UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE CIÊNCIA DA COMPUTAÇÃO

ARTHUR LAGES LENZ

Development and Comparison of Machine Learning Methods for Subjective Refraction Prediction

Work presented in partial fulfillment
of the requirements for the degree of
Bachelor in Computer Science

Advisor: Prof. Dr. Leandro Wives
Coadvisor: Aline Lutz Araujo

Porto Alegre
December 2018

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my family and friends, who have always supported me. Love to you all.

Special thanks, also, to Prof. Leandro Wives for guiding me throughout the development of this work, and to Dra. Aline Lutz, for all the patience when explaining the refractive disorders, the data collected by the TelessaúdeRS-UFRGS group, and for answering all the (many) questions I had during the time of this work.

RESUMO

De acordo com a estimativa apresentada no estudo realizado pelo Global Burden of Disease (Carga Global de Doença, em tradução livre), de 2016, erros de refração e deficiências no processo de acomodação visual afetam um total de 886 milhões de pessoas (43% de todas as deficiências relacionadas a orgãos sensoriais).

A fim de tratar este problema, oftalmologistas prescrevem lentes corretivas, capazes de corrigir os diversos tipos de erros refrativos: Miopia, Hypermetropia, Astigmatismo, e Presbiopia.

Para a prescrição das lentes, é necessário determinar o erro refrativo dos pacientes. Este processo é realizado em duas fases: uma estimativa inicial (chamada Refração Objetiva), e um ajuste-fino (chamado Refração Subjetiva). Este trabalho, busca aprimorar a qualidade da refração estimada, reduzindo o tempo necessário para a realização da etapa de ajuste-fino.

Este trabalho propõe então a utilização de *machine learning* para a predição da refração mais adequada (i.e. Refração Subjetiva) de cada olho, considerando não apenas a estimativa inicial (i.e. Refração Objetiva), como também outras características do paciente, como idade, sexo, e sintomas apresentados.

Um conjunto de dados coletados pelo grupo de telemedicina TelessaúdeRS-UFRGS, durante um período de 13 meses, foi utilizado. Dentre as três técnicas investigadas (Regressão Linear, Máquinas de Vetores de Suporte, e Redes Neurais), a melhor performance geral apresentada foi produzida pelo modelo preditivo baseado em redes neurais (0.406 dioptrias), reduzindo em até 70% o erro inicial produzido pelo equipamento Autorrefrator (de 1.381 dioptrias).

Palavras-chave: Inteligência Artificial, Aprendizado Supervisionado, Regressão Linear, Máquina de Vetores de Suporte, Redes Neurais, Oftalmologia, Optometria, Erros Refrativos, Vetores de Poder.

ABSTRACT

According to the estimate in the Global Burden of Disease study of 2016, visual impairments such as refractive errors and deficiencies in the visual accommodation process affect a total of 886 million people (43% of all sensory organ deficiencies).

To address this problem, clinicians prescribe corrective lenses, capable of correcting the many types of refractive errors: Myopia, Hypermetropia, Astigmatism, and Presbyopia.

In order to prescribe the lenses, clinicians must first determine the patient's refractive error. This process is performed in two steps: an initial estimate called Objective Refraction, and a fine-tuning step called Subjective Refraction. The current work seeks to improve the quality of the estimated refraction, reducing the time required to perform the fine-tuning step.

In order to reduce the time necessary for this process, this work proposes the use of *machine learning* for predicting the most adequate refractive power (i.e. the *Subjective Refraction*) of each eye, considering not only the subject's *Objective Refraction* information but also other characteristics such as the age, sex, and presented symptoms.

A data set containing information of over 3600 patients, collected by the telemedicine group TelessaúdeRS-UFRGS during a period of 13 months, was used.

Among the three techniques investigated (Linear Regression, Support Vector Regression, and Neural Networks), the best overall performance was presented by the predictive model based on neural networks (0.406 diopters), reducing by 70% the initial error produced by the Autorefractor (1.381 diopters).

Keywords: Artificial Inteligence. Supervised Learning. Linear Regression. Support Vector Regression. Neural Network. Ophthalmology. Optometry. Refractive Errors. Power Vectors.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ANN	Artificial Neural Network
API	Application Programming Interface
CNTK	Microsoft Cognitive Toolkit
GBD	Global Burden of Disease Study
GPU	Graphical Processing Unit
LogMAR	Logarithm of the Minimum Angle of Resolution
MAE	Mean Absolute Error
MLP	Multilayer Perceptron
MOD	Magnitude of Differences
OLS	Ordinary Least Squares
ReLU	Rectified Linear Unit
RBF	Radial Basis Function
RMSE	Root-Mean-Square Error
SGD	Stochastic Gradient Descent
SVM	Support Vector Machine
SVR	Support Vector Regression
UFRGS	Federal University of Rio Grande do Sul
VA	Visual Acuity
WHO	World Health Organization

CONTENTS

1 INTRODUCTION

The World Health Organization published in 2010 a study evaluating the severity of visual impairment and its causes. The organization reviewed publications over a period of 10 years, covering 39 countries, and estimated a total of 285 million people to be visually impaired. Additionally, the study demonstrated that nearly 80% of the causes are preventable, and that a total of 43% of them belong to uncorrected refractive errors (PASCOLINI; MARIOTTI, 2012).

Most recently, the Global Burden of Disease (GBD) study, from 2016, performed a systematic analysis and ranked the prevalence of 328 diseases and injuries over a total of 195 countries. In this study, refraction and accomodation disorders ranked 9th place, and affected an estimated total of 886 million people (accounting for 47% of all sense organ diseases) (VOS et al., 2017).

The GBD study also reviewed and ranked the diseases by a custom metric named Years Lived with Disability (YLD), and revealed sense organ diseases to be ranked second, with refraction and accomodation disorders ranking among the top 15 disorders, indicating that subjects are unable to reach treatment - perhaps due to its access not being adequately provided - and tend to live with the disorders for many years.

These studies make evident the major health issue that are sense organ diseases, particularly refraction and accomodation disorders. In order to correct for these refractive errors, clinicians can prescribe spectacles.

The assessment of refraction is usually made up of multiple measuring processes: characteristics such as *visual acuity*, *objective refraction*, and *subjective refraction* need to be evaluated so as to determine the need for corrective lenses.

According to the American Academy of Ophthalmology, due to the rapid advances in tecnology in the past years, newer diagnostic tools have been emerging, facilitating and accelerating the assessment of refractive errors by ophthalmologists and optometrists (AAO, 2014). However, the determination of subjective refraction, even with automated machines, still lacks accuracy, underperforming manual evaluation when performed by an experienced clinician (JORGE et al., 2005).

1.0.1 Goals

In an effort to help clinicians, this work proposes the application of machine learning techniques, so as to improve the accuracy and facilitate the process for subjective refraction determination of visually impaired patients, reducing the time required to perform the procedure, and consequently, reducing its associated costs.

For this, the current study will adhere to the following methodology: (i) application of machine learning algorithms for predicting subjective refraction of patients affected with refraction disorders, using data collected by the telemedicine group *Telessaúde RS-UFRGS*; (ii) evaluation of the performance of the predictive methods, in order to determine the most appropriate method for the previously mentioned data; (iii) validation of the results, determining whether they fall within a clinically acceptable range; and finally, (iv) discussion about the approaches taken in this study, commenting on obstacles and threats to validity.

1.1 Work Structure

This work is structured as follows: the next chapter (2) provides the medical concepts underlying the data used for training the supervised learning models; chapter three (**??**) discusses the use of AI in healthcare, and proposes the application of machine learning for subjective refraction prediction, while diving into the data itself, the problem, and the learning models used; chapter four (5) delves into the development process, explaining the chosen programming language and libraries, the environment configuration and problems faced during the implementation and application of the algorithms; chapter five (6) discusses the metrics chosen to evaluate model performance, and compares the results produced by the different learning models; lastly, chapter six (7) concludes this work commenting on results, limitations, and future work.

2 OPHTHALMOLOGY AND OPTOMETRY

For a complete understanding of this work, some important concepts must be covered. This chapter presents the medical concepts related to the visual system, the many anomalies that can be manifested in the human eye - in particular, refractive errors - and their treatment using corrective lenses, while also covering the mathematical models used to represent the refractive powers associated with such conditions.

2.1 Ophthalmology and Refractive Errors

The following subsections introduce the visual system, the refractive disorders that can be manifested in the human eye and how these are treated using corrective lenses.

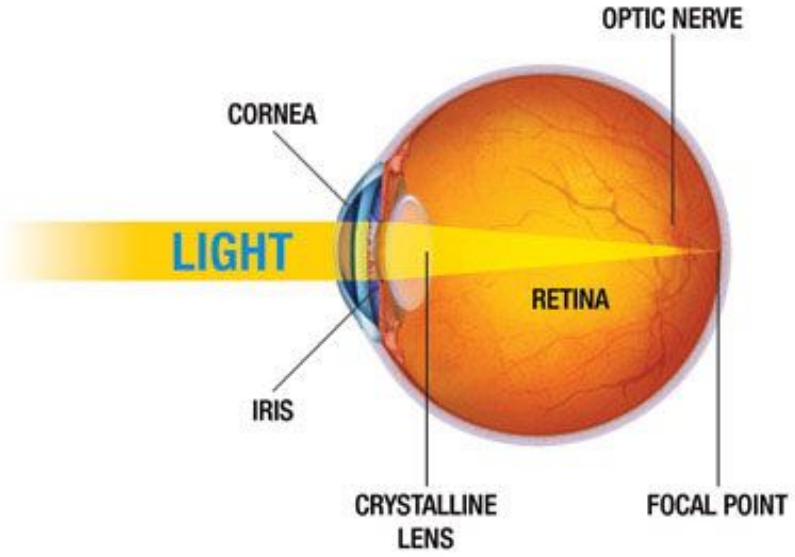
2.1.1 Light, the Human Eye, and Refractive Errors

The human eye is composed of three layers: the outermost layer, the middle layer and the innermost layer.

As shown in Figure 2.1, light rays enter the eye through the *cornea*, a transparent component of the outermost layer. The cornea's refractive power bends the rays towards the *pupil* (an opening present in the center of the *iris*) and through the *crystalline lens* (an element capable of changing its shape to adjust the focal distance of the eye), both in the middle layer. Next, the light reaches the innermost layer of the eye, also called *retina*. The two-dimensional image formed in the retina is captured by photoreceptors cells (cones and rods, for colors and luminance, respectively) and nerve impulses are then sent through the optic pathway to the brain, responsible for the visual perception process (PROBST; TSAI; GOODMAN, 2012).

The absence of refractive errors is called *emmetropia*, and an eye without refractive errors is said to be *emmetropic*. If refractive defects are present, however, the eye is said to have *ametropia* - or to be *ametropic*. These refractive errors are caused by anomalies in the shape of the eye, cornea or crystalline lens and lead to incorrect convergence points (i.e light rays not converging directly on the retina). These are categorized in the literature into the following groups: *Myopia*, *Hyperopia*, *Astigmatism*, and *Presbyopia* (PROBST; TSAI; GOODMAN, 2012).

Figure 2.1: Diagram showing how light enters the human eye.



From: National Keratoconus Foundation (2018)

2.1.2 Myopia

As aforementioned, myopia is associated with the incorrect convergence of light rays - specifically, when the point of convergence is located in front of the retina (as seen in Figure 2.2).

Myopia can be caused by multiple factors such as: eyes too long (relative to the focusing power of the cornea and lens); an overly curved cornea (causing light rays to bend more than they should, shortening the focal distance); or a crystalline lens too thick (resulting in a rounder shape, with higher refractive power, thus bending more the light - same as the cornea case).

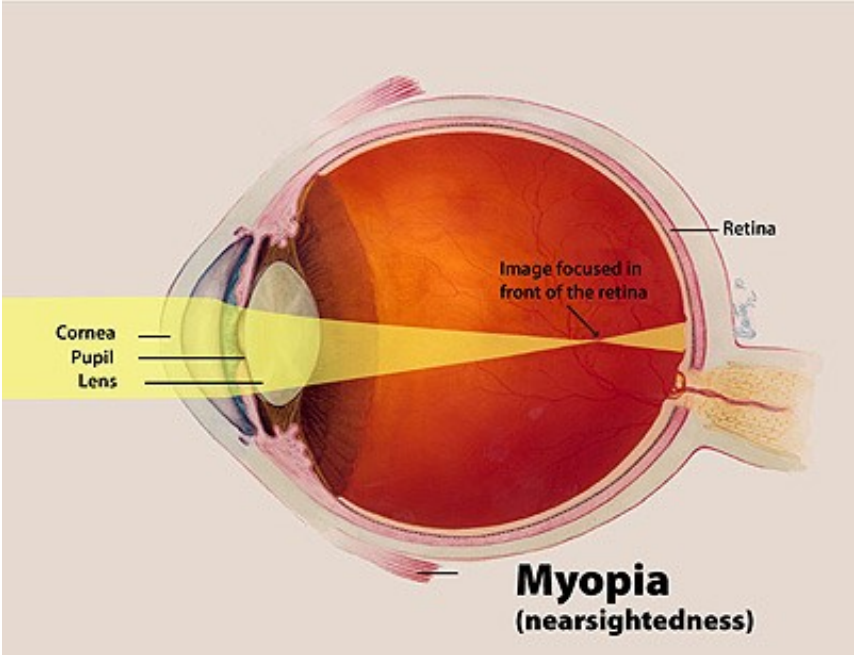
Subjects with such anomaly have problems to discern objects at distance. Hence, myopia is also referred as *Near-Sightedness* (PROBST; TSAI; GOODMAN, 2012).

2.1.3 Hyperopia

Hyperopia (also called *Far-Sightedness*, or *Hypermetropia*) is the anomaly of the eye responsible for increasing the eye's back focal distance, thus making light rays converge behind the retina (see Figure 2.3). Subjects affected with hyperopia have difficulties to distinguish shapes of objects at near distances (PROBST; TSAI; GOODMAN, 2012).

Contrary to the rest of the world, where other refractive errors are most common (S et al., 2008; KINGE; MIDELFART; JACOBSEN, 1998; MAVRACANAS et al., 2001), a

Figure 2.2: Diagram showing how light converges in front of the retina (Myopia).



From: National Eye Institute (2017)

study by Ferraz et al. (2015) shows that hyperopia is the predominant ametropia in Brazil. Such study is corroborated by others such as Garcia et al. (2005) (northeastern region) and Barros and Dias (2000) (center-west region). The latter, however, differs from the former as it ranks presbyopia as most predominant anomaly. Nonetheless, the study ranks hyperopia in second place, still prevailing myopia.

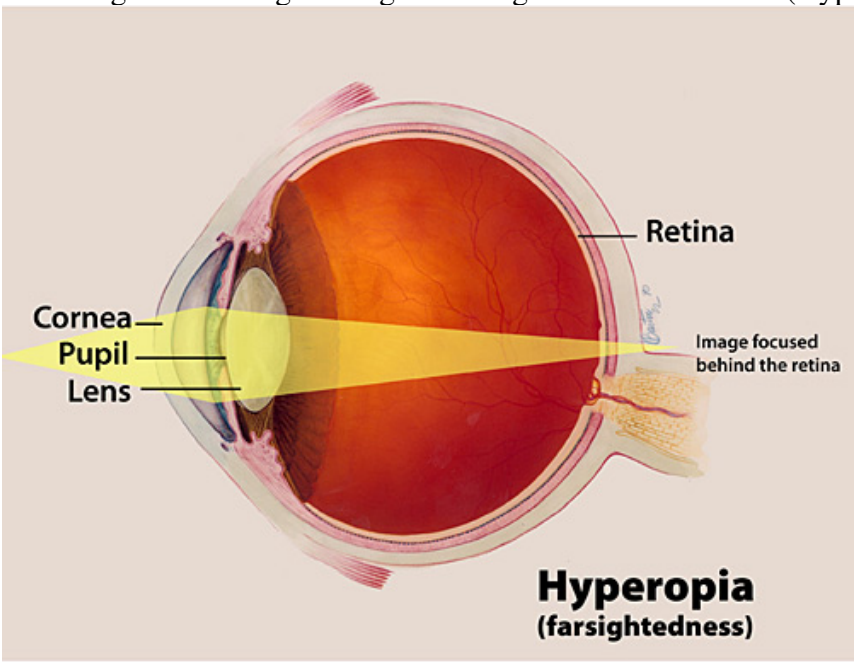
2.1.4 Astigmatism

If the cornea (or the crystalline lens) presents different refractive powers depending on the observed axis - thus, configuring an oval shape (see Fig. 2.4) - there is the occurrence of a condition called Astigmatism (PROBST; TSAI; GOODMAN, 2012).

Based on the axes of the steepest and the flattest meridians - i.e. axes with highest and lowest optical power, the anomaly is split into two groups: *regular* astigmatism - which consists of cases where the meridians are perpendicular to each other - and *irregular* astigmatism - when they are not. According to Probst, Tsai and Goodman (2012), regular astigmatism is also organized into sub-categories:

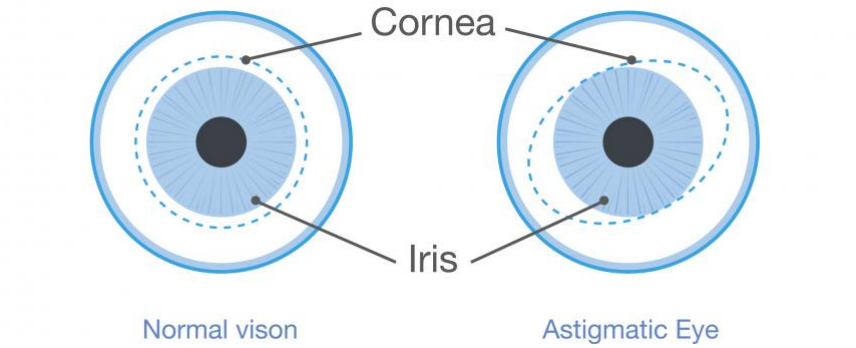
- *With-the-rule* astigmatism: when the vertical meridian has the highest refractive power.
- *Against-the-rule* astigmatism: when the horizontal meridian has the highest refrac-

Figure 2.3: Diagram showing how light converges behind the retina (Hyperopia).



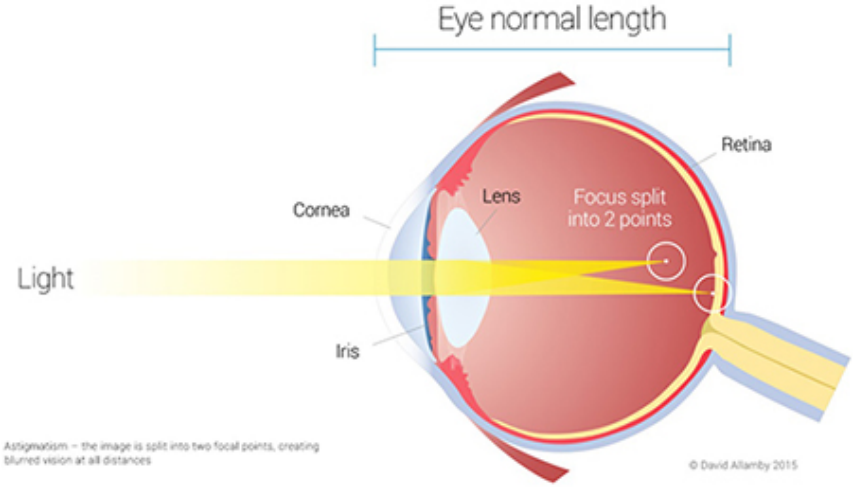
From: National Eye Institute (2016)

Figure 2.4: Comparison between a normal eye and an astigmatic one - notice the irregular shape of the cornea or lens.



From: Canadian Association of Optometrists (2017)

Figure 2.5: Convergence of light rays on two distinct focal points.



From: Anaheim Eye (2017)

tive power.

- *Oblique* astigmatism: when the highest refractive power lies between 120 and 150 degrees or 30 and 60 degrees.

As seen in Figure 2.5, each meridian bends the light differently, creating two distinct focal points, which then leads to blurry vision. Astigmatism can also be classified according to the position of each focal point: *simple* astigmatism - when only one focal point does not converge on the retina; *compound* astigmatism - when both focal points converge either in front of the retina or behind it; and *mixed* astigmatism - when one focal point converges in front of the retina, while the other converges behind.

2.1.5 Presbyopia

Presbyopia is the condition caused by the aging process of the eye, which leads to hardening of the crystalline lens, and therefore worsens the focusing capability of the lens (PROBST; TSAI; GOODMAN, 2012).

2.1.6 Corrective Lenses

In order to treat and correct the refractive errors caused by the aforementioned anomalies, corrective lenses can be prescribed.

Many characteristics are considered when crafting corrective lenses, such as *type* (how many distances does the lens correct), *optical profile* (the type of curvature for each

surface of the lens), and *materials* used in the manufacturing process (PROBST; TSAI; GOODMAN, 2012).

To clarify how lenses correct these refractive errors, *optical profiles* are briefly described next.

2.1.6.1 Optical Profiles

According to Probst, Tsai and Goodman (2012), each anomaly presents a different convergence error that needs to be compensated for:

- Myopia converges the light in front of the retina. In this case, is necessary to lenghten the focal distance, which can be done with divergent lenses - as indicated by the lenses' negative optical power.
- Now, for subjects with hyperopia, light rays are converged behind the retina. To compensate for this error, convergent (positive) lenses are then used, shortening the focal distance.
- Astigmatism presents a new scenario, where two different axes present different refraction powers each. Therefore, the prescribed corrective lenses must counterbalance each meridian's refractive error differently. The decision of which profile to use when correcting these errors follows the same principles described for myopia and hyperopia.

2.2 Diagnosis and Eyeglass Prescription

Ametropia diagnosis and lens prescription is done by evaluating both Central Visual Acuity and the refractive powers of the patient's eyes. Refractive power evaluation can be split into two processes: *Objective Refraction* and *Subjective Refraction* (KHURANA, 2003).

In the Objective Refraction process, the optical powers responsible for the incorrect convergence of light, caused by the anomalous conditions of the eyes (e.g. an elongated shape of the eye, in case of myopia) are estimated. This task can be accomplished either manually, by a procedure called *Retinoscopy* - which analyzes the reflection off the patient's retina - or automatically, using a machine called *Automatic Refractometer*.

Next, the ophthalmologist (or optometrist) combines relevant information such as clinical history, visual acuity, and objective refraction to determine a starting point for the

Subjective Refraction process. This process is responsible for defining the final and most adequate optical power needed to fix the subject's refractive errors, and therefore is the one used when prescribing corrective lenses to the patient.

2.2.1 Visual Acuity Assessment

Central Visual Acuity (VA) relates to the ability to distinguish different visual stimuli and covers a series of stages and components from the visual system - from the incidence of light rays on the cornea to the processing of these stimuli by the occipital lobe (LEVENSON, 1990). The assessment of such ability can be accomplished with the use of techniques such as the Snellen chart, and the LogMAR chart - either one can be used to measure near- and far-sightedness. In worse scenarios, alternative measurements such as finger counting, hand motion and light perception are applied (KHURANA, 2003).

2.2.1.1 The Snellen Chart

Herman Snellen introduced in (1862) the Snellen Chart (see Fig. 2.6), a tool intended for visual acuity assessment based on symbol recognition. The symbols - also called optotypes - are arranged in rows and must comply the following rules:

- Line thickness should be equal to the thickness of white spaces between lines.
- Optotype size must be five times the thickness of the line.
- Optotype size must also decrease for each subsequent row (increasing the difficulty in visual pattern recognition).

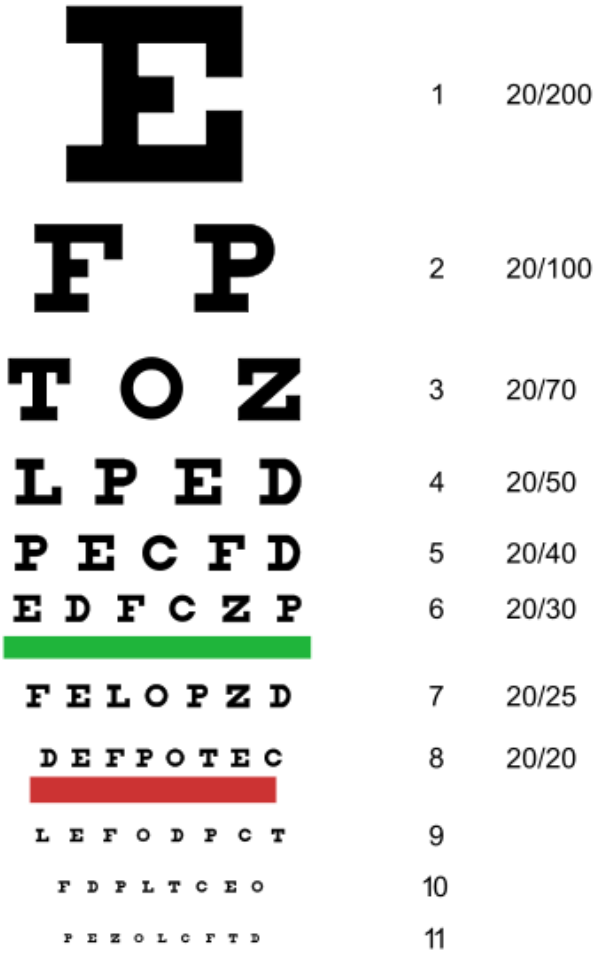
With exception of the optotype's size, each row must present the same difficulty level (i.e. the optotypes must be equally legible) (BAILEY; LOVIE, 1976).

The ability to discern visual patterns separated by a visual angle of one minute of arc, or optotypes that subtend an angle of five minutes of arc (since the optotype's size must be five times the thickness of the line) is named by Snellen as *Normal Vision*.

The standard distance for such test is 20 foot (or 6 meters) for distance acuity - as to approximate *optical infinity* (where parallel rays converge at the focal point of the lens - or in this case, direct on the retina, for normal vision).

Visual Acuity is then described using a nominator - which indicates the distance between the subject and the chart - and a denominator - which defines the distance in

Figure 2.6: Example of a Snellen Chart.



From: Wikipedia (2008)

which a subject with normal vision would discern the same optotype.

Therefore, subjects with normal vision are said to have *20/20 vision*.

Following this idea, 20/40 vision indicates lower visual performance: the smallest optotype discernible by the subject at 20 meters is discernible by someone with normal vision at 40 meters. On the other hand, 20/10 vision indicates a higher visual performance: the subject identifies at standard distance optotypes that can only be identifiable by subjects with normal vision at 10 meters.

In european countries, however, visual acuity description is required to be expressed as a decimal number (ISO Central Secretary, 2017).

2.2.1.2 LogMAR Chart

The LogMAR chart standardizes the design and use of optotype charts, eliminating some deficiencies from older charts (e.g. the Snellen Chart) and improving precision and reproducibility of measurements (BAILEY; LOVIE, 1976). For this reason, the use of the LogMAR chart is considered the gold standard technique for visual acuity measuring.

The chart is based on the *Logarithm of the Minimum Angle of Resolution* (i.e. the angle of the smallest visual pattern discernible by the subject's eye) - hence, LogMAR. Its usage of the minimum angle of resolution allows for conversion between Snellen and LogMar measurements (see Table 2.1).

Subjects capable of discerning visual patterns down to one minute of an arc (also called 20/20 vision, or normal vision) have a LogMAR score of 0 (since $log_{10}1 = 0$). Following this equation, patients with 20/40 vision have a LogMAR score of 0.3 ($log_{10}0.5 = 0.3$). Ergo, higher LogMAR scores indicate a lower visual performance, while lower ones indicate better performance (see Table 2.1).

In the LogMAR chart, both optotype sizing and spacing follow a logarithmic progression, decreasing for each row. In-between spacing is uniform, reducing errors caused by optotypes with reduced legibility (result of countour interaction - i.e having optotypes near one from another) (FLOM; WEYMOUTH; KAHNEMAN, 1963).

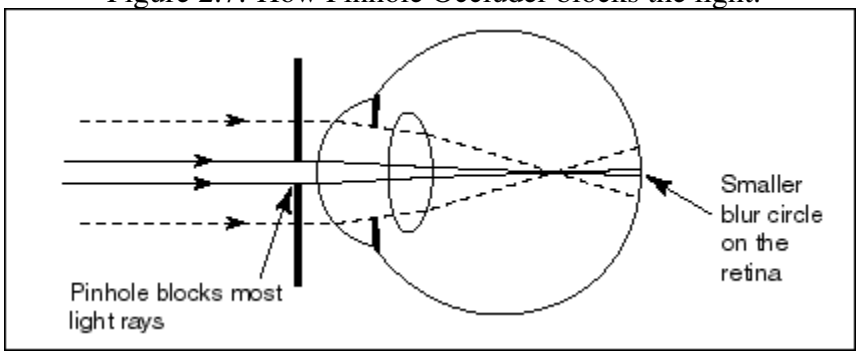
These characteristics improve the precision and reproducibility of measurements, and facilitate the use of non-standard distances for those cases where standard distance is not a viable option (BAILEY; LOVIE, 1976).

Table 2.1: Snellen - LogMAR Conversion Table

Foot	*Meter*	*Decimal*	*LogMAR*
20/200	6/60	0.1	1.0
20/160	6/48	0.125	0.9
20/125	6/38	0.16	0.8
20/100	6/30	0.2	0.7
20/80	6/24	0.25	0.6
20/63	6/19	0.32	0.5
20/50	6/15	0.4	0.4
20/40	6/12	0.5	0.3
20/32	6/9.5	0.63	0.2
20/25	6/7.5	0.8	0.1
20/20	6/6	1.0	0.0
20/16	6/4.8	1.25	-0.1
20/12.5	6/3.8	1.6	-0.2
20/10	6/3	2.0	-0.3

From: (KEIRL; CHRISTIE, 2007)

Figure 2.7: How Pinhole Occluder blocks the light.



From: International Myopia Prevention Association

2.2.1.3 Pinhole Occluder

Additionally, it is possible to measure the loss of visual acuity caused by such conditions by temporary correcting the subject's refractive errors with use of a *pinhole occluder* (BOWLING, 2015).

The pinhole occluder is a tool used to temporary correct refractive errors, and works by partially blocking the passage of light rays through the eye. It is an opaque tool, with a small circular hole in the middle - hence its name. As a result, when placed in front of the eye, only rays incident to its central portion are allowed through. Since these "central" rays are much less affected by the refractive power of the lens than those incident to peripheral regions, refractive error is reduced, causing the image to be formed correctly in the retina (see Fig. 2.7).

2.2.2 Retinoscopy and Autorefractor

To determine the *objective refractive error*, the ophthalmologist (or optometrist) analyzes how light is reflected off the patient's retina. This analysis can be done either manually - via *Retinoscopy* - or automatically, using an equipment called *Autorefractor* (KHURANA, 2003).

If done manually, the clinician uses an equipment called retinoscope to project light into the eye and observes the "behavior" of the reflection when the light is moved horizontally and vertically. Then, the clinician introduces correction lenses until the "reflection errors" caused by the refractive errors are corrected. When this is done, the clinician is able to estimate the optical power of the patient's eye.

As an alternative to the retinoscopy, an equipment called autorefractor can be used. The equipment emits infrared light to the eye and detects its reflection, analyzing the shape of the formed image in the retina - which can be either circular or oval, depending on the condition of the eye (see Figure 2.4). It automatically adjusts its magnification, simulating changes in the distance between the real image and the eye. As consequence, the equipment is capable of identifying the distance in which the projected image strikes the retina, and is able to approximate the eye's optical power, while doing it in a fast and repeatable manner.

The clinician then uses the measured objective optical power as a starting point to the *subjective refraction* process, in order to find the final and most adequade refractive power needed to correct the patient's refractive errors.

It is important to notice that the accuracy presented by the autorefractor can be outperformed by the manual alternative (when done by an experienced clinician) - presenting a trade-off between time and accuracy. (JORGE et al., 2005)

2.2.3 Subjective Refraction

The process of Subjective Refraction is responsible for determining the final and most adequate optical power needed to fix the subject's refractive errors - and therefore is the one used when prescribing corrective lenses to the patient (KHURANA, 2003).

Subjective Refraction relies on the patient's ability to *discern and communicate* possible improvements or distortions caused by corrective adjustments performed by the clinician, thus the naming of the procedure.

The clinician positions the subject in front of a visual acuity chart (e.g. the Snellen Chart shown in Fig. 2.6) and, introduces different corrective lenses, in order to balance out the subject's refractive errors. This procedure consists of three main steps: (i) correction of spherical error; (ii) correction of astigmatic error; and (iii) fine-tuning of the refractive correction, to provide the best optical performance while assuring the correction to be visually comfortable.

In order to facilitate the process of determinating the necessary refractive correction of steps (i) and (ii), objective refraction - measured via retinoscopy (see 2.2.2) - is used as starting point.

2.2.4 Static and Dynamic Measurements

The process of *visual accomodation* - i.e. the adjustment of the refractive power of the eye by altering the shape of the crystalline lens - occurs through contraction and relaxation of the *ciliary muscle*.

The ciliary muscle tone can influence the patient's focusing ability, and consequently, measurements such as central visual acuity and objective refraction may become inaccurate.

In order to avoid this issue, clinicians limit the patient's focusing ability by temporary reducing - or, in some cases, fully paralyzing - the movements of contraction and relaxation of the ciliary muscle, with the use of *cycloplegic eye drops*.

For this reason, measurements done with use of cycloplegic eye drops are named *static* measurements, while those that do not - and therefore allow for the visual accomodation process to happen - are said to be *dynamic* measurements (KHURANA, 2003).

2.3 Representational Models for Refractive Errors

2.3.1 Clinical Nomenclature

The standard notation for measuring refractions and prescribing correctors is called *Clinical Nomenclature* and its three components (S,C, and A) are directly related to the manufacturing of spectacles:

A spherical component (*S*) - meaning its value is equal in all meridians of the

eye, indicates the corrective power of the lens, in dipoters, used to correct near- and far-sightedness.

For those cases where there is presence of astigmatism, two extra components are specified: a cylindrical component (C), that corrects the uneven distribution of refractive error in the eye by introducing corrective power to a single axis, and consequently the degree associated with such axis (A).

Unfortunately, the sphero-cylindrical nomenclature is not suitable for graphical, numerical, and statistical analysis of optometric data (THIBOS; HORNER, 2001).

Several methods are available for the task of computing optical power of sphero-cylindrical lens combinations, such as the Power Cross, Prentice's formulas, Thompsons's graphical technique, Optometric Vectors (GARTNER, 1965), complex numbers, Power Vectors (HARRIS, 1997), and Power Matrices (KEATING, 1981).

It is shown in Thibos, Wheeler and Horner (1997), Harris (2007), that numerical and statistical analysis is a problematic task, that many of these tools are unable to perform (e.g. the computation of variance between multiple refractive powers). Such difficulty may be caused by the polar form used to describe astigmatism (HARRIS, 2007). In consequence of that, we explore two representational models that overcome this problem: *Power Vectors* and *Power Matrices*.

2.3.2 Power Vectors

Thibos, Wheeler and Horner (1997) attempt to solve the difficulties found in the traditional model (clinical nomenclature) through a representational model based on vectors, in which the components of the clinical nomenclature S (sphere), C (cylinder) and A (axis) are converted to a three-dimensional cartesian plane, and correspond to the dioptric power of three lenses (M: the spherical-equivalent; J_0: the vertical jackson-Cross-cylinder; and J_{45}: the oblique Jackson-Cross-cylinder).

The vector defined between the converted point and the origin is equivalent to the blurring strength of a sphero-cylindrical lens - hence the model's name: *Power Vectors*.

This vectorial representation allows for the use of conventional scalar methods on each component, facilitating statistical and numerical analysis of the refractive errors.

In consequence of that, Thibos shows that changes to the patient's refractive error can be measured by subtracting two vectors (one associated with the refractive power of the eye before the change, and one after).

Miller (2009) also contributes with two other important concepts: the MOD (Magnitude of Differences), related to the scalar resulting from the difference of two vectors - i.e. the magnitude of change of the optical blur; and the VDD (Vector difference in diopters), which normalizes the unit vector used by Harris (1997) and uses cylinder diopters instead of sphere diopters - which can be used to identify when spectacles should be changed, for example.

Representational models based on vectors, however, are applicable solely to *thin systems* (those in which there is no need to consider the refraction of each surface separately, since the distance between surfaces is neglegible when compared to the focal length), and are limited to simple problems that require only addition or multiplication with scalar values. To solve this limitation, the use of representational models based on matrices is suggested. These are able to represent both thin and thick systems, and are called Power Matrices (KEATING, 1981).

2.3.3 Power Matrices

Harris (1997) introduces a representational model of dioptric power in a three- and four-dimensional space called *Power Matrices*. The model - different from the vectorial model - is universal, since it can model both thin and thick systems. It does so by creating ray transfer matrices (i.e. a matrix that encapsulates all the optical characteristics of the system).

Although capable to generalize both thin and thick systems, the matricial system can be overly complex and not justifiable, in cases where only thin systems need to be modeled - as is the case of conventional spectacles. Therefore, this study opts to transform refraction data into Power Vectors notation.

Another factor that weighted the choice for the vectorial representation instead of the matricial one, was the lack of publications exploring practical applications of the Power Matrices model - in contrast to the vectorial model, present in Miller (2009), when is used for statistical analysis.

3 AI IN HEALTHCARE AND RELATED WORK

This chapter discusses the use of artificial intelligence algorithms in healthcare, as well as related publications.

3.1 Artificial Intelligence in Healthcare

Although present since the 60s, the field of Artificial Intelligence (AI) only became popular with the success of a supervised learning technique called *Multilayer Perceptron* (MLP) - a class of *Artificial Neural Network* (ANN) (MOOR, 2006). This rise in popularity was only possible due to a combination of two factors: (i) the development of faster Graphical Processing Units (GPUs) - which then improved drastically the performance of the learning algorithms, by allowing training on much deeper networks and (ii) the possibility to store much larger and richer labeled datasets - which also improved learning performance (JASON, 2017).

Nowadays, one of the most popular and well researched topics in AI for healthcare is medical imaging (GIGER, 2018) - notably, the dermatological classification of skin cancer (ESTEVA et al., 2017), and the detection of diabetic retinopathy (GULSHAN et al., 2016) and refractive errors (described in 3.2) through retinal fundus images. Other topics, however, have been emerging and getting attention in the field, such as robotic surgery (MAYER et al., 2008) and drug discovery (Benhenda, 2017).

3.2 Related Work

An exhaustive search - meaning that all results (described in Table 3.1) were inspected - on IEEExplore and MEDLINE, along with a 10-page search on Google Scholar returned, by the time of this writing, only three publications similar to this work's proposal: Varadarajan et al. (2018), Libralao et al. (2004) and Fageeri et al. (2017).

Other than these, no other publication targeting the problem of predicting subjective refraction was found.

Table 3.1 describes the searches performed on IEEExplore and MEDLINE, as well as the obtained results, and the similar publications found:

As previously mentioned, a similar work by Varadarajan et al. (2018) was found,

Table 3.1: Keywords used when searching for related work and number of obtained results.

Keywords	*IEEExplore*	*MEDLINE*
Artificial Intelligence Optometry	1/9	0/29
Artificial Intelligence Optometry Refraction	0/0	0/0
Artificial Intelligence Ophthalmology Refraction	0/0	0/15
Artificial Intelligence Lenses Refraction	0/2	0/6
Artificial Intelligence Refraction	0/42	0/25
Machine Learning Refraction	0/10	0/6
Machine Learning Optometry	1/10	0/16
Machine Learning Ophthalmology	1/38	0/204

Formatted as: Related Work / Results

proposing the application of deep learning algorithms for predicting refractive error. These, have shown a mean absolute error (MAE) of 0.56 diopters for estimating the spherical-equivalent (M) - presenting a reduction of almost 40% of the baseline expected MAE for the UK Biobank data set (0.91 diopters), and a reduction of approximately 65% of the baseline expected MAE computed for the AREDS data set (1.63 diopters).

This work, however, is not entirely comparable to the current work since its input data is made of retinal fundus images (while the current work deals with a csv formatted dataset of unstructured numerical and textual data). Also, the work proposed by VARADARAJAN et al. does not present astigmatic data, due to the lack of information about the toricity of the cornea or crystalline lens in retinal fundus images.

Lastly, Fageeri et al. (2017) investigates the performance of machine learning models for *classification* of refractive errors (myopia, hyperopia, and astigmatism) using techniques such as Decision Trees, SVM, and Naïve Bayes. As these models predict classes instead of the optical powers (the characteristics of the current work's problem will be further discussed in Sec. 4.2), direct comparison between works is not possible.

4 DATA AND PROPOSED APPROACH

This chapter discusses the current work's proposal: to investigate the applicability of three different techniques for subjective refraction prediction.

4.1 Data

This work uses data collected by the telemedicine group *Telessaúde-UFRGS* during a period of 13 months (from July 2017 to August 2018), and contains information of 8149 patients. Its access was approved by the local ethics committee.

From this dataset, the following features were extracted and used by the predictive models:

- Age
- Sex
- Race
- Symptoms [1]
- Visual Acuity data (LogMAR form - see 2.2.1):
  - Type (VA, CF, HM, LP, and NLP)
  - Acuity without correctors
  - Acuity with correctors (for patients that already use spectacles)
  - Acuity with Pinhole Occluder
- Objective Refraction measured by an autorefractor (Power Vector form - see 2.3.2):
  - Spherical Equivalent (M)
  - Vertical Jackson-Cross Cylinder (J_0)
  - Oblique Jackson-Cross Cylinder (J_{45})
- Subjective Refraction (Power Vector form - see 2.3.2):
  - Spherical Equivalent (M)
  - Vertical Jackson-Cross Cylinder (J_0)
  - Oblique Jackson-Cross Cylinder (J_{45})

[1] Symptoms were chosen by the clinicians from a predetermined set of options.

4.2 Problem and Supervised Learning

As explained previously (see 2.2.2), objective refraction is measured manually or automatically through the process of retinoscopy and is used as a starting point for the process of subjective refraction.

Subjective Refraction is necessary because manual and automatic retinoscopy are not always accurate and do not measure the most adequate optical power necessary to fix the subject's refractive errors. This procedure takes time, slows the process of diagnosis, and consequently reduces the rate of patients that can be diagnosed given a fixed amout of time.

In order to reduce the time necessary for this process, this work proposes the use of *machine learning* for predicting the most adequate refractive power (i.e. the subjective refraction) of each eye, considering not only the subject's objective refraction information but also other characteristics such as the age, sex, and presented symptoms.

Machine learning tasks are often classified into: (i) *Supervised Learning* - where the expected outcome (also called *target attribute*) is known and used in the learning process; (ii) *Unsupervised Learning* - models have no prior knowledge about the target attributes, and are commonly used to describe data; and (iii) *Reinforcement Learning* - in which the learning process happens through rewards, and depend on the actions taken by the algorithm (BISHOP, 2006).

Since the collected data (described in 4.1) provides subjective refraction measurements, supervised learning based methods can be applied.

As expressed by (THIBOS; HORNER, 2001), statistical analysis of directional data (e.g. astigmatism axis, one of the components of the clinical nomenclature - see 2.3.1) is fundamentally different from the analysis of nondirectional data, and the application of such conventional methods may lead to inaccurate results. In order to avoid this problem, refraction data is then converted from polar (Clinical Nomenclature) to cartesian coordinates (Power Vector).

The resulting Power Vector components (M, J_0, and J_{45}) are continuous variables. Being so, the predictive models proposed in this work must solve a regression problem.

4.2.1 Methods

To predict subjective refraction, the following methods were explored: (i) *Linear Regression* algorithms; (ii) a regression adaptation of Support Vector Machine called *Support Vector Regression*; and lastly, (iii) a simple feed-forward type of Neural Network called *Multilayer Perceptron*.

4.2.1.1 Linear Regression

Linear Regression models are those in which the predicted target value is a linear combination of the input variables, such as:

$$f(X) = \beta_0 + \sum_{j=1}^{p} X_j \beta_j$$

Being X the input vector, $f(X)$ the target value, and β_i the coefficients of the linear function (HASTIE; TIBSHIRANI; FRIEDMAN, 2001).

This study looks for the best linear method for each component in the Power Vector notation (M, J_0, and J_{45}). To do so, for each method presented below, multiple models are trained, validated, and tested.

The following linear regression methods were tested:

Ordinary Least Squares: fits a linear model minimizing the residual sum of squares between expected and predicted values - i.e. squared errors. Mathematically, it solves the following problem:

$$\arg\min_{\beta} \sum_{j=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} X_i j \beta_j)^2$$

Ordinary Least Squares (OLS) is the simplest method of linear regression and for that reason it does not address problems caused by more complex models, such as multicollinearity (a problem in which two features are highly linearly related, causing instability - i.e. making the model very sensitive to small changes - and consequently poorly generalizing the problem).

Model complexity can be viewed as: (i) a function of the weights of all features in the model; or (ii) a function of the total number of non-zero weighted features.

To address the complexity of our model and improve model generalization (thus, avoiding overfitting), two other methods were also explored: *Ridge Regression* and *Lasso*

Regression. Both performing *regularization*, a technique that introduces penalties to the estimated coefficients.

Ridge Regression: fits a linear model while minimizing the sum of the squared errors. Ridge differs from the simple OLS method by penalizing the coefficients through the use of *L2 regularization* - i.e. the sum of the *squared values* of all weights. Being λ the regularization parameter (i.e. how much weights are penalized), the Ridge method solves the problem below:

$$\arg\min_{\beta} \sum_{j=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} X_i j \beta_j)^2 + \lambda \sum_{j=1}^{p} \beta_j^2$$

Consequently, the Ridge method prefers solutions that do not depend too heavily on a specific feature.

Lasso Regression: fits a linear model while minimizing the sum of the squared errors. Like the Ridge method, its coefficients are also penalized. However, the Lasso method uses *L1 regularization* - i.e. the sum of the *absolute values* of all weights. Being λ the regularization parameter, the Lasso method solves the following problem:

$$\arg\min_{\beta} \frac{1}{2} \sum_{j=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} X_i j \beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j|$$

Consequently, the Lasso method produces sparse outputs. In other words, the method tends to prefer solutions with fewer non-zero weighted features. In other words, it performs *feature selection* (BISHOP, 2006).

A more stable variation of the Lasso algorithm, called *Lasso LARS*, is also explored. This variation implements the *Least-Angle Regression* proposed by Efron et al..

Lastly, a *bayesian ridge regression* method was also included in the experiment.

Bayesian Ridge Regression: like the previously described Ridge method, the bayesian alternative also fits a linear model minimizing the residual sum of squared errors, and also performs *L2 regulatization*. However, the Bayesian alternative includes the regularization parameter (λ) in the estimation process - in other words, it also learns how much the weights are penalized (BISHOP, 2006).

This model, however, presents a serious limitation: as the name itself states, linear models can only describe *linear* relationships between variables, meaning that nonlinear relationship (see Figure 5.8) won't be properly modeled, resulting in bad predictions.

The following methods deal with this problem by using *kernel tricks* - i.e. the

transformation of non-separable data into linearly separable data with help of nonlinear (*kernel*) functions, thus allowing for nonlinear relationships to be modeled (HASTIE; TIBSHIRANI; FRIEDMAN, 2001).

4.2.1.2 Support Vector Machines

A relevant machine learning model is the so called Support Vector Machine (SVM). First introduced by (CORTES; VAPNIK, 1995), the SVM model tries to separate linearly a set of data points by using a maximum-margin hyperplane, in which its distance to the nearest data point (on each side) is maximized.

A regression technique called Support Vector Regression Machine (SVR) was also included in the experiment. SVR (DRUCKER et al., 1997) is an alternative to the SVM model used in classification problems. The SVR model uses the same principles as the SVM one while including a margin of tolerance (ϵ), used to decide which samples to consider. In other words, the model tries to find a function $f(x_i)$ that differs no more than a margin of tolerance ϵ from the expected y_i (DEVELOPERS, 2007).

This study looks for the best SVR model for each component in the Power Vector notation (M, J_0, and J_{45}) - furtherly detailed in 6.2. To do so, multiple models are trained, validated, and tested (using the *RandomizedSearchCV* class of the *scikit-learn* library).

4.2.1.3 Neural Networks

Neural Network is a model first introduced by Rosenblatt, in 1958 (ROSENBLATT, 1958). It is inspired by biological neural networks, and is composed by layers of interconnected *neurons*. Each neuron receives one or more signals (the *input*), transforms it through the application of an *activation function* and sends the result forward to other neurons (just like synapses would in a biological setting) (HASTIE; TIBSHIRANI; FRIEDMAN, 2001).

The popularity of the model, however, only started to take off after the development of the *backpropagation* algorithm, which provided an efficient way to distribute the error term back through the layers, hence its name.

For this study, a Multilayer Perceptron (MLP) - a feedforward and fully-connected Neural Network - was implemented, and a Rectified Linear Unit (RELU) - shown by (GLOROT; BORDES; BENGIO, 2011) to enable better training when compared to hyperbolic tangent neurons - was selected as the activation function.

Contrary to the previously described models, Neural Networks can predict multiple outputs by having multiple neurons on the output layer. To that end, only one model needs to be trained, validated and tested, in order to predict all three components.

5 DEVELOPMENT

5.1 Language and Frameworks

Due to the existence of well documented and reliable machine learning libraries, *Python 3* (specifically v3.6.7) was selected as main programming language for this study.

For data manipulation and machine learning related tasks, the following tools were chosen: (i) *NumPy* and *pandas* were used during the preprocessing of the original dataset and plotting of the obtained results; (ii) the machine learning library *scikit-learn* was used for the development and application of machine learning methods, as it supplies an extensive collection of off-the-shelf models (accelerating the development phase) and provides methods for hyperparameter optimization (through *randomized search*) and overfitting prevention (via *cross-validation*); (iii) lastly, the *Keras* library was also selected, providing a high-level API that allows for an easy configuration and application of neural networks - these, implemented by low-level libraries (or *backends*) such as *TensorFlow* (the backend used in this study), *Theano*, or the *Microsoft Cognitive Toolkit - CNTK*.

For the sake of organization, each predictor was trained, validated, and tested within a different *jupyter notebook* and grouped into folders according to their respective methods (see 4.2.1). For those cases in which multiple predictors were needed (one for each Power Vector component), an extra notebook was created, combining the predictors with best performance, so as to compute an overall performance score (MOD - described in 6.1).

5.2 Environment

This study was performed in an Inspiron 14 7400 machine, under the following configuration: Intel® Core™ i5-7200U CPU @ 2.50GHz x 4, Memory 8 GB, SDD 120 GB, OS Ubuntu 18.04.1 LTS (64-bit).

5.3 Preprocessing

Due to the some missing or invalid values (common to unstructured data), a few transformations had to be done to the original dataset before its use in the learning phase:

First, each instance (patient data) was partitioned into two new instances, one for each eye - as these can present different refraction values, which are unrelated one to the other;

Secondly, all instances in which there was no need for corrective lenses to be prescribed were excluded from the dataset.

Then, eyes marked with Pterygium - a condition that causes tissue growth on the cornea - were removed, since this condition introduces distortions in autorefractor measurements, and would add noise to training data;

Before converting objective (autorefractor) and subjective refraction data from Clinical to Power Vector form, sphere and cylinder components were winsorized - i.e. extreme values were saturated, in order to mitigate the impact caused by outliers (e.g. mistyped values) in the learning process;

Axes were also rounded to their nearest multiple of 5. The decision to apply such transformation was based on the same action being performed by clinicians when prescribing corrective lenses - generally, the axis prescribed is rounded to the nearest multiple of 5 (or 20th), as to facilitate the manufacture process of spectacles. Being so, any measurement taken by the autorefractor that deviates from the expected subjective refraction by less than 5^{o} is not considered an error.

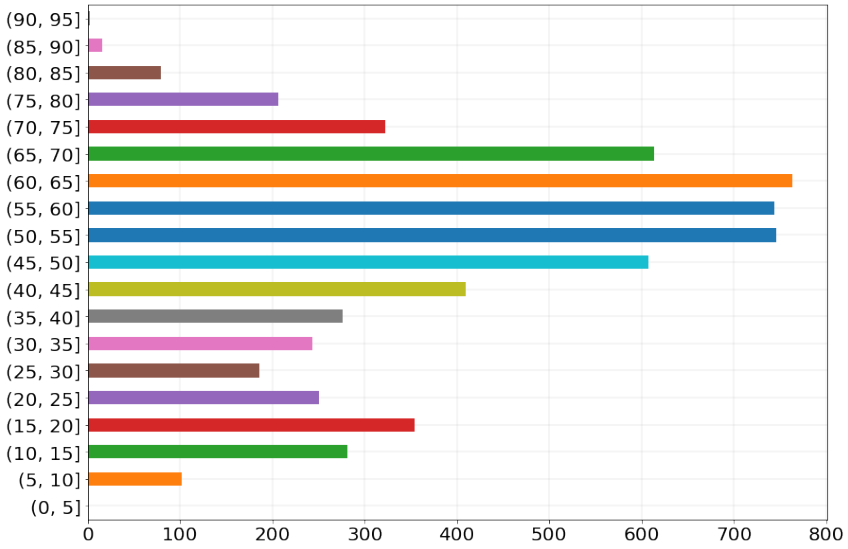
Visual Acuity data was converted from the traditional format (described in 2.2.1.1) to LogMAR format (see 2.2.1.2);

Patients with low visual acuity (CF and HM) were converted according to studies from Lange et al. (2009) and Schulze-Bonsel et al. (2006): entries with visual acuity of type *Counting Fingers* (CF) were converted to 0.014 LogMAR; while patients with VA of type *Hand Motion* (HM) had their visual acuity value set to 0.005 LogMAR. Entries with even lower visual acuity types - i.e. light perception (LP) and no light perception (NLP) - were excluded, since no direct conversion to LogMAR was found in the literature for such cases;

Categorical variables such as sex, race, and symptoms were one-hot-encoded, in order to be easily identifiable by the predictive models;

Both *static* and *dynamic* autorefractor measurements were stored into the original dataset. In order to predict the subjective refraction of each patient, the most adequate measurement for each patient had to be selected. This selection was performed manually by an experienced ophthalmologist.

Figure 5.1: Distribution of age between instances of the dataset.



From: Author

5.4 Processed Data

By the end of the preprocessing phase, the dataset presented the following characteristics:

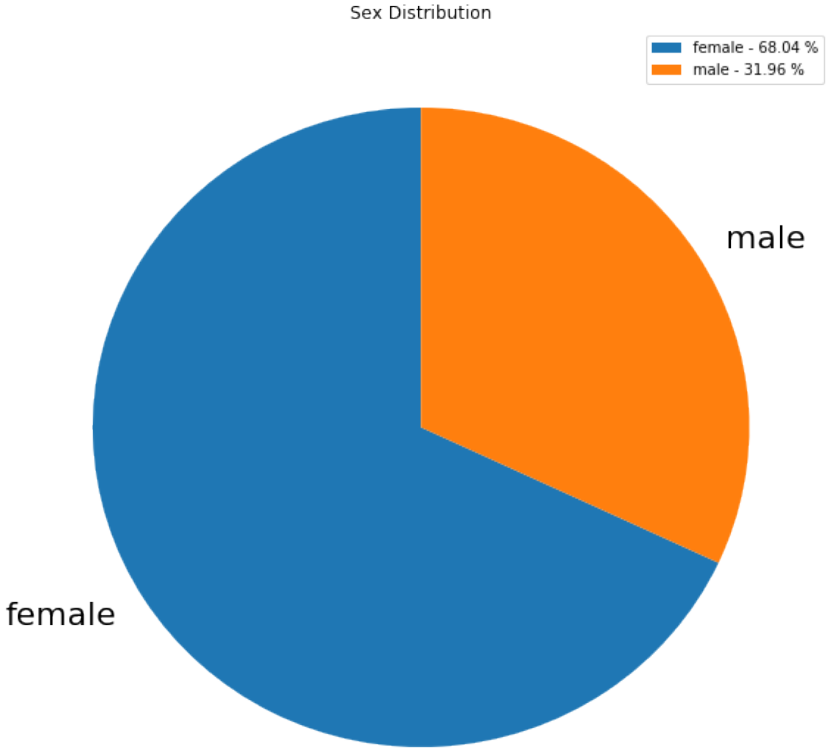
Table 5.1: Data collected by TelessaúdeRS-UFRGS from Jul/2017 to Aug/2018.

Characteristic	*Type*	
Number of instances (eyes)*		6,202
Mean Age (SD)		49.08 (18.60)
Sex (% female)		68.04
Ethnicity (%)	White	77.65
	Mixed (Pardo)	15.32
	Black	6.56
	Asian (Yellow)	0.32
	Indian	0.15
Already using correctors (%)		39.20

*From a total of 3,687 subjects.

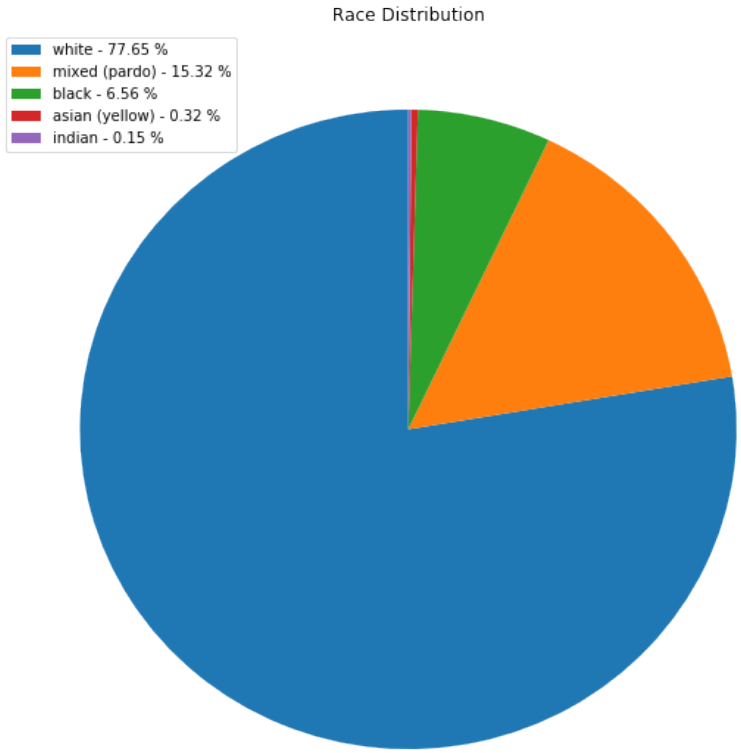
As figures 5.4 and 5.7 show, most of the intrisic error present in the autorefractor can be pointed to Spherical and Cylindrical components. This can be explained by the rounding of the Axis component, performed by the clinician when prescribing the corrective lens (described in 5.3).

Figure 5.2: Distribution of sex between instances of the dataset.



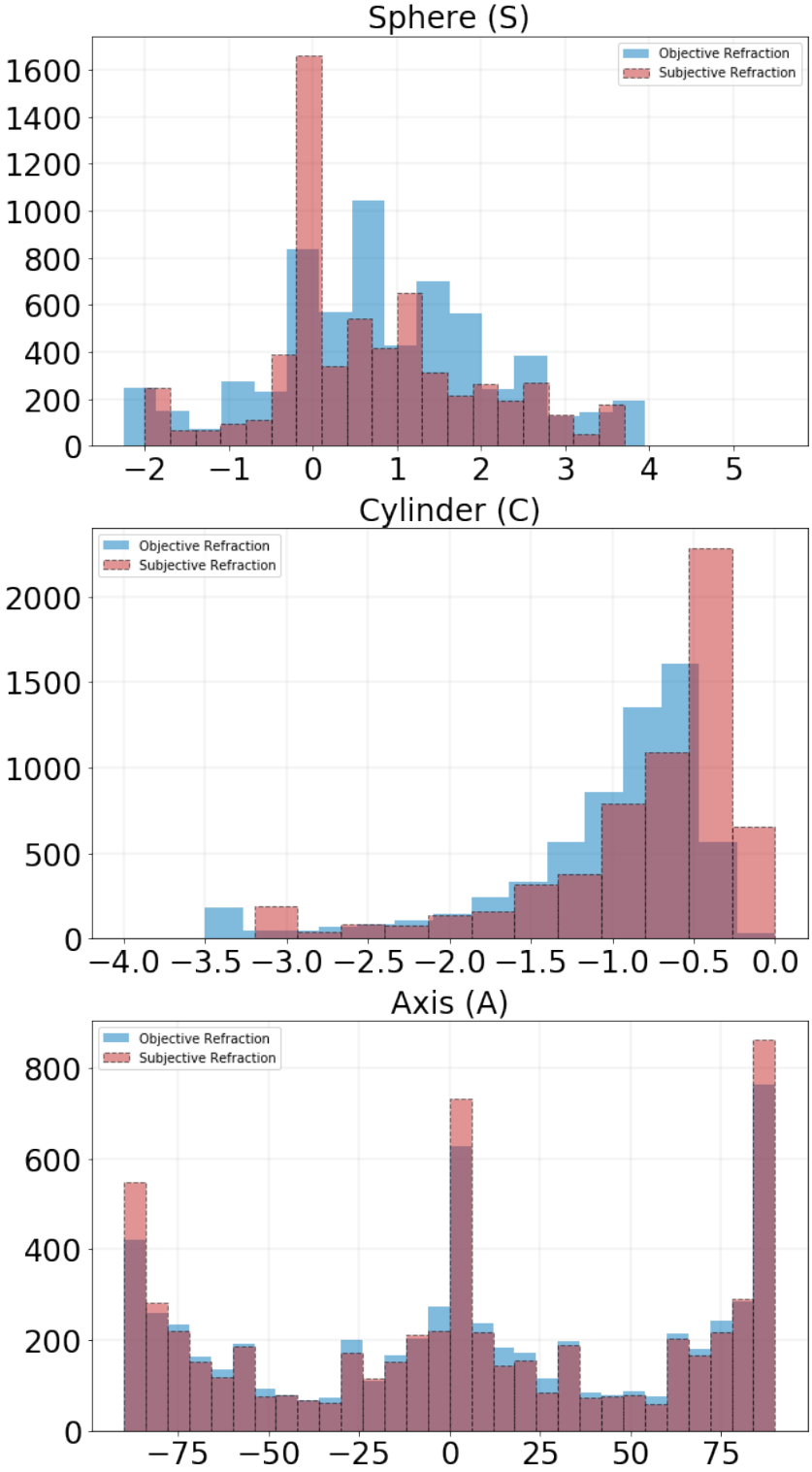
From: Author

Figure 5.3: Distribution of race between instances of the dataset.



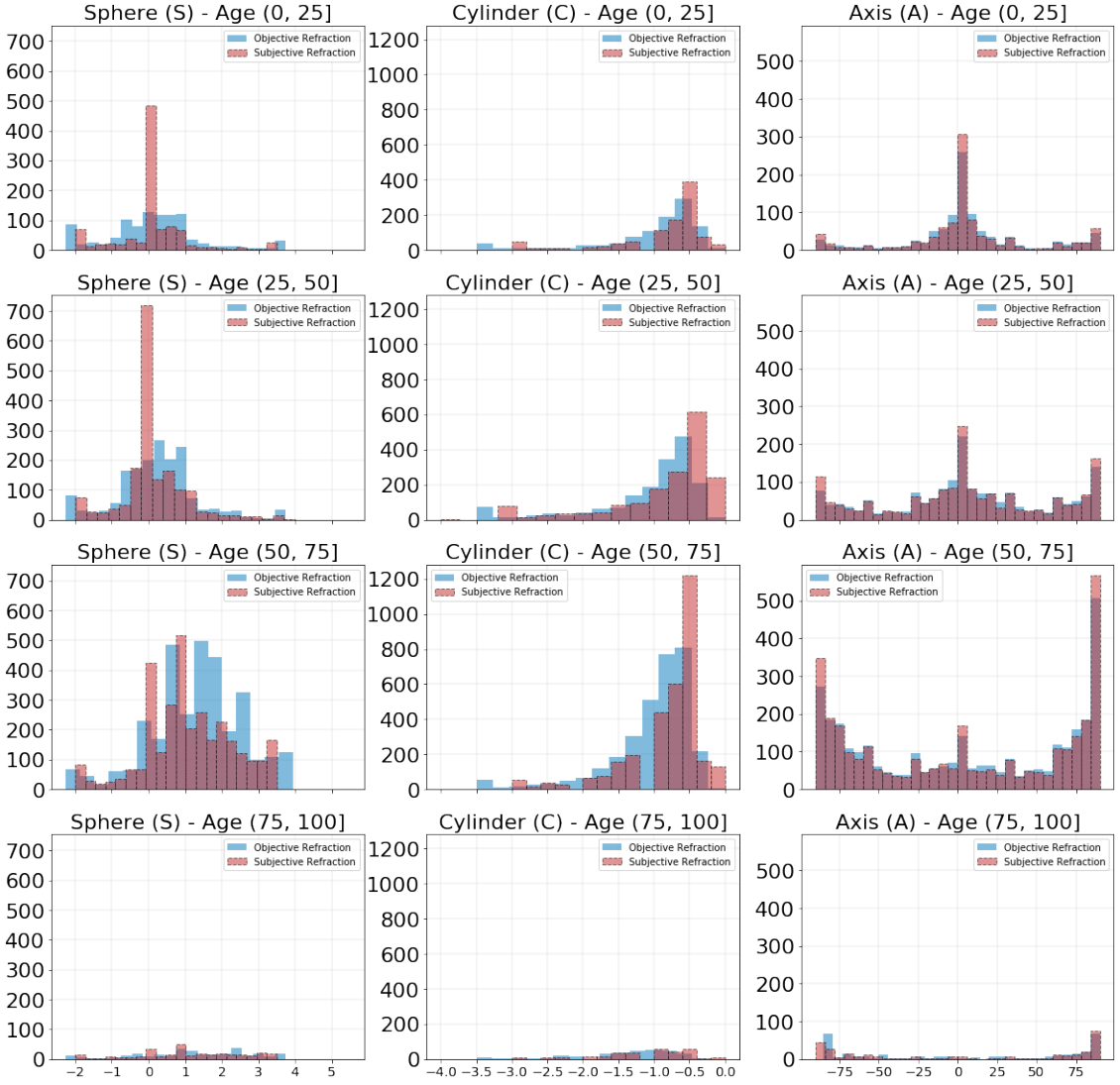
From: Author

Figure 5.4: Refraction data histograms (by component in Clinical Notation)



Objective refraction measured by the autorefractor (blue) and expected subjective refraction (red)
From: Author

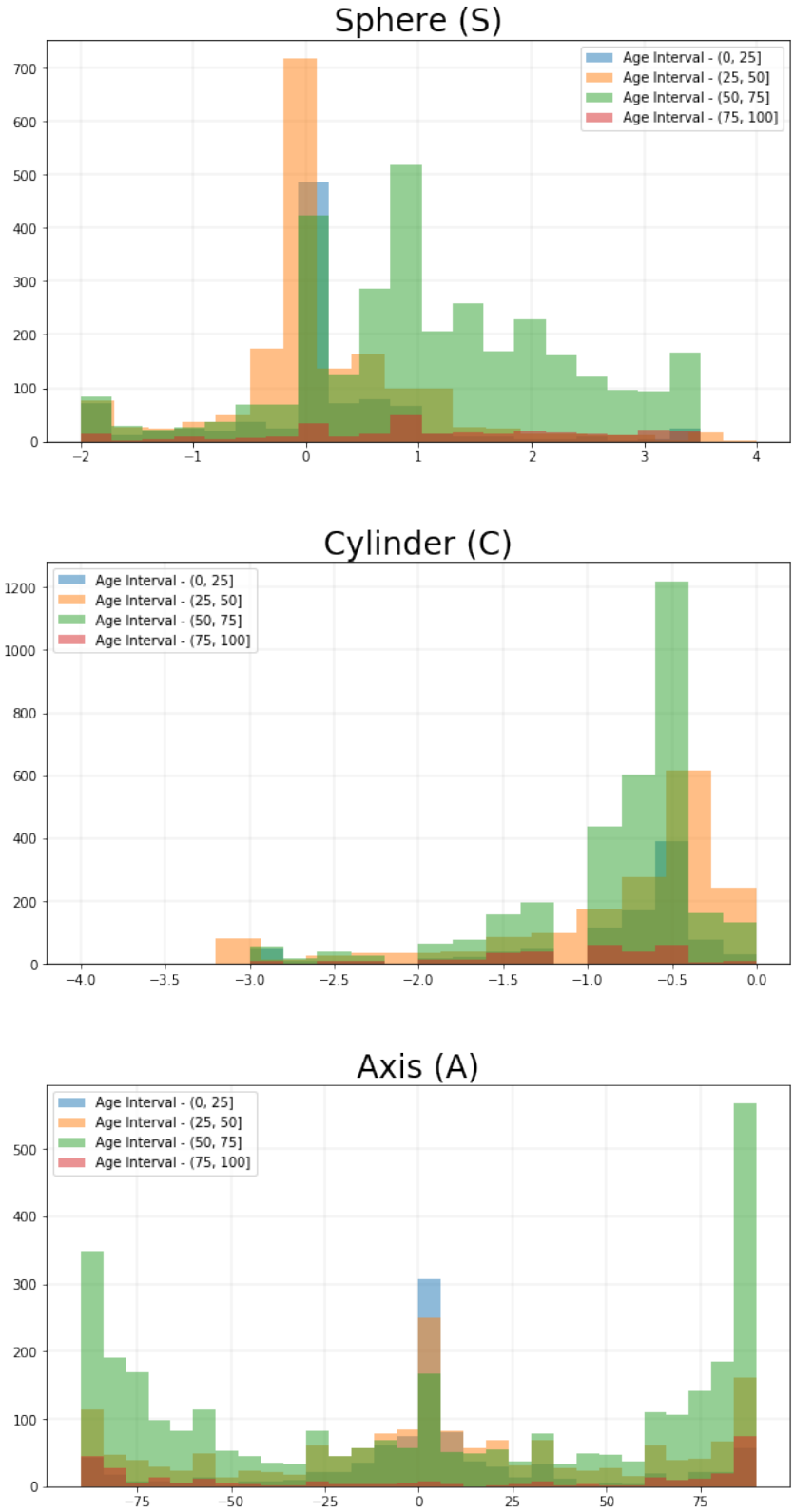
Figure 5.5: Refraction data histograms, grouped by Age (in Clinical Notation)



Objective refraction measured by the autorefractor (blue) and expected subjective refraction (red) Notice the drop in accuracy when measuring refraction for older patients and how both spherical and cylindrical components show similar skewness throughout the different age groups.

From: Author

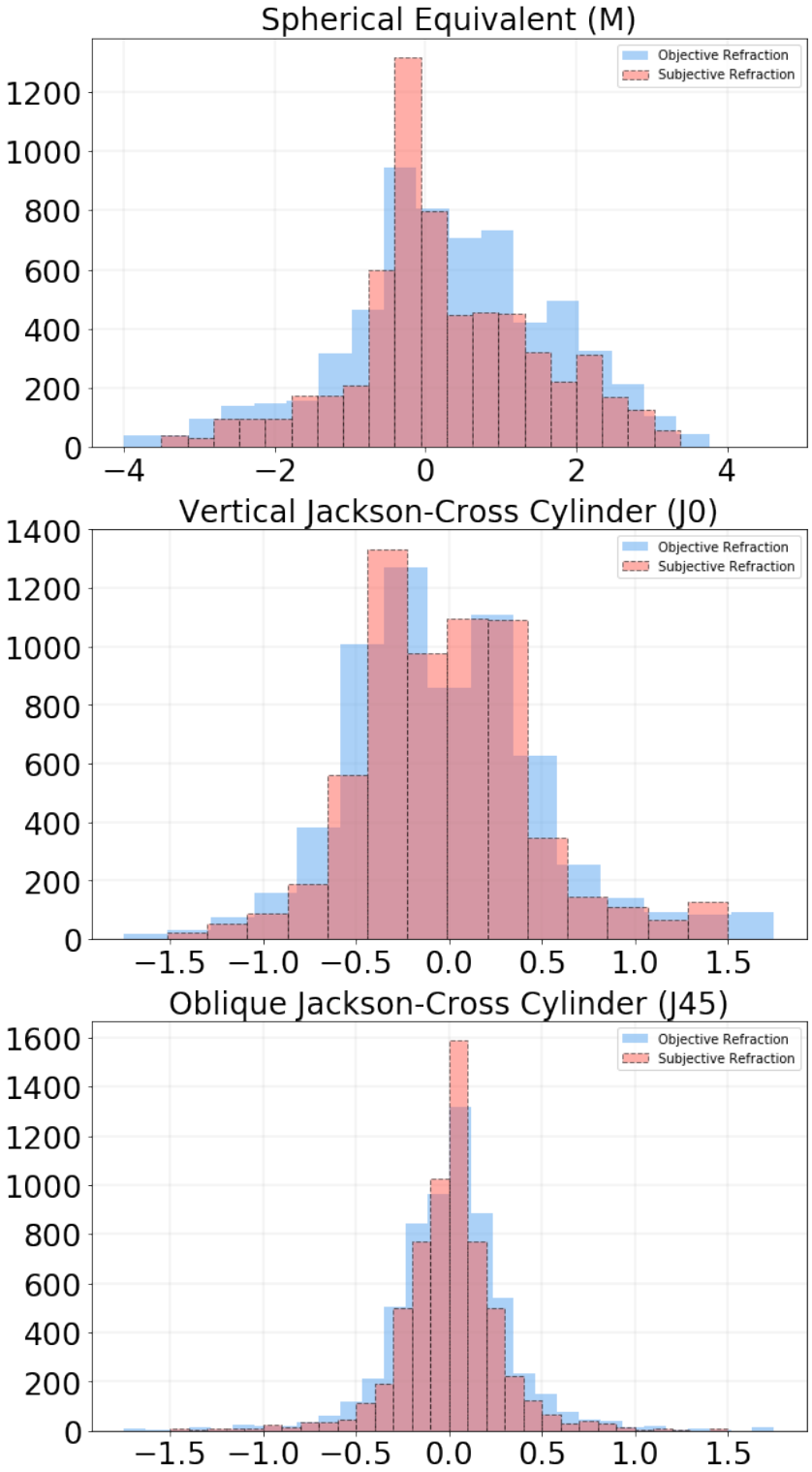
Figure 5.6: Distribution of Subjective Refraction, grouped by Age (Clinical Notation)



As expected, histograms show a clear difference between young and older subjects, as older patients present higher Spherical and Cylindrical refraction errors (distributions skewing towards higher values).

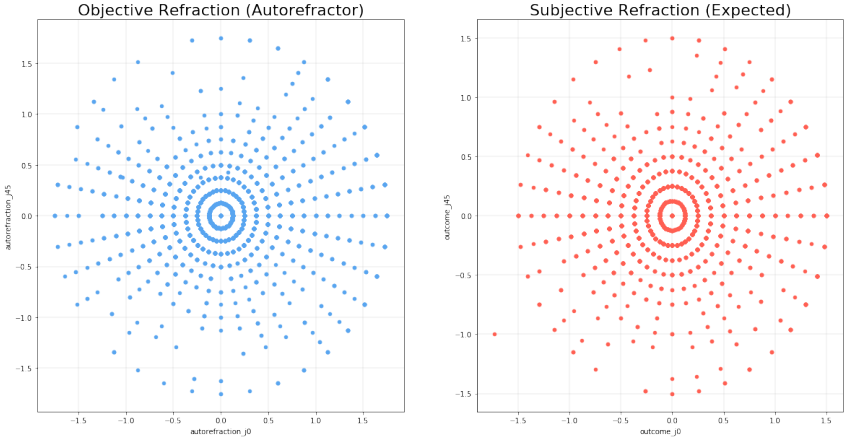
From: Author

Figure 5.7: Refraction data histograms (by Power Vector component)



Objective refraction measured by the autorefractor (blue) and expected subjective refraction (red)
From: Author

Figure 5.8: Refraction data (J_0 vs J_{45})



From: Author

Table 5.2: Difference between Objective and Subjective Refraction

Characteristic	*Type*	
Refraction Change (Power Vector)		
	Mean M (SD)	-0.0869 (0.4389)
	Mean J0 (SD)	-0.0043 (0.1720)
	Mean J45 (SD)	-0.0081 (0.1298)
Magnitude of Differences	Mean (SD)	0.3515 (0.3509)

Also presented in Figure 5.9 for better visualization.

5.5 Data Partitioning and Data Stratification

The preprocessed data was partitioned into two disjoint sets: (i) a Training and Validation set, used in the model selection phase; and a (ii) Test set (commonly named *Holdout-set*), used for a final evaluation of the best predictor (found in the model selection phase).

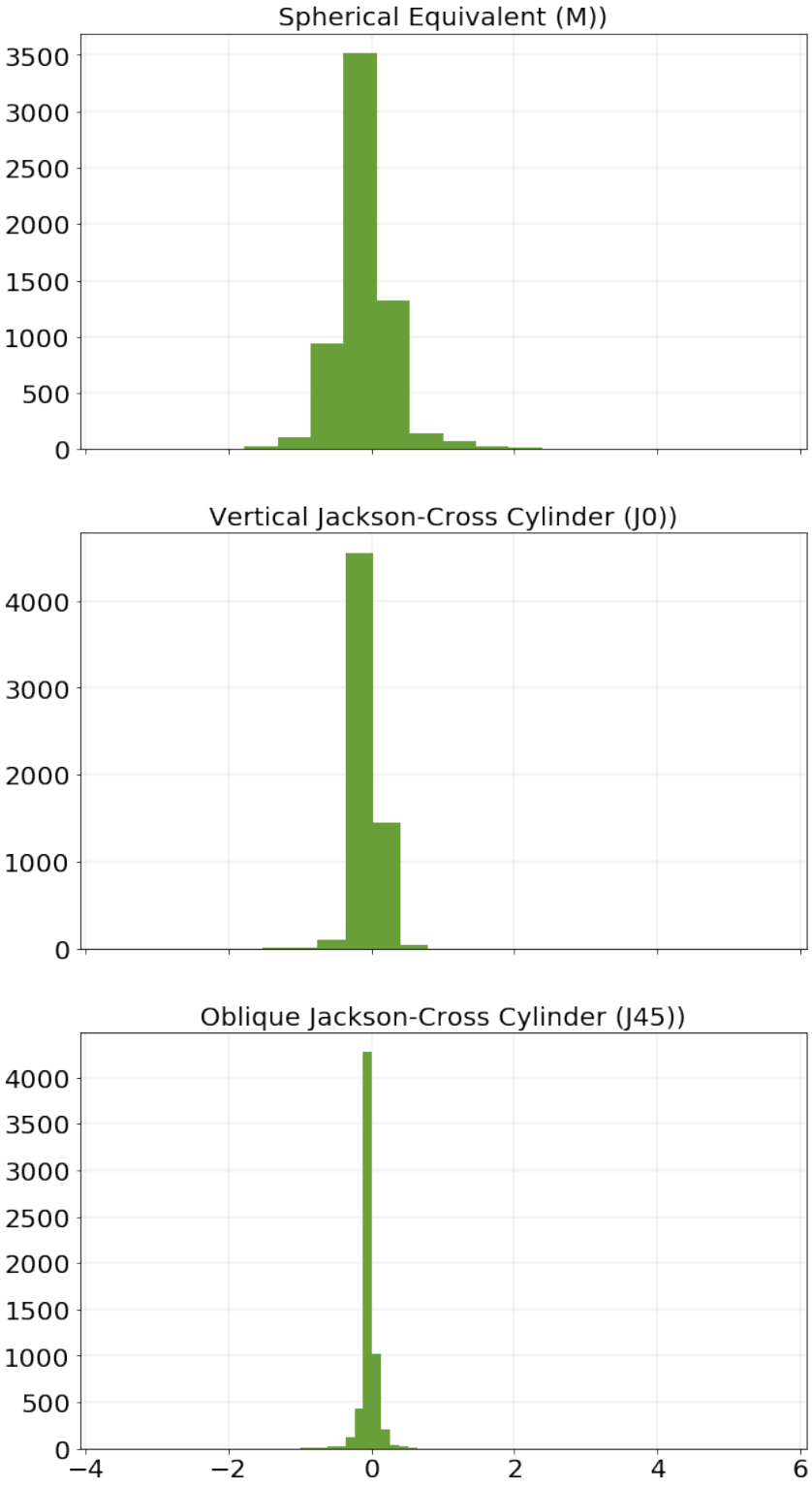
This partitioning is necessary in order to avoid the problem of overfitting: if there is no partitioning (i.e. model training and testing are performed on the same instances), or if the partitioned training and test sets are not mutually disjoint (i.e. sets share at least one element with each other), the resulting performance is most likely to be overestimated due to poor generalization. In other words, the trained model memorizes the information used during training and is, consequently, incapable of predict on unseen data.

After partitioning, sets consisted in: Training/Validation set (70% of the preprocessed dataset) and Test set (30% of the preprocessed dataset).

Additionally, so as to reduce variance between results during model selection (further explained in 6.2), *data stratification* is performed.

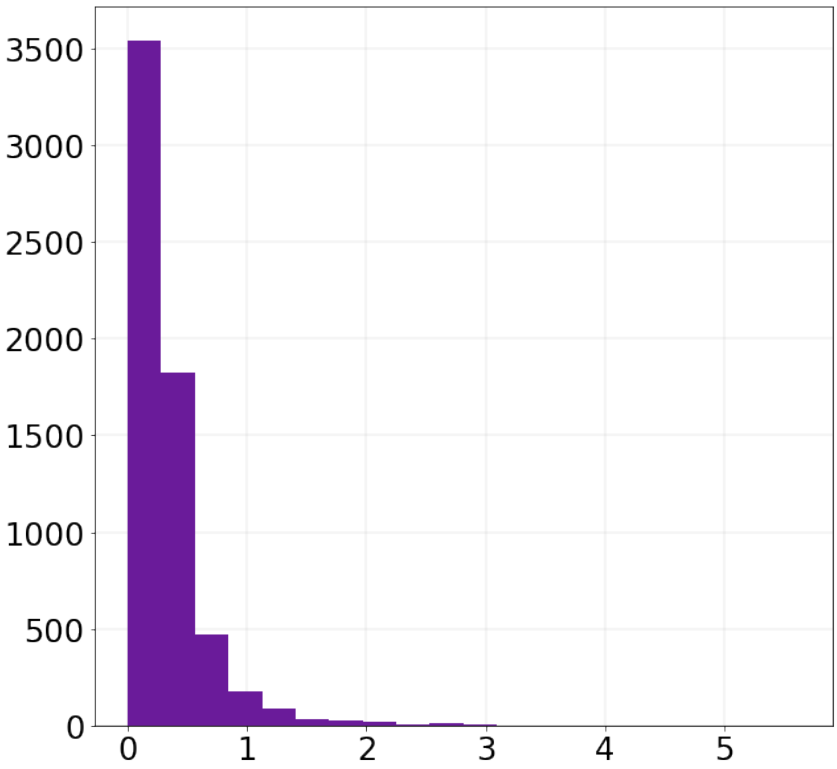
Data stratification is the process of separating data into smaller chunks, also called *strata* - hence the naming - and is used to balance out the distribution of classes between

Figure 5.9: Differences between objective refraction [Autorefractor] and subjective refraction [Expected] (by Power Vector component) - Histogram



Detailed in Table 5.2.
From: Author

Figure 5.10: Differences between objective refraction [Autorefractor] and subjective refraction [Expected] (MOD) - Histogram



From: Author

folds, producing similar results during cross-validation and, as mentioned before, reducing performance variance.

In order to stratify the multiple target attributes (the three Power Vector components), a few transformations were needed: first, the continuous components were discretized into bins and labeled accordingly; then, for each instance, labels were groupped together into triplets (a Python 3-tuple object); lastly, objects were then used for the partitioning of the preprocessed data (using the *train_test_split* method, provided by the *scikit-learn* library) and for model selection (using 10-fold cross validation).

6 RESULTS

The set of metrics chosen to evaluate model performance, as well as the performance results themselves - produced both during validation and test by the methods previously described in chapter **??** - are presented in the sections below, being followed by comments on the best overall model and its performance.

6.1 Performance Metrics

The performance evaluation of the predictive models is done through the use of the Root-Mean-Square Error (RMSE) metric. In other words, RMSE computes the average of the squared values of the prediction errors. Prediction error is the difference between the expected value and the value predicted by the model. Therefore, the closer to zero (meaning no prediction errors at all), the better the model.

Unlike the Mean Absolute Error (MAE) - another popular metric for performance evaluation of regression models - where each error contributes proportionally to its magnitude to the final result, RMSE squares all errors, causing larger errors to have a much greater influence to the final result.

It is important to point out that models trained with RMSE try to avoid predictions with large amounts of error and, in order to do so, might introduce some distortion to other predictions. The decision of using RMSE is then justified by the perspective of the patient, who desires the diagnosis process to be as quick as possible - hence, predictions with large amounts of error are avoided.

The metric is used to describe two different aspects of this work: (i) model's ability to learn and predict refractions in Power Vector form (by component); and (ii) overall model performance (using the blur strength metric proposed in Harris (1997)).

First, RMSE is calculated for each Power Vector component (M, J_0, and J_{45}), between the following pairs of vectors: (i) objective and subjective refraction - this is the error associated with the autorefractor, which we seek to reduce through the use of machine learning algorithms. Therefore, this will be considered as the *worst acceptable result*. Any predictive model with greater RMSE - i.e. worse performance - is automatically discarded; and (ii) predicted and subjective refraction - the error associated with the predictive model (or, how far are the model predictions from the actual subjective refractions).

Also, to facilitate the interpretation of the results, the *Magnitude of Differences* (MILLER, 2009) is calculated for both: (i) objective and subjective refraction; and (ii) predicted and subjective refraction. This single value (usually abbreviated as MOD) can be seen as the Euclidian distance between two points in the dioptric space, and represents the overall difference between two refractions.

6.2 Model Selection

A series of tests were performed, in order to investigate and find the best hyperparameter configuration for each method (described in 4.2.1).

Each method presented in Linear Regression (4.2.1.1) was tested (using a 10-fold cross-validation approach, to avoid overfitting) and ranked by their respective test scores.

Support Vector Regression Machine and Neural Network hyperparameters were optimized using a 10-fold cross-validation implemented with the help of the *RandomizedSearchCV* class, from the *scikit-learn* library.

In contrast to the time-consuming *Grid Search* approach, which tests all possible hyperparameter configurations given a set of parameter values, the *Randomized Search* approach only tests a fixed number of n different settings (which are created through random sampling - with replacement - of the given parameter values).

The metric used to score and rank each fold is the negative MSE - provided by the scikit-learn library as *neg_mean_squared_error*. However, to facilitate visualization and interpretation of results, these were converted to RMSE.

As explained in 4.2.1.1 and 4.2.1.2, in both Linear Regression and Support Vector Regression each Power Vector component is predicted independently - i.e. one predictor per component. As a result of that, the best predictors are - at the end of the model selection phase - grouped together, in order to predict the best overall refraction.

The parameters values provided to the Randomized Search algorithm during the optimization phase and its results are detailed in the subsections below.

Lastly, overall performance will be compared between the three approaches (Linear Regression, SVR, and Neural Network), so as to determine the best predictive model for the given dataset.

6.2.1 Linear Regression

In this subsection, we investigate the results produced by the different Linear Regression methods presented in 4.2.1.1.

Table 6.1 presents the average test scores of each method and vectorial component (produced through 10-fold cross validation), while highlighting the best performances. Figures 6.1, 6.2, and 6.3 detail the scores presented by the table, plotting the results produced for each fold in the cross validation.

The best models (for each power vector component) are then grouped toghether so as to make possible the computation of an overall performance score (MOD RMSE).

Table 6.1: Cross-Validation test scores for Linear Regression methods (per component)

Power Vector Component	*Method*	*Mean RMSE*	*Std RMSE*
M (Spherical-Equivalent)	Ordinary Least Squares	0.373670	0.016799
	Ridge Regression	0.373671	0.016777
	Lasso Regression	**0.373539**	0.017009
	Lasso Regression (LARS)*	0.373596	0.017035
	Bayesian Ridge Regression	0.373670	0.016792
J0 (Vertical Jackson-Cross Cylinder)	Ordinary Least Squares	0.146388	0.021926
	Ridge Regression	0.146391	0.021902
	Lasso Regression	**0.146275**	0.021973
	Lasso Regression (LARS)*	0.146318	0.021962
	Bayesian Ridge Regression	0.146389	0.021918
J45 (Oblique Jackson-Cross Cylinder)	Ordinary Least Squares	0.104175	0.011797
	Ridge Regression	0.104179	0.011738
	Lasso Regression	0.104145	0.011633
	Lasso Regression (LARS)*	**0.104093**	0.011610
	Bayesian Ridge Regression	0.104175	0.011786

Lower RMSE scores are better.

*An alternative implementation provided by the *scikit-learn* that uses Least-Angle Regression proposed in Efron et al. (2004).

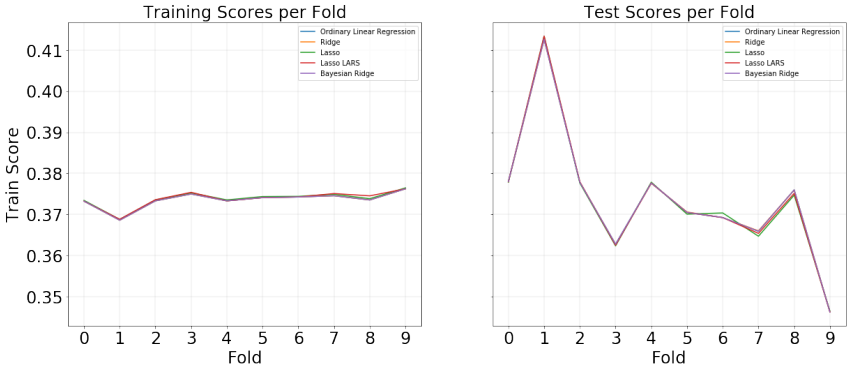
6.2.2 Support Vector Regression

In this subsection, we investigate the results produced by the Support Vector Regression method presented in 4.2.1.2.

Table 6.2 establishes the search space for hyperparameter optimization, while Table 6.3 presents the best hyperparameter configuration for the Support Vector Regression Machine, found through randomized search (with 10-fold cross validation).

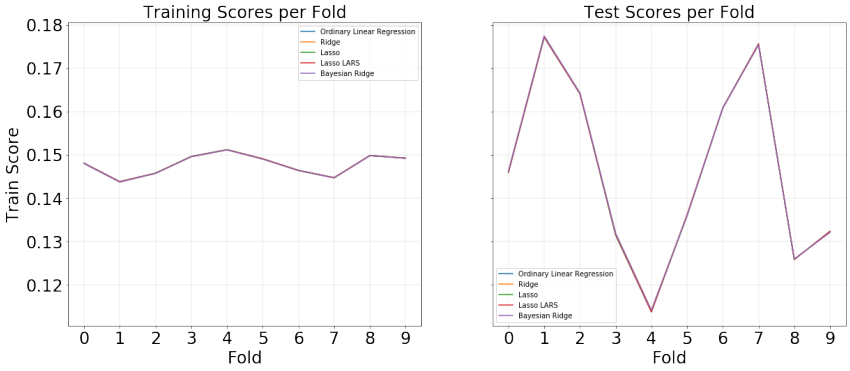
Table 6.4 presents the average test scores of all three power vector components

Figure 6.1: Linear Regression cross validation scores - Spherical Equivalent (M)



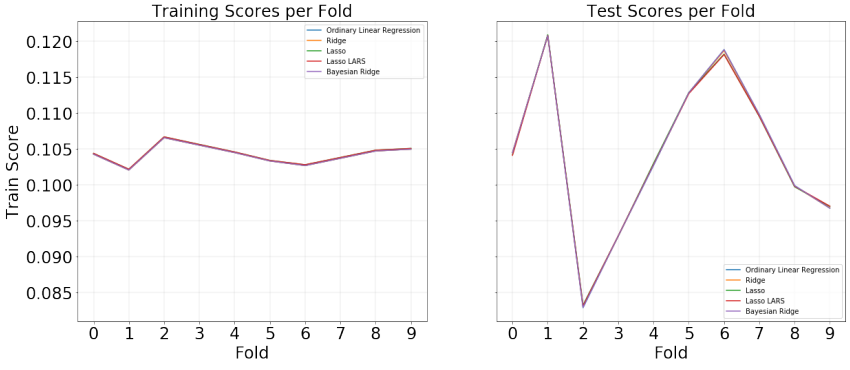
Training (left) and Test (right) scores.
From: Author

Figure 6.2: Linear Regression cross validation scores - J0 (Vertical Jackson-Cross Cylinder)



Training (left) and Test (right) scores.
From: Author

Figure 6.3: Linear Regression cross validation - J45 (Oblique Jackson-Cross Cylinder)



Training (left) and Test (right) scores.
From: Author

Figure 6.4: Linear Regression best model coefficients (per component)

	m	j0	j45
age_at_report_seconds	0.0445988	-0.0421928	0
sex_f	0.0547783	0.00802063	0
sex_m	-0	-0	0
race_amarela	-0	-0	0
race_branca	0.0370401	0	0
race_indigena	-0	0	0.00760746
race_parda	0	-0.010598	0
race_preta	-0.0279608	0.00684586	0
acuity_sc	0.236259	-0	-0.0155186
acuity_ph_sc	-0	-0.00287071	0
acuity_cc	0	-0	0
autorefraction_m	7.37274	0.0336707	0
autorefraction_j0	0.00450891	2.90586	-0.0365927
autorefraction_j45	0.0525117	-0	2.73818

Predictive model coefficients, showing selected features (non-zero values) and objective refraction importance to subjective refraction prediction. Notice also visual acuity sc's (without correctors) importance for predicting spherical-equivalent.

From: Author

(produced by a SVR model configured with the best hyperparameters, and using a 10-fold cross validation), while highlighting the best performances. Figures 6.5, 6.6, and 6.7 detail the scores presented by the table, plotting the results produced for each fold in the cross validation.

The best models (for each power vector component) are then groupped toghether so as to make possible the computation of an overall performance score (MOD RMSE).

Table 6.2: SVR hyperparameters optimized through Randomized Search (with cross validation).

Hyperparameter	*Value(s)*
Penalty parameter of error term (C)	[1;5]
Kernel	Linear, Polynomial, RBF*, Sigmoid
Degree of the polynomial kernel function**	[1;10]
Use of the Shrinking heuristic***	True, False

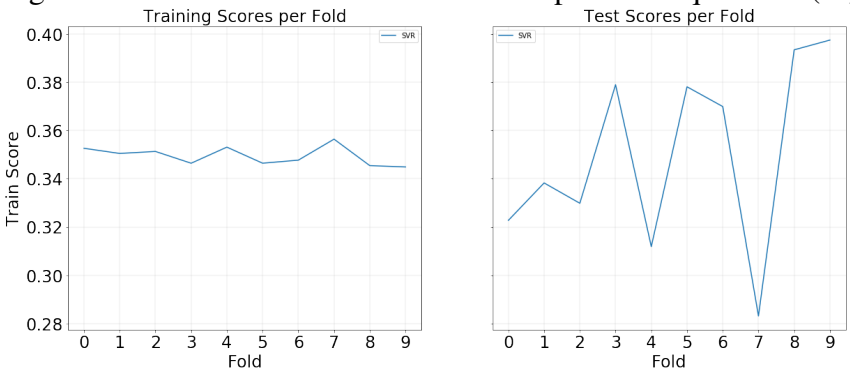
*Radial Basis Function. **Ignored by all other kernels. ***(CHANG; LIN, 2011)

6.2.3 Multilayer Perceptron

In this subsection, we investigate the results produced by the Multilayer Perceptron method presented in 4.2.1.3.

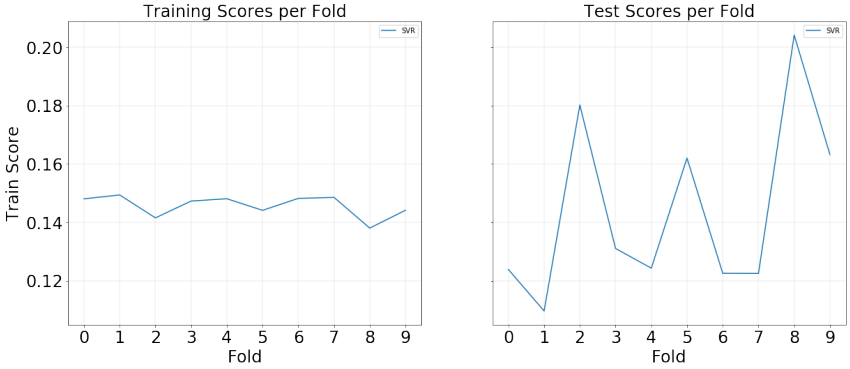
The search space for hyperparameter optimization is established in Table 6.5. The

Figure 6.5: SVR cross validation scores - Spherical Equivalent (M)



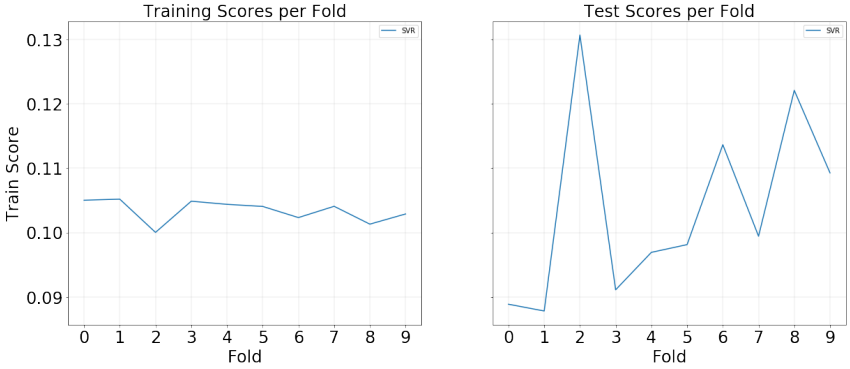
Training (left) and Test (right) scores.
From: Author

Figure 6.6: SVR cross validation scores - J0 (Vertical Jackson-Cross Cylinder)



Training (left) and Test (right) scores.
From: Author

Figure 6.7: SVR cross validation scores - J45 (Oblique Jackson-Cross Cylinder)



Training (left) and Test (right) scores.
From: Author

Table 6.3: Best Hyperparameters for SVR model.

Power Vector Component	*Hyperparameter*	*Value*
M	Penalty parameter of error term (C)	4
	Kernel	RBF
	Degree of the polynomial kernel function	8
	Use of the Shrinking heuristic	False
J0	Penalty parameter of error term (C)	4
	Kernel	RBF
	Degree of the polynomial kernel function	8
	Use of the Shrinking heuristic	False
J45	Penalty parameter of error term (C)	3
	Kernel	RBF
	Degree of the polynomial kernel function	3
	Use of the Shrinking heuristic	True

Optimization performed through Randomized Search (with built-in cross validation). Total of 30 iterations.

Table 6.4: Best SVR model - Cross-Validation test scores (per component)

Power Vector Component	*Mean RMSE*	*Std RMSE*
M (Spherical-Equivalent)	0.350399	0.038521
J0 (Vertical Jackson-Cross Cylinder)	0.144324	0.031066
J45 (Oblique Jackson-Cross Cylinder)	0.103799	0.014601

Lower RMSE scores are better.

best hyperparameter configuration for the Multilayer Perceptron, found through randomized search (with 10-fold cross validation) is then shown in Table 6.6.

Figure 6.8 details the cross validation scores (fold-by-fold) produced by the five best hyperparameter configurations.

Different from Linear Regression and SVR based models, the Multilayer Perceptron is able to predict multiple target attributes at once, by having multiple neurons in the last layer (also called *output layer*). Being so, there is no need to train and combine multiple models in order to compute an overall score (MOD RMSE).

Table 6.5: Multilayer Perceptron hyperparameters optimized through Randomized Search (with cross validation).

Hyperparameter	*Value(s)*
Hidden layers	[1;4]
Neurons per hidden layer	[10*;400]
Optimization heuristic**	SGD***, RMSProp, Adam
Epochs	[16;128]

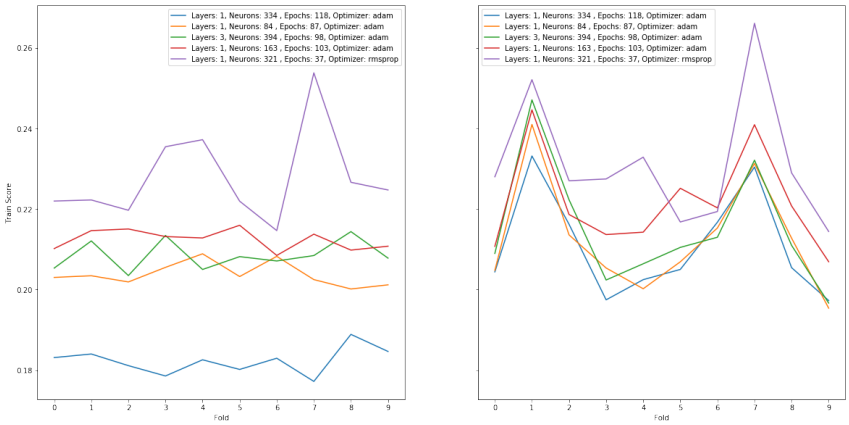
*Encoder behavior. **(Ruder, 2016). ***Stochastic Gradient Descent.

Table 6.6: Best Hyperparameters for Multilayer Perceptron model.

Hyperparameter	*Value*
Hidden layers	1
Neurons per hidden layer	334
Optimization heuristic	Adam
Epochs	118

Optimization performed through Randomized Search (with built-in cross validation).
Total of 30 iterations.

Figure 6.8: Multilayer Perceptron randomized search cross validation scores (top-5 hyperparameter configurations)



Training (left) and Test (right) scores.
From: Author

6.3 Comparison Between Techniques

The performance scores produced by the predictive models are displayed below in tables 6.7 (for each power vector component) and 6.8 (overall performance). The following figures 6.9 and 6.10 are shown so as to facilitate the visualization of the results.

6.4 Best Model Results

In this section, the model with best overall performance is further explored, in order to evaluate the applicatibilty of artificial intelligence algorithms and answer the questions raised in 1.0.1.

According to performance results shown in Table 6.8 (and Fig. 6.10), the best overall performance is produced by the Multilayer Perceptron model.

Figure 6.11 illustrates the model's learning curves, demonstrating no underfitting (when models perform well on training set but poorly on test set) nor overfitting (displayed when the test set performance improves only until reaching a certain point, where then

Table 6.7: Performance of predictive models (per component).

Method	*Component*	*RMSE*
Worst Case Scenario*	M (Spherical-Equivalent)	1.119009
	J0 (Vertical Jackson-Cross Cylinder)	0.608390
	J45 (Oblique Jackson-Cross Cylinder)	0.535598
Linear Regression	M (Spherical-Equivalent)	0.413058
	J0 (Vertical Jackson-Cross Cylinder)	**0.147472**
	J45 (Oblique Jackson-Cross Cylinder)	0.111887
SVR	M (Spherical-Equivalent)	0.393489
	J0 (Vertical Jackson-Cross Cylinder)	**0.146996**
	J45 (Oblique Jackson-Cross Cylinder)	**0.111782**
Multilayer Perceptron	**M (Spherical-Equivalent)**	**0.360249**
	J0 (Vertical Jackson-Cross Cylinder)	0.149369
	J45 (Oblique Jackson-Cross Cylinder)	0.116096

*Difference between objective refraction from autorefractor and expected subjective refraction (more details in 6.1).

Lower RMSE scores are better.

Table 6.8: Performance of predictive models (MOD RMSE)

Model	*RMSE*
Worst Case Scenario*	1.381732
Linear Regression	0.452641
Support Vector Regression	0.434669
Multilayer Perceptron	**0.406901**

*Difference between objective refraction from autorefractor and expected subjective refraction (more details in 6.1).

Lower RMSE scores are better.

pivots and starts to degrade).

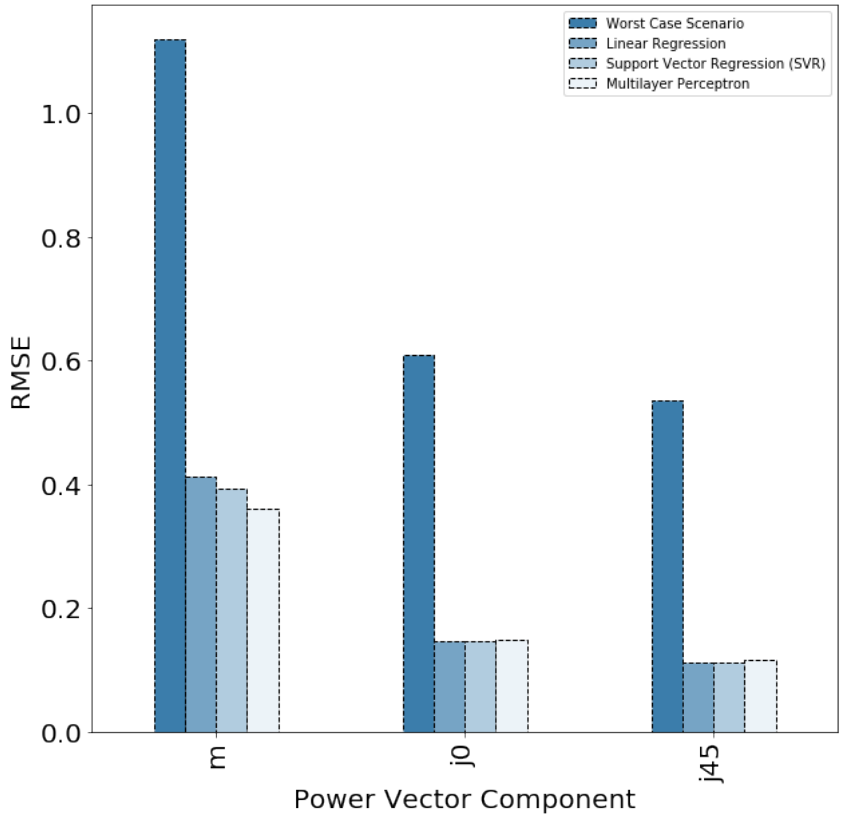
Figures 6.12, 6.13, and 6.14 provide insights on the accuracy of the predictions. The model exhibits bias when predicting subjective refractions for the M component. This can be observed by looking at the predicted values under the dotted lines, showing that the model has difficulty to properly predict higher refractive powers (i.e. it *undercorrects* the initial estimate). J_0 and J_{J45} components, however, do not share such bias.

Now for the variance, J_0 and J_{J45} components show a much higher spread between predictions when $outcome(j_i) = 0$. In other words, the model has difficulty to predict whether a patient has vertical or oblique astigmatic error or not.

In order to determine whether the best model provides an acceptable performance, two questions should be answered: (i) does the model perform better than the autorefractor? and (ii) is the error of the model within a clinically acceptable range?

(i) Does the best model perform better than the autorefractor? As explained in 6.1, the autorefractor's intrinsic error can be assessed by calculating its distance to the subjective refraction point within the dioptric space. This distance can be calculated either

Figure 6.9: Performance comparison between methods (per Power Vector component).



Lower RMSE scores are better.
From: Author

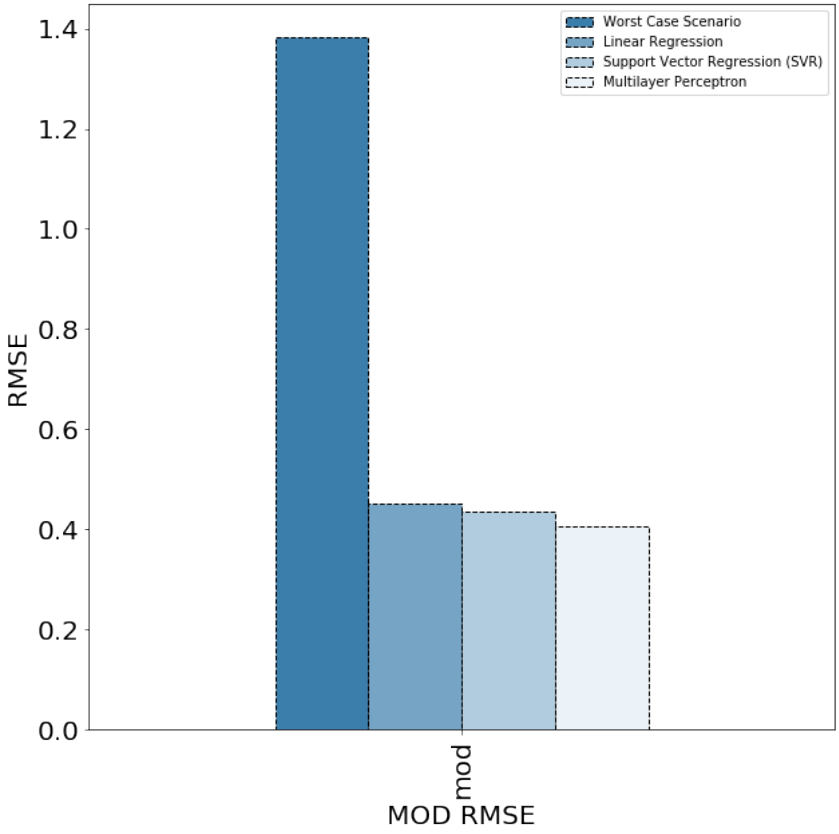
individually for each component of the Power Vector or as a single value, by computing the euclidean distance between the two points (also called Magnitude of Differences by Miller (2009)).

Both Table 6.8 and Figure 6.10 shows that **the predictive model do outperforms the autorefractor**, by scoring an RMSE of 0.406901 (in constrast to 1.381732, from the autorefractor), and reducing the error in approximately 70%.

(ii) Is the error of the best model within the clinically acceptable range? According to the literature, the maximum acceptable deviation for the Spherical Equivalent (S) component lies generally within a range of 0.2 to 0.6 diopters. (ROSENFIELD; CHIU, 1995; SMITH, 2006; GREIN; SCHMIDT; RITSCHE, 2014) Being so, any model performance with less than 0.6 diopters of error should be deemed as acceptable.

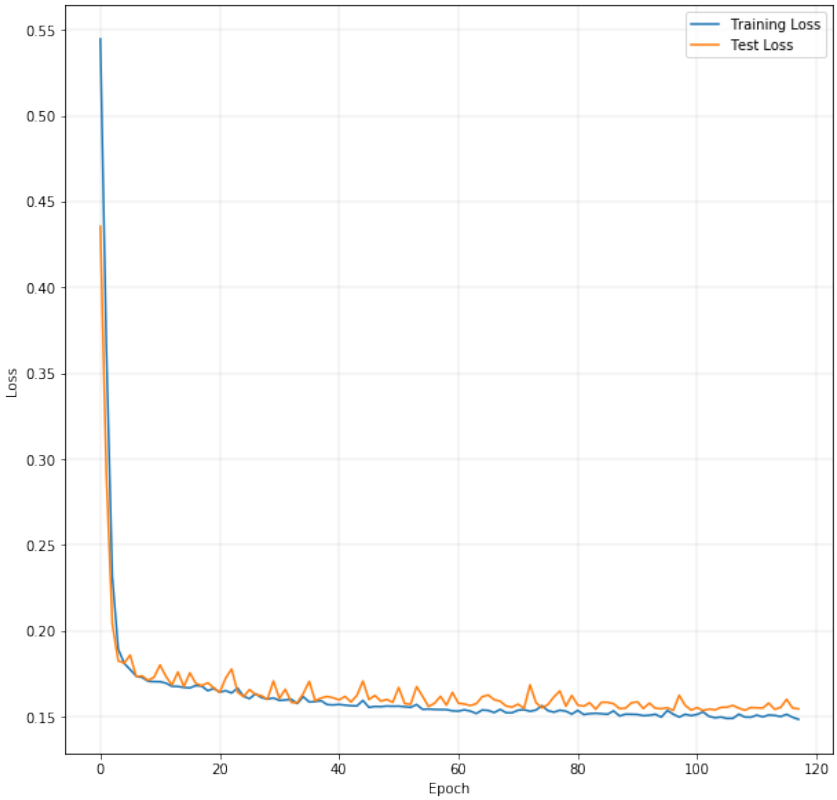
As Figure 6.15 shows, the average deviation for the Spherical Equivalent (M) is lower than the 0.6 diopter limit proposed in the literature (0.240229), and **can indeed be considerer clinically acceptable**.

Figure 6.10: Performance comparison between methods (MOD RMSE).



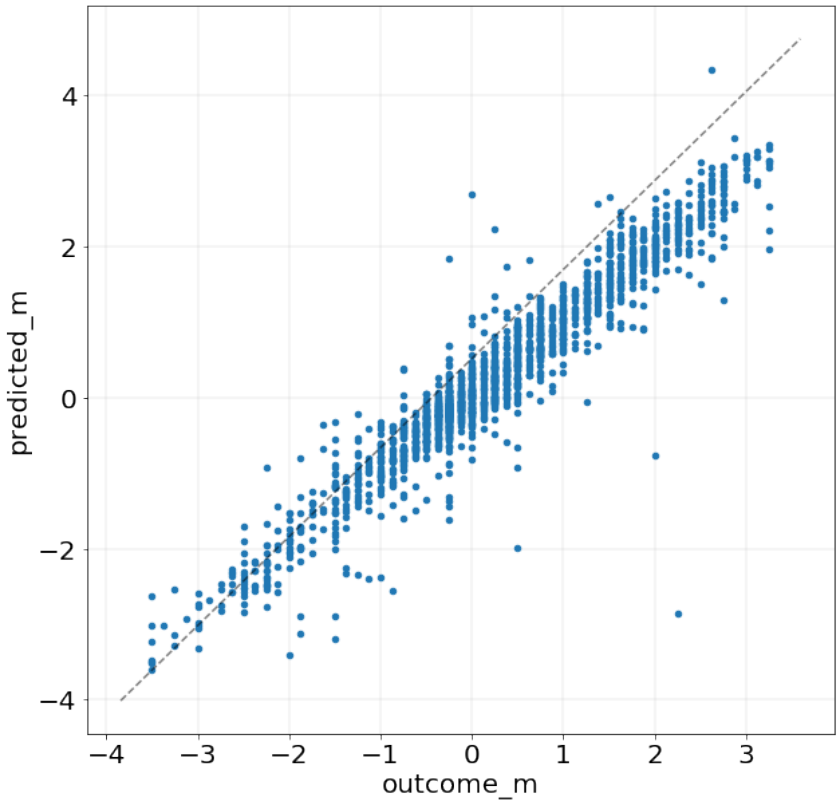
Lower RMSE scores are better.
From: Author

Figure 6.11: Loss history of Neural Network model (using best hyperparameter configuration)



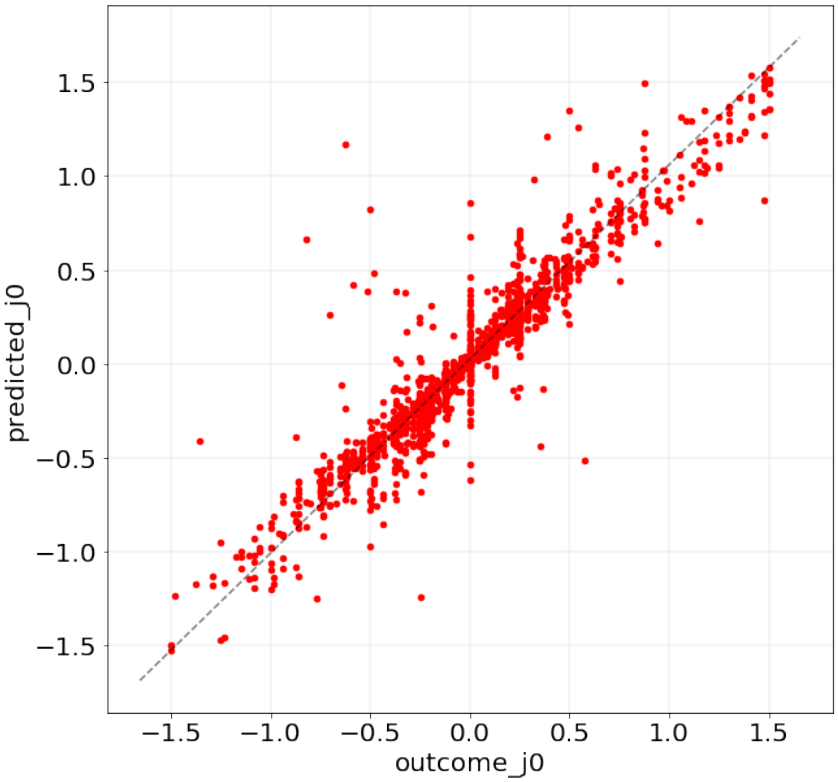
Training (blue) and Test (orange) scores.
From: Author

Figure 6.12: Predicted Refraction vs Expected Refraction - M



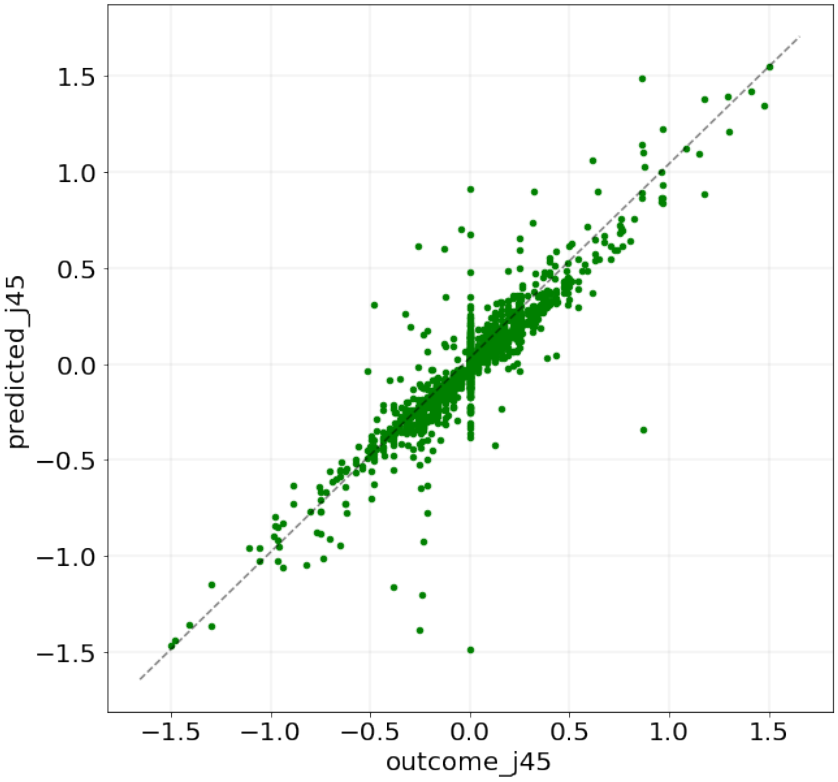
From: Author

Figure 6.13: Predicted Refraction vs Expected Refraction - J0



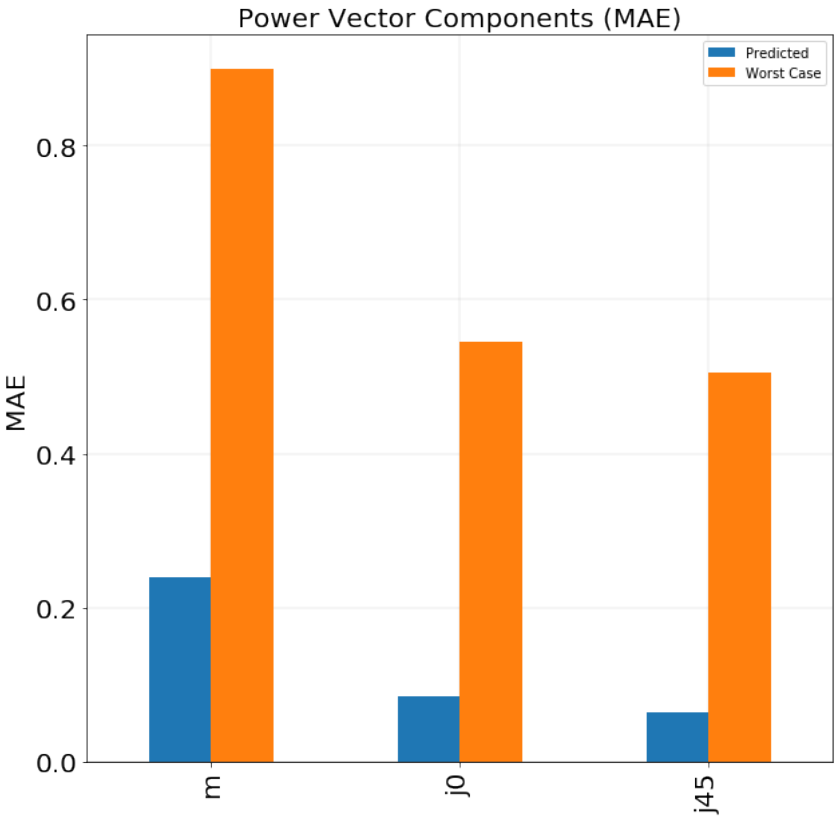
From: Author

Figure 6.14: Predicted Refraction vs Expected Refraction - J45



From: Author

Figure 6.15: Multilayer Perceptron Performance (MAE)



From: Author

7 CONCLUSION

In this work, several machine learning methods were explored, in order to predict subjective refraction - the refraction used when prescribing spectacles. This investigation was performed so as to determine whether ophthalmologists and optometrists would benefit from the use of AI in their workplace, by having predictive models to improving the quality of refractive error diagnosis and reducing the time needed to prescribe corrective lenses.

After the experiment, the best overall performance was produced by a class of Neural Network called *Multilayer Perceptron*.

The application of the technique proved to be beneficial, reducing the difference between objective refraction (measured by the autorefractor) and subjective refraction (used for prescribing spectacles). This improvement in prediction accuracy has the effect of reducing the time required for the diagnosis of refractive errors and the prescription of corrective lenses, causing a decrease in costs and allowing for more patients to be treated within the same time span.

As shown in the results chapter, the multilayer perceptron model was able to reduce the overall prediction error produced by the autorefractor down to 30 % - and approximately 32 % for the spherical equivalent error (shown in Fig. 5.9 to be the most inaccurate of all three components) - showing a significant improvement over the results produced through current practices.

It has also been shown that the performance produced by the best predictive model lies within a clinically acceptable range, and is suitable for practical use by ophthalmologists and optometrists.

7.1 Limitations of the Model

Although presenting a good performance, by reducing the original prediction errors up to approximatelly 70 %, the proposed approach bears limitations that may hinder the reproduction of the results presented. This section discusses some of these limitations.

As discussed in Section 5.3, the collected data had to be preprocessed, due to unstructured variables, that lead to the presence of invalid and missing patient data. During this stage, some assumptions have been made, and may or may not influence the quality of the final performance of the model:

Instances with low visual acuity (e.g. hand motion - HM) were converted to the LogMAR scale, using values found in Schulze-Bonsel et al. (2006). However, the work of Schulze-Bonsel et al. (2006) assumes that VA measurements are taken at a fixed distance of 50 cm. Therefore, if future work decides to use visual acuity measurements in predictive models, care should be taken during the conversion of these values, by taking into account the distance used during the measurement - thus avoiding distorted LogMAR values being used in the process of learning.

To avoid possible interference of the outliers in the final result, these went through a process of winsorization - i.e. saturation of extreme values based on quantiles. This practice was adopted so that instances with incorrect values of objective refraction do not influence the performance of the model too much. Therefore, the presence of extreme values in future work may negatively impact the training of models.

It should also be noted, that because of the winsorization process, the model trained in this work may present some instability - i.e. not generalize well - for instances with very high (but correct) optical power.

Regarding the reproducibility of results, another important point should be highlighted:

Due to the impossibility of differentiating the many types of astigmatism: simple, compound, and mixed (the clinical notation "combines" the optical powers of the axes into a single value - the cylindrical component), it was not possible to characterize the distribution of the anomalies in the data collected by the TelesaúdeRS-UFRGS group - i.e. how many instances were used during the training of the model, for each type of anomaly. Being so, some anomalies may not be well represented within the dataset, causing the trained model to be somewhat instable when predicting these specific cases.

Lastly, the selection of subjective refraction between static and dynamic refractions (described in Section 5.3) may present an obstacle and cause selection bias. To avoid this problem, it is strongly advised to have experienced clinicians performing such task.

7.2 Future Work

This work performed an investigation of three different methods of machine learning: (i) Linear Regression; (ii) SVR; and a simple neural network architecture, (iii) Multilayer Perceptron.

However, due to lack of time and computational resources, it was not possible to explore other models.

Therefore, it is suggested an investigation of the applicability of other models, such as nearest neighbors and decision trees - or even more robust (and generally better) alternatives like random forests, as well as their performance, and how these compare against the models presented here. Additionally, the use of different metrics for performance evaluation (e.g. MAE) is very much welcomed.

It is also suggested the introduction of new features in order to reduce the bias shown to be present in the M component predictions, or the use of a larger dataset, in order to improve the performance of predictive models by reducing the inconsistencies caused by variance. An investigation on the use of techniques such as boosting is also recommended, so as to identify and reduce possible model bias (such as the one shown in Fig. 6.12).

A deeper analysis of the prediction errors is also suggested for future works, in order to investigate their causes and possible improvements to the models.

Another topic to be explored is whether it is possible, using the Matrix representation (described in Subsec. 2.3.3), to produce satisfactory (i.e. clinically acceptable) results - such as the results presented in this paper. And if so, a discussion on issues such as the interpretability of the results produced and their application in thin and thick systems is highly encouraged.

As an additional and final suggestion, an extension of the work proposed by Varadarajan et al. (2018) is proposed, with the addition of features such as sex, gender, race, or even visual acuity measurements to the training variables (as it is done in the current work). A more accurate application of the power vector is also incentived, through the inclusion of axial data, which would then allow for the computation of both J_0 and $J45$ components, improving prediction of refractive errors in astigmatic eyes.

REFERENCES

AAO. **Innovative Technologies in Diagnosing Eye Diseases**. 2014. Available from Internet: <http://www.aao.org/clinical-statement/innovative-technologies-in-diagnosing-eye-diseases>.

Anaheim Eye. **Astigmatism**. 2017. Available from Internet: <http://www.anaheimeyemd.com/astigmatism-anaheim/>.

BAILEY, I.; LOVIE, J. E. New design principles for visual acuity letter charts. v. 53, p. 740–5, 12 1976.

BARROS, E. V. d.; DIAS, V. G. Incidência das ametropias no Hospital Universitário em Campo Grande (MS) entre 1996 e 1998. **Arquivos Brasileiros de Oftalmologia**, scielo, v. 63, p. 203 – 208, 06 2000. ISSN 0004-2749. Available from Internet: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492000000300006&nrm=iso>.

Benhenda, M. ChemGAN challenge for drug discovery: can AI reproduce natural chemical diversity? **ArXiv e-prints**, aug. 2017.

BISHOP, C. M. **Pattern Recognition and Machine Learning (Information Science and Statistics)**. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN 0387310738.

BOWLING, B. **Kanski's Clinical Ophthalmology, International Edition: A Systematic Approach**. Elsevier Health Sciences, 2015. ISBN 9780702055744. Available from Internet: <https://books.google.com.br/books?id=D9GfBwAAQBAJ>.

Canadian Association of Optometrists. **What is astigmatism?** 2017. Available from Internet: <https://opto.ca/health-library/astigmatism>.

CHANG, C.-C.; LIN, C.-J. LIBSVM: A library for support vector machines. **ACM Transactions on Intelligent Systems and Technology**, v. 2, p. 27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

CORTES, C.; VAPNIK, V. Support-vector networks. **Mach. Learn.**, Kluwer Academic Publishers, Hingham, MA, USA, v. 20, n. 3, p. 273–297, sep. 1995. ISSN 0885-6125. Available from Internet: <https://doi.org/10.1023/A:1022627411411>.

DEVELOPERS, S. learn. **SVM**. 2007. Available from Internet: <http://scikit-learn.org/stable/modules/svm.html#regression>.

DRUCKER, H. et al. Support vector regression machines. In: MOZER, M. C.; JORDAN, M. I.; PETSCHE, T. (Ed.). **Advances in Neural Information Processing Systems 9**. MIT Press, 1997. p. 155–161. Available from Internet: <http://papers.nips.cc/paper/1238-support-vector-regression-machines.pdf>.

Efron, B. et al. Least Angle Regression. **ArXiv Mathematics e-prints**, jun. 2004.

ESTEVA, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, v. 542, n. 7639, p. 115–118, 02 2017.

FAGEERI, S. O. et al. Eye refractive error classification using machine learning techniques. p. 1–6, Jan 2017.

FERRAZ, F. H. et al. Refractive errors in a brazilian population: age and sex distribution. **Ophthalmic and Physiological Optics**, v. 35, n. 1, p. 19–27, 2015.

FLOM, M. C.; WEYMOUTH, F. W.; KAHNEMAN, D. Visual resolution and contour interaction∗. **J. Opt. Soc. Am.**, OSA, v. 53, n. 9, p. 1026–1032, Sep 1963. Available from Internet: <http://www.osapublishing.org/abstract.cfm?URI=josa-53-9-1026>.

GARCIA, C. A. d. A. et al. Prevalence of refractive errors in students in Northeastern Brazil. **Arquivos Brasileiros de Oftalmologia**, scielo, v. 68, p. 321 – 325, 06 2005. ISSN 0004-2749. Available from Internet: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27492005000300009&nrm=iso>.

GARTNER, W. F. ASTIGMATISM AND OPTOMETRIC VECTORS. **Am J Optom Arch Am Acad Optom**, v. 42, p. 459–463, Aug 1965.

GIGER, M. L. Machine Learning in Medical Imaging. **J Am Coll Radiol**, v. 15, n. 3 Pt B, p. 512–520, Mar 2018.

GLOROT, X.; BORDES, A.; BENGIO, Y. Deep sparse rectifier neural networks. In: GORDON, G.; DUNSON, D.; DUDíK, M. (Ed.). **Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics**. Fort Lauderdale, FL, USA: PMLR, 2011. (Proceedings of Machine Learning Research, v. 15), p. 315–323. Available from Internet: <http://proceedings.mlr.press/v15/glorot11a.html>.

GREIN, H.-J.; SCHMIDT, O.; RITSCHE, A. Zur reproduzierbarkeit von refraktionsbestimmungen. **Der Ophthalmologe**, v. 111, n. 11, p. 1057–1064, Nov 2014. ISSN 1433-0423. Available from Internet: <https://doi.org/10.1007/s00347-014-3064-6>.

GULSHAN, V. et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. **JAMA**, v. 316, n. 22, p. 2402–2410, 12 2016.

HARRIS, W. F. Dioptric power: its nature and its representation in three- and four-dimensional space. **Optom Vis Sci**, v. 74, n. 6, p. 349–366, Jun 1997.

HARRIS, W. F. Power vectors versus power matrices, and the mathematical nature of dioptric power. **Optom Vis Sci**, v. 84, n. 11, p. 1060–1063, Nov 2007.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. New York, NY, USA: Springer New York Inc., 2001. (Springer Series in Statistics).

International Myopia Prevention Association.

ISO Central Secretary. **Ophthalmic optics – Visual acuity testing – Standard and clinical optotypes and their presentation**. [S.l.], 2017.

JASON. **Artificial Intelligence for Health and Health Care**. 2017. Available from Internet: <https://www.healthit.gov/sites/default/files/jsr-17-task-002_aiforhealthandhealthcare12122017.pdf>.

JORGE, J. et al. Retinoscopy/autorefraction: which is the best starting point for a noncycloplegic refraction? **Optom Vis Sci**, v. 82, n. 1, p. 64–68, Jan 2005.

KEATING, M. P. A system matrix for astigmatic optical systems: I. introduction and dioptric power relations. v. 58, p. 810–9, 11 1981.

KEIRL, A.; CHRISTIE, C. **Clinical Optics and Refraction: A Guide for Optometrists, Contact Lens Opticians and Dispensing Opticians**. Baillière Tindall Elsevier/Butterworth-Heinemann, 2007. ISBN 9780750688895. Available from Internet: <https://books.google.com.br/books?id=-9ftET_IDkYC>.

KHURANA, A. **Ophthalmology**. New Age Books, 2003. ISBN 9788122414714. Available from Internet: <https://books.google.com.br/books?id=tRzi3sYBInIC>.

KINGE, B.; MIDELFART, A.; JACOBSEN, G. Refractive errors among young adults and university students in Norway. **Acta Ophthalmologica Scandinavica**, v. 76, n. 6, p. 692–695, 1998.

LANGE, C. et al. Resolving the clinical acuity categories "hand motion" and "counting fingers" using the Freiburg Visual Acuity Test (FrACT). **Graefes Arch. Clin. Exp. Ophthalmol.**, v. 247, n. 1, p. 137–142, Jan 2009.

LEVENSON, A. K. J. H. **Clinical Methods: The History, Physical, and Laboratory Examinations.** third. [S.l.: s.n.], 1990.

LIBRALAO, G. L. et al. Machine learning techniques for ocular errors analysis. p. 569–578, Sept 2004. ISSN 1551-2541.

MAVRACANAS, T. A. et al. Prevalence of myopia in a sample of Greek students. **Acta Ophthalmologica Scandinavica**, v. 78, n. 6, p. 656–659, 2001.

MAYER, H. et al. A system for robotic heart surgery that learns to tie knots using recurrent neural networks. **Advanced Robotics**, Taylor Francis, v. 22, n. 13-14, p. 1521–1537, 2008. Available from Internet: <https://doi.org/10.1163/156855308X360604>.

MILLER, J. M. Clinical applications of power vectors. **Optom Vis Sci**, v. 86, n. 6, p. 599–602, Jun 2009.

MOOR, J. H. The dartmouth college artificial intelligence conference: The next fifty years. **AI Magazine**, v. 27, p. 87–91, 2006.

National Eye Institute. **Facts About Hyperopia**. 2016. Available from Internet: <https://nei.nih.gov/health/errors/hyperopia>.

National Eye Institute. **Facts About Myopia**. 2017. Available from Internet: <https://nei.nih.gov/health/errors/myopia>.

National Keratoconus Foundation. **How Does The Human Eye Work?** 2018. Available from Internet: <https://www.nkcf.org/about-keratoconus/how-the-human-eye-works/>.

PASCOLINI, D.; MARIOTTI, S. P. Global estimates of visual impairment: 2010. **Br J Ophthalmol**, v. 96, n. 5, p. 614–618, May 2012.

PROBST, L.; TSAI, J.; GOODMAN, G. **Ophthalmology: Clinical and Surgical Principles**. Slack, 2012. ISBN 9781556427350. Available from Internet: <https://books.google.com.br/books?id=QilKd2nvl8cC>.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. **Psychological Review**, p. 65–386, 1958.

ROSENFIELD, M.; CHIU, N. N. Repeatability of subjective and objective refraction. **Optom Vis Sci**, v. 72, n. 8, p. 577–579, Aug 1995.

Ruder, S. An overview of gradient descent optimization algorithms. **ArXiv e-prints**, sep. 2016.

S, V. et al. Prevalence of refractive error in the united states, 1999-2004. **Archives of Ophthalmology**, v. 126, n. 8, p. 1111–1119, 2008. Available from Internet: <+http://dx.doi.org/10.1001/archopht.126.8.1111>.

SCHULZE-BONSEL, K. et al. Visual acuities “hand motion” and “counting fingers” can be quantified with the freiburg visual acuity test. **Investigative Ophthalmology Visual Science**, v. 47, n. 3, p. 1236, 2006. Available from Internet: <+http://dx.doi.org/10.1167/iovs.05-0981>.

SMITH, G. Refraction and visual acuity measurements: what are their measurement uncertainties? **Clin Exp Optom**, v. 89, n. 2, p. 66–72, Mar 2006.

SNELLEN, H. **Probebuchstaben zur Bestimmung der Sehschärfe**. Van de Weijer, 1862. Available from Internet: <https://books.google.com.br/books?id=LEYJAAAAIAAJ>.

THIBOS, L. N.; HORNER, D. Power vector analysis of the optical outcome of refractive surgery. **J Cataract Refract Surg**, v. 27, n. 1, p. 80–85, Jan 2001.

THIBOS, L. N.; WHEELER, W.; HORNER, D. Power vectors: an application of Fourier analysis to the description and statistical analysis of refractive error. **Optom Vis Sci**, v. 74, n. 6, p. 367–375, Jun 1997.

VARADARAJAN, A. V. et al. Deep learning for predicting refractive error from retinal fundus images. **Investigative Ophthalmology Visual Science**, v. 59, n. 7, p. 2861, 2018. Available from Internet: <+http://dx.doi.org/10.1167/iovs.18-23887>.

VOS, T. et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. **Lancet**, v. 390, n. 10100, p. 1211–1259, Sep 2017.

WIKIPEDIA. **Snellen Chart**. 2008. Available from Internet: <https://en.wikipedia.org/wiki/Snellen_chart>.