Universiti Teknologi MARA

Sentiment Analysis for Mobile Brands Reviews using Convolutional Neural Networks (CNN)

Puteri Ika Shazereen Ibrahim

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

AUGUST 2020

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 goes to my supervisor, Ts. Prof Madya Dr. Hamidah Binti Jantan, who giving me the guidance and information throughout the whole supervision process in completing the research. Moreover, I would like to convey my gratitude to my examiner, Dr. Norlina Binti Mohd Sabri for her well-worth comments and suggestions regarding this project. Special appreciation also goes to my beloved parents that who gave a lot of emotional support and never ending prayers for me throughout this project. Finally, I would like to give my gratitude to my dearest classmates, seniors and friends for sharing the knowledge that help me during development of the project. Last but not least, thank you to those who were involved directly or indirectly in making this project a success.

ABSTRACT

Sentiment analysis is widely used to mine the web content on web, including texts, blogs, reviews, and comments. It is used to analyse the people opinion. As for product company to monitor their product is by reading all the reviews which is the tedious task and take a lot of time. Sentiment analysis can help to analyze these opinioned data and extract some important insights which will help other users to make the decision. However, because a lack of sufficient labeled data in the field of Natural Language Processing (NLP) it is becoming challenging. To solve that, the CNN approach has been proposed in sentiment analysis because of it learning ability and extract the important feature. The methodology for the project has been designed for CNN implementation that consists of five phases which are preliminary study, data collection, system design, system implementation, and result. The CNN model has shown great accuracy. The parameter tuning has been done to get the great accuracy of the model. The result of the experiment shows that CNN gives high accuracy when the train and test split is 90:10, with 500 input vector, feature map = 200, filter size = 8-12, dropout= 0.1, and batch size = 64. This combination has produced a higher accuracy of the CNN model in this project which is 95%. It is shown that this model needs more input or data to make it more accurate. The evaluation result of this model also has shown that this model is good for sentiment analysis. For future work to improve this project, more data will be used for training and Malay reviews also will be include. After that, real-time data also will be available for users monitoring their products.

TABLE OF CONTENTS

CONTENT	PAGE
SUPERVISOR APPROVAL	**II**
STUDENT DECLARATION	**III**
ACKNOWLEDGEMENT	**IV**
ABSTRACT	**V**
TABLE OF CONTENTS	**VI**
LIST OF FIGURES	**IX**
LIST OF TABLES	**XI**
CHAPTER ONE: INTRODUCTION	
1.1 Background of Study	1
1.2 Problem Statement	3
1.3 Objective	4
1.4 Scope	4
1.5 Significance of Study	5
1.6 Project Framework	5
1.7 Conclusion	6
CHAPTER TWO: LITERATURE REVIEWS	
2.1 Sentiment Analysis	7
2.1.1 Overview of Sentiment Analysis	7
2.1.2 Issues of Sentiment Analysis	8
2.1.3 Sentiment Analysis Technique and Approach	8
2.2 Text mining in Machine Learning	9
2.2.1 Overview of Text Mining	9
2.2.2 Text Mining Technique	10
2.2.3 Text Mining in Sentiment Analysis Applications and Approach	11
2.3 Convolutional Neural Network	12

2.3.1 Overview of Convolutional Neural Network 12
2.3.1 Fundamental of CNN 13
2.3.2 Advantages of CNN 15
2.3.3 Related studies in CNN 16
2.4 Similar Applications of CNN in Sentiment Analysis 17
2.4.1 Issues in CNN Application 20
2.5 Conclusion 21

CHAPTER THREE: METHODOLOGY

3.1 Sentiment Analysis using CNN Methodology Framework 22
3.2 Analysis Phase 25
3.2.1 Data Collection 25
3.2.2 Data pre-processing and preparation 27
3.3 Design Phase 31
3.3.1 Sentiment Analysis with CNN Algorithm 31
3.3.2 CNN Algorithm for Sentiment Analysis 33
3.3.3 Experiment Setup 36
3.4 Implementation Phase 37
3.4.1 Implement the CNN Algorithm 38
3.4.1 Prototype Development 39
3.4.2 Debugging and testing 40
3.5 Evaluation Phase 40
3.6 Conclusion 42

CHAPTER FOUR: RESULTS AND FINDINGS

4.1 Sentiment analysis for mobile brands reviews using convolutional neural network (CNN) conceptual framework 43
4.1.1 Data analysis and representation 44
4.1.2 CNN Classifier Development 47
4.2 Result Analysis of Experiments 49
4.2.1 Experimental Analysis 49
4.3 Evaluation on CNN model and Prototype of Sentiment Analysis using CNN Classifier. 52
4.3.1 Computational Analysis based on Recall and precision measure 53