

PREDIKSI *RATING* DRAMA KOREA DENGAN *NAÏVE BAYES*

SKRIPSI



disusun oleh

Nurul Hidayati

17.12.0404

**PROGRAM SARJANA
PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2021

PREDIKSI *RATING* DRAMA KOREA DENGAN *NAÏVE BAYES*

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Sistem Informasi



disusun oleh

Nurul Hidayati

17.12.0404

**PROGRAM SARJANA
PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2021

PERSETUJUAN

SKRIPSI

PREDIKSI *RATING* DRAMA KOREA DENGAN *NAÏVE BAYES*

yang dipersiapkan dan disusun oleh

Nurul Hidayati

17.12.0404

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 6 Juli 2021

Dosen Pembimbing,

Norhikmah, M.Kom.

NIK. 190302245

PENGESAHAN

SKRIPSI

PREDIKSI RATING DRAMA KOREA DENGAN NAÏVE BAYES

yang dipersiapkan dan disusun oleh

Nurul Hidayati

17.12.0404

telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Juli 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Asro Nasiri, Drs. M.Kom
NIK. 190302152

Hartatik, S.T., M.Cs.
NIK. 190302232

Norhikmah, M.Kom.
NIK. 190302245

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 22 Juli 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, S.Kom., M.Kom
NIK. 190302096

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 06 Agustus 2021



Nurul Hidayati

NIM. 17.12.0404

MOTTO

"Cara terbaik untuk memprediksi masa depan adalah dengan menciptakannya"

-Alan Kay-

"Orang pesimis melihat kesulitan di setiap kesempatan. Orang optimis melihat kesempatan di setiap kesulitan"

-Winston Churchill-

"Cara memulai itu adalah berhenti berbicara dan mulailah beraksi"

-Walt Disney-

PERSEMBAHAN

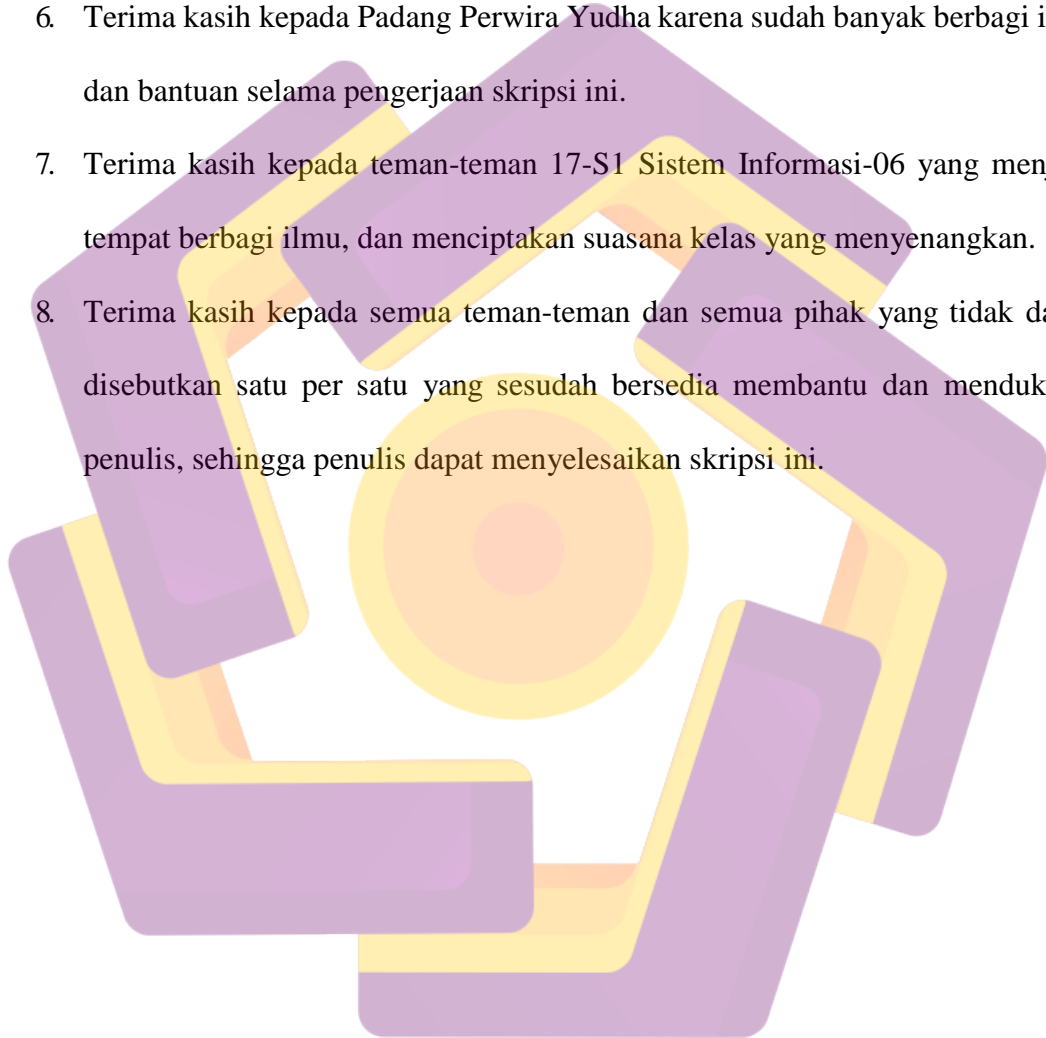
Segala puji dan syukur penulis panjatkan kepada Allah SWT yang telah memberikan rahmat, taufik, hidayah, serta inayah-Nya kepada penulis sehingga dapat menyelesaikan skripsi ini sesuai dengan target dan mendapatkan hasil yang terbaik.

Dalam kesempatan kali ini, penulis ingin mengucapkan banyak terimakasih kepada semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini, yaitu:

1. Terima kasih kepada orangtua saya, yang telah banyak mendoakan dan memberikan saya dukungan dan motivasi untuk selalu belajar selama perkuliahan dan pengerjaan skripsi ini, dan juga memberikan banyak bantuan secara moril maupun materi.
2. Terima kasih kepada Ibu Norhikmah, M.Kom, selaku dosen pembimbing. Terimakasih atas bimbingan dan dukungannya sehingga saya dapat menyelesaikan skripsi dengan baik.
3. Terima kasih kepada Deva Octaviano karena sudah memberi semangat hingga pada tahap ini, canda tawa, serta bersedia menemani saya selama mengerjakan skripsi ini.
4. Terima kasih saya ucapkan kepada teman seperjuangan “Lucknut Squad” yaitu Kekek, Amandel, Giden, Hillal, Aziz, Cahyo, Faisol, Jelembung, Topek, Sigid, dan Gagas. Terimakasih atas canda tawa, semangat, motivasi, solidaritas, dan perjuangan yang sangat panjang yang telah dilewati bersama. Terimakasih atas

bantuan kalian semua.

5. Terima kasih kepada Siti Roisadul Nisok karena sudah memberi semangat, dukungan serta do'a, dan sudah bersedia mendengar keluh kesah saya selama perkuliahan dan pembuatan skripsi ini.
6. Terima kasih kepada Padang Perwira Yudha karena sudah banyak berbagi ilmu dan bantuan selama pengerjaan skripsi ini.
7. Terima kasih kepada teman-teman 17-S1 Sistem Informasi-06 yang menjadi tempat berbagi ilmu, dan menciptakan suasana kelas yang menyenangkan.
8. Terima kasih kepada semua teman-teman dan semua pihak yang tidak dapat disebutkan satu per satu yang sesudah bersedia membantu dan mendukung penulis, sehingga penulis dapat menyelesaikan skripsi ini.



KATA PENGANTAR

Assalamu'alaikum Wr. Wb

Puji syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, hidayah, serta inayah-Nya, sehingga penulis dapat menyelesaikan skripsi ini sesuai dengan waktu yang diinginkan penulis. Tidak lupa sholawat serta salam penulis haturkan kepada Nabi Agung Muhammad SAW.

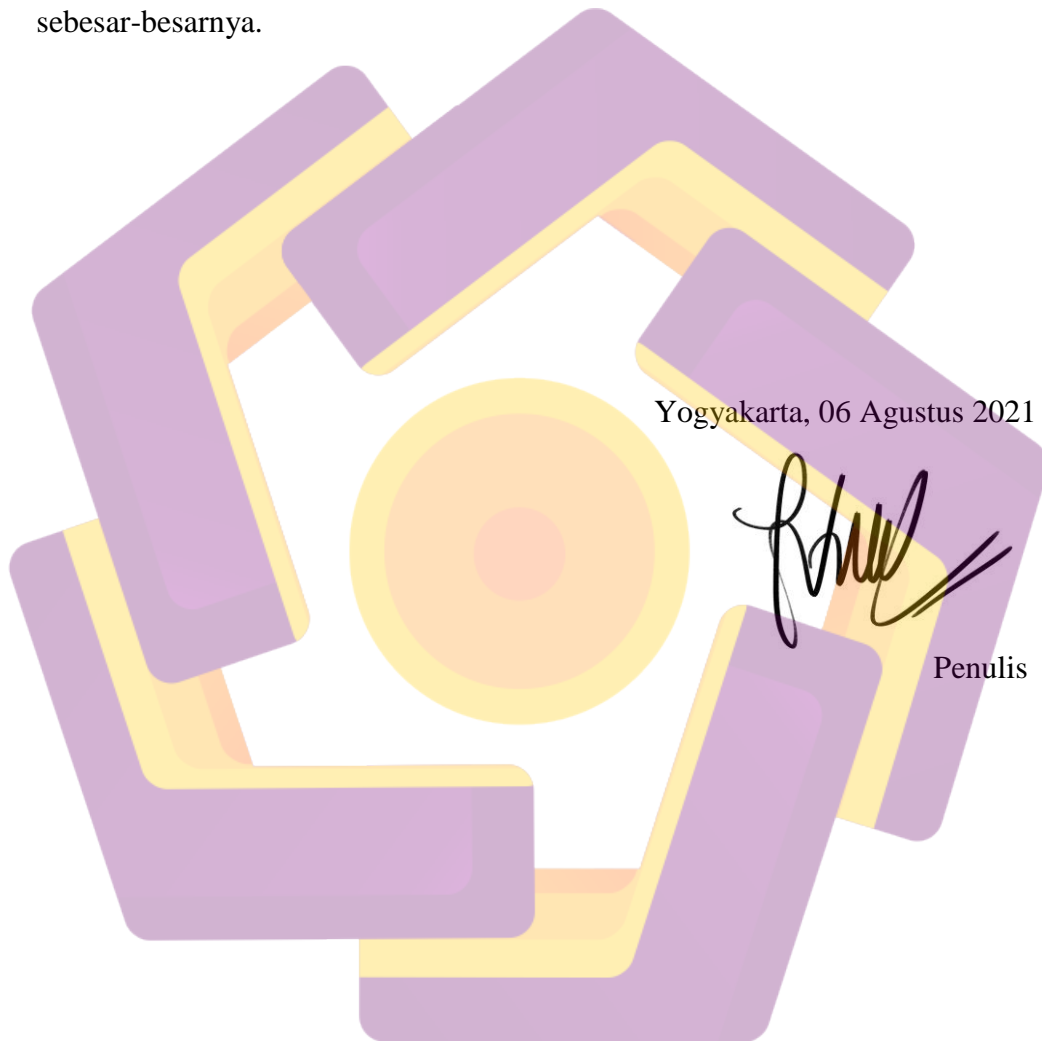
Skripsi ini disusun dalam rangka memenuhi salah satu syarat kelulusan jenjang Program Sarjana pada Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.

Proses penyusunan hingga selesainya laporan skripsi ini tidak lepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak baik secara langsung maupun tidak langsung. Maka dari itu, sebagai rasa hormat penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M, selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Norhikmah, M.Kom, M.Kom, selaku dosen pembimbing penulis yang telah memberikan petunjuk, bimbingan dan nasihatnya dalam proses penulisan skripsi ini.
4. Kedua orang tua dan serta saudara dan saudari yang telah memberikan motivasi, do'a, masukan, dan semangat.
5. Teman-teman seperjuangan 17-S1 Sistem Informasi-06, yang telah banyak

berdiskusi dengan penulis dalam masa pendidikan.

Akhir kata, semoga skripsi ini dapat bermanfaat bagi dunia pendidikan. Kritik dan saran sangat membantu perkembangan dan penyempurnaan karya tulis ini. Sekian dari penulis, apabila terdapat kesalahan dan kekurangan mohon maaf sebesar-besarnya.



Yogyakarta, 06 Agustus 2021

A handwritten signature in black ink, appearing to be 'Rahm' or similar, written over the decorative graphic.

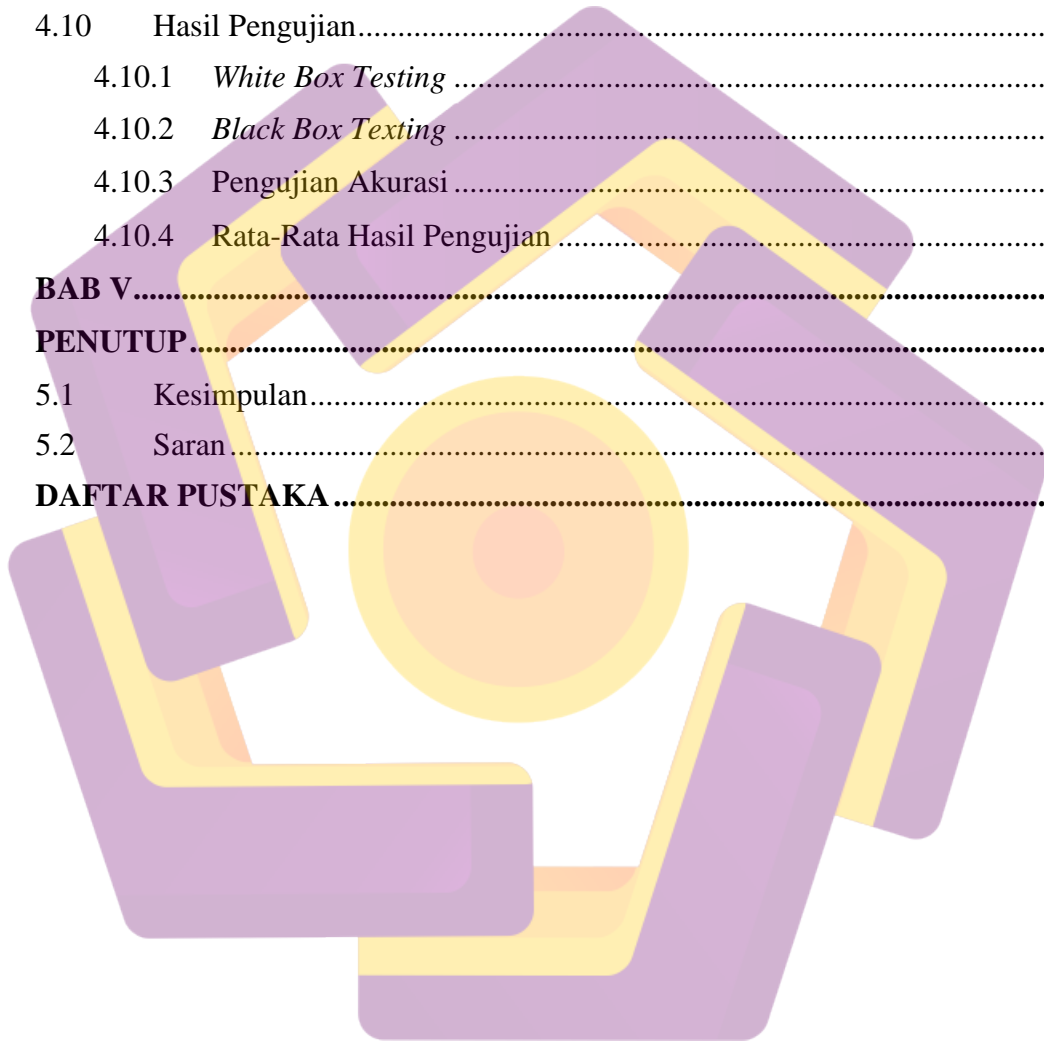
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN	iv
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
INTISARI	xvi
ABSTRACT	xvii
PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Maksud dan Tujuan Penelitian	5
1.5 Manfaat Penelitian.....	5
1.6 Metode Penelitian.....	5
1.6.1 Identifikasi Perumusan Masalah	6
1.6.2 Pengumpulan Data	6
1.6.3 Analisis Kebutuhan	7
1.6.4 Metode Perancangan	7
1.6.5 Metode Pengembangan Sistem	7
1.6.6 Metode Testing.....	7
1.6.7 Penarikan Kesimpulan	8
1.7 Sistematika Penulisan.....	9
BAB II	10
LANDASAN TEORI	10
2.1 Kajian Pustaka	10

2.2	Landasan Teori	13
2.2.1	Drama Korea	13
2.2.2	<i>Data Mining</i>	14
2.2.3	<i>Dataset</i>	19
2.2.4	<i>Naïve Bayes</i>	20
2.2.5	Metode Pengembangan Sistem	22
2.2.6	<i>Testing</i>	23
BAB III	28
METODE PENELITIAN	28
3.1	Tahapan Penelitian	28
3.2	Pengumpulan Data	29
3.2.1	Melakukan Studi Pustaka	29
3.2.2	Dokumentasi Data	29
3.3	Analisa dan Perancangan Model Klasifikasi	30
3.3.1	<i>Preprocessing Data</i>	30
3.3.2	Modelling <i>Naïve Bayes</i>	34
3.4	Analisis Kebutuhan	36
3.4.1	Analisis Kebutuhan Fungsional	36
3.4.2	Analisis Kebutuhan Non-Fungsional	37
3.5	Unified Modeling Language (UML)	38
3.5.1	<i>Use Case Diagram</i>	38
3.5.2	<i>Activity Diagram</i>	39
3.5.3	<i>Sequence Diagram</i>	45
3.5.4	<i>Class Diagram</i>	51
3.6	Perancangan Sistem	52
3.6.1	<i>Flowchart Naïve Bayes Classifier</i>	52
3.6.2	<i>Flowchart Sistem</i>	53
3.6.3	Perancangan Tampilan	54
BAB IV	60
HASIL DAN PEMBAHASAN	60
4.1	Implementasi dan Pembahasan Halaman <i>Login</i>	60
4.2	Implementasi dan Pembahasan Halaman <i>Dashboard</i>	61
4.3	Implementasi dan Pembahasan Halaman <i>Upload Data</i>	61

4.4	Implementasi dan Pembahasan Halaman Master Data	62
4.5	Implementasi dan Pembahasan Halaman Cleaning Data	62
4.6	Implementasi dan Pembahasan Halaman Transformasi Data	63
4.7	Implementasi dan Pembahasan Persebaran Data	64
4.8	Implementasi dan Pembahasan Pengujian Akurasi	64
4.9	Halaman Uji Tunggal	65
4.10	Hasil Pengujian.....	65
4.10.1	<i>White Box Testing</i>	65
4.10.2	<i>Black Box Texting</i>	66
4.10.3	Pengujian Akurasi	68
4.10.4	Rata-Rata Hasil Pengujian	74
BAB V	76
PENUTUP	76
5.1	Kesimpulan.....	76
5.2	Saran	76
DAFTAR PUSTAKA	77



DAFTAR TABEL

Tabel 2.1 Tabel Penelitian Terdahulu	12
Tabel 2.2 Confusion Matrix Klasifikasi Biner.....	24
Tabel 3.1 Hasil Scraping Data	29
Tabel 3.2 Representasi Data.....	31
Tabel 3.3 Cleaning	31
Tabel 3.4 Transformasi Status Klasifikasi	33
Tabel 3.5 Contoh Kasus Data Training.....	34
Tabel 4.1 Pengujian Black Box.....	67
Tabel 4.2 Fold Satu	69
Tabel 4.3 Fold Dua.....	69
Tabel 4.4 Rata-rata hasil pengujian.....	74

DAFTAR GAMBAR

Gambar 2.1 Tahapan Data Mining.....	17
Gambar 2.2 Contoh iterasi data dengan k-fold cross validation	26
Gambar 3.1 Tahapan Penelitian	28
Gambar 3.2 Use Case Diagram Aplikasi	38
Gambar 3.3 Activity Diagram Login	39
Gambar 3.4 Activity Diagram Upload Data	40
Gambar 3.5 Activity Diagram Master Data	41
Gambar 3.6 Activity Diagram Cleaning Data.....	41
Gambar 3.7 Activity Diagram Transformasi Data.....	42
Gambar 3.8 Activity Diagram Persebaran Data.....	42
Gambar 3.9 Activity Diagram Pengujian Akurasi	43
Gambar 3.10 Activity Diagram Pengujian Tunggal	44
Gambar 3.11 Sequence Diagram Login	45
Gambar 3.12 Sequence Diagram Unggah Data	46
Gambar 3.13 Sequence Diagram Master Data	47
Gambar 3.14 Sequence Diagram Cleaning Data.....	47
Gambar 3.15 Sequence Diagram Transformasi Data.....	48
Gambar 3.16 Sequence Diagram Pesebaran Data.....	49
Gambar 3.17 Sequence Diagram Pengujian Akurasi	49
Gambar 3.18 <i>Sequence Diagram</i> Pengujian Tunggal	50
Gambar 3.19 Class Diagram Model.....	51
Gambar 3.20 <i>Flowchart Naive Bayes Classifier</i>	52
Gambar 3.21 Flowchart Sistem.....	53
Gambar 3.22 Rancangan Tampilan Form Login.....	54
Gambar 3.23 Rancangan Tampilan Halaman Dashboard	55
Gambar 3.24 Rancangan Tampilan Halaman Upload Data	55
Gambar 3.25 Rancangan Tampilan Halaman Master Data.....	56
Gambar 3.26 Rancangan Tampilan Halaman Cleaning Data	57

Gambar 3.27 Rancangan Tampilan Halaman Transformasi Data	57
Gambar 3.28 Rancangan Tampilan Halaman Persebaran Data	58
Gambar 3.29 Rancangan Tampilan Halaman Pengujian Akurasi.....	58
Gambar 3.30 Rancangan Tampilan Halaman Pengujian Tunggal	59
Gambar 4.1 Halaman Login	60
Gambar 4.2 Halaman Dashboard	61
Gambar 4.3 Halaman Upload Data	61
Gambar 4.4 Halaman Master Data.....	62
Gambar 4.5 Halaman Cleaning Data	62
Gambar 4.6 Halaman Transformasi Data.....	63
Gambar 4.7 Halaman Persebaran Data	64
Gambar 4.8 Halaman Pengujian Akurasi.....	64
Gambar 4.9 Halaman Pengujian Tunggal	65
Gambar 4.10 White Box Testing	66
Gambar 4.11 Pengujian $k=2$	69
Gambar 4.12 Pengujian $k=3$	70
Gambar 4.13 Pengujian $k=4$	71
Gambar 4.14 Pengujian $k=5$	71
Gambar 4.15 Pengujian $k=6$	72
Gambar 4.16 Pengujian $k=7$	72
Gambar 4.17 Pengujian $k=8$	73
Gambar 4.18 Pengujian $k=9$	73
Gambar 4.19 Pengujian $k=10$	74

INTISARI

Pada saat ini drama Korea menjadi salah satu media penyebaran *Korean wave* yang memiliki jumlah penggemar atau penonton yang tidak sedikit. Salah satu faktor yang menentukan dalam pemilihan drama yaitu *rating*. *Rating* adalah nilai yg mendeskripsikan popularitas suatu acara pada media pada saat tertentu.

Penelitian terkait prediksi *rating* drama Korea menggunakan perhitungan Algoritma *Iterative Dichotomiser 3* (ID3). Dalam penelitian ini perhitungan dilakukan dengan menggunakan Algoritma *Naïve Bayes* yang merupakan metode berbasis peluang. Beberapa variabel yang mempengaruhi *rating* pada penelitian ini antara lain stasiun, *genre*, *writer*, *director*, dan *actor*. Penelitian ini bertujuan untuk membangun sistem yang dapat memprediksi *rating* yang dinyatakan dengan status *rating* (tinggi atau rendah) dengan menggunakan Algoritma *Naïve Bayes*.

Hasil pengujian yang dilakukan terhadap 633 data menggunakan *confusion matrix* dan *K-fold cross validation* dengan nilai k diatur dari 2 sampai 10 mendapatkan nilai akurasi tertinggi adalah pada k=10 dengan nilai akurasinya 65.26%.

Kata kunci : drama Korea, *rating*, prediksi, *Naïve Bayes*.



ABSTRACT

At this time Korean drama is one of the media for spreading the Korean wave which has a large number of fans or viewers. One of the determining factors in choosing a drama is rating. Rating is a value that describes an event in the media at a certain time.

Research related to Korean drama rating predictions uses the Iterative Dichotomiser 3 (ID3) algorithm. In this study, the calculations were performed using the Naïve Bayes Algorithm which is an opportunity-based method. Several variables that affect the rating in this study include stations, genres, writers, directors, and actors. This study aims to build a system that can predict the rating expressed by the rating status (high or low) using the Naïve Bayes Algorithm.

The results of tests carried out on 633 data using a confusion matrix and K-fold cross validation with the value of k set from 2 to 10 get the highest accuracy value at $k = 10$ with an accuracy value of 65.26%.

Keywords : *Korean drama, rating, prediction, Naïve Bayes.*

