S	
E1	
E2	
For Office Use Only	

Masters Project Final Report

(MCS)

2019

Project Title	A computational model using artificial neural networking for predicting astigmatism following corneal surgery
Student Name	Vibhavi Yugapamoda Attigala
Registration No. **& Index No.**	2016/mcs/011 16440114
Supervisor's Name	Mr G.P Senevirathna

For Office Use ONLY

A computational model using artificial neural networking for predicting astigmatism following corneal surgery

A dissertation submitted for the Degree of Master of Computer Science

V.Y Attigala

University of Colombo School of Computing

2020

Declaration

The thesis is my original work and has not been submitted previously for a degree at this or any other university/institute.

To the best of my knowledge it does not contain any material published or written by another person, except as acknowledged in the text.

Student Name: V Y Attigala

Registration Number: 2016/mcs/011

Index Number: 16440114

21/06/2020

---------- ----------

Signature: Date:

This is to certify that this thesis is based on the work of

Mr. V Y Attigala

under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Certified by:

Supervisor Name: Mr G.P Senevirathne

Abstract

Corneal ecstatic disorder is a corneal thinning condition and Keratoconus is a specific form of ectasia when the cornea thins and starts to bulge. Corneal cross linking is the only available accepted treatment for prevent or decrease the progressive Keratoconus condition. Although this treatment named as a successful surgery, numerically it has a 7.6% failure rate and 2.9% complication rate according to the post-operative statistics. However, there is no prediction mode available to measure or capture the post-operative results before the surgery occurs. Accordingly, this research is framed by the pre-operative and post-operative variables in order to develop a computation model for future predictions before the surgery occurs. The cornea reassembles a watch glass, which is transparent, tissue. The cornea covers a one fifth of the eye globe, which is the main refractive element and contribute to two third of the refractive power for the eye There are several refractive errors related to the eye, such as Myopia, Hyperopia, Presbyopia and astigmatism. The research is focus about the astigmatism as the refractive error. Astigmatism is a type of a refractive error caused by the irregularity in the shape of a person's cornea. In this condition, the eye fails to focus the light equally on the retina leading to blurred or distorted vision causing either myopia or hyperopia. The astigmatism can be categorized into two main sections. Regular astigmatism and the Irregular astigmatism. Most of the astigmatism is regular and can be corrected with the spectacles. Irregular astigmatism is the worst scenario resulted from scarring of the cornea following injury, due to eye surgery or corneal ectasia like keratoconus.

The keratoconus is a progressive, noninflammatory, bilateral, asymmetric disease, characterized by paraxial stromal thinning irregular protrusion. One of the acceptable ways to prevent progression of the keratoconus is Corneal collagen crosslinking, which strengthens corneal tissue to halt progression of the eye's surface in keratoconus by removing the outer layer of the cornea (called the epithelium) is removed to allow entry of riboflavin, a type of B vitamin, into the cornea, which then is activated with UV light.

According to the past literature, outcome of the corneal collagen crosslinking displaying a nonlinear behavior. The researcher used a novel idea of predicting the factors associated with collagen crosslinking for the outcome of pre-operative and post-operative astigmatism with the aid of an artificial neural network model.

The researchers were able to discover the age and the gender as a related factor to the astigmatism changes of the post-operative outcome of the surgery. The researchers were also able to apply radial basis functions to suit neural network model. The dataset used to train the neural network is needed to be increased and improved to suit a better fitting of the model.the covid-19 outbreak has restricted the researchers from collecting and finetuning the dataset from the hospitals. The researchers have to use alternative source from the public domain to improve the error rate. The premilitary data visualization techniques and statistic showed that Cornea Cross link might Fail after year of the surgery for a age group of 15 to 19 years and after 30 years of age. More parameter testing is recommended for the future researchers to improve the outcome of the selected domain.

Acknowledgement

My heartfelt gratitude extends to Mr. Gihan Seneviratne, my research supervisor and mentor, for his continuous direction and support, in aligning me towards preparing this research study. His comprehensive knowledge, constant guidance and valued constructive criticism has assisted me in better understanding the core foundations of the study and aided me with compiling the research to the best of my ability.

A special thanks extends to Prof Maduwanthi Dissanayake for supervising her medical expertise.

A special thanks extends Dr. Kusum Ratnayake. and his medical staff, for spending their valuable time in providing their genuine thoughts and opinions on the research.

Also, I would like to thank Ms Nirosha Karunarathne and my colleagues who help me in numerous ways to complete this project successfully.

Lastly, I'd like to thank my family for shedding their moral support and encouragement in completing the research.

Table of Contents

List of Figures

CHAPTER ONE
INTRODUCTION

1.1 Introduction to the Problem Statement

The cornea is the outer surface covering of the eye. Protecting the inner content, maintaining the shape and archiving light surface of the eye are essentials for complete the vision. Eye cornea does not have any blood vessels like other bodily tissues. Tears and aqueous liquid provide the nutrients to the cornea. The cornea is composed of six layers as illustrated in figure 1.

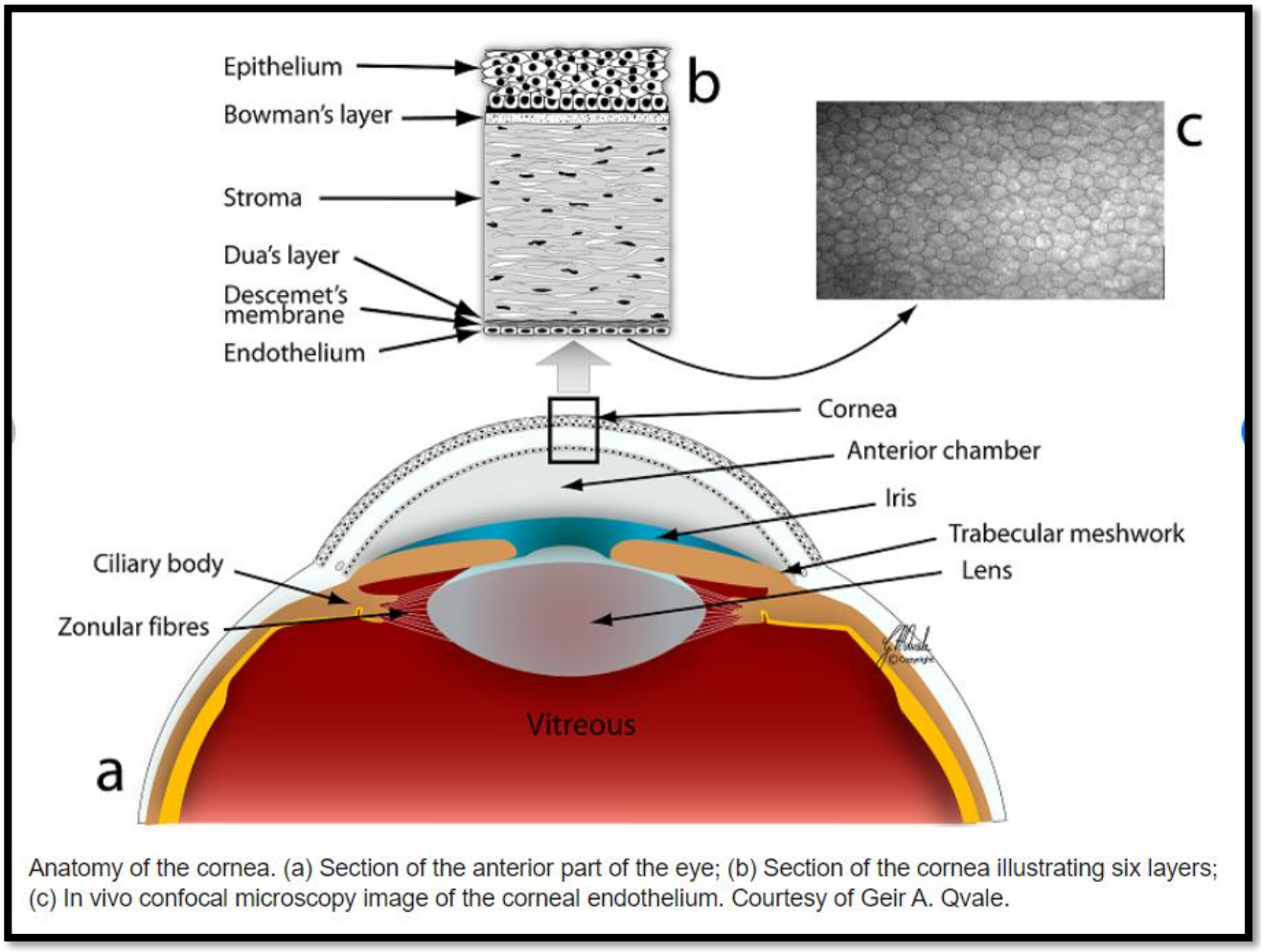


Figure 1 - Layers of the eye cornea
Source: Navaratnam, Jesintha & Utheim, Tor & Rajasekhar, Vinagolu & Shahdadfar, Aboulghassem. (2015). Substrates for Expansion of Corneal Endothelial Cells towards Bioengineering of Human Corneal Endothelium. Journal of functional biomaterials.

The first layer is Endothelium which absorbs the oxygen from tears and passes to the rest of the eye. This layer act as covering layer for the cornea. Keratoconus is the condition where the progressive thinning, deformation and scarring of the cornea. As per the literature, Keratoconus emerged due to genetic and some environmental factors [12] However, most of the patients who with allergies and eye rubbing frequently affect the Keratoconus conditions.

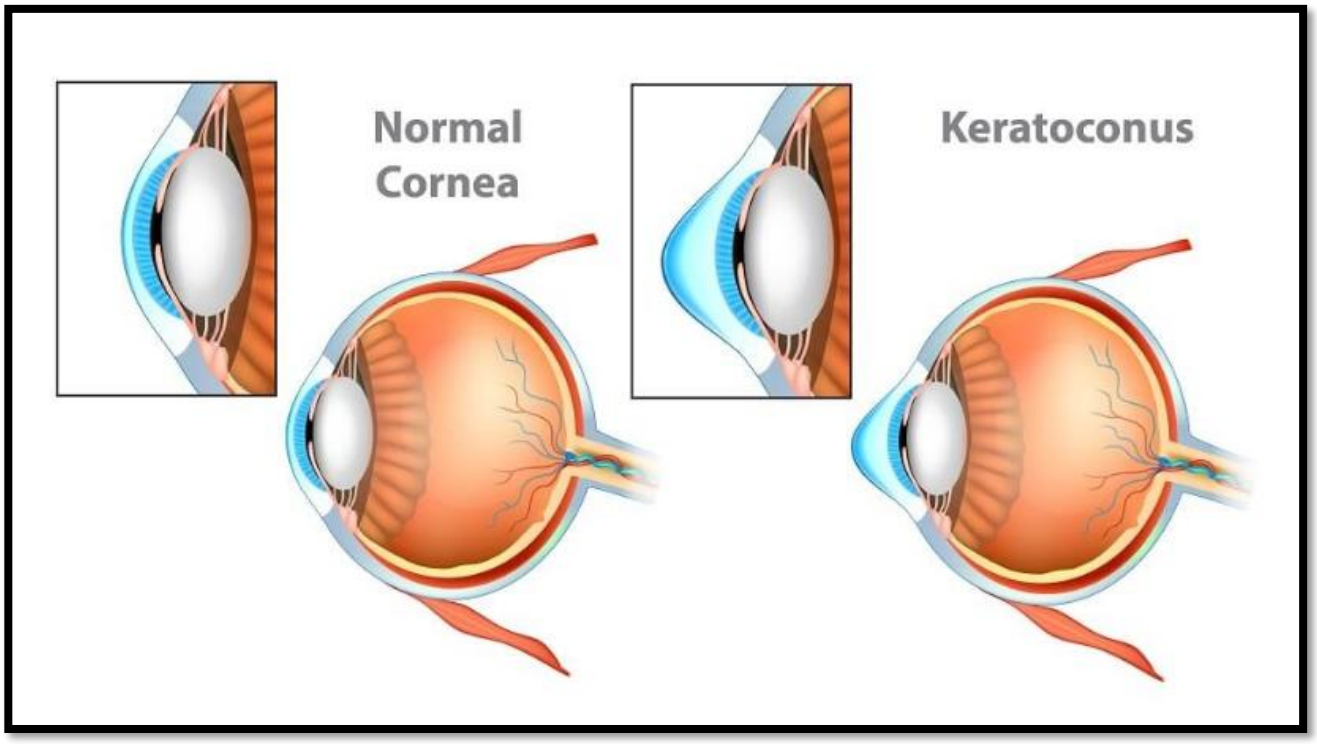


Figure 2 - Normal cornea Vs Keratoconus cornea
Source: https://www.oclvision.com/conditions/keratoconus/

Recently most of the researchers [7],[8],[9] attend to the studies and researchers which related to this disease as it continually growing. Over the years, several treatments proposed to correct the Keratoconus condition. But corneal cross linking is the only accepted treatment to prevent or decrease the Keratoconus condition. Although this treatment named as the accepted treatment, numerically it has 7.6% failure rate and 2.9% complication rate based on patients` pre- and post-operative results. But still there is no specific mode to predict the post-operative results before the corneal cross-linking surgery occurs. So this research is to develop a computation mode to predict the post-operative results before the surgery occurred.

1.2 Motivation

Now a day's computational predictions are used for most industries for their operations. But still there are some conditions where to be developed and implement prediction mode for more accurate results. The subjected eye disease is continuously growing disease and still researchers

failed to come up with prediction mode with more accurate details. The main motivation for this study is to identify the correlation between the factors which affect the Keratoconus disease and predict the post-operative results before the surgery occurred. The researcher also affected the Keratoconus condition from his childhood and still he is not in the perfect condition with his vision. Though the researcher already gone through the corneal cross-linking surgery he could not achieve the perfect vision of one eye. As per the professors and eye surgeons, corneal cross-linking surgery is succeeded only for one eye. So, the main motivation for the study is to identify the correlation between the pre- and post-operative factors and predict the post-operative results before it occurred. Then the patient can decide on the other available short-term treatments such as contact lenses and spectacles.

1.3 Objectives of the study

- To identify pre-operative and post-operative variables associated with astigmatism.

Currently most of the consultants record each patients pre and post-operative conditions, but unable to correlate each variable for astigmatism predictions. Therefore, in this research, researcher going to analyses each variable to find the correlation with the astigmatism. In current practice the data set or the experience of each surgeon is more limited to specific time or area. But since the study is based from a significant data set , researcher identifies the most relevant and effective factors and the correlation with the astigmatism level.

- To design a computational model using Artificial Neural Networking which can predict the post-operative astigmatism.

Once identifying the correlation between pre-operative factors with the post-operative results, researcher objects to design computation model using appropriate artificial neural networking algorithms. When considering the available techniques for the data analysis ANN is the most suitable method for data analysis when the data set is a significant amount.

- To implement an End user application which can transfer the knowledge of the neural network to the end users.

A web-based application will be developed which facilitate patients and doctors to predict the post-operative results according to the pre-operative factors. Since this model gave an upright prediction, the module needs to develop and implement as end user application where even the patient get a self-prediction over his/her ectasia condition.

- To identify the controlling factors for Keratoconus ectasia

Since Keratoconus is a specific eye ectasia with limited treatments. This research is effort to identifying the most effective factors which positively affect the astigmatism level and based on the identified factors

researcher can open opportunity for future researchers to implement a controlling module for Keratoconus disease before it occurred.

1.4 Scope of the study

The main scope of the study is to develop a computation model to predict astigmatism. The computation model is developed through the artificial neural network algorithms based on the machine learning techniques. The training data set designed through the periodical data which received from the eye consultants about their patient clinical history.The designed web-based solution will provide access to consultant to insert patients preoperative condition in order to get the post-operative astigmatism values changing over the period of regular time interval predicted by the Artificial neural networking model. Accordingly, researcher aims to predict the error rate of cross-linking treatment using the machine learning and artificial neural networking algorithms. This study provides a correlation between the level of astigmatism and other factors. When identifying the relevant factors, researcher can build up a space for future researchers to implement a controlling or preventive action for Keratoconus illness for future generation.

1.5 Structure of The Thesis

The thesis contains four chapters describing structure of each chapter as follows.

- Chapter One-describes the thesis motivation of the problem and the formal details related to the decertation, the scope of the project and the objective is discussed under the chapter one. The purpose of the chapter one is to guide the reader about the thesis roadmap.

- Chapter Two-describes the critical analysis of the existing literatures.it reviews the existing datasets and researchers conducted on keratoconus with astigmatism ,the neural networking models and statistical methods, the literature review helps the researcher to identify the variables used by the similar research projects.it helps the researcher as a guide to build the hypothesis.

- Chapter Three-describes the method methodology of the research.it start by describing the research hypothesis and step by step approaches with the description of the theoretical framework when developing the neural networking model. It describes the process of knowledge discovery. The statistical steps taken into the consideration is described under this section. The parameters and algorithms used to discover the patterns are justified with the reasons in the section,

- Chapter Four-describes the results the researcher achieved by applying the neural networking model. It represents the graphical representation of the visualized dataset. This chapter describes how the dataset has been controlled by the neural networking algorithm by mapping the inputs parameters to the output parameters.

- Chapter Five-describe the conclusion and the future works, the summery of the work that the researcher has conducted is described under this chapter. This chapter reviews whether the research objectives and aims are met. And what are the works to be done for future researchers by following the dissertation.

CHAPTER TWO
Literature Review

2.1 Introduction

In this chapter researcher outline the study by comparing the literature data which studied the same related concepts and theories. This chapter build the foundation for the research study and discussed the drawbacks and improving areas based on the given literature.

2.2 Factors Affects the Problem Statement

According to [1], corneal cross-linking is one of the most attained treatment for progressive corneal ectasia. Further, [1] states this treatment can stabilize the progression of corneal ectasia while delaying the requirement of corneal surgeries. Although this treatment named as a successful treatment for corneal ectasia [2] mentioned there is a 7.6% of failure rate and 2.9% of complication rate with the patient post-operative results. Further [2] mentioned the reasons for the failure and complication rates as interaction of factors mechanisms behind the treatment. Further [3] evidenced current studies of corneal collagen cross-linking is insufficient/ to determine the postoperative results.

Number of researches [7][8][2][9][10] studies categorized the factors which affect to the corneal cross-linking treatment into four aspects as follows.

1. Refractive Variables [7] [8]
2. Bio-mechanical Variables [2]
3. Surgical Variables [7][8]
4. Temporal Time Duration Variables [9][10]

[4] has discussed on the modern technical solutions to which use to diagnose on clinical decisions. In the discussion [4] states that machine learning is the optimum solution for clinical diagnoses which disrupted by higher input of factors and their correlations. Further [4] mentioned the main barrier of technological diagnosis is overloaded information which affect to the final clinical diagnosis.

Artificial Neural Networking (ANN)

Artificial Neural Networks are mathematical processes stimulated in certain aspects of bio-logical neurons through the acquired knowledge of learning process and stored data which called synaptic Weight [5]. [5] evidenced ANN (Artificial Neural Networks) are the optimum mathematical algorithm which inspired in certain aspects of biological neurons. Further [6] mentioned ANN has been widely using for medical applications in last two decades due to its learning and predicting relationship ability. Further [6] proved corneal surgery predictions didn't follow any linear behavior and ANN is the optimum non-linear analyzing tool which obtain best results. According [6] the error which occurred in ANN application was 0.98D and further he mentioned ANN method is more reliable since it provided the lowest error rate.

[6] mentioned the multi-layer perception (MLP) is the most widely used ANN in many studies and situation analysis. Further he mentioned MLP has several non-linear elements which arranged in layers. The first and last layers called as input and output layers and rest layers named as hidden layers.

The closest application build using ANN for the keratoconus-based prediction is [6]. According to the ANN model they have designed [6], the prediction of post-operative astigmatism is 0.93 diopter of error for the INTACTS surgery. Further they have approached the problem with 288 keratoconus patients' eyes preoperative factors feed to the ANN using multilayer perceptron, implemented in MATLAB. One of the limitations the author of the proposed research see is, they have only considered surgical and refractive variables in their model with different corneal surgery than the surgery the proposed research authors are going to address. They haven considered the biomechanical aspect of the cornea into the model.

When finding preoperative and post-operative variables for the input of the neural networking model

The past research papers conducted can be categorized into following sections

- Epithelium-on vs Epithelium-off surgeries factors. [10]
- Factors affecting to the cross-linking surgeries.[11]

Further Statistical models also being developed. [12]

The issue with the [10][11][12] is that they address the factors in the statistical method which address the problem in a linear behavior.[6] suggest that the usage of ANN in Corneal surgery can be used to solve the non-linearity of the problem.

The biomechanical aspect of corneal cross linking is vital when determining the success rate of the corneal cross-linking surgery.[2] proposed an algorithm to predict the corneal stiffening effects' have not predicted the CXl efficacy in human patients but used data in porcine, murine, and lapine

corneas. apart from [2] model designed a biomechanical corneal model using surgical factors.[2] also stated the demarcation line after CXL which a biomechanical perspective of the cornea is. they also support the idea of to date, the effectiveness of treatment could be monitored only indirectly by postoperative follow-up corneal topographies or using corneal confocal microscopy.by conducting the proposed research using ANN the researchers can open new possibilities using ANN rather than traditional follow-up corneal topographies after the surgery.

2.3 Possible Solution and Limitations according to the past Researches

According to [6] ANN provides a predictive model for surgeons, when the classical data is insufficient on the patients who have progressive keratoconus. Further he suggests a user-friendly prediction module using ANN to predict aid ophthalmologists ok K1 and astigmatism for KC corneas. Surgical variables (Incision and depth) and ring characteristics should input to find out the optimum position which give the highest level of vision. However, this research bounds to a limited number of considered variables where there are more reasons which improve the KC but not considered to this module. [6] ends up with an opportunity to develop a new module by considering more factors which affect the K1. Further 1 proved that ANN is adequate technique to predict pre- and post-operative keratoconus levels.

However, [12] implement a mathematical model using multiple regression model to astigmatism prediction. In that, he mentioned number of incisions and incision length is the surgical factor which effect for the astigmatism level after the required surgeries. But still he proved that age and gender also have a strong affective power over the astigmatism level. Further he explained the implemented mathematical model cannot predict the astigmatism as the result can be change due to the age and gender.

Limitations

- Bio mechanical factors can be changed according to the region or the country. But for this research, researcher select the sample as only three hospitals which located in Sri Lanka.
- Since this research only focus on one specific treatment, the subjected treatment can be outdated due to the technical development.
- This module limits to one specific eye disease.
- Since this is only based on the past records of patients`, actual result can be different.

Conclusion.

There exist researches [3],[4][5][6],[7] which shows several factors affects the astigmatism of keratoconus and their conclusion was to test the dataset with the more geographical data since the problem behavior nonlinear. The artificial neural networking models are the novel idea given by the past researchers [1] [2] shows promising results in corneal surgeries .The identified research gap is there is no neural network model designed to predict the astigmatism in the corneal cross linking surgery before doing the surgery.

This will make the researcher move towards developing the neural network that can connect the doctor and patients using the past data and the machine learning algorithms to predict the outcome of the surgery.

CHAPTER THREE
Methodology

3.1 INTRODUCTION

The methodology is the path that the research addressed the subjected research problem with the gained knowledge of Literature review. As above mentioned, the scope of the research is to implement a computational model to predict the astigmatism level. This chapter provides an overview of the implementation process with detailed analysis which have been carried out through the project to make the project successful. This chapter is a detailed described conversion process of raw data base into a computation model.

3.2 Research Strategy

According to the literature researcher identified four main factors which have higher effect for the pre and post operational astigmatism level.

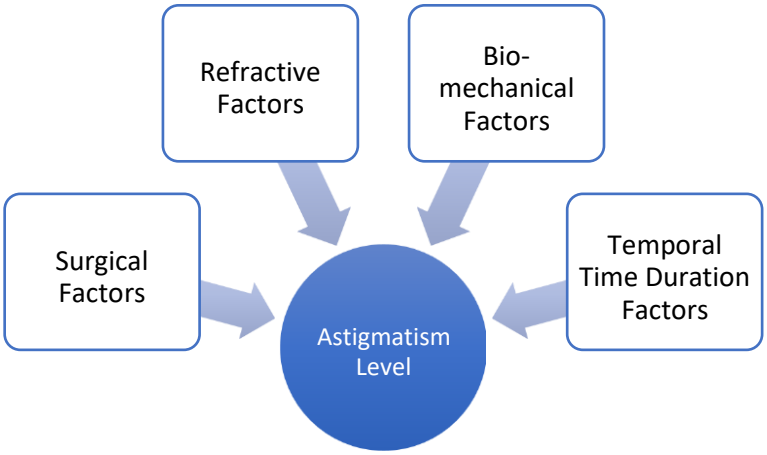


Figure 3 Factors Associated to Astigmatism

By nature, this study can be grouped as quantitative as more quantitative data and areas to be focus on. Hence the required data is collected through the past records of patient diagnosis reports. Sample for this study is the patients who visits below three hospitals for cross linking. during 01.01.2017 to 31.12.2019 period, among the target population of Keratoconus patients.

- National Eye Hospital of Sri Lanka.

Above hospital have the highest number of Keratoconus patient details within the Sri Lanka. That was the main root cause to the sample selection. Annexure 1 is the data sheet which use for data collection.

Due to the issue of low amount of dataset presence

The dataset that was collected from the National Eye hospitals Clinic Contains 84 records of patient's details varied from pre-operative to the period of one year.

The provided data does not contain the critical data that will expose the privacy of the patients and only the patients' medical history will be stored. The researcher aims to track the patient by only using a reference number associated with each case. If there are new anomalies detected by the prediction model the research can use the reference number to follow up the patients.

3.3 Conceptual Framework

The researcher needs to conduct a pilot research about the nature of the data set at national eye hospitals Record Room. The following variable types are expected to collect from the national eye hospitals Record Room, to conduct the pilot research.

1.Refractive Factors

Variables collected by refractive eye examination related to keratoconus will be collected before and after conducting cross linking surgeries by examining the past literature related to factors affecting astigmatism will be considered when choosing variables to make the input layer of the neural networking model.

Following variables has been used from the corneal clinic dataset.

- Age.
- Sex.
- Eye.
- Keratometry Values (K1, K2).
    - flat keratometry-K1
    - steep keratometry k2
- Cylinder.
- Best corrected vision acuity.
- Pachymetry values-Related to the corneal thickness
- Spherical Equivalent.

2.Bio-mechanical Factors

The biomechanical factors from the past literature will be collected. Factors related to corneal cross-linking surgeries and the treatment protocol will be collected.

3. Surgical Factors

Data set recorded during the cross-linking procedure will be recorded by the research team.

Because of the covid-19 outbreak the access to the surgical and biomechanical Factors could not be obtained because of the quarantine restrictions. The researcher had to use the initial dataset given by the cornea clinic of the eye hospital to conduct the research in the limited amount of time and limited amount of data quantity.

Regression Analysis with the Deep Neural Networks.

The objective of the research is to predict the astigmatism value follows the surgery. Which maps inputs variables into output variables to determine a Based on the previous researches, it is suitable to use multivariable regression models to develop the deep neural network model.

Implementation stages in the Neural Network Design (Knowledge discovery process).

The supervised learning approach is used in the proposed research. The neural network which is implemented in the research uses a learning a function that maps an input to an output based on example input-output pairs. It infers a function from labeled training data consisting of a set of training examples.

The research is step by step divided in to the seven steps as mentioned in the below figure, in each step the raw data set obtained from the cornea clinic will be analyzed by the researcher to extract the meaningful contents to develop the model. Under the preparation stage the data set can be transformed into one of the following data categories.

- Categorical Data
- Nominal Data

The processing steps of above two categories is discuss under the analysis of the neural network implementation.

Creation of the neural network model is the next step that research follows.in this section the researcher maps inputs layers and hidden layers along with the output layer to create the unknown relationship between the dataset. After learning throw the iterations and the learning rates to map the objective in the research. Once the layers neurons and the learning rate is identified the research can compile and fit the dataset with the testing datasets.

Once the training has been done the research. The researcher can use evaluation approach to evaluate the model to investigate the anomalies.

The evaluation techniques used in the research is as of follows.

1. Predicted astigmatism values against real values(scatterplots).
2. Confusion matrix for false and true positive predictions.

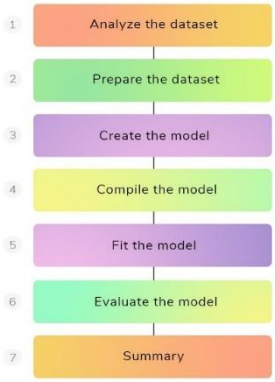


Figure 4 The Development Process of ANN

3.3 Development of artificial Neural network

The cases indicated by the cornea consultant of the national eye hospitals which has keratoconus with cornea cross linking is approximately 500 cases. According to the morgens table of sample size estimation there needs to be at least 300 cases to conduct a proper Artificial neural network model design. But in order to improve the accuracy of the model the researcher request from the ethics clearance commit to have a data set of at least 300 cases or more.

Due to the situation of covid-19 outbreak in Sri Lanka, the researcher couldn't obtain the desired dataset quantity. From the cornea clinic of the eye hospital in Sri Lake the researcher was able to collect the dataset of 26 patients with 84 cases for pre- and post-operative refractive factors.

TensorFlow with keras will be used to develop the ANN model. The main reason for choosing TensorFlow is because be an open source library and flexible low-level function calls with high community support.

Structure of the Raw Data Set

L

			PRE OP							1 MONTH							
AGE	SEX	EYE	BSCVA	SPHERE	CYLINDER	SE	K 1	K 2	PACHYMETRY	1 UCVA	1 BSCVA	1 SPHERE	1 CYLINDER	1 SE	1 K1	1 K2	1 PACHYME
24	F	OS	6/9			3.25			465		6/12			3.5			461
21	F	OS	6/12		-5.25	4.25	52	57	476		6/12		-4.25	2.5	53.25	58	452
15	M	OD	6/12		-5.25	1.5	48.5	53.75	505		6/9		-4.75	4	48.5	53.25	490
27	M	OD	6/18		-2.5	1.5	46.4	48.5	473		6/9		-2.5	1.25	47.25	49.15	435
27	F	OS	6/24		-3.25	-6	45.75	49.75	440		6/24		-4.25	-2.37	47.25	51.5	404
25	F	OS	6/9		-2.5	2	52.25	54.75	504		6/60		-3.75	3.5	53	56.75	416
17		OS	6/36		-5.54	3.37	45.37	50.91	459		6/12		-6.5	3	45.5	51.75	455
17	M	OS	6/9			3.5			508		6/9		-5.25	2.67	43.75	49	492
28	F	OS	6/18		-6	4.5	53	59	434		6/18		-5.5	5	54.75	60	455
22	M	OD	6/9			5.5			491		6/9			6			412
22	M	OS	6/12		-4.5	3.75	51	55.5	525		6/12		-4.75	3.75	52	56.75	493
17	M	OD	6/9		-2.75	4.09	46.75	49.75	483		6/9		-3	1.12	46.25	49.25	493
19	F	OS	6/36			0	43.06	43.64	542		6/18			2.75	45.29	49.96	543
22	M	OD	6/18		-6.5	2.75	44.75	50.25	497		6/18		-7	4.75	50.25	57.25	499
19	F	OD	6/12		-5.25	1.75	44.75	50	436		6/12		-6.25	1	45.25	51.5	432
16	M	OS	6/9		-7.5	2.75	48.75	59	483		6/9		-7.75	1.5	48.75	57.5	450
23	M	OD	6/9		-4	1.75	43	47	435		6/12		-5.25	1.75	41.5	47.75	513
18	F	OD	6/12		-5.75	2.67	45	51	456		6/9		-7.5	0.5	45	52.5	430
18	F	OS	6/12		-7.5	1.75	44.25	52	443		6/9		-8.75	1.25	44.5	53.25	451
24	M	OD	6/6		-6.75	3	46.75	53.5	447		6/6		-7.25	3.25	46.75	54	418
30	M	OD	6/9		-2.25	0.5	45.25	47.5	454		6/9		-2.5	0.5	45.75	48.25	464
30	M	OS	6/9		-3.5	2.5	48	51.5	433		6/9		-3.5	1.5	49	52.5	432
20	M	OD	6/15		-0.75	1.5	46.7	47.6	459		6/12		-3.5	2.5	47.25	50.75	457
20	M	OS	6/24		-5.5	1.75	48.2	48.7	458		6/12		-4.75	1	46.5	51.25	447
25	M	OD	6/9			0.5	44	46	545		6/6		-1.75	0.5	43.25	45	545
25	M	OS	6/24			2.5	50.47	54.8	466		6/9		-8.25	1.5	48.75	57	455
17	M	OS	6/9		-3.25	1.12	45.75	49	427		6/9		-3.8	1.62	46.5	50.3	401

Figure 5 Raw Dataset from National Eye Hospital

The above raw data set was a snapshot obtained from the national eye hospitals of 26 patients. The dataset contains several issues that needs to be address before feeds into the neural network model.

- Missing values

- Table is not normalized
- Categorical data cannot be applied to some neural network models
- Patients id and the name which can be sensitive to the patients

The preprocessing stage in the knowledge discovery process can address the mentioned issues.

Pre-processing

The initial step when developing the artificial neural network is to pre-process the data that will be feed to the neural network. pre-processing step involve handling missing dataset, check for categorical values and mainly the researcher must split the data set into testing data set and training data set. The final stage of the data pre-processing is Feature Scaling.

For the data pre-processing stages python panda's library is recommended based on the features and statistical capability to handle missing values.

Principal component analysis should be conducted to find the correlation between the features of the dataset.

Dealing with the Missing values.

Problem with the missing data is that it reduces the statistical power of a study Most statistical procedures automatically eliminate cases with missing data. This means that in the end, you may not have enough data to perform the analysis

The researcher observed following types of missing data in the raw dataset.

- Accidentally skip
    - Where Several items of the data cell are missing

CYLINDER	SE
	3.25
-5.25	4.25
-5.25	1.5
-2.5	1.5
-3.25	-6
-2.5	2
-5.54	3.37
	3.5
-6	4.5
	5.5
-4.5	3.75
-2.75	4.09

Figure 6 Missing Values in the Dataset

Some of the cylindrical values are missing from the patient's dataset. The researcher used to subsitite with the mean value can improve the accuracy of the dataset.

The following code segment was used to approximate with the mean values of the rows.

```
data['k1'].fillna((data['bcva'].mean()), inplace=True)
data['k2'].fillna((data['k2'].mean()), inplace=True)
data['pachymetry'].fillna((data['pachymetry'].mean()), inplace=True)
data['se'].fillna((data['se'].mean()), inplace=True)
data['cylinder'].fillna((data['cylinder'].mean()), inplace=True)
```

12 SE
3.5
2.75
2.25
5
3.25
0.87

Figure 7 Missing variables which can be calculated using mathematical equations

$$\textbf{Spherical Equivalent} = \frac{\boldsymbol{Cylinder}}{\mathbf{2}} + \textbf{Sphere}$$

Data Transformation.

The purpose of the data transformation stage is converting the data rows or column from one format into a desired format. The main reason for the data transformation is because of the existing format the data cannot be feed into the input layers of the neural networking model.the main example for the data transformation that the researcher as to do was that best corrected vision acuity was in the fraction format. The fraction format cannot be feed into the deep learning network. The chart that was used to record the best corrected visual acuity was in meter scale.by referring to the notation charts described by UK Civil Aviation Authority, Medical Department the researcher had to convert them into the decimal format as follows.

Distance Visual Acuity Conversion Chart

UK	Decimal	5m	US
6/3	2.0	5/3	20/10
6/4	1.5		20/13
6/5	1.2		20/17
6/6	1.0	5/5	20/20
6/9	0.7	5/7.5	20/30
6/12	0.5	5/10	20/40
6/18	0.3	5/15	20/60
6/24	0.25	5/20	20/80
6/36	0.2	5/30	20/120
6/60	0.1	5/50	20/200

Figure 8 Data Transformation

Building the neural network.

After the pre-process stage the researcher need to specify number of hidden layers and neurons in each layers and transfer function of each layers along with training function and the weight learning functions. After obtaining the data set and feature extraction only the researcher can indicate the exact configuration of the neural network.

Layers used python Keras Neural Network Models.

The following layers are available to use in the keras library to develop the neural networking model for the regression prediction.

• Input

• Dense

• Convolution1D and convolution2D

• Embedding

• LSTM

The input layers and the dense layers was used to develop the neural network since the other layers are effective when processing images related problems. The dense layer uses a matrix multiplication method. The following parameters has been used in the dense layer matrix multiplication which get updated by the backpropagation in the deep learning network training.

$u^T.W,\ W \in R^{nxm}$

according to the keras documentation the dense class contain following parametria needs to be implemented on the neural networking model.

- units: defines the number of neurons in the dense layer. This parameter must be an integer number which is not zero
- activation: defines an activation function which can be used. If activation=None the linear activation is used.
- use_bias: the researcher can use a bias vector by making this parameter to True. Bias can be

Initializers Parameters

The following parameters has been used to initialize the weights related matrix or the bias bias.

- kernel_initializer: Initializer for the kernel weights matrix.
- bias_initializer: Initializer for the bias vector.

Regularization techniques in the Neural Network model.

The dataset contains only 200 records due to the issue of not been able to collect the complete dataset. The effect of overfitting, which can be expected from the low quantity of data rows can be minimized with the help of a regularize.in keras the following can be set as a parameter to set the kernel level weight regulation and bias vector regulation. In this researcher the attempt to regulation is discuss in the evaluation section of the thesis.

- kernel_regularizer: Regularizer function applied to the kernel weights matrix.
- bias_regularizer: Regularizer function applied to the bias vector.
- activity_regularizer: Regularizer function applied to the output of the layer (its "activation").
- kernel_constraint: Constraint function applied to the kernel weights matrix.
- bias_constraint: Constraint function applied to the bias vector.

Training the neural network

The weights of the neural network are adjusted using an optimization algorithm to match inputs and outputs ideally. The commonly used Optimization algorithms are mentioned as follows.

- First Order Optimization Algorithms
- Second Order Optimization Algorithms

Selecting the ideal algorithm for the proposed research problem is based on the nature of the data set.

In keras the training can be done by splitting the dataset into two sections. Which is for training and testing.

In the research that can be done using the following code segment.

```
model.fit(train_X,train_y,validation_split=0.2,epochs=30,
callbacks=[early_stopping_monitor],shuffle=False)
```

in the above code segment, validation_split function guarantees that the dataset is divided into testing and training based on the same segment of the data for each epoch. In the above case there is 20 % from the dataset is separated for the validation and 80% for the training. TheW model has a callback function which can be triggered on a specific event. The callback function allow us to control the training process.in the research the callback has been used to minimize the effect of overfitting by controlling with a monitor parameter.The user can specifiy the monitor parameter by setting it as follows,

```
es = EarlyStopping(monitor='val_loss', mode='min')
```

by setting validation loss as the monitor parameter and the mode to minimizing the early stopping callback will trigger the event of stopping the training process when the mean square error is minimized.

Activating functions in the Neural Network.

The purpose of an activation function is to introduce nonlinearity to a complex problem in machine learning.

Most of the real-world problems are not following linear behaviors with the input variables. If the researchers didn’t use any activation function by specifying none in the dense layers as follows,

```
model.add(Dense(5470, activation=None))
```

the keras backend will interpret as linear activation and provides simple linear regression, the learning ability of using a linear activating function are really limited because of the linear behavior. Since the refractive measurements related to surgery are nonlinear, the main goal of the researcher was to use nonlinear activators in the neural network.by looking at the documentation of the keras the researcher was able to discover the following activation functions supported by the platform to be tested with.

- rectified linear unit activation function
- Sigmoid activation function
- SoftMax activation function
- Hyperbolic tangent activation function.

In sigmoid activation function small changes in x axis are casing higher increase or decrease in the y axis this causes the sigmoid activators are good at classifying. Hyperbolic tangent is tents to show a similar behavior as sigmoid but because of it derivative is steeper than the sigmoid activator it can get more values.this makes the learning rate of the hyperbolic tangent faster.

Rectified linear unit activation function doesn’t provide negative values. This is a good use for a estimator candidate. The researcher first used the Rectified linear unit activation function in the deep layers of the neural network because the aim of the research is to estimate the astigmatism value as a problem of regression.

Testing the neural network

Performance of the neural network can be measured by exposing unseen data or testing data set which is separated in the data pre-processing stage. The accuracy and the performance of the training data set can be measured using the following statistical factors.

- coefficient of determination
- mean bias error.
- root mean square

a measure that can estimate short term performance, and which can be described as the variance of predicated data around the measured data which was collected from the training dataset. The lower root means square error value determine the higher accuracy of the developed neural networking model.

In the testing stage the researcher used a mix method of testing, which can be described as follows.

1. Training with the dataset from the cornea clinic in sri lanka national eye hospital.
2. Testing with the dataset from the Melendy Dataset for kerataconus
3. Training with the dataset from the Melendy Dataset for kerataconus
4. Testing with the the dataset from the cornea clinic in sri lanka national eye hospital.

Results has been discussed under the results and discussion section.

Calculating how many hidden layers needed for the neural network.

The following formula was used to calculate the hidden layers of the neural network.

How many Hidden Layers

- $Nh = Ns/(\alpha * (Ni + No))$
- Ni = number of input neurons.
- No = number of output neurons.
- Ns = number of samples in training data set.
- α = an arbitrary scaling factor usually 2-10

Proposed Model/Design

According to the methodology of the research thesis, the researcher aims to develop a web-based system which can provide facility to doctors to enter the clinical details of patients before conducting the corneal cross-linking surgery. The surgeon can then get a prediction of astigmatism level by the artificial neural networking model. The researcher aims to provide periodic changes to the astigmatism level after the surgery.

The initial data set used to train the artificial neural network will no longer require storing after the learning has completed. Therefore, the system will not expose confidential details of the patients to the outside.

The system will be developed using TensorFlow library for the neural network model building and a MySQL database to store the patients records to prioritize the relationships between patients' medical records.

The frontend of the web application will be developed using angular and typescripts as the technologies and a NodeJS web api will be developed to communicate with the neural network prediction and the front-end web application.

The actors that uses the system can be a doctor or a optometrist who is authorized to store patients details into the system. The proposed system will be only accessible to the consultants practice hospital to keep the system from outside interference.

Web Service API

The web application developed for the use of the doctors will use the following code segment to get the result from the python script which has the training neural network.

```
var express = require('express');
var app = express();
var Astigmatism=0;
app.route('/getPredictions').get(function(req,res)
{
    res.send("Astigmatism After Value is "+Astigmatism);
    if(Astigmatism<0.6)
    {
        res.send("your Astigmatism is normal");
    }else if(Astigmatism>=0.6 && Astigmatism<=2.0)
    {
        res.send("you have small degree of Astigmatism ");
    }else if(Astigmatism>2.0 && Astigmatism<=4.0)
    {
        res.send("you have moderate degree of Astigmatism ");
    }else if(Astigmatism<4.0)
    {
        res.send("you have high degree of Astigmatism ");
    }

});

app.get('/',function(req,res){
    res.send('Training the Model');
    const { spawn } = require('child_process');
    const pyProg = spawn('python', ['./../predictastgmatism.py']);

    pyProg.stdout.on('data', function(data) {

        console.log(data.toString());
        Astigmatis=data;

    });
}));
```

Figure 9 Web Service to feed the dataset

Figure 10 Web Service

Web Application for Patients and Doctors

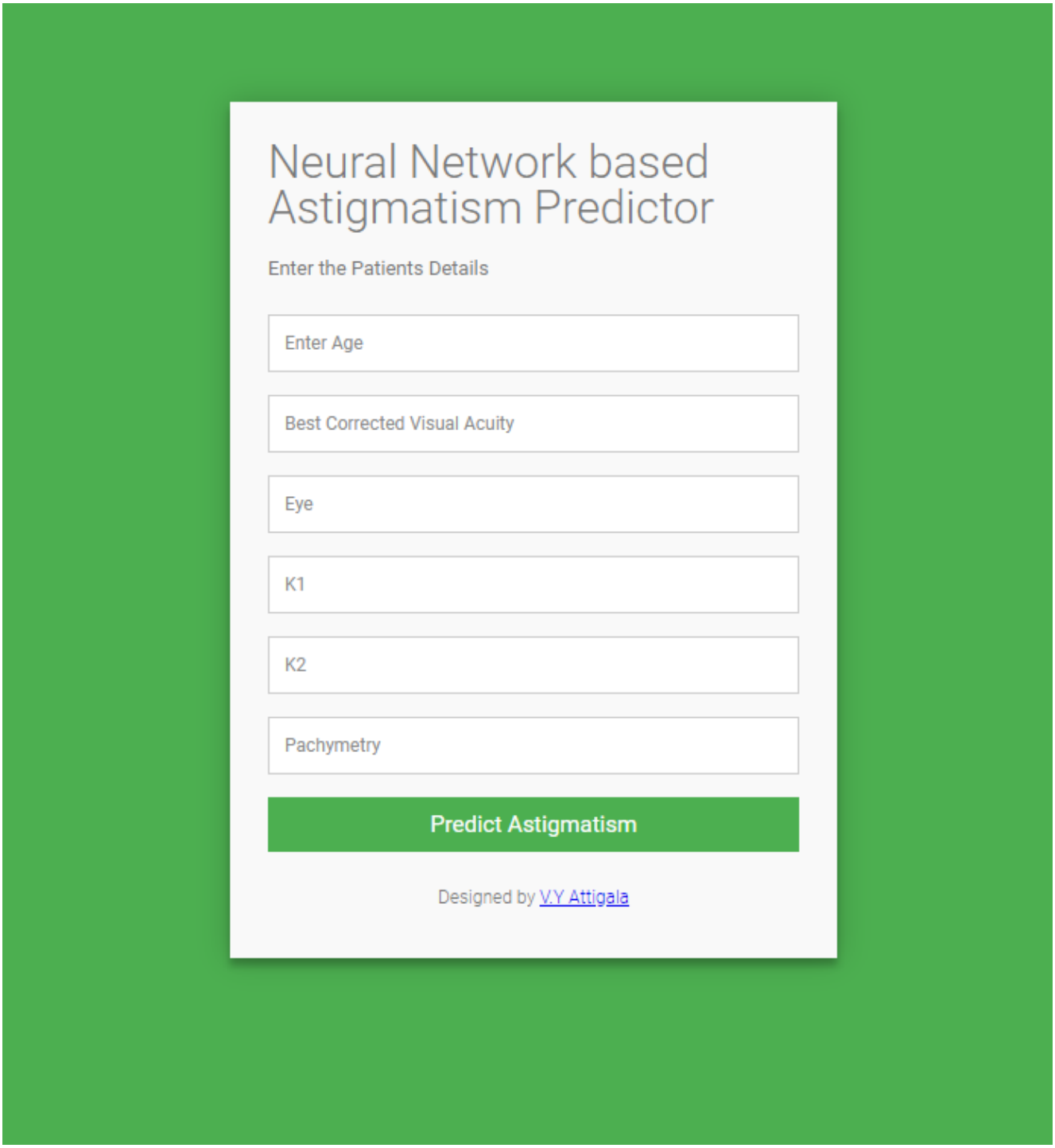


Figure 11 Front End App for the Doctor

The following diagram indicates the high-level architecture of the system.

High Level Architecture diagram

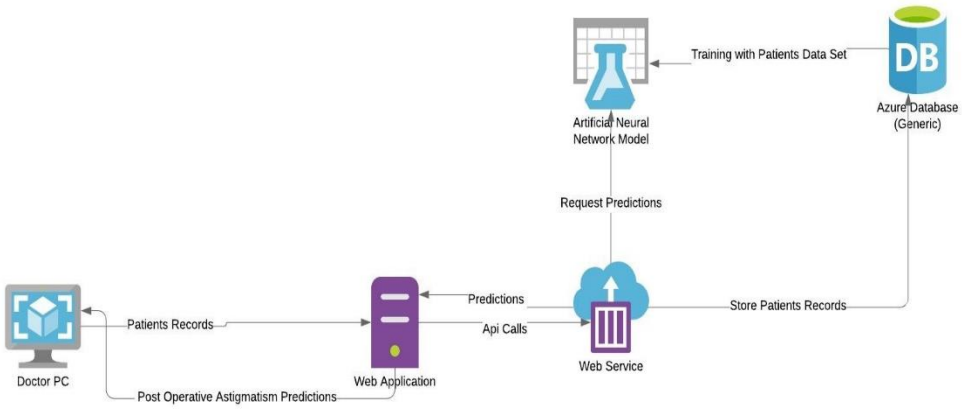


Figure 12 Architecture Diagram

Use Case Diagram

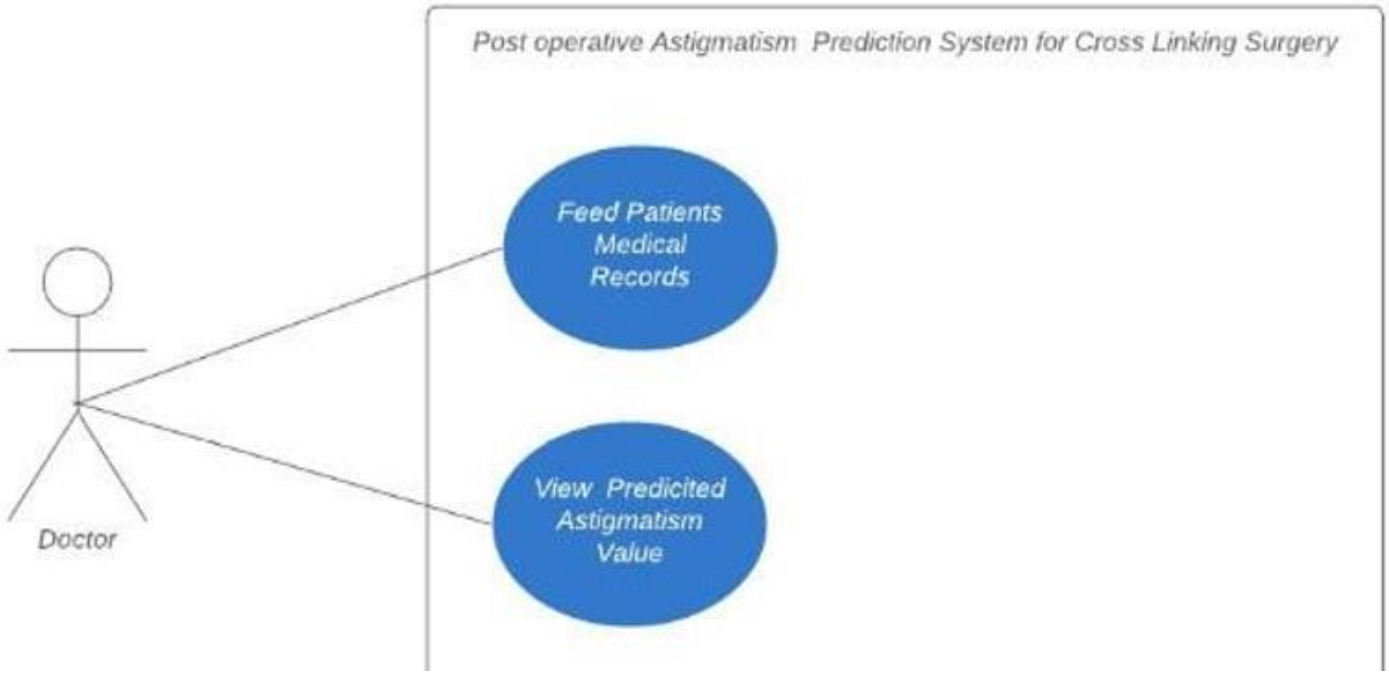


Figure 13 Use Case Diagram

Chapter 04
Evaluation and Results

The research is proposed to develop a computational model, using artificial neural networking to predict the astigmatism level following the corneal cross-linking surgery. the researcher must identify what preoperative factors and post-operative factors can be used to accurately predict the astigmatism following the corneal surgery. The following research objectives are needed to satisfy to evaluate the outcome of the research.

1. To identify pre-operative and post-operative variables associated with astigmatism.
2. To design a computational model using Artificial Neural Networking which can predict the postoperative astigmatism.
3. To implement an End user application which can transfer the knowledge of the neural network to the end users.
4. To identify the controlling factors for Keratoconus ectasia

The evaluation approach that the research team is going to address is experimental based evaluation approach. The reason to choose the mentioned approach is, the researcher can conduct experiments on the given dataset using machine learning models and the researcher can strictly follow the scientific method. Another reason to select the approach is the researcher can control the pre-operative variables and post-operative variables related to the corneal cross-linking surgery to calculate predict or compare with astigmatism value.

Two datasets have been used in the research to build the neural network and identify the correlations between pre-operative and post-operative variables.

5. Dataset from the corneal unit National Eye Hospital Sri Lanka (Dataset 1)
6. Replication Data for the "Keratoconus severity identification using unsupervised machine learning", Siamak Yousefi, Ebrahim Yousefi, Hidenori Takahashi, Takahiko Hayashi, Hironobu Tampo, Satoru Inoda, Yusuke Arai, and Penny Asbell, PLOS One 2018 (Dataset 2)
7. Melendy Dataset for kerataconus

The initial step is to identify correlations between given dataset variables in order to satisfy the first objective in the research. According to the both datasets following pre-operative variables are mentioned to find a correlation between variables

- Age
- Sex
- Eye (OD/OS)
- K1
- K2
- Cylinder
- Spherical Equivalent

The issue needs to be address is that comparing the validity of the two datasets. The comparison can be discussed with the statistical regression analysis using Microsoft excel as the analysis tool

First the research team has used Microsoft excel pivot table has been used to identify the correlations. They can be mentioned as follows.

Analysis of the Dataset from the corneal unit National Eye Hospital Sri Lanka

Initial dataset contains missing values the average missing values substitute method had to be used to improve the accuracy of the dataset. The best corrected visual acuity is the clearest vision that a patient can obtain using the spectacles or contact lenses. According to the opinion of the eye surgeons in sri lanka the best corrected visual acuity can be taken as the measure of success of the surgery. The researcher used the best corrected visual accuracy as the dependent variable of the surgery in the initial data analysis.

The following result was obtained from the excel Data charts analysis.

Average of BSCVA Column Labels						
Row Labels	Pre OP	1M	3M	6M	12M	Grand Total
F	0.461111111	0.42777778	0.525	0.5	0.55	0.48421053
M	0.576470588	0.66470588	0.65294118	0.62352941	0.68333333	0.63378378
Grand Total	**0.536538462**	**0.58269231**	**0.612**	**0.584**	**0.63**	**0.58303571**

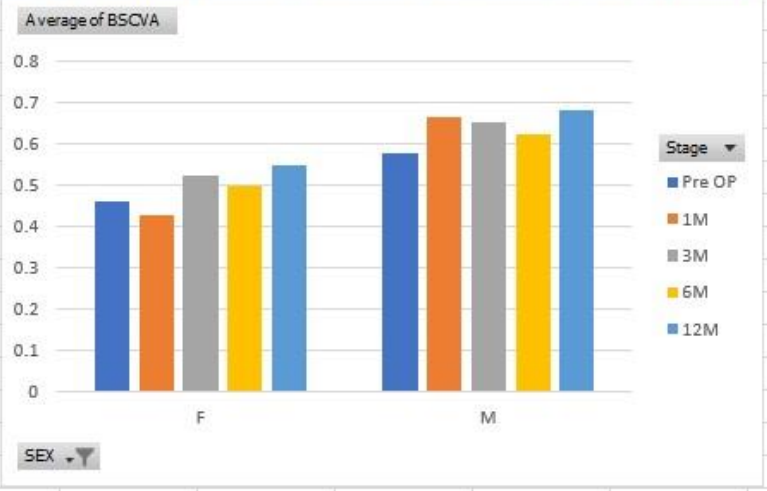


Figure 14 BCVA with Gender

From this result set the average success rate of male patients is higher compared to the female patients can be mentioned as an anomaly. The periodic change before the surgery and follow ups after one month, three months, six months and twelve months is compared along with the males and female gender.

Next the age of the patient is compared with the analysis. The following result was obtained.

Age

Average of BSCVA	Column Labels		
Row Labels	Pre OP	1M	Grand Total
15	0.5	0.7	0.6
16	0.7	0.7	0.7
17	0.575	0.65	0.6125
18	0.5	0.7	0.6
19	0.35	0.4	0.375
20	0.325	0.5	0.4125
21	0.5	0.5	0.5
22	0.5	0.5	0.5
23	0.7	0.5	0.6
24	0.85	0.75	0.8
25	0.55	0.6	0.575
27	0.275	0.475	0.375
28	0.3	0.3	0.3
30	0.7	0.7	0.7
Grand Total	**0.524074074**	**0.57962963**	**0.551851852**

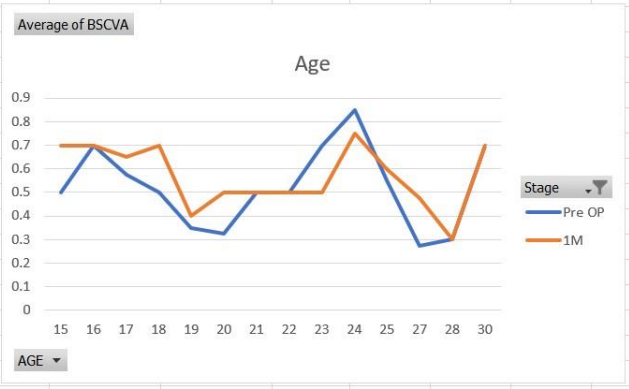


Figure 15 BCVA with Age

From the graph, the researcher can say the success rate of the patients between 15-20 is higher compared to the other age categories the is an outlier detected at the age 25 mark there is only one patient is listed in the given dataset.

The cases in the dataset for all 12 months also plotted as below.

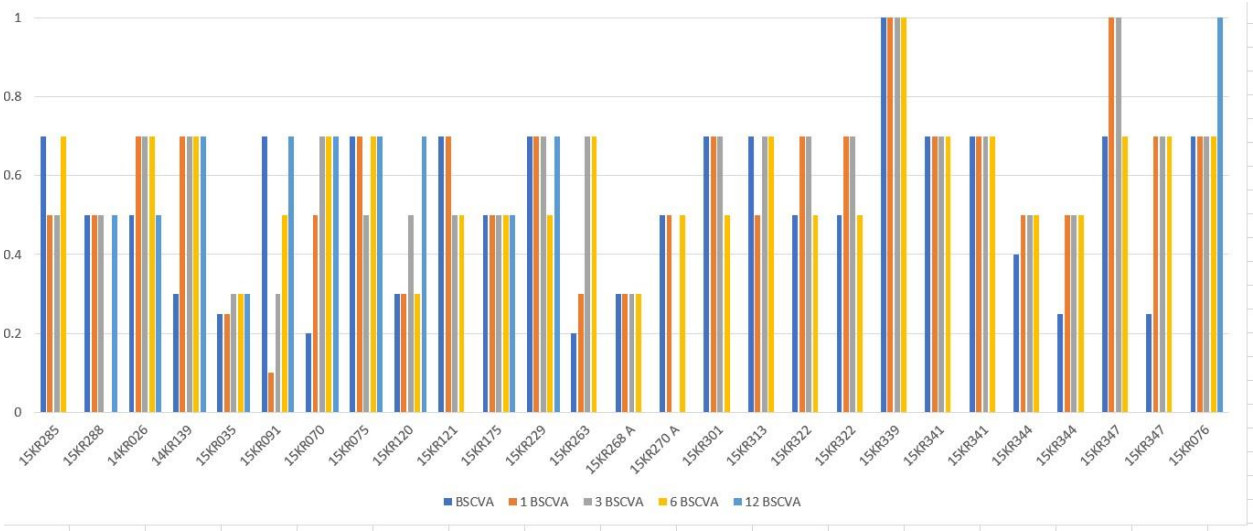


Figure 16 BCVA with Case

Correlation Analysis between the Datasets.

From the dataset obtained from the national eye hospitals and the dataset from the melendy dataset for keratoconus has been analyzed with the statistical approaches. This is to ensure the variables that is used

for the neural network has a dependency from each other. The following shows the correlation between the variables from before the surgery up to the 12 months after the surgery.

1. Correlation between variables in the pre-Operative patients.

	eye	bcva	k1	k2	se	cylinder	pachymetry
eye	1						
bcva	-0.25442	1					
k1	-0.08641	0.11388	1				
k2	-0.04481	0.270798	-0.22103	1			
se	0.466285	-0.02507	-0.01029	0.424059	1		
cylinder	0.409028	0.106303	-0.52932	0.485275	0.815011	1	
pachymetry	-0.03864	-0.0309	0.000514	0.101417	0.048913	-0.10284	1

Figure 17 Correlation 1

From the result set there is a negative correlation between k1 and the cylinder. Also, there is a positive correlation between k2 and the cylinder.one month after the surgery the correlation changed as follows.

	eye	bcva	k1	k2	se	cylinder	pachymetry
eye	1						
bcva	-0.376996588	1					
k1	0.364957322	0.101257	1				
k2	0.006324697	0.065335	-0.16473	1			
se	-0.024025267	-0.15986	-0.17358	-0.14831	1		
cylinder	-0.210589968	0.000892	0.213597	-0.27107	-0.18919	1	
pachymetry	-0.053303543	-0.06934	-0.15061	0.089352	0.292877	0.170832	1

Figure 18 Correlation 2

Now the correlation between k1 and cylinder changed from negative correlation into a positive correlation. The k2 which had a positive correlation has a negative correlation after the surgery. After six month the result changed as follows.

	eye	bcva	k1	k2	se	cylinder	pachymetry
eye	1						
bcva	0.146853	1					
k1	0.270821	0.002892	1				
k2	0.079063	-0.13231	-0.43815	1			
se	-0.08608	-0.38226	-0.02376	0.513699	1		
cylinder	-0.20569	-0.33603	-0.54748	0.64585	0.850331	1	
pachymetry	0.035281	0.318673	0.334681	0.003066	-0.34636	-0.47621	1

Figure 19 Correlation 3

Now there is a new positive correlation between best corrected visual acuity and the pachymetry. The pachymetry is related to the corneal thickness. The surgery data that was used in the research was a epi-off surgery that means the part of the cornea get pealed off. Once it starts to grow back the visual acuity is going to go high, the results with the statistical analysis confirms that because of the positive correlation.

After 1 year following the surgery there is a negative correlation between pachymetry and the best corrected visual acuity. And the k2 and the cylinder value correlation has been stable after one year following the crosslinking surgery. The following is the analysis result of the correlation.

	eye	bcva	k1	k2	cylinder	pachymetr
eye	1					
bcva	-0.05967	1				
k1	-0.73361	-0.32825	1			
k2	-0.24278	-0.14149	0.59999	1		
cylinder	-0.21847	-0.43849	0.690169	0.911303	1	
pachymet	-0.17565	-0.36859	0.07147	-0.20276	-0.12054	1

Figure 20 Correlation 5

Outcome of the Deep Learning Neural Network Model in keras.

The data from the other two datasets needs to be feeds to the neural network model and predict the astigmatism and k readings. Currently the accuracy of the developed neural network is as follows

Number of neurons in input layer 5.

Hidden layers 2 each layer contains 150 neurons

One output layer to predict k1 k2 values. The k1 and k2 values are associated with the astigmatism for predicting the astigmatism k1 and k2 prediction will do the stuffiest.

The following expected and the actual result was identified

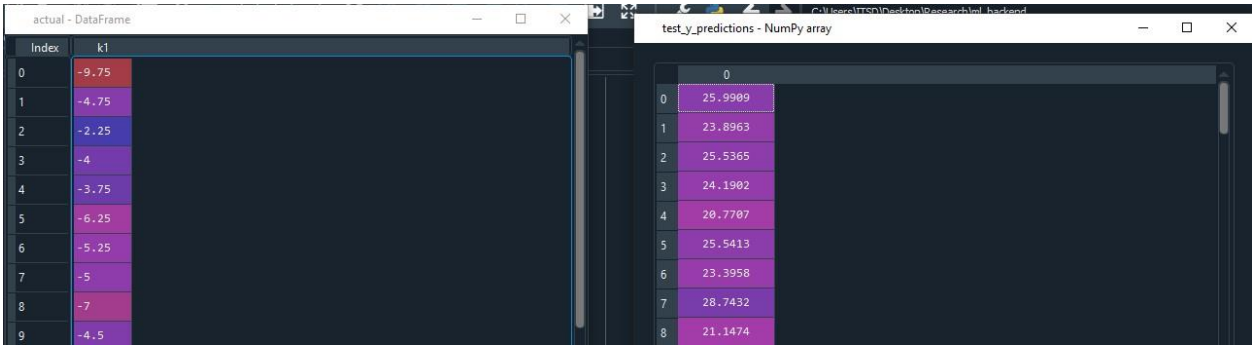


Figure 21 Prediction 1

The overall loss of the neural networking model is as follows. According to the following output the loss should be reduced to improve the accuracy of the hypothesis model.

```
In [19]: runfile('C:/Users/ITSD/Desktop/Research/ml_backend/predict_ast.py', wdir='C:/Users/ITSD/Desktop/Research/ml_backend')
Train on 64 samples, validate on 16 samples
Epoch 1/30
64/64 [==============================] - 0s 2ms/step - loss: 2303.7705 - val_loss: 472.0075
Epoch 2/30
64/64 [==============================] - 0s 244us/step - loss: 937.5914 - val_loss: 49.7642
Epoch 3/30
64/64 [==============================] - 0s 0us/step - loss: 1132.6436 - val_loss: 80.7744
Epoch 4/30
64/64 [==============================] - 0s 0us/step - loss: 867.8234 - val_loss: 18.0066
Epoch 5/30
64/64 [==============================] - 0s 242us/step - loss: 434.5393 - val_loss: 378.3542
Epoch 6/30
64/64 [==============================] - 0s 0us/step - loss: 425.2029 - val_loss: 824.8141
Epoch 7/30
64/64 [==============================] - 0s 0us/step - loss: 586.5900 - val_loss: 737.3311

In [20]:
```

Figure 22 Prediction 2

The accuracy of the dataset is low due to lack of more dataset which can be improved using the dataset from the online resource since the coved 19 outbreaks in Sri Lanka made the researcher impossible to access the patient's clinic files from the eye hospital. Once the situation is under control the issue of the dataset can be solved to improve the model.

Then the researcher tried with the way of minimizing the overfitting with early stopping monitor. The following results was obtained.

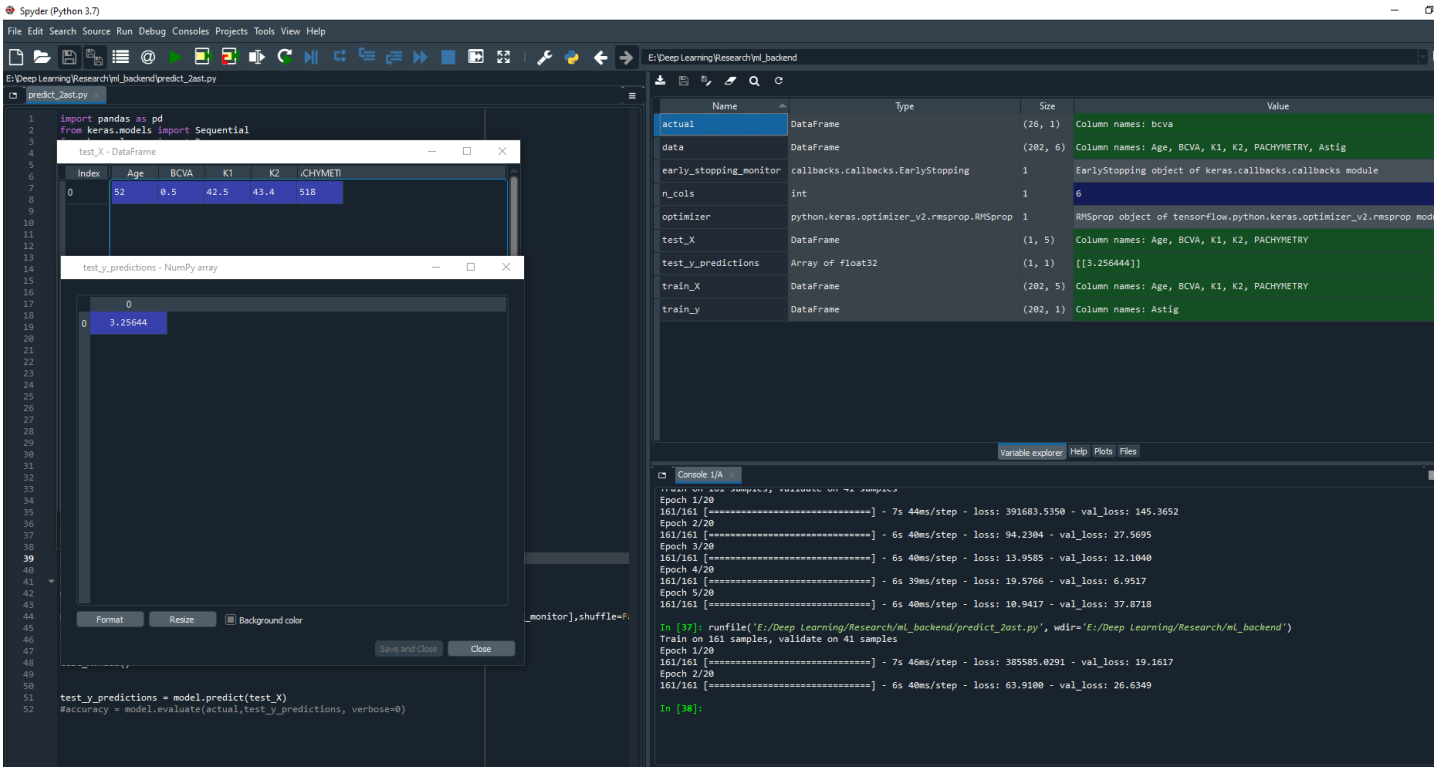


Figure 23 Predcition 3

The accuracy of the model was not up to a level which could be the reason due to low amount of dataset.in order to see the problem in different perspective the researcher used weka classifying algorithms to get more insights.

Random tree analysis.

Results obtained from the random forest tree is as follows.

```
|   |   |   |   PACHYMETRY >= 540.5 : 40.6 (1/0)
|   |   K2 >= 43.7
|   |   |   K2 < 48.35
|   |   |   |   Age < 53
|   |   |   |   |   Age < 50
|   |   |   |   |   |   Astig < -3.48
|   |   |   |   |   |   |   K2 < 46.8 : 41.05 (2/0)
|   |   |   |   |   |   |   K2 >= 46.8
|   |   |   |   |   |   |   |   K2 < 47.65 : 42.5 (1/0)
|   |   |   |   |   |   |   |   K2 >= 47.65 : 41.6 (2/0.01)
|   |   |   |   |   |   Astig >= -3.48
|   |   |   |   |   |   |   K2 < 44.35
|   |   |   |   |   |   |   |   Age < 25 : 41.3 (1/0)
|   |   |   |   |   |   |   |   Age >= 25
|   |   |   |   |   |   |   |   |   Age < 28.5 : 42.1 (1/0)
|   |   |   |   |   |   |   |   |   Age >= 28.5 : 41.65 (2/0)
|   |   |   |   |   |   |   K2 >= 44.35
|   |   |   |   |   |   |   |   K2 < 45.85
|   |   |   |   |   |   |   |   |   BCVA < 0.55
|   |   |   |   |   |   |   |   |   |   Astig < -2.87 : 41.3 (0.67/0)
|   |   |   |   |   |   |   |   |   |   Astig >= -2.87
|   |   |   |   |   |   |   |   |   |   |   Age < 22 : 42.45 (4/0.01)
|   |   |   |   |   |   |   |   |   |   |   Age >= 22 : 42.12 (2.67/0.01)
|   |   |   |   |   |   |   |   |   BCVA >= 0.55
|   |   |   |   |   |   |   |   |   |   Astig < -2.5 : 41.3 (0.33/0)
|   |   |   |   |   |   |   |   |   |   Astig >= -2.5
|   |   |   |   |   |   |   |   |   |   |   Astig < -1.79 : 43.8 (1/0)
|   |   |   |   |   |   |   |   |   |   |   Astig >= -1.79
|   |   |   |   |   |   |   |   |   |   |   |   Age < 30 : 43 (1/0)
```

Figure 24 Prediction 4

According to the above scenario the patients who has undergone the corneal cross-linking surgery has the following pattern to be discussed.

Age Group	Best Corrected Visual acuity	K1	K2	Astigmatism
50-53	<0.55	46.8	41.05	<-3.14

Microsoft neural network algorithm also used to verify the findings in the research. This proves the relationship between the age group and the best corrected vision acuity..

Results with the Microsoft Neural Network

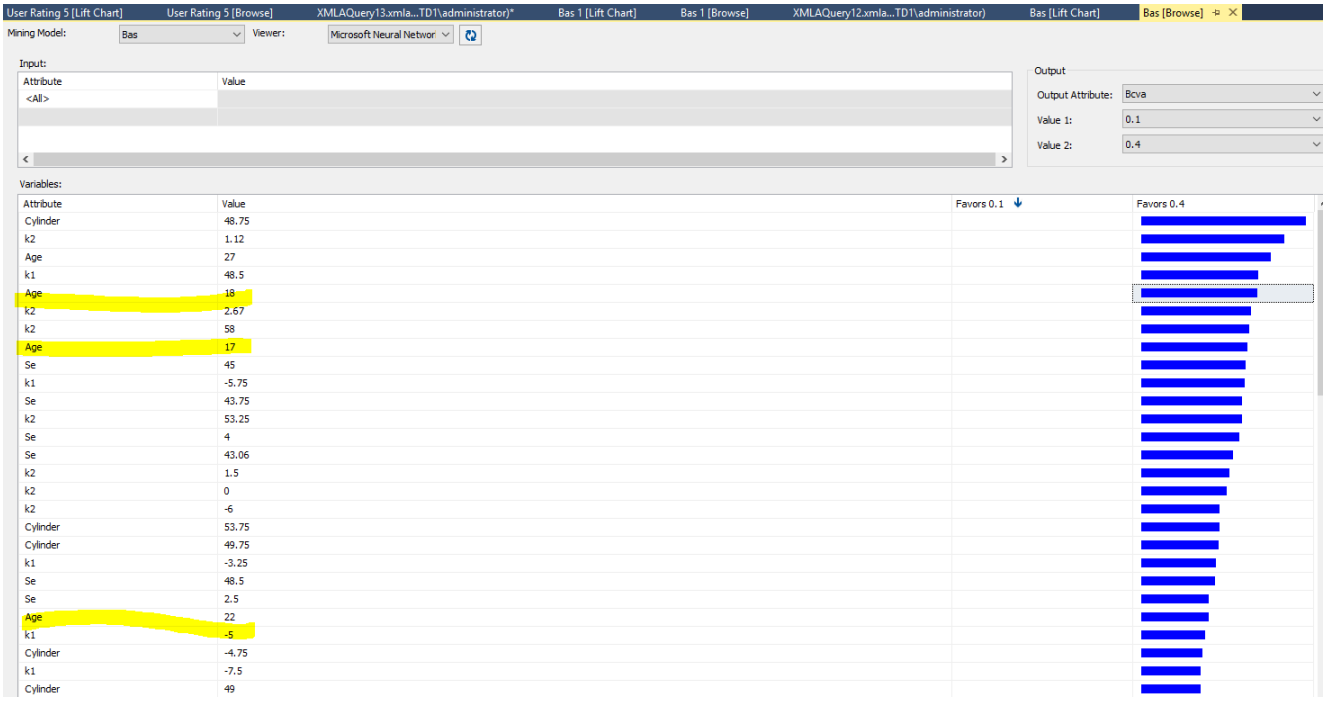


Figure 25 Age Comparison 1

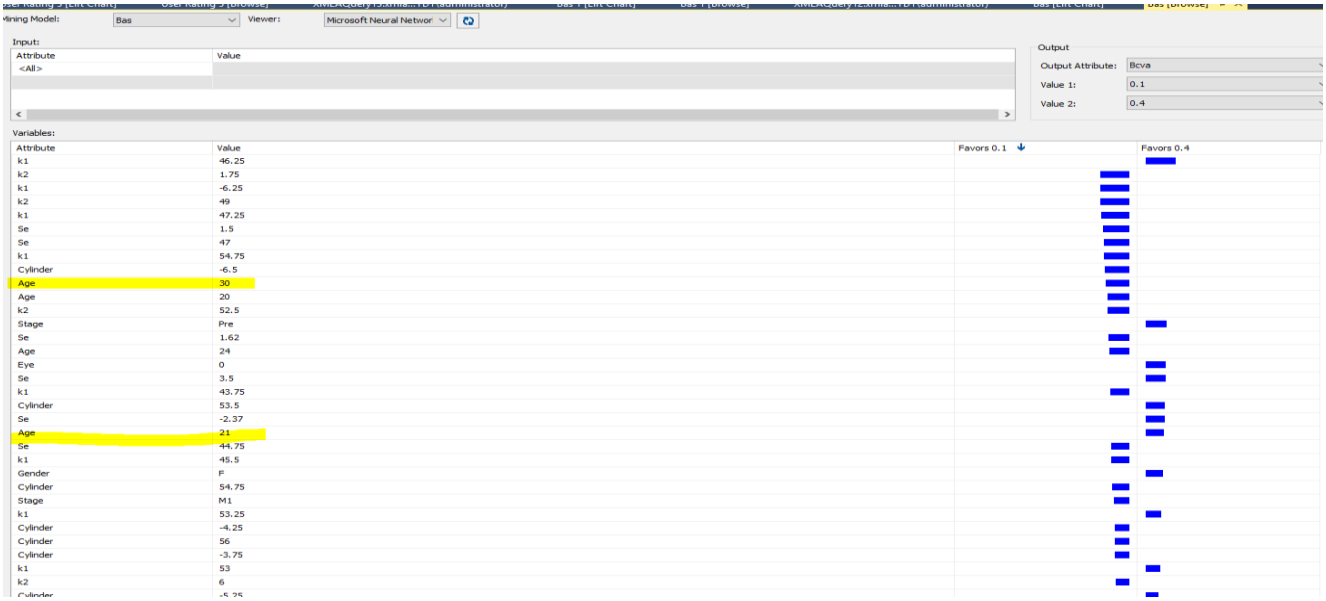


Figure 26 Age Comparison 2

The Results without calculated neural network and Transformation functions.

The results tested with the following set of parameters to the neural networking model/

```
model.add(Dense(5, activation='relu', input_shape=(n_cols,)))
model.add(Dense(100, activation='relu'))
model.add(Dense(150, activation='relu'))
model.add(Dense(1))
model.compile(optimizer='adam', loss='mean_squared_error')
early_stopping_monitor = EarlyStopping(patience=3)
model.fit(train_X, train_y, validation_split=0.2, epochs=30,
callbacks=[early_stopping_monitor],shuffle=False)
```

the results were not correlated with the actual dataset as given in the following output.

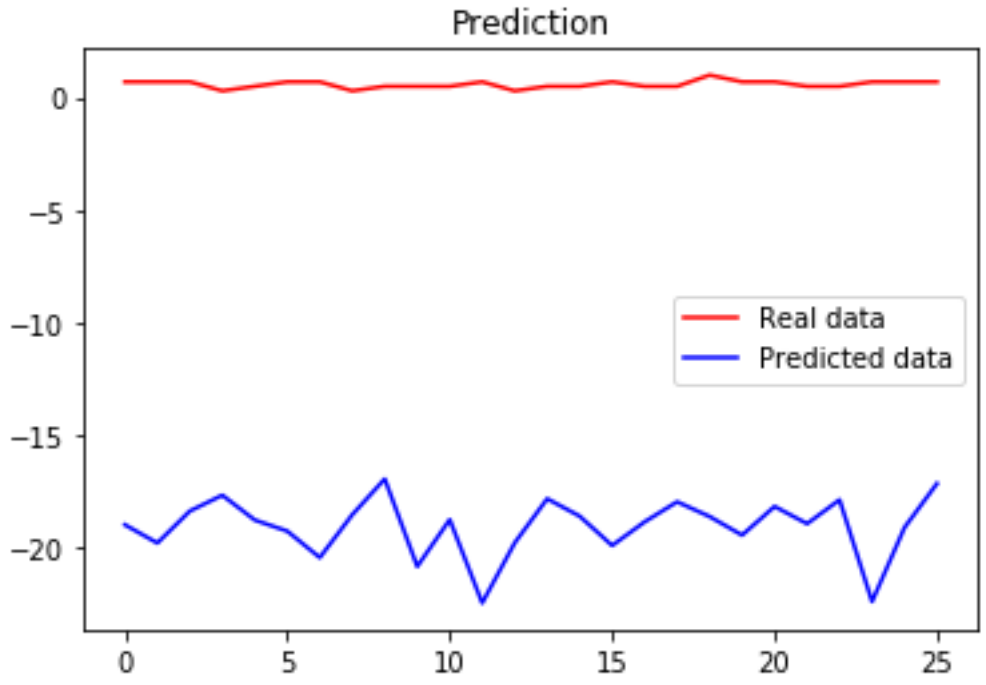


Figure 27 Uncalculated Result

Since the accuracy of the model was not up to the mark by evaluating the results the researcher tried to apply the calculation of the hidden layers stated in the methodology section of the documentation. The following result was obtained according to the figure 28 and 29

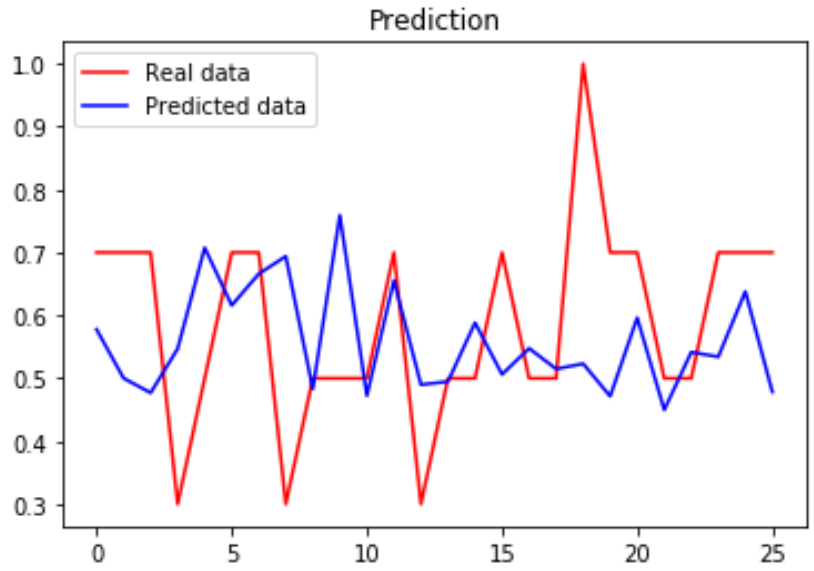


Figure 28 transformed model without Calculated Layers

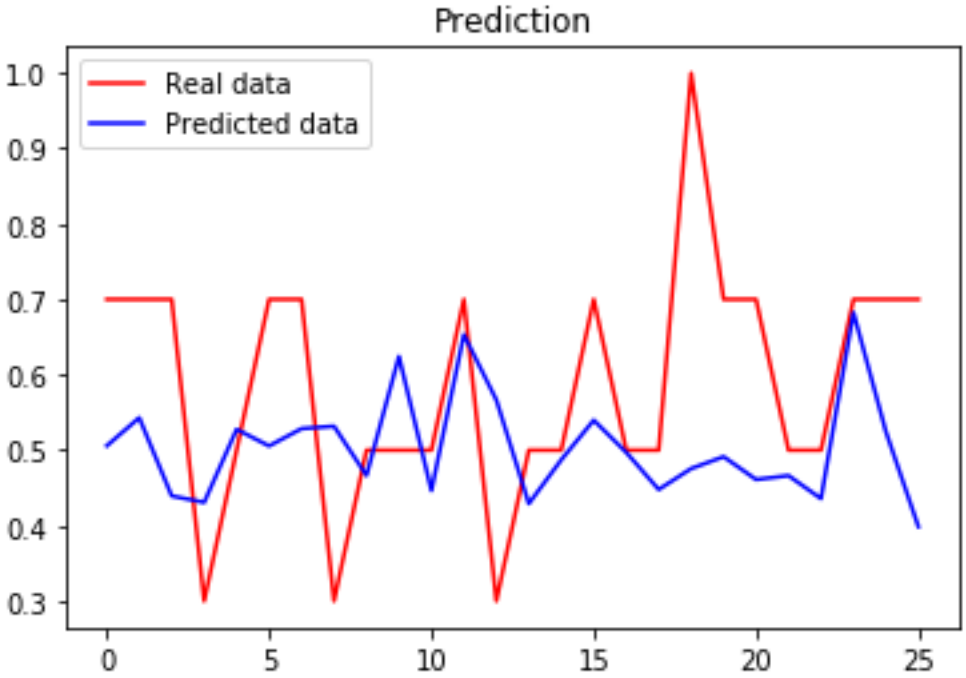


Figure 29 Transformed data with calculated layers

Usage of radial basis activation function.

The researcher has experimented with the radial basis activation function to improve the accuracy of the model. And was able to discover close relationship between with relu activation function and the radial basis activation function. Given in the figure 30.

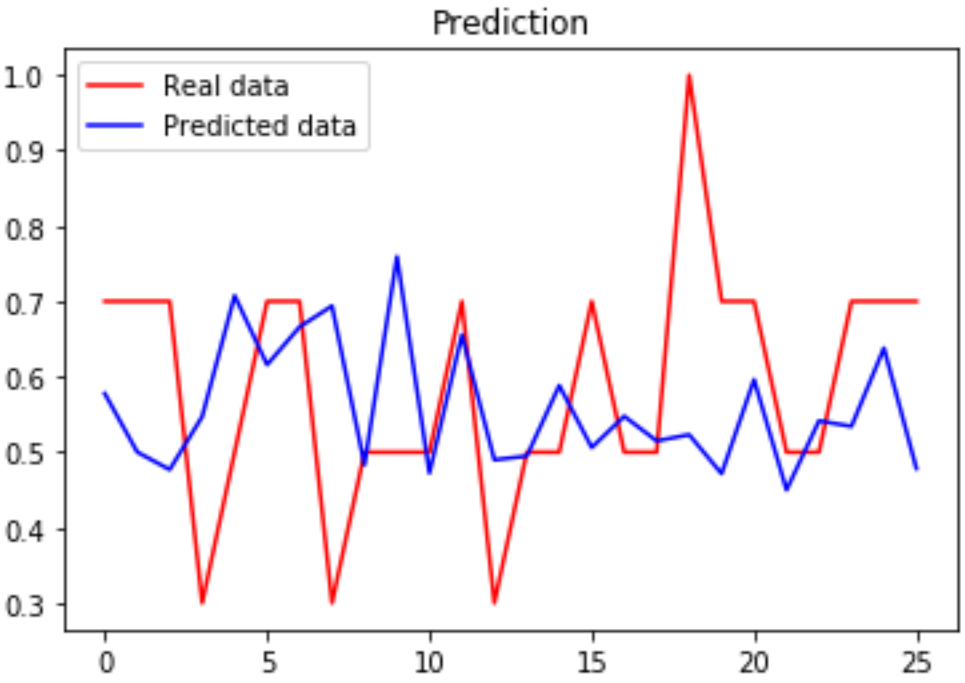
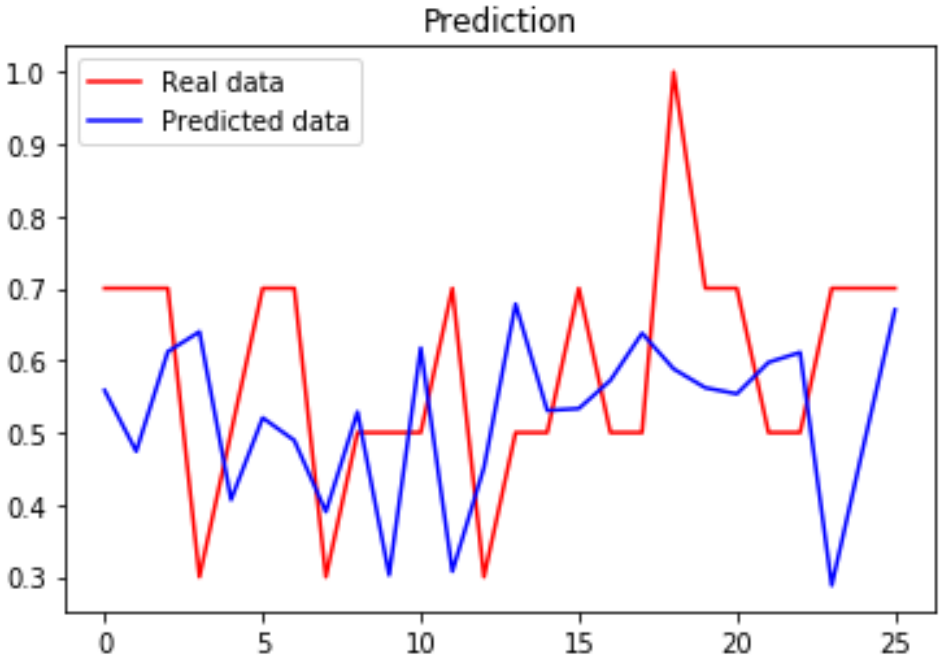


Figure 30 without radial basis and with radial basis.



Chapter 05
Conclusion and future work

To conclude the research, the researcher can express that the number of variables and amount of dataset was not enough to provide an accurate neural network. But the researcher used different machine learning techniques to verify the building of the neural network.by looking at the neural network having three intermediate dense layers with the relu activation function leads to the insights that the researcher discussed in the evaluation section. This research encourage the fellow researcher to having a larger dataset of keratoconus patients who are needs to undergo cornea cross linking surgery from the statistical methods and the neural network model the researcher can say under epi-off method a improvement of the vision and the astigmatism can be seen from the k1 and the k2 parameters between 1 to 6 months. The researcher had to use the dataset from the melendy to train the dataset with high accuracy. The model created from the neural network provides promising results. because of the covid-19 outcome the researcher wasn't able to complete the research with the high accuracy. The researcher has satisfied the following research objectives.

- To identify pre-operative and post-operative variables associated with astigmatism.
- To design a computational model using Artificial Neural Networking which can predict the post-operative astigmatism.
- To implement an End user application which can transfer the knowledge of the neural network to the end users

Identifying the controlling factors for kerataconus ectasia is inconclusive due to the accuracy issue of the dataset. As a limitation the missing values the researcher replaced using the mean value and a mathematical formula related to the dataset. Replacing missing values with the statistical dataset is not a recommended way to handle the missing data issue.

For the future work, this research can be polished with higher accuracy. if more parameters and dataset could be obtained from the patients in the cornea clinic. The researcher has to refer to the patient's clinic files. The bio mechanical and the surgical factors could not be validated due to the limited number of parameters obtained before the covid-19 outbreak. The researchers can apply Bio Mechanical and surgical variables into the neural network if the access to the dataset was available without the crisis of the COVID-19Different parameters can be used to improve the accuracy of the neural networkEnd user application can be enhanced to show graphical representations

Keratoconus has becoming increasingly popular in sri lanka if a valid dataset from several hospitals and eye clinics could be obtained the researchers can

References

[1] M. Hovakimyan, R. F. Guthoff and O. Stachs, “Collagen cross-linking: current status and future directions,” *Journal of Ophthalmology,* p. 12, 2012.

[2] S. Kling and Farhad Hafezi, “An Algorithm to Predict the Biomechanical Stiffening Effect in Corneal Cross-linking,” *Journal of Refractive Surgery,* 2017.

[3] Sykakis, Karim, Bunce, Amissah-Arthu, Patwary and Hamada S, “Corneal collagen cross-linking for treating keratoconus,” *Database of Systematic Reviews,* 2015.

[4] B. T. Lopes1, Ashkan Eliasy and Renato Ambrosio, “Artificial Intelligence in Corneal Diagnosis: Where Are we?,” *Current Ophthalmology Reports (2019),* 2019.

[5] S. Haykin, “Neural Networks and Learning Machines,” 2009.

[6] M. Valdés-Mas, J.D. Martín-Guerrero, M.J. Rupérez, F. Pastor, C. Dualde, C. Monserrat and C. PerisMartínez, “A new approach based on Machine Learning forpredicting corneal curvature (K1) and astigmatismin patients with keratoconus after intracorneal ringimplantation,” 2014.

[7] P. Vinciguerra, Elena Albè, Silvia Trazza, Theo Seiler and Daniel Epstein, “Intraoperative and Postoperative Effects of Corneal Collagen Cross-linking on Progressive Keratoconus,” 2009.

[8] O. Sarac, Mehtap Caglayan, Hasan Basri Cakmak and Nurullah Cagil, “Factors Influencing Progression of Keratoconus 2 Years After Corneal Collagen Cross-Linking in Pediatric Patients,” 2016.

[9] R. P. WISSE, DANIE¨L A. GODEFROOIJ, NIENKE SOETERS, SASKIA M. IMHOF and ALLEGONDA VAN DER LELIJ, “A Multivariate Analysis and Statistical Model for Predicting Visual Acuity and Keratometry One Year After Cross-linking for Keratoconus,” 2014.

[10] A. Cantemir, Anisia-Iuliana Alexa, Bogdan Gabriel Galan, Nicoleta Anton, Roxana Elena Ciuntu, Ciprian Danielescu, Dorin Chiselita and Danut Costin, “Iontophoretic collagen cross-linking versus epithelium-off collagen cross-linking for early stage of progressive keratoconus – 3 years follow-up study,” 2017.

[11] O. Sarac, Mehtap Caglayan, Hasan Basri Cakmak and Nurullah Cagil, “Factors Influencing Progression of Keratoconus 2 Years After Corneal Collagen Cross-Linking in Pediatric Patients,” 2010.

[12] A. Gordon-Shaag, M. Millodot, E. Shneor and Y. Liu, "The Genetic and Environmental Factors for Keratoconus," 2015.

Annexure

Data Collection Sheet Form Section-A

(for Patients Follow Up)

Case Number

Project - A computational model using artificial neural networking for predicting astigmatism following corneal surgery

1. DEMOGRAPHICS

Age	
Sex	
Town	
Consultant	
Town	
Ethnic category	

2. History

Scheme	**OD**	**OS**
Previous CXL		
Previous refractive surgery		
Previous diseases of the ocular surface		
Does the patient have atopic conditions?		

3. ASSESSMENTS

Date of visit	**OD**	**OS**
Vison/Refraction		
Reason for visit (new patient, follow-up)		
Vision correction used by patient (unaided, spectacles, SCL, RGP, hybrid, scleral)		
Chart used to record patient's visual acuity (Snellen, EDTRS)		
Uncorrected distance visual acuity		
Corrected distance visual acuity		
Refractive spherical error		
Refractive cylinder		
Axis of cylinder		
Topography		
Steepest keratometry over the whole cornea (KMax)		
Flattest front curvature in the 3 mm zone Front k1		

Steepest front curvature in the 3 mm zone Front k2		
Flattest back curvature in the 3 mm zone Back k1		

Steepest back curvature in the 3 mm zone Back k2		
Thinnest pachymetry		
Belin Ambrosio (Enhanced ectasia score)		
Quality score front (OK, Borderline yellow, Poor red)		
Quality score back OK, Borderline yellow, Poor red)		
CL removed prior to scan?		
Specular microscopy		
Endothelial cell density /mm2		
Coefficient of variation		
Slit-lamp		
Ocular surface inflammation		

Ocular surface inflammation = other		
Cornea Any other relevant signs?		
Outcome		
The diagnosis reached at the end of the assessment		
Is the ectasia stable or has it progressed(Status)		
Reason for deciding status		
What decision is made at the end of each assessment		
Comments		

4. CORNEAL CROSS-LINKING SURGERY

Date of surgery	
Grade of person performing CXL (Optometrist, nurse, consultant, fellow, ST)	
UVA device used IROC UVX-1000, IROC UVX 2000, Avedro KXL I, Avedro KXL II, Peschke CCL-Vario, other)	
Was the epithelium removed? (On, off, partial disruption)/ Epithelial status	
How was the epithelium removed? (manual, ETOH-assisted, transPTK, transPRK) /Epithelial debridement	

Debridement size/ Diameter of epithelial defect	
Iontophoresis Was iontophoresis used?	
Iontophoresis settings Iontophoresis current	
Iontophoresis (Iontophoresis settings) duration	
Name of solution Riboflavin preparation	
Riboflavin duration	
Total riboflavin soak time/Riboflavin duration	
UV power/UV irradiance	
Duration of UV exposure	
UV continuous or pulsed Continuous or on/off phase of pulsed exposure (seconds)	
Topography guided?	
UV total energy (Energy (J/cm2) = power (W/cm2) x duration (seconds)	
Duration of BCL wear How many days?	
Post-operative therapy Drops used following CXL	
Steroids before epithelium healed	

Additional surgery comments	

Microsoft Neural Network Model

<DataSourceView xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:ddl2="http://schemas.microsoft.com/analysisservices/2003/engine/2" xmlns:ddl2_2="http://schemas.microsoft.com/analysisservices/2003/engine/2/2" xmlns:ddl100_100="http://schemas.microsoft.com/analysisservices/2008/engine/100/100" xmlns:ddl200="http://schemas.microsoft.com/analysisservices/2010/engine/200" xmlns:ddl200_200="http://schemas.microsoft.com/analysisservices/2010/engine/200/200" xmlns:ddl300="http://schemas.microsoft.com/analysisservices/2011/engine/300" xmlns:ddl300_300="http://schemas.microsoft.com/analysisservices/2011/engine/300/300" xmlns:ddl400="http://schemas.microsoft.com/analysisservices/2012/engine/400" xmlns:ddl400_400="http://schemas.microsoft.com/analysisservices/2012/engine/400/400" xmlns:ddl500="http://schemas.microsoft.com/analysisservices/2013/engine/500" xmlns:ddl500_500="http://schemas.microsoft.com/analysisservices/2013/engine/500/500" xmlns:dwd="http://schemas.microsoft.com/DataWarehouse/Designer/1.0" dwd:design-time-name="ba65f82f-c9de-44c7-85bb-695e5f3adbd6" xmlns="http://schemas.microsoft.com/analysisservices/2003/engine">

<ID>Testg</ID>

<Name>Testg</Name>

<CreatedTimestamp>0001-01-01T00:00:00Z</CreatedTimestamp>

<LastSchemaUpdate>0001-01-01T00:00:00Z</LastSchemaUpdate>

<Annotations>

<Annotation>

<Name>http://schemas.microsoft.com/DataWarehouse/Designer/1.0:ShowFriendlyNames</Name>

<Value>true</Value>

</Annotation>

<Annotation>

<Name>http://schemas.microsoft.com/DataWarehouse/Designer/1.0:CurrentLayout</Name>

<Value>_ALL_TABLES_</Value>
</Annotation>
<Annotation>
<Name>http://schemas.microsoft.com/DataWarehouse/Designer/1.0:SchemaRestriction</Name>
<Value></Value>
</Annotation>
<Annotation>
<Name>http://schemas.microsoft.com/DataWarehouse/Designer/1.0:RetrieveRelationships</Name>
<Value>true</Value>
</Annotation>
<Annotation>
<Name>http://schemas.microsoft.com/DataWarehouse/Designer/1.0:Layouts</Name>
<Value>
<Layouts xmlns="">
<Diagram>
<Name>_ALL_TABLES_</Name>
<DiagramLayout />
<ShowRelationshipNames>False</ShowRelationshipNames>
<UseDiagramDefaultLayout>True</UseDiagramDefaultLayout>
<DiagramViewPortLeft>0</DiagramViewPortLeft>
<DiagramViewPortTop>0</DiagramViewPortTop>
<DiagramBoundingLeft>0</DiagramBoundingLeft>
<DiagramBoundingTop>0</DiagramBoundingTop>
<DiagramZoom>100</DiagramZoom>
</Diagram>

```
      </Layouts>
    </Value>
   </Annotation>
  </Annotations>
  <DataSourceID>Testg</DataSourceID>
  <Schema>
   <xs:schema id="Testg" xmlns="" xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns:msdata="urn:schemas-microsoft-com:xml-msdata" xmlns:msprop="urn:schemas-microsoft-com:xml-msprop">
    <xs:element name="Testg" msdata:IsDataSet="true" msdata:UseCurrentLocale="true" msprop:design-time-name="b28a2710-e990-42da-bff7-1a66bc132a70">
     <xs:complexType>
      <xs:choice minOccurs="0" maxOccurs="unbounded">
       <xs:element name="dbo_bas" msprop:design-time-name="18b05bf4-820b-4b95-9338-5591d5f08b17" msprop:DbTableName="bas" msprop:FriendlyName="bas" msprop:DbSchemaName="dbo" msprop:TableType="Table">
        <xs:complexType>
         <xs:sequence>
          <xs:element name="clinic_id" msprop:design-time-name="87b2cc87-a15f-49ca-b102-a6367327964e" msprop:DbColumnName="clinic_id" msprop:FriendlyName="clinic_id">
           <xs:simpleType>
            <xs:restriction base="xs:string">
             <xs:maxLength value="50" />
            </xs:restriction>
           </xs:simpleType>
          </xs:element>
          <xs:element name="age" msprop:design-time-name="d9081ddc-ecbd-4e53-9a6b-421ce5001365" msprop:DbColumnName="age" msprop:FriendlyName="age" type="xs:int" />
```

<xs:element name="gender" msprop:design-time-name="3fd54928-681d-4dea-a3e8-102425a0e12e" msprop:DbColumnName="gender" msprop:FriendlyName="gender">

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:maxLength value="50" />

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="eye" msprop:design-time-name="ce453104-0cb7-456b-a08f-4b374a775e22" msprop:DbColumnName="eye" msprop:FriendlyName="eye">

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:maxLength value="50" />

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="bcva" msprop:design-time-name="dbdebc87-6a47-407d-a4a7-820c8baa8144" msprop:DbColumnName="bcva" msprop:FriendlyName="bcva">

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:maxLength value="50" />

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="k1" msprop:design-time-name="ea321f55-c359-4ba5-94be-3b94f191c001" msprop:DbColumnName="k1" msprop:FriendlyName="k1" type="xs:double" minOccurs="0" />

<xs:element name="k2" msprop:design-time-name="9e6f50c1-fe89-4b28-a8da-32deb54d2c96" msprop:DbColumnName="k2" msprop:FriendlyName="k2" minOccurs="0">

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:maxLength value="50" />

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="se" msprop:design-time-name="79dc3058-ff19-415c-a60f-aad10704a012" msprop:DbColumnName="se" msprop:FriendlyName="se" minOccurs="0">

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:maxLength value="50" />

</xs:restriction>

</xs:simpleType>

</xs:element>

<xs:element name="cylinder" msprop:design-time-name="72283637-08b7-4890-9eba-2a1a26d220e8" msprop:DbColumnName="cylinder" msprop:FriendlyName="cylinder" type="xs:double" minOccurs="0" />

<xs:element name="pachymetry" msprop:design-time-name="d1cbf6e0-d9e6-41d5-ae3a-4abe9745f0db" msprop:DbColumnName="pachymetry" msprop:FriendlyName="pachymetry" type="xs:int" />

<xs:element name="stage" msprop:design-time-name="37a4c149-6e3b-41d8-833f-f5be50129099" msprop:DbColumnName="stage" msprop:FriendlyName="stage">

<xs:simpleType>

<xs:restriction base="xs:string">

<xs:maxLength value="50" />

</xs:restriction>

</xs:simpleType>

</xs:element>

```
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:choice>
    </xs:complexType>
  </xs:element>
 </xs:schema>
 <Testg xmlns="" />
</Schema>
</DataSourceView>
```