University of East London

Using Machine Learning to Determine the Significance of Risk Factors in Diagnosing Keratoconus

A project submitted in partial fulfilment of the requirements for the award of MSc Computer Science

Abin Zorto

Student Number: 2091940

Module Number: CN7000

Project Supervisor: Dr. Saeed Sharif

May /2022

Acknowledgements

Many people contributed and extended their support in various ways to make this dissertation possible. Without their direction and assistance, this dissertation would not have been conceivable.

My sincere gratitude to Dr. Saeed Sharif, my supervisor, for his understanding and patience.

I want to express my gratitude to everyone who has helped me in any way during the course of the project's completion

Abstract

Keratoconus is a non-inflammatory corneal disease that affects both eyes and affects one out of every 2000 people worldwide. The cornea deforms into a conical shape and thins, resulting in high order aberrations and gradual vision loss. Analysis of the risk factors involved in the degradation of keratoconus is under-researched. The goal of this project is to use machine learning models to gain a better understanding of how much the risk factors of a patient contribute towards the progressive stages of keratoconus as well as how significant these factors are in the creation of an accurate prediction model. This report also tries to demonstrate the value of shallow machine learning techniques over deep learning techniques on smaller datasets. This paper uses several machine learning algorithms to diagnose the severity of keratoconus using real-world medical data, such as corneal topography, elevation, and pachymetry parameters obtained from pentacam instruments at Sydney's Vision Eye Institute Chatswood. 8 different machine learning techniques were implemented in python and achieved an average accuracy of 68%, 80%, and 90% for the risk factor, pentacam and cumulative data models respectively. The research shows a 12.5% increase in accuracy and a 5% increase in AUC upon addition of risk factor data. This study can assist clinics in deciding which types of data are of most importance when assessing patients and making correct diagnoses. It also has the ability to alert those without a diagnosis on the urgency of an evaluation.

Table of Contents

1 Introduction

1.1 Introduction to the study

We have observed an increase in the use of machine learning and artificial intelligence for the diagnosis and monitoring of diseases in recent years. It has established itself as an indispensable resource for detecting and assessing trends in medicine and research.

Keratoconus is one such condition within the ophthalmic industry that ought to be examined using Machine Learning approaches. Keratoconus is a non-inflammatory corneal disease that affects both eyes and is present in 1 out of every 2000 patients globally. It is the deformation of the cornea into a conical shape, in which it thins, causing high order aberrations and progressive vision loss. (Hallett, N. *et al.* 2020).

The disease frequently manifests itself from adolescence and progresses to a state of stabilization by the time the patient reaches his or her forties, resulting in a significant reduction in quality of life. The most often used diagnostic procedures for keratoconus are corneal topography and corneal tomography. This, in conjunction with clinical evaluation parameters and a systematic classification approach, such as the Amsler-Krumeich classification, has been standard practice in the diagnosis of keratoconus and other corneal illnesses over the past several decades (Al-Timemy, A.H. *et al.* 2021).

The Placido ring and Scheimpflug imaging are utilized to acquire the topographic results. Aside from that, optical coherence tomography (OCT) can be used to give additional quantitative and qualitative information that can be used in the diagnosis process. When combined with other tomographic information, such as corneal thickness, this information can be used to classify the patient's stage of degeneration (Hallett, N. *et al.* 2020).

Although there is currently no consensus on the exact cause of keratoconus, there are several risk factors, both environmental and genetic, that have been identified as having high importance in the development of keratoconus in individuals.

Gender, race, general health, atopy hypertension, hay fever, known eye history, eye rubbing, family history of keratoconus, and primary optical aid are just a few of the risk factors to consider (Hallett, N. *et al.* 2020). It is possible that these factors will provide more insight into what is most important in recognizing and diagnosing not only keratoconus but also other corneal disorders in the future.

1.2 Background of the problem

When compared to other ocular disorders, there are far fewer machine learning experiments conducted on the cornea (Cao, K. *et al.* 2022). The majority of the research available employs deep learning models, specifically convolutional neural networks, and applies them to corneal topography and tomography maps with elevation, curvature, and thickness metrics in order to detect the presence and severity of keratoconus in patients.

The current research typically comprises of image datasets with fewer than 400 images in spite of the fact that deep learning models require a robust dataset in order to create a generalizable model (Cao, K. *et al.* 2022). This shows the benefit of using predictive algorithms that do not require too large a dataset to find a relationship between its input data and its output. Another thing considered is the computational cost of training and optimizing accurate deep neural networks.

This project uses a range of less computationally intensive models to classify the data so as to produce a robust set of results, and increase the like likelihood of incorporation into the clinical sector.

1.3 Purpose of the study

The goal of this project is to use machine learning models to gain a better understanding of how much the risk factors of a patient contribute towards the progressive stages of keratoconus as well as how significant these factors are in the creation of an accurate prediction model.

This report also aims to show the utility of shallow machine learning models over deep learning models on smaller datasets

The predictive models used in this paper are verified with public medical data collected from Vision Eye Institute Chatswood, therefore these algorithms should benefit ophthalmologists by allowing them to observe and analyze specific corneal patterns that are not readily visible to humans.

The main contribution of this paper is to use machine learning to illustrate the most important factors playing a part in diagnosing the stage of degradation in keratoconus.

The risk factors that are looked at in this project are Gender, Race, General Health, Atopy, Hypertension, Hay fever, Known Eye History, Eye Rubbing, Family History of keratoconus, and Primary Optical Aid.

I believe this study can help guide clinics on what type of data would be of most importance when assessing a patient and giving an accurate diagnosis. it also has the potential to highlight to individuals without a diagnosis if there exists a need to go for an assessment.

1.4 Organization of the study

The paper is organized as follows.

In Section 2, The related work on machine learning and deep learning methods for keratoconus diagnosis, as well as review the literature on the risk factors involved and how much correlation they have to the development of keratoconus will be reviewed, and some light will be shed on the classification methods used to assess the stage of degradation in a patient.

Section 3 describes the public patient data set collected in a private ophthalmic clinic (Vision Eye Institute Chatswood) by Hallett, N. *et al*. (2020), and how it was manipulated for this study's use case. I also look at the limitations and delimitations of the dataset.

In Section 4, The machine learning algorithms that are used for the classification problem will be described and the architecture of the models used will be illustrated in this section.

In Section 5, the results and findings of the report will be detailed and in Section 6 these results will be analyzed and discussed to give a better picture of the findings of the study.

Lastly, in Section 7, I will present a set of concluding statements and give recommendations based on the findings of the report.

2 Literature Review

This section will look at the impact machine learning has made on keratoconus diagnosis as well as what the goal of diagnosis is in each paper. Some comparisons will be made about the classification schemes used in current practice. Finally, the literature on risk factors and how they relate to the prevalence of keratoconus will be explored

2.1 Related work on Machine Learning

Over the last few decades, machine learning has become increasingly popular in Keratoconus, mostly for the detection of Keratoconus and early Keratoconus. Machine learning has the effect of providing reliable and unbiased diagnosis, which is important when diagnosing patients early on since early intervention with treatments like corneal crosslinking can slow down the degradation of the cornea, avoiding the need for a corneal transplant altogether (Yousefi, S. *et al.* 2018).

Zéboulon, P. *et al.* (2020) uses a large dataset with 3000 maps and achieved a highly accurate predictive model in order to classify patients, but the model's capability to predict different stages of progression of the disease has not yet been examined.

There are no successful instances of machine learning models in medical research at this time (Yousefi, S. *et al.* 2018). This could be attributed to a lack of large patient populations to confirm results, the use of diverse imaging technologies, a local participant group made up of people of various ethnicities, medical professionals' general acceptability of predictive modeling for detection of keratoconus, and their relative consistency to clinicians.

To distinguish clinically obvious keratoconus from the normal cornea, Chastang, P.J. *et al.* (2000) developed a binary decision trees technique based on corneal topography indices. Smolek, M.K. and Klyce, S.D. (1997) suggested a neural network based on corneal topography indices for keratoconus diagnostics. A similar approach was used to distinguish keratoconus from normal corneas by modeling the corneal surface with a seventh-order Zernike polynomial (Twa, M.D. *et al.* 2005). All of these approaches relied solely on the cornea's anterior topography. There is a shortage of data on the impact of merging data from multiple devices on machine learning models used to identify Keratoconus.

Data on posterior corneal curvature and pachymetry were obtained and utilized to evaluate corneal features as technology advanced (Ambrósio, R., Jr *et al.* 2006). In clinical and subclinical keratoconus, Pinero et al. measured corneal volume, pachymetry, and the relationship between the anterior and posterior shape of the cornea (Piñero, D.P. *et al.* 2010).

Fernández Pérez, J., Valero Marcos, A. and Martínez Peña, F.J. (2014) show that using corneal equipment and techniques such as Videokeratography, Orbscan, and Pentacam in combination with risk factors can lead to detection of subclinical keratoconus, but at the cost of a higher false-positive detection rate.

Other types of information, such as the corneal epithelial thickness map obtained by optical coherence tomography (OCT) are rapidly becoming acknowledged as important in the diagnosis of keratoconus (Silverman, R.H. *et al.* 2017), particularly early keratoconus (Venkateswaran, N. *et al.* 2018)

This illustrates that combining data from various devices and taking into account a wider range of parameters may help to enhance the early diagnosis of keratoconus.

Machine learning models that used data from the Pentacam had greater collective sensitivity and specificity in detecting clinical keratoconus and subclinical keratoconus from control eyes than models that used data from other imaging devices, according to studies. This is possibly due to the Pentacam's ability to produce a larger set of data than other devices (Chen, D. and Lam, A.K.C. 2009).

2.2 Classification Methods for Staging

Along with detecting KC eyes as a distinct category, some studies grouped their KC eyes into clinical stages and used machine learning algorithms to identify each stage separately.

The research shows a range of criteria to categorize keratoconus eyes into different groups. Kamiya et al. (2019) used the Grades 1–4 classification. according to the Amsler–Krumeich classification scheme, which is largely reliant on keratometry but often integrates refraction and pachymetry metrics (Alió, J.L. *et al.,* 2015)

Using a separate categorization system called RETICS, Blazquez et al. (2020) and Bolarin et al. (2020) categorized eyes into Grade I–IV plus. It is largely reliant on corrected distance visual acuity (Alió, J.L. *et al.,* 2015).

The Amsler–Krumeich classification can be used to detect keratoconus and track its progression. The Sandali classification scheme and this classification method are the two most widely utilized models in the corneal industry (Macsai, M.S., Varley, G.A. and Krachmer, J.H. 1990). There are other grades and methods used in current research for delineation of severity but to be consistent with the data samples collected, this paper uses Grades 1–4 according to the Amsler–Krumeich (AK) grading scheme.

In patients with keratoconus, the Amsler–Krumeich and Sandali classifications are the most widely used systems; nevertheless, they are not completely interchangeable since they identify different changes. The Amsler–Krumeich classification, in particular, is better at identifying patients in the early stages of the disease and tracking their progress over time. In contrast, the Sandali classification, which is based on the analysis of AS-OCT pictures, is effective for the diagnosis and follow-up of patients with advanced stages of KC, especially when a surgical plan needs to be selected.

This illustrates the utility of the Amsler–Krumeich model as a better classification model for machine learning algorithms, which are also reliable in early-stage diagnosis (Macsai, M.S., Varley, G.A. and Krachmer, J.H., 1990).

2.3 Risk Factors

Keratoconus diagnosis may also be aided by demographic data such as age and gender, as well as other potential risk factors for keratoconus such as eye rubbing and family history (Yousefi, S. *et al.,* 2018).

Most of the current research such as Sharif, R. *et al.,* (2018) is unsure exactly how these critical factors might affect the diagnostic accuracy in the machine learning models that are currently in use. Because there are few recognized risk factors for keratoconus. It should be reasonable to start examining potential risk factors and incorporating them into future machine learning models to determine their usefulness in diagnosis.

Gender, Race, General Health, Atopy, Hypertension, Hayfever, Known Eye History, Eye Rubbing, Family History of Keratoconus (Genetics), and Primary Optical Aid are the risk factors investigated in this study.

2.3.1 Gender

There is no conclusive evidence that the patient's sex influences the progression of keratoconus.

Differences between the sexes in keratoconus have been studied in several ways. The results differ across studies, with some demonstrating female prevalence, others demonstrating higher male prevalence, and still, others demonstrating no significant gender prevalence (Alió, J.L., 2017).

Certain studies, such as Amsler's (1961) with 65 percent, Laqua (1971) with 57 percent, Hammerstein (1972) with 66 percent, and Jonas (2009) with 53 percent reveal a female prevalence. Contrary to this, other studies found a prevalence of male cases. This is the case for Pouliquen (1981) with 57 percent, Street (1991) with 62 percent, Pobelle-frasson (2002) with 62 percent Owen (2003) with 59 percent and Fatima with 53 percent, and Ertan (2009) with 62 percent, while other studies did not find any significant differences between the sexes.

Research exists that illustrates a difference in the prevalence of hereditary family keratoconus (17% female versus 11% male) in another investigation (Fink, B.A. *et al.,* 2005).

2.3.2 Race

Keratoconus is a worldwide condition with varying levels of frequency among ethnic communities in the same geographical region. A large number of studies have shown that ethnicity and environmental conditions are likely to have a significant role in the disease's development.

Until recently, it was thought that keratoconus afflicted people of all ethnicities impartially (Rabinowitz, Y.S. 1998). Various studies, however, have shown that there is a disparity in prevalence based on ethnicity. For example, Lebanon has a high frequency of 3.3 percent, India's rural parts have a prevalence of 2.3 percent, while the north of Denmark, Finland, and Russia have a lower rate (Alió, J.L., 2017).

In another investigation, Druze and Bedouins were found to have a higher predominance than Muslims and Christians. However, the changes were not statistically meaningful due to the small sample size used in the study (Pan, C.-W. *et al.* 2014). Keratoconus was found to be substantially more common among Indians in Singapore than among Malays or Chinese (Shneor, E. *et al.,* 2014). Another study revealed a 3.18 percent prevalence of Keratoconus among Young Arab Students in Israel (Pan, C.-W. *et al.,* 2014).

2.3.3 General Health

2.3.3.1 Ehlers-Danlos Syndrome (also known as EDS).

Kumming (1977) asserts that because EDS is a collagen abnormality of the body, it has the potential to disrupt corneal collagen and, as a result, have an impact on the progression of keratoconus. Keratoconus is quite infrequent in these patients. A total of 72 patients with Ehler -Danlos syndrome were studied by McDermoitt (1998), and only one patient manifested keratoconus.

2.3.3.2 Marfan Syndrome

Because Marfan syndrome is an X-linked condition, there is a higher prevalence of keratoconus in these patients than in the general population. However, the association of Marfan syndrome with keratoconus is still rather infrequent and does not account for a significant proportion of cases (Maumenee, I.H., 1981)

2.3.3.3 Lonstein's Disease

Osteogenic imperfecta, also known as Lobstein disease, is a hereditary bone disorder that affects bone collagen. A higher incidence of keratoconus has been reported in some Lobstein disease families as a result of the disease (Beckh, U., Schönherr, U. and Naumann, G.O. 1995).

2.3.3.4 Prolapse of the Mitral Valve

In this condition, the mitral valve of the heart bulges outward. The exact cause is unknown, but it is generally assumed that there is an abnormality in the collagen fibers of the mitral valve and that this abnormal collagen fiber may be present in other fibrous tissue in the human body, such as corneal collagen.

According to some literature, mitral valve prolapse is more largely linked with keratoconus (Sharif, K.W., Casey, T.A. and Coltart, J. 1992). While other literature, such as Street (1991), has found no link between keratoconus and mitral valve prolapse, this is not the case for Street. Sharif (2013) investigated the prevalence of mitral valve prolapse in 50 keratoconus patients, finding that it was much greater at 53 percent than in the control group at 7 percent. These findings are in contrast to those published by Street (1991).

Frasson (2003) presented a study in which a cardiac echography was performed on 50 keratoconus patients. The findings were as follows: Mitral valve anomalies were found in 22 percent of patients (prolapse was found in 4 percent, mitral insufficiency was found in 10 percent, and mitral regurgitation was found in 8 percent), which is four times greater than the general population. Within this group of patients, the real prolapses that were discovered were in patients who were in an advanced stage of the disease (Barbara, A., 2012).

A possible explanation for the relationship between keratoconus and mitral valve prolapse is the presence of some common collagen types between the two illnesses (Alió, J.L., 2017), the most notable of which is collagen type 1, which is found in both the corneal stroma and the anatomy of the mitral valve.

2.3.3.5 Down Syndrome

Cullen (1963) found that 5.5 percent of the 143 Down syndrome patients he studied had keratoconus, and other writers, including Pierse, Slusher, and Kenyon, have found that patients who have experienced ocular traumatism more frequently have a greater frequency of acute keratoconus. When compared to the general population, the rate of postoperative problems after transplantation (stitch breakage, traumatism, rejection, etc.) is higher (Barbara, A., 2012).

Keratoconus prevalence in patients with Down syndrome appears to differ according to the ethnic group analyzed, according to the results of this study. A study conducted in Malaysia that included 60 Down syndrome patients discovered a higher frequency of various eye disorders, such as epicanthus, and nystagmus in the Down syndrome population. Despite this, no incidences of keratoconus were discovered (Liza-Sharmini, A.T., Azlan, Z.N. and Zilfalil, B.A., 2006). A study conducted by Shapiro, M.B. and France, T.D. (1985). indicated that the prevalence of keratoconus was 15 percent in a group of 53 Down syndrome patients who were of Caucasian descent.

2.3.3.6 Diabetes

The relationship between keratoconus and diabetes has been investigated, however, the prevalence of diabetes in people with keratoconus is lower than in the overall population. Those suffering from diabetes, on the other hand, experience a less severe version (Kuo, I.C. *et al.,* 2006).

It has been proposed that biochemical property alterations of the cornea in diabetics are caused by abnormal glycosylation of stromal collagen fibers in the cornea (Barbara, A., 2012). The activation of auto cross-linking in the corneal stroma, which prevents biomechanical weakening of the cornea, could explain the beneficial impact of diabetes in patients with keratoconus.

2.3.4 Hay fever

When the prevalence of asthma, hay fever, and eczema was compared between the control group and the keratoconus group, the latter group showed a statistically greater frequency of hay fever (P=0.007) and asthma (P=0.0002). The number of times people wipe their eyes increased among those who had asthma, hay fever, or eczema keratoconus disease. TMS-2, on the other hand, showed no significant connection with either of the atopic demographics when measured against corneal power. TMS-2, on the other hand, showed no significant connection with either of the atopic demographics when measured against corneal power (Weed, K.H. *et al.* 2008).

2.3.5 Hypertension

The findings of Naderan, M. *et al.* (2015), imply that keratoconus may be associated with other disorders often including vernal keratoconjunctivitis and hypertension.

2.3.6 Known eye History

2.3.6.1 Floppy eyelid syndrome

A possible link between keratoconus and this condition has been suggested by some studies (Negris, R., 1992), although there is currently no conclusive evidence to support this, and further research or data may be required. A possible link between keratoconus and this condition has been suggested by some studies (Negris, R., 1992), although there is currently no conclusive evidence to support this, and further research or data may be required.

2.3.6.2 Leber Congenital Amaurosis

It is a congenital illness that causes poor vision from infancy onward. In comparison to other blinding diseases, this condition appears to be more related to keratoconus (Galin, M.A. and Berger, R. 1958). Although a genetic component could be the source of this condition, repeated ocular rubbing is unquestionably a contributing element. (Barbara, A., 2012).

2.3.7 Corneal Dystrophies

Corneal Dystrophies are a type of eye disease that affects the cornea. Keratoconus is only seldom related with corneal dystrophies, according to the literature (Barbara, A., 2012).

2.3.7.1 Macular Dystrophy

Several publications have reported an association between keratoconus and macular dystrophy. Two examples of macular dystrophy linked with keratoconus were documented by (Javadi, 2007).

2.3.7.2 Granular Dystrophy

This disorder is linked with keratoconus, in which case a lamellar transplant is preferentially performed for keratoconus in the event of dystrophy worsening (Barbara, A., 2012).

2.3.7.3 Posterior Polymorphous Corneal Dystrophy

This is a rare autosomal dominant bilateral decrement membrane dystrophy that occurs in only a few people. In addition, certain cases of keratoconus linked with PPCD have been recorded in the literature (Barbara, A., 2012).

2.3.8 Eye Rubbing

Eye rubbing is also a primary environmental stressor that contributes to the development of keratoconus.

It is believed to produce corneal damage by one of the following mechanisms: Repeated epithelial trauma results in the release of matrix metalloproteinase 1 and 13, interleukin-1, and tumor necrosis factor-alpha, which cause stromal remodeling and keratocyte death (Alió, J.L., 2017). This, in turn, may cause epigenetic modification and affect genetically predisposed individuals by altering the expression of genes associated with the development of keratoconus in those individuals.

In addition to increased intraocular pressure (IOP), which has been linked to the advancement of keratoconus in atopy patients, particularly in young individuals (McMonnies, C.W. and Boneham, G.C., 2010), there is an increase in IOP in atopy patients.

The duration of eye rubbing in individuals with keratoconus appears to be greater than that associated with atopy Krachmer, J.H. (2004), which may explain why such a high percentage of atopic patients do not acquire keratectasia as a result of rubbing. It is critical to note how the patient rubs the eye, whether softly or violently as well as the duration of the rubbing. For example, the normal duration of rubbing in keratoconus patients is substantially longer within a range of 10 to 180 seconds, but it is often less than 15 seconds in allergic or infectious disorders and less than 5 seconds in healthy individuals without any eye illness (Weed, K.H. *et al.,* 2008).

A study by Kennedy, R.H., Bourne, W.M. and Dyer, J.A. (1986) demonstrated that keratoconus patients rub their eyes more frequently than normal participants and the study revealed that 48 percent of keratoconus patients stroked both eyes forcefully while only 2.2 percent rubbed one eye strongly.

In a recent study by Balasubramanian, S.A., Pye, D.C. and Willcox, M.D.P. (2013), the researchers established strong empirical evidence of a link between keratoconus and the habit of scratching one's eyes. A group of healthy volunteers who were not infected with the disease and who did not wear contact lenses were asked to massage their eyes in a controlled manner for 60 seconds each. In a subsequent study, tears were collected before and after eye rubbing and it was discovered that the levels of MMP-13, IL-6, and TNF- were dramatically elevated after rubbing the eyes. Several studies have established that chronic eye rubbing, which is highly prevalent in keratoconus patients, may contribute to the course of the illness by maintaining continually elevated levels of these proteases, inflammatory mediators, and protease activity.

2.3.9 Family History (Genetics)

It has not been reported that genetic effects in the pathogenesis of keratoconus have been studied in detail using current methodologies, but a review of the published literature gives strong evidence that genetic influences are involved in the pathogenesis of this illness. Among these are at least eight reports

of its occurrence in both identical twins, and multiple reports of its occurrence in family members over two and three generations. The literature contains nine occurrences of keratoconus in monozygous twins; in all but one of these cases, both twins had keratoconus. Videokeratography was not conducted on the twin who did not have keratoconus. Genetic factors in keratoconus are well supported by research observations; however, a formal prospective twin study comparing monozygotic with dizygotic twins without ascertainment bias is required to corroborate the conclusions gained from these observations. According to the majority of documented research, the illness is inherited in an autosomal dominant manner with varying expression and includes modest variants of the disorder such as keratoconus fruste or mild irregular astigmatism, among others. The literature has several accounts suggesting recessive inheritance, but none provide conclusive proof that three generations of the disorder were evaluated or that subtle variants of the disorder were sought to be considered for inclusion in the pedigree study. Formal genetic analyses are required to precisely characterize hereditary patterns for the many subtypes of keratoconus and to understand the role that genetic effects may have in the disease's etiology, even though the majority of research shows a dominant mode of inheritance. (Rabinowitz, Y.S., 1998).

2.3.10 Primary Optical Aid

A total of 199 patients were evaluated; 53 (27 percent) had a history of contact lens wear before the diagnosis of keratoconus, 68 (34 percent) had no history of contact lens wear prior to or after the diagnosis of keratoconus, and 78 (39 percent) had an existing history of keratoconus prior to wearing contact lenses (Macsai, M.S., Varley, G.A. and Krachmer, J.H., 1990).

When it came to contact lens wear before being diagnosed with keratoconus, the 53 patients had an average of 12.2 years (with a range of 5.5 to 22 years) of experience, averaging 15.2 hours each day.

Although the study recognizes that it is conceivable that keratoconus would still have progressed regardless of contact lens use, it also acknowledges that exceedingly mild symptoms of keratoconus may have been overlooked on the first test.

It is considered that the wearing of contact lenses causes trauma to the cornea that, in most circumstances, is not severe enough to cause keratoconus to develop. The patient is not prone to developing keratoconus after a lengthy period of time wearing the lenses but the wearing of contact lenses may pose the possibility of elevated degradation of keratoconus in patients.

3 Dataset

Between 2014 and 2017, data on 124 keratoconus patients were gathered at the Vision Eye Institute Chatswood in Sydney, Australia. Each patient's medical data were reviewed and analyzed, with 79 percent (63.7 percent) of those who participated being male. Table I and Table II contain information that has been extracted to be used as variables of consideration.

Based on clinical expertise, all patients were classified using the Amsler-Krumeich (A-K) classification, which includes mean-K values on the anterior curvature sagittal map, thickness at the thinnest location, and the patient's refractive error, as well as bio-microscopy. Table III illustrates how patients are classified using the A-K categorization system. (Hallett, N. *et al.* 2020).

Variable	Source
Gender	Patient
Race	Patient
Diabetes	Patient
Atopy	Patient
Allergy	Patient
Hypertension	Patient
Other disease presence	Patient
Known eye history	Patient
Family history	Patient
Eye rubbing	Patient
Primary optical aid	Patient

Table 1 Risk Factor Data (Hallett, N. *et al.* 2020).

Variable	Source
Refractive sphere	Instrument (Pentacam)
Refractive cylinder	Instrument (Pentacam)
Refractive axis	Instrument (Pentacam)
Flat keratometry	Instrument (Pentacam)
Steep keratometry	Instrument (Pentacam)
Thinnest pachymetry	Instrument (Pentacam)
Location X-axis	Instrument (Pentacam)
Location y-axis	Instrument (Pentacam)
Central pachymetry	Instrument (Pentacam)
Amsler-Krumeich (AK) classification	Clinician

Table 2 Pentacam Data as well as Clinician classification (Hallett, N. *et al.* 2020).

3.1 Data Classification Scheme

Grade	Characteristic
1	Eccentric steeping Myopia and astigmatism < 5.00D Mean central K readings < 48.00D
2	Myopia and astigmatism 5.00-8.00D Mean central K readings < 53.00D Absence of scarring Minimum corneal thickness > 400µm
3	Myopia and astigmatism 8.00-10.00D Mean central K readings > 53.00D Absence of scarring Minimum corneal thickness 300-400µm
4	Refraction not measurable Mean central K readings > 55.00D Central corneal scarring Minimum corneal thickness 200µm

Table 3 Amsler–Krumeich classification scheme (Hallett, N. *et al.* 2020).

3.2 Dataset Manipulation and Delimitation

The collected dataset contains metrics that are collected according to the assessment of the clinician, for this paper these data points are not advisable and were removed in the analysis as they do not fall in line with the goal of using data readily available from the patient and the pentacam (or any other appropriate topography device).

3.3 Data Analysis

Using python, quantitative analysis Is carried out on the dataset to assess the correlation between the various attributes within the dataset. These correlations will be separated according to the source of the data so as to provide useful analysis on the significance of each attribute.

The dataset will be analyzed as three different sections which are the Risk Factor data readily available from the patient, the Pentacam data readily available from the device, and dataset comprising of both parts so as to see the relative effect of the dataset sections on their own.

3.4 Issues of Trustworthiness

It is important to note that this dataset was not collected by myself so the validity of the data as well as the diligence in collecting it are assumed to be up to appropriate standards for the sake of completing this particular research project.

3.5 Data Limitation

Within this field of research, the sample size of data for applications of this nature is usually much larger so as to make the predictive model more generalizable. However, this project works on a relatively small dataset with the assumption that the machine learning techniques will offset this drawback.

4 Methodology

In this section, we will take a deeper look at the algorithms used to create the predictive models.

4.1 Multinomial Logistic Regression

The multinomial logistic regression (MLR) model is a relatively basic modification of the binary model, and both models rely mostly on logit analysis or logistic regression to achieve their results, respectively. Logit analysis is, in many respects, the natural complement to simple linear regression when the response is a categorical variable, and it is particularly useful when the response is a categorical variable. In the case of discrete variables that exist among the explanatory factors, they are handled with through the use of one or more (0,1) dummy variables; however, when the response variable is of this type, the regression model fails completely. Logit analysis is a readily available option.

For a response variable Y with two measurement levels (dichotomous) and an explanatory variable X, consider the following:

let: $\pi(x) = p(Y = 1 \mid X = x) = 1 - p(Y = 0 \mid X = x)$, the logistic regression model has linear form for logit of this probability

$$\text{Logit}\,[\pi(x)] = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x, \text{ where the odds } = \frac{\pi(x)}{1-\pi(x)},$$

The odds $= \exp(\alpha + \beta x)$, and the logarithm of the odds is called logit, so $\text{Logit}\,[\pi(x)] = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \log[\exp(\alpha + \beta x)] = \alpha + \beta x$

The logit has linear approximation relationship, and logit = logarithm of the odds. The parameter β is determined by the rate of increase or decrease of the Sshaped curve of $\pi(x)$. The sign of β indicates whether curve ascends ($\beta > 0$) or descends ($\beta < 0$), and the rate of change increases as $|\beta|$ increases (El -Habil, 2012).

4.2 Nearest Neighbor Classification

A set of n pairs $(x_1, \theta_2), \cdots, (x_n, \theta_n)$ is given, where the x_i 's take values in a metric space X upon which is defined a metric d, and the θ_i 's take values in the set $\{1,2, \cdots, M\}$. Each θ_i is considered to be the index of the category to which the i th individual belongs, and each x_i is the outcome of the set of measurements made upon that individual. For brevity, we shall frequently say " x_i belongs to θ_i " when we mean precisely that the i th individual, upon which measurements x_i have been observed, belongs to category θ_i.

A new pair (x, θ) is given, where only the measurement x is observable by the statistieian, and it is desired to estimate θ by utilizing the information contained in the set of correctly classified points. We shall call

$$x'_{\approx}\varepsilon\{x_1, x_2, \cdots, x_n\}$$

a nearest neighbor to x if

$$\min\ d(x_i, x) = d(x'_n, x)\ i = 1,2, \cdots, n.$$

The nearest neighbor rule decides x belongs to the category θ_n' of its nearest $'$ neighbor x_n'. A mistake is made if $\theta_n' \neq \theta_{\cdot}$. Notice that the NN rule utilizes only the classification of the nearest neighbor. The $n-1$ remaining classifications θ_i are ignored (Cover, T. and Hart, P., 1967).

4.3 Support Vector Machine for classification

In a high-dimensional feature space, an SVC model is like this:

$$\min_{\omega,b,\xi} J(\boldsymbol{\omega}, b, \xi) = \frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} + C\sum_{i=1}^{n} \xi_i$$

$$\begin{aligned} &s.t. \\ &y_i[\boldsymbol{\omega}^T \boldsymbol{\Phi}(\mathbf{x}_i) + b] \geq 1 - \xi_i, \quad i = 1, \cdots, n \\ &\xi_i \geq 0, \qquad\qquad\qquad\quad i = 1, \cdots, n \end{aligned}$$

A support vector machine won't solve the above model directly. Instead, kernels $\mathbf{K}(\mathbf{x}, \mathbf{x}') = \cdot \left(\Phi(\mathbf{x}) \cdot \Phi(\mathbf{x}')\right)$ and Lagrange multipliers $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)^T$ are introduced to build the dual model (3).

$$\begin{aligned} &\max_{a} \mathrm{W}\,(\boldsymbol{\alpha}) = \sum_{j=1}^{n} \alpha_j - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j\, \mathrm{K}(\mathbf{x}_i, \mathbf{x}_j) \\ &\text{s.t.} \sum_{i=1}^{n} \alpha_i y_i = 0 \\ &0 \leq \alpha_i \leq C \ \text{ for } \ i = 1, \cdots, n \end{aligned}$$

Suppose $\boldsymbol{\alpha}^0$ is the best one, among which $\alpha_i^0 \neq 0$ corresponds to support vector $\mathbf{x}_i$. Then the SVC classifier is

$$y = \mathrm{sgn}\,(f(\mathbf{x})) = \mathrm{sgn}\left(\sum_{i=1}^{n} \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b\right)$$

(Chen, Z.-J., Liu, B. and He, X.-P. 2007).

4.4 Decision Tree

A decision tree is a hierarchical tree structure resembling a flowchart that is formed of three fundamental elements: decision nodes that correspond to attributes, edges or branches that correspond to the various possible attribute values. The third component is leaves, which often consist of objects that are members of the same class or are extremely similar. This format enables us to generate decision rules for classifying new instances. Indeed, each path between the root and a leaf corresponds to a conjunction of test qualities, and the tree is regarded as a disjunction of these conjunctions.

To categorize a new instance that contains only values for all of its attributes, we begin at the root of the created tree and follow the path that corresponds to the attribute's observed value in the tree's interior node. This procedure is repeated until the presence of a leaf is detected. Lastly, we use the resulting label to determine the instance's anticipated class value. (Jenhani, I., Amor, N.B. and Elouedi, Z. 2008)

4.5 Random Forest

Breiman, L. (2001) invented the Random Forest algorithm for ensemble learning. An approach called ensemble learning generates a large number of individual learners and aggregates the results. Random Forest is a variation on the bagging technique. Each classifier in Bagging is constructed independently using a bootstrap sample of the input data. In a conventional decision tree classifier, a decision is made at each node split based on all feature properties.

However, in Random Forest, the optimal parameter at each node in a decision tree is chosen randomly from a set of features. This random selection of features not only enables Random Forest to scale effectively when there are a large number of features per feature vector, but also reduces the interdependence (correlation) between feature properties making it less susceptible to intrinsic noise in the data (Criminisi, A., Shotton, J. and Konukoglu, E. 2012).

As Breiman, L. (2001) mentions, the amount of random features m selected per decision node in a tree determines the forest classification's error rate. The Random Forest classifier's error rate is proportional to the correlation between any two trees and the classification strength of each individual tree. Reducing the number of randomly picked characteristics m decreases both the correlation between classification trees and the classification strength of each particular tree. Grows both the correlation between the trees and the strength of each tree as m increases.

4.6 Multilayer Perceptron

An MLP of depth K is a multivariate function in the form of

$$x_i^{(j+1)} = g_{ij}\left(x^{(j)} \cdot w_{ij} + b_{ij}\right),$$

for $j = 1,2, \dots, K$ and $j = 1,2, \dots, N_j$, where the $x^{(j)} = \left(x_1^{(j)}, \dots, x_{N_j}^{(j)}\right)$ is the input of j th layer with N_j neurons, and g_{ij} is the activation function on $\mathbb{R}$ which connects the input and the i th neuron in the j th layer, and w_{ij} and b_{ij} are the trainable weights and biases. For classification task, the last MLP is trained by back-propagation using a stochastic gradient descent optimization strategy. (Hallett, N. *et al.* 2020).

The MLP model used in this paper has 100 layers which is technically a deep learning model but its relatively shallow compared to most deep learning algorithms.

4.7 AdaBoost Decision Tree

AdaBoost's central idea is to fit a sequence of weak learners (i.e., models that are just marginally better than random guessing, such as tiny decision trees) to repeatedly updated versions of the data. To get the final prediction, the predictions from all of them are combined using a weighted majority vote (or sum). Each so-called boosting iteration modifies the data by assigning weights to each of the training samples. Initially, all of those weights are set to zero, which means that the first step merely trains a weak learner on the original data. Each subsequent iteration modifies the sample weights separately and reapplies the learning algorithm to the reweighted data. At a given step, the weights of those training examples that were mistakenly predicted by the boosted model induced in the previous step are increased, while the weights of those that were properly predicted are dropped. As iterations progress, difficult-to-predict cases gain more influence. Each future poor learner is thus compelled to concentrate on the examples that were overlooked by the preceding weak learners. (Freund, Y. and Schapire, R.E., 1997).

4.8 Quadratic Discriminant Analysis

Let us assign a test object $\boldsymbol{x}_i$ based on d measurements $\boldsymbol{x}_i = (x_{i1}, \dots, x_{id})^{\mathrm{T}}$ to one of K a priori defined classes. The classification score is defined as

$$\mathrm{cf}\,(x_i) = (x_i - \mu_{\mathrm{k}})^{\mathrm{T}} \Sigma_{\mathrm{k}}^{-1} (x_i - \mu_{\mathrm{k}}) + \ln |\Sigma_{\mathrm{k}}| - 2\ln\, \pi_{\mathrm{k}}$$

where Σ_{k} is the class covariance matrix of class k, μ_{k} is the mean vector of class k, and π_{k} is the prior probability of class k. They are estimated by

$$\hat{\mu}_{\mathrm{k}} = 1/n_{\mathrm{k}} \sum_{i=1}^{n_{\mathrm{k}}} \boldsymbol{x}_i$$

$$\hat{\Sigma}_{\mathrm{k}} = 1/n_{\mathrm{k}} \sum_{i=1}^{n_{\mathrm{k}}} (\boldsymbol{x}_i - \mu_{\mathrm{k}})(\boldsymbol{x}_i - \mu_{\mathrm{k}})^{\mathrm{T}}$$

$$\hat{\pi}_{\mathrm{k}} = n_{\mathrm{k}}/n$$

where n_{k} is the number of objects in class k, and n is the total number of objects in the training set. Object x_i is assigned to the class for which it has the lowest classification score. Based on this quadratic classification rule the boundary of classes is of quadratic shape. (Wu, W. *et al.,* 1996)

5 Results

In this chapter the results of the quantitative analysis as well as those produced by the machine learning model are presented in the form of charts and tables.

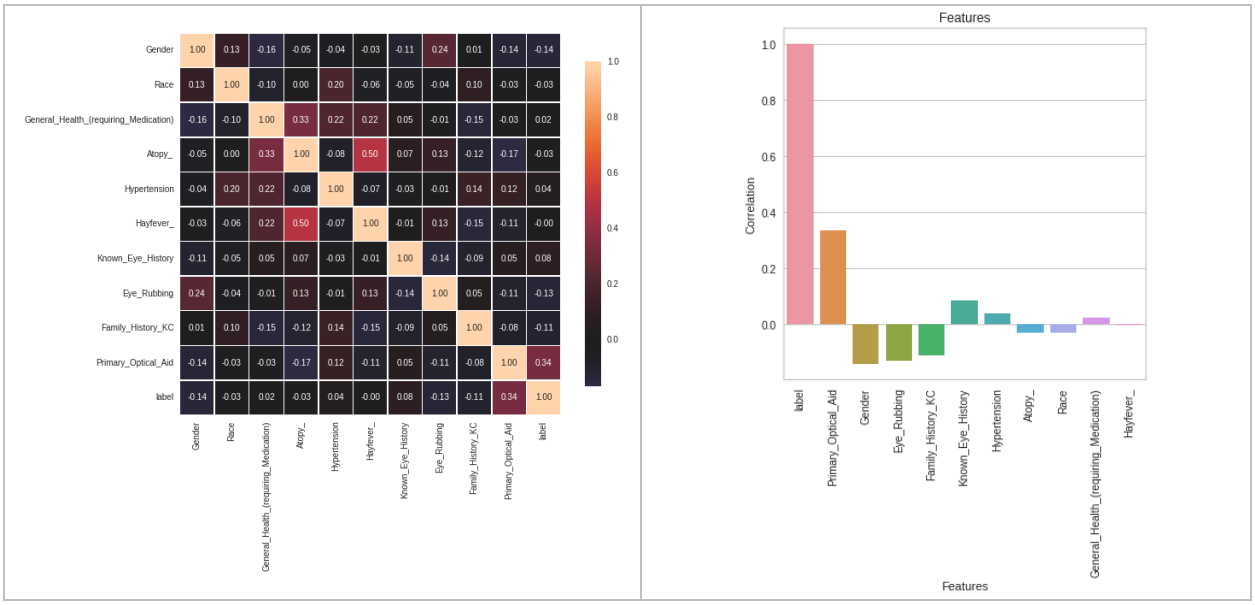


Figure I Risk Factor Correlation Data (Heatmap Left, Top features bar chart Right)

5.1 Risk Factor Correlation Data

The correlation heatmap shown in Figure I highlights how strongly all the risk factors correlate to the resulting class of the diagnosis as well as to each other feature.

There are some notable data points within the heatmap with a correlation of 0.2 and above

- The high correlation of 0.5 between atopy and hay fever
- The correlation between atopy, hypertension, hay fever, and family history of keratoconus
- Hypertension has a notable link to Race with a value of 0.2
- There is also a notable correlation between Gender and Eye rubbing (0.24)

5.2 Most Significant Risk Factor attributes

The most significant contributor to the class assessment process according to the data is the primary optical aid of the patient, this is followed by the gender of the patient and runner up to this is eye rubbing. The only truly notable contributor according to the data is what the primary optical aid of the patient is.

The average correlation of the attributes to the assessment is 0.092 which can be seen as insignificant

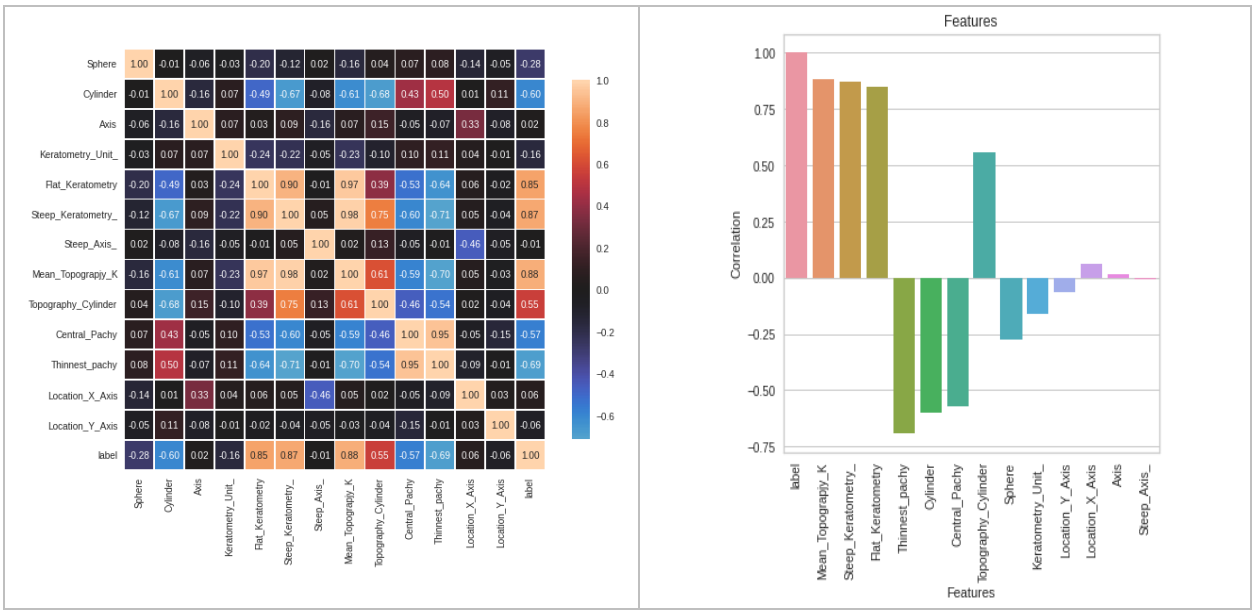


Figure II Pentacam Correlation Data (Heatmap Left, Top features bar chart Right)

5.3 Pentacam Correlation Data

The correlation heatmap shown in Figure II highlights how strongly all the pentacam features correlate to the resulting class of the diagnosis as well as to each other feature.

We can see from the data that there is a higher overall correlation between the attributes in this section of the dataset compared to the risk factors section of the dataset.

5.4 Most Significant Pentacam attributes

Within this section of the dataset there are more notable correlations between the attributes and the assessed class. The Keratometry readings are the highest correlators, but it is important to note the corneal thickness measurements (Pachymetry Readings) and refractive measurements also show high correlation.

The average correlation of the attributes to the assessment is 0.43 which can be seen as very significant.

5.5 Logistic Regression performance

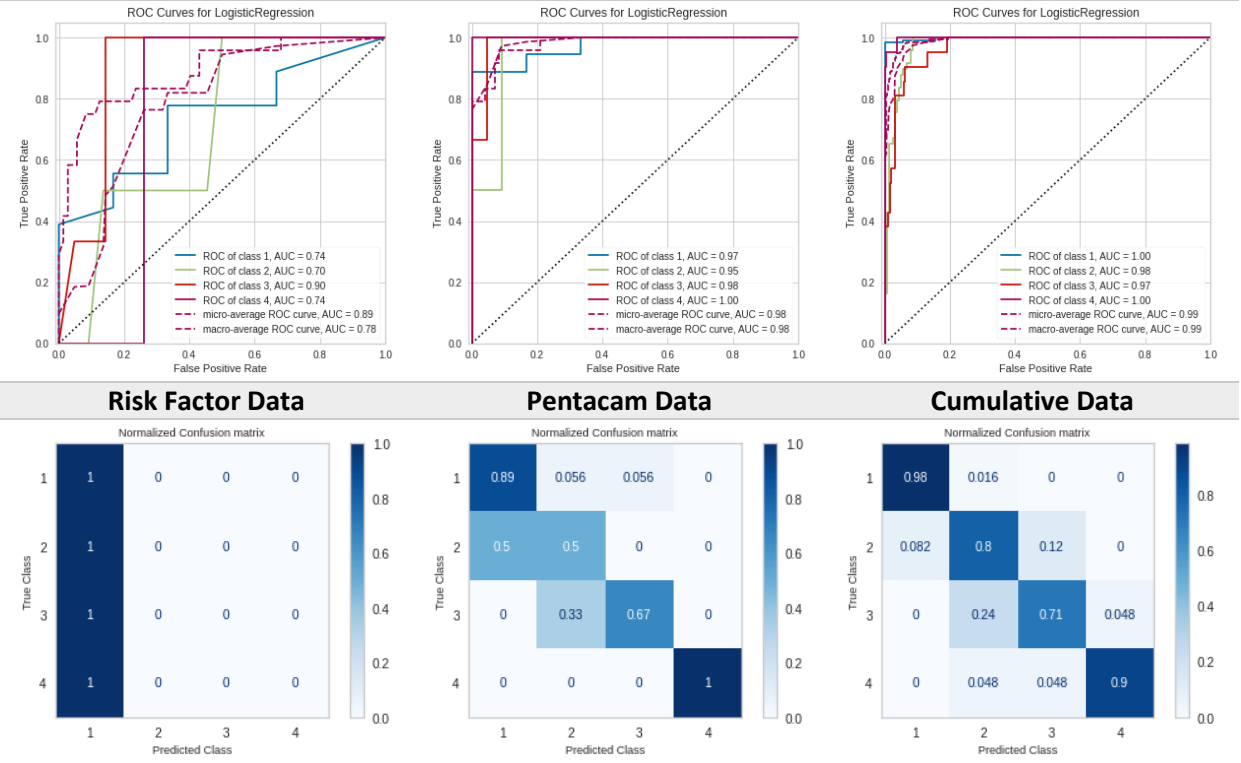


Figure III ROC curves and Confusion Matrices for Logistic Regression classifier.

Figure III depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the Logistic Regression classifiers.

5.5.1 Risk Factor Data

The ROC curves for all classes are closer to the middle, and classes 1-4 obtained the AUC values of 0.74, 0.70, 0.90, and 0.74, respectively. In this case, it demonstrates that the Logistic Regression classifier performs fairly well when it comes to classifying using the risk factor dataset.

5.5.2 Pentacam Data

The ROC curves for all classes are around the top left, and classes 1-4 obtained high AUC values 0.97, 0.95, 0.98, and 1.00, respectively. In this case, it demonstrates that the Logistic Regression classifier performs excellently when it comes to classifying using the pentacam dataset.

5.5.3 Cumulative Data

The ROC curves for all classes are at the top left, and classes 1-4 obtained high AUC values 1.00, 0.98, 0.97, and 1.00, respectively. In this case, it demonstrates that the Logistic Regression classifier performs excellently when it comes to classifying using the cumulative dataset.

5.5.4 Confusion Matrix

Figure III also presents the confusion matrices of the Logistic Regression classifiers we obtained a train accuracy of 59% and test accuracy of 75% for the classifier using the risk factor dataset, a train accuracy of 91% and test accuracy of 83% for the classifier using the pentacam dataset, a train accuracy of 91% and test accuracy of 91% for the classifier using the cumulative dataset.

5.6 K Neighbors performance

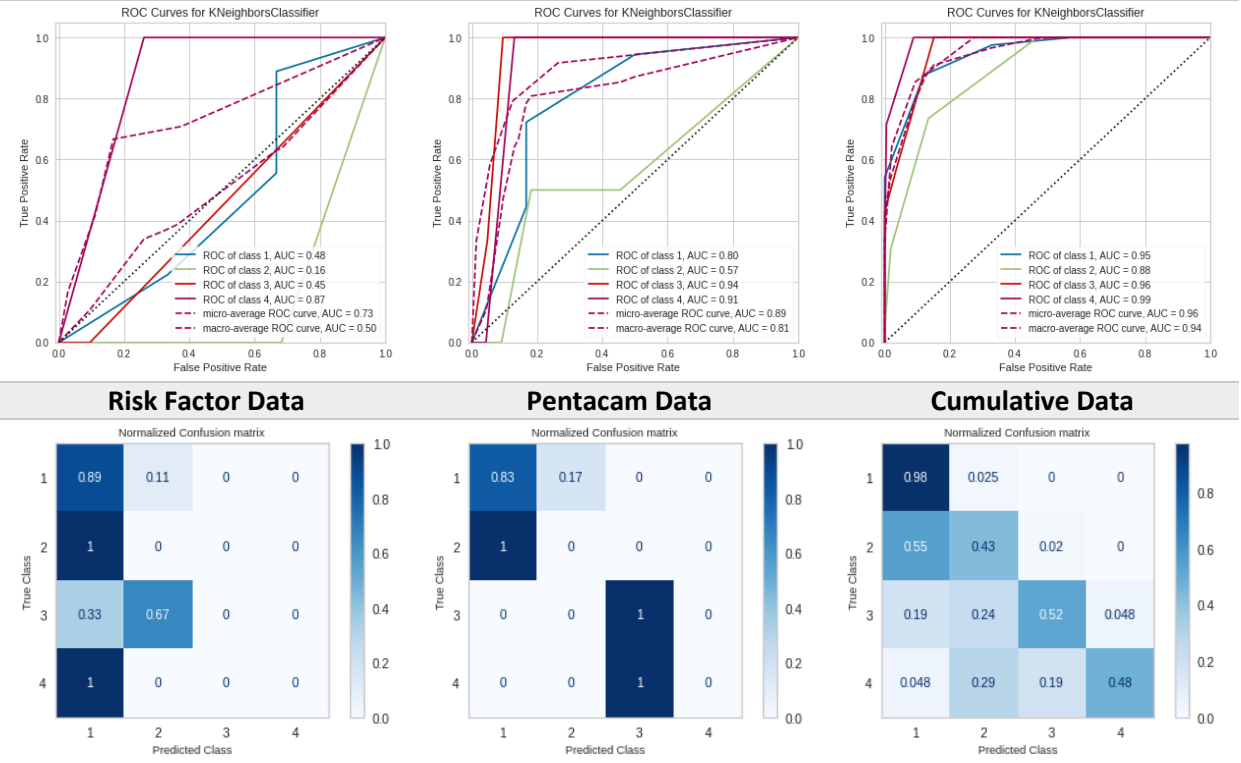


Figure IV ROC curves and Confusion Matrices for K Neighbors classifier.

Figure IV depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the K Neighbors classifiers.

5.6.1 Risk Factor Data

The ROC curves for all classes are around the middle, and classes 1-4 obtained the AUC values 0.48, 0.16, 0.45, and 0.87, respectively. In this case, it demonstrates that the K Neighbors classifier performs fairly well when it comes to classifying using the risk factor dataset.

5.6.2 Pentacam Data

The ROC curves for all classes are closer to the top left, and classes 1-4 obtained high AUC values 0.80, 0.57, 0.94, and 0.91, respectively. In this case, it demonstrates that the K Neighbors classifier performs fairly well when it comes to classifying using the pentacam dataset.

5.6.3 Cumulative Data

The ROC curves for all classes are around the top left, and classes 1-4 obtained high AUC values 0.95, 0.88, 0.96, and 0.99, respectively. In this case, it demonstrates that the K Neighbors classifier performs excellently when it comes to classifying using the cumulative dataset.

5.6.4 Confusion Matrix

Figure IV also presents the confusion matrices of the K Neighbors classifiers we obtained a train accuracy of 65% and test accuracy of 67% for the classifier using the risk factor dataset, a train accuracy of 76% and test accuracy of 75% for the classifier using the pentacam dataset, a train accuracy of 76% and test accuracy of 76% for the classifier using the cumulative dataset.

5.7 SVC performance

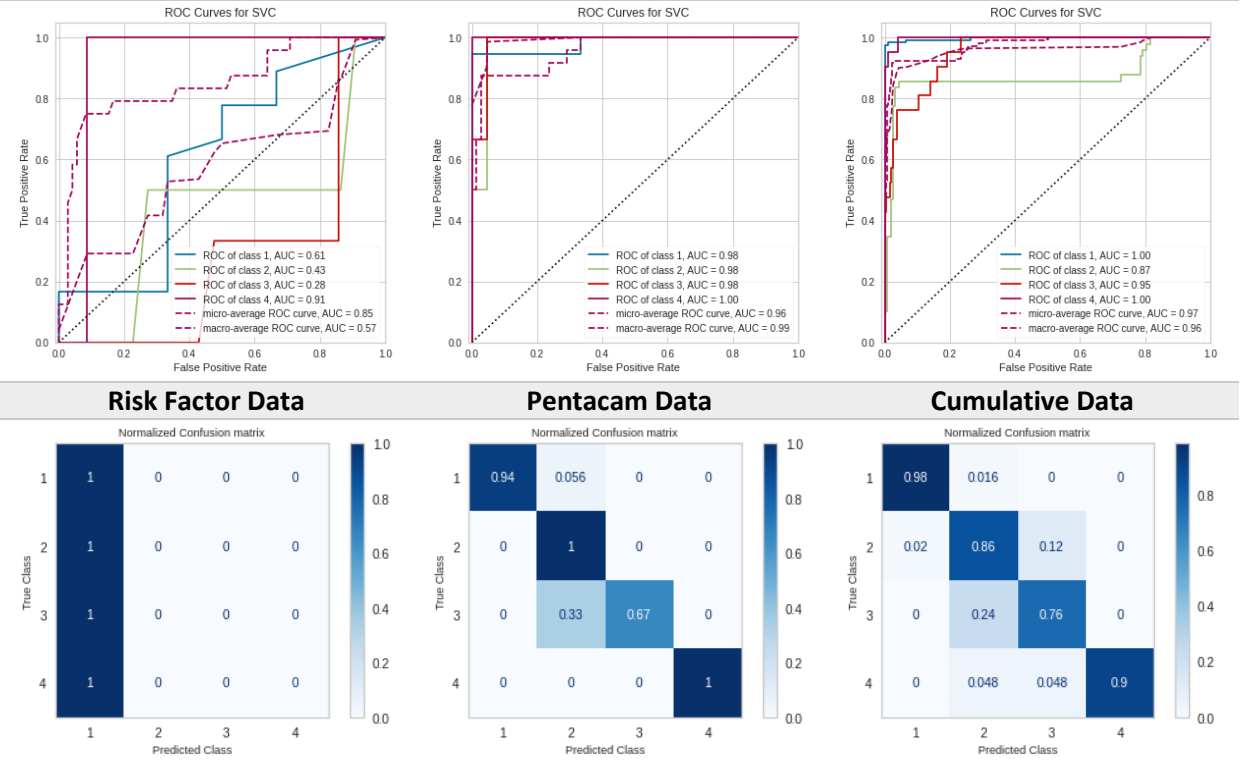


Figure V ROC curves and Confusion Matrices for SVC classifier.

Figure V depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the SVC classifiers.

5.7.1 Risk Factor Data

The ROC curves for all classes are around the middle, and classes 1-4 obtained the AUC values 0.61, 0.43, 0.28, and 0.91, respectively. In this case, it demonstrates that the SVC classifier performs fairly well when it comes to classifying using the risk factor dataset.

5.7.2 Pentacam Data

The ROC curves for all classes are around the top left, and classes 1-4 obtained high AUC values 0.98, 0.98, 0.98, and 1.00, respectively. In this case, it demonstrates that the SVC classifier performs excellently when it comes to classifying using the pentacam dataset.

5.7.3 Cumulative Data

The ROC curves for all classes are around the top left, and classes 1-4 obtained high AUC values 1.00, 0.87, 0.95, and 1.00, respectively. In this case, it demonstrates that the SVC classifier performs excellently when it comes to classifying using the cumulative dataset.

5.7.4 Confusion Matrix

Figure V also presents the confusion matrices of the SVC classifiers we obtained a train accuracy of 57% and test accuracy of 75% for the classifier using the risk factor dataset, a train accuracy of 92% and test accuracy of 92% for the classifier using the pentacam dataset, a train accuracy of 92% and test accuracy of 92% for the classifier using the cumulative dataset.

5.8 Decision Tree performance

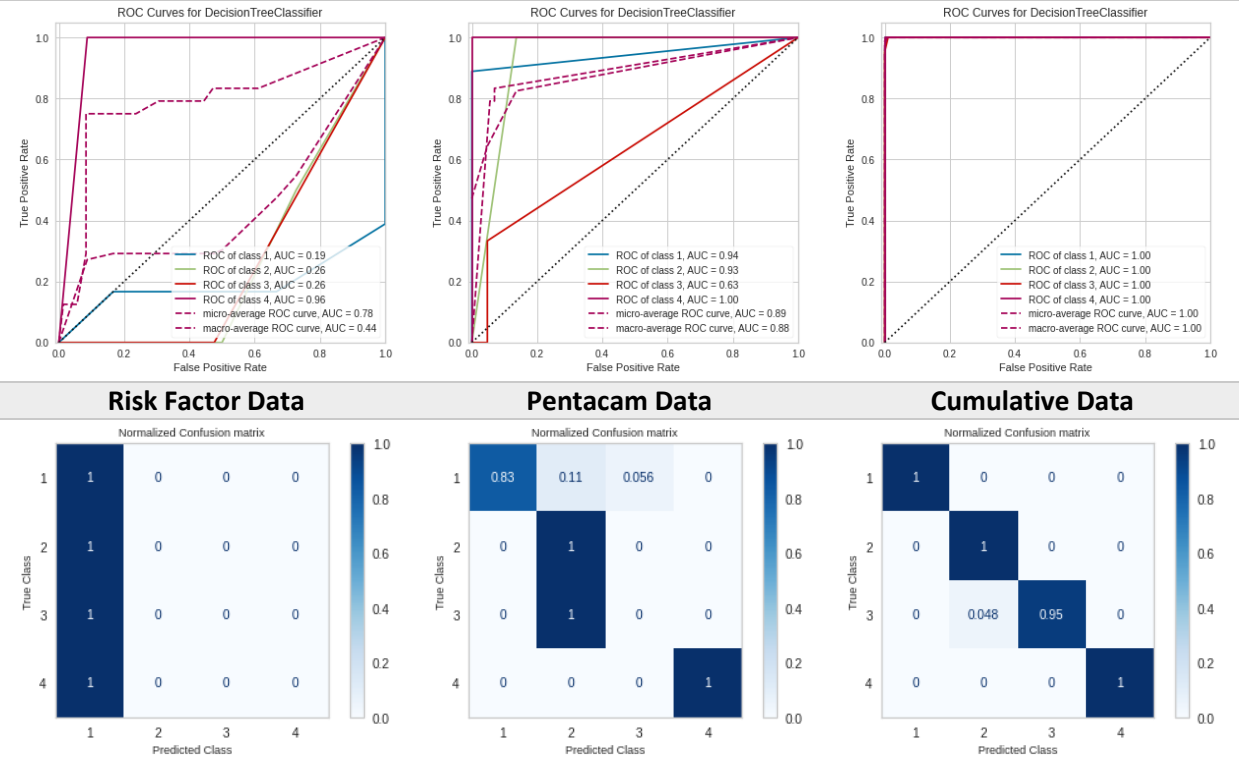


Figure VI ROC curves and Confusion Matrices for Decision Tree classifier.

Figure VI depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the Decision Tree classifiers.

5.8.1 Risk Factor Data

The ROC curves for all classes are all over the chart, and classes 1-4 obtained the AUC values 0.19, 0.26, 0.26, and 0.96, respectively. In this case, it demonstrates that the Decision Tree classifier performs poorly when it comes to classifying using the risk factor dataset.

5.8.2 Pentacam Data

The ROC curves for all classes are closer to the top left, and classes 1-4 obtained high AUC values 0.94, 0.93, 0.63, and 1.00, respectively. In this case, it demonstrates that the Decision Tree classifier performs fairly well when it comes to classifying using the pentacam dataset.

5.8.3 Cumulative Data

The ROC curves for all classes are at the top left, and classes 1-4 obtained high AUC values 1.00, 1.00, 1.00, and 1.00, respectively. In this case, it demonstrates that the Decision Tree classifier performs excellently when it comes to classifying using the cumulative dataset.

5.8.4 Confusion Matrix

Figure III also presents the confusion matrices of the Decision Tree classifiers we obtained a train accuracy of 65% and test accuracy of 67% for the classifier using the risk factor dataset, a train accuracy of 99% and test accuracy of 79% for the classifier using the pentacam dataset, a train accuracy of 100% and test accuracy of 100% for the classifier using the cumulative dataset.

5.9 Random Forest performance

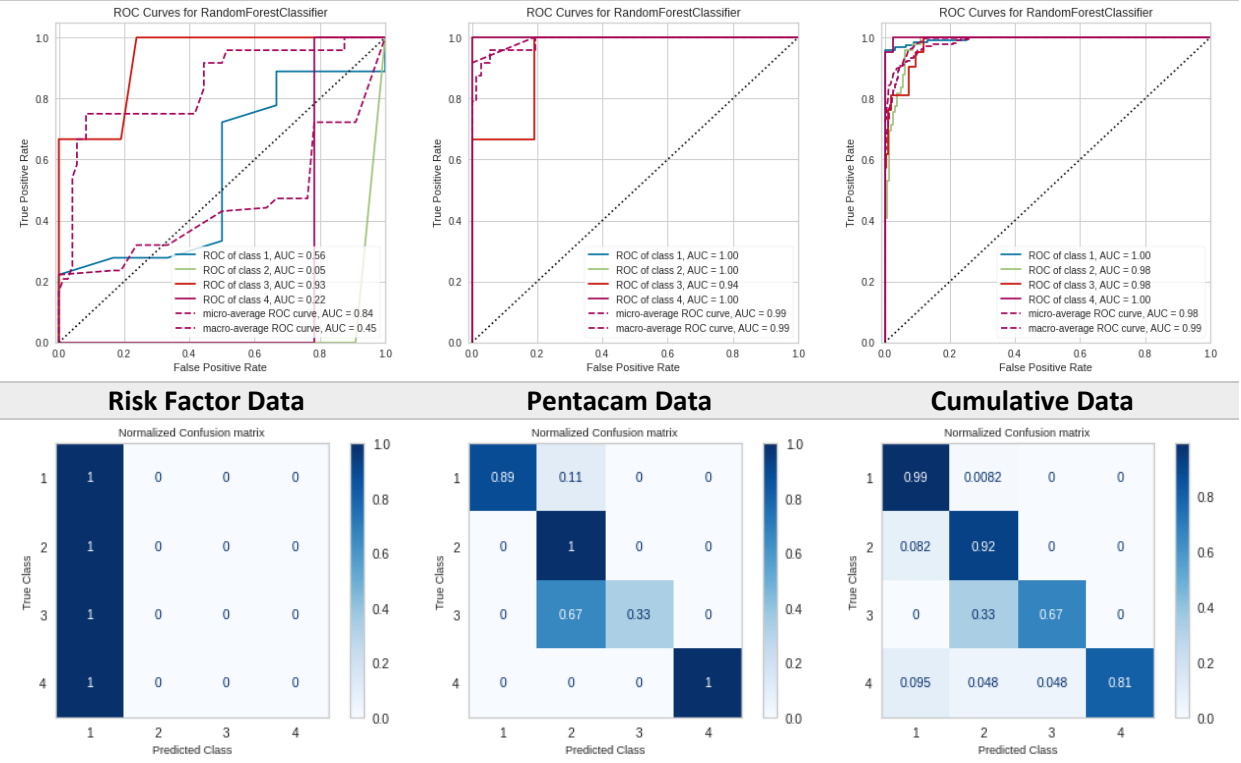


Figure VII ROC curves and Confusion Matrices for Random Forest classifier.

Figure VII depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the Random Forest classifiers.

5.9.1 Risk Factor Data

The ROC curves for all classes are all over the chart, and classes 1-4 obtained the AUC values 0.56, 0.05, 0.93, and 0.22, respectively. In this case, it demonstrates that the Random Forest classifier performs poorly when it comes to classifying using the risk factor dataset.

5.9.2 Risk Factor Data

The ROC curves for all classes are close to the top left, and classes 1-4 obtained high AUC values 1.00, 1.00, 0.94, and 1.00, respectively. In this case, it demonstrates that the Random Forest classifier performs excellently when it comes to classifying using the pentacam dataset.

5.9.3 Cumulative Data

The ROC curves for all classes are at the top left, and classes 1-4 obtained high AUC values 1.00, 0.98, 0.98, and 1.00, respectively. In this case, it demonstrates that the Random Forest classifier performs excellently when it comes to classifying using the cumulative dataset.

5.9.4 Confusion Matrix

Figure III also presents the confusion matrices of the Random Forest classifiers we obtained a train accuracy of 64% and test accuracy of 75% for the classifier using the risk factor dataset, a train accuracy of 98% and test accuracy of 96% for the classifier using the pentacam dataset, a train accuracy of 92% and test accuracy of 92% for the classifier using the cumulative dataset.

5.10 MLP performance

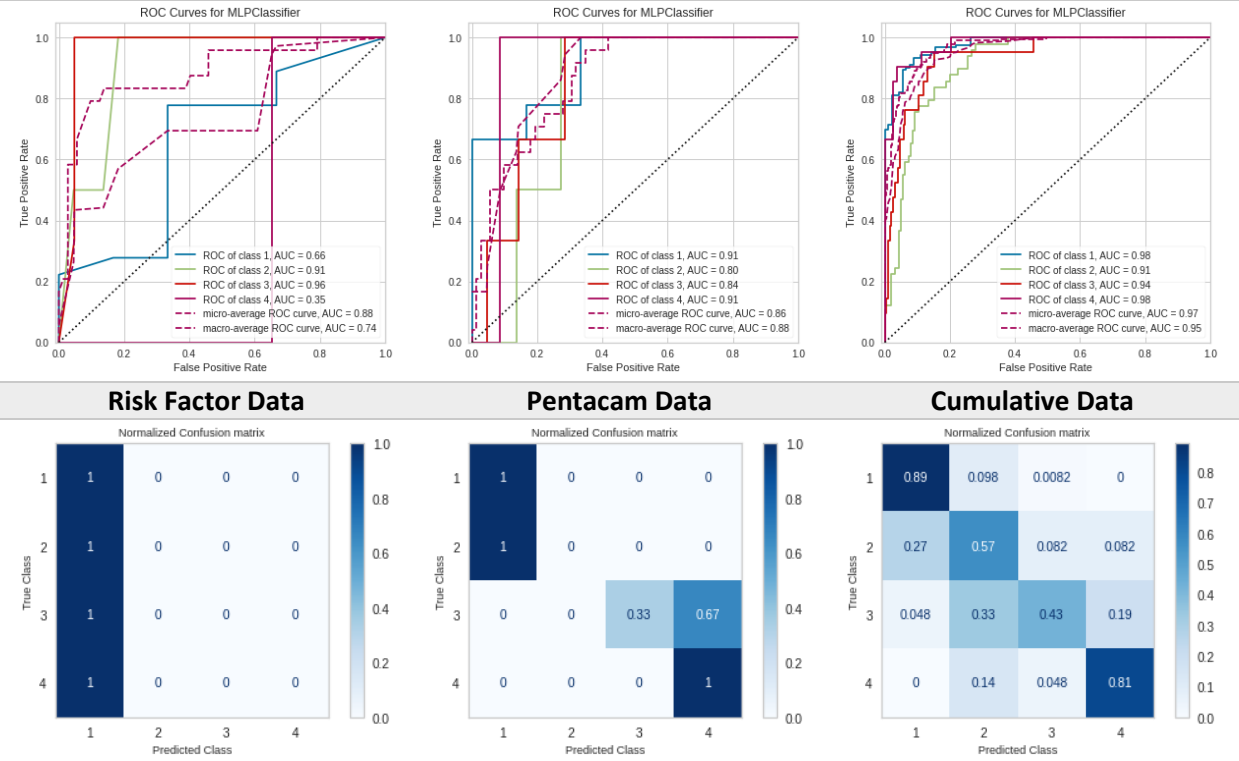


Figure VIII ROC curves and Confusion Matrices for Multi-Layer Perceptron classifier.

Figure III depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the Multi-Layer Perceptron classifiers.

5.10.1 Risk Factor Data

The ROC curves for all classes are around the middle, and classes 1-4 obtained the AUC values 0.66, 0.91, 0.96, and 0.35, respectively. In this case, it demonstrates that the MLP classifier performs fairly well when it comes to classifying using the risk factor dataset.

5.10.2 Pentacam Data

The ROC curves for all classes are closer to the top left, and classes 1-4 obtained high AUC values 0.91, 0.80, 0.84, and 0.91, respectively. In this case, it demonstrates that the MLP classifier performs decently when it comes to classifying using the pentacam dataset.

5.10.3 Cumulative Data

The ROC curves for all classes are around the top left, and classes 1-4 obtained high AUC values 0.98, 0.91, 0.94, and 0.98, respectively. In this case, it demonstrates that the MLP classifier performs excellently when it comes to classifying using the cumulative dataset.

5.10.4 Confusion Matrix

Figure III also presents the confusion matrices of the MLP classifiers we obtained a train accuracy of 59% and test accuracy of 75% for the classifier using the risk factor dataset, a train accuracy of 80% and test accuracy of 71% for the classifier using the pentacam dataset, a train accuracy of 79% and test accuracy of 79% for the classifier using the cumulative dataset.

5.11 ADA Boost performance

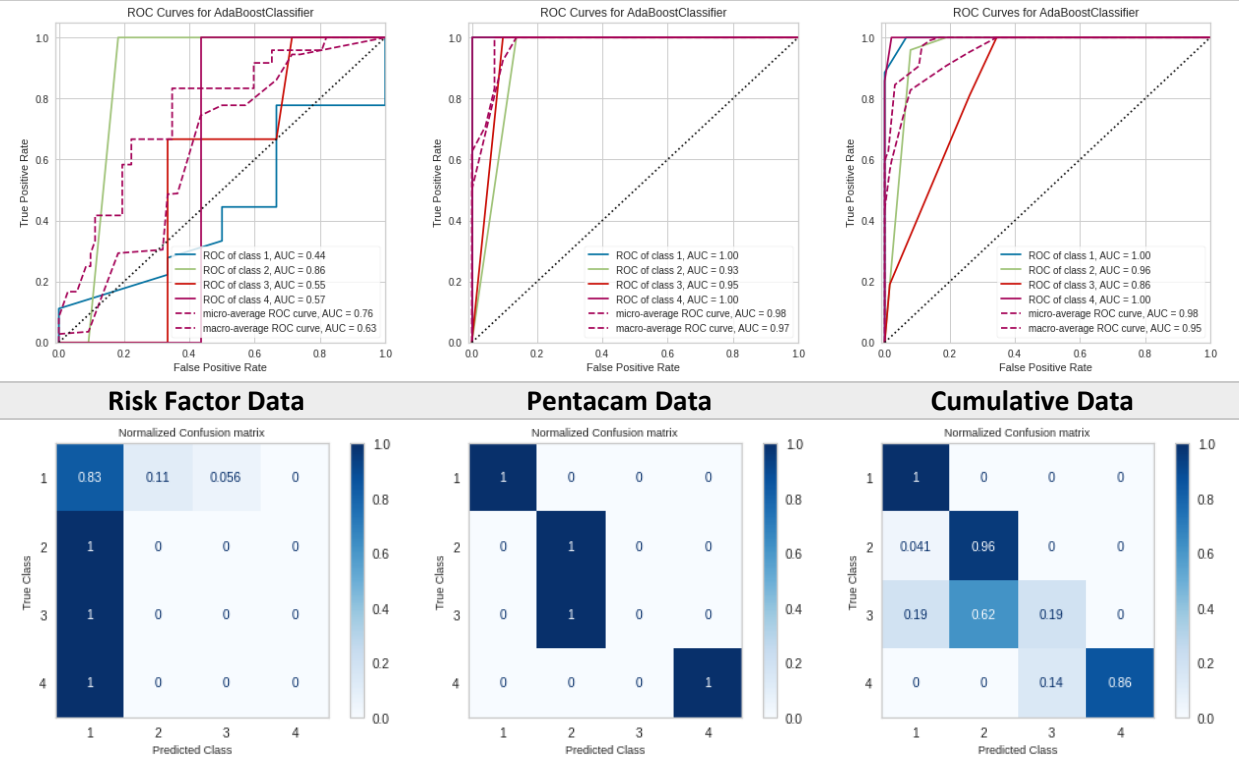


Figure IX ROC curves and Confusion Matrices for ADABoost classifier.

Figure IX depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the ADABoost classifiers.

5.11.1 Risk Factor Data

The ROC curves for all classes are around the middle, and classes 1-4 obtained the AUC values 0.44, 0.86, 0.55, and 0.57, respectively. In this case, it demonstrates that the ADABoost classifier performs fairly well when it comes to classifying using the risk factor dataset.

5.11.2 Pentacam Data

The ROC curves for all classes are closer to the top left, and classes 1-4 obtained high AUC values 1.00, 0.93 0.95, and 1.00, respectively. In this case, it demonstrates that the ADABoost classifier performs decently when it comes to classifying using the pentacam dataset.

5.11.3 Cumulative Data

The ROC curves for all classes are around the top left, and classes 1-4 obtained high AUC values 1.00, 0.96, 0.86, and 1.00, respectively. In this case, it demonstrates that the ADABoost classifier performs decently when it comes to classifying using the cumulative dataset.

5.11.4 Confusion Matrix

Figure III also presents the confusion matrices of the ADABoost classifiers we obtained a train accuracy of 49% and test accuracy of 63% for the classifier using the risk factor dataset, a train accuracy of 90% and test accuracy of 88% for the classifier using the pentacam dataset, a train accuracy of 90% and test accuracy of 90% for the classifier using the cumulative dataset.

5.12 Quadratic Discriminant performance

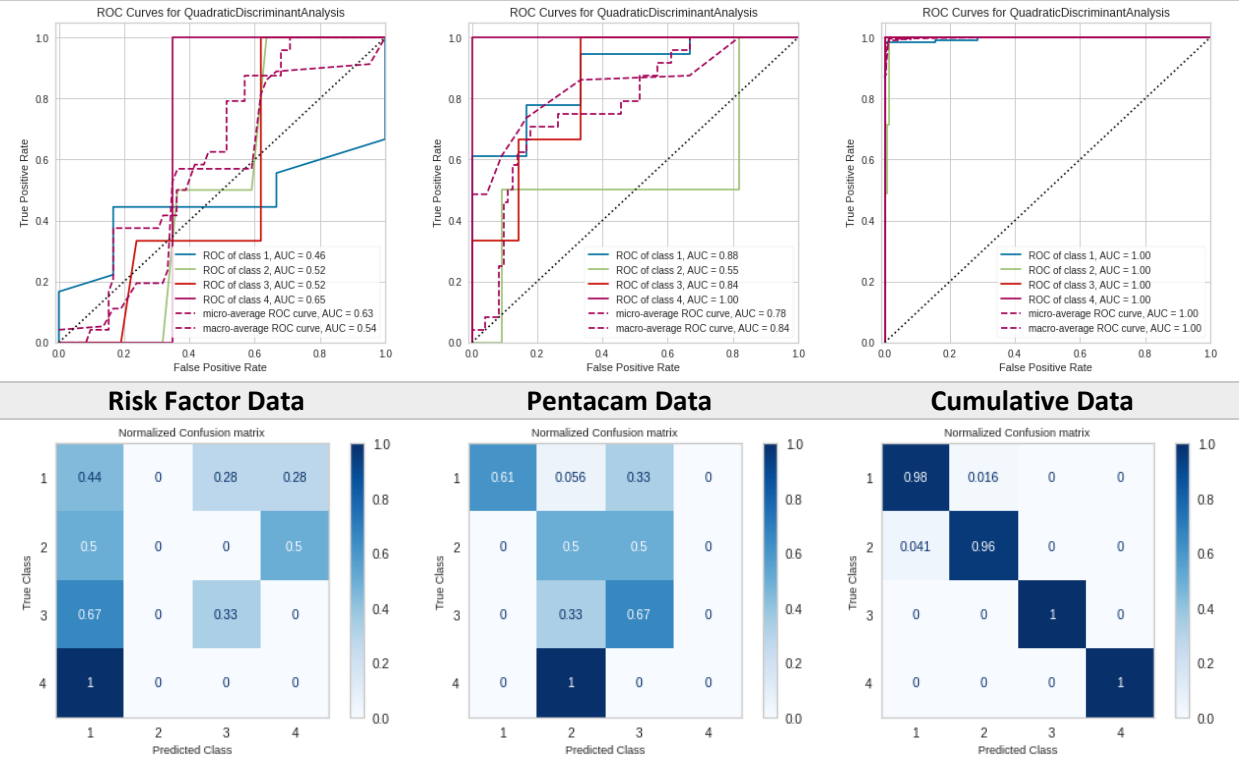


Figure X ROC curves and Confusion Matrices for Quadratic Discriminant classifier.

Figure X depicts the receiver operating characteristic (ROC) curves and area under the curve (AUC) for the Quadratic Discriminant Analysis classifiers.

5.12.1 Pentacam Data

The ROC curves for all classes are around the middle, and classes 1-4 obtained high AUC values 0.46, 0.52, 0.52, and 0.65, respectively. In this case, it demonstrates that the Quadratic Discriminant Analysis classifier performs fairly well when it comes to classifying using the risk factor dataset.

5.12.2 Pentacam Data

The ROC curves for all classes are closer to the top left, and classes 1-4 obtained high AUC values 0.88, 0.55 0.84, and 1.00, respectively. In this case, it demonstrates that the Quadratic Discriminant Analysis classifier performs decently when it comes to classifying using the pentacam dataset.

5.12.3 Cumulative Data

The ROC curves for all classes are at the top left, and classes 1-4 obtained high AUC values 1.00, 1.00, 1.00, and 1.00, respectively. In this case, it demonstrates that the Quadratic Discriminant Analysis classifier performs excellently when it comes to classifying using the cumulative dataset.

5.12.4 Confusion Matrix

Figure III also presents the confusion matrices of the Quadratic Discriminant Analysis classifiers we obtained a train accuracy of 42% and test accuracy of 38% for the classifier using the risk factor dataset, a train accuracy of 61% and test accuracy of 58% for the classifier using the pentacam dataset, a train accuracy of 98% and test accuracy of 98% for the classifier using the cumulative dataset.

6 Discussion

6.1 Risk Factor Correlation

6.1.1 Correlation between risk factors

There are some notable data points (above 0.2 correlation) within the heatmap that follow the general convention of what attributes we would expect to share some correlation, these datapoints include:

- The high correlation of 0.5 between atopy and hay fever
- Atopy, hypertension, hay fever, and family history of keratoconus are all highly correlated with General Health as would be expected considering these are all forms of illness.

Two of these points however bring forth interesting propositions:

- Hypertension has a notable link to Race with a value of 0.2, This is an idea that is supported by current research (Cardiol, 2016)
- There is also a notable correlation between Gender and Eye rubbing (0.24), but I was unable to find research to corroborate that position.

6.1.2 Most Significant Risk Factor Attributes

If we take a look at the data in the bar chart in Figure I, we can see which risk factors had the greatest link with the assessed class. In the data, the primary optical aid of the patient is the most significant contribution to the class evaluation process, followed by the gender of the patient and eye rubbing as the second most significant contributor. However, according to the data, the only really important factor is what the patient's primary eye aid is.

This does not follow closely with prevailing research in which eye rubbing is regarded as the most detrimental factor in the severity of keratoconus and gender has an inconclusive role in keratoconus diagnosis. The data also places a lot of weight on the impact of primary optical aid which the research has shown to be insignificant in the diagnosis of keratoconus. More research may be necessary to see if this analysis holds any weight.

6.2 Pentacam Correlation

6.2.1 Correlation between Pentacam Data

Figure II's correlation heatmap depicts the correlation strength between each pentacam characteristic. Due to the fact that a large number of the readings in this area of the dataset are a product of other attributes in both datasets, we can predict that there would be a larger overall correlation between these attributes than there would be between the risk factors.

6.2.2 Most Significant Pentacam Attributes

If we look at the bar chart in Figure II the data shows which risk factors had the most correlation to the assessed class. Within this section of the dataset there are more notable correlations between the attributes and the assessed class.

The Keratometry readings are the highest correlators as they track most closely to the grading system of the AK scheme, which is heavily reliant on keratometry readings. The corneal thickness measurements (Pachymetry Readings) and refractive measurements also show high correlation as they are both important metrics in the assessment of severity of keratoconus.

6.3 Overall Comparison

Model	Risk Factor	Pentacam	Cumulative
Logistic Regression	**75**	83	91
K Neighbors	67	75	76
SVC	**75**	**92**	92
Decision Tree	**75**	79	**100**
Random Forest	**75**	88	96
MLP	**75**	79	74
ADABoost	63	88	90
QDA	38	58	98
Average	*68*	*80*	*90*

Table 4 Accuracy Comparison

As we can see in Table 4 the prediction capability of the Risk factor dataset has a lower average accuracy than that of the pentacam dataset which itself has a lower accuracy than the cumulative dataset. This is to be expected as the addition of useful data to a model will generally result in better performance. The data shows that the addition of risk factor data improved model performance by 12.5%, which is a notable performance boost.

The MLP model is the only model to drop in accuracy when combining both data sections, this may be due to the depth of the network and its incompatibility with a high number of attributes but relatively little amount of data. We can say this illustrates the suitability of machine learning models for fields in which there is a small amount of data available for training.

Several models performed fairly well at diagnosing with only risk factor data but when including average AUC in our analysis as well as pentacam attributes we can have more insight into what the most suitable model to use would be in any given scenario.

Model	Risk Factor	Pentacam	Cumulative
Logistic Regression	**0.78**	0.98	0.99
K Neighbors	0.50	0.81	0.94
SVC	0.57	0.99	0.96
Decision Tree	0.44	0.88	**1.00**
Random Forest	0.25	**1.00**	0.99
MLP	0.76	0.88	0.90
ADABoost	0.63	0.97	0.95
QDA	0.54	0.84	**1.00**
Average	*0.56*	*0.92*	*0.97*

Table 5 AUC Comparison

In Table 5 the results show a lower average AUC for the risk factor versions of the models compared to the pentacam versions of the models, this phenomenon has similarly been observed in a recent study, where Fernández Pérez, J., Valero Marcos, A. and Martínez Peña, F.J. (2014) demonstrated that the use of corneal equipment and techniques such as Pentacam in combination with risk factors can result in the detection of subclinical keratoconus, but at the expense of a higher rate of false-positive detections.

This is a limitation of the risk factor data in which it has a tendency to underdiagnose severity in patients. More work can be done in tuning the algorithms towards a lower false positive detection rate.

Section	Risk Factor	Pentacam	Cumulative
Correlation	0.09	**0.43**	-
Macro Average AUC	0.56	0.92	**0.97**
Accuracy	68	80	**90**

Table 6 Overall Comparison

Table 6 illustrates the three key factors in determining the significance of each section to the clinical assessment.

Looking at the correlation data we see that the Pentacam data has a very significant correlation to the assessment, compared to the risk factor data which shows little relevance. However, as previously mentioned, the 12.5% overall increase in performance upon adding risk factor data indicates that there is a noticeable relevance to the assessment.

The results in Table 6 also shows a slight increase of 5% in average AUC with the addition of risk factor data illustrating how little the risk factor helps in improving specifity and sensitivity of a predictive model.

Model	Hallett	Cumulative
Macro Average AUC	0.90	**0.97**
Accuracy	77	**90**

Table 7 Hallett et al., (2020) vs Cumulative Model

Comparing the performance of the deep learning models used by Hallett et al., (2020) to the machine learning models used in this research paper illustrate that machine learning models have a better overall performance than the deep learning models when using the same dataset of small sample size.

7 Conclusion

7.1 Implications

7.1.1 Most Significant Risk Factors

Discrepancies arose with regards to the most significant risk factors i.e., the significance of primary optical aid, gender in assessing the severity of keratoconus, and current research. The data contrasts with the current research on these factors which may lead us to believe that dataset bias may have played a part in these results. More clinical research on these factors may be necessary to debunk or support any correlation between these factors and the assessment of the disease.

7.1.2 Risk Factor Significance

After conducting this study, I can comfortably say that the addition of risk factor data to the overall dataset increases the performance of the predictive models by a significant amount but may also bring about a higher false positive detection rate. The study has illustrated the necessity of employing risk factor data in the training process but more work needs to be done on the false positive detection rate, before this risk factor data can be used solely.

Overall, the risk factors average correlation to the assessment is 0.09 which seems insignificant but still brought about a 12.5% increase in accuracy when employed, this highlights a need to call into question the correlation value as a useful metric for measuring data relevance.

7.1.3 Machine Learning vs Deep Learning

Comparing the results gotten in this paper to those in the original paper by Hallett et al., (2020), there is a notable average overall increase in performance which illustrates the effectiveness of machine learning algorithms on small datasets over deep learning algorithms.

7.2 Limitations of the study

The dataset for this study was collected by a different author, so the veracity of the data and the diligence with which it was obtained are assumed to be up to adequate standards for the purposes of completing this research project.

Most research is focused on detection of keratoconus and not the diagnosis of stage of severity so the literature used to support this study may not be wholly applicable but may be sufficient to serve the purpose of this paper.

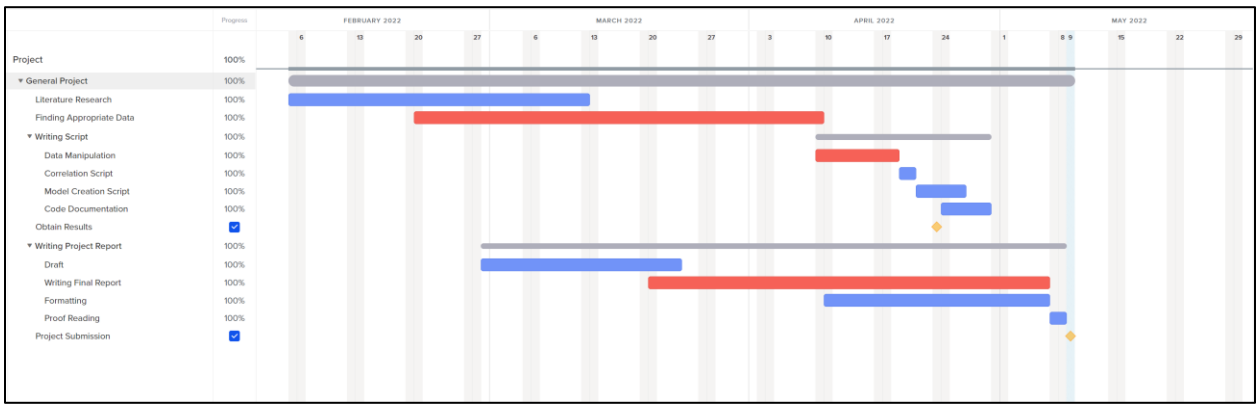
7.3 Recommendations

There should be more effort towards collecting risk factor data when assessing keratoconus and it should be a somewhat necessity to include them in the creation of a predictive model.

Finally, As the data for the diagnosis of this disease remains scarce, machine learning techniques should receive more favor as they demonstrate a notable increase in performance compared to deep learning techniques.

With tuning, the efficacy of these models can significantly improve and moreover increase the likelihood of implementation in real world diagnosis.

8 Gantt Chart



The figure above denotes the project timeline as well as what tasks were subject to delays, milestones and when they were achieved. The blue bars denote the tasks that were done on time while the red bars illustrate delayed tasks.

Due to a drop in communication with the original ophthalmologist for this project, the data collection stage and report writing stage suffered delays.

9 Appendix (Code)

```
# -*- coding: utf-8 -*-
"""Code.ipynb

Automatically generated by Colaboratory.

Original file is located at
    https://colab.research.google.com/drive/1sVXLwQnsHflvkbBsD58A3RUGVDxF4rlx
"""

import sys
!{sys.executable} -m pip install pandas seaborn

#Visualisation Libraries
import matplotlib.pyplot as plt
import seaborn as sns
#Array and DataStructures libraries
import numpy as np
import pandas as pd

pd.set_option("display.max_columns", 101)

# Dataset is already loaded below
data = pd.read_csv("train.csv")
#Explore columns
data.columns = data.columns.str.replace("'","")
data.columns = data.columns.str.replace(" ","_")
data.columns;

"""### Dimensionality Reduction

I am manually reducing the number of columns to be used for training the model so as to simplify the dataset and produce a better fit predictive model
"""

#Reducing number of columns in dataset
train_data_reduced_patient = data.drop(['Eye','LogMAR_UDVA',
       'LogMAR_CDVA', 'Sphere', 'Cylinder', 'Axis', 'Keratometry_Unit_',
       'Flat_Keratometry', 'Steep_Keratometry_', 'Steep_Axis_',
       'Mean_Topograpjy_K', 'Topography_Cylinder', 'Central_Pachy',
       'Thinnest_pachy', 'Location_X_Axis', 'Location_Y_Axis',
       'SLE_Corneal_Scarring', 'SLE_Vogts_Striae_', 'Fleischers_ring_'], axis = 1)

#Reducing number of columns in dataset
train_data_reduced_pentacam= data.drop(['Gender', 'Race',
'General_Health_(requiring_Medication)', 'Atopy_',
       'Hypertension', 'Hayfever_', 'Known_Eye_History', 'Eye_Rubbing',
       'Family_History_KC', 'Primary_Optical_Aid', 'Eye', 'LogMAR_UDVA',
       'LogMAR_CDVA', 'SLE_Corneal_Scarring', 'SLE_Vogts_Striae_', 'Fleischers_ring_'],
axis = 1)

train_data_reduced_all = data.drop([ 'Eye', 'LogMAR_UDVA',
       'LogMAR_CDVA', 'SLE_Corneal_Scarring', 'SLE_Vogts_Striae_', 'Fleischers_ring_'],
axis = 1)
```

```
"""### Data Wrangling and Cleanup for Test Dataset

"""

# Dataset is already loaded below
test_data = pd.read_csv("test.csv")
#Explore columns
test_data.columns = test_data.columns.str.replace("'","")
test_data.columns = test_data.columns.str.replace(" ","_")
test_data = test_data.astype({"Race":'str', "Primary_Optical_Aid":'str'})
test_data.dtypes

"""### Dimensionality Reduction

"""

#Reducing number of columns in dataset
test_data_reduced_patient = test_data.drop(['Eye','LogMAR_UDVA',
       'LogMAR_CDVA', 'Sphere', 'Cylinder', 'Axis', 'Keratometry_Unit_',
       'Flat_Keratometry', 'Steep_Keratometry_', 'Steep_Axis_',
       'Mean_Topograpjy_K', 'Topography_Cylinder', 'Central_Pachy',
       'Thinnest_pachy', 'Location_X_Axis', 'Location_Y_Axis',
       'SLE_Corneal_Scarring', 'SLE_Vogts_Striae_', 'Fleischers_ring_'], axis = 1)

#Reducing number of columns in dataset
test_data_reduced_pentacam= test_data.drop(['Gender', 'Race',
'General_Health_(requiring_Medication)', 'Atopy_',
       'Hypertension', 'Hayfever_', 'Known_Eye_History', 'Eye_Rubbing',
       'Family_History_KC', 'Primary_Optical_Aid', 'Eye', 'LogMAR_UDVA',
       'LogMAR_CDVA', 'SLE_Corneal_Scarring', 'SLE_Vogts_Striae_', 'Fleischers_ring_'],
axis = 1)

test_data_reduced_all = data.drop([ 'Eye', 'LogMAR_UDVA',
       'LogMAR_CDVA', 'SLE_Corneal_Scarring', 'SLE_Vogts_Striae_', 'Fleischers_ring_'],
axis = 1)
```

```
"""#### Illustrating the correlation between all the attributes"""
correlations_patient = train_data_reduced_patient.corr()
fig, ax = plt.subplots(figsize=(10,10))
#making heatmap with customised visualisation settings
sns.heatmap(correlations_patient, vmax=1.0, center=0, fmt='.2f', square=True,
linewidths=.5, annot=True, cbar_kws={"shrink": .70})
plt.show()
correlations_pentacam = train_data_reduced_pentacam.corr()
fig, ax = plt.subplots(figsize=(10,10))
#making heatmap with customised visualisation settings
sns.heatmap(correlations_pentacam, vmax=1.0, center=0, fmt='.2f', square=True,
linewidths=.5, annot=True, cbar_kws={"shrink": .70})
plt.show()
# creating the barplot with customised visualisation settings
correlation_viz_patient = correlations_patient
correlation_viz_patient['feature'] = correlation_viz_patient.index

ix = (abs(train_data_reduced_patient.corr()).sort_values('label', ascending =
False).index)
order = train_data_reduced_patient.loc[:, ix]

plt.figure(figsize=(6,6))
sns.barplot(x='feature', y="label", data=correlation_viz_patient, order=order)
plt.xlabel("Features")
plt.ylabel("Correlation")
plt.title("Features")
plt.xticks(rotation=90)
plt.show()

# creating the barplot with customised visualisation settings
correlation_viz_pentacam = correlations_pentacam
correlation_viz_pentacam['feature'] = correlation_viz_pentacam.index

ix = (abs(train_data_reduced_pentacam.corr()).sort_values('label', ascending =
False).index)
order = train_data_reduced_pentacam.loc[:, ix]
plt.figure(figsize=(6,6))
sns.barplot(x='feature', y="label", data=correlation_viz_pentacam, order=order)
plt.xlabel("Features")
plt.ylabel("Correlation")
plt.title("Features")
plt.xticks(rotation=90)
plt.show()

x1train = train_data_reduced_patient.drop(['label'], axis = 1)
y1train = train_data_reduced_patient['label']
x1test = test_data_reduced_patient.drop(['label'], axis = 1)
y1test = test_data_reduced_patient['label']

x2train = train_data_reduced_pentacam.drop(['label'], axis = 1)
y2train = train_data_reduced_pentacam['label']
x2test = test_data_reduced_pentacam.drop(['label'], axis = 1)
y2test = test_data_reduced_pentacam['label']

x3train = train_data_reduced_all.drop(['label'], axis = 1)
y3train = train_data_reduced_all['label']
x3test = test_data_reduced_all.drop(['label'], axis = 1)
y3test = test_data_reduced_all['label']
```

```
from sklearn.metrics import confusion_matrix
from sklearn.metrics import ConfusionMatrixDisplay
from yellowbrick.classifier import ROCAUC
from sklearn.linear_model import LogisticRegression
from sklearn.neural_network import MLPClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.svm import SVC
from sklearn.gaussian_process import GaussianProcessClassifier
from sklearn.gaussian_process.kernels import RBF
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier, AdaBoostClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.discriminant_analysis import QuadraticDiscriminantAnalysis
import warnings

def CreateModel(model, xtrain, xtest, ytrain, ytest, Name):
    train_acc = model.score(xtrain, ytrain)
    test_acc= model.score(xtest, ytest)

    print('\n')
    print(f"{Name}")
    print(f"Train Accuracy = {train_acc*100} %")
    print(f"Test Accuracy = {test_acc*100} %")

    fig, ax = plt.subplots(figsize = (6,6))
    visualizer = ROCAUC(model)
    visualizer.fit(xtrain, ytrain)        # Fit the training data to the visualizer
    visualizer.score(xtest, ytest)        # Evaluate the model on the test data
    visualizer.finalize()
    ax.set_xlim([-0.01, 1])
    ax.set_ylim([0.0, 1.05])

    disp = ConfusionMatrixDisplay.from_estimator(model.fit(xtrain,
ytrain),xtest,ytest,cmap=plt.cm.Blues,normalize='true',)
    disp.ax_.set_title("Normalized Confusion matrix", fontsize=9)
    disp.ax_.set_xlabel("Predicted Class", fontsize=9)
    disp.ax_.set_ylabel("True Class", fontsize=9)
    plt.grid(False)
    plt.show()

    return
```

```
MODEL1 = LogisticRegression(solver='lbfgs', max_iter=10000)
MODEL2 = KNeighborsClassifier(4)
MODEL3 = SVC(kernel="linear", C=0.025)
MODEL4 = DecisionTreeClassifier(max_depth=5)
MODEL5 = RandomForestClassifier(max_depth=5, n_estimators=10, max_features=1)
MODEL6 = MLPClassifier(alpha=1, max_iter=1000)
MODEL7 = AdaBoostClassifier()
MODEL8 = QuadraticDiscriminantAnalysis()

MODELS = [MODEL1, MODEL2, MODEL3, MODEL4, MODEL5, MODEL6, MODEL7, MODEL8]
MODELNAMES = ["LogisticRegression", "KNeighborsClassifier", "SVC",
"DecisionTreeClassifier", "RandomForestClassifier", "MLPClassifier",
"ADABoostClassifier", "QuadraticDiscriminantAnalysis"]

warnings.filterwarnings('ignore', category=DeprecationWarning)
for i in range(len(MODELS)):
    CreateModel(MODELS[i].fit(x1train, y1train), x1train, x1test, y1train, y1test,
MODELNAMES[i])

for i in range(len(MODELS)):
    CreateModel(MODELS[i].fit(x2train, y2train), x2train, x2test, y2train, y2test,
MODELNAMES[i])

for i in range(len(MODELS)):
    CreateModel(MODELS[i].fit(x3train, y3train), x3train, x3test, y3train, y3test,
MODELNAMES[i])
```

10 References

Al-Timemy, A.H. *et al.* (2021) 'A Hybrid Deep Learning Construct for Detecting Keratoconus From Corneal Maps', *Translational vision science & technology*, 10(14), p. 16.

Alam, M.S. and Vuong, S.T. (2013) 'Random Forest Classification for Detecting Android Malware', in *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pp. 663–669.

Alió, J.L. *et al.* (2015) 'Keratoconus Management Guidelines', 4(1), pp. 1–39.

Alió, J.L. (ed.) (2017) *Keratoconus: Recent Advances in Diagnosis and Treatment*. Springer, Cham.

Ambrósio, R., Jr *et al.* (2006) 'Corneal-thickness spatial profile and corneal-volume distribution: tomographic indices to detect keratoconus', *Journal of cataract and refractive surgery*, 32(11), pp. 1851–1859.

Balasubramanian, S.A., Pye, D.C. and Willcox, M.D.P. (2013) 'Effects of eye rubbing on the levels of protease, protease activity and cytokines in tears: relevance in keratoconus', *Clinical & experimental optometry: journal of the Australian Optometrical Association*, 96(2), pp. 214–218.

Barbara, A. (2012) 'Textbook on Keratoconus: New Insights', in *jaypee*. jaypee.

Beckh, U., Schönherr, U. and Naumann, G.O. (1995) '[Autosomal dominant keratoconus as the chief ocular symptom in Lobstein osteogenesis imperfecta tarda]', *Klinische Monatsblatter fur Augenheilkunde*, 206(4), pp. 268–272.

Bennett, A., Parto, P. and Krim, S.R. (2016) 'Hypertension and ethnicity', *Current opinion in cardiology*, 31(4), pp. 381–386.

Bolarín, J.M. *et al.* (2020) 'A Machine-Learning Model Based on Morphogeometric Parameters for RETICS Disease Classification and GUI Development', *NATO Advanced Science Institutes series E: Applied sciences*, 10(5), p. 1874.

Breiman, L. (2001) 'Random Forests', *Machine learning*, 45(1), pp. 5–32.

Cao, K. *et al.* (2022) 'Accuracy of Machine Learning Assisted Detection of Keratoconus: A Systematic Review and Meta-Analysis', *Journal of clinical medicine research*, 11(3). doi:10.3390/jcm11030478.

Chastang, P.J. *et al.* (2000) 'Automated keratoconus detection using the EyeSys videokeratoscope', *Journal of cataract and refractive surgery*, 26(5), pp. 675–683.

Chen, D. and Lam, A.K.C. (2009) 'Reliability and repeatability of the Pentacam on corneal curvatures', *Clinical & experimental optometry: journal of the Australian Optometrical Association*, 92(2), pp. 110–118.

Chen, Z.-J., Liu, B. and He, X.-P. (2007) 'A SVC Iterative Learning Algorithm Based on Sample Selection for Large Samples', in *2007 International Conference on Machine Learning and Cybernetics*, pp. 3308–3313.

Cover, T. and Hart, P. (1967) 'Nearest neighbor pattern classification', *IEEE transactions on information theory / Professional Technical Group on Information Theory*, 13(1), pp. 21–27.

Criminisi, A., Shotton, J. and Konukoglu, E. (2012) 'Decision Forests: A Unified Framework for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning', *Found. Trends. Comput. Graph. Vis.*, 7(2–3), pp. 81–227.

Cullen, J.F. and Butler, H.G. (1963) 'MONGOLISM (DOWN'S SYNDROME) AND KERATOCONUS', *The British journal of ophthalmology*, 47, pp. 321–330.

Cutler, A., Cutler, D.R. and Stevens, J.R. (2012) 'Random Forests', in Zhang, C. and Ma, Y. (eds) *Ensemble Machine Learning: Methods and Applications*. Boston, MA: Springer US, pp. 157–175.

Fernández Pérez, J., Valero Marcos, A. and Martínez Peña, F.J. (2014) 'Early diagnosis of keratoconus: what difference is it making?', *The British journal of ophthalmology*, 98(11), pp. 1465–1466.

Fink, B.A. *et al.* (2005) 'Differences in keratoconus as a function of gender', *American journal of ophthalmology*, 140(3), pp. 459–468.

Freund, Y. and Schapire, R.E. (1997) 'A decision-theoretic generalization of on-line learning and an application to boosting', *Journal of Computer and System Sciences*, 55(1), pp. 119–139.

Galin, M.A. and Berger, R. (1958) 'Atopy and keratoconus', *American journal of ophthalmology*, 45(6), pp. 904–906.

Giannaccare, G. *et al.* (2021) 'Comparison of Amsler–Krumeich and Sandali Classifications for Staging Eyes with Keratoconus', *NATO Advanced Science Institutes series E: Applied sciences*, 11(9), p. 4007.

Hallett, N. *et al.* (2020) 'Deep Learning Based Unsupervised and Semi-supervised Classification for Keratoconus (CornealAI)', *arXiv [stat.ML]*. Available at: http://arxiv.org/abs/2001.11653.

Javadi, M.-A. *et al.* (2004) 'Concomitant keratoconus and macular corneal dystrophy', *Cornea*, 23(5), pp. 508–512.

Jenhani, I., Amor, N.B. and Elouedi, Z. (2008) 'Decision trees as possibilistic classifiers', *International journal of approximate reasoning: official publication of the North American Fuzzy Information Processing Society*, 48(3), pp. 784–807.

Kamiya, K. *et al.* (2019) 'Keratoconus detection using deep learning of colour-coded maps with anterior segment optical coherence tomography: a diagnostic accuracy study', *BMJ open*, 9(9), p. e031313.

Kennedy, R.H., Bourne, W.M. and Dyer, J.A. (1986) 'A 48-year clinical and epidemiologic study of keratoconus', *American journal of ophthalmology*, 101(3), pp. 267–273.

Krachmer, J.H. (2004) 'Eye rubbing can cause keratoconus', *Cornea*, pp. 539–540.

Kuming, B.S. and Joffe, L. (1977) 'Ehlers-Danlos syndrome associated with keratoconus. A case report', *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde*, 52(10), pp. 403–405.

Kuo, I.C. *et al.* (2006) 'Is there an association between diabetes and keratoconus?', *Ophthalmology*, 113(2), pp. 184–190.

Lavric, A. *et al.* (undefined 2020) 'Detecting Keratoconus From Corneal Imaging Data Using Machine Learning', *IEEE Access*, 8, pp. 149113–149121.

Liza-Sharmini, A.T., Azlan, Z.N. and Zilfalil, B.A. (2006) 'Ocular findings in Malaysian children with Down syndrome', *Singapore medical journal*, 47(1), pp. 14–19.

Macsai, M.S., Varley, G.A. and Krachmer, J.H. (1990) 'Development of keratoconus after contact lens wear. Patient characteristics', *Archives of ophthalmology*, 108(4), pp. 534–538.

Maumenee, I.H. (1981) 'The eye in the Marfan syndrome', *Transactions of the American Ophthalmological Society*, 79, pp. 684–733.

McDermott, M.L. *et al.* (1998) 'Corneal topography in Ehlers-Danlos syndrome', *Journal of cataract and refractive surgery*, 24(9), pp. 1212–1215.

McMonnies, C.W. and Boneham, G.C. (2010) 'Corneal responses to intraocular pressure elevations in keratoconus', *Cornea*, 29(7), pp. 764–770.

Naderan, M. *et al.* (2015) 'Characteristics and associations of keratoconus patients', *Contact lens & anterior eye: the journal of the British Contact Lens Association*, 38(3), pp. 199–205.

Negris, R. (1992) 'Floppy eyelid syndrome associated with keratoconus', *Journal of the American Optometric Association*, 63(5), pp. 316–319.

Omer, K. (2018) 'Epidemiology of keratoconus worldwide', *The open ophthalmology journal*, 12(1), pp. 289–299.

Pan, C.-W. *et al.* (2014) 'Ethnic variation in central corneal refractive power and steep cornea in Asians', *Ophthalmic epidemiology*, 21(2), pp. 99–105.

Piñero, D.P. *et al.* (2010) 'Corneal volume, pachymetry, and correlation of anterior and posterior corneal shape in subclinical and different stages of clinical keratoconus', *Journal of cataract and refractive*

surgery, 36(5), pp. 814–825.

Rabinowitz, Y.S. (1998) 'Keratoconus', *Survey of ophthalmology*, 42(4), pp. 297–319.

Shapiro, M.B. and France, T.D. (1985) 'The ocular features of Down's syndrome', *American journal of ophthalmology*, 99(6), pp. 659–663.

Sharif, K.W., Casey, T.A. and Coltart, J. (1992) 'Prevalence of mitral valve prolapse in keratoconus patients', *Journal of the Royal Society of Medicine*, 85(8), pp. 446–448.

Sharif, R. *et al.* (2018) 'Pathogenesis of Keratoconus: The intriguing therapeutic potential of Prolactin-inducible protein', *Progress in retinal and eye research*, 67, pp. 150–167.

Shneor, E. *et al.* (2014) 'Prevalence of Keratoconus among Young Arab Students in Israel', 3(1), pp. 9–14.

Silverman, R.H. *et al.* (2017) 'Combined tomography and epithelial thickness mapping for diagnosis of keratoconus', *European journal of ophthalmology*, 27(2), pp. 129–134.

Smolek, M.K. and Klyce, S.D. (1997) 'Current keratoconus detection methods compared with a neural network approach', *Investigative ophthalmology & visual science*, 38(11), pp. 2290–2299.

Street, D.A. *et al.* (1991) 'Lack of Association Between Keratoconus, Mitral Valve Prolapse, and Joint Hypermobility', *Ophthalmology*, 98(2), pp. 170–176.

Twa, M.D. *et al.* (2005) 'Automated decision tree classification of corneal shape', *Optometry and vision science: official publication of the American Academy of Optometry*, 82(12), pp. 1038–1046.

Velázquez-Blázquez, J.S. *et al.* (2020) 'EMKLAS: A New Automatic Scoring System for Early and Mild Keratoconus Detection', *Translational vision science & technology*, 9(2), p. 30.

Venkateswaran, N. *et al.* (2018) 'Optical coherence tomography for ocular surface and corneal diseases: a review', *Eye and vision (London, England)*, 5, p. 13.

View of An Application on Multinomial Logistic Regression Model (no date). Available at: https://pjsor.com/pjsor/article/view/234/206 (Accessed: 28 April 2022).

Weed, K.H. *et al.* (2008) 'The Dundee University Scottish Keratoconus study: demographics, corneal signs, associated diseases, and eye rubbing', *Eye* , 22(4), pp. 534–541.

Wu, W. *et al.* (1996) 'Comparison of regularized discriminant analysis linear discriminant analysis and quadratic discriminant analysis applied to NIR data', *Analytica chimica acta*, 329(3), pp. 257–265.

Yousefi, S. *et al.* (2018) 'Keratoconus severity identification using unsupervised machine learning', *PloS one*, 13(11), p. e0205998.

Zéboulon, P. *et al.* (2020) ‘Corneal Topography Raw Data Classification Using a Convolutional Neural Network’, *American journal of ophthalmology*, 219, pp. 33–39.