

**PERBANDINGAN ALGORITMA INDOBERT, NAIVE BAYES, SVM,  
DAN KNN TERHADAP ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI M-PASPOR DI GOOGLE PLAY STORE**

**SKRIPSI**

Program Studi Sistem Informasi (Bilingual)

Jenjang Sarjana



Oleh:

**Sindu Prakoso**

**09031382126160**

**PROGRAM STUDI SISTEM INFORMASI BILINGUAL**

**JURUSAN SISTEM INFORMASI**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS SRIWIJAYA**

**2024**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**PERBANDINGAN ALGORITMA INDOBERT, NAIVE BAYES, SVM, DAN  
KNN TERHADAP ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI M-PASPOR DI GOOGLE PLAY STORE**



Sebagai salah satu syarat untuk menyelesaikan studi  
di Program Studi Sistem Informasi S1

Oleh :

**Sindu Prakoso      09031382126160**

**Palembang, 27 Desember 2024**

Mengetahui,  
Ketua Jurusan Sistem Informasi,



**Ahmad Rifai, M.T.**  
NIP. 197910202010121003

Pembimbing,



**Ken Ditha Tania, M.Kom., Ph.D.**  
NIP. 198507182012122003

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

NIM : 09031382126160  
Nama : Sindu Prakoso  
Prodi/Jurusan : Sistem Informasi (Bilingual)  
Judul Skripsi : Perbandingan Algoritma IndoBERT, Naïve Bayes, SVM,  
dan KNN Terhadap Analisis Sentimen Ulasan Pengguna  
Aplikasi M-Paspor di Google Play Store

Hasil Pengecekan Turnitin : 6%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya saya sendiri dan bukan penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi saya ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, 27 Desember 2024






**Sindu Prakoso**  
NIM. 09031382126160

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Jumat  
Tanggal : 27 Desember 2024  
Nama : Sindu Prakoso  
NIM : 09031382126160  
Judul : Perbandingan Algoritma IndoBERT, Naïve Bayes, SVM,  
dan KNN Terhadap Analisis Sentimen Ulasan Pengguna  
Aplikasi M-Paspor di Google Play Store

Komisi Penguji :

1. Ketua : Apriansyah Putra, M.Kom. : 
2. Pembimbing : Ken Ditha Tania, M.Kom., Ph.D. : 
3. Penguji : Mira Afrina, M.Sc., Ph.D. : 

Mengetahui  
Ketua Jurusan Sistem Informasi

  
  
Ahmad Rifai, M.T.  
NIP. 197910202010121003

## HALAMAN PERSEMBAHAN

### **Motto:**

“Kehidupan yang sebenarnya adalah hari ini. Jangan menyibukkan diri dengan mengenang masa lalu dan mencemaskan masa depan yang tidak kita ketahui.”

Skripsi ini dipersembahkan kepada:

- ▶ Kedua orangtua, saudara, dan keluarga besar
- ▶ Seluruh teman dan sahabat
- ▶ Seluruh pihak yang membantu selama ini
- ▶ Universitas Sriwijaya

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa karena atas rahmat dan karunia-Nya penulis dapat menyusun dan menyelesaikan tugas akhir yang berjudul “Perbandingan Algoritma IndoBERT, Naïve Bayes, SVM, dan KNN Terhadap Analisis Sentimen Ulasan Pengguna Aplikasi M-Paspor di Google Play Store”. Tugas akhir ini merupakan salah satu syarat yang harus dipenuhi Penulis untuk memperoleh gelar sarjana di Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Sriwijaya.

Selama menyelesaikan tugas akhir ini, bimbingan, arahan, dukungan, bimbingan dan kesempatan, bantuan, dukungan dan doa dari berbagai pihak yang berdedikasi sehingga dapat memperlancar pembuatan tugas akhir ini. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

1. Allah SWT yang telah memberikan rahmat, ilmu, kesempatan dan kesehatan sehingga penulis mampu melaksanakan dan menyelesaikan Tugas Akhir ini.
2. Orang tua dan saudara yang selalu senantiasa memberikan doa dan dukungan, serta saran yang sangat bermanfaat bagi penulis untuk terus melakukan yang terbaik.
3. Bapak Prof. Dr. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Ahmad Rifai, M.T. selaku Ketua Jurusan Sistem Informasi.
5. Ibu Ken Ditha Tania, M.Kom., Ph.D. selaku Dosen Pembimbing Tugas Akhir yang selalu memberikan Bimbingan, saran dan arahan dalam penulisan dan pengerjaan

serta pelaksanaan tugas akhir, hingga penulis dapat membuat Tugas Akhir ini dengan baik.

6. Mba Rifka, selaku Admin Program Studi Sistem Informasi Bilingual yang telah menjadi sumber informasi dan membantu seluruh proses pemberkasan selama mengerjakan Tugas Akhir ini.
7. Sahabat seperjuangan selama perkuliahan dan Terima kasih banyak selama perkuliahan ini sudah mau berbagi informasi penting maupun tidak penting baik itu tentang perkuliahan maupun non-perkuliahan. Dan terima kasih banyak juga telah banyak membantu dan saling support mulai dari saat kerja praktik sampai mengerjakan Tugas Akhir ini yaitu Ahmad Syakir Muliana, Muhammad Ali Lubis, Muhammad Hanura Nooryadi, dan Mutsaqqaful Fadhiil Yusrizal.
8. Terima kasih banyak juga untuk semua teman-teman di kelas SIBIL B angkatan 21 yang tidak bisa disebutkan namanya satu persatu terima kasih banyak untuk semuanya atas segala kebaikannya dan semua hal yang tidak bisa diungkapkan dan disebutkan satu persatu .

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari kesempurnaan, apabila terdapat kekeliruan dalam penulisan laporan ini penulis sangat mengharapkan masukan, kritik dan sarannya sehingga penulis bisa membuat Tugas Akhir ini menjadi lebih baik lagi.

Akhir kata dengan segala keterbatasan, penulis mengharapkan semoga Tugas Akhir ini Penulis dapat bermanfaat bagi seluruh pihak terutama mahasiswa secara langsung ataupun tidak langsung sebagai peningkatan mutu pembelajaran.

Palembang 27 Desember 2024

A handwritten signature in black ink, appearing to be 'Sindu Prakoso', written in a cursive style.

Sindu Prakoso  
NIM. 09031382126160



**PERBANDINGAN ALGORITMA INDOBERT, NAIVE BAYES, SVM,  
DAN KNN TERHADAP ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI M-PASPOR DI GOOGLE PLAY STORE**

Oleh:

**Sindu Prakoso      09031382126160**

**ABSTRAK**

Perkembangan teknologi informasi yang sangat cepat seiring dengan kemajuan zaman membuat semua aspek kehidupan beralih ke digital. Aplikasi M-Paspor merupakan aplikasi yang dapat digunakan oleh masyarakat untuk pengajuan permohonan paspor baru dan penggantian paspor secara online. Aplikasi M-Paspor memiliki beragam tanggapan, baik positif, negatif dan netral, dari pengguna yang tercantum dalam ulasan di Google Play Store. Dengan banyaknya data ulasan, diperlukan analisis sentimen untuk memahami pola sentimen secara efisien. Penelitian ini membandingkan kinerja algoritma *Naïve Bayes*, SVM, KNN, dan IndoBERT dalam analisis sentimen terhadap ulasan aplikasi M-Paspor. Hasil penelitian menunjukkan bahwa algoritma IndoBERT unggul dengan akurasi 99%, diikuti oleh SVM dengan akurasi 98%, sementara *Naïve Bayes* dan KNN masing-masing memiliki akurasi 91% dan 67%.

**Kata Kunci :** Analisis Sentimen, M-Paspor, IndoBERT, *Naïve Bayes*, SVM, KNN

**COMPARISON OF INDOBERT, NAIVE BAYES, SVM, AND KNN  
ALGORITHMS ON SENTIMENT ANALYSIS OF USER REVIEWS OF M-  
PASPOR APPLICATION ON GOOGLE PLAY STORE**

Oleh:

**Sindu Prakoso      09031382126160**

**ABSTRACT**

The development of information technology is very fast along with the progress of the times making all aspects of life switch to digital. The M-Passport application is an application that can be used by the public to apply for new passports and passport replacements online. The M-Passport application has various responses, both positive, negative and neutral, from users listed in reviews on the Google Play Store. With the large amount of review data, sentiment analysis is needed to understand sentiment patterns efficiently. This study compares the performance of the Naïve Bayes, SVM, KNN, and IndoBERT algorithms in sentiment analysis of M-Passport application reviews.. The results show that the IndoBERT algorithm excels with 99% accuracy, followed by SVM with 98% accuracy, while Naïve Bayes and KNN have 91% and 67% accuracy respectively.

**Keywords :** *Sentiment Analysis*, M-Paspor, IndoBERT, *Naïve Bayes*, SVM, KNN

## DAFTAR ISI

HALAMAN PENGESAHAN.....	i
HALAMAN PERNYATAAN .....	ii
HALAMAN PERSETUJUAN.....	ii
HALAMAN PERSEMBAHAN .....	iv
KATA PENGANTAR .....	v
ABSTRAK .....	viii
ABSTRACT.....	ix
DAFTAR ISI .....	x
DAFTAR GAMBAR .....	xiii
DAFTAR TABEL.....	xiv
DAFTAR LAMPIRAN .....	xv
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah .....	4
1.5 Manfaat Penelitian.....	4
BAB II TINJAUAN PUSTAKA.....	<b>Error! Bookmark not defined.</b>
2.1 Analisis Sentimen .....	<b>Error! Bookmark not defined.</b>
2.2 M-Paspor .....	<b>Error! Bookmark not defined.</b>
2.3 Google Play Store .....	<b>Error! Bookmark not defined.</b>
2.4 IndoBERT.....	<b>Error! Bookmark not defined.</b>
2.5 <i>Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
2.6 <i>Support Vector Machine</i> .....	<b>Error! Bookmark not defined.</b>
2.7 <i>K – Nearest Neighbor</i> .....	<b>Error! Bookmark not defined.</b>
2.8 <i>Confusion Matrix</i> .....	<b>Error! Bookmark not defined.</b>
2.9 <i>Receiver Operating Characteristic</i> .....	<b>Error! Bookmark not defined.</b>
BAB III METODOLOGI PENELITIAN .....	<b>Error! Bookmark not defined.</b>

3.1 Pengambilan Data.....	<b>Error! Bookmark not defined.</b>
3.2 Pelabelan Data .....	<b>Error! Bookmark not defined.</b>
3.3 <i>Preprocessing</i> .....	<b>Error! Bookmark not defined.</b>
3.4 <i>Balancing Data</i> .....	<b>Error! Bookmark not defined.</b>
3.5 <i>Feature Extraction</i> .....	<b>Error! Bookmark not defined.</b>
3.6 <i>Split Data</i> .....	<b>Error! Bookmark not defined.</b>
3.7 Pemodelan .....	<b>Error! Bookmark not defined.</b>
3.8 Evaluasi .....	<b>Error! Bookmark not defined.</b>
BAB IV HASIL DAN PEMBAHASAN .....	<b>Error! Bookmark not defined.</b>
4.1 Pengambilan Data.....	<b>Error! Bookmark not defined.</b>
4.2 Pelabelan Data .....	<b>Error! Bookmark not defined.</b>
4.3 <i>Preprocessing</i> .....	<b>Error! Bookmark not defined.</b>
4.3.1 <i>Cleaning</i> .....	<b>Error! Bookmark not defined.</b>
4.3.2 <i>Case Folding</i> .....	<b>Error! Bookmark not defined.</b>
4.3.3 <i>Tokenizing</i> .....	<b>Error! Bookmark not defined.</b>
4.3.4 <i>Stopwords</i> .....	<b>Error! Bookmark not defined.</b>
4.3.5 <i>Stemming</i> .....	<b>Error! Bookmark not defined.</b>
4.4 <i>Balancing Data</i> .....	<b>Error! Bookmark not defined.</b>
4.5 <i>Feature Extraction</i> .....	<b>Error! Bookmark not defined.</b>
4.6 <i>Split Data</i> .....	<b>Error! Bookmark not defined.</b>
4.7 Pemodelan .....	<b>Error! Bookmark not defined.</b>
4.7.1 Pemodelan IndoBERT .....	<b>Error! Bookmark not defined.</b>
4.7.2 Pemodelan <i>Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
4.7.3 Pemodelan SVM .....	<b>Error! Bookmark not defined.</b>
4.7.4 Pemodelan KNN .....	<b>Error! Bookmark not defined.</b>
4.8 Evaluasi .....	<b>Error! Bookmark not defined.</b>
4.8.1 Evaluasi IndoBERT .....	<b>Error! Bookmark not defined.</b>
4.8.2 Evaluasi <i>Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
4.8.3 Evaluasi SVM .....	<b>Error! Bookmark not defined.</b>
4.8.4 Evaluasi KNN .....	<b>Error! Bookmark not defined.</b>
4.8.5 Evaluasi Perbandingan .....	<b>Error! Bookmark not defined.</b>

BAB V KESIMPULAN DAN SARAN.....	<b>Error! Bookmark not defined.</b>
5.1 Kesimpulan.....	<b>Error! Bookmark not defined.</b>
5.2 Saran .....	<b>Error! Bookmark not defined.</b>
DAFTAR PUSTAKA .....	35

## DAFTAR GAMBAR

Gambar 2.1 Kurva ROC-AUC .....	<b>Error! Bookmark not defined.</b>
Gambar 3.1 Tahapan Penelitian (Asnawi et al., 2021).....	<b>Error! Bookmark not defined.</b>
Gambar 4.1 Contoh Hasil Pengambilan Data.....	<b>Error! Bookmark not defined.</b>
Gambar 4.2 Distribusi Label Sentimen .....	<b>Error! Bookmark not defined.</b>
Gambar 4.3 Hasil Balancing Data .....	<b>Error! Bookmark not defined.</b>
Gambar 4.4 Hasil <i>Split Data</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.5 Kode Proses Pemodelan IndoBERT (1) .....	<b>Error! Bookmark not defined.</b>
Gambar 4.6 Kode Proses Pemodelan IndoBERT (2) .....	<b>Error! Bookmark not defined.</b>
Gambar 4.7 Kode Proses Pemodelan IndoBERT (3) .....	<b>Error! Bookmark not defined.</b>
Gambar 4.8 Kode Proses Pemodelan IndoBERT (4) .....	<b>Error! Bookmark not defined.</b>
Gambar 4.9 Kode Proses Pemodelan IndoBERT (5) .....	<b>Error! Bookmark not defined.</b>
Gambar 4.10 Kode Proses Pemodelan <i>Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.11 Kode Proses Pemodelan SVM.....	<b>Error! Bookmark not defined.</b>
Gambar 4.12 Kode Proses Pemodelan KNN.....	<b>Error! Bookmark not defined.</b>
Gambar 4.13 Hasil Confusion Matrix IndoBERT.....	<b>Error! Bookmark not defined.</b>
Gambar 4.14 Nilai Kinerja IndoBERT.....	<b>Error! Bookmark not defined.</b>
Gambar 4.15 Kurva ROC-AUC IndoBERT.....	<b>Error! Bookmark not defined.</b>
Gambar 4.16 Hasil <i>Confusion Matrix Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.17 Nilai Kinerja <i>Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.18 Kurva ROC-AUC <i>Naïve Bayes</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.19 Hasil Confusion Matrix SVM.....	<b>Error! Bookmark not defined.</b>
Gambar 4.20 Nilai Kinerja SVM.....	<b>Error! Bookmark not defined.</b>

**Gambar 4.21** Kurva ROC-AUC SVM.....**Error! Bookmark not defined.**  
**Gambar 4.22** Hasil Confusion Matrix KNN.....**Error! Bookmark not defined.**  
**Gambar 4.23** Nilai Kinerja KNN.....**Error! Bookmark not defined.**  
**Gambar 4.24** Kurva ROC-AUC KNN.....**Error! Bookmark not defined.**

### DAFTAR TABEL

**Tabel 2.1** *Confusion Matrix*.....**Error! Bookmark not defined.**  
**Tabel 2.2** Kategori Hasil Klasifikasi Berdasarkan Nilai ROC ... **Error! Bookmark not defined.**  
**Tabel 4.1** Hasil *Cleaning*.....**Error! Bookmark not defined.**  
**Tabel 4.2** Hasil *Case Folding*.....**Error! Bookmark not defined.**  
**Tabel 4.3** Hasil *Tokenizing*.....**Error! Bookmark not defined.**  
**Tabel 4.4** Hasil *Stopwords*.....**Error! Bookmark not defined.**  
**Tabel 4.5** Hasil *Stemming*.....**Error! Bookmark not defined.**  
**Tabel 4.6** Hasil Pembobotan Kata .....**Error! Bookmark not defined.**  
**Tabel 4.7** Perbandingan Kinerja Algoritma.....**Error! Bookmark not defined.**

**DAFTAR LAMPIRAN**

Lampiran 1. Surat Kesediaan Membimbing .....	A-1
Lampiran 2. Surat Keputusan Tugas Akhir.....	B-1
Lampiran 3. Kartu Konsultasi .....	C-1
Lampiran 4. Form Perbaikan.....	D-1



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Perkembangan teknologi informasi di Indonesia telah membantu mewujudkan berbagai aspek yang mendukung kemajuan negara, khususnya di sektor pemerintahan. Penerapan teknologi informasi di sektor ini mencakup distribusi surat secara digital, pengelolaan administrasi, penggunaan website dan aplikasi untuk pembuatan dokumen pribadi, sistem pelaporan pemerintah, media komunikasi untuk menyalurkan aspirasi masyarakat, layanan perizinan online, serta media informasi lokal yang meningkatkan efisiensi dan kualitas pelaksanaan sistem pemerintahan. Selain itu, penerapan teknologi informasi juga membawa manfaat lain, seperti meningkatkan kualitas pelayanan pemerintah kepada masyarakat.

Jika berbicara mengenai pelayanan, pelayanan keimigrasian merupakan salah satu yang perlu diutamakan untuk memberikan pelayanan sebaik-baiknya karena keimigrasian memiliki tugas di bidang perlintasan negara. Pelayanan pembuatan paspor merupakan hal yang sangat dibutuhkan oleh masyarakat, dan pelayanan ini memerlukan perhatian khusus karena berkaitan langsung dengan mobilitas masyarakat yang hendak bepergian ke luar negeri. (Sukmawati et al., 2020).

Pada tanggal 27 Januari 2022 di Jakarta, Menteri hukum dan HAM RI resmi meluncurkan aplikasi M-Paspor sebagai media untuk membantu masyarakat dalam mengajukan permohonan pembuatan paspor. Masyarakat cukup mengunggah scan berkas data yang dibutuhkan ke aplikasi, sehingga tidak perlu datang keimigrasi

untuk sekedar mengirimkan berkas, aplikasi ini telah dilengkapi fitur validasi sehingga dapat memastikan dokumen yang diunggah sudah benar. Tidak hanya itu pengguna juga dapat memantau status permohonan paspor melalui aplikasi (Az-Zahra & Pambudi, 2024).

Aplikasi M-Pasport tersedia untuk pengguna Android melalui platform digital yaitu Google Play Store. Berbagai informasi dan layanan yang diberikan oleh aplikasi M-Paspor menghasilkan beragam umpan balik, baik positif maupun negatif, dari pengguna yang tercantum dalam ulasan di Google Play Store. Ulasan ini dapat digunakan sebagai referensi untuk mengevaluasi aplikasi berdasarkan pengalaman langsung pengguna. Namun, dengan jumlah data yang besar dan beragam, pengambilan keputusan berdasarkan pengalaman pengguna menjadi sulit dan memakan banyak waktu. Oleh karena itu perlu dilakukan analisis sentimen untuk mempermudah dalam memahami ulasan pengguna secara efisien.

Analisis sentimen adalah proses mengekstraksi, memproses, dan memahami informasi sentimen secara otomatis dari data teks yang tidak terstruktur. (Jauhary, 2023). Beberapa pendekatan yang sering digunakan peneliti lain untuk analisis sentimen yaitu IndoBERT, *Naïve Bayes*, SVM, dan KNN. Setiap algoritma ini memiliki keunggulan masing-masing dimana, IndoBERT dilatih untuk bahasa Indonesia, sehingga dapat memahami ulasan bahasa Indonesia dengan baik (Khofifah, 2023), Naive Bayes dapat bekerja dengan baik bahkan dengan dataset yang relatif kecil (Liem, 2024), SVM dapat bekerja secara efektif dalam klasifikasi ketika data memiliki banyak atribut (Lombu, 2022), dan KNN memiliki ketahanan terhadap data latih yang memiliki noise (Mutrofin, 2014). Adapun hasil penelitian menggunakan algoritma IndoBERT, *Naïve Bayes*, SVM, dan KNN sebagai berikut.

Pradhisa dan Fajriyah melakukan penelitian tentang analisis sentimen ulasan pengguna e-commerce Shopee dan Bukalapak menggunakan metode IndoBERT melibatkan 2500 data ulasan Shopee dan 2500 data ulasan Bukalapak, diperoleh akurasi sebesar 89% pada data ulasan Shopee dan 88% pada data ulasan Bukalapak (Pradhisa & Fajriyah, 2024). Adapun penelitian yang berjudul "Analisis Sentimen Menggunakan Algoritma Naive Bayes Terhadap Review Restoran di Singapura" mendapatkan nilai akurasi sebesar 73% (Permadi, 2020). Arsi dan Waluyo melakukan penelitian tentang analisis sentimen diskusi mengenai pemindahan ibu kota Indonesia, memanfaatkan algoritma SVM mencapai tingkat akurasi sebesar 96%. (Arsi & Waluyo, 2021). Adapun penelitian yang berjudul "Penerapan Algoritma KNN Pada Analisis Sentimen Review Aplikasi Peduli Lindungi" mendapatkan nilai akurasi sebesar 81% (Astuti & Nuris, 2022).

Berdasarkan penelitian sebelumnya diatas, algoritma *Naïve Bayes*, SVM, KNN, dan IndoBERT memberikan hasil akurasi yang baik dalam melakukan analisis sentimen. Oleh karena itu, penulis memiliki ketertarikan untuk membandingkan keempat algoritma tersebut untuk mengetahui algoritma yang paling terbaik pada analisis sentimen ulasan pengguna aplikasi M-Paspor.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang dijelaskan di atas, maka rumusan masalah penelitian ini, yaitu: "Bagaimana membandingkan algoritma IndoBERT, Naïve Bayes, SVM, dan KNN pada analisis sentimen ulasan pengguna aplikasi M-Paspor?"

### **1.3 Tujuan Penelitian**

Penelitian ini membandingkan algoritma IndoBERT, Naïve Bayes, SVM, dan KNN, dengan tujuan untuk mengetahui algoritma yang paling terbaik pada analisis sentimen ulasan pengguna aplikasi M-Paspor.

### **1.4 Batasan Masalah**

Agar permasalahan dapat lebih terfokus dan mudah dipahami, maka permasalahan dibatasi pada beberapa hal, yaitu:

1. Dalam penelitian ini, peneliti melakukan klasifikasi data dengan algoritma IndoBERT, Naïve Bayes, SVM, dan KNN untuk mengetahui algoritma yang paling akurat pada analisis sentimen ulasan pengguna aplikasi M-Paspor.
2. Data yang digunakan adalah ulasan pengguna aplikasi M-Paspor di Google Play Store pada 1 Januari 2023 hingga 7 Oktober 2024.
3. Klasifikasi dari data yang digunakan terdiri dari 3 kelas, yaitu positif, negatif, dan netral.

### **1.5 Manfaat Penelitian**

Adapun manfaat dari penelitian ini sebagai berikut:

1. Memberikan wawasan mengenai proses penerapan algoritma IndoBERT, Naïve Bayes, SVM, dan KNN pada analisis sentimen ulasan pengguna aplikasi M-Paspor.
2. Dapat digunakan sebagai referensi pada penelitian selanjutnya mengenai kinerja dari tiap algoritma dalam melakukan analisis sentim

**DAFTAR PUSTAKA**

- Af'idah, D. I., Dairoh, D., Handayani, S. F., Pratiwi, R. W., & Sari, S. I. (2022). Sentimen Ulasan Destinasi Wisata Pulau Bali Menggunakan Bidirectional Long Short Term Memory. *Matrik: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 21(3), 607–618.
- Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (Svm). *J. Teknol. Inf. Dan Ilmu Komput*, 8(1), 147.
- Artanti, V. (2024). Klasifikasi Penyakit Kardiovaskular Menggunakan Algoritma K-Nearest Neighbors (Knn) (Doctoral Dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).
- Asnawi, M. H., Firmansyah, I., Novian, R., & Pontoh, R. S. (2021). Perbandingan Algoritma Naïve Bayes, K-Nn, Dan Svm Dalam Pengklasifikasian Sentimen Media Sosial. *E-Prosiding Seminar Nasional Statistika| Departemen Statistika Fmipa Universitas Padjadjaran*, 10, 20.
- Astuti, P., & Nuris, N. (2022). Penerapan Algoritma Knn Pada Analisis Sentimen Review Aplikasi Peduli Lindungi. *Computer Science (Co-Science)*, 2(2), 137–142.
- Az-Zahra, F. F., & Pambudi, A. (2024). Fine-Tuning Model Indobert (Indonesian Bidirectional Encoder Representations From Transformers) Untuk Analisis Sentimen Berbasis Aspek Pada Aplikasi M-Paspor. *Joutica*, 9(2), 183–195.
- Chairina, N. (2023). Penggunaan Metode Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Prodi Pendidikan Teknologi Informasi Menggunakan Rapidminer (Doctoral Dissertation, Uin Ar-Raniry Banda Aceh).
- Faiz, A. Y. (2022). Aplikasi Klasifikasi Durian Menggunakan Algoritma K-Nearest Neighbors (Knn) Berbasis Android (Doctoral Dissertation, Universitas Pgrri Semarang).
- Fajar, F. M. (2024). Analisis Sentimen Kurikulum Merdeka Dengan Penerapan Convolutional Neural Network. *Journal Automation Computer Information System*, 4(1), 1-11.
- Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis Of Tweets Before The 2024 Elections In Indonesia Using Indobert Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI)*, 9(3), 746–757.
- Jauhary, H. (2023). Perbandingan Metode Analisis Sentimen Support Vector Machine, Naïve Bayes, Dan Logistic Regression (Studi Kasus: Ulasan Google Playstore Aplikasi LinkedIn) (Bachelor'S Thesis, Fakultas Sains Dan Teknologi Uin Syarif Hidayatullah Jakarta).

- Khofifah, T. N. (2024). Analisis Sentimen Terhadap Kebijakan Perubahan Skripsi Menggunakan Metode Indobert (Studi Kasus: Respon Masyarakat Terhadap Kebijakan Perubahan Skripsi Di Media Sosial Youtube) (Doctoral Dissertation, Universitas Islam Indonesia).
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). Indolem And Indobert: A Benchmark Dataset And Pre-Trained Language Model For Indonesian Nlp. Arxiv Preprint Arxiv:2011.00677.
- Lasulika, M. E. (2019). Komparasi Naïve Bayes, Support Vector Machine Dan K-Nearest Neighbor Untuk Mengetahui Akurasi Tertinggi Pada Prediksi Kelancaran Pembayaran Tv Kabel. *Ilkom Jurnal Ilmiah*, 11(1), 11–16.
- Liem, S., Setiawan, T., & Pribadi, M. R. (2024). Perbandingan Algoritma Svm Dan Naïve Bayes Berbasis Smote Dalam Analisis Sentimen Komentar Tiktok Pada Produk Skincare. *Applied Information Technology And Computer Science*, 3(1), 28-32.
- Lombu, A. S., Hidayat, S., & Hidayatullah, A. F. (2022). Pemodelan Klasifikasi Gaji Menggunakan Support Vector Machine. *Journal Of Computer System And Informatics (JOSYC)*, 3(4), 363-370.
- Lubis, R. A. (2024). Inovasi Aplikasi M-Paspor Di Kantor Imigrasi Kelas I Tpi Kota Pekanbaru Dalam Meningkatkan Pelayanan Publik (Doctoral Dissertation, Universitas Islam Negeri Sultan Syarif Kasim Riau).
- Mahatiara, R. E. (2023). Analisis Kinerja Algoritma Data Mining K-Nearest Neighbor (KNN) Dan Naive Bayes Untuk Klasifikasi Keluhan Masyarakat Terhadap Solusi Penanganan Sampah.
- Mutrofin, S., Izzah, A., Kurniawardhani, A., & Masrur, M. (2014). Optimasi Teknik Klasifikasi Modified K Nearest Neighbor Menggunakan Algoritma Genetika. *Jurnal Gamma*, 10(1).
- Nabiilah, G. Z., Alam, I. N., Purwanto, E. S., & Hidayat, M. F. (2024). Indonesian Multilabel Classification Using Indobert Embedding And Mbert Classification. *International Journal Of Electrical & Computer Engineering (2088-8708)*, 14(1).
- Natasuwarna, A. P. (2020). Seleksi Fitur Support Vector Machine Pada Analisis Sentimen Keberlanjutan Pembelajaran Daring. *Techno. Com*, 19(4).
- Permadi, V. A. (2020). Analisis Sentimen Menggunakan Algoritma Naïve Bayes Terhadap Review Restoran Di Singapura. *Jurnal Buana Informatika*, 11(2), 141–151.
- Pradhisa, K. C., & Fajriyah, R. (2024). Analisis Sentimen Ulasan Pengguna E-Commerce Di Google Play Store Menggunakan Metode Indobert. *Building Of Informatics, Technology And Science (BITS)*, 6(1), 92–104.

- Putra, K. T., Hariyadi, M. A., & Crysdiyan, C. (2023). Perbandingan Feature Extraction Tf-Idf Dan Bow Untuk Analisis Sentimen Berbasis Svm. *Jurnal Cahaya Mandalika* Issn 2721-4796 (Online), 3(2), 1449–1463.
- Rachma, C. A. (2022). Implementasi Algoritma K-Nearest Neighbor Dalam Penentuan Klasifikasi Tingkat Kedalaman Kemiskinan Provinsi Jawa Timur (Doctoral Dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).
- Romadhoni, Y. (2022). Klasifikasi Kalimat Perbincangan Masyarakat Terhadap Pandemi Covid-19 Pada Twitter Dengan Metode Long Short-Term Memory (Doctoral Dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).
- Suadana, L. H., Santoso, I., & Panjaitan, A. T. B. (2021). Transfer Learning Of Pre-Trained Transformers For Covid-19 Hoax Detection In Indonesian Language. *IJCCS (Indonesian Journal Of Computing And Cybernetics Systems)*, 15(3), 317–326.
- Sukmawati, S., Rijal, R., & Iriawan, H. (2020). Kualitas Pelayanan Pembuatan Paspor Pada Kantor Imigrasi Kelas Ii Tempat Pemeriksaan Imigrasi Biak. *Matra Pembaruan: Jurnal Inovasi Kebijakan*, 4(1), 37–46.
- Wardhana, S. R., Purwitasari, D., & Rochimah, S. (2016). Analisis Sentimen Pada Review Pengguna Aplikasi Mobile Untuk Evaluasi Faktor Usability. *Jurnal Sistem Dan Informatika (JSI)*, 11(1), 128-136.
- Yahya, S. A. (2018). Klasifikasi Ketepatan Lama Studi Mahasiswa Menggunakan Metode Support Vector Machine Dan Random Forest (Studi Kasus: Data Lama Studi Alumni Universitas Islam Indonesia Tahun Kelulusan 2000-2017).
- Zhafirah, D. (2023). Penanganan Imbalance Data Dengan Random Oversampling (ROS) Pada Klasifikasi Penderita Diabetes Menggunakan Support Vector Machine (SVM).