



PROJECT REPORT
COMPARING THE COLLABORATIVE FILTERING
ALGORITHM WITH NAIVE BAYES ON THE FILM
RECOMMENDATION SYSTEM

LIEM, DANIEL ADITYA SANTOSO
18.K1.0036

Faculty of Computer Science
Soegijapranata Catholic University
2022

HALAMAN PENGESAHAN



Judul Tugas Akhir: : Comparing The Collaborative Filtering Algorithm with Naive Bayes on The Film Recommendation System

Diajukan oleh : Liem, Daniel Aditya S.

NIM : 18.K1.0036

Tanggal disetujui : 14 Januari 2022

Telah setuju oleh

Pembimbing : Hironimus Leong S.Kom., M.Kom.

Penguji 1 : Hironimus Leong S.Kom., M.Kom.

Penguji 2 : R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D

Penguji 3 : Yonathan Purbo Santosa S.Kom., M.Sc

Penguji 4 : Yulianto Tejo Putranto S.T., M.T.

Penguji 5 : Y.b. Dwi Setianto S.T., M.Cs.

Penguji 6 : Rosita Herawati S.T., M.I.T.

Ketua Program Studi : Rosita Herawati S.T., M.I.T.

Dekan : Dr. Bernardinus Harnadi S.T., M.T.

Halaman ini merupakan halaman yang sah dan dapat diverifikasi melalui alamat di bawah ini.

sintak.unika.ac.id/skripsi/verifikasi/?id=18.K1.0036

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : LIEM, DANIEL ADITYA SANTOSO

ID : 18.K1.0036

declare that this work, titled "COMPARING THE COLLABORATIVE FILTERING ALGORITHM WITH NAÏVE BAYES ON THE FILM RECOMMENDATION SYSTEM", and the work presented in it is my own. I confirm that:

- 1 This work was done wholly or mainly while in candidature for a research degree at Soegijapranata Catholic University
- 2 Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- 3 Where I have consulted the published work of others, this is always clearly attributed.
- 4 Where I have quoted from the work of others, the source is always given.
- 5 Except for such quotations, this work is entirely my own work.
- 6 I have acknowledged all main sources of help.
- 7 Where the work is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Semarang, January, 20, 2022



LIEM, DANIEL ADITYA SANTOSO

18.K1.0036

STATEMENT PAGE OF SCIENTIFIC PUBLICATIONS FOR ACADEMIC INTEREST

I, the undersigned:

Name : LIEM, DANIEL ADITYA SANTOSO

Study Program : Informatics Engineering

Faculty : Computer Science

Type of Work : Thesis

Approve to grant Soegijapranata Catholic University Semarang Non-exclusive RoyaltyFree Rights for the scientific works entitled “COMPARING THE COLLABORATIVE FILTERING ALGORITHM WITH NAIVE BAYES ON THE FILM RECOMMENDATION SYSTEM” along with existing tools (if needed). With these rights, database form, maintain, and publish this final project as long as it includes my name as the writer/creator and as the copyright owner.

This statement was made truthfully.

Semarang, January, 20, 2022



LIEM, DANIEL ADITYA SANTOSO

18.K1.036

ACKNOWLEDGMENT

First, I would like to say a big thank you to all the lecturers in the computer science faculty at Soegijapranata Unika Semarang who have accompanied the entire learning process so that I can be at this point. Especially to Mr. Hironimus Leong who has been a supervisor for the past few months who continues to motivate and provide.

I also thank my family who always support me in completing this thesis. And also my friends who always accompany me and always give me motivation.

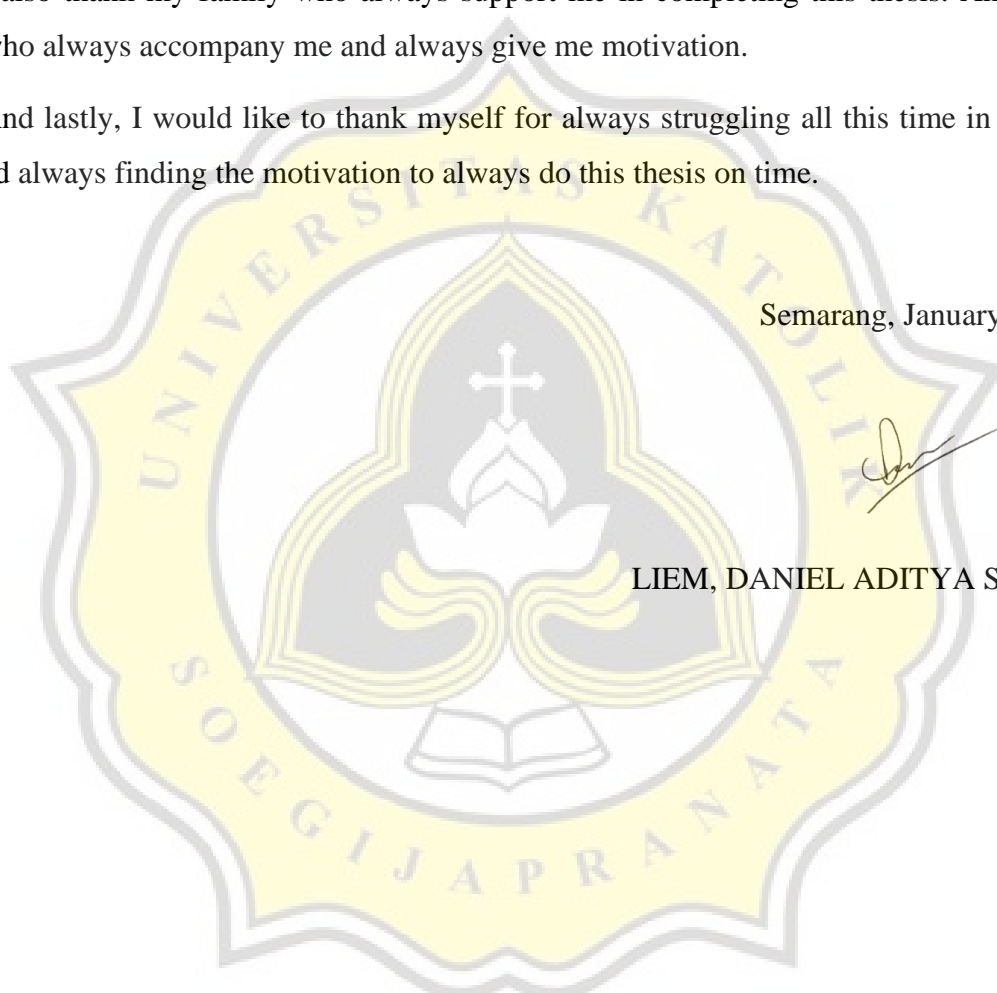
And lastly, I would like to thank myself for always struggling all this time in doing this thesis and always finding the motivation to always do this thesis on time.

Semarang, January, 20, 2022



LIEM, DANIEL ADITYA SANTOSO

18.K1.036



ABSTRACT

The many movies that are circulating and the many platforms that provide movie streaming platforms raise a question, namely what algorithm is the most suitable for use in providing movie recommendations. Of course, each of these streaming platforms uses different algorithms and factors.

In this study the author tries to compare two algorithms in providing movie recommendations based on the rating factor. The algorithm used is Collaborative Filtering with Cosine Similarity and also naive Bayes. Both authors tested using a dataset from movieLens.org as much as 10,000 data.

And in the results, Collaborative Filtering got better results through MSE and RMSE testing than naive Bayes. But the prediction score of each movie in each algorithm has a similar and the same score because it only uses the rating factor.

Keyword: naive bayes, collaborative filtering, cosine similarity

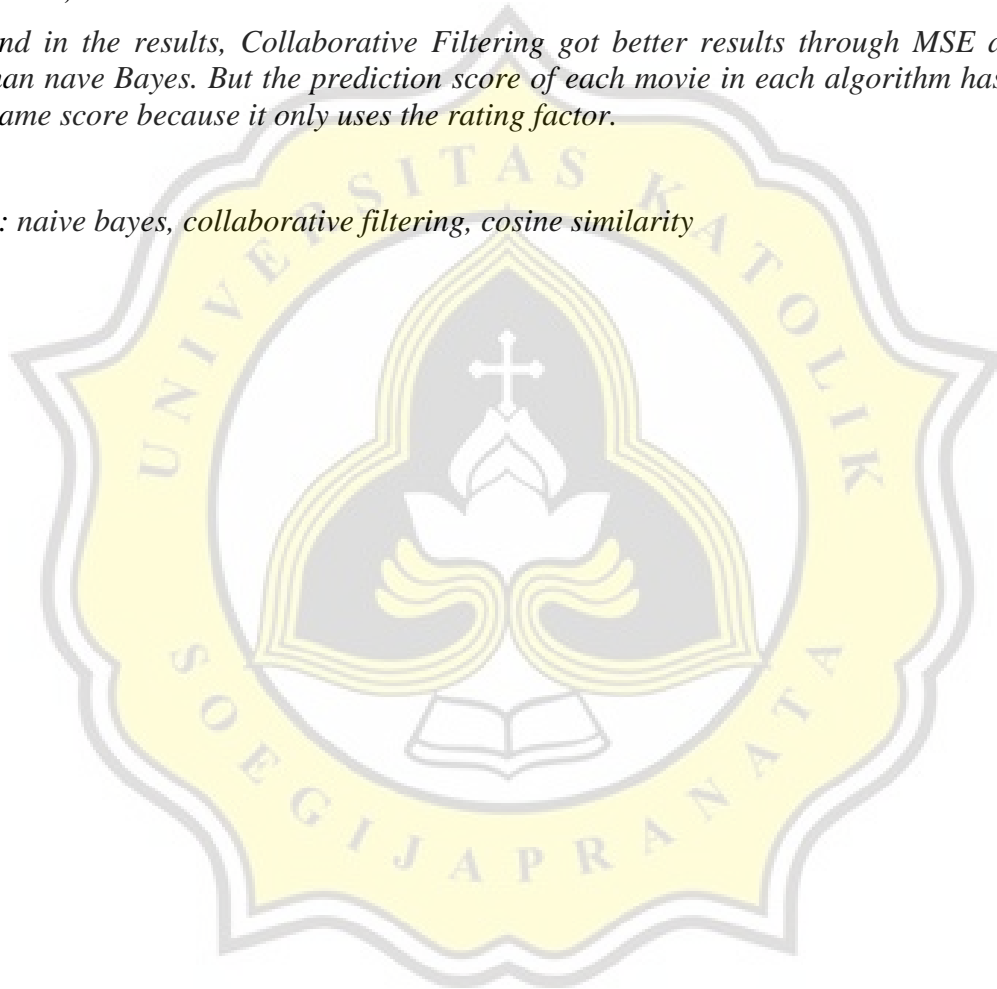


TABLE OF CONTENTS

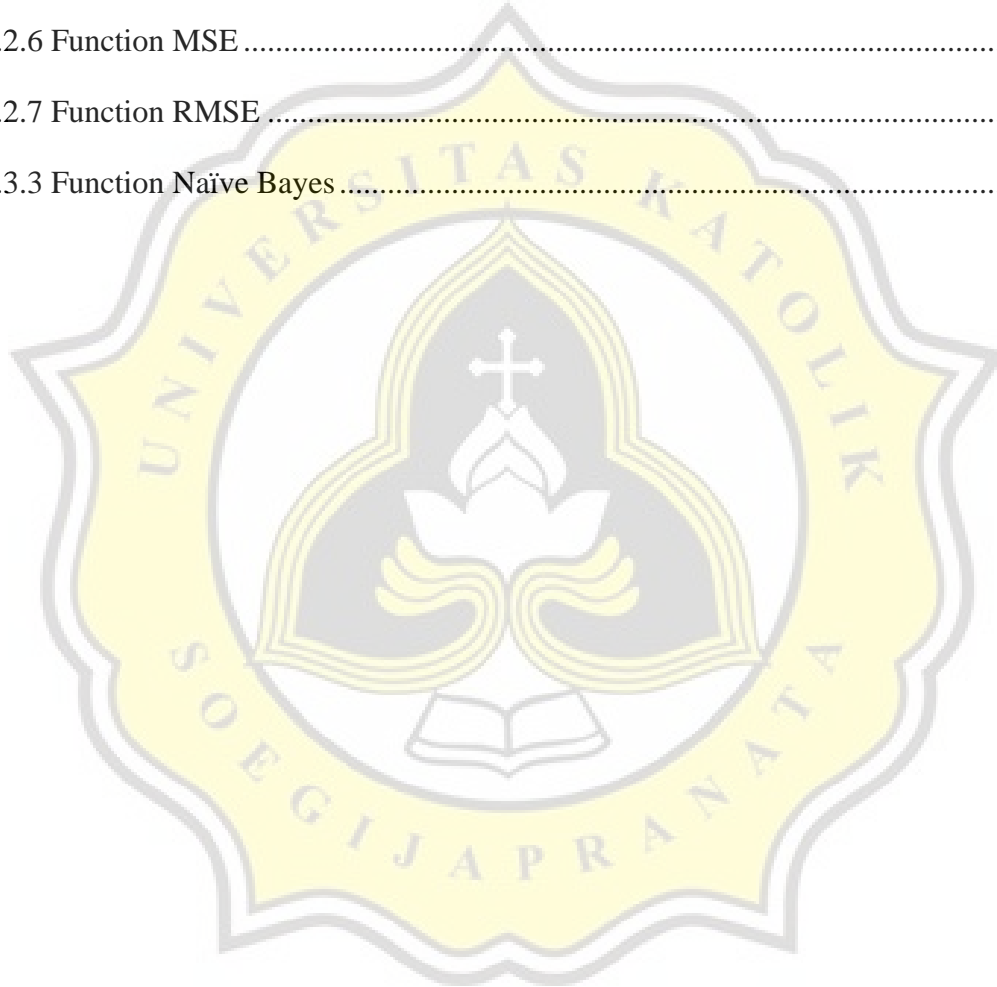
COVER	i
APPROVAL AND RATIFICATION PAGE	ii
DECLARATION OF AUTHORSHIP	iii
STATEMENT PAGE OF SCIENTIFIC PUBLICATIONS FOR ACADEMIC INTEREST iv	
ACKNOWLEDGMENT	v
ABSTRACT	vi
TABLE OF CONTENTS	vii
LIST OF FIGURE	1
LIST OF TABLE	2
CHAPTER 1 INTRODUCTION	3
1.1 Background.....	3
1.2 Problem Formulation	4
1.3 Scope.....	4
1.4 Objective	4
CHAPTER 2 LITERATURE STUDY	5
CHAPTER 3 RESEARCH METHODOLOGY	9
3.1 Data Collection	9
CHAPTER 4 ANALYSIS AND DESIGN	11
4.1 Pre Processing Dataset	11
4.2 Collaborative Filtering	15
4.3 Naïve Bayes	19
4.4 Perbandingan Collaborative Filtering dengan Naïve Bayes	22
CHAPTER 5 IMPLEMENTATION AND RESULTS	24
5.1 Collaborative Filtering	24

5.2	Naïve Bayes.....	28
5.3	Result.....	32
5.4	Analysis.....	33
CHAPTER 6 CONCLUSION.....		a
REFERENCES.....		b
APPENDIX.....		c



LIST OF FIGURE

4.1.3 Function vLookUp Microsoft Excel	12
4.1.5 Query untuk mendapatkan ranking movie data	14
4.2.1 Flowchart collaborative Filtering.....	16
4.2.3 Function Cosine Similarity	16
4.2.6 Function MSE	18
4.2.7 Function RMSE	18
4.3.3 Function Naïve Bayes.....	20



LIST OF TABLE

4.1.1 Table Rating Dataset.....	11
4.1.2 Table Movie Dataset	11
4.1.4 Table ratingMovie.....	13
4.1.6 Ranking Dataset	14
4.2.2 Table Preprocessing Data Collaborative Filtering.....	16
4.2.4 Table Cosine Similarity Score	17
4.2.5 Table Cosine Similarity Sorted.....	17
4.3.2 Movie Data.....	19
4.3.3 Table User Data	19

