Faculty of Computer Science and Information Technology

Mobile Legend: Bang Bang (MLBB) Win-Lose Prediction by Using Machine Learning

SITI RUBIAH BINTI MUSLIM

Bachelor of Computer Science with Honours

(Information System)

2023

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE   MOBILE LEGEND: BANG BANG (MLBB) WIN-LOSE PREDICTION BY USING MACHINE LEARNING

ACADEMIC SESSION: _____2022/2023_____

_____SITI RUBIAH BINTI MUSLIM_____
(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

DR. SHAPIEE ABD RAHMAN
Senior Lecturer (DS52)
Computational Science Programme
Faculty of Computer Science & Information Technology
Universiti Malaysia Sarawak

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

Permanent Address
NO 15, LORONG 10,
KAMPUNG JAYA BAKTI
JALAN MATANG
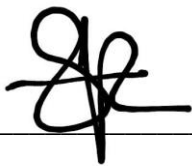
Date: 30/06/2023   Date: 30/06/2023

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

DECLARATION

I hereby declare that the work has been done by myself and no portion of the work contained in this thesis has been submitted in support of any application for any other degree or qualification of this or any other university or institute of learning.

Siti Rubiah binti Muslim

Faculty of Computational Science and Information Technology

University Malaysia Sarawak

Date: 25/01/2023

ACKNOWLEDGEMENT

To begin this endeavour on such a positive tone is a blessing. Thank you so much to everyone who has assisted me on this incredible academic path. My research would not have been feasible without the guidance and support of Dr. Shapi-ee bin Abd Rahman and Mr. Ahmad Alif bin Kamal, whom I would like to publicly thank and honour for their invaluable contributions. Their comments and suggestions helped me through the entire process of drafting my project. What is more, they sacrificed their own time in order to read my work, offer feedback, and engage in a fruitful dialogue that lighted next steps and clarified my responsibilities for this study. I would also like to express my gratitude to Dr. Wang Yin Chai, coordinator of the final year project, for holding lecture classes and providing us with helpful guidance during the course of our work.

I would also like to express my gratitude to my parents, especially for their unwavering encouragement, financial support, and patience as I worked on this project. I would like to thank you for praying for me; it has kept me going. I would like to mention my support system, my best friend in my university life: Nurin Alya, Syahazwani Nurain, Nur'Azra Alia Nisa, Nurfazlina, Afiqah Liyana, and Fateen Ezzani. Thank you for all your support, memorable moments, laughter, and tears that we shared together on this journey. You all have been really helpful in providing me with feedback, ideas, and inspiration as I worked on this study. I could not imagine going through college without you guys. Without you guys, my university life would be dull and plain.

Last but not least, I would like to thank myself for being strong, for believing in yourself throughout this research study, for all of the hard work done, for all of the stay-ups; you did an excellent job!

ABSTRACT

Mobile Legends: Bang Bang achieved the highest number of global downloads among free multiplayer online battle arena (MOBA) games, with an impressive count of over 4.7 million downloads across both Google Play and the Apple App Store combined in July 2022. Since then, a number of studies incorporating machine learning have been conducted for this mobile game, mostly focused on attempting to anticipate the actions of the players and predict the outcome of the match. This research study's goal is to propose a machine learning approach to win and loss prediction in MLBB by using Logistic Regression and to measure the performance of the machine learning model. This study involves collecting and pre-processing game statistics data, such as hero, role, kill, death, gold gained, hero damage, turret damage and damage taken from 30 gameplay where Google Colab will be used to develop and testing the model. According to the findings, the logistic regression models had achieved 0.8 accuracy indicates the model is capable of generalizing well to new data and has a reasonably good predictive performance. Overall, this study contributes to the growing body of knowledge in e-sports analytics and showcases the power of machine learning in revolutionizing competitive gaming.

ABSTRAK

Mobile Legends: Bang Bang (MLBB) ialah permainan *multiplayer online battle arena* (MOBA) yang paling banyak telah dimuat turun di seluruh dunia setakat Julai 2022, dengan lebih 4.7 juta muat turun di dalam gabungan Google Play dan Apple App Store. Sejak itu, beberapa kajian yang berkaitan dengan *machine learning* telah dijalankan untuk permainan mudah alih ini, kebanyakannya tertumpu pada percubaan untuk menjangka tindakan pemain dan meramalkan keputusan perlawanan. Objektif utama dalam kajian ini adalah untuk mencadangkan pendekatan *machine learning* untuk ramalan menang dan kalah dalam MLBB dengan menggunakan *Logistic Regression* dan mengukur prestasi model *machine learning*. Kajian ini melibatkan pengumpulan dan pemprosesan data statistic permainan seperti *hero, role, kill, deaths, gold gained, hero damage, turret damage* dan *damage taken* daripada 30 permainan dimana Google Colab akan digunakan untuk membina dan menguji model tersebut. Mengikut hasil kajian, model *logistic regression* telah mencapai 0.8 ketepatan menunjukkan model ini mampu untuk mengeneralisasikan data baru dan mempunyai prestasi ramalan yang bagus. Secara keseluruhannya, kajian ini telah menyumbang kepada peningkatan pengetahuan dalam analisis e-sukan dan mempamerkan kuasa pembelajaran mesin dalam merevolusikan permainan kompetitif.

Table of Contents

List of Figures

List of Tables

List of Equations

Chapter 1 : INTRODUCTION

1.1 Introduction

In the digital world of the 21st century, gaming has emerged as one of the most prominent forms of entertainment, with multiplayer online competitive games developing into both a social phenomenon and an activity that is increasingly acknowledged as a sport or also known as esports. In simple terms, esports refer to video games played in a well-structured competitive setting, observed by spectators, and offering rewards to the participants upon concluding the contest, following intense battles to achieve victory (Schubert et al., 2016; James, 2020). Esports encompass a wide range of games, ranging from widely recognized and favoured multiplayer online battle arenas (MOBAs) to single-player first-person shooters, thrilling survival battle royales, and even virtual simulations of real-world sports (Leroux-Parra, 2020).

MOBA, a widely recognized subgenus of real-time strategy video games (Toth et al., 2021), has gained significant popularity in recent times due to its challenging yet accessible gameplay (Segal, 2014). In a typical MOBA battle, the primary objective for both sides is to destroy the major structure of the opposing team, usually located on the opposite side of the playing area. In some variations of MOBA games, the goal can also be to eliminate every player on the opposing team (Li et al., 2016). In certain MOBA games, the goal is can be to defeat every player on the opposing team. Some of the most well-known and lucrative multiplayer online battle arenas (MOBAs) are Defense of the Ancients II (DotA2) (Porokhnenko et. al, 2019), League of Legends (LoL) (Kim et. al, 2017) and Honor of Kings 3 (HoK), with Mobile Legends: Bang Bang (Safitri et. al, 2019) also being among the most played in its genre on mobile devices.

Mobile Legends: Bang Bang, commonly referred to as ML or MLBB (Loa & Berlianto, 2022), is a mobile MOBA created and developed by Shanghai Moontoon Technology back on 2016 (Shabati, 2019) that can be played either on Android or Ios platform (Andreas & Arymami, 2022). After its initial release, this mobile game quickly gained worldwide popularity, particularly in Southeast Asia (Loa & Berlianto, 2022). In July 2022, Mobile Legends: Bang Bang achieved the highest number of global downloads among free multiplayer online battle arena (MOBA) games, with an impressive count of over 4.7 million downloads across both Google Play and the Apple App Store combined (Clement, 2022). Since then, a number of studies incorporating machine learning have been conducted for this mobile game, mostly focused on attempting to anticipate the actions and decisions of players.

Machine learning (ML) is an aspect of artificial intelligence (AI) (Nichols et. al, 2019) that enables software to enhance its predictive accuracy without explicit programming for specific outcomes (El Naqa & Murphy, 2015). ML algorithms use data from the past as input to make predictions about what the potential output will be like (Burns, 2021). When it comes to predictive modelling using ML, various number of algorithms that are available for use such as the random forest algorithm (Speiser et. al, 2019), neural network algorithm (Rodriguez-Galiano et. al, 2015), decision tree algorithm (Matzavela & Alepis, 2021) and classification algorithm.

In this research project, Logistic Regression (LR) will serve as one of the classification algorithms (Kambria, 2019). As a supervised learning method (Nurunnabi & Nasser, 2011), LR is utilized to predict the probability of a binary outcome, such as yes or no, 0 or 1, in a given event (Thanda, 2022). The subsequent sections of this project will cover the problem statement (Section 1.3), the project's scope (Section 1.4), its goals and objectives (Section 1.5), the methodology (Section 1.6), its significance (Section 1.7), the project schedule (Section 1.8), and the expected project outcome (Section 1.9).

1.2 Problem Statement

Winning is the desired goal for any game or esports that is being contested. To win the game entails defeating the player or team that is competing with, battling against, or performing better than the others involved. Esports, especially in MOBAs, have seen a significant focus on win prediction as the primary area of analytics research in both industry and academia (Hodge et al., 2019). In recent years, there has been a notable surge in popularity for match-winning prediction in Mobile Legends. This prediction of match outcomes has attracted various types of fans, including expert analysts, team managers, and others, who attempt to foresee the results before the matches commence (Abdul Rahman, 2018). In order to gain victory for the battle, it is really important to know what is the best strategy or the significant factors for a player to focus on that leads to winning the game and gaining victory.

There are few Mobile Legends predictions that have been conducted, and most of the research uses a machine learning algorithm called Naïve Bayes algorithm. Since the method used in the previous research conducted is the same, there is no variety of method in this study area. Therefore, another ML method, the LR algorithm, which is considered a classic in the area of ML, is going to be the method that will be use and explored to solve this issue by generating the outcome of the win and loss prediction of the game based on the individual performance.

1.3 Scope

In this research project, the primary focus will be on using the logistic regression approach to train machine learning models for win-lose prediction in Mobile Legends: Bang Bang, a mobile multiplayer online battle arena game. The main objective of the study is to evaluate the performance of these machine learning models in accurately predicting match outcomes. As of the end of December 2022, there are a total of 112 different sorts of heroes based on the Mobile Legend: Bang Bang Wiki websites, and the primary function of heroes can be broken down into six distinct roles: tank, marksman, assassin, fighter, mage, and support. In this application of Logistic Regression algorithm, it uses the results of game statistics as the basis of algorithm calculation, which is can be obtained after a match is done.

For the purpose of this research, a machine learning analysis will be carried out using data taken 30 gameplays that submitted from 10 respondents. Next, Google Colab is the software programme that will be utilised to perform data analysis and visualization. Within this platform, it enables importing our own dataset, using it to train an image classifier, and then evaluating the model. Lastly, for evaluating the performance of the machine learning models, the confusion matrix and F1 score will be employed, which combines the precision and recall of a classifier into a single metric by computing their harmonic mean (Wood, 2019).

1.4 Aims and Objectives

1.4.1 Aim

The main aim for this project research is to develop prediction on the winning and loss of Mobile Legends Bang Bang (MLBB) by using logistic regression algorithm as a method in machine learning.

1.4.2 Objective

The objectives of this research are as followed:

1. To propose machine learning approach on the win and lose prediction in mobile legends by using Logistic Regression.
2. To measure the performance of machine learning model.

1.5 Brief Methodology

1.5.1 Datasets

The research methods that will be used in this study are quantitative research methods to predict the winning and lose of Mobile Legends Bang Bang by adopting the Logistic Regression algorithm as a method of prediction in machine learning. Research instruments that will be used for this study are the datasets that obtained from 30 gameplays of MLBB.

1.5.2 Data pre-processing

Data pre-processing is a crucial step in machine learning, as it improves the data quality and facilitates the extraction of meaningful insights (Goyal, 2021). In the context of Mobile Legends win-lose prediction, missing values and inconsistencies in the raw data could adversely affect the accuracy of the predictions. To enhance and preserve the data quality and prediction outcomes, the raw data will undergo data preparation, including data cleaning.

Data cleaning involves addressing missing values, smoothing noisy data, and handling data discrepancies. Several methods, such as using the attribute mean or the mean of the same class for missing values, can be employed to replace missing data. Smoothing techniques like binning, clustering, regression, and human inspection can be used to handle noisy data.

For this study, the extracted dataset will undergo both data pre-processing procedures mentioned above to ensure that only relevant features are utilized for forecasting the win-lose rate of the game and to minimize the time required for the experiment.

1.5.3 Training and Testing

The training data will be utilised in order to train the machine learning model so that it can accurately predict the winning rate based on the data that is provided. In this phase, Logistic Regression algorithm is used to calculate and predict output of the win and lose prediction of data that has been processed before. The Logistic Regression algorithm is selected for its simplicity and effectiveness in supervised learning. It proves to be a powerful technique, particularly well-suited for binary and linear classification problems (Edgar & Manz, 2017). To evaluate the machine learning's models performance, confusion matrix and F1 score is

utilize where it takes the contribution of recall and precision mean which means the higher the F1 score, the more accurate the model (Meel, 2021).

1.6 Significance of Project

This research study has the potential to determine the win and lose prediction of Mobile Legends future matches by using machine learning that applies the Logistic Regression algorithm as the method for prediction, and it can also discover the features that are important to win games. With this project, it can provide a solution for players who are still confused about determining the influential factors that affect their winning, which will lead to winning the game and gaining victory. This will also make it less likely that they will lose, which will make the players enjoy the game more because they will feel more challenged once they know their strategy, which will help the team play the game more effectively and win. Aside from that, by predicting the winning and losing, players will be able to stay engaged in the game and make connections between their prior knowledge and the information gained from the prediction. It ensures that the players are able to understand their decision as they try to make logical and reasonable assumptions about what could happen next if they kill or gained gold during the game. Besides, this will also help current and aspiring professional esports athletes better gauge their chances of winning with the help of the right hero.

1.7 Project Schedule

Throughout the first and second semesters of the academic year 2022/23, the project schedule outlines the time necessary to complete the overall proposed research project. In this project, a Gantt chart is used to depict the overall schedule for FYP1 and FYP2. Figures 1.0 and 1.1 demonstrate the approximate progression of the Gantt charts for Final Year Projects 1 and 2, respectively.
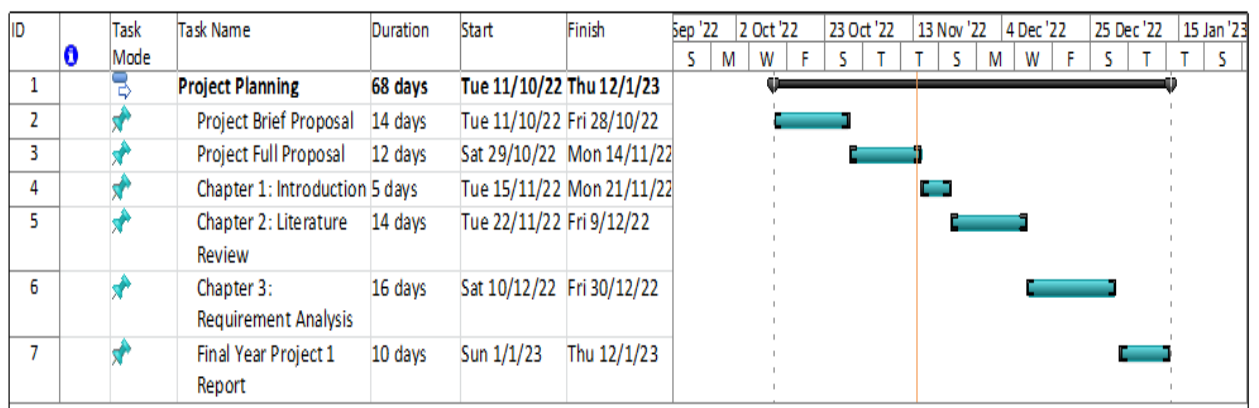

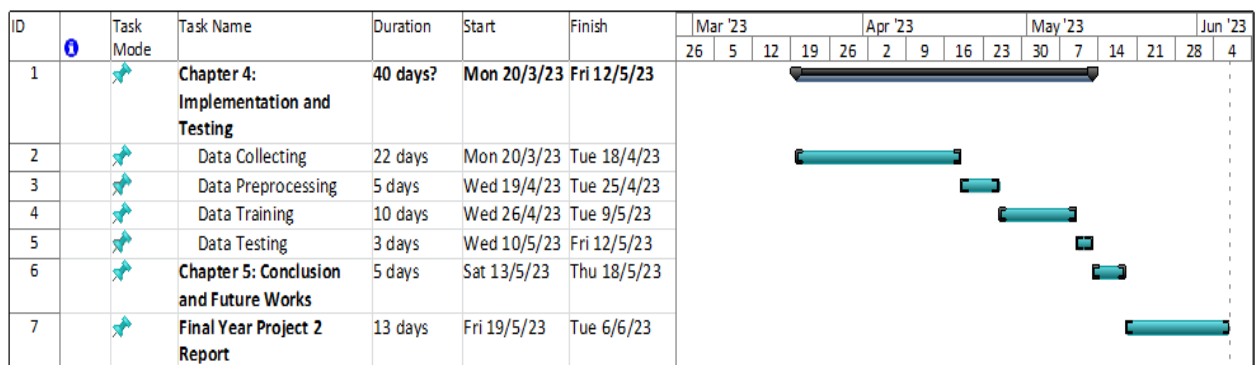*Figure 1.1* Project schedule for FYP1 Semester 2022/2023

Figure 1.2 Project schedule for FYP2 Semester 2022/2023

1.8 Expected Outcome

At the completion of this project, the winning and losing prediction of the Mobile Legends game will have been successfully obtained. Other than that, a machine-learning model using the logistic regression algorithm to predict the winning prediction will be developed. The influential factors that are significant for winning will also be determined. The logistic regression algorithm as a method in machine learning is used for prediction in this study, and in the future, it may be adapted for use in Decision Support Systems in order to improve the capabilities of the platform in terms of decision-making. In addition, the information that was gleaned from this study can be helpful for Mobile Legends stakeholders and the community that uses the Mobile Legends platform, particularly its players, in configuring their game strategies.

1.9 Summary

As a summary, this research project aims to develop a win-loss prediction model for Mobile Legends: Bang Bang (MLBB) using the Logistic Regression algorithm as a method in machine learning. The increasing popularity of esports and competitive gaming has led to a growing interest in predicting match outcomes to improve strategies and decision-making.

The significance of this project lies in its potential to provide players with a valuable tool for team strategic planning, leading to better chances of victory and a more enjoyable gaming experience. By accurately predicting match outcomes, players can stay engaged, make informed decisions, and further enhance their gameplay.

The methodology involves using quantitative research methods and datasets from 30 MLBB gameplays to train and test the Logistic Regression model. Data pre-processing is conducted to ensure data quality and consistency. The model's performance will be assessed using metrics like the F1 score, which will offer valuable insights into its accuracy.

The successful completion of this project will lead to a win-loss prediction model for MLBB, benefiting players, professional esports athletes, and the Mobile Legends community as a whole. The insights gained from this study can also be applied in Decision Support Systems to further improve gameplay strategies. By incorporating machine learning into the gaming world, this research contributes to the ever-evolving landscape of competitive gaming and its potential for innovation and growth.

Chapter 2 : LITERATURE REVIEW

2.1 Introduction

Chapter 2 of this project will concentrate on the literature review, which involves an academic piece of writing that provides context and demonstrates a comprehensive understanding of the existing academic literature on the specific subject (University of Edinburgh, 2022). Firstly, this chapter will be explaining more details on the roles and heroes that exist in the Mobile Legends Bang Bang. Secondly, this chapter will go through Mobile Legends gameplay to understand more on this research project subject matter. Besides, in this literature review, it will be explaining more on machine learning algorithm which is logistic regression and how to calculate its accuracy by using F1 score and confusion matrix. The advantages of using logistic regression as a machine learning method will also be discussed. Finally, this literature review chapter will examine the current research projects related to Mobile Legends prediction using machine learning methods.

2.2 Mobile Legends Bang Bang

Mobile Legends: Bang Bang, also referred as ML or MLBB, is a MOBA game (Adiarto, 2020) released in 2016, developed and published by Shanghai Moontoon Technology and is playable on both Android and iOS (Suprapto, 2020). It provides a vast array of features that gamers might anticipate from any other multiplayer online battle arena game. In addition to a variety of heroes to pick from, players can engage in rapid matches supported by a fast-paced fight that can keep them engaged for hours. Additionally, it is perfect for a single session on the go.