Universiti Teknologi MARA

Prediction of MUET result based on KNN algorithm

Siti Fatimah Azzahra Binti Hamrizan

Thesis submitted in fulfilment of the requirements for Bachelor of Computer Science (Hons.) Faculty of Computer Science and Mathematical Sciences

February 2021

ACKNOWLEDGEMENT

Alhamdulillah, praises and thanks to Allah because of His Almighty and His utmost blessings, I was able to finish the research within the time duration given. Firstly, my special thanks go to my supervisor, Dr Norlina Mohd Sabri for providing invaluable guidance throughout this research. Special appreciation also goes to my beloved parents Hamrizan Abd Hamid and Aida Mustapha. Finally, I would like to give my gratitude to my dearest friend Siti Syazwana Jafri.

ABSTRACT

Malaysian University English Test (MUET) has been used to measure students' English proficiency level before entering bachelor programmes in Malaysian universities. UiTM offers English language proficiency courses (ELC) for semester 1 to semester 3 diploma students to prepare them for MUET. The exam results from the ELC can be the features to help the students to know if they can pass the MUET exam and the lecturers can figure out which language skills need to be improved to help preparing the students to sit for MUET. The study aims to explore the K-Nearest Neighbour (KNN) algorithm in solving the MUET result prediction problem, to develop a prototype of MUET result prediction based on the KNN algorithm and to evaluate the accuracy of the KNN algorithm in MUET result prediction. The machine learning technique used to develop the prediction prototype is the KNN algorithm. The results show that the highest average accuracy was at 65.294% and the percentage error was at 25.490%. Some additional future works should be applied to improve the algorithm performance are to add more data, use other related results as the features and compare the KNN algorithm to other algorithm.

TABLE OF CONTENTS

CONTENT		PAGE		
SUPERVISOR APPROVAL	i			
STUDENT DECLARATION	ii			
ACKNOWLEDGEMENT		iii		
ABSTRACT		iv		
TABLE OF CONTENTS	v			
LIST OF FIGURES		viii		
LIST OF TABLES		x		
LIST OF ABBREVIATIONS	xi			
CHAPTER ONE: INTRODUCTION				
1.1. Background of Study		1		
1.2. Problem Statement		2		
1.3. Objective		4		
1.4. Project Scope		4		
1.5. Project Significance		5		
1.6. Overview of Research Framework		5		
1.7. Conclusion		6		
CHAPTER TWO: LITERATURE REVIEW				
2.1. Machine Learning		8		
2.1.1. Prediction in Various Area		9		
2.2. Prediction of Exam Result		10		
2.3. K-Nearest Neighbour (KNN) 12				
2.3.1. KNN and how does it work		13		
2.3.2. Choosing the value of K		14		
r.				

2.	.3.3.	Pseudocode of KNN algorithm		14	
2.	.3.4.	Flowchart of KNN algorithm		15	
2.4. Ir	mplei	mentation of KNN algorithm		16	
2.5. S	Simila	r works		21	
2.6. T	The In	nplication of Literature Review	24		
2.7. C	Conclu	usion		26	
CHAPTER	R THR	EE: RESEARCH METHODOLOGY			
3.1. C	Dverv	iew of Research Methodology		27	
3.	.1.1.	Detailed of Research Framework		27	
3.2. P	Prelim	ninary Phase		30	
3.	.2.1.	Literature study		30	
3.	.2.2.	Data Pre-processing		30	
	3.2	.2.1. Data Collection		31	
	3.2	.2.2. Data Cleaning			31
3.3. D	Desigr	n Phase		32	
3.	.3.1.	System Architecture		32	
3.	.3.2.	Flowchart		33	
3.	.3.3.	User Interface Design	35		
3.	.3.4.	Pseudocode of Selected Algorithm		35	
3.	.3.5.	Prototype Implementation		36	
3.4. P	Perfor	mance Evaluation		36	
3.	.4.1.	k-Fold Cross-Validation		37	
3.	.4.2.	Prediction Percentage Error, Mean Squared Error (N	∕ISE)		
		and Accuracy		38	
3.5. C	Conclu	usion		39	
СНАР	PTER I	FOUR: RESULT AND FINDING			
4.1. C	Conce	ptual Framework		41	
4.2. Program Codes for Prototype			42		
4.	.2.1.	Calculate Euclidean Distance	43		
4.	.2.2.	Get Nearest Neighbour		43	