

UNIVERSITI TEKNOLOGI MARA

**ANALYZING CUSTOMER REVIEWS FOR ARBA
TRAVEL USING SENTIMENT ANALYSIS**

NURULAIN BINTI ABDULLAH

**BACHELOR OF INFORMATION TECHNOLOGY (Hons.)
BUSINESS COMPUTING**

JULY 2025

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. First and foremost, I would like to express my deepest gratitude to my supervisor, Ts. Dr. Hasiah Binti Mohamad @ Omar, for her continuous support, invaluable guidance, and constructive feedback throughout the development of this Final Year Project. Her encouragement and professional insight have been important to the completion of this research. Special appreciation also goes to my lecturer, Miss Hasnul Azirah Binti Abdul Hamid who taught and guided me in managing the Final Year Project process. Her structured teaching and support have helped me stay focused and organized throughout the semester. To my friends and coursemates, thank you for being by my side, sharing knowledge, and making this journey less stressful and more memorable. Your companionship throughout this degree has meant a lot to me. Lastly, I would like to take a moment to sincerely thank myself for not giving up, for showing up every day even when things got overwhelming, for pushing through sleepless nights, for learning from mistakes, and for staying committed to finishing what I started. This journey has taught me patience and the value of self-belief, and for that, I am truly grateful.

ABSTRACT

Sentiment analysis is a natural language processing (NLP) that categorizes customer feedback into as positive, negative, or neutral. The project entails gathering a dataset of customer reviews from Google Reviews and Facebook, cleaning the text to eliminate any noise, and analyzing sentiments using three machine learning algorithms; Naïve Bayes, Support Vector Machine, and Logistic Regression. Key phrases are extracted from 400 customer reviews on Google Reviews and 300 from Facebook to reveal common themes in customer experiences. Data visualization techniques are used to present trends, such as the frequency of positive and negative reviews over time and the most commonly mentioned aspects of ARBA Travel's services. The findings offer actionable recommendations for ARBA Travel to improve its offerings, address issues, and reinforce strengths. The integration of sentiment analysis into the feedback evaluation process at ARBA Travel will enhance its decision-making processes, refinement in customer engagement approaches, and sustain competitiveness within the travel industry. This research illustrates the potential of sentiment analysis to convert unstructured customers' feedback into meaningful insights that drive improved business results and stronger customer relationships. Three machine learning algorithms which are Naïve Bayes, Logistic Regression, and Support Vector Machine, were implemented and evaluated using cross-validation and performance metrics such as accuracy, precision, recall, and F1-score. Among these, Naïve Bayes achieved the highest performance with an accuracy of 93.67% and an F1-score of 93.54%, making it the most effective model for sentiment classification in this context. The results were visualized using an interactive dashboard developed in Power BI, allowing users to explore sentiment trends, keyword frequency, and review distributions by gender, platform, and time. Future enhancements include expanding the analysis to other platforms like TripAdvisor and Instagram for broader sentiment coverage. Implementing real-time sentiment analysis would enable faster responses to customer feedback. Additionally, enhancing the dashboard with features like trend forecasting, customer segmentation, and automated alerts can support more proactive and data-driven decision-making for ARBA Travel.

TABLE OF CONTENTS

CONTENT	PAGE
SUPERVISOR APPROVAL	ii
STUDENT DECLARATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	x
LIST OF TABLES	xii
LIST OF ABBREVIATIONS	xiii

CHAPTER ONE: INTRODUCTION

1.1	Background of Study	1
1.2	Current Business Process	3
1.3	Problem Statement	4
1.4	Objectives.....	5
1.5	Scope.....	5
1.6	Significance.....	6
1.7	Project Framework.....	7
1.8	Gantt Chart.....	9
1.9	Conclusion	9

CHAPTER TWO: LITERATURE REVIEW

2.1	Introduction.....	11
2.2	Business Analytics	11
2.2.1	Data Mining	13
2.2.2	Data Mining in Tourism Industry	15

2.3	Sentiment Analysis.....	16
2.3.1	Sentiment Classification Approaches.....	17
2.3.2	Aspect-Based Sentiment Analysis.....	19
2.3.3	Advantages of Sentiment Analysis.....	21
2.3.4	Sentiment Analysis in Tourism Industry.....	23
2.4	Cross-Industry Standard Process for Data Mining.....	24
2.4.1	Business Understanding.....	26
2.4.2	Data Understanding.....	27
2.4.3	Data Preparation.....	29
2.4.4	Modelling.....	30
2.4.5	Evaluation.....	31
2.4.6	Deployment.....	32
2.5	Similar Existing Works.....	33
2.6	Implication of Literature Review.....	35
2.7	Conclusion.....	37

CHAPTER THREE: METHODOLOGY

3.1	Introduction.....	38
3.2	Cross-Industry Standard Process for Data Mining (CRISP-DM) ..	38
3.3	Business Understanding.....	40
3.4	Data Understanding.....	41
3.5	Data Preparation.....	43
3.5.1	Data Format Conversion.....	43
3.5.2	Text Data Cleaning.....	44
3.5.3	Remove Stop Words.....	45
3.5.4	Tokenization.....	46
3.5.5	Stemming.....	47
3.5.6	Text Labelling.....	48
3.5.7	TF-IDF.....	49
3.6	Modeling.....	52
3.6.1	VADER.....	53
3.7	Evaluation.....	54