

UNIVERSITI TEKNOLOGI MARA

**ONLINE COURSES
RECOMMENDATION SYSTEM
USING CONTENT-BASED
FILTERING**

**NIK MUHAMMAD IRFAN BIN SHAMSUL
ZUHAIRI**

BACHELOR OF COMPUTER SCIENCE (Hons.)

JANUARY 2025

ACKNOWLEDGEMENT

Above all, I want to say a big thank you to my supervisor, Miss Nik Marsyahariani Binti Nik Daud, who has been with me every step of the way for two semesters. I also owe a lot to Madam Ummu Fatimah Binti Mohd Bahrin for helping me understand how to evaluate the recommendation system, which was crucial for my research. I cannot forget to mention my family and friends, especially those who are also researching recommendation systems. They've been my rock, sharing knowledge and cheering me on through this whole journey of writing my thesis. I'm also grateful to everyone who had a hand, big or small, in helping me with my thesis. As I wrap up this thesis, I look back at the challenges and successes that have shaped my academic path. Thank you all for being a part of this important milestone in my learning journey.

ABSTRACT

A content-based filtering approach serves as the primary basis for developing this project's main recommendation functionality for online courses. Users currently face challenges in finding proper online courses from numerous choices because the new system design establishes an improvement. The content-based filtering system personalizes course ideas based on the attributes of the courses and the user's preferences, improving the learning experience by making relevant and efficient recommendations. In this project, the objectives are to study the implementation of content-based filtering, develop a prototype system, and evaluate its performance. This approach enables users to find courses matching their interests with ease, hence enhancing their learning experience. The growing number of Massive Open Online Courses (MOOCs) available online makes it difficult for students to choose courses that align with their interests and learning objectives. To solve this, an online course recommendation system that employs Content-Based Filtering (CBF) is presented. The system personalizes recommendations based on course attributes and user preferences, improving the learning experience. The methodology involves collecting data from Kaggle, preprocessing it, then using TF-IDF and Cosine Similarity to calculate course similarity scores. A prototype system was developed which allowed users to provide their course option and receive personalized recommendations. The system's performance was evaluated using Mean Absolute Error (MAE) and Root Mean Square Error. The results show that the system has an MAE of 0.25 and an RMSE of 0.35, showing its ability to provide recommendations. Future work will include implementing hybrid filtering techniques, real-time user feedback, and increasing the dataset with additional diverse course sources to increase recommendation performance.

TALE OF CONTENTS

CONTENT	PAGE
SUPERVISOR APPROVAL	i
STUDENT DECLARATION	ii
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
LIST OF FIGURES	viii
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	xi
 CHAPTER 1	 1
1.1 Background Study	1
1.2 Problem Statement	2
1.3 Objective	4
1.4 Project Scope	4
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 The Use of Recommendation System	8
2.2.2 Recommendation Systems Technique	8
2.2.2.1 Collaborative Filtering Technique	8
2.2.2.2 Content-Based Filtering Technique	9

2.2.2.3 Hybrid Filtering Technique	10
2.2.3 Advantages and Disadvantages of Recommendation System Technique	10
2.3 Content-Based Filtering Algorithm	12
2.3.1 Text Preprocessing	13
2.3.2 Term Frequency-Inverse Document Frequency (TF-IDF)	14
2.3.3 Cosine Similarity	15
2.3.4 Implementation of Content-Based Filtering in Various Problem	15
2.4 Online Courses Recommendation	21
2.4.1 Similar Works in Online Courses	23
2.5 Related Works	26
2.6 The Implication of Literature Review	30
2.7 Conclusion	31
CHAPTER 3	33
3.1 Overview of Research Methodology Framework	33
3.1.1 Detailed of Research Framework	34
3.2 Preliminary Phase	35
3.2.1 Literature Study	36
3.2.2 Data Pre-processing	36
3.2.2.1 Data Collection	37
3.2.2.2 Data Cleaning	37
3.3 Design Phase	38
3.3.1 System Architecture	38
3.3.2 Flowchart	39
3.3.3 User Interface Design	41
3.3.4 Pseudocode of Selected Algorithm	43
3.3.5 Prototype Implementation	44