UNIVERSITI TEKNOLOGI MARA

SENTIMENT ANALYSIS REGARDING MARITAL ISSUES USING NAÏVE BAYES ALGORITHM

FARAH NABILA BINTI MOHD RAZALI

BACHELOR OF COMPUTER SCIENCE (Hons.)

Jan 2025

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to Allah, the Most Gracious and the Most Merciful, for granting me the strength, knowledge, and perseverance to complete this final year project. Without His blessings, none of this would have been possible.

I would also like to extend my heartfelt thanks to my parents, whose unwavering support and encouragement have been my pillars of strength throughout this journey. Their constant prayers, sacrifices, and belief in my abilities have been a source of immense motivation for me.

I am deeply indebted to my supervisor, Ts. Dr. Hasiah Binti Mohamed @ Omar, whose expertise, patience, and constructive feedback have been crucial in shaping this project. Your continuous support and willingness to help have made this challenging task much more manageable. Your mentorship has not only guided me academically but has also played a vital role in my personal growth.

My sincerest appreciation goes to my lecturer, Madam Ummu Fatihah Binti Mohd Bahrin, for providing me with invaluable guidance and insights that have significantly contributed to the success of this project. Your dedication to teaching and your passion for the subject matter have inspired me to strive for excellence.

In conclusion, this project is a culmination of the support, guidance, and encouragement I have received from many individuals. To all who have contributed in one way or another, I extend my deepest gratitude.

ABSTRACT

This study explores the application of sentiment analysis using the Naïve Bayes algorithm to understand public perceptions of marital issues, particularly factors contributing to the rising divorce rate. By analyzing data from social media, primarily Twitter, the research identifies key challenges in marriages, including communication breakdowns, financial stress, and infidelity. The Naïve Bayes algorithm was chosen for its efficiency in text classification and ability to handle large volumes of unstructured data. The results indicate that financial instability and poor communication are the most prevalent issues, with the overall sentiment being predominantly negative. The model's performance was evaluated using accuracy, precision, recall, and F1-score, demonstrating its reliability in sentiment classification. Future research can enhance classification accuracy by incorporating advanced machine learning techniques and expanding the dataset to include diverse social media platforms. These insights can assist policymakers, mental health professionals, and marriage counselors in developing targeted interventions to support healthier relationships and strengthen societal well-being.

Table of Contents

ACKNOWLEDGEMENT **4**

ABSTRACT **5**

CHAPTER 1: INTRODUCTION **8**

1.1 Background of Study 8

1.2 Problem Statement 11

1.3 Objectives 12

1.4 Project Scope 13

1.5 Significance 14

1.6 Conclusion 15

CHAPTER 2: LITERATURE REVIEW **16**

2.1 Sentiment Analysis 16

2.1.1 Overview of Sentiment Analysis 17

2.1.2 Application in Sentiment Analysis 18

2.1.3 Method and Technique Used in Sentiment Analysis 19

2.2 Naïve Bayes Algorithm 23

2.2.1 Overview of Naïve Bayes Algorithm 23

2.2.2 Application of Naïve Bayes Algorithm in Sentiment Analysis 24

2.2.3 Related Works using Naïve Bayes Algorithm in Sentiment Analysis 26

2.3 Marital Issues and Public Perceptions 34

2.3.1 Overview of Marital Issues 34

2.3.2 Public Perception Regarding Marital Issues 34

2.3.3 Factors Affecting Public Perception Regarding Marital Issues 36

2.3.4 Implications of Public Perception Regarding Marital Issues 38

2.4 Similar Application 40

2.5 Implication Of Literature Review 45

2.6 Conclusion 46

CHAPTER 3: METHODOLOGY **47**

3.1 Overview of Research Framework Methodology 47

3.2 Preliminary Study 51

3.2.1 Problem Statement Identification 51

3.2.2 Knowledge Acquisition 52

3.3 Data Collection and Preparation 53

3.3.1 Data Collection 53

	3.3.2 Data Preprocessing	54
3.4	Design	56
	3.4.1 Design Prototype	56
3.5	Development	63
	3.5.1 Model Training	63
	3.5.2 Development of Sentiment Analyzer	64
3.6	Test and Evaluation	65
	3.6.1 Testing	65
	3.6.1 Evaluation	65
3.7	Conclusion	68
CHAPTER 4: RESULTS AND DISCUSSION		**69**
4.1	Program Code	69
	4.1.1 Data Description	69
	4.1.2 Data Pre-Processing	70
	4.1.3 Feature Extraction	78
	4.1.4 Naïve Bayes Algorithm Implementation	80
	4.1.5 K-Fold Cross-Validation	82
4.2	User Interface of sentiment Analyzer Dashboard	84
4.3	Evaluation Results	87
	4.3.1 Word Cloud Visualization	87
	4.3.2 Confusion Matrices for Each Fold	90
4.4	Discussion	96
4.5	Research Summary	98
4.6	Conclusion	100
CHAPTER 5: CONCLUSION		**101**
5.1	Research Summary	101
5.2	Project Limitation	103
5.3	Future Recommendations	104
5.4	Conclusion	106
REFERENCES		**107**