

PERBANDINGAN KEAKURATAN METODE *NAÏVE BAYES* DAN *K-NEAREST NEIGHBORS* DALAM ANALISIS SENTIMEN APLIKASI HALODOC DI GOOGLE *PLAY STORE*

TUGAS AKHIR

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana
Program Studi S-1 Sistem Informasi



Disusun oleh:

AYU WULANSARI

202103025

**PROGRAM STUDI S-1 SISTEM INFORMASI
FAKULTAS TEKNIK & TEKNOLOGI INFORMASI
UNIVERSITAS JENDERAL ACHMAD YANI YOGYAKARTA
2024**

HALAMAN PENGESAHAN

TUGAS AKHIR

PERBANDINGAN KEAKURATAN METODE *NAIVE BAYES* DAN *K-NEAREST NEIGHBORS* DALAM ANALISIS SENTIMEN APLIKASI HALODOC DI GOOGLE *PLAYSTORE*

Diajukan oleh:

AYU WULANSARI
202103025

Telah dipertahankan di depan dewan penguji dan dinyatakan sah sebagai salah satu syarat untuk memperoleh gelar Sarjana di Fakultas Teknik & Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta

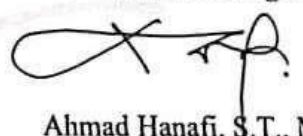
Tanggal: 02 Juli 2024

Mengesahkan:

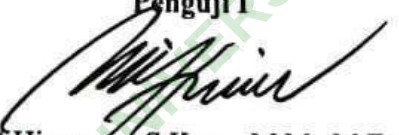
Pembimbing I


Ulfi Saadeta Aesyi, S.Kom., M.Cs.
NIDN: 0515129002

Pembimbing II


Ahmad Hanafi, S.T., M.Eng.
NIDN: 0528088301

Penguji I


Arif Himawan, S.Kom., M.M., M.Eng.
NIDN: 0517127402

Penguji II


Aris Wahyu Murdiyanto, S.Kom., M.Cs.
NIDN: 0502098501

Ketua Program Studi S-1 Sistem Informasi
Fakultas Teknik & Teknologi Informasi
Universitas Jenderal Achmad Yani Yogyakarta


Ahmad Hanafi, S.T., M.Eng.
NPP: 2008.13.0020

PERNYATAAN

Saya yang bertanda tangan di bawah ini, adalah mahasiswa Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta.

Nama : Ayu Wulansari
NPM : 202103025
Program Studi : S-1 Sistem Informasi
Judul Tugas Akhir : Perbandingan Keakuratan Metode *Naive Bayes* dan *K-Nearest Neighbors* Dalam Analisis Sentimen Aplikasi Halodoc di Google *Play store*

Menyatakan bahwa hasil penelitian dengan judul tersebut di atas adalah asli karya saya sendiri dan bukan hasil plagiarisme. Semua referensi dan sumber terkait yang dikutip dalam karya ilmiah ini telah ditulis sesuai kaidah penulisan ilmiah yang berlaku. Dengan ini, saya menyatakan untuk menyerahkan hak cipta penelitian kepada Universitas Jenderal Achmad Yani Yogyakarta guna kepentingan ilmiah.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya tanpa ada paksaan dari pihak mana pun. Apabila terdapat kekeliruan atau ditemukan adanya pelanggaran akademik di kemudian hari, maka saya bersedia menerima konsekuensi yang berlaku sesuai ketentuan akademik.

Yogyakarta, 13 Juni 2024



Ayu Wulansari

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul: “Perbandingan Keakuratan Metode NB dan *K-Nearest Neighbors* Dalam Analisis Sentimen Aplikasi Halodoc Di *Google Play Store*”. Penyusunan laporan ini merupakan salah satu persyaratan untuk menyelesaikan studi di Program Studi S-1 Sistem Informasi Fakultas Teknik & Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta. Laporan ini dapat diselesaikan atas bimbingan, arahan, dan bantuan dari berbagai pihak. Pada kesempatan ini penulis dengan rendah hati mengucapkan terima kasih dengan setulus-tulusnya kepada:

1. Bapak Aris Wahyu Murdiyanto, S.Kom., M.Cs. selaku Dekan Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta.
2. Bapak Ahmad Hanafi, S.T., M.Eng. selaku Ketua Program Studi S-1 Sistem Informasi Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
3. Ibu Ulfi Saidata Aesyti, S.Kom., M.Cs. dan Bapak Ahmad Hanafi, S.T., M.Eng. selaku Dosen Pembimbing Tugas Akhir;
4. Para dosen yang telah memberikan banyak bekal ilmu pengetahuan kepada penulis selama menjadi mahasiswa di Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
5. Alm. Ayah Priyanto, ayahanda tercinta saya yang selalu saya rindukan dan menjadi alasan utama saya untuk bangkit dari keterpurukan, terus menjadi inspirasi dalam setiap langkah saya. Kehilangan beliau merupakan pukulan berat, namun semangat dan kasih sayangnya tetap hidup dalam hati saya, memberi saya kekuatan untuk terus berjuang dan tidak pernah menyerah.
6. Pintu surgaku, Ibu Sumini, yang tidak henti-hentinya memberikan kasih sayang dengan penuh cinta, selalu memberikan dukungan, motivasi, serta do'a yang dipanjatkan selama ini sehingga penulis mampu menyelesaikan karya tulis sederhana ini. Terima kasih telah merawat dan melahirkan saya. Semoga Ibu sehat

selalu dan hidup lebih lama sehingga Ibu selalu berada di setiap perjalanan dan pencapaian saya.

7. Adikku, Partner dan Sahabatku Ini Ganisa, Paskalis Rivaldianus, Bagas, Zul, Dini, Nika yang telah memberikan semangat dan doa kepada saya.
8. Rekan-rekan mahasiswa Prodi S-1 Sistem Informasi di Universitas Jenderal Achmad Yani Yogyakarta yang sudah memberi dukungan dan kerja sama selama pembuatan tugas akhir.

Penulis menyadari bahwa laporan tugas akhir ini masih jauh dari kata sempurna. Maka dari itu dengan segala kerendahan hati penulis sangat menghargai adanya kritik dan saran yang membangun dari semua pihak yang bersedia meluangkan waktu untuk membaca laporan tugas akhir ini.

Yogyakarta, 13 Juni 2024



Ayu Wulansari

DAFTAR ISI

Halaman Pengesahan	ii
Pernyataan	iii
Kata Pengantar	iv
Daftar Isi	vi
Daftar Tabel	ix
Daftar Gambar	x
Daftar Lampiran	xii
Daftar Singkatan	xiii
Intisari	xiv
Abstract	xv
Bab 1 Pendahuluan	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Pertanyaan Penelitian	3
1.5 Tujuan Penelitian	3
1.6 Manfaat Hasil Penelitian	4
Bab 2 Tinjauan Pustaka dan Landasan Teori	5
2.1 Tinjauan Pustaka	5
2.2 Landasan Teori	10
2.2.1 Analisis Sentimen	10
2.2.2 Data Mining	10
2.2.3 Naive Bayes	11
2.2.4 K-Nearest Neighbor	11
2.2.5 K- Fold Cross Validation	12
Bab 3 Metode Penelitian	13
3.1 Bahan dan Alat Penelitian	14
3.2 Jalan Penelitian	14
3.2.1 Pengumpulan Data	16

3.2.1.1	Web Scraping.....	16
3.2.2	Preprocessing.....	16
3.2.3	Labelling.....	18
3.2.4	K-Fold Cross Validation.....	18
3.2.5	TF IDF.....	18
3.2.6	Klasifikasi.....	18
3.2.7	Evaluasi.....	20
Bab 4	Hasil Penelitian	21
4.1	Ringkasan Hasil Penelitian.....	21
4.2	Hasil Penelitian.....	21
4.2.1	Hasil Pengumpulan Data	21
4.3	Hasil Preprocessing Data.....	23
4.3.1	Case Folding.....	23
4.3.2	Cleaning.....	24
4.3.3	Tokenizing.....	26
4.3.4	Stopwords.....	26
4.3.5	Stemming.....	27
4.3.6	Normalization.....	28
4.3.7	Menghapus Nan.....	29
4.4	Hasil Labeling	30
4.5	K-Fold Validation.....	31
4.6	Fitur Ekstraksic.....	34
1.	TF – IDF	34
4.7	Naïve Bayes.....	35
4.8	K-Nearest Neighbors	36
4.9	Hasil Analisis Perbandingan.....	38
4.10	Hasil Analisis Sentimen Naïve Bayes dan KNN.....	40
4.11	Hasil Perbandingan Penelitian Terdahulu Dan Terbaru	42
Bab 5	Kesimpulan dan Saran.....	43
5.1	Kesimpulan.....	43
5.2	Saran.....	44

Daftar Pustaka.....	45
Lampiran.....	48

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR TABEL

Tabel 1 Tinjauan Pustaka.....	6
Tabel 2 Hasil Scrapping.....	21
Tabel 3 Perbandingan Hasil Penelitian Terdahulu dan Terbaru	42

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR GAMBAR

Gambar 1.1 Ulasan Pengguna Halodoc.....	1
Gambar 3.1 Metode Penelitian	13
Gambar 3.2 Flowchart Jalan Penelitian.....	15
Gambar 3.3 Proses Web Scrapping.....	16
Gambar 4.1 Kode Program Case Folding.....	23
Gambar 4.2 Hasil Case Folding	23
Gambar 4.3 Kode Program Special Removal.....	24
Gambar 4.4 Kode Program Number Removal	24
Gambar 4.5 Kode Program Punctuation Removal	24
Gambar 4.6 Kode Program Whitespaces Removal	25
Gambar 4.7 Kode Program Single Char.....	25
Gambar 4.8 Hasil Cleaning	25
Gambar 4.9 Kode Program Tokenizing.....	26
Gambar 4.10 Hasil Tokenizing.....	26
Gambar 4.11 Kode Program Stopwords.....	27
Gambar 4.12 Hasil Stopwords.....	27
Gambar 4.13 Kode Program Stemming	28
Gambar 4.14 Hasil Stemming	28
Gambar 4.15 Kode Program Normalization.....	29
Gambar 4.16 Hasil Normalization.....	29
Gambar 4.17 Kode Penghapus NaN.....	30
Gambar 4.18 Hasil Pelabelan	30
Gambar 4.19 Jumlah Label	31
Gambar 4.20 Keterangan Split Data K-Fold Naive Bayes.....	32
Gambar 4.21 Split Data K-Fold Naive Bayes	32
Gambar 4.22 Split Data K-Fold KNN.....	33
Gambar 4.23 Keterangan Split Data K-Fold KNN	33
Gambar 4.24 Hasil Perhitungan TF-IDF	34
Gambar 4.25 Hasil Prediksi Naive Bayes	35

Gambar 4.26 Confusion Matrix Naive Bayes	36
Gambar 4.27 Classification Report Naive Bayes	36
Gambar 4.28 Nilai k Terbaik.....	37
Gambar 4.29 Hasil Prediksi KNN.....	37
Gambar 4.30 Confusion Matrix KNN	38
Gambar 4.31 Classification Report KNN.....	38
Gambar 4.32 Perbandingan Klasifikasi NB dan KNN.....	39
Gambar 4.33 Perbandingan Klasifikasi K-Fold Cross Validation	40
Gambar 4.34 Sentimen Positif.....	40
Gambar 4.35 Sentimen Negatif	41

DAFTAR LAMPIRAN

LAMPIRAN 1 Jadwal Penelitian	48
LAMPIRAN 2 Lembar Bimbingan Dosen.....	49
LAMPIRAN 3 Data Hasil Penelitian	50
LAMPIRAN 4 Hasil Cek Plagiarisme	51

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA

DAFTAR SINGKATAN

NB	<i>Naive Bayes</i>
KNN	<i>K-Nearest Neighbors</i>
TF	<i>Term Frequency</i>
IDF	<i>Inverse Document Frequency</i>
TP	<i>True Positif</i>
TN	<i>True Negatif</i>
FP	<i>False Positif</i>
FN	<i>False Negatif</i>

PERPUSTAKAAN
UNIVERSITAS JENDERAL ACHMAD YANI
YOGYAKARTA