

**Universiti Teknologi MARA**

**Sentiment Analysis of Restaurant Reviews in Kuala  
Terengganu Based on K-Nearest Neighbor**

**Siti Syazwana Binti Jafri**

**Thesis submitted in fulfilment of the requirements for  
Bachelor of Computer Science and Mathematics (Hons.)  
Faculty of Computer 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 this research within the time duration given. Firstly, my special thanks go to my supervisor, Dr. Norlina binti Mohd Sabri who has given guide for this project from start to the end of project and helped a lot about this project. Special appreciation also goes to my beloved parents that put a lot of trust on doing this Final Year Project. Last but not least, I would like to give my gratitude to my dearest friend that always give support in completing this project and shared the knowledge about the related studies in this project..

## **ABSTRACT**

Sentiment analysis nowadays being the important role in many of industries that especially for the things or works that related to the review or feedback from the individual in the cyberspace. The people may review for the places, products and other things by expressing their emotion or opinions into a sentence. This action may lead to the problem of the understanding the meaning or description behind the texts and it also difficult to discover the sentiment polarity of the certain words. Some of the restaurant in Kuala Terengganu may be lacks in the term of the promotional the restaurants and any related activities often overlooked the reviews from the customers about the many different aspects of the restaurant. However, negative reviews will be affected the image of the restaurants. This study will perform the sentiment analysis on the restaurant reviews in Kuala Terengganu on the TripAdvisor. The study will be identified sentiment analysis tasks based on the classification model. A classifier will be designed and developed which is K- Nearest Neighbor (KNN). Lastly, the accuracy of the proposed classifier will be tested. The chosen technique is classification and the algorithm that will be applied in the classification process is K- Nearest Neighbor (KNN). The output will be the accuracy of the KNN model and the visualization of sentiment analysis of the new data that user will choose in the prototype. The accuracy achieved is 83%. In the future, it is very recommended to experiment with different algorithm. The volume of data should be large as it can generate better result of classification method.

## TABLE OF CONTENT

<b>CONTENT</b>	<b>PAGE</b>
<b>SUPERVISOR APPROVAL</b>	<b>ii</b>
<b>STUDENT DECLARATION</b>	<b>iii</b>
<b>ACKNOWLEDGEMENT</b>	<b>iv</b>
<b>ABSTRACT</b>	<b>v</b>
<b>TABLE OF CONTENTS</b>	<b>vi</b>
<b>LIST OF FIGURES</b>	<b>ix</b>
<b>LIST OF TABLES</b>	<b>xi</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xii</b>
<b>CHAPTER ONE: INTRODUCTION</b>	<b>1</b>
1.1 Background of Study	1
1.2 Problem Statement	2
1.3 Objective	3
1.4 Project Scope	3
1.5 Project Significance	3
1.6 Overview of Research Framework	4
1.7 Summary	5
<b>CHAPTER 2: LITERATURE REVIEW</b>	<b>6</b>
2.1 Sentiment Analysis	6
2.2 K-Nearest Neighbor Algorithm (KNN)	7
2.2.1 KNN and How Does It Works	7

2.2.2	Choosing value of K	7
2.2.3	Pseudocode of KNN Algorithm	8
2.3	Implementation KNN in Various of Problems	10
2.4	Sentiment Analysis in Various Fields	12
2.5	Sentiment Analysis based on KNN	15
2.6	Implication of Literature Review	16
2.7	Conclusion	18
 <b>CHAPTER 3: RESEARCH METHODOLOGY</b>		 <b>19</b>
3.1	Overview of Research Methodology	19
3.2	Preliminary Study	21
3.2.1	Literature Study	21
3.2.2	Data Collection	21
3.2.3	Data Preparation	22
3.2.3.1	Data pre-processing	22
3.3	Design Phase	23
3.3.1	Interface Design	23
3.4	Classifier and Development	24
3.4.1	System Architecture	24
3.4.2	Pseudocode	26
3.5	Prototype Implementation	27
3.6	Hardware and software requirement	27
3.7	Performance Evaluation	27
3.7.1	10-Fold Cross- Validation Technique	28
3.7.2	Accuracy Evaluation	28
3.8	Conclusion	28
 <b>CHAPTER 4: RESULT AND FINDINGS</b>		 <b>31</b>
4.1	Conceptual Framework	31
4.2	Data Collection	32