

Universiti Teknologi MARA

**Sentiment Analysis Regarding Childcare Issues
Using Naïve Bayes Algorithm**

ALIS FARHANA BINTI ZULKIPELI

BACHELOR OF COMPUTER SCIENCE (HONS.)

Jan 2025

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious, the Most Merciful. All praises to Allah and His blessing for the completion of this research and granting me the opportunity to pursue higher education.

My sincere appreciation to my supervisor, Ts. Dr. Hasiah Binti Mohamed @ Omar, for her patience, invaluable guidance, and efforts in steering me through the challenges of this research project. Special thanks to Madam Ummu Fatimah Mohd Bahrin for their insightful lessons and comment throughout this project. All of their encouragement has significantly contributed to the development of my understanding in the field.

I offer my special thanks to all my family members. It would not be possible to finish this project without support from them. I would like to thank my dearest father Zulkipeli Ibrahim, my mother Kamariah Ibrahim, and all my beloved siblings. It was indeed helpful to have family members who were able to understand my situation while writing this research.

I want to extend my deepest appreciation to all the beloved friends who stand by me no matter the challenges we face together. I benefited from getting guidance and encouragement during my research time. The growth with important lessons learned fills me with gratitude. All those who support me form the network that achieved this great accomplishment through their peaceful actions. I feel incredibly fortunate and express my sincere thanks to everyone once again.

ABSTRACT

This study applies the Naïve Bayes algorithm for sentiment analysis to assess public perceptions of childcare issues, particularly child abandonment and accidents. With the growing volume of childcare-related discussions on social media, efficient sentiment analysis tools are essential for extracting insights. However, the lack of comprehensive methodologies poses challenges for policymakers, childcare providers, and researchers in understanding public concerns and developing effective interventions. To address this, a dataset of 1,079 tweets from X (formerly Twitter) is analyzed. The data undergoes preprocessing steps such as stop-word and emoji removal, tokenization, and feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF). VADER is used for initial sentiment labeling, and the Naïve Bayes classifier categorizes sentiments into positive and negative classes. The motivation behind this project is to leverage sentiment analysis for enhancing childcare policies and public awareness. The project aims to enhance childcare policies and public awareness by leveraging sentiment analysis to bridge the gap between public sentiment and policy decisions. The Naïve Bayes model achieves 87% accuracy with high precision, recall, and F1 scores using 10-fold cross-validation, demonstrating its effectiveness in classifying social media sentiments. Future research could explore advanced techniques like Bidirectional Encoder Representations from Transformers (BERT) or Recurrent Neural Networks (RNNs) to improve classification accuracy and contextual understanding. Expanding the dataset to include multilingual content and incorporating topic modeling techniques would further enhance sentiment analysis in childcare-related discourse.

TABLE OF CONTENTS

CONTENT	PAGE
SUPERVISOR APPROVAL	ii
STUDENT DECLARATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
LIST OF FIGURES	ix
LIST OF TABLES	xi
LIST OF EQUATIONS	xii
LIST OF ABBREVIATIONS	xiii
CHAPTER 1: INTRODUCTION	1
1.1 Background of Study	1
1.2 Problem Statement	3
1.3 Objectives	5
1.4 Scope	6
1.5 Significance	7
1.6 Conclusion	9
CHAPTER 2: LITERATURE REVIEW	10
2.1 Sentiment Analysis	10
2.1.1 Overview of Sentiment Analysis	11
2.1.2 Application in Sentiment Analysis	11
2.1.3 Method and Technique Used in Sentiment Analysis	14
2.2 Naïve Bayes Algorithm	17
2.2.1 Overview of Naïve Bayes Algorithm	17
2.2.2 Application of Naïve Bayes Algorithm in Sentiment Analysis	18
2.2.3 Related Works using Naïve Bayes Algorithm in Sentiment Analysis	20

2.3	Childcare Issues And Public Perceptions	28
2.3.1	Overview of Childcare Issues	28
2.3.2	Public Perception Regarding Childcare Issues	29
2.3.3	Factors Affecting Public Perception Regarding Childcare Issues	31
2.3.4	Implications of Public Perception Regarding Childcare Issues	33
2.4	Similar Application	35
2.5	Implication Of Literature Review	40
2.6	Conclusion	42
CHAPTER 3: METHODOLOGY		43
3.1	Overview of Research Framework Methodology	43
3.2	Preliminary Study	46
3.2.1	Problem Statement Identification	46
3.2.2	Knowledge Acquisition	47
3.3	Data Collection and Preparation	48
3.3.1	Data Collection	48
3.3.2	Data Preprocessing	49
3.4	Design	52
3.4.1	Design Sentiment Analyzer	52
3.5	Development	60
3.5.1	Model Training	60
3.5.2	Testing	62
3.5.3	Development of Sentiment Analyzer	63
3.6	Evaluation	64
3.6.1	Evaluation	64
3.7	Conclusion	67
CHAPTER 4: RESULTS AND DISCUSSION		68
4.1	Program Codes	68
4.1.1	Data Description	68
4.1.2	Data Pre-Processing Results	69
4.1.3	Feature Extraction	77