LORENZONI, Guido; SCHMUKLER, Sergio L. Why do emerging economies borrow short term?. **Journal of the European Economic Association**, v. 11, n.1, p. 67-100, 2013. Disponível em: [https://academic.oup.com/jeea/article-abstract/11/suppl\\_1/67/2316398](https://academic.oup.com/jeea/article-abstract/11/suppl_1/67/2316398). Acesso em: 28 nov. 2020.
- CALVO, Guillermo A. Capital flows and capital-market crises: the simple economics of sudden stops. **Journal of applied Economics**, v. 1, n. 1, p. 35-54, 1998. Disponível em: <https://bit.ly/2ThGrVU>. Acesso em: 20 out. 2020.
- CAMILO, Cássio Oliveira; SILVA, João Carlos da. **Mineração de dados: Conceitos, tarefas, métodos e ferramentas**. Goiânia: Universidade Federal de Goiás, 2009. Disponível em: <https://bit.ly/3dNTND1>. Acesso em: 20. out. 2020.
- CATÃO, Luis AV. Sudden stops and currency drops: a historical look. *In*: EDWARDS, Sebastian; ESQUIVEL, GERARDO; MARQUES, Graciela. **The Decline of Latin American Economies: growth, institutions, and crises**. Chicago: University of Chicago Press. p. 243-290, 2007. Disponível em: <https://www.nber.org/system/files/chapters/c10658/c10658.pdf>. Acesso em 20 abr. 2021.
- CATÃO, Luis AV; MILESI-FERRETTI, Gian Maria. External liabilities and crises. **Journal of International Economics**, v. 94, n. 1, p. 18-32, 2014.



CAVALLO, Eduardo; EICHENGREEN, Barry; PANIZZA, Ugo. Can countries rely on foreign saving for investment and economic development?. **Review of World Economics**, v. 154, n. 2, p. 277-306, 2018.

CHAMON, Marcos; MANASSE, Paolo; PRATI, Alessandro. Can we predict the next capital account crisis?. **IMF Staff Papers**, v. 54, n. 2, p. 270-305, 2007. Disponível em: <https://link.springer.com/article/10.1057/palgrave.imfsp.9450012> . Acesso em: 17 mar. 2021.

CIARLONE, Alessio; TREBESCHI, Giorgio. Designing an early warning system for debt crises. **Emerging Markets Review**, v. 6, n. 4, p.376-395, 2005. Disponível em: <https://bit.ly/3gsWEno>. Acesso em: 09 dez. 2020.

CLARK, John. Debt reduction and market reentry under the Brady plan. **Federal Reserve Bank of New York Quarterly Review**, v. 18, n. 4, p.38-62, 1993. Disponível em: <https://nyfed.org/3B8nHfk> . Acesso em: 11 mar. 2021.

COLE, Harold; KEHOE, Timothy. "Self-Fulfilling Debt Crises". **The Review of Economic Studies**, v. 67, n. 1, p.91-116, 2000. Disponível em: <https://academic.oup.com/restud/article-abstract/67/1/91/1576276> . Acesso em 23 jul. 2021.

DAWOOD, Mary; HORSEWOOD, Nicholas; STROBEL, Frank. Predicting sovereign debt crises: an early warning system approach. **Journal of Financial Stability**, v. 28, p. 16-28, 2017. Disponível em: <https://ideas.repec.org/a/eee/finsta/v28y2017icp16-28.html>. Acesso em: 20 out. 2020.

DETRAGIACHE, Enrica; SPILIMBERGO, Antonio. **Crises and Liquidity: Evidence and Interpretation**. Documento de trabalho do FMI, n.01/2, 2001. Disponível em: <https://www.imf.org/external/pubs/ft/wp/2001/wp0102.pdf>. Acesso em: 15 out. 2020.

DUTTAGUPTA, Rupa; CASHIN, Paul. Anatomy of banking crises in developing and emerging market countries. **Journal of International Money and Finance**, v. 30, n. 2, p. 354-376, 2011. Disponível em: <https://bit.ly/3jcO51P>. Acesso em: 21 dez. 2020.

DOMAR, Evsey D. The "burden of the debt" and the national income. **The American Economic Review**, p. 798-827, 1944. Disponível em: <https://www.jstor.org/stable/1807397>. Acesso em: 20 nov. 2020.

DORNBUSCH, Rudiger; GOLDJAHN, Ilan; VALDÉS, Rodrigo. Currency crises and collapses. **Brookings papers on economic activity**, p. 219-293, 1995. Disponível em: <https://brook.gs/2Ht3EBL>. Acesso em: 20 out. 2020.

EATON, Jonathan; GERSOVITZ, Mark. Debt with Potential Repudiation: Theoretical and Empirical Analysis Source. **The Review of Economic Studies**, v. 48, n. 2, p. 289-309, 1981. Disponível em: <https://www.jstor.org/stable/2296886>. Acesso em: 20 out. 2020.

EICHENGREEN, Barry; HAUSMANN, Ricardo. **Exchange rates and financial fragility**. 1999. Disponível em: <https://www.nber.org/papers/w7418>. Acesso em: 20 nov. 2020.

ELREEDY, Dina; ATIYA, Amir F. A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance. **Information Sciences**, p. 32-64, 2019. Disponível em: <https://bit.ly/34jk8pdf> Acesso em: 20 out. 2020.

FERRI, Piero; MINSKY, Hyman P. Market processes and thwarting systems. **Structural Change and Economic Dynamics**, v. 3, n. 1, p. 79-91, 1992. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0954349X92900274>. Acesso em: 20 jan. 2021.

FIORAMANTI, Marco. Predicting sovereign debt crises using artificial neural networks: A comparative approach. **Journal of Financial Stability**, v. 4, n. 2, p. 149-164, 2008.

FUERTES, Ana-Maria; KALOTYCHOU, Elena. Optimal design of early warning systems for sovereign debt crises. **International Journal of Forecasting**, v. 23, n. 1, p. 85-100, 2007. Disponível em: <https://bit.ly/2XImehv>. Acesso em: 09 fev. 2021.

FUERTES, Ana-Maria; KALOTYCHOU, Elena. Early warning systems for sovereign debt crises: The role of heterogeneity. **Computational statistics & data analysis**, v. 51, n. 2, p. 1420-1441, 2006. Disponível em: <https://bit.ly/3yfl15X> . Acesso em: 08 fev. 2021.

FUNDO MONETÁRIO INTERNACIONAL - FMI. **The liberalization and management of capital flows: an institutional view**. Washington: IMF, 2012. Disponível em: <https://www.imf.org/external/np/pp/eng/2012/111412.pdf> . Acesso em: 01 nov. 2021.

FUNDO MONETÁRIO INTERNACIONAL - FMI. **Debt Maturity and the Use of Short-Term Debt: Evidence from Sovereigns and Firms** IMF Departmental Papers. 2019. Policy Papers 2019/001. Disponível em: <https://ideas.repec.org/p/imf/imfdps/2019-001.html> Acesso em: 09 jun. 2021.

GHOSH, Atish R., et al. Surges. **Journal of International Economics** v. 92, n. 2, p. 266-285, 2014. Disponível em: <https://bit.ly/3DaitSj> . Acesso em 13 abr. 2021.

HAUSMANN, Ricardo; EICHENGREEN, Barry; PANIZZA, Ugo. Currency mismatches, debt intolerance and original sin: Why they are not the same and why it matters. **NBER Working Paper**, n. 10036, 2003. Disponível em: <https://bit.ly/3kgITKK> . Acesso em: 07 mar. 2021.

KAMINSKY, Graciela Laura. **Currency and banking crises: the early warnings of distress**. International Monetary Fund, 1999.

KUHN, Max; JOHNSON, Kjell. **Applied predictive modeling**. New York: Springer, 2013.

KINDLEBERGER, Charles P.; ALIBER, Robert Z. **Manias, panics and crashes: a history of financial crises**. Palgrave Macmillan, 2011.

KOTSIANTIS, Sotiris B. et al. Supervised machine learning: A review of classification techniques. **Emerging artificial intelligence applications in computer engineering**, v. 160, n. 1, p. 3-24, 2007.

LEMOS, Eliane Prezepiorski; STEINER, Maria Teresinha Arns; NIEVOLA, Julio César. Análise de crédito bancário por meio de redes neurais e árvores de decisão: uma aplicação simples de data mining. **Revista de Administração - RAUSP**, São Paulo, v.40, n.3, p. 225-234, 2005. Disponível em: <https://bit.ly/3m9D67N>. Acesso em: 20 out. 2020.

MANASSE, Paolo; ROUBINI, Nouriel; SCHIMMELPFENNIG, Axel. **Predicting Sovereign Debt Crises**. 2003. Documento de trabalho do FMI n.03/221. Disponível em: <https://bit.ly/3jk7t9R>. Acesso em: 21 out. 2020.

MANASSE, Paolo; ROUBINI, Nouriel. “Rules of thumb” for sovereign debt crises. **Journal of International Economics**, v. 78, n. 2, p. 192-205, 2009. Disponível em: <https://bit.ly/3mscdzw>. Acesso em: 27 set. 2020.

MANASSE, Paolo; SAVONA, Roberto; VEZZOLI, Marika. **Rules of thumb for banking crises in emerging markets**. , Bologna: Universidade de Bologna, 2013. Disponível em: <https://bit.ly/3jdK0qQ>. Acesso em: 20 out. 2020.

MENDOZA, Enrique G. Credit, prices, and crashes: Business cycles with a sudden stop. *In*: EDWARDS, Sebastian; FRANKEL, Jeffrey A. **Preventing currency crises in emerging markets**. University of Chicago Press, p. 335-392, 2002. Disponível em: <https://www.nber.org/papers/w8338>. Acesso em: 20 out. 2020.

MCFADDEN, Daniel *et al.* Is there life after debt? An econometric analysis of the creditworthiness of developing countries. **International debt and the developing countries**, p. 179-209, 1985. Disponível em: <https://bit.ly/34jSN6d> . Acesso em: 20 out. 2020.

NYMAN, Rickard; ORMEROD, Paul. **Predicting Economic Recessions Using Machine Learnings Algorithms**. New York: Cornell University, 2017. Disponível em: <https://bit.ly/388eSpG>. Acesso em: 15 mar. 2021.

PAINCEIRA, Juan Pablo; KALTENBRUNNER, Annina. “Developing countries’ changing nature of financial integration and new forms of external vulnerability: the Brazilian experience”. **Cambridge Journal of Economics**, v. 39, p. 1281–1306, 2015. Disponível em: <https://academic.oup.com/cje/article-abstract/39/5/1281/1686544>. Acesso em: 28 nov. 2020.

PASINETTI, Luigi L. The social ‘burden’ of high interest rates. *In*: SAWYER, Malcom. **Capital Controversy, Post Keynesian Economics and the History of Economic Thought**. London: Routledge, 2005. p. 173-180. Disponível em: <https://bit.ly/3y5DyTv> Acesso em: 25 nov. 2020.

PETER, Marcel. **Estimating default probabilities of emerging market sovereigns: A new look at a not-so-new literature**. Genebra: Instituto Universitário de Altos Estudos Internacionais, 2002. Disponível em: <https://bit.ly/3jfEO5D>. Acesso em: 20 out. 2020.

QUINLAN, J. Ross. **C4. 5: Programming for machine learning**. Burlington: Morgan Kauffmann, 1993.

REINHART, Carmen M.; ROGOFF, Kenneth S. Financial and sovereign debt crises: Some lessons learned and those forgotten. **Journal of Banking and Financial Economics**, v.4, n. 2 , p. 5-17, 2015. Disponível em: <https://bit.ly/35htst0>. Acesso em: 20 out. 2020.

- REINHART, Carmen M.; ROGOFF, Kenneth S.; SAVASTANO, Miguel. **Debt intolerance**. 2003. Disponível em: <https://www.nber.org/papers/w9908>. Acesso em: 20 nov. 2020.
- REINHART, Carmen M.; ROGOFF, Kenneth S. **Serial Default and the "Paradox" of Rich to Poor Capital Flows**. NBER Working Paper, n. 10296, 2009a. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/0002828041302370>. Acesso em: 20 jun. 2021.
- REINHART, Carmen M.; ROGOFF, Kenneth S. **This time is different**. Princeton University Press, 2009b.
- REINHART, Carmen M. Default, Currency Crises and Sovereign Credit Ratings. **The World Bank Economic Review**, v.16, n.2, p. 151–70, 2002. Disponível em: <https://bit.ly/2HnQZ39>. Acesso em: 20 out. 2020.
- RODRIK, Dani. **The Globalization Paradox: Democracy and the Future of the World Economy**, New York: W.W. Norton & Company. 2011.
- RODRIK, D.; VELASCO, A. "Short-Term Capital Flows". **Annual World Bank Conference on Development Economics**, Washington, D.C.: World Bank. 1999. Disponível em: <https://www.nber.org/papers/w7364>. Acesso em: 08 dez. 2020.
- SACHS, Jeffrey. **Theoretical issues in international borrowing**. National Bureau of Economic Research, 1983. Disponível em: <https://www.nber.org/papers/w1189>. Acesso em: 20 out. 2020.
- SAVONA, Roberto; VEZZOLI, Marika; CIAVOLINO, Enrico. **A data-driven explanation of country risk: emerging markets vs. eurozone debt crises**. Documento de trabalho da SYRTO - Tomografia de risco sistêmico da União Europeia, n.3, 2015. Disponível em: <https://bit.ly/3jhWUUQ>. Acesso em: 22 out. 2020.
- SAVONA, Roberto; VEZZOLI, Marika. Fitting and forecasting sovereign defaults using multiple risk signals. **Oxford Bulletin of Economics and Statistics**, v. 77, n. 1, p. 66-92, 2015. Disponível em: <https://bit.ly/3msd5Ei>. Acesso em: 05 maio 2020.
- SIEGMANN, Arjen. Policy Lessons from Systemic Risk Modeling and Measurement. In: **Systemic Risk Tomography**. Amsterdã: Elsevier, p. 239-273, 2017. Disponível em: <https://bit.ly/2TcgCXv>. Acesso em: 20 out. 2020.
- SILVA, Anderson Caputo; DE CARVALHO, Lena Oliveira; DE MEDEIROS, Otavio Ladeira (Ed.). **Dívida Pública: a experiência brasileira**. Washington: Banco Mundial, 2009.
- STURZENEGGER, Federico; ZETTELMEYER, Jeromin. **Debt defaults and lessons from a decade of crises**. Cambridge: MIT press, 2006.
- VARIAN, Hal R. Big Data: New Tricks for Econometrics. **The Journal of Economic Perspectives**, v. 28, n. 2, p. 3-27, 2014. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/jep.28.2.3> . Acesso em: 01 fev. 2021.

VASCONCELOS, Bruno Freitas Boynad de. **Poder preditivo de métodos de Machine Learning com processo de seleção de variáveis:** uma aplicação às projeções de produto de países. 2017. Tese (Doutorado em Economia) – Universidade de Brasília, Brasília, 2017. Disponível em: <https://repositorio.unb.br/handle/10482/23995> . Acesso em: 11 mar. 2021.

WIJAYANTI, Rani; SAGITA, Rachmanira. Early Warning System for Government Debt Crisis in Developing Countries. **Journal of Central Banking Theory and Practice**, v.9, p.103-124, 2020. Disponível em: <https://sciendo.com/article/10.2478/jcbtp-2020-0025> . Acesso em: 11 out. 2021.

WITTEN, Ian H. et al. Practical machine learning tools and techniques. *In*: WITTEN, Ian H. et al. **Data Mining**. 4. ed. Cambridge, United States: Elsevier, 2016.

## ANEXO A

Modelo Global - detalhamento da árvore gerada para previsão do ano de 1990

```

Histórico de Default = 0
  Variação IED <= 255.285266: 0 (263.45)
  Variação IED > 255.285266
    Variação IED <= 263.079219: 1 (5.07/0.01)
    Variação IED > 263.079219
      Variação IED <= 1600.0: 0 (27.23)
      Variação IED > 1600.0
        Supervariação Cambial <= -199.689677: 1 (5.06)
        Supervariação Cambial > -199.689677: 0 (2.79)
Histórico de Default != 0
  FMI = 0
    Saldo Conta Corrente/PIB <= -3.93744
      Região = MED: 1 (15.19)
      Região != MED
        Região = WHD: 1 (99.52/3.33)
        Região != WHD
          Supervariação Cambial <= -100.170523: 1 (45.56)
          Supervariação Cambial > -100.170523
            Dívida Pública/PIB <= 44.379301
              Histórico de Default = 1
                Variação Anual Exportações <= 7.928955: 1 (52.29/1.66)
                Variação Anual Exportações > 7.928955: 0 (2.22)
              Histórico de Default != 1: 1 (15.19)
            Dívida Pública/PIB > 44.379301
              Exportações <= 3.485187
                Variação das Reservas <= -18.844036: 0 (3.88)
                Variação das Reservas > -18.844036
                  Freedom House Index = 4: 1 (10.13)
                  Freedom House Index != 4
                    Variação Anual Exportações <= 1.558378: 1 (15.74/0.55)
                    Variação Anual Exportações > 1.558378: 0 (2.22)
              Exportações > 3.485187: 0 (4.99)
          Saldo Conta Corrente/PIB > -3.93744
            Variação PIB <= 8.828156
              Região = MED
                Dívida Externa/PIB <= 28.433306: 1 (5.06)
                Dívida Externa/PIB > 28.433306: 0 (7.77)
              Região != MED
                Abertura Comercial <= 21.507078
                  Dívida Externa/PIB <= 11.762697: 1 (5.06)
                  Dívida Externa/PIB > 11.762697: 0 (11.1)
                Abertura Comercial > 21.507078
                  Região = WHD
                    Freedom House Index = 2
                      Juros EUA <= 5.97: 1 (10.68/0.55)
                      Juros EUA > 5.97: 0 (2.22)
                    Freedom House Index != 2: 1 (25.31)
                  Região != WHD
                    Supervariação Cambial <= -99.423182: 0 (4.99)
                    Supervariação Cambial > -99.423182
                      M2/Reservas <= 4.990805: 0 (3.88)
                      M2/Reservas > 4.990805
                        Inflação anual <= 7.080715
                          Juros EUA <= 11.39: 0 (4.44)
                          Juros EUA > 11.39: 1 (5.06)
                        Inflação anual > 7.080715
                          Juros EUA <= 6.67
                            Freedom House Index = 5: 0 (2.22)
                            Freedom House Index != 5: 1 (11.23/1.11)
                          Juros EUA > 6.67: 1 (46.67/1.11)
                      Variação PIB > 8.828156: 0 (7.77)
            FMI != 0: 0 (4.99)

```

Fonte: Elaboração do autor

## ANEXO B

Distribuição das variáveis (vermelho sinaliza *default*)



Fonte: Elaboração do autor

## ANEXO C

Código Java – Configurações dos Modelos Regionais:

```
#Seleção do algoritmo com a opção de uso de filtros
weka.classifiers.meta.FilteredClassifier -F "weka.filters.MultiFilter -F

#Seleção do filtro de balanceamento de classes e de remoção dos atributos 1 e 2 (nome
dos países e ano, para que o algoritmo não use estas informações)
\"weka.filters.unsupervised.attribute.Remove -R 1,2\" -F
\"weka.filters.supervised.instance.ClassBalancer -num-intervals 10\" -S 1 -W

#Seleção do algoritmo J48 (C4.5 em Java)
weka.classifiers.trees.J48 -- -B -C 0.25 -M 1
```



## ANEXO D

Código Java – Configurações dos Modelos Globais:

```
#Seleção do algoritmo com a opção de uso de filtros
weka.classifiers.meta.FilteredClassifier -F "weka.filters.MultiFilter -F

#Seleção do filtro de balanceamento de classes e de remoção dos atributos 1 e 2 (nome
dos países e ano, para que o algoritmo não use estas informações)
\"weka.filters.unsupervised.attribute.Remove -R 1,2\" -F
\"weka.filters.supervised.instance.ClassBalancer -num-intervals 10\" -S 1 -W

#Seleção do algoritmo J48 (C4.5 em Java)
weka.classifiers.trees.J48 -- -B -C 0.25 -M 2
```