Faculty of Computer Science and Information Technology

A DEEP LEARNING APPROACH FOR MULTICLASS CLASSIFICATION OF RETINAL IMAGES

OON WENG WAI

Bachelor of Computer Science With Honours

(Multimedia Computing)

2023

A DEEP LEARNING APPROACH FOR MULTICLASS CLASSIFICATION OF RETINAL IMAGES

OON WENG WAI

This project is submitted in partial fulfillment of the
requirements for the degree of
Bachelor of Computer Science and Information Technology

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK
2023

PENDEKATAN PEMBELAJARAN MENDALAM UNTUK KLASIFIKASI BERBILANG KELAS IMEJ RETINAL

OON WENG WAI

Project ini merupakan salah satu keperluan untuk Ijazah
Sarjana Muda Sains Komputer dan Teknologi Maklumat

Fakulti Sains Komputer dan Teknologi Maklumat
UNIVERSITI MALAYSIA SARAWAK
2023

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE **A Deep Learning Approach for Multiclass Classification of Retinal Images**

ACADEMIC SESSION: 2022 / 2023

OON WENG WAI
(CAPITAL LETTERS)

hereby agree that this Thesis* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [or for the purpose of interlibrary loan between HLI]
5. ** Please tick (√)

	CONFIDENTIAL	(Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
	RESTRICTED	(Contains restricted information as dictated by the body or organization where the research was conducted)
√	UNRESTRICTED	

Validated by

(AUTHOR'S SIGNATURE) (SUPERVISOR'S SIGNATURE)

Permanent Address

15, LINTASAN BUNGA RAYA,
TAMAN ON,
31650, IPOH, PERAK.

Date: 30/6/2023 Date: 25/7/2023

Note * Thesis refers to PhD, Master, and Bachelor Degree
** For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to my final-year supervisor, Dr. Stephanie Chua Hui Li for constantly guiding me with the knowledge of data, text, and image mining. She constantly provides me with detailed instructions. Without the guidance of my supervisor, I truly believe the completion of my final-year project will become a chore. Next, I want to give a million thanks to Dr. Wang Yin Chai for giving me a plethora of useful information regarding the final-year project where the given information has helped me to complete my project in a timely and professional manner. I would like to give my thanks to my data mining lecturer, Dr. Bong Chih How who taught me the fundamentals of machine learning so that I could better grasp the entire data mining process. Besides, I did not forget to thank Guna Venkat Doddi, the author who provided ample eye diseases dataset on the Kaggle Website. Additionally, I would like to show my appreciation to my friends for their help. I am thankful to them as they are willing to sacrifice their time to have a discussion and some opinions on my final-year project. Last but not least, I would like to thank my beloved family for their encouragement and endless support, be it financial or emotional for me to complete this project.

ABSTRACT

The eyes fulfill a vital function in daily life as the primary sensory organs for perceiving the world. Neglecting proper attention and care for eye health can lead to vision loss or impairment, with severe eye diseases having the potential to cause complete blindness. The main problem addressed in this study is the global burden of eye diseases, worsened by a shortage of ophthalmologists and insufficient training in using essential instruments. These factors contribute to delays in diagnosis and treatment, underscoring the necessity for an automated approach to classify eye diseases from retinal images. Therefore, this study aims to compare various types of deep learning models, such as Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) to classify specific eye diseases, including cataracts, diabetic retinopathy, glaucoma, and normal classes. Based on a review of related research, it is evident that CNN is well-suited for image classification. Therefore, this study will train models using a proposed CNN, VGG19, and InceptionV3 architectures. In this study, the dataset collected from Kaggle underwent preprocessing involving histogram equalization and image segmentation techniques. It was then divided into training (80%), validation (10%), and testing (10%) sets. Performance evaluation of the models will be based on accuracy, precision, and recall. Techniques such as early stopping, model checkpointing, hyperparameter tuning, and fine-tuning were employed to optimize the models and improve their accuracy. VGG19 Model 3 with 16 convolutional layers, 3 dense layers (each with 300 neurons), a dropout rate of 0.5, and Adam optimizer with a learning rate of 0.0001, achieved impressive training accuracy of 90% and testing accuracy of 87%. The utilization of deep learning models in clinical settings could help improve the accuracy and speed of eye disease classification, leading to timely interventions and improved patient outcomes.

ABSTRAK

Mata memenuhi fungsi penting dalam kehidupan seharian sebagai organ deria utama untuk melihat dunia. Mengabaikan perhatian dan penjagaan yang sewajarnya terhadap kesihatan mata boleh memyebabkan kehilangan penglihatan atau kecacatan, dengan penyakit mata yang teruk berpotensi memyebabkan kebutaan sepenuhnya. Masalah utama yang ditangani dalam kajian ini ialah beban global penyakit mata, diburukkan lagi oleh kekurangan pakar oftalmologi dan latihan yang tidak mencukupi dalam menggunakan instrument penting. Faktor-faktor ini menyumbang kepada kelewatan dalam diagnosis dan rawatan, menekankan keperluan unutk pendekatan automatic untuk mengklasifikasikan penyakit mata daripada imej retina. Oleh itu, kajian ini bertujuan untuk membandingkan pelbagai jenis model pembelajaran mendalam, seperti Rangkaian Neural Buatan (ANN), Rangkain Neural Konvolusi (CNN), dan Rangkaian Neural Berulang (RNN) untuk mengklasifikasikan penyakit mata tertentu, termasuk katarak, retinopati diabetic, glaukoma, dan kelas biasa. Berdasarkan semakan penyelidikan berkaitan, terbukti bahawa CNN sangat sesuai untuk klasifikasi imej. Oleh itu, kajian ini akan melatih model menggunakan seni bina CNN, VGG19, dan InceptionV3 yang dicadangkan. Dalam kajian ini, set data yang dikumpul daripada Kaggle menjalani prapemprosesan yang melibatkan penyamaan histogram dan teknik pembahagian imej. Ia kemudiannya dibahagikan kepada set latihan (80%), pengesahan (10%), dan set ujian (10%). Penilaian prestasi model akan berdasarkan "Accuracy", "Precision", dan "Recall". Teknik seperti berhenti awal, titik Semak model, penalaan hyperparameter dan penalaan halus digunakan untuk mengoptimumkan model dan meningkatkan ketepatannya. VGG19 Model 3 dengan 16 lapisan konvolusi, 3 lapisan padat (masing-masing mempunyai 300 neuron), kadar keciciran 0.5m dan pengoptimum Adam dengan kadar pembelajaran 0.0001, mencapai ketepatan latihan yang mengagumkan sebanyak 90% dan ketepatan ujian sebanyak 87%. Penggunaan model pembelajaran mendalam dalam tetapan klinikal boleh membantu meningkatkan ketepatan dan kelajuan klasifikasi penyakit mata, yang membawa kepada campur tangan yang tepat pada masanya dan hasil pesakit yang lebih baik.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

Chapter 1
Introduction

1.1 Introduction

Our eyesight is one of the most important senses because it accounts for 80% of the information we took in (*Importance of Eye Care*, 2016). Having good eyesight will keep us safe and help maintain the sharpness of our minds (Victoria, 2019). Furthermore, good vision plays a vital role in compensating for impairments in taste, smell, hearing, and touch. The sense of taste is not solely responsible for deciding whether something tastes good or not as the eyes also play a major role in this process, with the ability to distinguish only five flavors (*Why Good Vision Is So Important*, 2017). For instance, yellow, orange, and red food are often thought to be sweeter. Even professional wine experts can be misled, as nine out of ten of them were unable to distinguish white wine that had been colored with red food coloring from regular red wine. However, eye diseases can significantly impact an individual's quality of life. This is because vision loss can affect our physical health and more seriously can lead to mental health such as loneliness, social isolation, feelings of worry, anxiety, and fear (*Vision Loss and Mental Health*, 2021). In the field of eye disease classification, multiclass classification is crucial for accurately identifying and distinguishing between various eye conditions.

A projected 596 million individuals globally, including 43 million blind persons, have distant vision impairment in 2020 (Burton, 2021). Due to the lack of reading glasses, another 510 million individuals have an uncorrected near-vision impairment. 90% of the afflicted individuals reside in low and middle-income nations (LMICs). According to the World Health Organization's 2022 (WHO) reports that out of the 1 billion people affected by vision impairment, 94 million have cataracts, 7.7 million have glaucoma, and 3.9 million have diabetic retinopathy.

A retinal photograph is another term for digital retinal imaging (Lazarus, 2020). According to Perez et al. (2012), digital photography has revolutionized retinal photography by enabling instantaneous image presentation without requiring a long development process. Explanation gives that it further allows real-time adjustments in composition, focus, and flash intensity to maximize the quality of the images, obtaining additional images without film has almost no incremental cost and quick and exact duplicates can be easily made and archived, which can be transmitted to remote locations. Retinal images are often utilized for eye disease screening and diagnoses, such as glaucoma, diabetic retinopathy, and age-related macular degeneration (Wan et al., 2022). This is because the optic cup, optic nerve, and blood vessels are frequently affected by these eye diseases' anomalies. Having high-quality retinal images is an important basis for ophthalmology precision medicine to detect eye diseases. In general, deep learning is a subset of machine learning and is sometimes the term used to describe algorithms that logically examine data, much like a person would (Wolfewicz, 2022). The difference between machine learning and deep learning is using machine learning, activities may be completed by computers without explicit programming, but the ability of machine learning to deal with images or videos still has many difficulties (Middleton, 2021). Due to its precise modeling of the human brain, deep learning models bring a very advanced approach to machine learning and are prepared to take on these difficulties. In short, it is always referred to as deep neural networks since the majority of deep learning techniques employ neural network concepts (*What Is Deep Learning? 3 Things You Need To Know, n.d.*). This study is focused on the development and comparison of deep learning models, including artificial neural network (ANN), convolutional neural network (CNN), and recurrent neural network (RNN). The objective is to effectively classify retinal images into multiple eye disease classes by harnessing the capabilities of these deep learning techniques. The primary aim is to improve the accuracy and efficiency of diagnosing eye diseases, leading to enhanced patient care and treatment outcomes. The project involves evaluating and comparing the performance of ANN, CNN, and RNN models to identify the best-performing model for multiclass eye disease classification. Furthermore, the selected model will be implemented into a web application for demonstration.

1.2 Problem Statement

Eye disease is a common problem. Millions around the world suffer from eye diseases at certain points in their lives. Except for those who were classified as being blind, more than 4.2

million Americans aged 40 and older are either legally blind or have low vision (*Common Eye Disorders and Diseases*, n.d.). Ophthalmology is the specialty in medicine that handles the diagnosis and treatment of eye diseases and disorders. It is regarded as a difficult specialty due to reasons such as lack of training in using the ophthalmoscope and other instruments, as well as difficulty in interpreting eye disease symptoms. Patients can also be difficult to examine as they may be overanxious or in pain. Digital retinal imaging (DRI) is non-invasive and uses a high-resolution imaging system to take colored images of the retina, optic nerve, and blood vessels in the back of our eyes (Lazarus, 2020). With retinal imaging, a digital image of the back of the eye can be viewed and this image can help doctors diagnose eye diseases. Retinal images can be taken using a fundus camera to photograph the fundus, which is the rear of the eye. This process only takes a few minutes and is painless. The retinal images can then be used by doctors to see signs of eye diseases that they could not see before. With the availability of these retinal images, many eye patients can be diagnosed quickly and given the appropriate, treatment. However, too many eye patients and too few doctors, in particular ophthalmologists, can cause a delay in the diagnosis and treatment of eye diseases. Besides, many students do not have much time to contact ophthalmology during training which also can result in students being unable to lay a solid foundation in ophthalmology (J. Khan et al., 2006). Therefore, an automated approach for eye disease classification proposed in this project can be used to aid classify different eye diseases from retinal images.

1.3 Scope

The scope of this project is to develop a deep learning approach for multiclass eye disease classification using retinal images. Specifically, the project focuses on differentiating between cataracts, diabetic retinopathy, glaucoma eye disease, and normal. Multiclass classification poses unique challenges that must be addressed to accurately classify retinal images into multiple disease classes. These challenges include handling variations in image features, distinguishing between similar disease classes, and optimizing the performance of the deep learning models for accurate multiclass predictions.

To ensure the dataset is appropriate for multiclass classification, the project utilized a dataset where each image is limited to a single eye disease. The dataset has been carefully organized by the dataset provider, placing retinal images with different diseases into separate folders.

The most frequent cause of blindness in the world is cataracts. According to Hildreth et al. (2009), cataracts also be called age-related eye disease because cataracts can develop at any age, but most commonly occur in those people over 40 years and above. A more scientific explanation given is it often results from a variety of interrelated environmental factors operating in combination with a genetic predisposition expressed in the lens protein genes (Shiels & Hejtmancik, 2019). The consequences can cause blurry vision or loss of vision.

The most common form of glaucoma is called Chronic Open Angle Glaucoma (COAG) or Primary Open Angle Glaucoma (POAG). Chronic open-angle glaucoma is defined by a loss of peripheral visual function and an excavated appearance of the optic disc on ophthalmoscopy. It is also progressing with slow atrophy of the optic nerve (Quigley, 1996). Typically, there is fluid often accumulates in the front of the eye. By increasing pressure on the eye, this additional fluid eventually harms the optic nerve and caused the person's vision loss.

With the global prevalence of diabetes increasing in many industrialized countries, diabetic retinopathy remains the leading cause of vision loss in the elderly (Lin et al., 2021). Usually, symptoms of diabetic retinopathy often occur in people with high blood sugar levels or people with Diabetic Mellitus (DM). As blood sugar levels rise, the sugar can clog and damage tiny blood vessels throughout the body, causing them to bleed or to occlude capillaries.

1.4 Aims and Objectives

The first aim and objective of this project are to study the background of eye disease, specifically Cataracts, Diabetic Retinopathy, and Glaucoma.

The second objective of this project is to evaluate and compare deep-learning models for multiclass eye disease classification. Various performance metrics, including accuracy, precision, and recall will be utilized as benchmarks to assess the accuracy and effectiveness of the deep learning models in classifying retinal images into multiple eye disease classes.

The last objective of this project is to implement the best eye disease classification model in a web application for demonstration, allowing users to conveniently and accurately diagnose and identify eye diseases.

1.5 Brief Methodology

Figure 1.1 below shows the full structure of the methodology where combined the KDD.

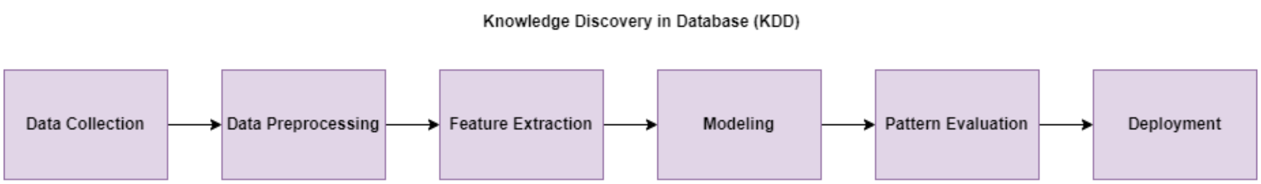


Figure 1.1 KDD (Mayo, 2016)

The data mining methodology employed in this project is Knowledge Discovery in Database (KDD), which is widely recognized and commonly used in data mining. The project is divided into distinct stages: data collection, data preprocessing, feature extraction, modeling, pattern evaluation, and deployment. In this project, the process will be repeated three times to compare three different types of deep-learning techniques such as artificial neural networks (ANN), convolutional neural networks (CNN), and recurrent neural networks (RNN). The aim is to identify the best model for multiclass eye disease classification. Once the best model is selected, a web application will be developed, integrating the model for demonstration purposes. While the project primarily focuses on data mining, the web application will not be the primary emphasis.

1.5.1 Data Collection

Firstly, the dataset will be collected and downloaded from the website Kaggle.com. The dataset includes example images of cataracts, diabetic retinopathy, glaucoma, and normal classes which have been divided into 4 different folders.

1.5.2 Data Preprocessing

According to the specification of this project, it is more appropriate to refer to the process as image preprocessing rather than data preprocessing, given that the collected data consists of images only. Typically, image preprocessing involves several steps, including resizing, orientation correction, and color adjustments, before feature selection and model training can be performed. As the dataset's image resolution has already been resized, noise is not a significant concern, although adjustments can still be made if necessary.

1.5.3 Feature Extraction

At this stage, the image datasets undergo a process to reduce redundant data, aiming to optimize the efficiency of the underlying machine operations. In deep learning, feature extraction primarily involves convolutional layers, which are also known as lower-level features in CNN models. These convolutional layers play a crucial role in extracting relevant features from the images for subsequent analysis and classification tasks. However, for models like ANN, manual feature engineering or traditional machine learning methods may still be employed to extract relevant features.

1.5.4 Modeling

The modeling process in deep learning constructs and trains neural network models for accurate predictions or classifications based on input data. These models automatically learn hierarchical representations, extracting relevant features at multiple levels. It involves defining the network architecture, including layer arrangement, neuron count, and activation functions. The model is trained using labeled data, adjusting weights and biases iteratively through optimizer algorithms like SGD, Adam, and so on. In this project, 80% portion of the dataset will be used for training the model, while 10% will be allocated for validation purposes.

1.5.5 Pattern Evaluation

Pattern evaluation in deep learning involves assessing the performance and effectiveness of trained models in making predictions or classifications. It focuses on measuring the model's accuracy, precision, recall, and other relevant metrics to evaluate its ability to correctly identify patterns in the data. Through this process, the model's performance is analyzed, and any areas of improvement (hyperparameter tuning) or fine-tuning are identified. In this project, an 10% portion of the dataset will be reserved for testing purposes.

1.5.6 Deployment

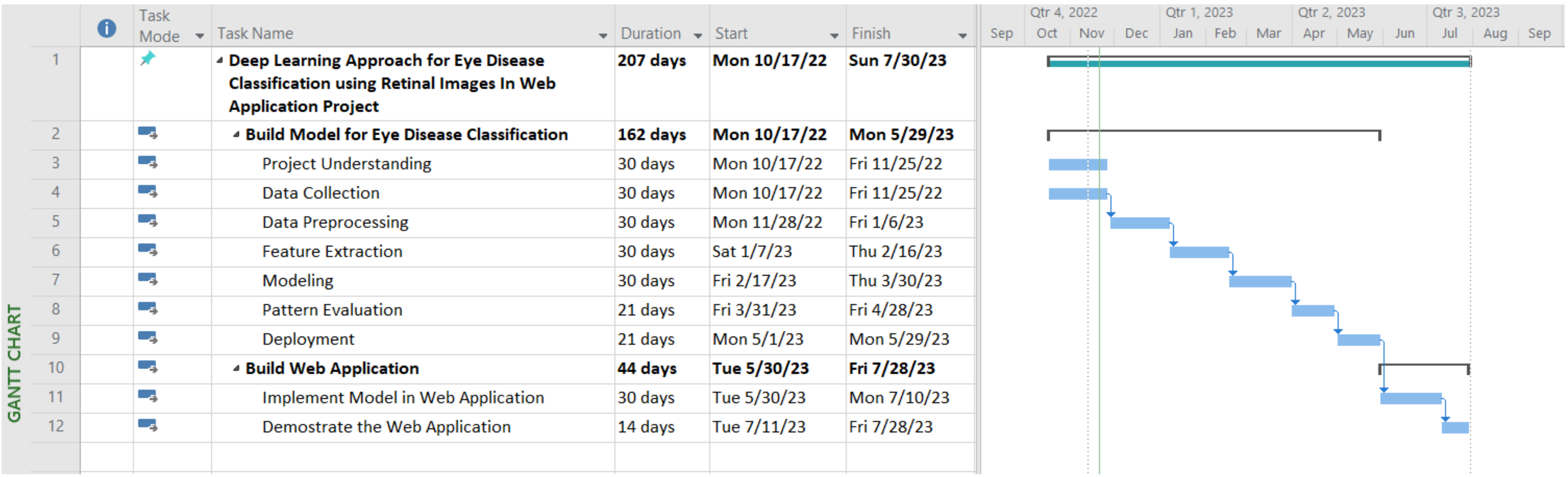
The deployment process in this project involves implementing the selected best-performing model into a web application for demonstration purposes. Once the model for multiclass eye disease classification is chosen based on its performance and accuracy, it is integrated into the web application to make it accessible to users.

1.6 Significance of Project

The significance of this project is to contribute to the field of healthcare more specifically to ophthalmology. It is expected that the development of this model will swiftly provide ease for ophthalmologists in diagnosing patients' symptoms of eye diseases. Ophthalmologists can better treat patients and offer more effective care if they can rapidly and properly diagnose symptoms. Also, given the lack of training in the use of ophthalmoscopes, and the difficulty for ophthalmologists to predict eye disease just by looking at the retinal images, this strategy can eventually offer simplicity and convenience for ophthalmologists while enhancing the standard of care for their patients by expediting the diagnosis process and reducing the possibility for the human mistake or error.

On the other hand, it is truly believed that once the model have being created, it will benefit all communities and government agencies, not only for ophthalmology or healthcare. This is because, as mentioned in the above data reports, people in today's society have a high probability of suffering from eye disease and the lack of doctors will prevent many patients from getting accurate treatment within a relatively short period. Thus, communities must have an awareness of eye diseases or acknowledgments. The project will help the public in general when it comes to eye health. It will raise awareness of eye disease among the public which help the communities to identify the possible symptoms of eye diseases early on. Furthermore, it can also help governments allocate resources more effectively to address the growing need for eye care services, as well as provide valuable data to researchers for developing better screening methods and treatments.

1.7 Project Schedule



	Task Mode	Task Name	Duration	Start	Finish
1		**Deep Learning Approach for Eye Disease Classification using Retinal Images In Web Application Project**	**207 days**	**Mon 10/17/22**	**Sun 7/30/23**
2		**Build Model for Eye Disease Classification**	**162 days**	**Mon 10/17/22**	**Mon 5/29/23**
3		Project Understanding	30 days	Mon 10/17/22	Fri 11/25/22
4		Data Collection	30 days	Mon 10/17/22	Fri 11/25/22
5		Data Preprocessing	30 days	Mon 11/28/22	Fri 1/6/23
6		Feature Extraction	30 days	Sat 1/7/23	Thu 2/16/23
7		Modeling	30 days	Fri 2/17/23	Thu 3/30/23
8		Pattern Evaluation	21 days	Fri 3/31/23	Fri 4/28/23
9		Deployment	21 days	Mon 5/1/23	Mon 5/29/23
10		**Build Web Application**	**44 days**	**Tue 5/30/23**	**Fri 7/28/23**
11		Implement Model in Web Application	30 days	Tue 5/30/23	Mon 7/10/23
12		Demostrate the Web Application	14 days	Tue 7/11/23	Fri 7/28/23

Figure 1.2 Gantt chart of the project