

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Pengguna internet di Indonesia terus mengalami peningkatan. Menurut data dari We Are Social, pada awal Januari 2024, jumlah penduduk Indonesia tercatat mencapai 278,7 juta orang, dengan 225,6 juta di antaranya merupakan pengguna internet, yang setara dengan sekitar 77% dari total populasi pada Desember 2023 [1]. Penggunaan internet memudahkan berbagai aktivitas, mulai dari komunikasi, akses informasi, pengembangan bisnis, hingga hiburan.

Pengguna smartphone di Indonesia juga menunjukkan peningkatan setiap tahunnya. Menurut Badan Pusat Statistik, pada tahun 2022, sebanyak 67,88% dari total penduduk Indonesia yang berumur 5 tahun ke atas memiliki smartphone [2].



Gambar 1.1 Persentase Penduduk yang Memiliki Smartphone (Survei Sosial Ekonomi Nasional, BPS)

Dapat dilihat dari grafik persentase tersebut, menunjukkan ketergantungan penduduk dengan smartphone masih sangat masif. Saat ini, smartphone telah berkembang menjadi perangkat serba bisa yang memudahkan berbagai pekerjaan, termasuk dalam kegiatan pembayaran pajak.

Kemajuan dalam teknologi informasi telah membawa dampak yang menguntungkan bagi beragam sektor dalam pemerintahan, salah satunya adalah perpajakan. Teknologi yang semakin canggih memungkinkan manajemen data dan informasi pajak dapat lebih cepat, tepat, dan efisien. Penerapan teknologi dalam administrasi perpajakan, yang diwujudkan melalui aplikasi perpajakan, telah menjadi salah satu solusi efektif untuk meningkatkan kepatuhan wajib pajak serta mengoptimalkan penerimaan negara.

Peningkatan penggunaan internet menggunakan smartphone telah membawa perubahan dalam perilaku wajib pajak. Mereka dapat menggunakan smartphone untuk mengakses informasi seputar perpajakan mulai dari peraturan perpajakan, pembayaran pajak dengan tempo pembayaran pajak. Mereka tidak harus mengunjungi kantor pajak, konsultasi dengan ahli pajak, atau bertanya kepada orang lain tentang pajak, cukup dengan satu ketukan di layar smartphone untuk mendapatkan informasi tersebut. Kemudahan ini tidak hanya menghemat waktu namun juga mengurangi kecemasan ketidaktahuan bagi para wajib pajak.

Aplikasi perpajakan menawarkan berbagai fitur yang memudahkan para wajib pajak dalam memenuhi kewajiban perpajakan, seperti pelaporan, pembayaran, dan pengarsipan pajak. Dengan adanya aplikasi ini, proses yang sebelumnya memakan waktu dan rentan terhadap kesalahan kini dapat dilakukan secara otomatis dan real-time, sehingga dapat mengurangi beban administrasi bagi lembaga pajak dan para wajib pajak.

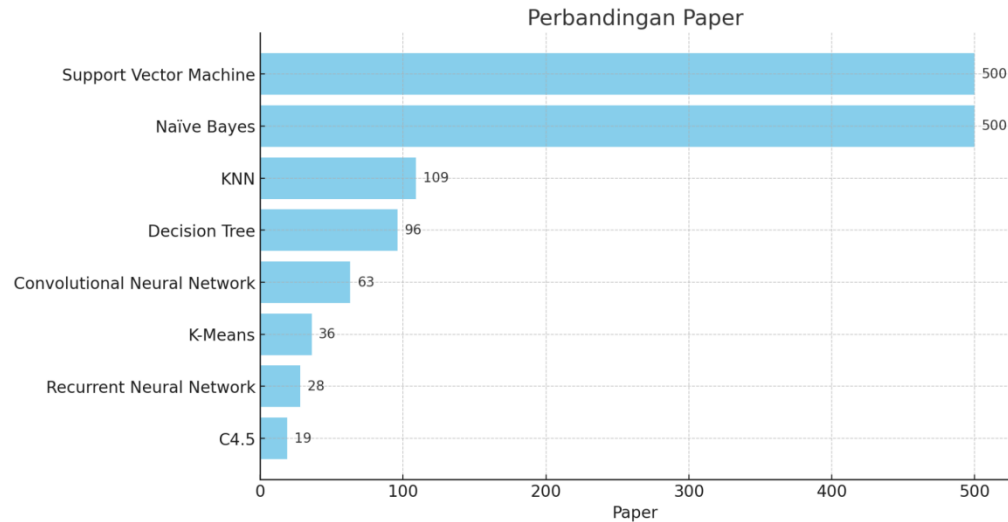
Direktorat Jenderal Pajak (DJP) adalah bagian dari Kementerian Keuangan Republik Indonesia yang bertugas mengelola dan mengumpulkan pajak, memastikan kepatuhan wajib pajak, serta menegakkan hukum perpajakan di Indonesia. Perkembangan pengguna internet mendorong DJP untuk menyesuaikan diri dengan meluncurkan aplikasi yaitu M-Pajak

M-Pajak merupakan aplikasi digital yang dibangun oleh Direktorat Jenderal Pajak (DJP) untuk membantu masyarakat dalam urusan perpajakan. Aplikasi ini menawarkan berbagai fitur menarik, seperti riwayat pajak, kartu NPWP digital, informasi seputar pajak, serta pengingat tenggat waktu pelaporan dan pembayaran pajak. Aplikasi ini pertama kali dirilis oleh DJP pada tahun 2021 dan telah tersedia di AppStore dan PlayStore. Hingga tahun 2024, total unduhan dari PlayStore lebih dari 1 juta. Namun, meskipun telah diunduh oleh 1 juta pengguna, aplikasi ini memiliki respon negatif di

PlayStore seperti mengeluh tentang bug, tidak dapat membayar pajak, atau kinerja aplikasi yang lambat.

Respon negatif di PlayStore dapat mempengaruhi calon pengguna baru dalam memutuskan untuk menggunakan aplikasi ini. Apabila jumlah teks yang besar akan membuat analisa menjadi sangat sulit. Oleh karena itu, penelitian ini akan mengkaji sentimen yang terkandung pada ulasan PlayStore pada aplikasi M-Pajak dengan menggunakan beberapa algoritma klasifikasi, yakni: algoritma Support Vector Machine, Naive Bayes, dan KNN. Penelitian ini mengikuti studi sebelumnya yang mencapai *accuracy* 92% menggunakan *Naive Bayes* pada aplikasi perpajakan motor [3].

Penelitian ini akan melakukan perbandingan akurasi pada klasifikasi *SVM*, *Naive Bayes*, dan *KNN* dengan menggunakan ulasan aplikasi M-Pajak sebagai bahan untuk melakukan analisis sentimen. Data penelitian diambil dari Kaggle dan akan diklasifikasikan menjadi dua kategori sentimen: positif dan negatif. Hasil akhir dari penelitian ini akan mengetahui tingkat akurasi, precision, dan recall dari ketiga algoritma klasifikasi tersebut. Peneliti memiliki algoritma klasifikasi *SVM*, *Naive Bayes*, dan *KNN* karena algoritma tersebut paling banyak digunakan dalam penelitian analisis sentimen di Indonesia. Data ini dikumpulkan dari hasil pencarian di *Google Scholar* menggunakan aplikasi *Publish or Perish* pada Agustus 2024. Pencarian dilakukan dengan kata kunci seperti Support Vector Machine sentimen, Naïve Bayes sentimen, KNN sentimen, Convolutional Neural Network sentimen, Decision Tree sentimen, K-Means sentimen, C4.5 sentimen, dan Recurrent Neural Network sentimen. Pada saat pencarian peneliti memberi batasan 500 pencarian. Berikut grafik jumlah perbandingan paper algoritma dalam waktu 5 tahun terakhir (2024-2020):



Gambar 1.2 Perbandingan paper Algoritma Sentimen

1.2. Perumusan Masalah

1.2.1. Permasalahan

Aplikasi M-Pajak yang dikembangkan oleh Direktorat Jenderal Pajak (DJP) bertujuan untuk mempermudah wajib pajak dalam mengelola kewajiban perpajakan, mulai dari pelaporan hingga pembayaran. Meski demikian, setelah peluncurannya, aplikasi ini menerima beberapa ulasan negatif di PlayStore, yang berpotensi menurunkan minat pengguna baru. Oleh karena itu, diperlukan sebuah analisis sentimen guna mengklasifikasikan ulasan-ulasan ini menjadi sentimen positif atau negatif, dengan tujuan untuk memahami lebih dalam persepsi pengguna terhadap aplikasi tersebut.

1.2.2. Metode Usulan

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap ulasan aplikasi M-Pajak. Proses klasifikasi sentimen akan dibagi menjadi dua kategori, yaitu sentimen positif dan negatif. Dalam penelitian ini, tiga algoritma akan digunakan untuk klasifikasi, yakni *Support Vector Machine* (SVM), *K-Nearest Neighbors* (KNN), dan *Naive Bayes* (NB). Evaluasi kinerja algoritma akan dilakukan dengan menggunakan metode validasi silang *10-Fold Cross Validation*, yang bertujuan untuk mengukur tingkat *accuracy*, *precision*, dan *recall* dari tiap algoritma yang digunakan.

1.2.3. Pertanyaan Penelitian

Berdasarkan latar belakang yang telah dipaparkan sebelumnya, maka rumusan masalah yang diangkat dalam penelitian ini adalah:

1. Bagaimana hasil analisis sentimen ulasan aplikasi M-Pajak berdasarkan algoritma *Support Vector Machine*, *Naive Bayes*, dan *KNN* dalam membagi ulasan ke dalam sentimen positif dan negatif?
2. Bagaimana algoritma *Support Vector Machine*, *Naive Bayes*, dan *KNN* mampu mengklasifikasikan ulasan M-Pajak berdasarkan sentimen yang diekspresikan oleh pengguna?
3. Apa saja faktor yang memengaruhi kinerja algoritma *Support Vector Machine*, *Naive Bayes*, dan *KNN* dalam melakukan analisis sentimen terhadap ulasan M-Pajak?
4. Bagaimana tingkat akurasi dari hasil analisis sentimen aplikasi M-Pajak dengan menggunakan algoritma *Support Vector Machine*, *Naive Bayes*, dan *KNN*?

1.3. Tujuan dan Manfaat

1.3.1. Tujuan

Tujuan yang diperoleh dalam penelitian ini adalah mengetahui tingkat akurasi dari algoritma klasifikasi *Support Vector Machine*, *Naive Bayes*, dan *KNN* dalam analisis sentimen ulasan aplikasi M-Pajak

1.3.2. Manfaat

Adapun manfaat penelitian, yaitu:

1. Memberikan pengetahuan mengenai analisis sentimen ulasan Aplikasi M-Pajak menggunakan algoritma *Support Vector Machine*, *Naive Bayes*, dan *KNN*.
2. Melihat dan melakukan analisis sentimen secara cepat ke dalam kelas positif dan negatif.
3. Menjadi referensi untuk penelitian mendatang mengenai analisis sentimen menggunakan algoritma *Support Vector Machine*, *Naive Bayes*, dan *KNN*.

1.4. Batasan Masalah

Adapun batasan-batasan penelitian ini adalah:

1. Pada penelitian ini, data yang digunakan diambil dari Kaggle yang berjudul “Sentimen Aplikasi M-Pajak”, data diambil pada tanggal 26 Agustus 2024.

2. Analisis sentimen menggunakan algoritma klasifikasi *Naive Bayes* (NB), *K-Nearest Neighbors* (KNN), dan *Support Vector Machine* (SVM) dengan menggunakan aplikasi RapidMiner.
3. Analisis sentimen mencakup kelas positif dan negatif serta perbandingan accuracy dari algoritma klasifikasi *Naive Bayes* (NB), *K-Nearest Neighbors* (KNN), dan *Support Vector Machine* (SVM) dengan menggunakan aplikasi RapidMiner.

1.5. Sistematika Penulisan

Penulisan skripsi ini disusun secara sistematis dan dibagi menjadi beberapa BAB, seperti berikut:

1. BAB I Pendahuluan

Bab ini menjelaskan tentang latar belakang penelitian, rumusan masalah, tujuan yang ingin dicapai, manfaat penelitian, serta batasan masalah yang dihadapi. Selain itu, diuraikan pula sistematika penulisan sebagai panduan bagi pembaca.

2. BAB II Tinjauan Pustaka

Pada bab ini, dibahas teori-teori yang relevan dan berkaitan langsung dengan penelitian. Dasar-dasar teori ini menjadi pijakan untuk memahami konteks penelitian yang dilakukan.

3. BAB III Metodologi Penelitian

Bab ini menguraikan secara rinci mengenai lokasi dan waktu penelitian, serta metode yang digunakan. Prosedur dan teknik yang dipilih dalam melakukan analisis sentimen juga dijelaskan di sini.

4. BAB IV Hasil Penelitian dan Analisis

Bab ini memaparkan hasil penelitian berupa implementasi metode klasifikasi sentimen yang telah diuji. Analisis terhadap akurasi algoritma *Support Vector Machine*, *Naive Bayes*, dan *K-Nearest Neighbors* juga dijelaskan di bagian ini.

5. BAB V Penutup

Kesimpulan dan saran dari hasil penelitian.

1.6. Jadwal Penelitian

Penelitian ini dilaksanakan pada semester ganjil 2023/2024 di Program Studi Sistem Informasi, Jurusan Teknik Informasi, Fakultas Teknik, Universitas Trunojoyo Madura. Jadwal penelitian akan dijelaskan secara rinci pada BAB III, pada poin bagian Tempat dan Waktu Penelitian.