

**ANALISIS SENTIMEN ULASAN PENGGUNA GAME HAY DAY DI
GOOGLE PLAY STORE MENGGUNAKAN MODEL
KLASIFIKASI**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

HANNY OLIVIA

21.12.2005

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN ULASAN PENGGUNA GAME HAY DAY DI
GOOGLE PLAY STORE MENGGUNAKAN MODEL
KLASIFIKASI**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

HANNY OLIVIA

21.12.2005

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN ULASAN PENGGUNA GAME HAY DAY DI
GOOGLE PLAY STORE MENGGUNAKAN MODEL
KLASIFIKASI**

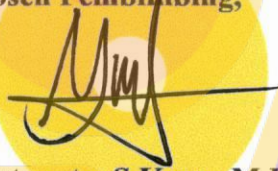
yang disusun dan diajukan oleh

Hanny Olivia

21.12.2005

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Februari 2025

Dosen Pembimbing,



Yoga Pristyanto, S.Kom., M.Eng.

NIK. 190302412

HALAMAN PENGESAHAN
SKRIPSI
ANALISIS SENTIMEN ULASAN PENGGUNA GAME HAY DAY DI
GOOGLE PLAY STORE MENGGUNAKAN MODEL
KLASIFIKASI

yang disusun dan diajukan oleh

Hanny Olivia

21.12.2005

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Februari 2025

Susunan Dewan Penguji

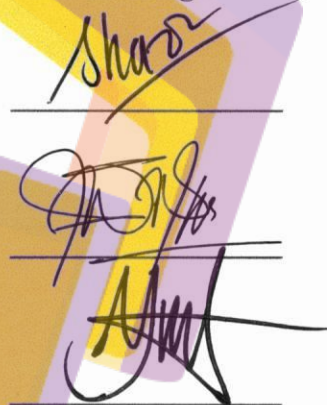
Nama Penguji

Tanda Tangan

Sharazita Dyah Anggita, S.Kom., M.Kom.
NIK. 190302285

Wiwi Widayani, S.Kom., M.Kom.
NIK. 190302272

Yoga Pristyanto, S.Kom., M.Eng.
NIK. 190302412



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hanny Olivia
NIM : 21.12.2005

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Sentimen Ulasan Pengguna Game Hay Day di Google Play Store
Menggunakan Model Klasifikasi**

Dosen Pembimbing : Yoga Pristyanto, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 februari 2025

Yang Menyatakan,



Hanny Olivia

HALAMAN PERSEMBAHAN

Alhamdulillah, 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 baik.

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

1. Terimakasih tak terhingga kepada kedua orang tua tercinta yakni Bapak Recim dan Ibu Ai Hisnawati, yang dengan segala pengorbanan kerja keras dan kasih sayang tulusnya selalu mendukung saya dalam setiap langkah. Meski tidak pernah sempat merasakan pendidikan sampai bangku perkuliahan, mereka tidak pernah leleh mengusahakan yang terbaik untuk saya. Doa, Motivasi, Kerja keras dan dukungan mereka menjadi kekuatan terbesar bagi saya hingga saya berhasil menyelesaikan skripsi ini. Terimakasih telah berhasil mengantarkan saya berada di tempat ini. Bapak, Mamah, Putri kecilmu kini telah dewasa dan siap menggapai mimpi yang lebih tinggi.
2. Adiku tersayang, Aditya Naufal Abyyu. Penulis persembahkan karya ini sebagai bentuk rasa cinta dan tanggung jawab untuk menjadi contoh dan inspirasi. Terimakasih atas doa, dukungan, serta keceriaan yang di berikan, yang membuat setiap perjalanan menjadi lebih berarti.
3. Terimakasih kepada bapak Yoga Pristyanto, S.Kom, M.Eng selaku dosen pembimbing. Terimakasih atas bimbingan dan dukungannya sehingga dapat menyelesaikan skripsi dengan baik.
4. Terimakasih kepada semua ibu bapak dosen Sistem Informasi yang telah banyak memberi ilmu dan pengalaman selama perkuliahan.
5. Kepada teman seperjuangan (Fanny Novatriana Kristianti, Jahnawi Kesianata, Aulia Putri Firdauhsya, Az Zahra Hijriah, Devi Andiska P)

terimakasih sudah kebersamai penulis dari awal perkuliahan hingga akhir.

6. Kepada ibu bapak dan teman teman humas penulis ucapkan terimakasih untuk semua dukungan dan motivasi yang diberikan.
7. Terimakasih penulis ucapkan kepada teman teman 21 Sistem Informasi 03 untuk semua canda dan tawa serta berbagai ilmu yang kita bagi, semoga kalian semua sukses di jalannya masing masing.
8. Seluruh pihak yang tidak bisa saya sebutkan satu persatu. Terimakasih atas segala bantuan dan doanya sehingga terselesaikannya skripsi ini.
9. Kepada sang penulis karya tulis ini, diri saya sendiri. Terimakasih atas keberanian untuk bermimpi, keteguhan untuk berjuang dan ketangguhan untuk tidak menyerah. Dari sekian banyaknya air mata yang jatuh, dari semua pikiran yang berantakan setiap malam, dari setiap renungan disudut kamar, terimakasih sudah memutuskan untuk tetap berusaha dan tidak menyerah. Rayakan setiap prosesmu, Berbahagialah dimanapun kapanpun kamu berada.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta
2. Bapak Hanif Al Fatta, S.Kom., M.Kom selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartanto, M.Kom selaku Ketua Program Studi Sistem Informasi
4. Bapak Yoga Pristyanto, S.Kom, M.Eng selaku dosen pembimbing yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan, saran, dan motivasi kepada penulis.
5. Bapak / Ibu Dosen Universitas Amikom Yogyakarta yang telah membekali penulis dengan berbagai ilmu dan pengalaman yang bermanfaat.
6. Kedua orang tua dan adik yang selalu memberikan dukungan semangat, material, dan kepada penulis.
7. Kepada semua pihak yang tidak bisa penulis sebutkan satu per satu, yang selalu memberikan semangat dan dukungannya kepada penulis dalam menyelesaikan skripsi ini.

Yogyakarta, 18 Februari 2025

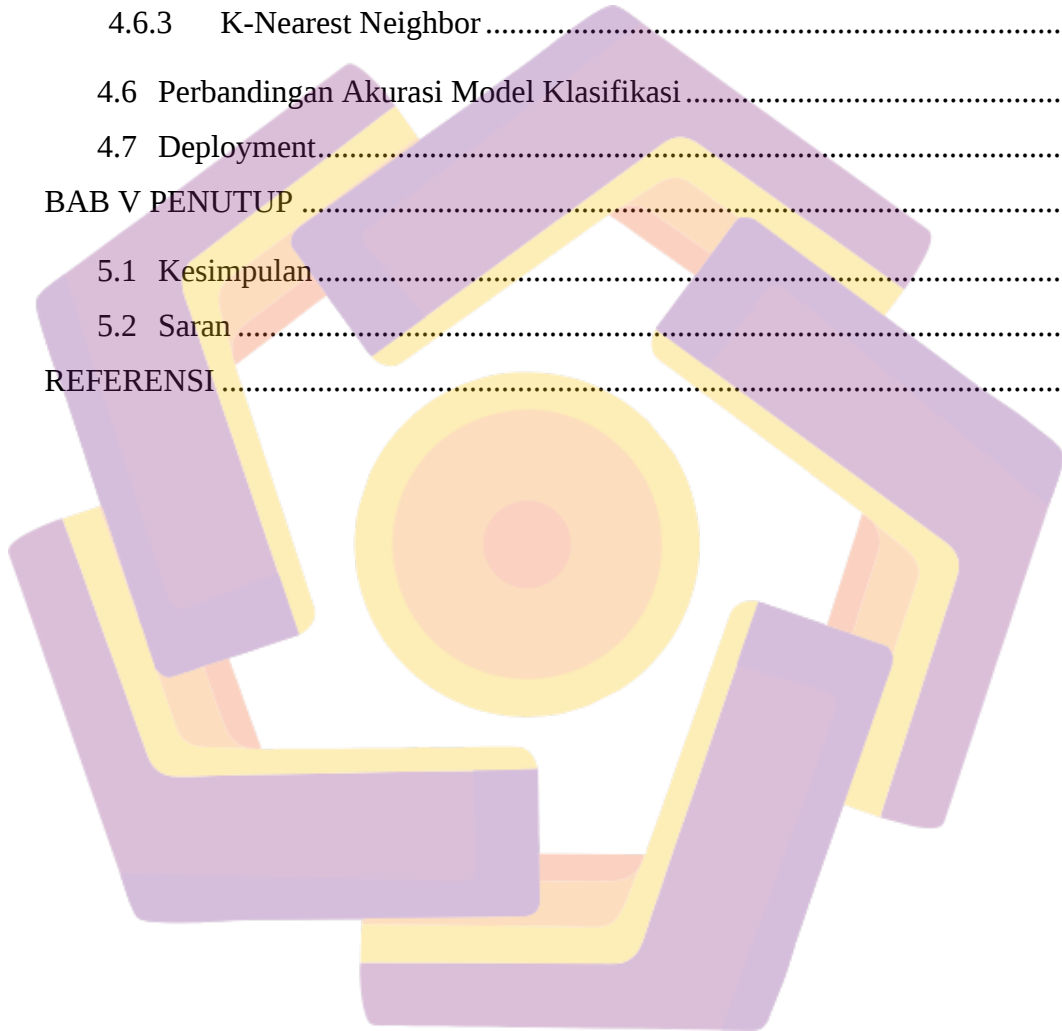
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	17
2.2.1 Analisis Sentimen	17
2.2.2 Google Play Store	17

2.2.3	Hay Day	17
2.2.4	Web Scraping.....	17
2.2.5	Klasifikasi	17
2.2.6	PreProcessing.....	18
2.2.7	Naïve bayes.....	19
2.2.8	Ter Frequency – Inverse Document Frequency (TF-IDF).....	19
2.2.9	Support Vector Machine	20
2.2.10	K-Nearest Neighbor	20
2.2.11	Confusion matrix	21
BAB III METODE PENELITIAN		23
3.1	Objek Penelitian.....	23
3.2	Alur Penelitian	23
3.2.1	Pengumpulan data.....	24
3.2.2	Pelabelan.....	24
3.2.3	Preprocessing.....	24
3.2.4	TF-IDF	29
3.2.5	Pembagian Data	29
3.2.6	Model Klasifikasi.....	29
3.2.7	Evaluasi model klasifikasi	29
3.2.8	Deployment.....	30
3.3	Alat dan Bahan.....	30
3.3.1	Data Penelitian	30
3.3.2	Alat/instrumen.....	30
BAB IV HASIL DAN PEMBAHASAN		32
4.1	Pengumpulan Data.....	32
4.2	Pelabelan.....	32

4.3 Preprocessing	33
4.4 TF-IDF	40
4.5 Model Klasifikasi	40
4.6.1 Naïve Bayes	41
4.6.2 Support Vector Machine	44
4.6.3 K-Nearest Neighbor	48
4.6 Perbandingan Akurasi Model Klasifikasi	51
4.7 Deployment	55
BAB V PENUTUP	56
5.1 Kesimpulan	56
5.2 Saran	56
REFERENSI	58



DAFTAR TABEL

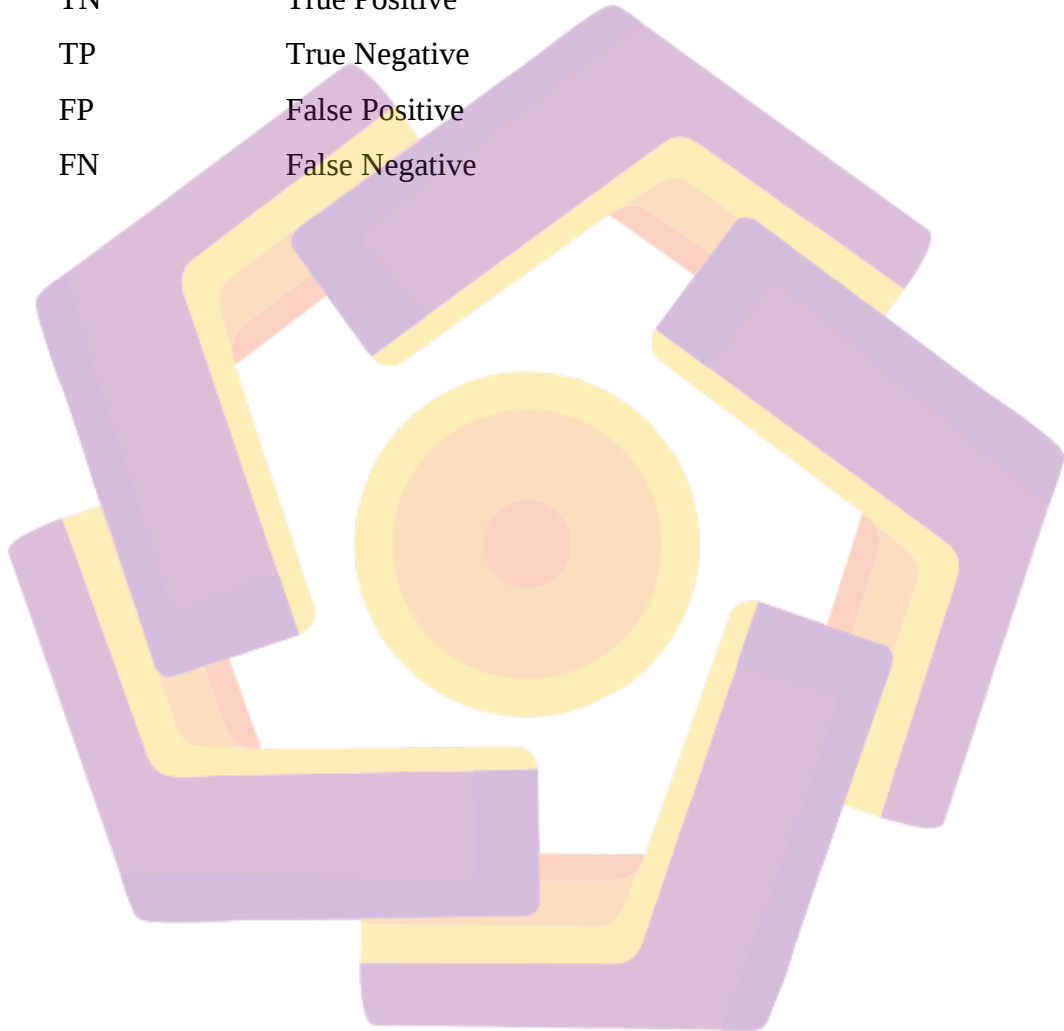
Tabel 2. 1 keaslian penelitian.....	9
Tabel 2. 2 Confusion Matrix	21
Tabel 4.1 tahap Cleaning ulasan	34
Tabel 4.2 tahap case folding ulasan	35
Tabel 4.3 tahap normalization ulasan	36
Tabel 4.4 tahap tokenization ulasan.....	37
Tabel 4.5 tahap stopword ulasan.....	38
Tabel 4.6 tahap stemming ulasan.....	39
Tabel 4.7 Hasil pembobotan TF-IDF	40
Tabel 4.8 hasil confusion matrix naïve bayes skenario 1 (90:10).....	41
Tabel 4.9 hasil confusion matrix naïve bayes skenario 1 (80:20).....	41
Tabel 4.10 hasil evaluasi model naïve bayes skenario 1.....	42
Tabel 4.11 hasil confusion matrix naïve bayes skenario 2 (90:10).....	43
Tabel 4.12 hasil confusion matrix naïve bayes skenario 2 (80:20).....	43
Tabel 4.13 Hasil evaluasi model Naïve Bayes Skenario 2	43
Tabel 4.14 hasil confusion matrix model SVM skenario 1 (90:10).....	44
Tabel 4.15 hasil confusion matrix model SVM skenario 1 (80:20).....	45
Tabel 4.16 hasil evaluasi model SVM skenario 1.....	45
Tabel 4.17 hasil confusion matrix model SVM skenario 2 (90:10).....	46
Tabel 4.18 hasil confusion matrix model SVM skenario 2 (80:20).....	47
Tabel 4.19 hasil evaluasi model SVM skenario 2.....	47
Tabel 4.20 hasil confusion matrix KNN skenario 1 (90:10).....	48
Tabel 4.21 hasil confusion matrix KNN skenario 1 (80:20).....	48
Tabel 4.22 hasil evaluasi model KNN skenario 1.....	49
Tabel 4.23 hasil confusion matrix KNN skenario 2 (90:10).....	50
Tabel 4.24 hasil confusion matrix KNN skenario 2 (80:20).....	50
Tabel 4.25 hasil evaluasi model KNN skenario 2.....	51

DAFTAR GAMBAR

Gambar 3.1 Alur penelitian.....	23
Gambar 4.1 hasil scraping data.....	32
Gambar 4.2 dataset final	32
Gambar 4.3 hasil pelabelan dataset.....	33
Gambar 4.4 penghapusan data duplikat	34
Gambar 4.5 Hasil Visualisasi Perbandingan Accuracy Tiap Model Klasifikasi ...	52
Gambar 4.6 Hasil Visualisasi Perbandingan Precision Tiap Model Klasifikasi	52
Gambar 4.7 Hasil Visualisasi Perbandingan Recall Tiap Model Klasifikasi	53
Gambar 4.8 Hasil Visualisasi Perbandingan F1-Score Tiap Model Klasifikasi	53
Gambar 4.9 Halaman deployment prediksi sentimen	55

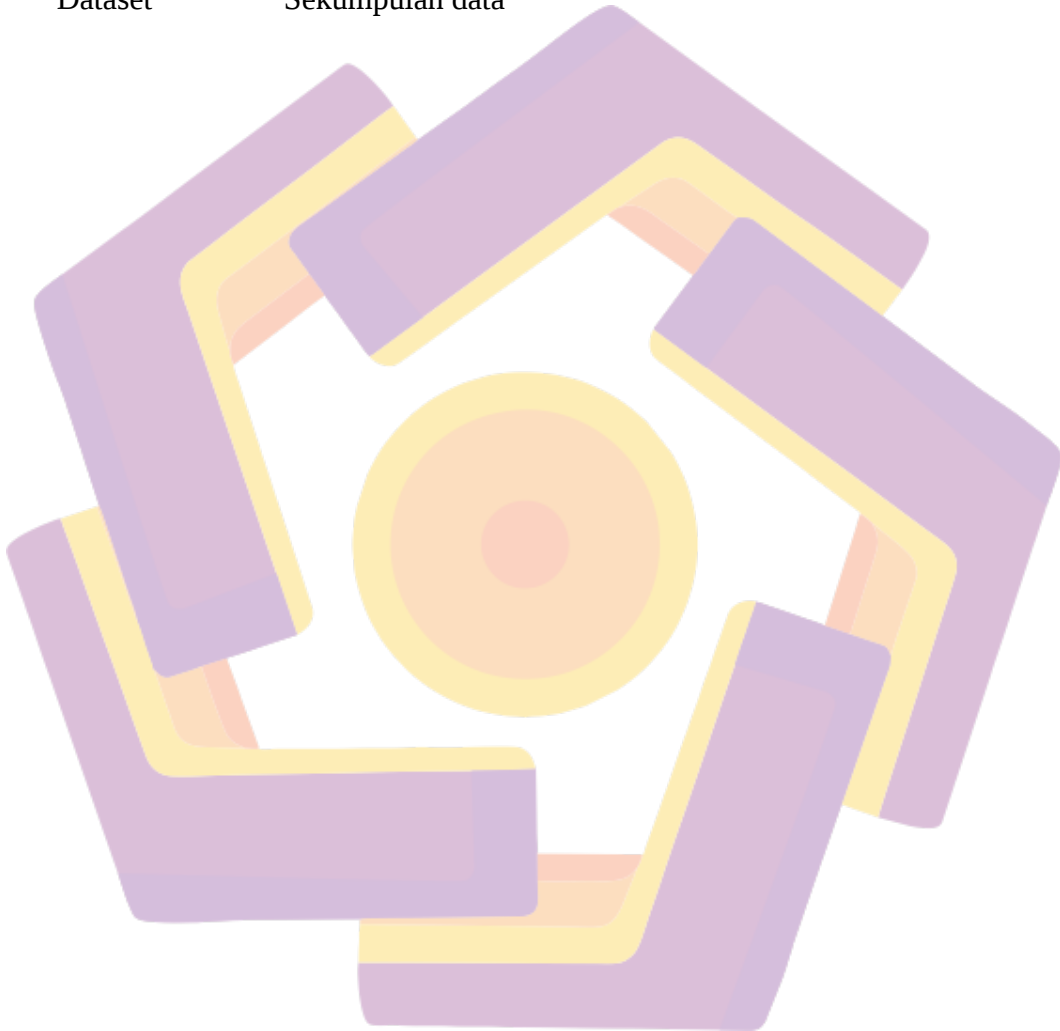
DAFTAR LAMBANG DAN SINGKATAN

SVM	Support Vector Machine
KNN	K-Nearest Neighbor
TF-IDF	Term Frequency Inverse Document Frequency
TN	True Positive
TP	True Negative
FP	False Positive
FN	False Negative



DAFTAR ISTILAH

Library	Perpustakaan
Scraping	mengumpulkan data dari web secara otomatis menggunakan kode program
Dataset	Sekumpulan data



INTISARI

Perkembangan teknologi yang pesat di era modern ini telah memberikan dampak signifikan pada berbagai sektor, termasuk industri game online. Hay Day, sebagai salah satu game simulasi pertanian yang populer di Google Play Store, telah diunduh oleh lebih dari 100 juta kali dan mendapat beragam macam ulasan pengguna. Ulasan-ulasan ini mempengaruhi keputusan calon pengguna dalam mengunduh dan memainkan aplikasi tersebut.

Penelitian ini dilakukan menggunakan tiga model klasifikasi yaitu Naïve Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN) untuk mengetahui model mana yang memiliki performa terbaik dalam mengklasifikasikan sentimen ulasan pengguna. Menggunakan dua skenario dengan penggunaan 3493 data dan 1000 data yang telah melalui proses text preprocessing dan konversi fitur menggunakan TF-IDF. Evaluasi dilakukan dengan membandingkan akurasi, presisi, recall, dan F1-score berdasarkan variasi jumlah data (1000 dan 3493) serta proporsi data latih (80% dan 90%).

Hasil penelitian menunjukkan bahwa model Naïve Bayes memiliki performa terbaik dengan akurasi tertinggi sebesar 86% pada dataset 1000 ulasan dan 90% data latih. Model Support Vector Machine dengan kernel RBF menghasilkan akurasi maksimal sebesar 81% pada kondisi yang sama, sementara model K-Nearest Neighbor mencapai akurasi terbaik sebesar 74%. Dengan melihat tingkat akurasi tertinggi, kestabilan performa pada berbagai skenario data, serta efisiensi proses pelatihan dan prediksi, model Naïve Bayes dinilai sebagai algoritma yang paling optimal dalam mengklasifikasi sentimen ulasan pengguna aplikasi Hay Day.

Kata kunci: Analisis Sentiment, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Hay Day.

ABSTRACT

The rapid development of technology in this modern era has had a significant impact on various sectors, including the online game industry. Hay Day, as one of the popular farming simulation games on the Google Play Store, has been downloaded more than 100 million times and received various user reviews. These reviews influence the decision of potential users in downloading and playing the application.

This study was conducted using three classification models, namely Naïve Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN) to find out which model has the best performance in classifying user review sentiments. Using two scenarios with the use of 3493 data and 1000 data that have gone through text preprocessing and feature conversion using TF-IDF. The evaluation was carried out by comparing accuracy, precision, recall, and F1-score based on variations in the amount of data (1000 and 3493) and the proportion of training data (80% and 90%).

The results showed that the Naïve Bayes model had the best performance with the highest accuracy of 86% on the 1000 review dataset and 90% training data. The Support Vector Machine model with the RBF kernel produced a maximum accuracy of 81% under the same conditions, while the K-Nearest Neighbor model achieved the best accuracy of 74%. By looking at the highest level of accuracy, performance stability on various data scenarios, and the efficiency of the training and prediction process, the Naïve Bayes model is considered the most optimal algorithm in classifying the sentiment of Hay Day application user reviews.

Keywords: Sentiment Analysis, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Hay Day.