UNIVERSITI TEKNOLOGI MARA

MACHINE LEARNING MODEL FOR PERFORMANCE PREDICTION IN MOBILE NETWORK MANAGEMENT

MUHAMMAD HAZIM BIN WAHID

Thesis submitted in fulfillment
of the requirements for the degree of
Master of Science
(Electrical and Electronic Communication)

College of Engineering

December 2022

ABSTRACT

The current practice of mobile network management involves human interaction and human work ranging from conducting drive tests that can evaluate mobile network performance and coverage to identify customer complaints. Conducting drive test causing extended human work and time-consuming when processing a large and complex data. Predictive network analytics related to network management involves predicting network performance at locations or times in which no direct measurement data is available or solving missing values in the data collection when conducting the drive test. One of the major challenges when applying machine learning is to identify the best algorithm from a variety of algorithms to solve a problem. This study aims to propose the best machine learning algorithm for predicting mobile network performance. This can be achieved by comparing several types of machine learning algorithms in predicting mobile network performance. The methodology includes drive test measurement for data collection, exploratory data analysis, data preparation, and applying machine learning algorithms to predict mobile network performance. Since throughput feature has a strong correlation with the signal strength performance, it is the targeted parameter in the network performance prediction. Three machine learning algorithms were tested in this study which are random forest, Gaussian process regression, and K-Nearest Neighbor (KNN) for throughput prediction. Based on the results and analysis of the evaluation metric comparison, it shows that the random forest model is the best model that comes with the highest performance prediction with the R^2 score of 0.7919 followed by KNN 0.66 and Gaussian process regression 0.34. The random forest also gets the lowest value for the evaluation metric error. Random forest achieved the best result because of an additional layer of randomness that can lessen the variance thus increasing the model accuracy. Using the hyperparameter tuning technique to adjust the number of trees in the forest and the value for the depth of each tree in the forest, it will increase the random forest model performance and accuracy by 0.57%. Based on the feature importance list of the random forest model, the location of the measurement and signal-to-noise ratio (SNR) feature plays an important role that affecting network performance prediction.

ACKNOWLEDGEMENT

Firstly, I wish to thank God for giving me the opportunity to embark on my Master's and for completing this long and challenging journey successfully. My gratitude and thanks go to my supervisor Dr. Nur Idora Binti Abdul Razak and my co-supervisor Dr. Syahrul Afzal Bin Che Abdullah. This project is funded by UiTM under Geran Penyelidikan Khas (600-RMC/GPK 573 (209/2020))

My appreciation goes to the staff and lab assistant of the Wireless Communication lab who provided the facilities and assistance during data measurement and collection. Also lending me equipment for my research study. Special thanks to my colleagues and special friend for helping me with this project.

Finally, this thesis is dedicated to the loving that is for my very dear father and mother for the vision and determination to educate me and put a very high hope on me to finish my Master's journey. That always supports me mentally and financially till the very end. This piece of victory is dedicated to both you and my family. Alhamdulilah.

TABLE OF CONTENTS

		Page
CONFIRMATION BY PANEL OF EXAMINERS		**ii**
AUTHOR'S DECLARATION		**iii**
ABSTRACT		**iv**
ACKNOWLEDGEMENT		**v**
TABLE OF CONTENTS		**vi**
LIST OF TABLES		**ix**
LIST OF FIGURES		**x**
LIST OF SYMBOLS		**xi**
LIST OF ABBREVIATIONS		**xii**
CHAPTER ONE INTRODUCTION		**1**
1.1	Introduction	1
1.2	Research Background	1
1.3	Problem Statement	6
1.4	Objectives	8
1.5	Scope and Limitation of Study	8
1.6	Significance of Study	9
1.7	Organization of the Thesis	10
1.8	Conclusion	10
CHAPTER TWO LITERATURE REVIEW		**11**
2.1	Introduction	11
2.2	Machine learning Challenges in Mobile Network	11
2.3	Key Factors in Performing Machine Learning	14
2.4	Dataset Used in Mobile Network Performance Prediction	16
2.5	Parameter Used in Mobile Network Performance Prediction	19
2.6	Machine Learning algorithm in Mobile Network Prediction	21
2.7	Machine Learning Evaluation	25
2.8	Conclusion	27

CHAPTER THREE RESEARCH METHODOLOGY **29**
3.1 Introduction 29
3.2 Flowchart of Research Activities 29
3.3 Dataset Collection and Measurement 31
3.4 Exploratory Data Analysis 33
3.4.1 Data visualization 33
3.5 Machine Learning Algorithm 35
3.5.1 Random Forest 36
3.5.2 Gaussian Process Regression 39
3.5.3 K-Nearest Neighbor (KNN) 42
3.6 Evaluation Metrics 45
3.6.1 Mean Absolute Error 46
3.6.2 Mean Squared Error and Root Mean Squared Error 46
3.6.3 Coefficient of Determination 47
3.7 Conclusion 47

CHAPTER FOUR RESULTS AND DISCUSSION **48**
4.1 Introduction 48
4.2 Variable Importance List 48
4.3 Predictive Model Performance 51
4.3.1 Random Forest 51
4.3.2 Gaussian Process Regression 52
4.3.3 K Nearest Neighbour 54
4.4 Comparison Analysis 55
4.5 Random Forest Hyperparameter Tuning 58
4.6 Conclusion 60

CHAPTER FIVE CONCLUSION **61**
5.1 Introduction 61
5.2 Summary 61
5.3 Potential Future Research 63