

**SISTEM REKOMENDASI FILM DENGAN MENGGUNAKAN
ALGORITMA *FP-GROWTH* DAN *K-NEAREST NEIGHBOR***

SKRIPSI



Oleh:

MOHAMMAD AMIR FANANI

17081010071

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2024**

LEMBAR PENGESAHAN SKRIPSI

**Judul : SISTEM REKOMENDASI FILM DENGAN MENGGUNAKAN
ALGORITMA *FP-GROWTH* DAN *K-NEAREST NEIGHBOR***

Oleh : MOHAMMAD AMIR FANANI

NPM : 17081010071

Telah Diseminarkan Dalam Ujian Skripsi Pada :

Hari Jumat , Tanggal 5 Januari 2024

Mengetahui

Dosen Pembimbing

1.

Yisti Vita Via, S.ST, M.Kom

NIP. 19860425 2021212 001

2.

Eva Yulia Puspaningrum, S.Kom, M.Kom

NIP. 19890705 2021212 002

Dosen Penguji

1.

Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.

NPT. 222198 60 816400

2.

Firza Prima Aditiawan, S.Kom., MTL

NIP. 19860523 2021211 003

Menyetujui

**Dekan
Fakultas Ilmu Komputer**

Prof. Dr. Ir. Novirina Hendrasarie. MT.

NIP. 19681126 199403 2 001

**Koordinator Program Studi
Informatika**

Fetty Tri Anggraeny, S.Kom, M.Kom

NIP. 19820211 2021212 005

SURAT PERNYATAAN ANTI PLAGIAT

Saya mahasiswa Program studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Mohammad Amir Fanani

NPM : 17081010071

Dengan ini menyatakan bahwa judul skripsi atau tugas akhir yang saya ajukan dan kerjakan, yang berjudul:

“SISTEM REKOMENDASI FILM DENGAN MENGGUNAKAN ALGORITMA *FP-GROWTH* DAN *K-NEAREST NEIGHBOR*”

Bukan merupakan plagiat dari tugas akhir atau skripsi maupun penelitian orang lain dan juga bukan merupakan produk atau *software* yang saya beli dari pihak lain. Saya juga menyatakan bahwa skripsi ini merupakan pekerjaan Saya sendiri, kecuali yang dinyatakan dalam daftar pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional “Veteran” Jawa Timur maupun institusi Pendidikan lainnya.

Jika ternyata kemudian hari pernyataan ini terbukti tidak benar, maka Saya siap menerima segala konsekuensinya.

Surabaya, 5 Januari 2024

Hormat Saya,



Mohammad Amir Fanani

NPM. 17081010071

SISTEM REKOMENDASI FILM MENGGUNAKAN ALGORITMA *FP-GROWTH* DAN *K-NEAREST NEIGHBOR*

Nama Mahasiswa : Mohammad Amir Fanani
NPM : 17081010071
Program Studi : Informatika
Dosen Pembimbing : 1. Yisti Vita Via, S.ST, M.Kom
2. Eva Yulia Puspaningrum, S.Kom, M.Kom

Abstrak

Sistem rekomendasi merupakan sebuah sistem yang dapat merekomendasikan atau menyarankan pengguna dengan berbagai faktor yang dapat ditentukan sesuai kebutuhan. Sistem rekomendasi dapat dibuat dengan menggunakan salah satu cabang Kecerdasan Buatan yaitu Data Mining. Tanpa menggunakan Sistem Rekomendasi, maka tampilan atau rekomendasi dari sistem akan acak dan tidak sesuai dengan keinginan pengguna.

Terdapat berbagai macam metode atau algoritma untuk menerapkan sebuah sistem rekomendasi salah satunya dengan menggunakan *FP-Growth* dan *K-Nearest Neighbor*. Dari dua algoritma ini dapat digunakan untuk memprediksi rekomendasi dari sebuah sistem dengan contoh sistem rekomendasi film. Data yang digunakan adalah penilaian dari pengguna lain dalam data film tersebut. Data film yang akan digunakan memiliki lebih dari 100.004 data dengan penilaian skala 1 sampai 5. Data tersebut akan diproses dengan metode *cascade*.

Dari pengujian menggunakan nilai $k = 5$ pada data tes sebanyak 30%, dan 20% menghasilkan nilai 1.0 atau 100% pada *Precision*. Begitu pula pengujian menggunakan nilai $k = 8$ pada data tes sebanyak 30%, dan 20% menghasilkan nilai 1.0 atau 100% pada *Precision*. Pengujian terakhir menggunakan nilai $k = 10$ pada data tes sebanyak 30%, dan 20% menghasilkan nilai 1.0 atau 100% pada *Precision*. Hal ini menunjukkan bahwa pengujian *K-Nearest Neighbor* sudah memenuhi kriteria untuk dijadikan sistem rekomendasi.

Kata Kunci: Sistem rekomendasi, Cascade, *FP-Growth*, *K-Nearest Neighbor*

KATA PENGANTAR

Segala puji hanya untuk Allah Subhanahu Wa Ta'ala, Rabb semesta alam yang telah yang telah memberikan rahmat dan hidayah-Nya karena atas izin dan keridhoan-Nya penulis mampu untuk menyelesaikan penelitian skripsi dan laporan skripsi yang berjudul “Sistem Rekomendasi Menggunakan Metode *FP-Growth* Dan *K-Nearest Neighbor*”. Shalawat serta salam senantiasa tercurahkan kepada nabi Muhammad Shallallahu ‘alaihi wa sallam, keluarga, para sahabat, serta umat beliau yang senantiasa mengikuti petunjuk beliau.

Banyak dukungan serta bantuan yang didapatkan selama melakukan penelitian hingga menyelesaikan laporan ini. Dengan rasa hormat serta banyak terima kasih penulis ucapkan kepada seluruh pihak yang telah membantu proses penyelesaian penelitian dari awal hingga akhir.

Penulis menyadari masih terdapat kekurangan karena keterbatasan pengetahuan dan pemikiran. Oleh karena itu, penulis mengharapkan dan menerima kritik dan saran yang membangun dari semua pihak dalam penyempurnaan laporan ini. Semoga skripsi ini menjadi berkah, bermanfaat, dan berguna serta mampu memenuhi harapan berbagai pihak.

Surabaya, 5 Januari 2024

Penulis,

Mohammad Amir Fanani

UCAPAN TERIMA KASIH

Penyusunan skripsi ini melalui tahapan dan pengorbanan baik harta, jiwa, serta waktu. Banyak hambatan yang dihadapi, namun berkat kehendak Allah subhanahu wa ta'ala sehingga penulis berhasil menyelesaikan penyusunan skripsi ini. Penulis menyadari bahwa penyelesaian skripsi ini tidak terlepas dari dukungan oleh semua pihak yang terlibat di dalamnya. Oleh karena itu, pada kesempatan ini dengan segala kerendahan hati dan penuh rasa hormat penulis mengucapkan terima kasih yang tidak terhingga bagi semua pihak yang telah memberi bantuan baik secara moril maupun materiil, baik secara langsung maupun tidak langsung dalam penyusunan skripsi ini hingga selesai, khususnya kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom. selaku Koordinator Program Studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Yisti Vita Via, S.ST, M.Kom., selaku Dosen Pembimbing I yang telah bersedia membimbing dan mengarahkan penulis selama menyusun skripsi serta memberikan banyak ilmu serta solusi pada setiap permasalahan atas kesulitan dalam penulisan skripsi ini.
5. Ibu Eva Yulia P., S.Kom, M.Kom, selaku Dosen Pembimbing II yang telah sabar, serta ikhlas meluangkan waktu dalam memberikan bimbingan, motivasi, dan arahan yang berharga kepada penulis selama penyusunan skripsi ini.
6. Dosen serta staf pada Program Studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan ilmu, arahan, serta pengalaman selama perkuliahan.

7. Orang Tua, Saudara, Istri, serta Anak saya yang telah memberikan doa, kasih sayang, dan berbagai dukungan selama proses penelitian dan penyusunan laporan.
8. Teman-teman penulis yang tergabung dalam Kelompok Pencari Makna Hidup (PMH) yang telah memberikan inspirasi serta saran dalam penyelesaian skripsi ini.
9. Teman-teman Program Studi Informatika angkatan 2017 yang telah memotivasi penulis untuk menyelesaikan skripsi ini.
10. Teman dekat penulis yang pada pembuatan skripsi ini masih menjalani proses studi di Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan semangat.
11. Semua pihak yang penulis tidak dapat sebutkan satu persatu.

Semoga Allah Subhanahu Wa Ta’ala memberikan balasan yang berlipat ganda atas kebaikan yang telah diberikan.

Surabaya, 5 Januari 2024

Penulis

Mohammad Amir Fanani

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	i
SURAT PERNYATAAN ANTI PLAGIAT	ii
ABSTRAK	iii
KATA PENGANTAR.....	iv
UCAPAN TERIMA KASIH.....	v
DAFTAR ISI	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL.....	xi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Perumusan Masalah	4
1.3. Tujuan Penelitian	5
1.4. Manfaat Penelitian	5
1.5. Batasan Masalah	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Penelitian Sebelumnya	6
2.2. Sistem Rekomendasi	8
2.3. <i>Hybrid Recommendation System</i>	9
2.4. <i>Cascade</i>	9
2.5. Film.....	10
2.6. <i>Cleaning Data</i>	10
2.7. <i>Frequent Pattern Tree</i>	12
2.8. <i>Frequent Pattern Growth</i>	13
2.9. <i>Association Rule</i>	15
2.10. <i>K-Nearest Neighbor</i>	16
2.11. <i>Scikit-Learn</i>	19
BAB III METODOLOGI PENELITIAN.....	21
3.1 Tahapan Penelitian	21
3.2 <i>Dataset</i>	22
3.3 <i>Cleaning Data</i>	23
3.4 <i>FP-Growth</i>	28

3.5	<i>Association Rule</i>	32
3.6	KNN.....	34
3.6	<i>Training Data</i>	37
3.7	Lingkungan pengembangan.....	43
3.8	Skenario uji coba.....	44
BAB IV HASIL DAN PEMBAHASAN		45
4.1.	Implementasi Program.....	45
4.1.1.	<i>Import Library</i>	45
4.1.2.	<i>Cleaning Data</i>	46
4.1.3.	Arsitektur <i>FP-Growth</i>	52
4.1.4.	Arsitektur <i>Association Rule</i>	53
4.1.5.	Arsitektur KNN.....	55
4.2.	Pengujian skenario	58
4.2.1.	Pengujian menggunakan data tes 30% dengan k=5	59
4.2.2.	Pengujian menggunakan data tes 20% dengan k=5	60
4.2.3.	Pengujian menggunakan data tes 30% dengan k=8	60
4.2.4.	Pengujian menggunakan data tes 20% dengan k=8	61
4.2.5.	Pengujian menggunakan data tes 30% dengan k=10	61
4.2.6.	Pengujian menggunakan data tes 20% dengan k=10	62
4.3.	Penerapan metode Cascade.....	62
BAB V KESIMPULAN DAN SARAN.....		76
5.1	Kesimpulan.....	76
5.2	Saran	77
DAFTAR PUSTAKA.....		78

DAFTAR GAMBAR

Gambar 2.1. Skema Metode <i>Cascade</i>	9
Gambar 2.2. Skema Pembentukan <i>FP-tree</i>	13
Gambar 3.1. Tahapan Penelitian	21
Gambar 3.2. Tampilan Visual dari sampel <i>dataset</i>	22
Gambar 3.3 Proses <i>Cleaning Data</i>	23
Gambar 3.4. Metode <i>FP-tree</i>	29
Gambar 3.5. <i>Flowchart FP-Growth</i>	30
Gambar 3.6. Pemrosesan dari <i>Association Rules</i>	33
Gambar 3.7. Pemrosesan Metode KNN.....	35
Gambar 3.8. Konsep perhitungan KNN.....	35
Gambar 3.9. <i>Training Data</i>	38
Gambar 4.1. <i>Dataset</i> dari <i>movies_metadata.csv</i>	47
Gambar 4.2. <i>Dataset</i> dari <i>Rating_small.csv</i>	48
Gambar 4.3. Hasil <i>Cleaning Data</i> dari <i>movies_metadata.csv</i>	49
Gambar 4.4. Hasil <i>Cleaning</i> dan diubah ke bentuk <i>pivot</i>	50
Gambar 4.5. Hasil perubahan kedalam bentuk <i>boolean</i>	51
Gambar 4.6. Hasil Pemrosesan metode <i>FP-Growth</i>	52
Gambar 4.7. Hasil Pemrosesan <i>Assosiation Rule</i>	54
Gambar 4.8. Hasil Pengujian Metode KNN dengan data tes 30%.....	59
Gambar 4.9. Hasil Pengujian Metode KNN dengan data tes 20%.....	60
Gambar 4.10. Hasil Pengujian Metode KNN dengan data tes 30%.....	60
Gambar 4.11. Hasil Pengujian Metode KNN dengan data tes 20%.....	61
Gambar 4.12. Hasil Pengujian Metode KNN dengan data tes 30%.....	61
Gambar 4.13. Hasil Pengujian Metode KNN dengan data tes 20%.....	62
Gambar 4.14. Hasil dari metode <i>FP-Growth</i>	64
Gambar 4.15. Hasil dari proses <i>FP-Growth</i>	65
Gambar 4.16. Hasil dari metode KNN.....	67
Gambar 4.17. Ringkasan hasil dari proses KNN	67
Gambar 4.18. Hasil Rekomendasi dari <i>User 30</i>	68
Gambar 4.19. Data asli untuk <i>movieId 858</i>	69
Gambar 4.20. Data asli dari <i>movieId 296</i>	69

Gambar 4.21 Data asli dari <i>movieId</i> 527.....	70
Gambar 4.22 Data asli dari <i>movieId</i> 318.....	70
Gambar 4.23 Data asli dari <i>movieId</i> 1213.....	71
Gambar 4.24 Data asli dari <i>movieId</i> 608.....	71
Gambar 4.25 Data asli dari <i>movieId</i> 2028.....	72
Gambar 4.26 Data asli dari <i>movieId</i> 2762.....	73
Gambar 4.27 Data asli dari <i>movieId</i> 595.....	73
Gambar 4.28 Data asli dari <i>movieId</i> 110.....	74

DAFTAR TABEL

Tabel 3.1. Isi <i>dataset Rating_small.csv</i>	22
Tabel 3.2. Contoh untuk penghapusan data.....	24
Tabel 3.3. Contoh penghapusan data	24
Tabel 3.4. Contoh sebelum pergantian data	25
Tabel 3.5 Contoh sesudah pergantian data.....	25
Tabel 3.6 Contoh tabel untuk pergantian data	26
Tabel 3.7. Hasil setelah data sudah terisi	26
Tabel 3.8. Contoh dari Pivot data.....	27
Tabel 3.9. Hasil dari proses perubahan Pivot data.....	28
Tabel 3.10. Contoh data untuk <i>FP-Growth</i>	30
Tabel 3.11 Tabel Pivot hasil dari contoh sebelumnya.....	31
Tabel 3.12 Tabel Pivot setelah disortir	31
Tabel 3.13 Perhitungan nilai <i>support</i> pada contoh tabel	31
Tabel 3.14. Hasil perhitungan metode <i>Confidence</i>	34
Tabel 3.15. Contoh data untuk metode KNN	37
Tabel 3.16. Hasil sortir dari Metode <i>Cosine similarity</i>	39
Tabel 3.17. Hasil sortir 3 tetangga terdekat.....	39
Tabel 3.18. Hasil sortir Metode MSD.....	40
Tabel 3.19. Hasil Data dari 3 Terdekat	40
Tabel 3.20. Hasil sortir dari perhitungan metode <i>Pearson</i>	41
Tabel 3.21. Hasil dari 3 terdekat dalam metode <i>Pearson</i>	41
Tabel 3.22. Hasil dari metode <i>Pearson</i> sebelumnya.....	42
Tabel 3.23. Hasil dari sortir nilai terdekat.....	42
Tabel 3.24. Hasil 3 nilai teratas	43
Tabel 4.25. Perbandingan Data Asli dan Hasil Rekomendasi	75