

UNIVERSITI TEKNOLOGI MARA

**PERFUME RECOMMENDATION
SYSTEM USING CONTENT-BASED
FILTERING ALGORITHM**

MUHAMMAD BAIHAQI BIN BUKHORI

BACHELOR OF COMPUTER SCIENCE (Hons.)

JANUARY 2025

ACKNOWLEDGEMENT

In the name of Allah, the Most Gracious and the Most Merciful. All praises to Allah for His blessings that enabled the completion of this research. I am grateful for all the opportunities, trials, and strength granted to me in finishing this report.

First and foremost, I sincerely thank my lecturer, Madam Ummu Fatimah binti Mohd Bahrin, for her guidance, understanding, and patience. Her positive encouragement and thorough explanations have been invaluable. It has been a great honor and pleasure to work under her direction.

My deepest gratitude goes to Dr. Najiahtul Syafiqah Binti Ismail, my supervisor for this project, for her full cooperation and willingness to review my research. Her wealth of knowledge and expertise, especially in the computer science industry, has been a tremendous help in completing my research.

I extend my special thanks to all my family members. Their support made it possible to complete this case study. I am particularly grateful to my father Bukhori Bin Mohd Daud, my mother Habisah Binti Ismail, and all my beloved siblings for their understanding during this period.

Lastly, I sincerely thank all my dear friends who have supported me through thick and thin. Their advice and encouragement during the research sessions were invaluable.

May God shower all the mentioned individuals with success and honour in their lives.

ABSTRACT

This research investigates the development of a perfume recommendation system using a content-based filtering approach. The system is designed to provide personalized recommendations by matching perfume characteristics, including scent, concentration, and department, with user preferences. The methodology involves processing a curated dataset from Kaggle, applying TF-IDF vectorization to analyze perfume attributes, and utilizing cosine similarity to generate recommendations. The system was tested across three evaluations, achieving an average Precision of 0.77 (77%), Recall of 0.68 (68%), and an F1-Score of 0.72 (72%). The results indicate that content-based filtering identifies relevant perfumes while improving user satisfaction and reducing decision-making time. The study focuses on Malaysian users who seek personalized perfume recommendations suited to the country's hot climate. These findings demonstrate the potential of content-based filtering in revolutionizing the perfume discovery process.

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 ABBREVIATIONS	xii
CHAPTER 1	1
1.1 Background of Study	1
1.2 Problem Statement	2
1.3 Objective	3
1.4 Project Scope	3
1.5 Project Significance	4
1.6 Overview of Research Framework	5
1.7 Conclusion	6
CHAPTER 2	7
2.1 Introduction	7
2.2 Recommendation System	7
2.2.1 Overview of Recommender System	8
2.2.2 Phase of Recommendation Process	8
2.2.2.2 Learning Phase	10
2.2.3 Recommendation System Technique	11
2.2.4 Advantages and Disadvantages of Recommendation System Technique	13
2.3 Perfume Recommendation	15
2.3.1 The Recommendation in Getting Perfumes	16

2.3.2	Problem with Perfume Recommendation	16
2.3.3	Benefits of Perfume Recommendation Existence	17
2.4	Understanding Content-Based Filtering Algorithms	17
2.4.1	Cosine Similarity	17
2.4.2	Term Frequency-Inverse Document Frequency (TF-IDF) Similarity	18
2.5	Similar Works	25
2.6	The Implication of Literature Review	30
2.7	Conclusion	31
CHAPTER 3		32
3.1	Overview of Research Methodology	32
3.1.1	Detailed of Research Framework	32
3.2	Preliminary Phase	35
3.2.1	Literature Study	35
3.2.2	Data Pre-processing	35
3.3	Design Phase	37
3.3.1	Logical Design	37
3.3.2	Flowchart	39
3.3.3	Interface Design	41
3.3.4	Pseudocode of Selected Algorithm	42
3.3.5	Prototype Implementation	42
3.4	Performance Evaluation	44
3.5	Gantt Chart	46
3.6	Conclusion	48
CHAPTER 4		49
4.1	System Architecture	49
4.2	Program Codes for Algorithm	50
4.2.1	Import Libraries	50
4.2.2	Setting the Working Directory	51
4.2.3	Load the Dataset	52
4.2.4	Data Preprocessing	52