

**PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBORS*
(KNN) DAN *NAÏVE BAYES* DALAM ANALISIS SENTIMEN
PENGGUNA CAPCUT DI ULASAN *PLAYSTORE***



Skripsi

Untuk memenuhi syarat memperoleh Derajat
Sarjana Teknik (S.T.)

Oleh:

TETI HERLIANA

NIM 2001020017

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG
2024**

**PERBANDINGAN ALGORITMA *K-NEAREST NEIGHBORS*
(KNN) DAN *NAÏVE BAYES* DALAM ANALISIS SENTIMEN
PENGGUNA CAPCUT DI ULASAN *PLAYSTORE***



Pembimbing I,

Nurfalinda S.T., M.Cs.
NIP. 198401212019032013

Pembimbing II

Muhamad Radzi Rathomi, S.Kom., M.Cs.
NIP. 198903252019031014

HALAMAN PENGESAHAN

Judul Skripsi : Perbandingan Algoritma K-Nearest Neighbor (KNN) dan *Naïve Bayes* dalam Analisis Sentimen Pengguna CapCut di Ulasan *PlayStore*
Nama Mahasiswa : Teti Herliana
NIM : 2001020017
Jurusan : Teknik Informatika

Telah dipertahankan di depan Dewan Penguji dan dinyatakan lulus
pada tanggal 22 Juli 2024

Susunan Tim Pembimbing dan Penguj

Jabatan	Nama Dosen	Tanda Tangan	Tanggal
Pembimbing I	: Nurfalinda S.T., M.Cs.		29/07-2024
Pembimbing II	: Muhamad Radzi Rathomi, S.Kom., M.Cs.		29/7/2024
Ketua Penguji	: Ferdi Chahyadi, S.Kom., M.Cs.		29/7/24
Anggota Penguji I	: Nola Ritha, S.T., M.Cs.		29/07/2024
Anggota Penguji II	: Novrizal Fattah Fahmitra, S.Kom., M.Cs.		29/07/2024

Tanjungpinang, 30 Juli 2024
Universitas Maritim Raja Ali Haji
Dekan Fakultas Teknik dan Teknologi Kemaritiman



Ir. Sapti Nugraha, S.T., M.Eng
NIP. 198904132015041005

PERNYATAAN ORISINALITAS

Dengan ini saya menyatakan bahwa skripsi saya yang berjudul Perbandingan Algoritma *K-Nearest Neighbor* (KNN) dan *Naïve Bayes* dalam Analisis Sentimen Pengguna CapCut di Ulasan *PlayStore* adalah benar karya saya dengan arahan dari komisi pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Jika kemudian hari ternyata terbukti pernyataan saya ini tidak benar dan melanggar peraturan yang sah dalam karya tulis dan hak intelektual maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Maritim Raja Armali Haji.

Tanjungpinang, 30 Juli 2024

Yang menyatakan



(Teti Herliana)

HALAMAN PERSEMBAHAN

Dengan segenap rasa syukur, kerendahan hati dan sebagai bentuk penghargaan serta terima kasih, saya mempersembahkan karya skripsi ini kepada :

1. Allah Subhanahu Wata'ala, yang telah memberikan segala bentuk nikmat dan rezeki serta kesempatan yang luar biasa sehingga saya dapat menyelesaikan tahap pendidikan ini.
2. Kedua orang tua saya, Ibu Yuli Dormauli Simarmata dan Bapak Amlan yang selalu berjasa melebihi siapapun dan dalam hal apapun. Serta adik-adik saya Wilyan Pirgiansyah dan M. Rio Saputra yang selalu menjadi penyemangat dalam menempuh dunia pendidikan.
3. Keluarga besar penulis yang senantiasa memberikan support dalam perjalanan menempuh dunia pendidikan.
4. Ibu Nurfalinda S.T., M.Cs dan Bapak Muhamad Radzi Rathomi, S.Kom., M.Cs. selaku dosen pembimbing yang telah memberikan segala bentuk ilmu pengetahuan, arahan dan dukungannya selama proses penyelesaian skripsi.
5. Teman – teman Teknik Informatika angkatan 20 yang selalu membantu dan berproses bersama selama masa perkuliahan.
6. Ejika, Nuralifa, Nifia, Farida, Trinanda, M.Irfan Raif, Boyke, yang senantiasa membantu, mendukung dan memotivasi penulis dalam menyelesaikan tahap skripsi ini.
7. Seluruh pihak yang telah membantu, memberikan support, memberikan doa, serta memberikan pertanyaan kapan wisuda.

HALAMAN MOTO

“ Cukup bagi kami Allah, dan Dia sebaik-baik penolong “

“ No matter who you are, where u from, your skin color, your gender identity.
Just speak yourself ” (Kim Namjoon)



KATA PENGANTAR

Puji syukur kehadirat Allah Subhanahu Wa Ta'ala, atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul Perbandingan Algoritma K-Nearest Neighbor (KNN) dan *Naïve Bayes* dalam Analisis Sentimen Pengguna CapCut di Ulasan *PlayStore*. Skripsi ini penulis susun untuk memenuhi syarat memperoleh gelar sarjana pada Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi Kemaritiman UMRAH.

Selama proses penulisan skripsi ini, penulis mendapatkan banyak doa, dukungan, bimbingan serta bantuan dari banyak pihak. Oleh karenanya, penulis menyampaikan rasa terima kasih yang sangat tulus kepada :

1. Allah Subhanahu Wata'ala, yang telah memberikan segala bentuk rahmat dan ridho-Nya sehingga saya dapat menyelesaikan skripsi ini.
2. Orang Tua dan keluarga besar, yang selalu memberikan dukungan dalam berbagai bentuk kepada penulis.
3. Bapak dan Ibu Dosen Pembimbing yang senantiasa memberi arahan, bimbingan, serta solusi kepada penulis dalam menyelesaikan skripsi ini.
4. Sahabat seperjuangan Teknik Informatika 20 yang telah menemani, memberi semangat dan motivasi serta kebersamaan selama masa perkuliahan dan seluruh pihak yang telah membantu penulis dalam proses skripsi maupun proses perkuliahan.
5. *Last but not least, I wanna thank Me. I wanna thank me for believing in me. I wanna thank me for doing all this hard work. I wanna thannk me for having no days off. I wanna thank me for, for never quitting.*

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan untuk perbaikan di masa mendatang

Tanjungpinang, 03 Juli 2024



(Teti Herliana)

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERETUJUAN	ii
HALAMAN PENGESAHAN	iii
PERNYATAAN ORISINALITAS	iv
HALAMAN PERSEMBAHAN	v
HALAMAN MOTO	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xii
GLOSARIUM	xiii
ABSTRAK	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
★ BAB II KAJIAN LITERATUR	5
2.1 Kajian Literatur	5
2.2 Landasan Teori	9
2.2.1 Analisis Sentimen	9
2.2.2 Algoritma <i>K-Nearest Neighbors</i>	9
2.2.3 Algoritma Naïve Bayes	11
2.2.4 Pre-Processing Data	12
2.2.5 Pembobotan TF-IDF	13
2.2.6 Confussion Matrix	14
2.2.7 Indonesia Sentiment (InSet) Leksikon	15
2.2.8 Python	16
2.2.9 Aplikasi CapCut	16
2.2.10 <i>PlayStore</i>	17
2.2.11 Web Scraping	17
BAB III METEDO PENELITIAN	19
3.1 Sub Waktu Penelitian	19
3.2 Jenis Penelitian	19
3.3 Instrumen Penelitian	19
3.4 Prosedur Penelitian	19
3.5 Analisis dan Perancangan	22
3.5.1 Analisis data	22
3.5.2 Pengolahan Data	23
3.5.3 Pemodelan <i>K-Nearest Neighbor (KNN)</i>	33

3.5.4	Pemodelan Naïve Bayes	35
3.6	Perancangan Model	39
3.7	Flowchart Model <i>Naïve Bayes</i>	41
3.8	Perancangan Sistem.....	43
3.8.1	Flowchart sistem.....	43
3.8.2	Diagram Arsitektur	44
3.8.3	Use Case Diagram	45
3.8.4	Sequence Diagram	45
3.8.5	Activity Diagram	47
3.8.6	Perancangan UI.....	47
BAB IV HASIL DAN PEMBAHASAN		49
4.1	Pelatihan dan Akurasi Model KNN	49
4.2	Pelatihan dan Akurasi Model <i>Naïve Bayes</i>	51
4.3	Pengujian Model KNN	52
4.3.1	Pengujian Performa Waktu Model KNN.....	52
4.3.2	Pengujian Akurasi Model KNN	55
4.4	Pengujian Model Naïve Bayes	58
4.4.1	Pengujian Performa Waktu Model Naïve Bayes.....	58
4.4.2	Pengujian Akurasi Model Naïve Bayes.....	60
4.5	Pembahasan	61
4.6	Implementasi Sistem	63
4.6.1	Halaman Beranda	63
4.6.2	Halaman Klasifikasi Sentimen	63
★	BAB V PENUTUP	65
5.1	Kesimpulan.....	65
5.2	Saran.....	65
DAFTAR PUSTAKA		66

DAFTAR TABEL

Tabel 3.1 <i>Data Latih</i>	22
Tabel 3.2. <i>Data Uji</i>	22
Tabel 3.3. <i>Hasil cleaning data latih</i>	23
Tabel 3.4. <i>Hasil cleaning data uji</i>	24
Tabel 3.5. <i>Case Folding data uji</i>	24
Tabel 3.6. <i>Case folding data latih</i>	25
Tabel 3.7. <i>Tokenizing data uji</i>	25
Tabel 3.8. <i>Tokenizing data latih</i>	25
Tabel 3.9. <i>Hasil normalisasi data latih</i>	26
Tabel 3.10. <i>Normalisasi data uji</i>	26
Tabel 3.11. <i>Stopword data latih</i>	27
Tabel 3.12. <i>Stopword data uji</i>	27
Tabel 3.13. <i>Pelabelan data latih</i>	28
Tabel 3.14. <i>Pelabelan data uji</i>	28
Tabel 3.15. <i>TF data latih</i>	29
Tabel 3.16. <i>TF data uji</i>	29
Tabel 3.17. <i>Pembobotan TF data latih</i>	30
Tabel 3.18. <i>Pembobotan IDF</i>	31
Tabel 3.19. <i>TF-IDF</i>	32
Tabel 3.20. <i>Jumlah bobot TF-IDF</i>	33
Tabel 3.21. <i>Penjumlahan bobot data latih</i>	34
Tabel 3.22. <i>IDF data uji</i>	36
Tabel 3.23. <i>TF-IDF data uji</i>	36
Tabel 3.24. <i>IDF kelas sentiment</i>	37
Tabel 4.1. <i>Akurasi Pengujian dari Model KNN</i>	49
Tabel 4.2. <i>Performa elatihan Algoritma Naive Bayes</i>	51
Tabel 4.3. <i>Performa waktu KNN berdasarkan kata input</i>	53
Tabel 4.4. <i>Performa waktu KNN berdasarkan kata input</i>	54
Tabel 4.5. <i>Performa waktu Naive Bayes berdasarkan kata input</i>	58

Tabel 4.6. *Performa waktu Naïve Bayes berdasarkan data inputan..... 59*



DAFTAR GAMBAR

Gambar 2.1. <i>Flowchart knn</i>	10
Gambar 2.2. <i>Confussion matrix 3x3</i>	15
Gambar 3.1 <i>Flowchart penelitian</i>	20
Gambar 3.2. <i>Flowchart model KNN</i>	40
Gambar 3.3. <i>Flowchart model Naïve Bayes</i>	42
Gambar 3.4. <i>Flowchart sistem</i>	43
Gambar 3.5. <i>Diagram Arsitektur</i>	44
Gambar 3.6. <i>Usecase Diagram</i>	45
Gambar 3.7. <i>Sequence diagram KNN</i>	46
Gambar 3.8. <i>Sequence diagram Naïve Bayes</i>	46
Gambar 3.9. <i>Activity Diagram</i>	47
Gambar 3.10. <i>Perancangan halaman beranda</i>	48
Gambar 3.11. <i>Perancangan halaman kalsifikasi</i>	48
Gambar 4.1. <i>Grafik Akurasi Pelatihan dari Model KNN</i>	51
Gambar 4.2. <i>Confussion Matrix Algoritma KNN</i>	56
Gambar 4.3. <i>Confussion matriks Naïve Bayes</i>	60
Gambar 4.4. <i>Halaman beranda</i>	63
Gambar 4.5. <i>Halaman klasifikasi</i>	64

GLOSARIUM

Scraping	proses untuk melakukan pengambilan informasi dari situs website yang disimpan dalam format yang dapat dianalisis
Accuracy	seberapa sering model membuat prediksi yang benar dengan menghitung jumlah prediksi yang benar dibagi dengan total semua prediksi
<i>Precision</i>	seberapa banyak prediksi positif yang benar, dibandingkan dengan semua prediksi positif
Recall	seberapa banyak kasus positif yang ditemukan oleh model yang dibuat
F1 Score	rata-rata dari presisi dan recall.

