

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan data Badan Pusat Statistik (BPS), jumlah kendaraan bermotor berbagai jenis di Indonesia tahun 2021 adalah 141.992.573 dan setiap tahun akan mengalami peningkatan. Banyak Masyarakat yang memiliki kendaraan pribadi lebih dari satu karena kendaraan pribadi dapat membuat perjalanan lebih mudah dan lebih cepat kemanapun kita inginkan [1]. Setiap Masyarakat yang memiliki kendaraan pribadi diwajibkan untuk membayar pajak kendaraan bermotor. Pajak kendaraan bermotor ini adalah pembayaran Surat Tanda Nomor Kendaraan bermotor yang dibayarkan setiap satu tahun sekali atau lima tahun STNK [2].

Berdasarkan peraturan Menteri Dalam Negeri Republik Indonesia Nomor 82 Tahun 2022 tentang dasar pengenaan pajak kendaraan bermotor, maka semua kendaraan bermotor dikenakan pajak kendaraan bermotor. Pembayaran pajak kendaraan bermotor sebelumnya dilakukan dengan cara datang ke Samsat atau Unit Pelayanan Samsat yang tersebar di seluruh kabupaten/kota di Indonesia. Dengan perkembangan teknologi yang semakin pesat, hampir semua aspek kehidupan Masyarakat menjadi lebih mudah, hal ini dapat dilakukan karena adanya kemajuan teknologi digital [3]. Pemerintah memanfaatkan teknologi dalam pelayanan publik yang bisa diakses kapan pun dan dari mana pun pengguna berada dengan menerapkan sistem online sehingga masyarakat bisa lebih mudah melakukan pembayaran pajak karena bisa dilakukan secara online [4].

Saat ini muncul aplikasi yang dapat digunakan di *Smartphone* yang dinamakan SIGNAL. Aplikasi SIGNAL adalah aplikasi Samsat Digital Nasional, sebuah aplikasi yang berguna memudahkan Masyarakat dalam melakukan pembayaran pajak kendaraan bermotor dengan aman dan mudah. Signal bisa digunakan Masyarakat untuk membayar pajak kendaraan dari rumah [5].

Pengguna aplikasi SIGNAL mengalami peningkatan seiring berjalannya waktu, karena semakin banyak masyarakat yang mengetahui tentang adanya aplikasi tersebut. Masyarakat menggunakan aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL dengan mengunduh melalui Google Play Store dan menceritakan pengalamannya pada komentar di Google Play Store. Untuk

mengetahui optimal atau tidak aplikasi tersebut, dapat diketahui dengan menggunakan metode yang cepat dan tepat yaitu melakukan analisis sentimen ulasan komentar pengguna aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL dari Google Playstore. Ada beberapa algoritma klasifikasi yang bisa digunakan untuk melakukan analisis sentimen salah satunya seperti *Support Vector Machine* (SVM) [6].

Penelitian pertama yang menjadi referensi adalah penelitian yang dilakukan oleh E. R. Kaburuan and N. R. Setiawan dengan judul “Sentimen Analisis Review Aplikasi Digital Korlantas Pada Google Play Store Menggunakan Metode SVM” menggunakan metode *Support Vector Machine* mendapatkan hasil yang terbaik dengan rasio data 90:10 mendapatkan nilai *accuracy* 0.82, sedangkan hasil yang terburuk dengan rasio 80:20 dan 60:40 mendapatkan nilai *accuracy* sama yaitu 0.74 [3].

Penelitian kedua yang menjadi referensi adalah penelitian yang dilakukan oleh M. Diki Hendriyanto, A. A. Ridha, and U. Enri yang berjudul “Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine” dilakukan klasifikasi sentimen dengan tiga skenario pembagian data dan menggunakan empat kernel pada algoritma *Support Vector Machine* dan mendapatkan hasil yang diperoleh pada skenario 90:10 menggunakan kernel RBF (*Radial Basis Function*) menghasilkan *accuracy* 92,31%, *precision* 96,3%, *recall* 89,66%, dan *f1-score* 92,86% [7].

Penelitian ketiga yang menjadi referensi adalah penelitian yang dilakukan oleh Shahmirul Hafizullah Imanuddin, Kusworo Adi, dan Rahmat Gernowo yang berjudul “Sentiment Analysis on Satusehat Application Using Support Vector Machine Method” menggunakan metode *Support Vector Machine* dengan membagi kategori menjadi dua yaitu negatif dan positif, didapat hasil 73.4% sentimen negatif, dan 26.6% sentimen positif dengan akurasi pengujian sebesar 91% [8].

Penelitian keempat yang menjadi referensi adalah penelitian yang dilakukan oleh Z. A. Diekson, M. R. B. Prakoso, M. S. Q. Putra, M. S. A. F. Syaputra, S. Achmad, and R. Sutoyo yang berjudul “Sentiment analysis for customer review: Case study of Traveloka” menggunakan tiga metode untuk melakukan sentimen analisis, salah satunya adalah metode *Support Vector Machine* dan mendapatkan hasil evaluasi *Confusion Matrix* terbaik diantara ketiga metode yang digunakan [9].

Penelitian kelima yang menjadi referensi adalah penelitian yang dilakukan oleh R. AL Anshari *et al* yang berjudul “Komparasi payment digital untuk analisis sentimen berdasarkan ulasan di google playstore menggunakan metode support vector machine” menggunakan metode *Support Vector Machine* untuk melakukan analisis sentimen ulasan dua aplikasi yaitu Dana dan Ovo, dari kedua aplikasi tersebut mendapatkan hasil *accuracy* yang berbeda, aplikasi Dana mendapatkan *accuracy* sebesar 92%, sedangkan Ovo 90% [10].

Penelitian keenam yang menjadi referensi adalah penelitian yang dilakukan oleh Cholid Fadilah Hasri dan Debby Alita yang berjudul “Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter” menggunakan metode *Naïve Bayes* dan *Support Vector Machine* dengan membagi kategori menjadi tiga yaitu positif, negatif, dan netral. Hasil yang diperoleh untuk algoritma *Naïve Bayes Classifier* menghasilkan akurasi sebesar 81.07%, dan untuk algoritma *Support Vector Machine* menghasilkan akurasi sebesar 79.96% [11].

Penelitian ketujuh yang menjadi referensi adalah penelitian yang dilakukan oleh Ratih Puspitasari, Y. Findawati, and M. A. Rosid dengan judul “Sentiment Analysis Of Post-Covid-19 Inflation Based On Twitter Using The K-Nearest Neighbor And Support Vector Machine Classification Methods” menggunakan dua metode yaitu *K-nearest Neighbor* dan *Support Vector Machine* dengan membagi kategori menjadi tiga yaitu positif, negatif, dan netral. Terkumpul 5989 data dengan 2508 data positif, 1804 data negatif, dan 1677 data netral. Hasil yang didapat untuk metode *K-Nearest Neighbor* mendapat akurasi sebesar 54%, dan untuk *Support Vector Machine* mendapat akurasi sebesar 79% [12].

Penelitian kedelapan yang menjadi referensi adalah penelitian yang dilakukan oleh A. W. Sari, T. I. Hermanto, and M. Defriani yang berjudul “Sentiment Analysis Of Tourist Reviews Using K-Nearest Neighbors Algorithm And Support Vector Machine” menggunakan metode *K-Nearest Neighbors* dan *Support Vector Machine* dengan membagi kategori menjadi dua yaitu positif dan negatif. Penelitian ini menggunakan data yang diperoleh dari hasil scraping dari situs Trip Advisor dengan lima (5) tempat wisata. Diperoleh hasil akurasi menggunakan metode *Support Vecrtor Machine* pada Waterbom Bali sebesar 88%, Mandala Suci Wenara

Wana sebesar 83%, Tegallalang Terraces sebesar 89%, Pura Tanah Lot sebesar 88%, dan Pura Uluwatu sebesar 95%. Sedangkan dengan menggunakan *metode K-Nearest Neighbors* pada Waterbom Bali sebesar 67%, Mandala Suci Wenara Wana sebesar 60%, Tegallalang Terraces sebesar 69%, Pura Tanah Lot sebesar 71%, dan Pura Uluwatu sebesar 81%. Dapat disimpulkan bahwa algoritma *Support Vector Machine* lebih baik daripada algoritma *K-Nearest Neighbors* [13].

Penelitian kesembilan yang menjadi referensi adalah penelitian yang dilakukan oleh M. N. Muttaqin and I. Kharisudin yang berjudul “Analisis Sentimen Aplikasi Gojek Menggunakan Support Vector Machine Dan K Nearest Neighbor” menggunakan metode *Support Vector Machine* dan *K-Nearest Neighbor* dengan membagi kategori menjadi dua yaitu sentimen positif dan negatif. Pada metode KNN mendapatkan hasil terbaik dengan nilai $K=22$ dengan nilai akurasi, presisi, dan recall sebesar 82,14%, 82,28%, dan 95,43%. Sedangkan metode SVM didapat hasil bahwa kernel *linear* dengan parameter $C=1$ mendapatkan nilai terbaik dengan akurasi, presisi, dan recall berturut-turut sebesar 87,98%, 88,55%, dan 95,43% [14].

Penelitian kesepuluh atau penelitian terakhir yang menjadi referensi adalah penelitian yang dilakukan oleh F. Fitriana, E. Utami, and H. Al Fatta yang berjudul “Analisis Sentimen Opini Terhadap Vaksin Covid-19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes” menggunakan dua metode yaitu *Support Vector Machine* dan *Naive Bayes* dengan membagi kategori menjadi tiga yaitu sentimen netral, negatif, dan positif. Hasil performa menggunakan algoritma *Support Vector Machine* memperoleh akurasi sebesar 90.47%, presisi 90.23%, dan recall 90.78%, sedangkan algoritma *Naive Bayes* memperoleh akurasi sebesar 88.64%, presisi 87.32%, dan recall 88.13% [15].

Berdasarkan beberapa penelitian yang ada sebelumnya, pada penelitian ini dilakukan analisis sentimen ulasan komentar pengguna aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL dari Google Playstore yang dimana penelitian ini bertujuan untuk melakukan klasifikasi terhadap ulasan Masyarakat ke dalam dua kategori sentimen, yaitu sentimen positif dan negatif. Dalam penelitian ini, peneliti menggunakan metode *Support Vector Machine* untuk melakukan klasifikasi ulasan Masyarakat mengenai aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian, maka rumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana implementasi sentimen Masyarakat terhadap aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL menggunakan metode *Support Vector Machine*?
2. Bagaimana hasil evaluasi metode *Support Vector Machine* pada analisis sentimen?

1.3 Tujuan Penelitian

Adapun tujuan pada penelitian ini adalah :

1. Melakukan implementasi analisis sentimen ulasan dari Google Play Store terhadap komentar pengguna aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL menggunakan metode *Support Vector Machine*.
2. Mengetahui tingkat akurasi *Support Vector Machine* menggunakan Confusion Matrix pada aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL.

1.4 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah :

1. Penelitian ini dilakukan untuk mengetahui ulasan aplikasi SIGNAL – SAMSAT DIGITAL NASIONAL pada Google Play Store.
2. Data diperoleh dari hasil scraping pada Google Play Store
3. Rentang waktu pengambilan data yang digunakan mulai tanggal 19 November 2021 – 16 Oktober 2023 dengan jumlah data yaitu 1000 ulasan.
4. Bahasa yang digunakan adalah Bahasa Indonesia
5. Metode yang digunakan adalah klasifikasi dengan algoritma *Support Vector Machine*.