UNIVERSITI TEKNOLOGI MARA

DYSLEXIC HANDWRITING DETECTION USING MACHINE LEARNING

ADAM FAHIM BIN JALALUDIN

FEBRUARY 2024

Universiti Teknologi MARA

Dyslexic Handwriting Detection Using Machine Learning

Adam Fahim bin Jalaludin

Report submitted in partial fullfillment
of the requirements for the degree of
Bachelor of Information Systems (Hons.)
(Intelligent Systems Engineering)
College of Computing, Informatics and Media

February 2024

SUPERVISOR'S APPROVAL

This project was prepared under the supervision of the project supervisor, Dr. Norzehan Sakamat. It was submitted to the College of Computing, Informatics and Media and was accepted in partial fulfilment of the requirements for the degree of Bachelor of Information System (Hons.) Intelligent System Engineering.

Approved by:

....
Dr. Norzehan Sakamat
Project Supervisor

January 23, 2024

AUTHOR'S DECLARATION

I certify that this project to which it refers is the product of my own work and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledge in accordance with the standard referring practices of the discipline.

....

Adam Fahim bin Jalaludin

2022970835

January 23, 2024

ABSTRACT

The project aims to develop a machine learning model that can be applied in a system, to detect potential dyslexia through handwriting images. The approach taken was using the LeNet-5 model. The first objective in this study is to identify the features of dyslexia handwriting in images. This is achieved by performing a thorough literature review on the features of dyslexia handwriting, which spans from various journals, research articles, and books published by academics and researchers all over the world. The second objective of the study is to develop a machine learning model that distinguish dyslexic or non-dyslexic samples. This is achieved by using a modified version of the LeNet-5 model, which is a convolutional neural network (CNN) model that is used to classify handwritten digits. The model is modified to be able to classify dyslexic or non-dyslexic samples. The model is then trained using the dataset that was pre-processed. The third objective of the study is to evaluate the performance of the machine learning model. This is achieved by evaluating the performance of the model using the accuracy, precision, recall, and F1-score metrics. The model is also evaluated using the confusion matrix and classification report. The model is also tested using a sample of the dataset that was not used in the training process. However, the model is found to be overfitting, which can be related to the publicly available dataset, which is unoptimized for the model. In terms of recommendations, the first recommendation is to acquire a better dataset. The dataset used in this study is heavily imabalanced, which is not optimized for the model. This is evident in the overfitting of the model, where the model is not able to generalize well. Once the model is optimized, the model can be deployed in a system, where the system can be used to detect dyslexia in children. The system can be used to detect dyslexia in children at a young age, as early diagnosis of dyslexia can allow the children to be taught to read and write properly, and will allow them to be able to perform well in school, on par with their peers.

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to my supervisor, Dr. Norzehan Sakamat, for her unwavering support, invaluable guidance, and expert mentorship throughout the development of this study. The direction and quality of this study were strengthened by their knowledgeable contributions, ongoing support, and constructive feedback.

I would also like to extend my deepest appreciation to my CSP600: Project Formulation and CSP650: Project lecturers Dr. Azliza, Dr. Farah and Dr. Yuzi. Their valuable insights, constructive feedback, and academic expertise have significantly enriched the content and scope of this study.

I am immensely grateful to my friends and family for their unwavering support and understanding. Their love, encouragement, and belief in my abilities have been a constant source of motivation and strength throughout this study.

Lastly, I would like to acknowledge the contributions of all the individuals who have directly or indirectly supported me during this journey. Your presence, encouragement, and words of wisdom have been invaluable in shaping my academic and personal growth.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviations	
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
DCDC2	Doublecortin Domain Containing 2
DL	Deep Learning
DT	Decision Tree
DYX1C1	Dyslexia Susceptibility 1 Candidate Gene 1
EEG	Electroencephalogram
K-NN	K-Nearest Neighbors
LSTM	Long Short-Term Memory
ML	Machine Learning
MRI	Magnetic Resonance Imaging
NB	Naïve Bayes
NN	Neural Network
RF	Random Forest
RNN	Recurrent Neural Network
ROBO1	Roundabout Guidance Receptor 1
SVM	Support Vector Machine

CHAPTER ONE
INTRODUCTION

1.1 Background of Study

Dyslexia, a common learning disorder, has a profound impact on an individual's life, particularly in the context of writing and reading. Studies reveal that its prevalence is estimated to be from as low as 5.37% (Ashraf & Majeed, 2011), to as high as 17.6% (Aboudan et al., 2011). However, these numbers may not fully capture the scope of this issue, as many cases remain undiagnosed due to the subtlety and variance of symptoms.

One of the key challenges faced by individuals with dyslexia is difficulty with written expression. This manifests as problems in spelling accuracy, handwriting legibility, and the speed of writing, all of which contribute to the overall impact on the quality of education and work life. It is not uncommon for these struggles to lead to reduced confidence and self-esteem, compounding the challenges faced.

However, dyslexia are not a determinant of intellectual capability. There are many successful individuals who have been diagnosed with this disorder, demonstrating that with the right tools and strategies, the impact can be managed effectively. These individuals include Steven Spielberg, known for his work on "Jurassic Park", Richard Branson, one of the most successful businessmen in the world with Virgin Group at his helm, and most significantly Steve Jobs, founder of Apple, where his contributions have paved the way for leaps of advancements in computing technology. The importance of early and accurate detection cannot be overstated, as it paves the way for appropriate intervention strategies and accommodations, which are instrumental in enabling individuals to reach their full potential.

In recent years, advancements in technology, particularly machine learning, offer promising new avenues for dyslexia detection. ML models can analyze and learn from patterns that may be too complex or subtle for human detection. In the context of dyslexia, these models can be trained to detect patterns in dyslexic handwriting that could indicate a diagnosis. This could potentially provide a more accessible, less invasive, and cost-effective tool for early detection.

In conclusion, the integration of ML in dyslexia detection presents a powerful opportunity to enhance the lives of those affected by dyslexia. This study aims to explore the potential of machine learning in detecting dyslexic handwriting and how it can be utilized to support individuals with this disorder.

1.2 Problem Statement

Dyslexia, affecting 5% to 17% of the population (Ramli et al., 2020), is a common learning difficulty that impairs an individual's ability to read, write, and spell. Current dyslexia detection methods are time-consuming, expensive, and require specialized personnel, limiting accessibility.

The aim of this study is to explore the use of machine learning techniques for the early and accurate detection of dyslexic handwriting. By investigating various feature extraction and classification algorithms, the study seeks to identify the most effective approach for distinguishing dyslexic handwriting from non-dyslexic handwriting.

The study's outcomes could significantly contribute to earlier interventions and advance the integration of machine learning in healthcare, leading to the development of more sophisticated diagnostic tools.

1.3 Research Questions

The research questions are as follows:

- How do we differentiate between dyslexic and non-dyslexic handwriting?
- How to create a machine learning model that can identify dyslexic and non-dyslexic handwriting?
- How do we view the performances of the machine learning model?

1.4 Research Objectives

The objectives of this research are as follows:

- To identify the features of dyslexia handwriting in images.
- To develop a machine learning model that distinguishes dyslexic or non-dyslexic samples.
- To evaluate the performance of the machine learning model.

1.5 Scope of Study

This study centers on the development of a dyslexic handwriting detection model by leveraging machine learning techniques. We will focus on both supervised and unsupervised learning techniques, exploring algorithms such as artificial neural network, convoluted neural network and recurring neural network to build and optimize the model.

The research will use a dataset of handwriting samples, obtained from Kaggle (D. I. S. Isa, 2022), which contains samples of handwriting images from individuals diagnosed with dyslexia and those without the disorder.

An integral part of the research's scope is to delve into various feature extraction methods that effectively highlight the distinguishing features of dyslexic handwriting. Further, we aim to test the efficacy of different classification algorithms in accurately segregating dyslexic handwriting from non-dyslexic handwriting.

The proposed model's performance will be rigorously evaluated using standard evaluation metrics. We will measure the accuracy, precision, recall, and F1 score of the model and compare the results with existing methods, if any, to determine the relative effectiveness of our approach.

1.6 Significance of Research

The significance of the research falls in these aspects:

1.6.1 Early Identification

Dyslexia's impact reaches far beyond academic performance; it also significantly influences individuals' self-esteem and mental health. Early identification of dyslexia is crucial as it allows for the implementation of targeted educational strategies and accommodations. This early intervention can help mitigate some of the challenges dyslexic individuals face, enabling them to better realize their potential. It can also alleviate some of the psychological stress associated with the condition, as understanding the root cause of their difficulties can provide a sense of relief and pave the way for the development of coping strategies.

1.6.2 Reliable and Cost-Effective Screening

Currently, dyslexia screening processes can be time-consuming, expensive, and require specialized personnel to conduct and interpret assessments. This makes it particularly challenging in resource-limited settings. A machine learning model that can accurately detect dyslexia through handwriting analysis has the potential to offer a more reliable, accessible, and cost-effective means of screening large populations for dyslexia. It could be particularly impactful in under-resourced areas where access to specialists is limited, ensuring that more individuals receive the support they need.

1.6.3 Assisting Educators and Clinicians

The model's ability to accurately identify dyslexic handwriting could serve as a powerful tool for educators and clinicians. It could assist in pinpointing students or patients who might be struggling due to undiagnosed dyslexia, facilitating early and targeted support. Furthermore, this tool could aid in tracking the progress of individuals with dyslexia over time, offering insights into the effectiveness of interventions and allowing for necessary adjustments.

1.6.4 Contributing to Knowledge

This research stands at the intersection of technology and healthcare, more specifically the application of machine learning in dyslexia detection. By exploring how machine learning techniques can be employed to detect dyslexic handwriting, this study will contribute significantly to the existing body of knowledge in both fields. It will add new dimensions to our understanding of dyslexia, and how technology can be harnessed to aid in its diagnosis and management. Furthermore, the methodologies and findings of this research could potentially be applied to other learning disorders, widening its scope and significance.

1.7 Summary

To summarize, this chapter provides an initial overview of the research, and heavily emphasizes on the domains and the techniques that will be implemented in the study. The problem statement taps into the issue, which leads to the research questions. The research too finds its scope of study that leads towards its significance.

CHAPTER TWO
LITERATURE REVIEW

2.1 Dyslexia

2.1.1 Definition and Prevalence

Dyslexia can be defined in so many ways. According to Merriam-Webster, dyslexia is defined as "a variable often familial learning disability involving difficulties in acquiring and processing language that is typically manifested by a lack of proficiency in reading, spelling, and writing" (“Dyslexia, n.”, 2023). The word’s history dates back to 1888, where it is initially defined as an “impairment in the ability to read due to a brain injury”. It is borrowed from French and German, where both it is written as *dyslexie*.

Estimates for the prevalence of dyslexia suggest that it affects around 5 to 17% (Ramli et al., 2020) of the global population, marking it as one of the most common learning disorders. However, these statistics are subject to considerable variance due to the numerous definitions of dyslexia, the diverse methodologies employed for its diagnosis, and the sampling of different populations in research studies.

Despite these disparities in prevalence rates, what remains consistent is the universal recognition of dyslexia’s significant impact on individuals’ educational journeys, their professional progress, and even their day-to-day lives. It is also evident that dyslexia do not discriminate between genders alike, according to Guerin et al. (1993).

2.1.2 Causes and Risk Factors

The causes of dyslexia are complex and multifaceted, involving an intricate interplay of genetic, neurobiological, and environmental factors. In fact, some of the causes are yet to be fully understood.

Genetic researches have identified several genomic regions that have been linked to dyslexia, including chromosomes 6 and 18 (Francks et al., 2002; Schumacher et al., 2007). Advancements in research has also found identifying candidate genes for dyslexia, which includes Dyslexia Susceptibility 1 Candidate Gene

1 (DYX1C1), Roundabout Guidance Receptor 1 (ROBO1), KIAA0319, and Doublecortin Domain Containing 2 (DCDC2) (Paracchini et al., 2007).

From a neurobiological standpoint, structural and functional differences have been identified in the brains of individuals with dyslexia. Neuroimaging studies have identified differences in brain structure and function in individuals with dyslexia, including disruptions in left hemisphere posterior brain systems and increased reliance on frontal lobe circuits (Kearns et al., 2019; Norton et al., 2015).

Another research has found that dyslexia is highly heritable and displays polygenic transmission, and adult neuroimaging studies have found structural, functional, and physiological changes in the parieto-occipital and occipito-temporal regions, and in the inferior frontal gyrus, in adults with dyslexia (Soriano-Ferrer & Martínez, 2017).

Another research also reviewed evidence of autopsy and structural imaging studies and found consistent evidence of symmetry of the planum temporale, thalamus, and cortical malformations in individuals with dyslexia (Wajuihian, 2011).

Environmental factors further contribute to the manifestation and development of dyslexia. found that there is a substantial social-cultural bias in the delineation of literacy skills and in the definitions of reading disabilities, and suggests that phonological deficits should be emphasized as the core component in defining dyslexia (Samuelsson & Lundberg, 2003).

2.1.3 Characteristics

Dyslexia encompasses a variety of symptoms, which generally become apparent once a child starts school and is confronted with the challenges of learning to read and write. Common characteristics of dyslexia include difficulty with phonological skills, low accuracy and fluency of reading, poor spelling, and/or rapid visual-verbal responding (Roitsch & Watson, 2019).

These signs can vary in intensity and nature among individuals, contributing to the spectrum of dyslexia manifestations. Additionally, dyslexia can extend to erratic eye movements during reading and other tasks (Pavlidis, 1981).

2.1.4 Impact on Writing and Handwriting

Dyslexia's impact extends noticeably to an individual's handwriting, with various studies indicating that individuals with dyslexia often grapple with handwriting fluency, neatness, and speed. A study found that children with dyslexia struggle with the graphomotor aspects of writing and are more impacted by the graphic complexity of words than typically developing children (Gosse & Reybroeck, 2020).

Another study has also found that spelling ability influences the rate of handwriting production in children with dyslexia, and that productivity relies on spelling capabilities (Sumner et al., 2014).

Writing speed also varies within children with dyslexia, as a study found that handwriting speed in Chinese children with dyslexia is related to deficits in rapid automatic naming, saccadic efficiency, and visual-motor integration (Cheng-Lai et al., 2013).

Besides that, spelling deficits associated with dyslexia affect the dynamics of the interaction between central and peripheral processes and the level of anticipation that can be observed in word spelling in the context of a sentence to dictation task (Suárez-Coalla et al., 2020).

2.1.5 Current Diagnostic Methods

The diagnosis of dyslexia usually involves a holistic evaluation of the individual's academic performance, cognitive and linguistic skills, and developmental history.

One study discusses the diagnostic assessment of dyslexia, which consists of standardized reading and spelling tests, evaluation of psychological state, and additional information from parents and teachers (Schulte-Körne, 2010). The diagnostic procedure may incorporate a battery of tests to assess reading, spelling, and writing skills.

However, in recent times, studies have discovered advancements in diagnostic procedures. One research proposes a diagnostic method based on involuntary neurophysiological responses to auditory stimuli, using electroencephalogram signals to analyze temporal behavior and spectral content (Ortiz et al., 2020).

Another study has reviewed different technology-based approaches for dyslexia detection, including eye-tracking and Electroencephalogram (EEG) devices, and statistical or machine learning algorithms (Jankovic, 2022).

2.1.6 Interventions and Treatments

Even though there is no known cure for dyslexia, early assessment and intervention, paired with supportive teaching strategies, can significantly improve success rates for individuals with dyslexia. These interventions often include multi-sensory, structured language programs that explicitly teach phonics, morphology, syntax, and semantics.

A study has found that multisensory, phonological, and cognitive training methods can be used to improve literacy and cognitive deficits among children with dyslexia in Malaysia (Yuzaidey et al., 2017).

Besides that, in Brazil, there are four key themes of interventions for dyslexia, which includes phonological-based intervention, computerized technology, auditory processing training, and visuomotor skills training (Signor et al., 2020).

However, on a national level, governments have to make strides in dyslexia intervention. Study suggests that early identification of children at risk of dyslexia followed by evidence-based interventions is a realistic aim for practitioners and policymakers (Snowling, 2012).

2.1.7 Challenges and Limitations of Current Practices

Current practices in dyslexia diagnosis and intervention, despite their effectiveness, present several challenges.

For one, the cognitive approach alone is not sufficient to address dyslexia, and that a combination of cognitive psychology, connectionism, and behaviorism is necessary (Tønnessen, 1999).

A local study also finds that there is a lack of comprehensive studies that combine interventions for both cognitive functions and literacy deficits (Yuzaidey et al., 2017).

Besides that, while appropriate instruction can help at-risk readers become accurate and fluent, intensive remedial interventions have been less effective in closing the fluency gap. (Alexander & Slinger-Constant, 2004).

2.2 Machine Learning

2.2.1 Definition and Overview

By definition, machine learning is the process by which a computer is able to improve its own performance by continuously incorporating new data into an existing statistical model.

It can also be defined as "the branch of computer science dealing with the creation and use of computer software that employs machine learning" (“Machine Learning, n.”, 2023).

However, researchers have a differing opinion on the definition. One has defined machine learning as a branch of computational algorithms that emulate human intelligence by learning from the environment (Naqa & Murphy, 2015), while another describes machine learning as a study of computational methods for improving performance by mechanizing the acquisition of knowledge from experience (Langley & Simon, 1995).

Besides that, machine learning has also been described as a way to address problems where phenomena are changing rapidly, or where applications need to be customized for each user separately (Dietterich, 1996).

Commonly, machine learning is classified into supervised, unsupervised, and reinforcement learning. However, there can also be semi-supervised, transduction and learning to learn algorithms (Oladipupo, 2010).

Supervised learning is typically used for prediction tasks, where a model is trained using labeled data meanwhile unsupervised learning involves finding hidden patterns or intrinsic structures from unlabeled data. Reinforcement learning enables a software agent to learn in an interactive environment by trial and error.

Within the broader machine learning landscape, a particularly noteworthy development is the emergence of deep learning. Deep learning is described as a network of nodes and edges that resemble the biological communication of brain neurons (Dinov, 2018).

Deep learning algorithms, often based on neural networks, can model high-level abstractions in data, thereby providing enhanced predictive accuracy in tasks such as image and speech recognition, natural language processing, and more.

2.2.2 Applications

Machine learning's potential for pattern recognition and predictive analysis has led to a broad array of applications across multiple domains. This includes cyber security, healthcare, and intelligent transportation systems, to name a few. (Jhaveri et al., 2022).

Socially, machine learning has made a big impact too. Social media has utilised machine learning to analyze large amounts of data, where for example Twitter, where it has been used to identify real and fake tweets (Arora & Gabrani, 2018).

The reality is, machine learning is almost applied everywhere. A mobile phone is now capable of contextual search, object recognition, intelligent control, speech recognition, natural language processing, and computer vision, which is all powered by machine learning itself. As time goes on, the importance keeps getting even more significant as it slowly integrates to everyone's daily lives.

2.2.3 Machine Learning in Healthcare

Healthcare stands out as a domain particularly poised to benefit from the advancements in machine learning. machine learning has been applied in a lot of areas, including medical imaging, natural language processing of medical documents, genetic information, disease prediction, disease detection, and personalized healthcare (Mana et al., 2022).

Another key applications of ML in healthcare is in the field of medical imaging. In terms of radiology, it has shown that machine learning has potential to improve various aspects of radiology, including detection and interpretation of findings, post-processing, and radiology reporting (Choy et al., 2018; Erickson et al., 2017).

Another significant area of application is in predictive healthcare. machine learning algorithms Naïve Bayes, Decision Tree, Random Forest, and K-Nearest Neighbors are being utilised to predict diseases based on symptoms and large datasets of medical procedures (Garg & Bansal, 2023).

2.2.4 Feature Extraction and Selection

Feature extraction and selection are crucial processes in any machine learning model. They involve identifying and selecting the most relevant information from raw data to be used for machine learning.

In the context of healthcare, feature extraction and selection could be applied to a variety of data, such as medical imaging, genomic data, or patient records. For instance, in Magnetic Resonance Imaging (MRI) feature extraction approach utilizes spatial filters, edge detection algorithms, and wavelet transform-based image fusion (Udomhunsakul & Wongsita, 2004).

In handwriting analysis, features could refer to the stroke width, stroke length, speed, pressure, and slant among others. These characteristics can be critical in distinguishing between dyslexic and non-dyslexic handwriting. Selecting the right features is crucial as it directly impacts the performance of the machine learning model (‘Arif et al., 2015).

2.2.5 Machine Learning Algorithms

Machine learning algorithms are the crux of any machine learning model, determining how it learns from data and makes predictions or decisions. The choice of algorithm often depends on the nature of the task and the data at hand, with a wide range of algorithms available each with unique strengths and limitations.

Some of machine learning algorithms include K-Nearest Neighbors, Naïve Bayes, Support Vector Machine, Decision Tree, and others. These have been widely used in various fields, including healthcare, due to their interpretability and robustness (Ray, 2019).

In recent years, deep learning algorithms, such as convolutional neural network and recurrent neural network, have gained popularity. These algorithms can learn complex patterns in data, making them highly effective for tasks that involve large amounts of data or that require the extraction of intricate features (Pamina & Raja, 2019).

2.2.6 Performance Evaluation

Evaluating the performance of machine learning models is essential to verify their effectiveness and reliability. Metrics such as accuracy, precision, recall, and F1 score are commonly used for this purpose. However, the choice of metrics should align with the problem at hand, as different tasks may require prioritizing different aspects of the model's performance.

In addition to these metrics, consideration of the model's robustness against variations in the data is crucial. This could involve testing the model under different conditions, or using different subsets of the data, to ensure that it performs consistently and does not overly rely on specific characteristics of the training data.

There are a bunch of methods to evaluate robustness, including prediction consistency between source and target data in the neighborhood of the source samples (Shi et al., 2019) and a framework for evaluating robustness to changes in setting or population using a single, fixed evaluation dataset (Subbaswamy et al., 2020).

Another important aspect of performance evaluation is the interpretability of the model. While some complex models, like deep learning models, may achieve high performance, they are often regarded as 'black boxes' due to their lack of interpretability (Zhang & Zhu, 2018). On the other hand, simpler models may offer more interpretability, but at the cost of performance. Striking a balance between these factors is an ongoing challenge in the field of machine learning.

2.2.7 Limitations and Challenges

Despite the promising advancements, machine learning is not without its challenges and limitations. The quality and diversity of data is one of the primary concerns. machine learning models thrive on large amounts of high-quality, representative data. However, in situations where data is scarce, imbalanced, or inherently biased, the model's performance may be compromised (Lum, 2017).

Another concern is overfitting, where the model learns the training data too well, to the point that it performs poorly on unseen data (Ying, 2019). This issue highlights the importance of proper model validation and testing methodologies.

Data privacy is a crucial challenge across many domains that use machine learning. Ensuring the protection and appropriate use of sensitive data is a complex issue that often requires navigating regulatory and ethical considerations (Strobel & Shokri, 2022).

2.3 LeNet-5

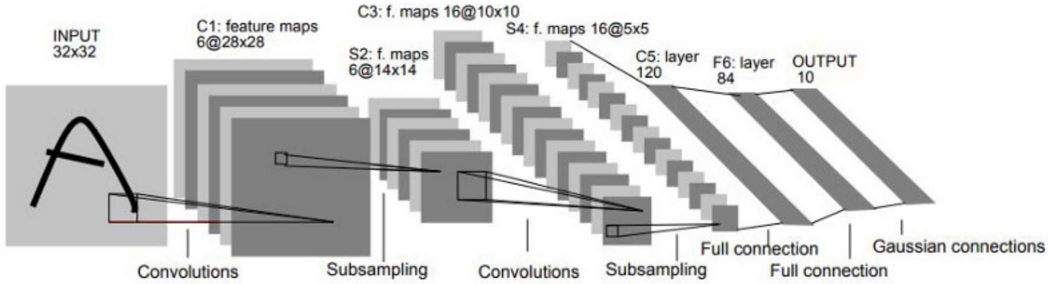

Figure 2.1 LeNet-5 architecture
Lecun et al. (1998)

Le-Net 5 is a convolutional neural network (CNN) architecture proposed by Lecun et al. (1998) for handwritten and machine-printed character recognition. It is a simple convolutional neural network and is widely used for image classification.

The neural network feautures a 7-level convolutional neural network. The input to the network is a 32x32 pixel image. The image is passed through the network and the output is a probability distribution over the 10 classes of digits (0-9). The architecture consists of two sets of convolutional and average pooling layers, followed by a flattening convolutional layer, then two fully-connected layers, and finally a softmax classifier (Lecun et al., 1998).

LeNet-5 has been used in several researches, such as in I. S. Isa et al. (2021) and I. S. Isa et al. (2019b).

2.4 Related Works

2.4.1 Existing Research

Over the past few years, several research efforts have been dedicated to the application of machine learning for dyslexia. These studies have largely focused on aspects such as speech and language patterns, cognitive assessments, and reading behaviors.

However, the use of handwriting as a potential source of data for dyslexia detection has been less explored. Early efforts have indicated the promise of this approach, leveraging machine learning techniques to identify unique features in dyslexic handwriting and assess their correlation with dyslexia.

Despite the potential, much of the existing research is in its early stages, and many studies have been conducted on a small scale or under controlled conditions. Thus, while the preliminary results are encouraging, there is still much to understand about the effectiveness of these techniques when applied to more diverse and larger-scale real-world contexts.

This emerging field presents an opportunity for new research to contribute to both the understanding of dyslexia through handwriting analysis and the development of effective machine learning techniques for dyslexic handwriting detection.

2.4.2 Data Used in Previous Studies

Previous studies investigating dyslexic handwriting detection with machine learning have made use of various types of data. In many cases, these include handwriting samples collected from both dyslexic and non-dyslexic individuals, often children.

Handwriting samples can come in various forms, such as isolated characters, words, or continuous texts. Each of these forms provides different levels of complexity and contextual information, which could influence the detection performance.

In addition to the handwriting samples themselves, some studies also utilize supplementary data, such as demographic information, educational history, and details of dyslexia severity or related symptoms. This additional information could help in contextualizing the handwriting features and improving the prediction models.

The choice of data, its collection, and pre-processing significantly impact the results of the machine learning models. Therefore, understanding the types and characteristics of data used in existing studies can provide valuable insights for future research.

Furthermore, some datasets, such as I. S. Isa et al. (2021), made their datasets publically available through Kaggle, a data science platform and online community of data scientists and machine learning practitioners made by Google.

2.4.3 Feature Extraction in Handwriting Analysis

Feature extraction plays a crucial role in the analysis of handwriting for dyslexia detection. This process involves identifying and quantifying characteristics in the handwriting samples that could potentially differentiate between dyslexic and non-dyslexic individuals.

Features extracted from handwriting can be categorized into static and dynamic features. Static features are derived from the visual appearance of the written text and include aspects such as letter shape, size, slant, and spacing.

On the other hand, dynamic features capture the process of writing and may include pen pressure, velocity, and acceleration, among others. In (Lam et al., 2011), the pressure of the pen was captured using the Wacom digitizer, allowing for more in-depth data regarding the handwriting.

Machine learning methods, particularly deep learning, have also been used to automatically extract features from handwriting samples. These techniques can learn to identify complex patterns in the data without the need for manual feature engineering, potentially uncovering novel features of dyslexic handwriting.

I. S. Isa et al. (2021) has used the convolutional layer work as the feature extractor where the layer studies the feature representations of input images. Then, the trainable convolutional kernel will regulate its kernel weights automatically in backpropagation training process.

However, the choice of features and extraction methods can significantly influence the performance of dyslexic handwriting detection. Hence, understanding the approaches used in previous studies can help guide the selection and development of feature extraction techniques in future research.

2.4.4 Techniques, Performance and Findings

Various machine learning techniques have been employed in previous studies to analyze handwriting for dyslexia detection. These range from traditional machine learning methods to more recent deep learning models.

Deep learning models, including artificial neural network and convolutional neural network, have also been applied to this task. artificial neural networks, in this context has been used in I. S. Isa et al. (2019b). In contrast, convolutional neural networks are utilised in I. S. Isa et al. (2021), in the form of LeNet-5.

Each of these techniques has its strengths and limitations and may be more or less suitable depending on the specific characteristics of the data and task. Therefore, a thorough understanding of the machine learning techniques used in existing studies is crucial for informing the choice of methods in future research.

The performance of machine learning techniques in detecting dyslexic handwriting varies widely across different studies, largely due to the differences in data used, features extracted, and the specific models applied.

Many of these studies report varying results. These findings suggest that machine learning has potential in identifying dyslexic handwriting and could aid in early and more accurate diagnosis.

In I. S. Isa et al. (2019a), the accuracy of the ANN model is reported to be at 73.33 percent. Later in 2021, having switched to the CNN model, I. S. Isa et al. (2021) has reported training accuracies as high as 0.985 on the CNN-1 model, while other models received slightly lower accuracies of 0.98 and 0.968.

On the other hand, through the random forest approach, Richard and Serrurier (2020) has achieved 90 percent accuracy using the model. However, not all studies has achieved high accuracy.

A study from Drotár and Dobeš (2020) has obtained only a 79.5 percent prediction accuracy using the AdaBoost algorithm. Furthermore, Spoon et al. (2019) has reported an accuracy of only 55.7 percent in its research.

It is also important to note that many of the reported studies have been conducted under controlled conditions and on small, often homogeneous datasets. Therefore, it remains to be seen whether these results can be generalized to more diverse and larger populations.

Considering the current state of research, while the initial results are encouraging, there is still much work needed to validate and refine these techniques for practical use. More comprehensive studies, involving larger and more diverse samples, are required to establish the reliability and applicability of these techniques.

2.4.5 Challenges and Limitations

While the use of machine learning for dyslexic handwriting detection presents exciting opportunities, it also comes with several challenges and limitations that have been identified in previous research.

One of the major challenges lies in the data itself. Handwriting can vary widely between individuals and can be influenced by numerous factors such as age, education, and cultural background. This variability makes it challenging to develop models that are generalizable across diverse populations.

Additionally, even though getting data sets are getting easier day by day, especially through open platforms, such as Kaggle, obtaining large quantities of labeled handwriting data from dyslexic individuals for training machine learning models can be difficult due to privacy concerns and the excruciatingly time-consuming nature of data collection and labeling.

From a technical perspective, the choice of features and machine learning algorithms can significantly impact the performance of the models. Yet, there is still currently no consensus at all on the optimal features or models for dyslexic handwriting detection, making it a challenging area of research.

Furthermore, while deep learning methods have shown promise in this field, they come with their own set of challenges. They typically require large amounts of data, significant computational resources, and often times their complex models are severely lacking in interpretability, which is crucial in healthcare applications.

Despite these challenges, all of the researches done so far has laid a solid foundation for future work in this area. By understanding and addressing these limitations, future research can contribute to the development of more effective and practical solutions for dyslexic handwriting detection.

2.4.6 Future Directions and Opportunities

As the research in dyslexic handwriting detection using machine learning continues to evolve, several promising directions and opportunities have been identified.

One such direction is the further exploration of deep learning techniques. Despite the challenges, their potential in automatically extracting complex features from handwriting samples and handling large-scale datasets make them a promising area for future research. There were also suggestions of trying out unsupervised approaches, which includes clustering to group the data into clusters that could then be inspected by human experts (Spoon et al., 2019).

Moreover, more comprehensive and diverse datasets can vastly enhance the generalizability and reliability of the models. Efforts to create larger, more diverse datasets that encompass a range of handwriting styles, demographic backgrounds, and dyslexia severities are needed. This claim has been generally voiced by many studies, that includes studies from I. S. Isa et al. (2019a), Richard and Serrurier (2020) and Usman et al. (2021).

Lastly, more research into the interpretability of machine learning models can also be valuable. Developing models that not only perform well, but also provide insights into their decision-making process can enhance trust and adoption of these techniques in real-world settings.

In conclusion, while the field of dyslexic handwriting detection using machine learning is still in its early stages, the initial findings are promising and point towards a future with more accurate, timely, and accessible dyslexia detection tools.

2.5 Summary

To summarise everything in Literature Review, this study reviews dyslexia and its impact on handwriting. Besides that, this study discovers machine learning and its use in dyslexic handwriting detection, where then a machine learning model, which is LeNet-5, is used to perform dyslexic handwriting detection.

A comprehensive review of previously done researches relevant to dyslexic handwriting detection using machine learning has been done. Finally, the spirit of research itself, there will definitely always be new discoveries to be found in the study.

CHAPTER THREE
RESEARCH METHODOLOGY

This chapter details the strategy adopted in this study to achieve the study objectives. The aim of the research method section is to explain the techniques and processes employed in the investigation, along with their operational mechanisms.

3.1 Research Framework

The Research Methodology is the orderly process and strategy employed in conducting research, gathering information, inspecting material, and drawing conclusions. It encompasses a range of methodologies, processes, tools, and tactics that researchers use to meet their research objectives or questions.

The methodology section of the research provides a rundown of the steps and procedures followed in the study, encompassing the research design, sampling strategies, data analysis methods, and ethical concerns. It offers a systematic framework for conducting and presenting research across diverse academic disciplines, assisting in the maintenance of rigor, validity, and reliability in the research results. The research framework for this study is designed to follow a systematic and iterative process.

The study has been conducted in multiple stages, progressing from initial preliminary studies, to acquisition of data, pre-processing of data, the development of the machine learning model, proceeded by the evaluation of the model and finally the documentation. The figure below presents a visualization of the research.

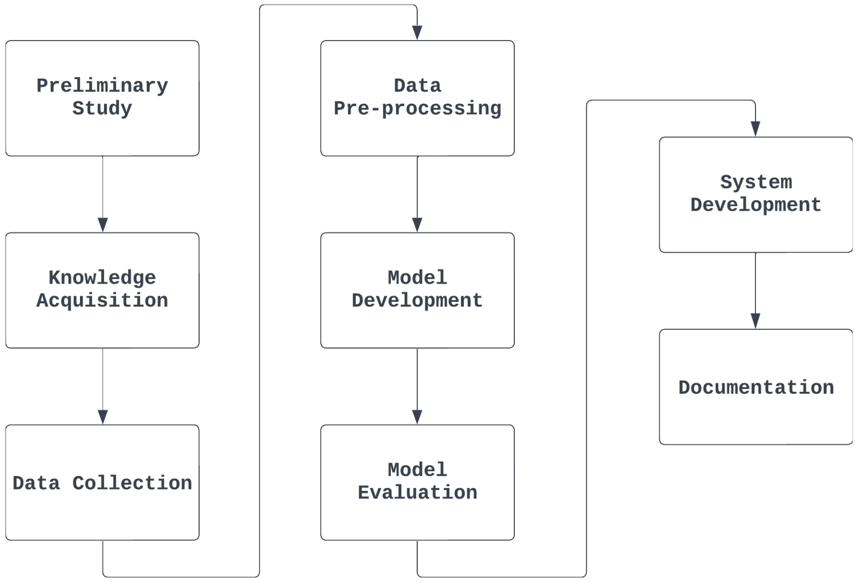

Figure 3.1 Methodology Phases

The following table presents a compilation of various research inquiries that have been raised to address the problem statement outlined in Chapter 1.

Table 3.1
Research Questions.

Questions	**Description**
How do we differentiate between dyslexic and non-dyslexic handwriting?	To identify the features of dyslexia handwriting in images.
How to create machine learning model that can identify dyslexic and non-dyslexic handwriting?	To develop a machine learning model to distinguish dyslexic or non-dyslexic samples.
How do we view the performances of the machine learning model?	To evaluate the performance of the machine learning model.

The problem statement discussed in Chapter 1 has resulted in several research questions. One of these questions explores the performance evaluation of a machine learning model developed to differentiate between dyslexic and non-dyslexic handwriting samples, using metrics such as accuracy, precision, recall, and F1-score. Another research question pertains to conducting a literature review on the detection of dyslexic handwriting, aimed at gaining a comprehensive understanding of dyslexia's impact on handwriting and identifying gaps in existing methodologies.

In Table 3.2, the table presents the methodology framework which includes the phases, activities, and deliverables for this report based on the study objectives.

Table 3.2
Methodology Framework.

OBJECTIVES	**PHASES**	**ACTIVITIES**	**DELIVERABLES**
To identify the features of dyslexia handwriting in images.	Preliminary Study	Examine and comprehend the study thoroughly. Pinpoint the research goal, determine its scope, and highlight the importance of the study.	Chapter 1: Background of Study, Problem Statement, Research Objectives, Scope of Study, Significance of Research
	Knowledge Acquisition	Review existing research on dyslexia, and its effect on handwriting. Discover relevant machine learning algorithms, and its tools.	Chapter 2: Literature Review
	Data Collection	Obtain data for the research from available resources.	Datasets to be used for dyslexic handwriting recognition.
To develop machine learning models to distinguish dyslexic or non-dyslexic samples.	Data Pre-processing	Process the raw datasets, and go through pre-processing procedures.	Cleaned data.
	Model Development	Apply the LeNet-5 model	LeNet-5 model to be evaluated.

Table 3.2 – continued from previous page

OBJECTIVES	PHASES	ACTIVITIES	DELIVERABLES
To evaluate the performance of the machine learning model.	Model Testing and Evaluation	Evaluate the performance of LeNet-5	Performance of the LeNet-5 model. Result comparison and analysis.
	System Development	Develop the system and present a usable interface.	System that can be used.
	Documentation	Document every step and process of the research.	Completed document to use as a reference.

3.2 Preliminary Study

Preliminary Study represents the initial phase where it involves a comprehensive exploration and understanding of the study's foundational elements, and in the context of this study, it primarily involved dyslexia and machine learning. This phase involved conducting domain and title searches for interested parties. To come up with a suitable and good title, it also entails having frequent interactions with supervisor(s). This is a safety measure done to prevent titles irrelevant to the subject matter being picked up. It is followed up by an in-depth research, carried out by reading credible articles and research papers with the aim of grasping and identifying the problem in the chosen domain.

3.2.1 Problem Identification

The problem identified in this study is that dyslexic handwriting detection currently relies on manual evaluation, which can be time-consuming and subjective. This approach results in potential delays in identifying dyslexic individuals and providing them with appropriate support. Therefore, there was a demand to implement machine learning techniques to accurately detect dyslexic handwriting.

3.2.2 Domains and Technique Understanding

3.2.2.1 Dyslexia

Dyslexia is a neuro-developmental condition that presents challenges in acquiring reading skills, even with standard teaching methods, sufficient intelligence, and a well-rounded socio-cultural environment. According to Wajuihian and Naidoo (2011), it is the most prevalent form of learning disorder, and its impact on reading abilities can significantly hinder a child's academic performance.

3.2.2.2 Machine Learning

Machine learning is a field that encompasses automated computational processes where logical or binary operations are utilized to enable computers to acquire knowledge and skills by studying numerous instances or examples. By employing sophisticated algorithms, machine learning systems can analyze and comprehend patterns, relationships, and data in order to enhance their performance and make informed predictions or decisions. This transformative technology, according to Fulkerson et al. (1995), has revolutionized various industries, especially in healthcare, by enabling computers to learn from experience and adapt their behavior accordingly.

3.3 Knowledge Acquisition

To gather a thorough understanding of the study, firstly knowledge must be acquired. Knowledge acquisition is the stage where information is gathered primarily through review methods. It refers to the process of deriving knowledge and information from a multitude of sources to enhance understanding or facilitate decision-making. This acquisition of knowledge was accomplished through various means such as articles, interviews, surveys, literature reviews, and experimental research.

To enhance the process of knowledge acquisition, a systematic approach had been developed, which involved reviewing previous research and studies. This approach ensured a structured method of acquiring knowledge by categorizing research topics according to the following table. Each domain and technique within these categories has its own set of concepts that need to be read and analyzed in order to grasp the fundamental principles of each specific area.

Once the fundamental concepts of each domain or technique were understood, the correlation and analysis of combining two domains or techniques come into play. This approach allowed for a deeper exploration of the subject matter by examining the connections and interactions between different domains or techniques.

For this project, a thorough study of dyslexia and machine learning was conducted. This involves performing a literature review on dyslexia, and its effect on handwriting, as well as discovering relevant machine learning algorithms, and its tools. The knowledge is acquired from online platforms, most notably Google Scholar, Semantic Scholar, and IEEE Xplore. As for IEEE Xplore, it is made accessible through UiTM's library website.

Related articles and research papers are then saved in Mendeley's reference manager, which is then used to organize the references, and to cite them in the documentation. In this context, the .bib file is exported from Mendeley, and is used in LaTeX to cite the references.

3.4 Data Collection

Data collection is a systematic process of gathering relevant information and data points from various sources or individuals to aid in research, analysis, or in decision-making. This process involves identifying the necessary data, designing data collection methods such as surveys, interviews, observations, or experiments, and obtaining the data through structured or unstructured techniques.

The nature of the data gathered can be either qualitative or quantitative, depending on the research objectives, and can originate from primary sources like direct participant interviews, or secondary sources such as pre-existing datasets, literature, or publicly available data.

In this study, a Kaggle dataset was being used for data collection, where there are a wide collection of datasets relevant to a sorts of studies being done. For this research, a dataset released from D. I. S. Isa (2022) has been used. The dataset was gathered from 3 sources, which were from Patricia A. Flanagan (2016) for uppercase letters, while using Sachin Patel (2017) for lowercase letters and several testing datasets were obtained dyslexic kids of Seberang Jaya primary school, Penang, Malaysia.

Figure 3.2 Kaggle logo

3.5 Data Pre-processing

The phase in focus is data pre-processing, a critical step used to transform raw data into an effective and efficient format for further use. Raw data from the dataset need to be processed prior to their application in training and testing procedures.

Initially, the dataset used was designed to classify three types of handwriting, which are normal, reversal, and corrected handwriting. The tree structure of the dataset can be seen in Figure 3.3. This dataset contains a total of 78,275 images for normal class while for reversal are 52,196 images and for corrected are 8,029 images, which totals up to 138,500 images.

- Dataset
  - Train
    - Normal
    - Corrected
    - Reversal
  - Test
    - Normal
    - Corrected
    - Reversal

Figure 3.3 Tree structure of the dataset

The dataset was then modified to only classify normal and potentially dyslexic handwriting, which consists of the combination of the correction and reversal class. The tree structure of the modified dataset can be seen in Figure 3.4. With that, the dataset contains a total of 78,275 images for normal class and a total of 60,225 images for potentially dyslexic class.

- Dataset
  - Train
    - Normal
    - Potential Dyslexia
  - Test
    - Normal
    - Potential Dyslexia

Figure 3.4 Tree structure of the modified dataset

As the dataset has been already pre-processed, the only pre-processing step taken was to apply the `rescale` function, in order to normalize the pixel values in the images to the range [0,1]. It is believed that it helps the model converge faster and prevents the gradients from becoming too large. Furthermore, as the training and testing dataset are already separated, no splitting of the dataset was required, with the exception of the validation dataset, which was 10% of the testing dataset. The dataset was then loaded into the model, and the model was trained using the dataset.

3.6 Model Development

For model development, Python was chosen as the preferred programming language. The reasoning behind this is as to utilise TensorFlow as the main framework. TensorFlow is an open-source software library for machine learning, which was developed by the Google Brain team. It is a symbolic math library, and is also used for machine learning applications such as neural networks.

Figure 3.5 TensorFlow logo

This is assisted by using Keras, a high-level API that is built on top of TensorFlow. Keras is a deep learning API written in Python, which runs on top of TensorFlow, CNTK, or Theano. It was developed with a focus on enabling fast experimentation, and is designed to be user-friendly, modular, and extensible.

To smoothen the process of developing the model, the development is done locally on a Windows Subsystem for Linux 2 (WSL2) environment, running Ubuntu 22.04, codenamed ‘Jammy Jellyfish’. With that, it opens up for the use of GPU acceleration in TensorFlow, which is provided by NVIDIA CUDA. The specifications of the hardware used can be seen in Table 3.3.

Table 3.3
Hardware Specifications.

Hardware	**Description**
CPU	AMD Ryzen 7 3700X 8-Core Processor
GPU	NVIDIA GeForce RTX 3060 12GB
RAM	16GB 3200MHz DDR4

In this study, a modified version of LeNet-5 has been used, a convolutional neural network (CNN) architecture that was developed by Lecun et al. (1998). The

LeNet-5 architecture consists of 7 layers, which are 2 convolutional layers, 2 subsampling layers, 2 fully connected layers, and 1 output layer.

However, for this study, the architecture has been modified to include BatchNormalization, where it improves accuracy, followed by applying MaxPooling2D, to produce smaller feature sizes. Besides that, a dropout layer is included to prevent overfitting. The architecture of the modified LeNet-5 can be seen in Figure 3.6.

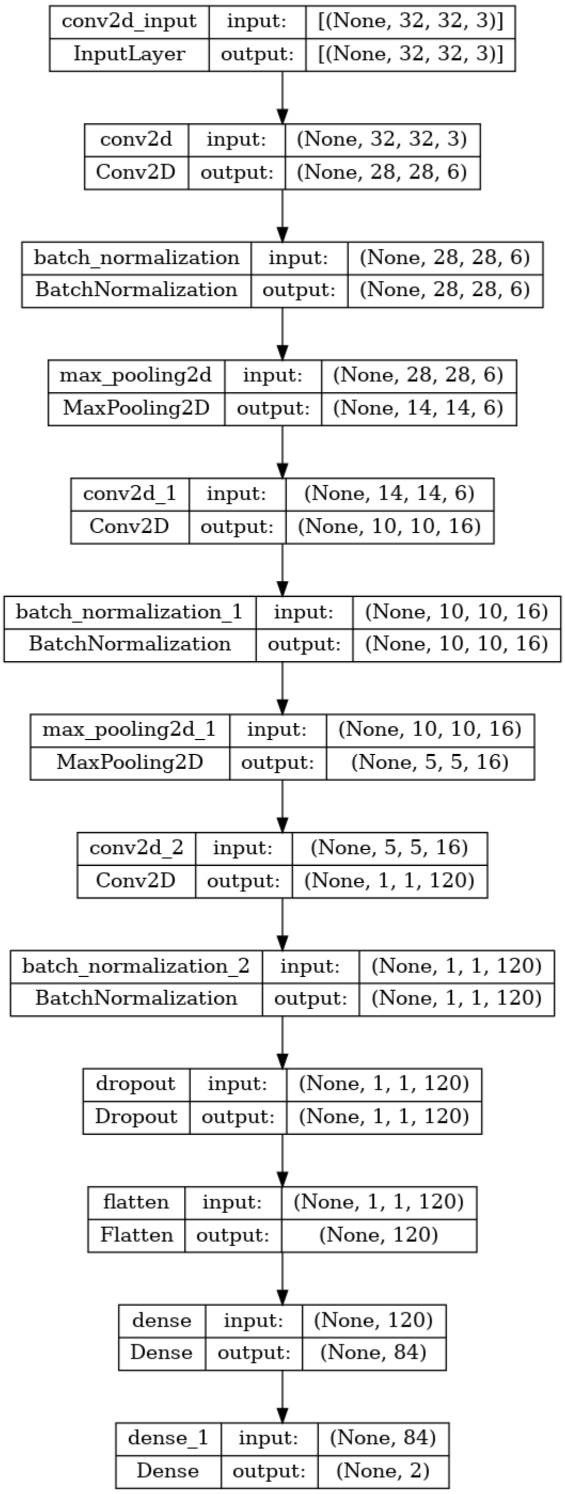

Figure 3.6 Modified LeNet-5 architecture

The model takes an input size of 32x32, which is the size of the images in the dataset. The breakdown of the layers can be seen in Table 3.4.

Table 3.4
Detailed description of the LeNet-5 model.

Layer	**Description**
Input	Input layer, takes an input size of 32x32.
Conv2D	This initial convolutional layer applies 6 filters of size 3x3, extracting basic features from the input images. It has 456 trainable parameters.
BatchNormalization	This layer normalizes the activations of the first convolutional layer, stabilizing training and potentially accelerating convergence.
MaxPooling2D	This layer downsamples the spatial dimensions of the feature maps by a factor of 2, reducing computation and promoting invariance to small translations.
Conv2D	The second convolutional layer applies 16 filters of size 3x3, capturing more complex features. It has 2,416 trainable parameters.
BatchNormalization	This layer normalizes the activations of the second convolutional layer, similarly aiding training.
MaxPooling2D	This layer further downsamples the feature maps, enhancing spatial invariance and reducing overfitting.
Conv2D	The third convolutional layer applies 120 filters, further refining feature extraction. It has 48,120 trainable parameters.
BatchNormalization	This layer normalizes the activations of the third convolutional layer.
Dropout	This layer randomly drops 20% of the activations during training, preventing overfitting and improving generalization.

Table 3.4 – continued from previous page

Layer	Description
Flatten	This layer reshapes the 3D output of the convolutional layers into a 1D vector for input to the dense layers.
Dense	This fully connected layer has 84 neurons, further processing the extracted features. It has 10,164 trainable parameters.
Dense	The final output layer has 2 neurons, corresponding to the number of classes in the problem. It has 170 trainable parameters.

In total, there are a total of 61,894 total parameters, with 61,610 of them being trainable parameters. The model was then compiled using the Adam optimizer, with a learning rate of 0.001, and a loss function of sparse categorical cross-entropy. The model was then trained using the dataset, with a batch size of 128, and a total of 20 epochs. The model was then saved, and the performance of the model was evaluated.

Besides that, the model was given two callbacks, which are the EarlyStopping callback, and the ReduceLROnPlateau callback. The EarlyStopping callback is used to stop the training of the model when a monitored metric has stopped improving, while the ReduceLROnPlateau callback is used to reduce the learning rate when a metric has stopped improving. This is to prevent overfitting, and to improve the performance of the model.

3.7 Model Testing and Evaluation

Model testing and evaluation is the process of assessing the performance and quality of the model by measuring its effectiveness, which was via evaluating the model performance with accuracy, precision, recall, and F1-score.

Figure 3.7 matplotlib logo

Obtaining the results of the model is made possible via TensorFlow, however it is further visualized using matplotlib, a plotting library for Python. matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. It is a multi-platform data visualization library built on NumPy arrays and designed to work with the broader SciPy stack.

With matplotlib, the results of the model can be visualized in the form of plot graphs, confusion matrix, and classification report. The plot graphs are used to visualize the accuracy and loss of the model, while the confusion matrix and classification report are used to visualize the performance of the model.

3.8 System Development

The system development phase spans from the data pre-processing phase until the end of this study, this phase covers the development of the system. A system was required to present the results of the findings the machine learning model. The system will be a culmination of all work that has been done from start to finish.

For the project, Streamlit was utilized to develop the system. Streamlit is an open-source Python library that makes it simple to create and share web apps for machine learning and data science. It is a Python library that allows for the creation of web applications, and is used to present the results of the model. With Streamlit, the requirement for a frontend and backend is eliminated, as it is a single-page application, and is easily integrated with the model, as it is written in Python.

Figure 3.8 Streamlit logo

With Streamlit, the requirement for a frontend and backend is eliminated, as it is a single-page application, and is easily integrated with the model, as it is written in Python. Streamlit uses Markdown, which is a lightweight markup language for creating formatted text using a plain-text editor. Markdown is often used to format readme files, for writing messages in online discussion forums, and to create rich text using a plain text editor. Its simplicity allows for the system to present itself gracefully, without the need for a complex frontend.

Figure 3.9 Markdown logo

The system features 4 pages, which includes:

- Home

  The landing page showcases a brief introduction of the system, and the purpose of the system.

- About Project

  This page details of the project, where it presents the overview of the project, its dataset, and the model.

- Findings

  This page presents the findings of the project, where it showcases the accuracy and loss of the model, as well as the confusion matrix of the project.

- Demo

  This page allows users to test the system out, by uploading an image of a handwriting sample, and the system will predict the class of the handwriting sample.

The system is presented in a custom dark-coloured theme, with lime green as the accent colour. The system is made responsive, where it can be viewed on mobile devices, as well as desktops. The system is hosted on Streamlit Community Cloud.

3.9 Documentation

The final phase of the study is the documentation phase, where all completed tasks were recorded. This involves the preparation of a comprehensive final report detailing each and every activities carried out during this research. The final report represents an culmination of all information and discoveries related to the project's implementation and development, serving as a complete record of the project's execution.

LaTeX

Figure 3.10 LaTeX logo

The documentation is made possible via LaTeX, a document preparation system for high-quality typesetting. It is most often used for medium-to-large technical or scientific documents but it can be used for almost any form of publishing. LaTeXis not a word processor, but rather a document markup language. For this project, a template made by Rizauddin (2023) was used as a base for the documentation, with a few changes made to adapt to the Final Year Project format.

3.10 Summary

In conclusion, this chapter goes through the research methodologies that has been applied throughout the research. This chapter details all the phases that will be involved from the start to finish. It also provides an mental image of how the flow of the study went.

CHAPTER FOUR
RESULTS AND FINDINGS

This chapter presents the results and findings of the study. Following the methodology outlined in Chapter 3, the data was collected, pre-processed, and utilized to train and evaluate the chosen machine learning model. This chapter will delve into the key findings obtained through this process, focusing on the model's performance and its implications for dyslexia diagnosis.

4.1 Training performance

Utilising the CUDA cores on the GPU, the model was trained for 20 epochs. However, due to the callbacks implemented into the training process, the model was able to stop training early. At epochs 10 and 12, the callback ReduceLROnPlateau was called, reducing the learning rate to 0.0005 and 0.00025 respectively, in order to reduce overfitting. At epoch 18, the callback EarlyStopping was called, stopping the training process early. The model was trained for 18 epochs, with a total training time of 1 minute and 46 seconds, making it 5.89 seconds per epoch.

4.2 Training results

Figure 4.1 shows a combined graph of both the accuracy and loss of the model during training and validation.

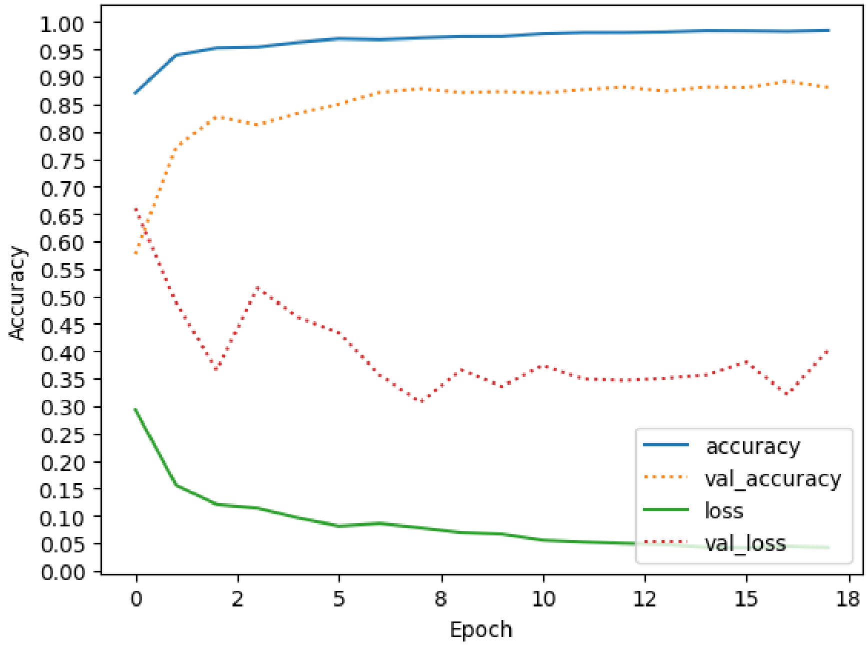

Figure 4.1 Graph accuracy and loss plot of the LeNet-5 model

4.2.1 Training accuracy

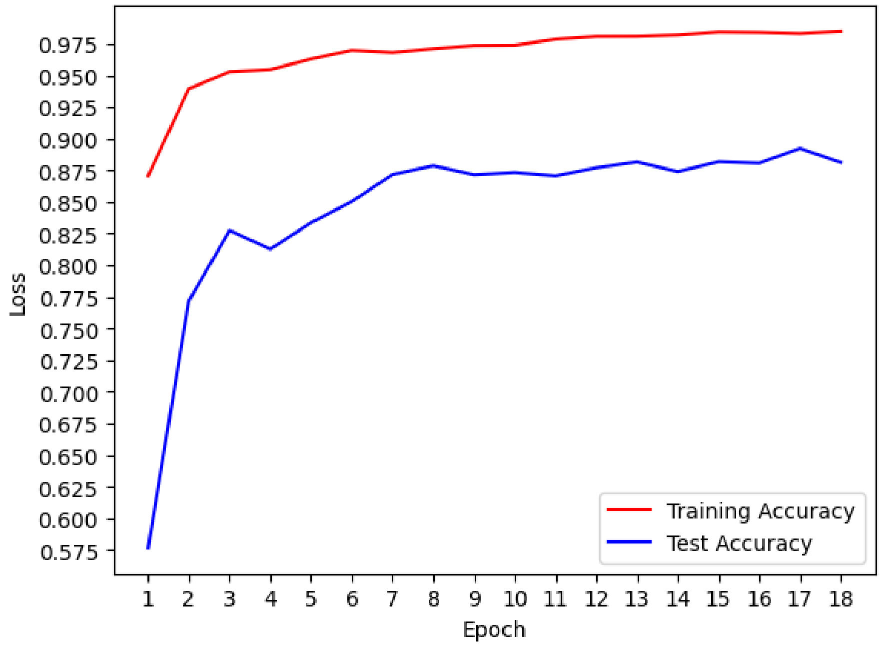

Figure 4.2 Graph accuracy plot of the LeNet-5 model

Figure 4.2 provides a graph representation of the accuracy of the model during training and validation. The model has reported an accuracy of 98.47%, with a validation accuracy of 88.12%. From this, we can infer that the model is slightly off in terms of accuracy, as the validation accuracy is 10.35% off the training accuracy.

This could mean that the model is overfitting, where the model is not able to generalize well. This is a problem because the model will not be able to perform well on the test data.

4.2.2 Training loss

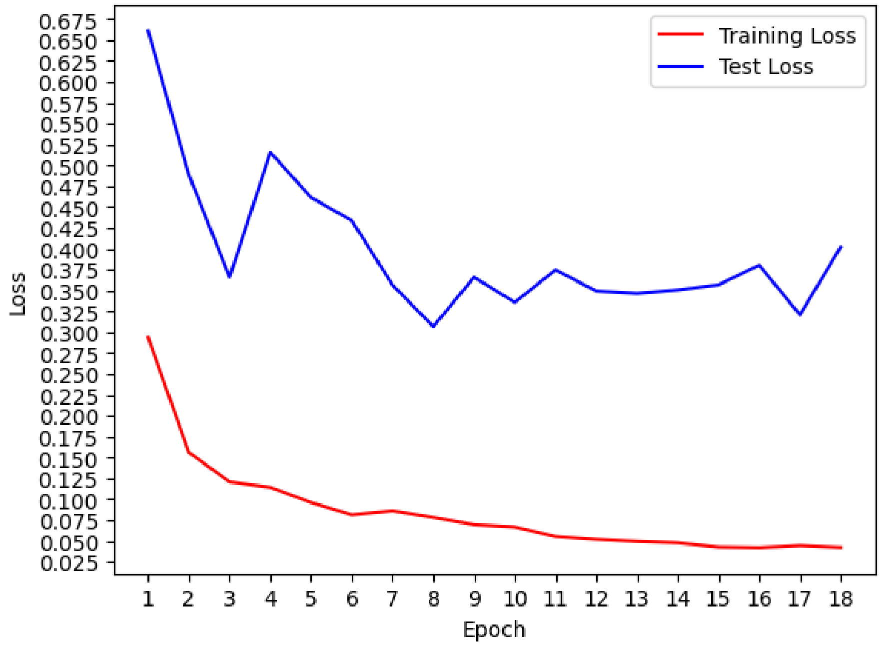

Figure 4.3 Graph loss plot of the LeNet-5 model

Figure 4.3 provides a graph representation of the loss of the model during training and validation. The model has reported loss values as low as 4.21%, with validation loss of 40.20%. From this, we can infer that the model is off in terms of loss, as the validation loss is 35.99% off the training loss. This carries over from the training accuracy results, where these values could mean that the model is overfitting.

4.3 Evaluation result

4.3.1 `model.evaluate`

Using the `model.evaluate` function, the model was evaluated on the test dataset. The `model.evaluate` function returns the loss value and metrics values for the model in test mode. The results of the evaluation process can be seen in Table 4.1.

Table 4.1
Evaluation results of the modified LeNet-5.

Test Accuracy (%)	88.23
Test Loss (%)	39.06

4.3.2 Confusion Matrix

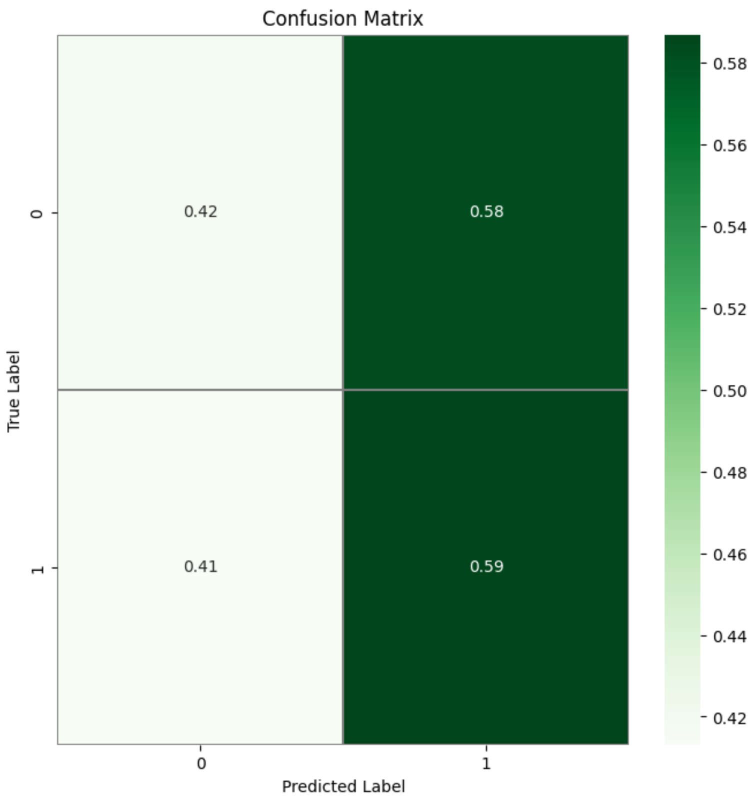

Figure 4.4 Confusion matrix of the LeNet-5 model

Figure 4.4 shows the confusion matrix for the evaluation process of the LeNet-5 model. The confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known. The confusion matrix itself is relatively simple to understand, but the related terminology can be confusing. The confusion matrix shows the ways in which the classification model is confused when it makes predictions. It gives us insight not only into the errors being made by a classifier but more importantly the types of errors that are being made.

Based off this confusion matrix, the model has predicted 42% as true negatives, while 59% as true positives. However, as the model has predicted 41% as false negatives, and 58% as false positives, this shows that the model is not able to classify the samples well.

4.3.3 Classification Report

The findings of the evaluation process can be seen in the classification report below, Table 4.2.

Table 4.2
Classification report of the modified LeNet-5 model.

Class Label	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
0	35	42	38
1	66	59	62
	Accuracy (%)		**53**
Macro average (%)	50	50	50
Weighted average (%)	55	53	54

Table 4.2 presents with the classification report for the evaluation process of the LeNet-5 model. From this, we can infer that the some highlights:

- The model has a higher precision for class 1 (0.66) than for class 0 (0.35). This means that the model is more accurate at identifying true positives for class 1 than for class 0.
- The model has a higher recall for class 0 (0.59) than for class 1 (0.42). This means that the model is more likely to correctly identify all true positives for class 0 than for class 1.

- The F1-score, a balanced measure of precision and recall, is higher for class 1 (0.62) than for class 0 (0.38). This suggests that the model performs better overall for class 1 than for class 0.

- The overall accuracy of the model is 53%. This is the average of the precision and recall for both classes. This is relatively low, and could mean that the model is not able to classify the samples well.

- The weighted average precision (0.55) is slightly higher than the macro average precision (0.50). This suggests that the model is biased towards class 1, as there are more instances in that class.

4.4 Summary

To summarise, the chapter goes through the findings of the model, spanning from its training and validation results to its evaluation results. As the findings show, it is rather concerning that the model is overfitting. This is a problem as the model is not able to generalize well, and will not be able to classify new samples that it has not seen before. This is a problem that is commonly found in machine learning models, and is a problem that is difficult to solve. The validation accuracy is reported to be 10.35% off the training accuracy, while the validation loss is reported to be off by a significant 35.99%. This could also mean that the altered dataset, may not suitable for the model.

CHAPTER FIVE
CONCLUSION AND RECOMMENDATIONS

The chapter presents the conclusion of the study and the recommendations for future works. This chapter goes through the objectives of the study, its strengths and limitations, and recommendations for future works.

5.1 Purpose of the Study

The purpose of the study is to develop a machine learning model that can be applied in a system, to detect potential dyslexia through handwriting images. The approach taken was using the LeNet-5 model.

5.1.1 First Objective

The first objective in this study is to identify the features of dyslexia handwriting in images. This is achieved by performing a thorough literature review on the features of dyslexia handwriting, which spans from various journals, research articles, and books published by academics and researchers all over the world. The information regarding the features of dyslexia handwriting is then used to develop a machine learning model that can be used to detect dyslexia handwriting in images, which was documented in the second chapter of this report. Then, the dataset is acquired through Kaggle and pre-processed to be applied in the machine learning model.

5.1.2 Second Objective

The second objective of the study is to develop a machine learning model that distinguish dyslexic or non-dyslexic samples. This is achieved by using a modified version of the LeNet-5 model, which is a convolutional neural network model that is used to classify handwritten digits. The model is modified to be able to classify dyslexic or non-dyslexic samples. The model is then trained using the dataset that was pre-processed in the previous objective.

5.1.3 Third Objective

The third objective of the study is to evaluate the performance of the machine learning model. This is achieved by evaluating the performance of the model using the accuracy, precision, recall, and F1-score metrics. The model is also evaluated using the confusion matrix. The model is also tested using a sample of the dataset that was not used in the training process. However, the model is found to be overfitting, which can be related to the publicly available dataset, which is unoptimized for the model.

5.2 Strengths and Limitations

In the study, there has been its strengths and limitations. Its discovery is detrimental to the study, as it indicates that the study is properly conducted, hence revealing areas of the study that could be further imrproved in the future.

5.2.1 Strengths

One of the strengths of this study is the development of the modified LeNet-5 model. Slight modifications made to the model has allowed it to be able to train the model to classify dyslexic or non-dyslexic samples at a high accuracy, reaching 98.11% accuracy. Furthermore, the loss of the model is impressively low, achieving a loss of 4.21%. This indicates that the model is able to classify the samples with a high accuracy, and the model is able to learn the features of dyslexia handwriting.

Besides that, the model is presented in an accessible manner. The model is utilized in a Streamlit application, which is a web application framework that is used to develop machine learning applications. Streamlit has allowed the project to be deployed anywhere, and it is compatible with most if not all devices. This allows the study to be accessible to everyone, and it can be used to detect dyslexia in children.

5.2.2 Limitations

There are various limitations found in this study. The first of many telling signs were the off-putting validation scores. The validation scores were significantly lower than the training scores, which indicates that the model is overfitting. This is a problem as the model is not able to generalize well, and will not be able to classify new samples that it has not seen before. This is a problem that is commonly found in machine learning models, and is a problem that is difficult to solve. The validation accuracy is reported to be 10.24% off the training accuracy, while the validation loss is reported to be off by a significant 35.9%.

Besides that, the evaluation scores for the model has reported rather worrying scores, where the classification report has reported only 53% accuracy. This all goes back to the issue of overfitting, where the model is not able to generalize well.

5.3 Recommendation and Future Works

In terms of recommendations, the first recommendation is to acquire a better dataset. The dataset used in this study is a publicly available dataset, which is not optimized for the model. This is evident in the overfitting of the model, where the model is not able to generalize well.

Besides that, the dataset used in this study is also not balanced, where the number of dyslexic samples is significantly lower than the number of non-dyslexic samples. This is a problem as the model will be biased towards the non-dyslexic samples, and will not be able to classify dyslexic samples well. Besides that, the model can be further optimized, where the model can be trained using different optimizers, different loss functions, and also different activation functions.

Once the model is optimized, the model can be deployed in a system, where the system can be used to detect dyslexia in children. The system can be used to detect dyslexia in children at a young age, where the system can be used to detect dyslexia in children before they enter primary school. This is important as dyslexia can be treated at a young age, where the children can be taught to read and write properly. This will allow the children to be able to read and write properly, and will allow them to be able to perform well in school.

REFERENCES

'Arif, M., Hassan, H., Nasien, D., & Haron, H. (2015). A review on feature extraction and feature selection for handwritten character recognition. *International Journal of Advanced Computer Science and Applications*, *6*(2). https://doi.org/10.14569/ijacsa.2015.060230

Aboudan, R., Eapen, V., Bayshak, M., Al-Mansouri, M., & Al-Shamsi, M. (2011). Dyslexia in the united arab emirates university – a study of prevalence in english and arabic. *International Journal of English Linguistics*, *1*, 64.

Alexander, A. W., & Slinger-Constant, A.-M. (2004). Current status of treatments for dyslexia: Critical review. *Journal of Child Neurology*, *19*(10), 744–758. https://doi.org/10.1177/08830738040190100401

Arora, N., & Gabrani, G. (2018). Significant machine learning and statistical concepts and their applications in social computing. *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*. https://doi.org/10.1109/icsccc.2018.8703329

Ashraf, M., & Majeed, S. (2011). Prevalence of dyslexia in secondary school students in lahore. *Pakistan Journal of Psychological Research*, *26*.

Cheng-Lai, A., Li-Tsang, C. W., Chan, A. H., & Lo, A. G. (2013). Writing to dictation and handwriting performance among chinese children with dyslexia: Relationships with orthographic knowledge and perceptual-motor skills. *Research in Developmental Disabilities*, *34*(10), 3372–3383. https://doi.org/10.1016/j.ridd.2013.06.039

Choy, G., Khalilzadeh, O., Michalski, M., Do, S., Samir, A. E., Pianykh, O. S., Geis, J. R., Pandharipande, P. V., Brink, J. A., & Dreyer, K. J. (2018). Current applications and future impact of machine learning in radiology. *Radiology*, *288*(2), 318–328. https://doi.org/10.1148/radiol.2018171820

Dietterich, T. (1996). Machine learning. *ACM Computing Surveys*, *28*(4es), 3. https://doi.org/10.1145/242224.242229

Dinov, I. D. (2018). Deep learning, neural networks. In *Data science and predictive analytics* (pp. 765–817). Springer International Publishing. https://doi.org/10.1007/978-3-319-72347-1_23

Drotár, P., & Dobeš, M. (2020). Dysgraphia detection through machine learning. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-78611-9

Dyslexia, n. (2023). In *Merriam-webster.com dictionary*. Merriam-Webster. Retrieved June 25, 2023, from https://www.merriam-webster.com/dictionary/dyslexia

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. *RadioGraphics*, *37*(2), 505–515. https://doi.org/10.1148/rg.2017160130

Francks, C., MacPhie, I. L., & Monaco, A. P. (2002). The genetic basis of dyslexia. *The Lancet Neurology*, *1*(8), 483–490. https://doi.org/10.1016/s1474-4422(02)00221-1

Fulkerson, B., Michie, D., Spiegelhalter, D. J., & Taylor, C. C. (1995). Machine learning, neural and statistical classification. *Technometrics*, *37*, 459. https://doi.org/10.2307/1269742

Garg, A., & Bansal, D. (2023). Application of machine learning in disease prediction. *International Journal for Research in Applied Science and Engineering Technology*, *11*(2), 47–51. https://doi.org/10.22214/ijraset.2023.48954

Gosse, C., & Reybroeck, M. V. (2020). Do children with dyslexia present a handwriting deficit? impact of word orthographic and graphic complexity on handwriting and spelling performance. *Research in Developmental Disabilities*, *97*, 103553. https://doi.org/10.1016/j.ridd.2019.103553

Guerin, D. W., Griffin, J. R., Gottfried, A. W., & Christenson, G. N. (1993). Dyslexic subtypes and severity levels: Are there gender differences? *Optometry and vision science : official publication of the American Academy of Optometry*, *70 5*, 348–51.

Isa, D. I. S. (2022, February). Dyslexia handwriting dataset. https://www.kaggle.com/datasets/drizasazanitaisa/dyslexia-handwriting-dataset

Isa, I. S., Rahimi, W. N. S., Ramlan, S. A., & Sulaiman, S. N. (2019a). Automated detection of dyslexia symptom based on handwriting image for primary school children. *Procedia Computer Science*.

Isa, I. S., Rahimi, W. N. S., Ramlan, S. A., & Sulaiman, S. N. (2019b). Automated detection of dyslexia symptom based on handwriting image for primary school children. *Procedia Computer Science*, *163*, 440–449.

Isa, I. S., Zahir, M. A., Ramlan, S. A., Li-Chih, W., & Sulaiman, S. N. (2021). Cnn comparisons models on dyslexia handwriting classification.

Jankovic, M. M. (2022). Biomarker-based approaches for dyslexia screening: A review. *2022 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC)*. https://doi.org/10.1109/zinc55034.2022.9840532

Jhaveri, R. H., Revathi, A., Ramana, K., Raut, R., & Dhanaraj, R. K. (2022). A review on machine learning strategies for real-world engineering applications (S. Hakak, Ed.). *Mobile Information Systems*, *2022*, 1–26. https://doi.org/10.1155/2022/1833507

Kearns, D. M., Hancock, R., Hoeft, F., Pugh, K. R., & Frost, S. J. (2019). The neurobiology of dyslexia. *TEACHING Exceptional Children*, *51*(3), 175–188. https://doi.org/10.1177/0040059918820051

Lam, S., Au, R. K. C., Leung, H., & Li-Tsang, C. W. P. (2011). Chinese handwriting performance of primary school children with dyslexia. *Research in developmental disabilities*, *32 5*, 1745–56.

Langley, P., & Simon, H. A. (1995). Applications of machine learning and rule induction. *Communications of the ACM*, *38*(11), 54–64. https://doi.org/10.1145/219717.219768

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Lum, K. (2017). Limitations of mitigating judicial bias with machine learning. *Nature Human Behaviour*, *1*(7). https://doi.org/10.1038/s41562-017-0141

Machine learning, n. (2023). In *Merriam-webster.com dictionary*. Merriam-Webster. Retrieved June 25, 2023, from https://www.merriam-webster.com/dictionary/machine%5C%20learning

Mana, S. C., Kalaiarasi, G., R, Y., Helen, L. S., & Selvi, R. S. (2022). Application of machine learning in healthcare: An analysis. *2022 3rd International Confer-*

ence on Electronics and Sustainable Communication Systems (ICESC). https://doi.org/10.1109/icesc54411.2022.9885296

Naqa, I. E., & Murphy, M. J. (2015). What is machine learning? In *Machine learning in radiation oncology* (pp. 3–11). Springer International Publishing. https://doi.org/10.1007/978-3-319-18305-3_1

Norton, E. S., Beach, S. D., & Gabrieli, J. D. (2015). Neurobiology of dyslexia. *Current Opinion in Neurobiology*, *30*, 73–78. https://doi.org/10.1016/j.conb.2014.09.007

Oladipupo, T. (2010, February). Types of machine learning algorithms. In *New advances in machine learning*. InTech. https://doi.org/10.5772/9385

Ortiz, A., Martinez-Murcia, F. J., Luque, J. L., Giménez, A., Morales-Ortega, R., & Ortega, J. (2020). Dyslexia diagnosis by EEG temporal and spectral descriptors: An anomaly detection approach. *International Journal of Neural Systems*, *30*(07), 2050029. https://doi.org/10.1142/s012906572050029x

Pamina, J., & Raja, B. (2019). Survey on deep learning algorithms. *CompSciRN: Other Machine Learning (Topic)*.

Paracchini, S., Scerri, T., & Monaco, A. P. (2007). The genetic lexicon of dyslexia. *Annual Review of Genomics and Human Genetics*, *8*(1), 57–79. https://doi.org/10.1146/annurev.genom.8.080706.092312

Patricia A. Flanagan. (2016). Nist handprinted forms and characters - nist special database 19. https://doi.org/10.18434/T4H01C

Pavlidis, G. T. (1981). Do eye movements hold the key to dyslexia? *Neuropsychologia*, *19*(1), 57–64. https://doi.org/10.1016/0028-3932(81)90044-0

Ramli, S., Idris, I. B., Omar, K., Harun, D., & Surat, S. (2020). Development of dyslexia health education module (dhem) for preschool teachers. *Malaysian Journal of Medicine and Health Sciences*, *16*, 2636–9346.

Ray, S. (2019). A quick review of machine learning algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. https://doi.org/10.1109/comitcon.2019.8862451

Richard, G., & Serrurier, M. (2020). Dyslexia and dysgraphia prediction: A new machine learning approach. *ArXiv*, *abs/2005.06401*.

Rizauddin, S. (2023). Uitmthesis.

Roitsch, J., & Watson, S. (2019). An overview of dyslexia: Definition, characteristics, assessment, identification, and intervention. *Science Journal of Education*, *7*(4), 81. https://doi.org/10.11648/j.sjedu.20190704.11

Sachin Patel. (2017). A-z handwritten alphabets in .csv format. https://www.kaggle.com/datasets/sachinpatel21/az-handwritten-alphabets-in-csv-format

Samuelsson, S., & Lundberg, I. (2003). The impact of environmental factors on components of reading and dyslexia. *Annals of Dyslexia*, *53*(1), 201–217. https://doi.org/10.1007/s11881-003-0010-8

Schulte-Körne, G. (2010). The prevention, diagnosis, and treatment of dyslexia. *Deutsches Ärzteblatt international*. https://doi.org/10.3238/arztebl.2010.0718

Schumacher, J., Hoffmann, P., Schmal, C., Schulte-Korne, G., & Nothen, M. M. (2007). Genetics of dyslexia: The evolving landscape. *Journal of Medical Genetics*, *44*(5), 289–297. https://doi.org/10.1136/jmg.2006.046516

Shi, Z., Zhong, C., Yokota, Y., Xia, W., & Sun, J. (2019). Robustness evaluation of deep learning models based on local prediction consistency. *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. https://doi.org/10.1109/icmla.2019.00224

Signor, R., Claessen, M., & Leitão, S. (2020). Interventions for dyslexia in brazil: A scoping review discussed within the perspective of international best practice. *Australian Journal of Learning Difficulties*, *25*(1), 1–30. https://doi.org/10.1080/19404158.2019.1709216

Snowling, M. J. (2012). Early identification and interventions for dyslexia: A contemporary view. *Journal of Research in Special Educational Needs*, *13*(1), 7–14. https://doi.org/10.1111/j.1471-3802.2012.01262.x

Soriano-Ferrer, M., & Martínez, E. P. (2017). Una revisión de las bases neurobiológicas de la dislexia en población adulta. *Neurología*, *32*(1), 50–57. https://doi.org/10.1016/j.nrl.2014.08.003

Spoon, K., Crandall, D. J., & Siek, K. (2019). Towards detecting dyslexia in children's handwriting using neural networks.

Strobel, M., & Shokri, R. (2022). Data privacy and trustworthy machine learning. *IEEE Security &amp Privacy*, *20*(5), 44–49. https://doi.org/10.1109/msec.2022.3178187

Suárez-Coalla, P., Afonso, O., Martínez-García, C., & Cuetos, F. (2020). Dynamics of sentence handwriting in dyslexia: The impact of frequency and consistency. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00319

Subbaswamy, A., Adams, R. J., & Saria, S. (2020). Evaluating model robustness to dataset shift. *ArXiv*, *abs/2010.15100*.

Sumner, E., Connelly, V., & Barnett, A. L. (2014). The influence of spelling ability on handwriting production: Children with and without dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(5), 1441–1447. https://doi.org/10.1037/a0035785

Tønnessen, F. E. (1999). Options and limitations of the cognitive psychological approach to the treatment of dyslexia. *Journal of Learning Disabilities*, *32*(5), 386–393. https://doi.org/10.1177/002221949903200504

Udomhunsakul, S., & Wongsita, P. (2004). Feature extraction in medical MRI images. *IEEE Conference on Cybernetics and Intelligent Systems, 2004*. https://doi.org/10.1109/iccis.2004.1460437

Usman, O. L., Muniyandi, R. C., Omar, K. B., & Mohamad, M. (2021). Advance machine learning methods for dyslexia biomarker detection: A review of implementation details and challenges. *IEEE Access*, *9*, 36879–36897.

Wajuihian, S. O. (2011). Neurobiology of developmental dyslexia: Part 1: A review of evidence from autopsy and structural neuro-imaging studies. *African Vision and Eye Health*, *70*(4). https://doi.org/10.4102/aveh.v70i4.117

Wajuihian, S. O., & Naidoo, K. S. (2011). Dyslexia: An overview. *African Vision and Eye Health*, *70*, 89–98.

Ying, X. (2019). An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, *1168*, 022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Yuzaidey, N. A. M., Normah Che Din, Mahadir Ahmad, Norhayati Ibrahim, Rogayah A. Razak, & Dzalani Harun. (2017). Interventions for children with dyslexia: A review on current intervention methods. https://doi.org/10.13140/RG.2.2.30068.04483

Zhang, Q.-s., & Zhu, S.-c. (2018). Visual interpretability for deep learning: A survey. *Frontiers of Information Technology &amp Electronic Engineering*, *19*(1), 27–39. https://doi.org/10.1631/fitee.1700808

APPENDICES

APPENDIX A
GANTT CHART

A Gantt chart visualizing the flow of the entire study was made, based on the phases of the research, as shown in the figure below:

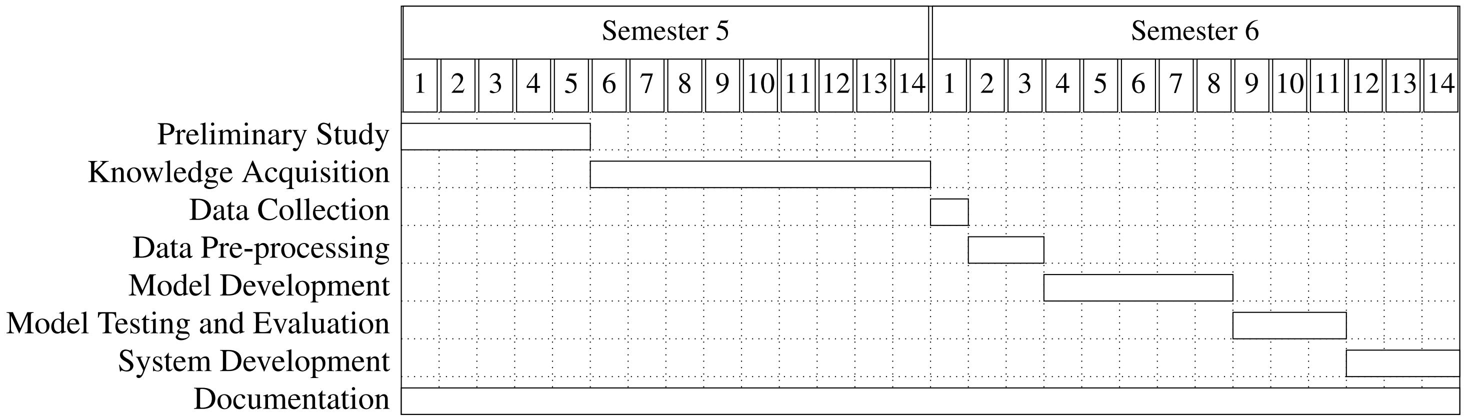

Figure A.1 Gantt Chart

APPENDIX B
THE DATA

This is some of the data that has been used in training and evaluating the data.

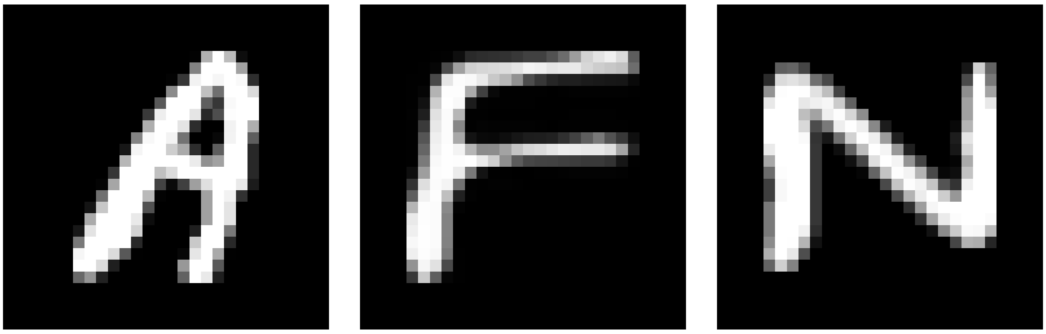

Table B.1
Non-dyslexic samples.

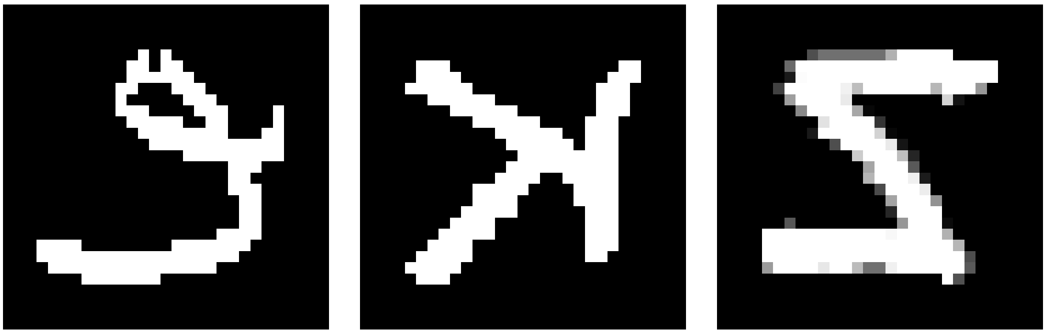

Table B.2
Potentially dyslexic samples.

APPENDIX C
THE SYSTEM

The following is the Streamlit web application interface. The interface contains 4 pages, which includes the homepage, the about project page, the findings page, and finally the demonstration page. The image is taken on a PC, on a window resolution of 1663 by 1301 pixels. The browser used was Brave Browser in guest mode, and the zoom level was set to 100%.

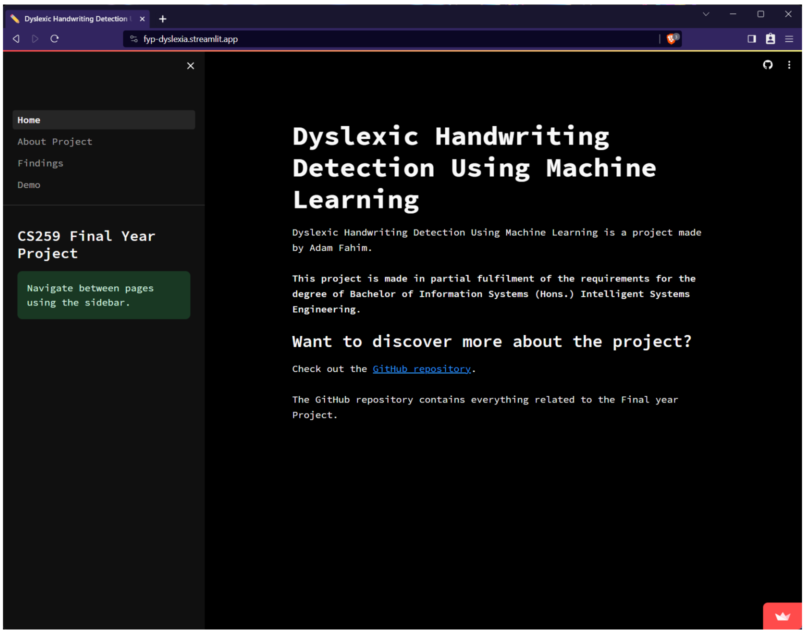

Figure C.1 The homepage of the Streamlit web application.

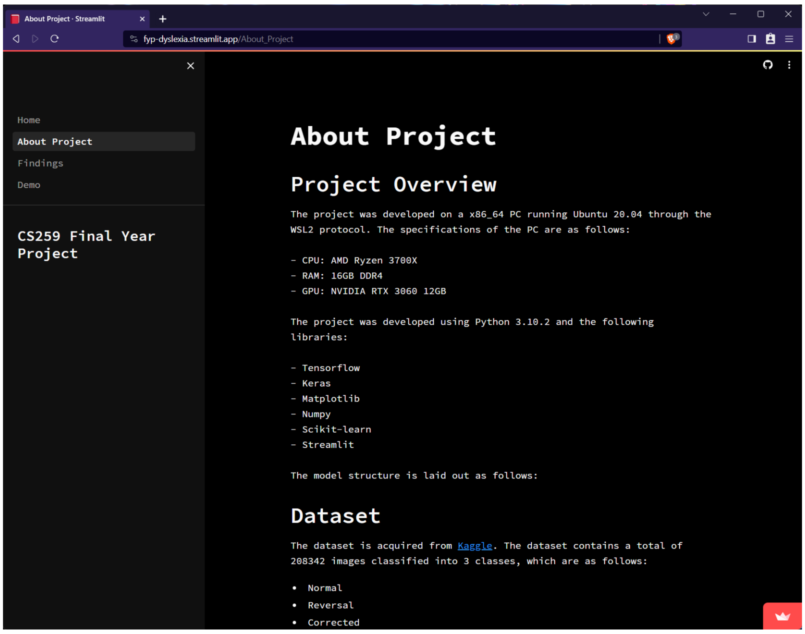

Figure C.2 The about project page of the Streamlit web application.

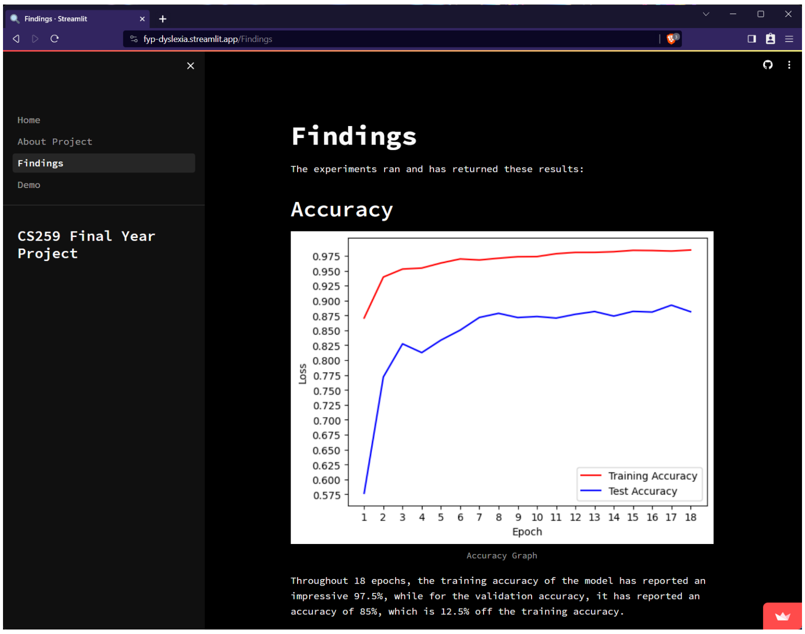

Figure C.3 The findings page of the Streamlit web application.

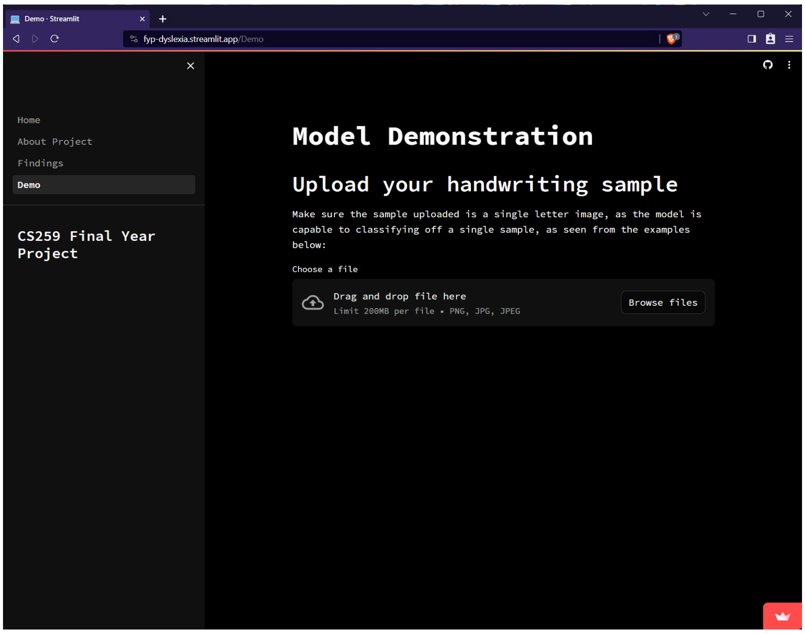

Figure C.4 The demonstration page of the Streamlit web application before running the model.

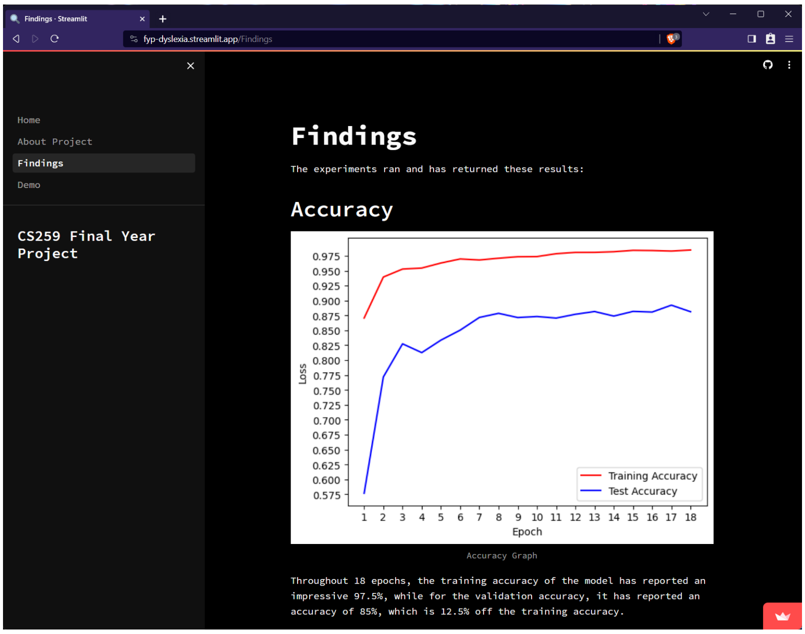

Figure C.5 The demonstration page of the Streamlit web application after running the model.