

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Transformasi yang cepat dalam pelayanan publik kini memungkinkan masyarakat untuk mengakses layanan tidak hanya manual, tetapi bisa melalui secara online (Permadi & Rokhman, 2023). Pesatnya perkembangan teknologi komputer pada sektor pelayanan publik, khususnya di sistem identitas digital mulai merasakan langsung manfaatnya, di mana proses yang sebelumnya secara manual kini perlahan beralih ke bentuk digital (Fatimah & Purwanti, 2023). Salah satu cara nyata untuk memperbaiki kualitas pelayanan kepada masyarakat adalah dengan menerapkan Teknologi Informasi dan Komunikasi secara maksimal dalam setiap proses pelayanan.

Sebagai bentuk inovasi, Dirjen Dukcapil Kementerian Dalam Negeri telah meluncurkan Identitas Kependudukan Digital (IKD), atau yang dikenal luas sebagai KTP Digital Elektronik. IKD merupakan langkah maju Dirjen Dukcapil dalam mendigitalkan dokumen kependudukan bagi seluruh masyarakat Indonesia yang mempunyai KTP elektronik IKD bertujuan untuk beradaptasi dengan kemajuan teknologi informasi dan komunikasi, mendorong masyarakat untuk mengadopsi digitalisasi kependudukan, serta mempercepat pertukaran data digital dalam pelayanan public. Selain itu, IKD dirancang untuk meningkatkan keamanan kepemilikan data melalui kerangka kerja verifikasi yang kuat, guna mencegah pemalsuan identitas dan kebocoran data. (Peraturan Kemendagri Nomor 72 Tahun 2022).

## Identitas Kependudukan Digital

DITJEN DUKCAPIL KEMENDAGRI

Layanan Dokumen Administrasi Kependudukan dan Pencatatan Sipil.



3,0★  
65,6 rb ulasan

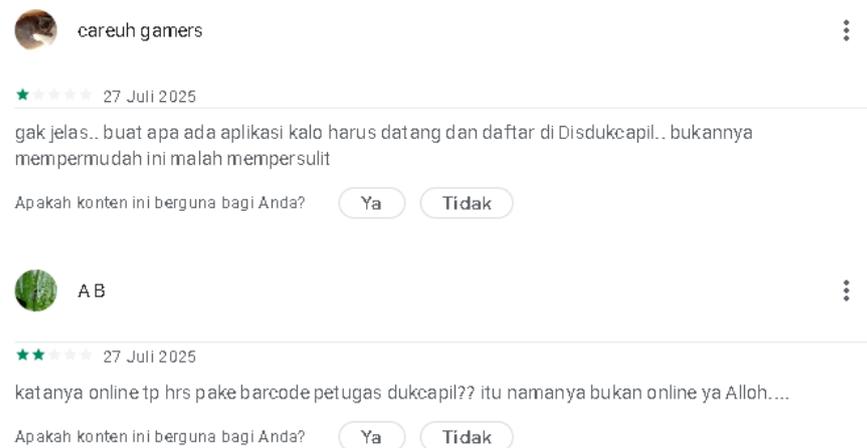
 Pemerintah

10 jt+  
Download

 3+  
Rating 3+ 

**Gambar 1. 1** *Rating* dari Aplikasi Identitas Kependudukan Digital

Namun, dalam implementasinya aplikasi ini masih mengalami sejumlah permasalahan yang berdampak pada tingkat kepuasan pengguna. Berdasarkan **Gambar 1.1** aplikasi ini memiliki rating sebesar 3,0 dari total lebih dari 65 ribu ulasan, dengan jumlah unduhan yang telah melebihi 10 juta. Capaian ini menunjukkan adanya gap antara ekspektasi pengguna dengan kinerja aplikasi.



### **Gambar 1.2** *Ulasan dari Aplikasi Identitas Kependudukan Digital*

Selain itu, banyaknya ulasan negatif terkait fitur layanan dalam aplikasi ini menunjukkan adanya kendala yang dirasakan langsung oleh pengguna, khususnya dalam hal kemudahan mengakses dan mengurus dokumen administrasi kependudukan. **Gambar 1.2** menunjukkan beberapa pengguna mengeluhkan bahwa proses pendaftaran masih mengharuskan kehadiran langsung dari kantor Disdukcapil, yang dinilai bertentangan dengan konsep digital. Upaya yang dilakukan untuk mendukung peningkatan tersebut adalah melakukan analisis sentimen terhadap ulasan pengguna, sehingga dapat dijadikan acuan dalam pengembangan aplikasi agar memperoleh rating yang lebih baik.

Analisis sentimen, atau sering disebut juga penambangan opini (*opinion mining*), adalah proses otomatis untuk memahami, mengekstrak, dan mengelola data tekstual guna memperoleh informasi yang terkandung dalam sebuah kalimat opini. Analisis sentimen merupakan metode yang digunakan untuk mengevaluasi dan mengidentifikasi opini terhadap suatu isu atau objek. Tujuannya untuk

menentukan apakah opini tersebut memiliki kecenderungan positif, negatif, netral (Luthfika Fairuz et al., 2021). Analisis sentimen terhadap ulasan pengguna aplikasi IKD sangat penting karena akan menghasilkan informasi berharga. Informasi ini berfungsi sebagai dasar perbaikan aplikasi IKD, yang diharapkan bisa untuk meningkatkan nilai rating aplikasi tersebut. Untuk mencapainya, ulasan pengguna, baik ulasan positif maupun negatif perlu dilakukan klasifikasi secara akurat. Beberapa metode yang umum digunakan dalam klasifikasi sentimen meliputi *Support Vector Machine (SVM)*, *Naive Bayes*, *Decision Tree*, dan lain sebagainya.

Penelitian ini memanfaatkan metode *Support Vector Machine (SVM)* untuk melakukan analisis sentimen. Metode ini dipilih karena landasan teori yang kuat dan kemampuannya untuk melakukan klasifikasi dengan akurasi tinggi daripada metode yang lainnya. SVM merupakan algoritma *machine learning* yang didasarkan pada prinsip *Structural Risk Minimization (SRM)*, dengan tujuan utama menemukan *hyperplane* terbaik yang dapat memisahkan dua kelas dalam ruang input (Gifari et al., 2022). Sebuah penelitian yang dilakukan oleh Rita Ajeng Lestari, et al. pada tahun 2023, yang berfokus pada penerapan algoritma *Support Vector Machine* pada analisis sentimen terhadap Identitas Kependudukan Digital, menunjukkan hasil yang signifikan. Dari pengujian dan evaluasi, algoritma SVM berhasil mencapai akurasi sebesar 77,0%. Hasil ini mengidentifikasi bahwa SVM memiliki akurasi yang tinggi dalam klasifikasi.

Penelitian oleh Azhari et al. (2021) membahas Perbandingan kinerja klasifikasi antara algoritma C4.5, *Random Forest*, SVM, dan *Naïve Bayes* berdasarkan metrik akurasi, *recall*, dan *presisi*. Studi ini menggunakan data dari kegiatan pramuka *Jogja International Scout Camp (JISC)* di Provinsi Sumatera Utara tahun 2020. Hasil penelitian menunjukkan bahwa masing-masing algoritma memberikan tingkat akurasi yang berbeda, yaitu C4.5 dan *Naïve Bayes* memperoleh akurasi sebesar 86,67%, *Random Forest* sebesar 83,33%, dan SVM mencapai akurasi tertinggi sebesar 95%. Dengan demikian, algoritma SVM menunjukkan performa paling baik, sementara *Random Forest* memiliki akurasi terendah diantara algoritma yang diuji.

## **1.2. Rumusan Masalah**

Merujuk pada uraian latar belakang yang telah dijelaskan sebelumnya, maka permasalahan dalam penelitian ini dapat dirumuskan ke dalam beberapa pertanyaan sebagai berikut:

1. Bagaimana penerapan algoritma Support Vector Machine dalam melakukan klasifikasi sentimen terhadap ulasan pengguna aplikasi Identitas Kependudukan Digital (IKD) di platform Google Play Store?
2. Seberapa baik kinerja model klasifikasi yang dibangun menggunakan metode Support Vector Machine, jika diukur berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score?
3. Bagaimana cara menyajikan hasil klasifikasi sentimen dari ulasan aplikasi IKD dalam bentuk visualisasi yang informatif dan mudah dipahami?

## **1.3. Tujuan Penelitian**

Penelitian ini bertujuan untuk mencapai hal-hal sebagai berikut:

1. Menerapkan algoritma Support Vector Machine dalam proses klasifikasi sentimen terhadap ulasan pengguna aplikasi Identitas Kependudukan Digital (IKD) di Google Play Store.
2. Melakukan evaluasi terhadap performa algoritma Support Vector Machine dalam mengklasifikasikan sentimen, dengan mengacu pada metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.
3. Menyajikan hasil klasifikasi sentimen dalam bentuk visualisasi yang informatif dan mudah dipahami sebagai bagian dari analisis data ulasan aplikasi IKD.

## **1.4. Batasan dan Asumsi Penelitian**

Beberapa batasan untuk menjaga fokus dan ruang lingkup masalah dari penelitian ini, yaitu:

1. Data yang digunakan hanya pada ulasan aplikasi Identitas Kependudukan Digital yang tersedia di *Play Store*.

2. Teks yang diambil hanya berbahasa Indonesia.
3. Pengambilan dataset dimulai dari 1 Agustus 2023 - 31 Januari 2025.

### **1.5. Manfaat Penelitian**

Berikut adalah manfaat yang diharapkan dari penelitian ini:

1. Menerapkan ilmu yang didapatkan untuk melakukan analisis sentimen dan mengklasifikasikannya menggunakan algoritma *Support Vector Machine* (SVM).
2. Memberikan wawasan tentang ilmu data mining dan mendapatkan gambaran langkah proses analisis sentimen menggunakan algoritma *Support Vector Machine* (SVM).
3. Dapat mengukur tingkat kemampuan mahasiswa selama belajar dalam perkuliahan dan dapat menambah referensi literatur terhadap algoritma *Support Vector Machine* (SVM).

### **1.6. Sistematika Penulisan**

Tugas akhir berjudul “Analisis Sentimen Pada Aplikasi Identitas Kependudukan Digital Menggunakan Metode *Support Vector Machine* (SVM)” ini adalah dengan tahapan sebagai berikut:

#### **1.6.1. Studi Literatur**

Tahap awal ini melibatkan penelusuran dan pemahaman konsep-konsep terkait analisis sentimen, *Support Vector Machine* (SVM), serta aplikasi Identitas Kependudukan Digital.

#### **1.6.2. Pengumpulan Data**

Tahap ini berfokus pada pengumpulan data berupa ulasan-ulasan penggunaan dari aplikasi Identitas Kependudukan Digital. Proses pengumpulan data ini dilakukan menggunakan teknik *scraping*.

### **1.6.3. Preprocessing**

Pada tahap ini digunakan untuk melakukan proses pada hasil dari pengumpulan data agar dapat diolah lebih baik pada tahap ekstraksi. *Preprocessing* ini dapat dilakukan setelah mempelajari beberapa jurnal artikel internet tentang tahap *preprocessing*.

### **1.6.4. Pelabelan**

Pada tahap ini digunakan untuk melakukan proses menetapkan atau memasang suatu objek atau data ke dalam kategori atau kelas tertentu. Ini merupakan bagian dari bidang pembelajaran mesin di mana model dilatih untuk mengenali dan memisahkan data menjadi kategori yang telah ditentukan sebelumnya. Pelabelan ini dapat dilakukan setelah mempelajari beberapa jurnal artikel internet tentang tahap pelabelan.

### **1.6.5. Pembobotan TF-IDF**

Pada tahap ini, dilakukan salah satu proses dalam teknik ekstraksi fitur, yaitu pemberian bobot pada tiap kata dalam data latih. Tujuannya untuk mengetahui sejauh mana suatu kata merepresentasikan sebuah kalimat. Pemberian bobot ini dilakukan berdasarkan pemahaman dari berbagai referensi, seperti jurnal ilmiah dan artikel di internet mengenai tahap pembobotan.

### **1.6.6. Klasifikasi SVM**

Tahap ini merupakan proses klasifikasi yang menggunakan algoritma machine learning berbasis pembelajaran terawasi (*supervised learning*) untuk memprediksi kelas dari suatu sampel atau model berdasarkan hasil pelatihan sebelumnya. *Support Vector Machine* bekerja dengan mencari garis pembatas yang paling optimal. Garis ini berfungsi untuk memisahkan dua kelas data sehingga menghasilkan akurasi model tertinggi. Klasifikasi SVM ini dapat dilakukan setelah mempelajari beberapa jurnal artikel internet tentang *Support Vector Machine* dan juga Kernel.

### **1.6.7. Evaluasi**

Tahap ini bertujuan untuk mengevaluasi sejauh mana performa suatu model dalam mengklasifikasi data ke dalam kelas yang sesuai. *Confusion matrix* adalah salah satu metode yang digunakan untuk melakukan evaluasi proses model klasifikasi berupa jumlah data uji yang benar dan salah. Evaluasi ini dapat dilakukan setelah mempelajari beberapa jurnal artikel internet tentang confusion matrix.