

**ANALISIS KOMPARASI MODEL KLASIFIKASI UNTUK ULASAN
APLIKASI SIREKAP 2024 MENGGUNAKAN CART DAN METODE
WORDNET**

SKRIPSI

Diajukan Oleh :

Muhammad Fariz Ijlal Rafi

2011102441124



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
SAMARINDA**

2024

**ANALISIS KOMPARASI MODEL KLASIFIKASI UNTUK ULASAN
APLIKASI SIREKAP 2024 MENGGUNAKAN CART DAN METODE
WORDNET**

SKRIPSI

Diajukan Sebagai Salah Satu Persyaratan Untuk Memenuhi Gelar Sarjana Fakultas Sains dan
Teknologi Universitas Muhammadiyah Kalimantan Timur

Diajukan Oleh :

Muhammad Fariz Ijlal Rafi

2011102441124



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
SAMARINDA**

2024

LEMBAR PERSETUJUAN

ANALISIS KOMPARASI MODEL KLASIFIKASI UNTUK ULASAN APLIKASI SIREKAP 2024 MENGGUNAKAN CART DAN METODE WORDNET

Diajukan Oleh :

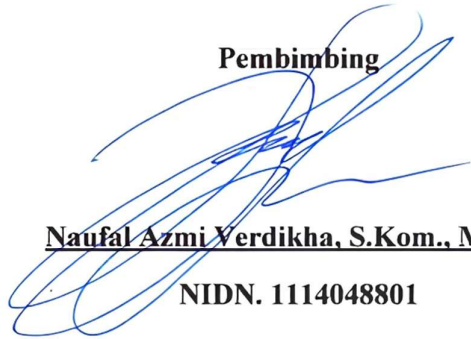
MUHAMMAD FARIZ IJLAL RAFI

2011102441124

Disetujui untuk diujikan

Pada tanggal 28 Juni 2024

Pembimbing



Naufal Azmi Verdikha, S.Kom., M.Eng.

NIDN. 1114048801

Mengetahui,

Koordinator Tugas Akhir/Skripsi/Tesis/Disertasi



Abdul Rahim, S.Kom., M.Cs.

NIDN. 0009047901

LEMBAR PENGESAHAN

ANALISIS KOMPARASI MODEL KLASIFIKASI UNTUK ULASAN APLIKASI SIREKAP 2024 MENGGUNAKAN CART DAN METODE WORDNET

SKRIPSI


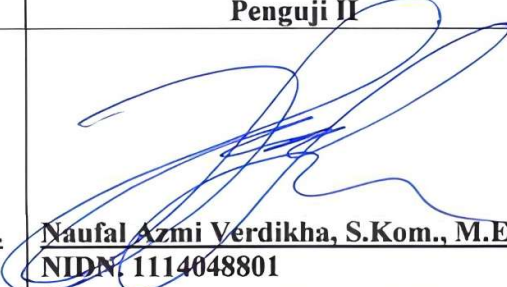
Diajukan Oleh:

Muhammad Fariz Ijlal Rafi

2011102441124

Diseminarkan dan Diujikan

Pada tanggal 11 Juli 2024

Penguji I	Penguji II
 <u>Taghfirul Azhima Yoga Siswa, S.Kom., M.Kom.</u> NIDN. 1118038805	 <u>Naufal Azmi Verdikha, S.Kom., M.Eng.</u> NIDN. 1114048801

Mengetahui,

Ketua

Program Studi Teknik Informatika



Arbansyah, S.Kom., M.T.I.

NIDN. 1118019203

PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan dibawah ini:

Nama : Muhammad Fariz Ijlal Rafi
NIM : 2011102441124
Program Studi : S1 Teknik Informatika
Judul Penelitian : ANALISIS KOMPARASI MODEL KLASIFIKASI UNTUK ULASAN APLIKASI SIREKAP 2024 MENGGUNAKAN CART DAN METODE *WORDNET*

Menyatakan bahwa **skripsi** yang saya tulis ini benar-benar hasil karya saya sendiri, dan bukan merupakan hasil plagiasi/falsifikasi/fabrikasi baik sebagian atau seluruhnya.

Atas pernyataan ini, saya siap menanggung resiko atau sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam **skripsi** saya ini, atau klaim dari pihak lain terhadap keaslian karya saya ini.

Samarinda, 18 Juni 2024
Yang membuat pernyataan



Muhammad Fariz Ijlal Rafi
NIM. 2011102441124

ABSTRAK

Model klasifikasi sentimen adalah model yang digunakan untuk mengkategorikan teks berdasarkan sentimen yang terkandung di dalamnya. Terdapat beberapa jenis model klasifikasi sentimen seperti model *Wordnet* dan algoritma klasifikasi *Machine Learning* seperti model *Decision Tree* CART. Penelitian ini bertujuan untuk membandingkan dua jenis model klasifikasi sentimen yaitu *Wordnet* dan *Decision Tree* dengan algoritma CART dalam mengklasifikasi sentimen dari sebuah ulasan yang dievaluasi menggunakan *F1-Score*. Dataset yang digunakan adalah data ulasan aplikasi SIREKAP 2024 dari *Google Playstore* dengan *rating* 1 hingga 5 bintang yang diambil pada awal bulan februari 2024. Hasil penelitian menunjukkan, model *Decision Tree* dengan algoritma CART memiliki nilai *F1-Score* yang lebih tinggi yaitu sebesar 31,62%, sedangkan model *Wordnet* memiliki nilai *F1-Score* sebesar 17,70%. Penelitian ini membuktikan bahwa kedua model dapat digunakan untuk mengklasifikasi sentimen terhadap ulasan aplikasi SIREKAP 2024, dan model *Decision Tree* dengan algoritma CART terbukti lebih baik dibandingkan model *Wordnet* dalam mengklasifikasi sentimen dari ulasan.

Kata Kunci: Analisis Sentimen, Ulasan, SIREKAP 2024, *Wordnet*, *Decision Tree*, CART

ABSTRACT

Sentiment classification model is a model used to categorize text based on the sentiment contained in it. There are several types of sentiment classification models such as the *Wordnet* model and *Machine Learning* classification algorithms such as the *Decision Tree* CART model. This study aims to compare two types of sentiment classification models, namely *Wordnet* and *Decision Tree* with the CART algorithm in classifying sentiment from a review evaluated using *F1-Score*. The dataset used is the SIREKAP 2024 application review data from *Google Playstore* with a *rating* of 1 to 5 stars taken in early February 2024. The results of the study showed that the *Decision Tree* model with the CART algorithm had a higher *F1-Score* value of 31.62%, while the *Wordnet* model had an *F1-Score* value of 17.70%. This study proves that both models can be used to classify sentiment towards the SIREKAP 2024 application review, and the *Decision Tree model* with the CART algorithm is proven to be better than the *Wordnet* model in classifying sentiment from reviews.

Keywords: Sentiment Analysis, Reviews, SIREKAP 2024, *Wordnet*, *Decision Tree*, CART

PRAKATA

Alhamdulillah, dengan mengucapkan puji syukur kehadiran Allah SWT karena atas petunjuk-Nya, penulis dapat menyelesaikan skripsi ini dengan judul “ANALISIS KOMPARASI MODEL KLASIFIKASI UNTUK ULASAN APLIKASI SIREKAP 2024 MENGGUNAKAN CART DAN METODE *WORDNET*” sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.

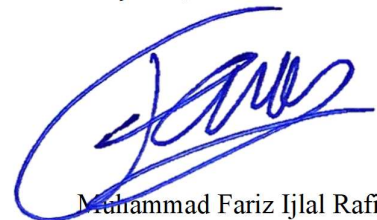
Segala kekurangan dalam penulisan skripsi ini sangat disadari oleh penulis, namun atas bantuan dan dukungan moral maupun materi dari sebagian pihak dapat terpenuhi segala kelengkapan dalam penulisan. Pada kesempatan ini penulis menyampaikan ucapan terimakasih serta penghargaan yang setinggi-tingginya untuk berbagai pihak yang telah memberikan bantuan baik secara langsung atau tidak langsung.

1. Kedua orang tua tercinta Ayahanda Kusnadi dan Ibunda Surtinah yang telah memberikan dukungan serta semangat, doa, nasehat, motivasi kepada penulis sehingga proposal ini dapat diselesaikan dengan baik.
2. Bapak Naufal Azmi Verdikha, S.kom., M.Eng sebagai dosen pembimbing yang telah memberikan waktu bimbingan baik masukan, saran dan pengetahuan yang sangat membantu dalam menyelesaikan skripsi ini.
3. Bapak Taghfirul Azhima Yoga Siswa, S.Kom., M.Kom., sebagai dosen penguji yang telah memberikan saran-saran bermanfaat, sehingga penulis dapat menyempurnakan isi proposal skripsi ini.
4. Bapak Arbansyah, S.Kom., M.TI selaku ketua Program Studi Teknik Informatika.
5. Teman-teman seperjuangan baik dari prodi Teknik Informatika maupun dari prodi lain yang tidak bisa disebutkan satu-satu.

Semoga semua do'a, bimbingan dan dukungan yang telah diberikan kepada penulis mendapat balasan dari Allah SWT. Penulis berharap semoga laporan ini memberikan kesan yang bermanfaat kepada pembaca.

Samarinda, 18 Juni 2024

Penyusun,



Muhammad Fariz Ijlal Rafi

DAFTAR ISI

	Halaman
HALAMAN JUDUL	ii
LEMBAR PERSETUJUAN	iii
LEMBAR PENGESAHAN	iv
PERNYATAAN KEASLIAN PENELITIAN	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
PRAKATA	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	14
1.1 Latar Belakang Masalah	14
1.2 Rumusan Masalah	15
1.3 Tujuan Penelitian	15
1.4 Manfaat Penelitian	15
BAB II METODE PENELITIAN	16
2.1 Objek Penelitian	16
2.2 Alat dan Bahan	16
2.3 Prosedur Penelitian	17
2.3.1 Pengambilan Data	17
2.3.2 <i>Pra-proses</i> Data	18
2.3.3 <i>Wordnet</i>	19
2.3.4 Klasifikasi <i>Decision Tree</i> CART	23
BAB III HASIL DAN PEMBAHASAN	27
3.1 Hasil	27
3.1.1 Pengambilan Data	27
3.1.2 <i>Pra-proses</i> Data	30
3.1.3 <i>Wordnet</i>	33
3.1.4 Klasifikasi <i>Decision Tree</i> CART	35
3.2 Pembahasan	40
3.2.1 Pengambilan dan Analisis Data	40
3.2.2 Analisis Hasil <i>Pra-proses</i>	40
3.2.3 Analisis Komparasi Model Klasifikasi <i>Wordnet</i> dan CART	40
BAB IV PENUTUP	42

4.1	Kesimpulan.....	42
4.2	Saran.....	42
	DAFTAR PUSTAKA.....	43
	LAMPIRAN.....	45
	RIWAYAT HIDUP.....	60

DAFTAR TABEL

Tabel	Halaman
2.1 Library <i>Python</i>	16
2.2 Pelabelan Kelas <i>Wordnet</i>	20
3.1 Jumlah Kemunculan Data Pada Setiap <i>Rating</i>	28
3.2 <i>Fold 1 Cross-Validation</i>	37
3.3 <i>Fold 10 Cross-Validation</i>	38
3.4 Hasil <i>F1-Score</i> pada Setiap <i>Fold</i>	38
3.5 Hasil <i>F1-Score</i> Kedua Metode	41

DAFTAR GAMBAR

Gambar	Halaman
2.1 Alur Penelitian.....	17
2.2 Penggunaan <i>Deep-translator</i>	20
2.3 Penggunaan <i>TextBlob</i> dan Klasifikasi <i>Wordnet</i>	20
2.4 <i>Confusion Matrix</i>	21
2.5 <i>Confusion Matrix Multiclass</i> Pada Kelas 1 <i>Star</i>	21
2.6 <i>K-fold Cross-Validation</i> K=10	26
3.1 Hasil Pengambilan Data Ulasan SIREKAP 2024	27
3.2 Hasil Input Data.....	27
3.3 Distribusi <i>Rating</i>	28
3.4 <i>Wordcloud Rating</i> 1.....	29
3.5 <i>Wordcloud Rating</i> 5.....	29
3.6 Hasil Input Id.....	30
3.7 Hasil <i>Lower Case</i>	30
3.8 Hasil <i>Remove Unnecessary Character</i>	31
3.9 Kamus Tidak Baku.....	31
3.10 Hasil <i>Spell Checker</i>	32
3.11 Hasil <i>Stemming</i>	32
3.12 Hasil Hapus Data Kosong	33
3.13 Hasil <i>Pra-proses</i>	33
3.14 Hasil <i>Deep-translator</i>	34
3.15 Hasil <i>Wordnet</i>	34
3.16 <i>Confusion Matrix</i> dengan <i>Wordnet</i>	35
3.17 komentar <i>_stemming</i>	36
3.18 <i>Rating</i>	36
3.19 Hasil <i>TF-IDF</i>	37
3.20 <i>F1-Score</i> 10 <i>Fold</i>	39
3.21 <i>Confusion Matrix Overall</i> dengan <i>Decision Tree</i> CART	39
3.22 Komparasi Hasil <i>F1-Score</i>	40

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Proses Pengambilan Data.....	45
Lampiran 2 Memasukan Data Ke <i>Dataframe</i> Pandas	45
Lampiran 3 Filtering dan Sorting Kolom	45
Lampiran 4 Ekspor Data Ke File CSV	45
Lampiran 5 Input File CSV	45
Lampiran 6 Proses Distribusi <i>Rating</i>	46
Lampiran 7 Proses <i>Wordcloud</i>	46
Lampiran 8 Proses Pemberian ID Disetiap Data	47
Lampiran 9 Proses <i>Lower Case</i>	47
Lampiran 10 Proses <i>Remove Unnecessary Character</i>	47
Lampiran 11 Source Code Kamus "4 lebih list kata KBBI.txt"	48
Lampiran 12 Cek Kata Tidak Baku.....	48
Lampiran 13 Cek List Kata Tidak Baku.....	49
Lampiran 14 Proses Input "kamus_tidak_baku"	49
Lampiran 15 Proses <i>Spell Checker</i>	49
Lampiran 16 Proses <i>Stemming</i>	49
Lampiran 17 Cek Data Kosong	50
Lampiran 18 Hapus Data Kosong	50
Lampiran 19 Proses Penerjemahan Data Ulasan ke dalam Bahasa Inggris.....	50
Lampiran 20 Fungsi Hitung Score.....	51
Lampiran 21 Proses <i>Wordnet</i>	51
Lampiran 22 Proses Evaluasi <i>F1-Score Wordnet</i>	52
Lampiran 23 Proses Penentuan Input dan Kelas	53
Lampiran 24 Proses Ekstraksi Fitur TF-IDF	53
Lampiran 25 Proses <i>Cross-Validation 10 Fold</i>	55
Lampiran 26 Menampilkan Jumlah Data Setiap Pembagian <i>Fold</i>	55
Lampiran 27 Proses Input Model Klasifikasi <i>Decision Tree</i> CART	55
Lampiran 28 Proses Visualisasi Model dengan Menjalankan <i>Cross Validation</i>	55
Lampiran 29 Surat Keterangan Pengambilan Data Sekunder	56
Lampiran 30 Kartu Kendali Bimbingan Laporan Karya Ilmiah.....	57