

**UNIVERSITI TEKNOLOGI MARA**

**A CONTENT-BASED MOVIE RECOMMENDATION  
SYSTEM USING COSINE SIMILARITY MEASURE**

**MUHAMMAD ASYRAF BIN ABD GHANI**

**Thesis submitted in fulfilment of the requirements  
for Bachelor of Computer Science (Hons.) Faculty  
of Computer and Mathematical Science**

**FEBRUARY 2021**

## **ACKNOWLEDGEMENT**

Alhamdulillah, and thanks to Allah because of His Almighty and His utmost blessings, I was able to finish the research within the time duration given. Firstly, my special thanks go to my supervisor, Madam Norizan Binti Mohamad for providing invaluable guidance throughout this research. Special appreciation also goes to my beloved parents for supporting me to complete this project. Finally, I would like to give my gratitude to my dearest friends that always give support in completing this project and shared the knowledge about the related studies in this project.

## **ABSTRACT**

Movie recommendation system can provide the entertainment needed by our social life due to the high accuracy of the recommended movie. The system can make a suggestions to all the user based on their movie preference. Every people will have their own preference when they want to go watching movie in the cinema or online streaming platform such as Netflix, Amazon Prime or Youtube. The existense of movie recommendation system will make the movie selection become easier and accurate to the user. In this paper, the proposed movie recommendation system by using Cosine Similarity measure will has the ability to give the user their own recommended movie. It calculate the content in the movie and calculate its similarity by the words appear in the movie information. It generates the recommended movie for the movie lovers and make it convenient for them to make a selection on the movie. The accuracy of the recommended movie also have been discuss in details in this project.

# Table of Contents

SUPERVISOR APPROVAL .....	ii
STUDENT DECLARATION .....	iii
ACKNOWLEDGEMENT.....	iv
CHAPTER 1 .....	3
INTRODUCTION.....	3
1.0 Background of Study.....	3
1.1 Problem Statement.....	5
1.2 Objective.....	5
1.3 Project Scope .....	6
1.4 Project Significance.....	7
1.5 Overview of Research Framework.....	8
1.6 Summary .....	9
CHAPTER 2 .....	10
LITERATURE REVIEW .....	10
2.0 Introduction .....	10
2.1 Movie Recommendation System.....	10
2.2 Content-Based Filtering in the recommendation system.....	11
2.2.1 How does Content-Based Filtering work?.....	11
2.3 Cosine Similarity.....	16
2.3.1 How Cosine Similarity works? .....	16
2.6 Similar Works.....	17
2.7 The implication of Literature Review.....	21
2.8 Conclusion.....	22
CHAPTER 3 .....	23
INTRODUCTION.....	23
3.0 Detailed of Research Methodology Framework.....	23
3.1 Preliminary Study.....	25
3.2 Literature Review .....	25
3.3 Data Analysis.....	26
3.3.1` Data Collection.....	26
3.4 Design Phase.....	28
3.4.1 System Architecture .....	28

3.4.2	Flow chart .....	29
3.4.3	User Interface .....	31
3.4.5	Pseudocode .....	33
3.5	Conclusion .....	35
CHAPTER 4	.....	36
INTRODUCTION	.....	36
4.1	Conceptual Framework.....	36
4.2	Program Codes .....	38
4.2.2	Data Pre-processing .....	39
4.2.2.1	Data Cleaning .....	40
4.2.2.1.2	Removing duplicate movie information .....	40
4.2.2.1.3	Removing missing value of the movie .....	41
4.3	The Rake and Bag of Words model.....	43
4.3.2	Write function code to sort recommended recommended .....	46
4.4	Prototype Interface.....	48
4.5	Evaluation Results .....	56
4.6	Discussion .....	60
4.7	Conclusion.....	60
CHAPTER 5	.....	61
INTRODUCTION	.....	61
5.1	Summary of the project.....	61
5.2	Project Contribution .....	62
5.3	Project Limitations .....	63
5.4	Project Recommendations.....	64
5.5	Conclusion.....	65
REFERENCES	.....	66
APPENDICES	.....	69