

**ANALISIS SENTIMEN KEPUASAN PENGGUNA APLIKASI
IDENTITAS KEPENDUDUKAN DIGITAL MENGGUNAKAN
ALGORITMA SUPPORT VECTOR MACHINE**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

MA'MUR ZAKY NURROKHMAN

20.11.3791

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS SENTIMEN KEPUASAN PENGGUNA APLIKASI
IDENTITAS KEPENDUDUKAN DIGITAL MENGGUNAKAN
ALGORITMA SUPPORT VECTOR MACHINE**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh
MA'MUR ZAKY NURROKHMAN
20.11.3791

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN KEPUASAN PENGGUNA APLIKASI IDENTITAS
KEPENDUDUKAN DIGITAL MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Ma'mur Zaky Nurrokhman

20.11.3791

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 12 Februari 2025

Dosen Pembimbing,



Yuli Astuti, M. Kom.

NIK. 190302146

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN KEPUASAN PENGGUNA APLIKASI IDENTITAS
KEPENDUDUKAN DIGITAL MENGGUNAKAN ALGORITMA
SUPPORT VECTOR MACHINE**

yang disusun dan diajukan oleh

Ma'mur Zaky Nurrokhman

20.11.3791

Telah dipertahankan di depan Dewan Penguji
pada tanggal 12 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Windha Mega Pradnya Dhuhiya, M.Kom.
NIK. 190302185

Ike Verawati, M.Kom.
NIK. 190302237

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393



Skrripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 12 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ma'mur Zaky Nurrokhman
NIM : 20.11.3791

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Kepuasan Pengguna Aplikasi Identitas Kependudukan Digital Menggunakan Algoritma Support Vector Machine

Dosen Pembimbing : Yuli Astuti, M.Kom.

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, 12 Februari 2025

Yang Menyatakan,



Ma'mur Zaky Nurrokhman

HALAMAN PERSEMBAHAN

Puji syukur ke hadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis bisa menyelesaikan skripsi yang berjudul *Analisis Sentimen Kepuasan Pengguna Aplikasi Identitas Kependudukan Digital Menggunakan Algoritma Support Vector Machine*. Dengan penuh rasa syukur yang mendalam, penulis persembahkan skripsi ini kepada:

1. Kedua orang tua tercinta yang selalu mendoakan serta memberikan kasih sayang tanpa henti kepada penulis dalam menyelesaikan studi
2. Kakak-kakak dan saudara-saudara penulis yang senantiasa memberikan dukungan moral dan kebahagiaan
3. Dosen pembimbing, bu Yuli Astuti, M.Kom., atas bimbingan, masukan, dan arahan serta kesabarannya selama proses penyelesaian skripsi ini
4. Teman-teman penulis baik itu teman seangkatan di sekolah, teman-teman dari Universitas Amikom Yogyakarta terutama dari kelas IF 08, maupun teman-teman dari kampus lain yang senantiasa memberikan motivasi, berbagi kebahagiaan, dan canda tawa serta mendampingi penulis selama masa-masa sulit
5. Almamater tercinta, Universitas Amikom Yogyakarta, yang menjadi tempat belajar, bertumbuh, dan berkembang menjadi pribadi yang lebih baik
6. Diri sendiri yang telah melewati berbagai tantangan, ujian, dan kesulitan sehingga mampu menyelesaikan perjalanan yang panjang ini

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas segala limpahan rahmat dan hidayah-Nya sehingga penulis bisa menyelesaikan skripsi yang berjudul *Analisis Sentimen Kepuasan Pengguna Aplikasi Identitas Kependudukan Digital Menggunakan Algoritma Support Vector Machine*. Skripsi ini disusun sebagai bagian dari tanggung jawab penulis sebagai mahasiswa informatika di Universitas Amikom Yogyakarta dan ditujukan sebagai tugas akhir serta syarat kelulusan dari program studi informatika.

Skripsi ini dapat diselesaikan tentunya tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Segala bentuk bantuan dan dukungan sangat membantu penulis untuk menyelesaikan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih yang tulus kepada pihak-pihak yang telah membantu, mendukung, dan membimbing penulis selama menyusun skripsi ini, yakni kepada:

1. Prof. Dr. M. Suyanto, MM., selaku Rektor Universitas Amikom Yogyakarta
2. Hanif Al Fatta, S.Kom., M.Kom., Ph.D., selaku Dekan Fakultas Ilmu Komputer
3. Windha Mega Pradnya Dhuhiita, M.Kom., selaku ketua program studi Informatika
4. Yuli Astuti, M.Kom., selaku dosen pembimbing skripsi yang senantiasa membimbing penulis selama penyusunan skripsi
5. Rumini, M.Kom., selaku dosen wali
6. Tim dosen penguji yang telah memberikan umpan balik berupa saran dan kritik yang membangun
7. Pihak Universitas Amikom Yogyakarta yang selalu memberikan dukungan
8. Rekan-rekan mahasiswa yang selalu memberikan semangat
9. Kedua orang tua penulis yang selalu mendoakan, memberi dukungan, dan semangat untuk menyelesaikan skripsi

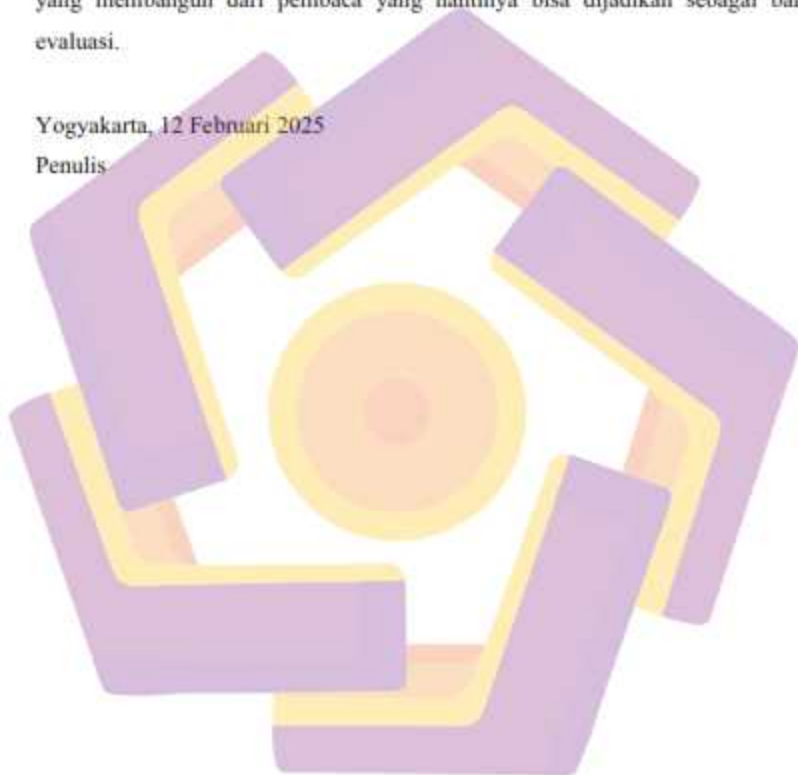
Semoga Allah SWT memberikan pahala dan berkah yang berlimpah atas segala bentuk bantuan dan dukungan yang telah diberikan kepada penulis. Penulis

juga berharap agar skripsi ini dapat bermanfaat bagi para pembaca dari berbagai kalangan khususnya para mahasiswa di Indonesia.

Penulis tentu menyadari bahwa masih terdapat banyak kesalahan dalam penyusunan skripsi ini, baik dari segi EBI, kosa kata, tata bahasa, etika maupun isi. Maka dari itu, penulis sangat mengharapkan umpan balik berupa kritik atau saran yang membangun dari pembaca yang nantinya bisa dijadikan sebagai bahan evaluasi.

Yogyakarta, 12 Februari 2025

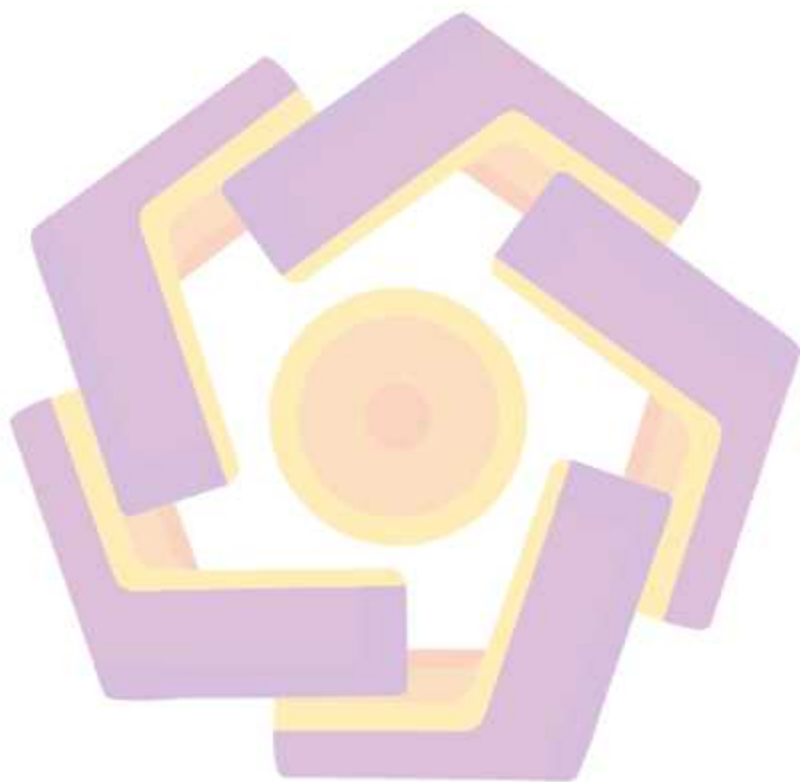
Penulis



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
DAFTAR LAMBANG DAN SINGKATAN	xiv
DAFTAR ISTILAH	xv
INTISARI	xix
ABSTRACT	xx
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	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	23

2.2.1	Analisis Sentimen	23
2.2.2	Identitas Kependudukan Digital	23
2.2.3	Support Vector Machine	23
2.2.4	Term Frequency-Inverse Document Frequency	25
2.2.5	Synthetic Minority Oversampling Technique	26
2.2.6	Bayesian Optimization	26
2.2.7	Word Cloud	27
2.2.8	Confusion Matrix	27
BAB III METODE PENELITIAN		30
3.1	Objek Penelitian	30
3.2	Alur Penelitian	30
3.2.1	Persiapan Data	32
3.2.2	Preprocessing	32
3.2.3	Pelatihan Model	34
3.2.4	Evaluasi	35
3.3	Alat dan Bahan	35
BAB IV HASIL DAN PEMBAHASAN		38
4.1	Persiapan Data	38
4.2	Preprocessing	38
4.3	Pelatihan Model	45
4.4	Evaluasi	45
BAB V PENUTUP		51
5.1	Kesimpulan	51
5.2	Saran	51
REFERENSI		53

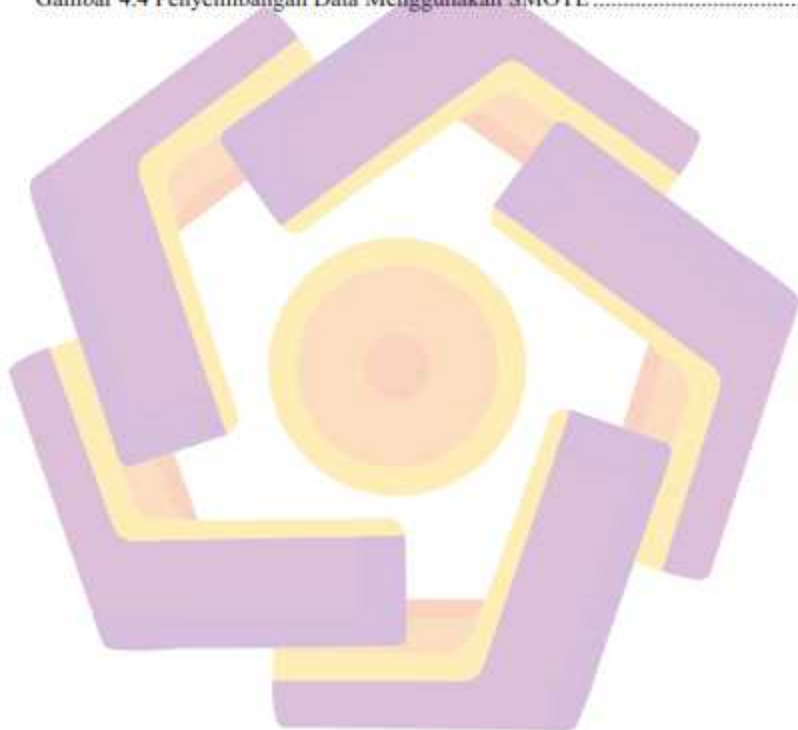


DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	7
Tabel 2.2 Struktur Confusion Matrix	28
Tabel 4.1 Hasil Case Folding	39
Tabel 4.2 Hasil Tokenization	40
Tabel 4.3 Hasil Penghapusan Stop Words	40
Tabel 4.4 Hasil Stemming	40
Tabel 4.5 Hasil Perbaikan Ejaan Kata	41
Tabel 4.6 Hasil Penghapusan Stop Words dan Stemming Ulang	41
Tabel 4.7 Hasil TF-IDF	44
Tabel 4.8 Kombinasi Metode Pelatihan Model	45
Tabel 4.9 Hyperparameter Terbaik	45
Tabel 4.10 Confusion Matrix	46
Tabel 4.11 Metrik Accuracy, Precision, Recall, dan F1-score	47

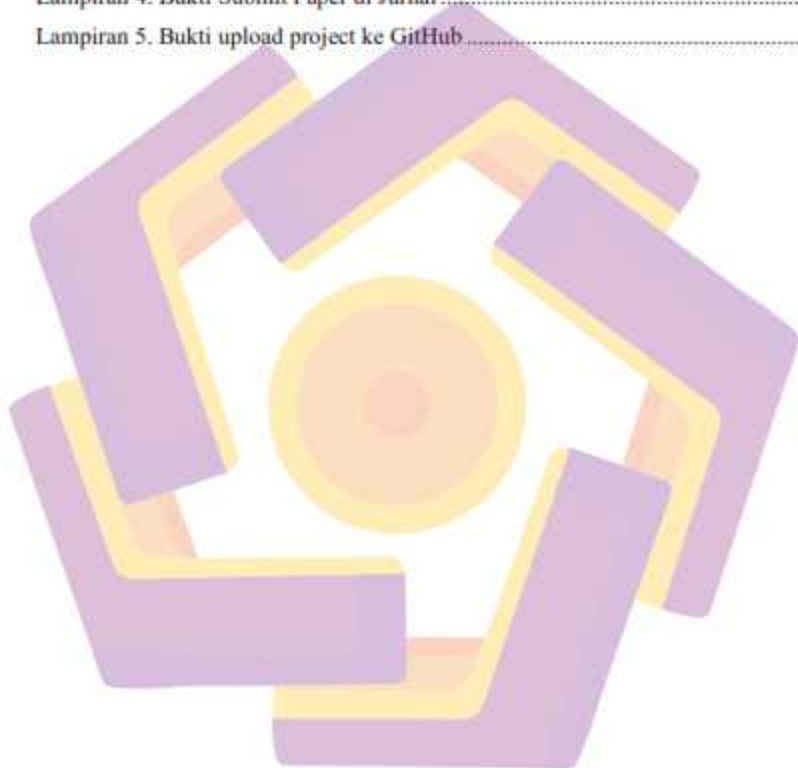
DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	31
Gambar 4.1 Visualisasi Word Cloud	42
Gambar 4.2 Visualisasi Word Cloud pada Kelas Sentimen Positif	42
Gambar 4.3 Visualisasi Word Cloud pada Kelas Sentimen Negatif	43
Gambar 4.4 Penyeimbangan Data Menggunakan SMOTE	44

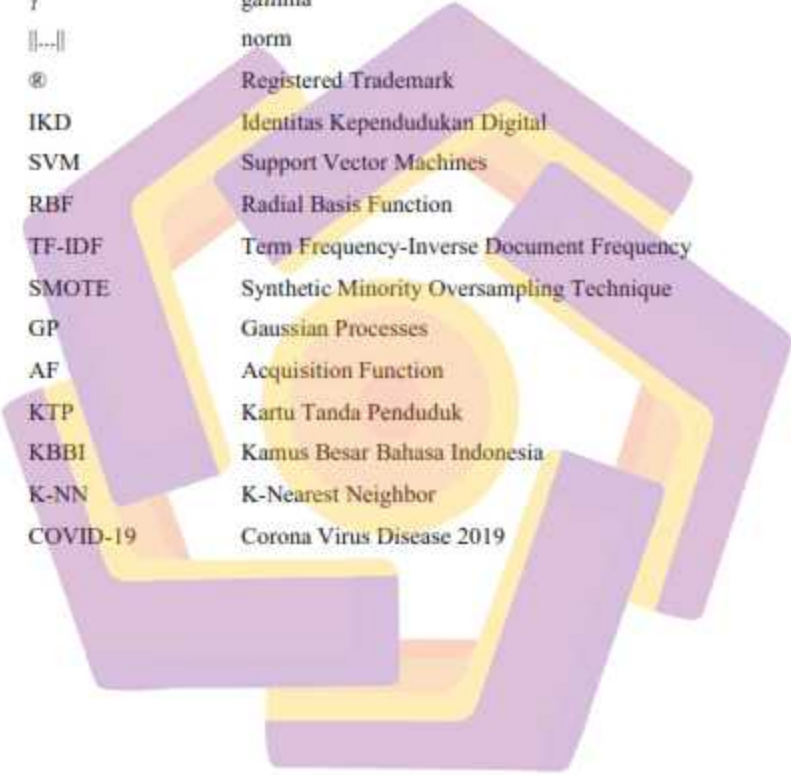


DAFTAR LAMPIRAN

Lampiran 1. Kode Program Python di Jupyter Notebook	59
Lampiran 2. Dataset Ulasan Aplikasi IKD di Play Store	72
Lampiran 3. Dataset Ulasan Aplikasi IKD di App Store	74
Lampiran 4. Bukti Submit Paper di Jurnal	75
Lampiran 5. Bukti upload project ke GitHub	76



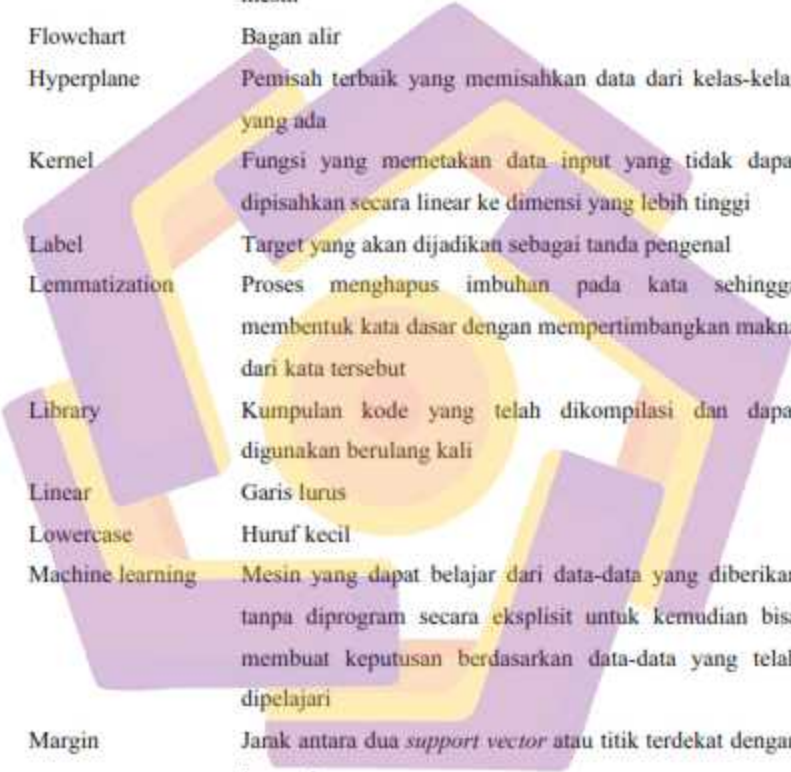
DAFTAR LAMBANG DAN SINGKATAN



+	tambah
-	kurang
>	lebih besar dari
γ	gamma
$\ \dots \ $	norm
®	Registered Trademark
IKD	Identitas Kependudukan Digital
SVM	Support Vector Machines
RBF	Radial Basis Function
TF-IDF	Term Frequency-Inverse Document Frequency
SMOTE	Synthetic Minority Oversampling Technique
GP	Gaussian Processes
AF	Acquisition Function
KTP	Kartu Tanda Penduduk
KBBI	Kamus Besar Bahasa Indonesia
K-NN	K-Nearest Neighbor
COVID-19	Corona Virus Disease 2019

DAFTAR ISTILAH

Accuracy	Akurasi model dalam memprediksi atau mengklasifikasikan data
Algoritma	Prosedur sistematis untuk memecahkan masalah matematis dalam langkah-langkah terbatas
Aplikasi mobile	Aplikasi yang digunakan di <i>smartphone</i>
Atribut	Kolom-kolom yang terdapat pada dataset
Case folding	Teknik mengubah huruf-huruf pada teks menjadi huruf kecil atau <i>lowercase</i> dan menghilangkan tanda baca serta <i>whitespace</i> (karakter kosong)
Confusion matrix	Metode untuk mengevaluasi model <i>machine learning</i> yang berbasis klasifikasi yang menghasilkan tabel dengan nilai True Positive, True Negative, False Positive, dan False Negative
COVID-19	Penyakit pernapasan yang disebabkan oleh virus corona, ditandai oleh demam dan batuk yang dapat memburuk menjadi pneumonia, gagal ginjal, gagal napas, koagulasi darah abnormal, bahkan kematian, kondisi akut dapat terjadi terutama pada lansia atau pasien dengan penyakit bawaan
Dataframe	Tabel/data tabular dengan <i>array</i> dua dimensi (baris dan kolom)
Dataset	Kumpulan data yang sudah terorganisasi, biasanya dalam format tabel atau <i>database</i> yang memiliki struktur baris dan kolom
Digital	Sesuatu yang berkaitan dengan atau menggunakan komputer atau internet
Digitalisasi	Proses pemberian atau pemakaian sistem digital
Document	Sebuah istilah pada metode TF-IDF yang menunjukkan satu baris data teks

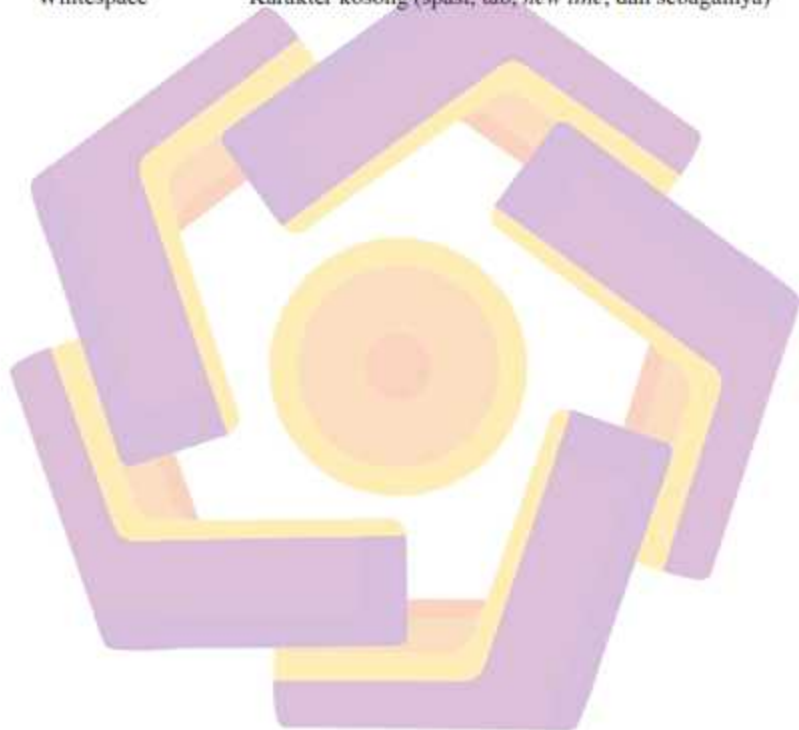


Emoji	Emotikon yang berasal dari Jepang
F1-score	Nilai rata-rata harmonik dari nilai <i>precision</i> dan <i>recall</i>
Feature selection	Proses seleksi fitur-fitur atau variabel-variabel yang relevan
Fitur/Feature	Variabel acuan yang akan dihitung/dikomputasi oleh mesin
Flowchart	Bagan alir
Hyperplane	Pemisah terbaik yang memisahkan data dari kelas-kelas yang ada
Kernel	Fungsi yang memetakan data input yang tidak dapat dipisahkan secara linear ke dimensi yang lebih tinggi
Label	Target yang akan dijadikan sebagai tanda pengenal
Lemmatization	Proses menghapus imbuhan pada kata sehingga membentuk kata dasar dengan mempertimbangkan makna dari kata tersebut
Library	Kumpulan kode yang telah dikompilasi dan dapat digunakan berulang kali
Linear	Garis lurus
Lowercase	Huruf kecil
Machine learning	Mesin yang dapat belajar dari data-data yang diberikan tanpa diprogram secara eksplisit untuk kemudian bisa membuat keputusan berdasarkan data-data yang telah dipelajari
Margin	Jarak antara dua <i>support vector</i> atau titik terdekat dengan <i>hyperplane</i>
Marketplace	Platform <i>online</i> yang mempertemukan antara penjual dan pembeli untuk melakukan transaksi jual beli barang baik barang fisik maupun digital dan jasa
Masif	Menggambarkan sesuatu yang kuat, sangat besar, dan memiliki nilai yang tinggi
Noise	Data dengan nilai yang salah atau rusak sehingga tidak



	memberikan pengaruh yang baik pada model <i>machine learning</i>
Online	Daring, terhubung ke internet
Outlier	Data pencilan (sebutan lain dari <i>noise</i>)
Overfit	Suatu kondisi ketika model machine learning terlalu detail dalam mengenali data
Oversampling	Sebuah teknik yang meningkatkan jumlah data pada kelas minoritas dengan tujuan menyeimbangkan jumlah data di semua kelas yang ada
Precision	Nilai yang menunjukkan persentase perbandingan antara kelas positif yang diprediksi dengan benar dengan jumlah keseluruhan kelas yang diprediksi positif
Preprocessing	Tahapan untuk menghilangkan beberapa permasalahan pada data
Rating	Peringkat
Recall	Nilai yang menunjukkan persentase perbandingan antara kelas positif yang diprediksi dengan benar dengan jumlah keseluruhan data kelas positif yang sebenarnya
Scraping	Proses mengambil data dari website
Sistem operasi	Perangkat lunak yang berfungsi untuk mengelola sumber daya perangkat keras (<i>hardware</i>) dan perangkat lunak (<i>software</i>) pada komputer
Smartphone	Telepon genggam yang memiliki fungsi dan fitur canggih
Stemming	Proses menghapus imbuhan pada kata sehingga membentuk kata dasar
Stop words	Kata-kata yang diabaikan karena tidak memiliki makna yang signifikan
Supervised learning	Jenis algoritma yang membutuhkan arahan berupa label atau kelas pada data agar model bisa mempelajari data yang diberikan
Support vector	Titik yang paling dekat dengan <i>hyperplane</i>

Term	Kata atau istilah yang terdapat pada <i>document</i> atau baris data teks
Text mining	Kegiatan menambang data teks
User interface	Tampilan antarmuka pada aplikasi
User experience	Pengalaman pengguna dalam mengakses atau menggunakan sebuah aplikasi
Whitespace	Karakter kosong (<i>spasi</i> , <i>tab</i> , <i>new line</i> , dan sebagainya)



INTISARI

Aplikasi Identitas Kependudukan Digital (IKD) merupakan sebuah aplikasi *mobile* yang dapat diakses menggunakan *smartphone* dengan sistem operasi Android dan iOS. Aplikasi ini berfungsi untuk mendigitalisasi dokumen kependudukan di Indonesia sehingga mempercepat akses terhadap data kependudukan. Namun, rating dan ulasan dari aplikasi ini cukup buruk yang menunjukkan bahwa aplikasi masih terdapat banyak kekurangan yang harus diperbaiki. Oleh karena itu, pada penelitian ini dilakukan analisis sentimen mengenai kepuasan pengguna aplikasi IKD. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat mengenai kepuasan penggunaan aplikasi IKD dengan menghasilkan model *machine learning*. Data yang digunakan berupa ulasan aplikasi IKD yang diambil dari Play Store dan App Store. Pelatihan model menggunakan empat kombinasi metode yang berbeda yang melibatkan teknik penyeimbangan data dengan metode SMOTE dan *hyperparameter tuning* dengan metode *bayesian optimization*. Hasil penelitian menunjukkan bahwa model yang menerapkan metode penyeimbangan data dan *hyperparameter tuning* menjadikan model tersebut sebagai model terbaik yang menghasilkan performa terbaik dibandingkan model-model yang tidak menerapkan kedua metode tersebut. Model terbaik menghasilkan akurasi sebesar 91,40% dengan *f1-score* pada kelas positif dan kelas negatif yang seimbang serta mendekati angka 1 yang menunjukkan bahwa model dapat mengenali kedua kelas sentimen dengan baik dan seimbang. Sedangkan model yang tidak menerapkan kedua metode tersebut menghasilkan performa paling buruk dengan akurasi 86,55% dan *f1-score* pada kelas positif dan kelas negatif yang tidak seimbang serta kurang mendekati angka 1 pada kelas positif. Dengan penelitian ini diharapkan dapat memberikan manfaat bagi pengembang aplikasi IKD maupun pemerintah sebagai bahan evaluasi untuk meningkatkan kualitas pelayanan publik.

Kata kunci: Analisis Sentimen, Identitas Kependudukan Digital, Kepuasan Pengguna, Support Vector Machine, Text Mining.

ABSTRACT

The *Identitas Kependudukan Digital (IKD)* application is a mobile application that can be accessed using smartphones with Android and iOS operating systems. This application digitizes population documents in Indonesia to accelerate access to population data. However, the ratings and reviews of this application are quite bad which shows that the application still has many shortcomings that must be improved. Therefore, in this study, a sentiment analysis was conducted regarding the satisfaction of IKD application users. This study aims to analyze public sentiment regarding the satisfaction of using the IKD application by generating a machine learning model. The data used are IKD application reviews taken from the Play Store and App Store. Model training uses four different combinations of methods involving data balancing techniques with the SMOTE method and hyperparameter tuning with the Bayesian optimization method. The results show that the model that applies the data balancing and hyperparameter tuning methods makes the model the best model that produces the best performance than models that do not apply these two methods. The best model produces an accuracy of 91.40% with f1-score on positive and negative classes that are balanced and close to 1, indicating that the model can recognize both sentiment classes well and balanced. Meanwhile, the model that does not apply the two methods produces the worst performance with an accuracy of 86.55% and f1-score on positive and negative classes that are not balanced and less close to 1 on the positive class. This research is expected to provide benefits for IKD application developers and the government as evaluation material to improve the quality of public services.

Keyword: *Identitas Kependudukan Digital, Sentiment Analysis, Support Vector Machine, Text Mining, User Satisfaction.*