

## **Analyzing User Sentiments in Motor Vehicle Tax Applications Using the Naïve Bayes Algorithm**

**Wahyudi Ariannor<sup>1\*</sup>, Erwin Arry Kusuma<sup>2</sup>, Fadilah<sup>3</sup>, Muhammad Arsyad<sup>4</sup>**  
 Teknik Informatika, STMIK Banjarbaru, Banjarbaru, Indonesia  
 \*e-mail *Corresponding Author*: wahyu.arian@gmail.com

### **Abstract**

*The South Kalimantan Tax Info application service can provide information about the amount of tax, when you have to pay tax, the due date, and so on. However, the South Kalimantan Tax Info application received many negative reviews from users. So it is necessary to analyze user sentiment using computational techniques. In this context, sentiment analysis is applied using the Naïve Bayes method to user reviews, assisting the Regional Revenue Agency in understanding perceptions and enhancing service quality. The literature review encompasses similar studies that employ the Naïve Bayes algorithm for sentiment analysis in e-government applications. The research methodology involves collecting review data from the Google Play Store through web scraping, labeling based on ratings, and pre-processing. The results of sentiment analysis, utilizing the confusion matrix, demonstrate the highest accuracy of 92% with a 10:90 data split. This study contributes to users' comprehension of public service applications, facilitating continuous improvement.*

**Keywords:** *Sentiment analysis; Confusion matrix; Text mining; Emoticon*

### **Abstrak**

Layanan aplikasi Info Pajak Kalsel dapat memberikan informasi tentang besaran pajak, waktu harus bayar pajak, jatuh tempo dan lain-lain. Namun, aplikasi Info Pajak Kalsel menerima banyak ulasan negatif dari pengguna. Maka diperlukan analisis sentimen pengguna menggunakan teknik komputasi. Dalam konteks ini, analisis sentimen diterapkan menggunakan metode *Naïve Bayes* pada ulasan pengguna, membantu Badan Pendapatan Daerah memahami persepsi dan meningkatkan kualitas layanan. Studi literatur mencakup penelitian sejenis yang menggunakan algoritma *Naïve Bayes* untuk analisis sentimen pada aplikasi *e-Government*. Metodologi penelitian melibatkan pengumpulan data ulasan dari *Google Playstore* melalui *web scraping*, pemberian label berdasarkan rating, dan proses *pre-processing*. Hasil analisis sentimen menggunakan *confusion matrix* menunjukkan akurasi tertinggi sebesar 92% pada pembagian data 10:90. Studi ini memberikan kontribusi pada pemahaman pengguna terhadap aplikasi pelayanan publik, memungkinkan perbaikan berkelanjutan.

**Kata kunci:** *Analisis sentimen; Matriks konfusi; Text mining; Emotikon*

### **1. Pendahuluan**

Pemanfaatan teknologi dan sistem informasi pada institusi pemerintahan melalui *e-Government* dapat meningkatkan kinerja dan kualitas pelayanan publik menjadi lebih akurat, bermanfaat, dan efisien [1]. Pelayanan masyarakat yang ideal tidak memerlukan birokrasi yang berbelit-belit karena hal ini dapat membuat masyarakat enggan untuk berurusan dengan layanan birokrasi, terutama bagi masyarakat yang memiliki kesulitan untuk mendapatkan akses dan jarak ke layanan birokrasi [2].

Badan Pendapatan Daerah Provinsi Kalimantan Selatan memiliki layanan berbasis teknologi yaitu Aplikasi Pajak Kendaraan Bermotor berbasis android yang diberi nama Info Pajak Kalsel. Aplikasi ini tersedia di *Google Playstore* dan telah diunduh lebih dari 100 ribu kali. Namun, dibalik banyaknya pengguna aplikasi tersebut masih terdapat komentar pengguna yang memberikan tanggapan negatif seperti aplikasi yang *error blank*, tidak bisa cek pajak, *error 404 found* dan lain sebagainya. Tanggapan negatif merepresentasikan bentuk ketidaknyamanan

terhadap layanan yang terdapat pada fitur aplikasi. Tanggapan negatif tersebut sudah dapat menjadi bukti bahwa ada ketidaksesuaian antara keadaan sebenarnya dan harapan pengguna.

Tanggapan dari setiap pengguna pada Google *Playstore* dapat berpengaruh pada calon pengguna sebagai bahan pertimbangan penggunaan suatu aplikasi [3]. Tanggapan pengguna yang bersifat tekstual dalam jumlah banyak, akan sangat menyulitkan jika diproses secara manual. Maka, diperlukan penerapan analisis sentimen dengan menerapkan metode komputasi. *Text Mining* adalah metode komputasi yang tepat untuk menyelesaikan studi kasus analisis sentimen. Analisis sentimen terhadap ulasan pengguna membantu pencari aplikasi menemukan aplikasi yang paling sesuai dengan kebutuhan mereka dan membantu pengembang aplikasi mengidentifikasi fitur-fitur lemah yang memerlukan perbaikan [4].

Analisis sentimen pengguna aplikasi Info Pajak Kalsel penting dilakukan untuk mengetahui persepsi pengguna terhadap aplikasi, sehingga dapat digunakan untuk meningkatkan kualitas aplikasi tersebut. Analisis sentimen digunakan untuk menentukan apakah data bersifat positif, negatif, atau netral, yang dapat digunakan untuk mengambil keputusan bisnis berdasarkan opini pengguna [5]. Analisis sentimen diterapkan di berbagai ranah bisnis dan sosial karena opini dapat memengaruhi perilaku individu hingga pengambilan keputusan suatu organisasi [6]. Analisis sentimen dapat dilakukan dengan menggunakan berbagai teknik, salah satunya adalah dengan menerapkan algoritma *Naïve Bayes*. Algoritma *Naïve Bayes* adalah teknik populer untuk analisis sentimen, menggunakan probabilitas bersyarat untuk menetapkan label pada *instance* dalam teks [7].

Metode *Naïve Bayes* yaitu sebuah teknik pengklasifikasian yang menggunakan probabilitas dan statistika memprediksi suatu peluang dengan asumsi yang kuat akan suatu kejadian atau kondisi masing-masing. *Naïve Bayes* menghasilkan suatu pengklasifikasian sebuah opini secara tepat. Opini tersebut dapat berupa paragraf dengan beberapa kalimat baik kalimat positif maupun kalimat negatif [8].

Dengan memahami bagaimana pengguna merasakan dan menilai aplikasi ini, langkah-langkah perbaikan dapat diambil untuk memastikan pelayanan publik yang lebih baik. Dengan hasil penelitian ini, Badan Pendapatan Daerah Provinsi Kalimantan Selatan dapat meningkatkan kualitas aplikasi mereka dan meningkatkan kepuasan pengguna. Selain itu, penelitian ini juga dapat memberikan pandangan yang lebih luas tentang bagaimana teknologi dan pelayanan publik dapat terus berintegrasi dengan baik.

## 2. Tinjauan Pustaka

Beberapa penelitian terdahulu yang relevan dengan penelitian ini yaitu penelitian yang dilakukan oleh [9] yang membahas analisis sentimen aplikasi *E-Government* pada Google *Play* menggunakan algoritma *Naïve Bayes*. Jumlah data yang telah melalui proses pembersihan sebanyak 642 data. Dengan 407 ulasan positif dan 235 ulasan negatif, proses pelabelan menunjukkan bahwa aplikasi Sentuh Tanahku memiliki banyak ulasan positif dari pengguna. Hasil analisis sentimen yang dilakukan menggunakan algoritma *Naïve Bayes* dengan pembobotan TF-IDF menunjukkan akurasi sebesar 89%, *precision* sebesar 83%, dan *recall* sebesar 87%. Perbandingan jumlah pembagian data dalam penelitian ini adalah 90:10 yang terdiri dari 90% data latih atau sebanyak 577 data dan 10% data uji.

Kemudian penelitian analisis sentimen terhadap aplikasi Sistem Administrasi Kendaraan Pajak *Online* Jawa Tengah oleh [10] menggunakan algoritma *Naïve Bayes*. *Dataset* dibagi menjadi 2 bagian 10% sebagai data testing, 90% sebagai data *training*. Simpulan penelitian ini, algoritma *Naïve Bayes* memiliki tingkat akurasi sebesar 88% dari 2414 komentar positif dan 2395 komentar negatif saat melakukan klasifikasi ulasan. Setelah dilakukan pengujian validasi *K-Fold* sebanyak sepuluh kali, ini menunjukkan bahwa algoritma ini dapat melakukan klasifikasi komentar untuk aplikasi yang ada di Google *Playstore*.

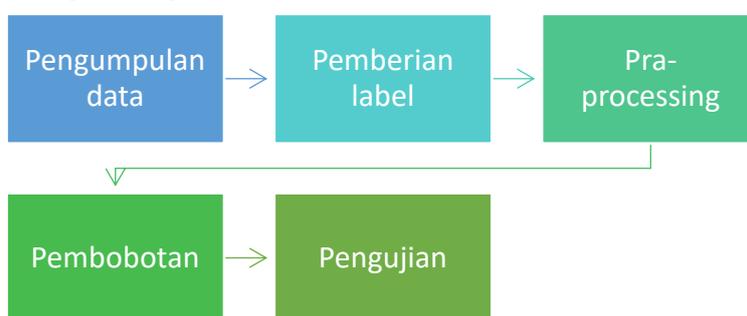
Penelitian yang dilakukan oleh [11] yang membahas analisis sentimen ulasan pengguna aplikasi e-samsat provinsi Jawa Barat menggunakan metode BiGRU. Studi ini melakukan analisis sentimen pada aplikasi Samsat Mobile Jawa Barat (SAMBARA) yang tersedia di Google *Playstore*. Ulasan pengguna dibagi menjadi tiga kelompok: Positif, Negatif, dan Netral. Distribusi data latih dan data pengujian pada penelitian ini adalah 70:30, sehingga 70% data dari setiap kelas digunakan untuk pelatihan sedangkan sisanya 30% digunakan untuk pengujian. *Bidirectional gated recurrent unit* (BiGRU) adalah metode yang dipilih karena mampu memprediksi ulasan pengguna dengan akurasi hingga 87,37%. Metode ini dianggap baik dan dapat membantu pengembangan aplikasi pelayanan di Jawa Barat.

Kemudian penelitian lain yang dilakukan oleh [12] yang membahas tentang analisis sentimen komentar SIAKAD menggunakan metode *Naive Bayes Classifier*. *Dataset* berupa komentar mahasiswa berjumlah 112 komentar. Komentar mahasiswa dikategorikan dan kemudian digunakan untuk melakukan analisis sentimen. Emosi positif dan negatif diklasifikasikan sebagai sentimen. Hasil pengujian akurasi menggunakan metode *Naive Bayes Classifier* menunjukkan bahwa tingkat akurasi sebesar 89,28%.

Dalam penelitian ini, untuk menganalisis ulasan dan *rating* yang diberikan oleh pengguna, algoritma *Naive Bayes* akan digunakan untuk mengetahui sentimen pengguna terhadap aplikasi Info Pajak Kalsel. Berbeda dengan beberapa penelitian terdahulu yang menghapus *emoticon* [9][10][11], pada penelitian ini *emoticon* tidak dihapus melainkan dikonversi menjadi teks dan diterjemahkan ke dalam bahasa Indonesia berbeda dengan penelitian [12].

### 3. Metodologi

Pada penelitian ini terdapat tahapan-tahapan yang dilakukan, yaitu pengumpulan data, pemberian label, *praprocessing* (*casefolding, cleaning, stemming, stopword, tokenizing*), pembobotan dan pengujian algoritma.



Gambar 1. Rahaan Penelitian

#### 3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data ulasan pengguna aplikasi Info Pajak Kalsel yang tersedia di halaman *website* Google *Playstore*. Data diperoleh dengan teknik *web scraping*. *Web scraping* adalah metode komputerisasi untuk memperoleh informasi dari situs Internet, mengubah HTML tidak terstruktur menjadi data terstruktur untuk digunakan dalam beragam aplikasi [13].

Berdasarkan informasi yang tertera pada halaman *website* Google *Playstore*, jumlah ulasan aplikasi sebanyak 1,24 ribu ulasan. Namun setelah dilakukan *scraping*, data ulasan yang diperoleh hanya sebanyak 755 data valid. Yang dimaksud data valid dalam penelitian ini adalah pengguna yang memberikan ulasan lengkap, yaitu mengisikan komentar dan memberikan nilai *rating*. Pengguna yang memberikan ulasan namun tidak memberikan komentar (kosong) tidak digunakan sebagai data penelitian.

#### 3.2. Pemberian Label

Data yang telah diperoleh kemudian diberi label berdasarkan komentar-komentar dari pengguna aplikasi. Pelabelan dalam analisis sentimen adalah proses menentukan pendapat atau perasaan suatu teks dan memberi label (positif, negatif, atau netral) [14].

Pada penelitian ini komentar pengguna diberi label positif dan negatif dengan memanfaatkan nilai *rating* yang diberikan oleh pengguna. Nilai *rating* 1 & 2 diberikan label negatif, sedangkan nilai *rating* 4 & 5 diberikan label positif. Sementara nilai *rating* 3 tidak digunakan karena sifatnya mengarah kepada netral.

#### 3.3. Pre-processing

Pada tahap ini dilakukan normalisasi data penelitian. Menormalkan data web selama *pre-processing* akan meningkatkan kinerja analisis sentimen dengan meningkatkan akurasi dan mengurangi pergeseran polaritas [15]. Tahapan *pre-processing* meliputi *case folding, cleaning text, tokenization, stopwords* dan *stemming*. Tahapan pada *pre-processing* dapat dilihat pada tabel berikut:

Tabel 1. Tahapan pre-processing

No.	Tahap	Keterangan
1	<i>Case folding</i>	Mengubah teks menjadi huruf kecil (lower case)
2	<i>Cleaning text</i>	Membersihkan teks dari karakter-karakter dan kalimat yang tidak diperlukan, seperti tanda baca, simbol, url, <i>emoticon</i> dan lain-lain. Karakter atau kalimat yang mengandung hal tersebut akan dihapus, kecuali <i>emoticon</i> . Untuk <i>emoticon</i> akan diubah atau dikonversi menjadi teks yang mewakili makna dari <i>emoticon</i> tersebut. Contoh: 😊 → senyum, 👍 → jempol ke atas, 😞 → sedih
3	<i>Tokenization</i>	Memisahkan setiap kata dalam kalimat menggunakan tanda koma. Contoh: <i>saya suka dengan aplikasi ini</i> → [ <i>saya, suka, dengan, aplikasi, ini</i> ]
4	<i>Stopwords</i>	Menghapus kata sambung dan kata yang tidak efektif pada kalimat. Contoh: <i>[saya, suka, dengan, aplikasi, ini]</i> → [ <i>saya, suka, aplikasi</i> ]
5	<i>Stemming</i>	Mengonversi kata menjadi kata dasar. Contoh: <i>aplikasi disarankan dikembangkan agar menjadi lebih baik</i> → <i>aplikasi saran kembang agar jadi lebih baik</i>

### 3.4. Pembobotan

Setiap kata kemudian diberikan bobot. Pembobotan menggunakan teknik *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF). TF-IDF dalam analisis sentimen membantu mengurangi jumlah fitur yang tidak relevan dengan memberikan nilai bobot pada fitur berdasarkan nilai korelasinya yang kuat [16].

TF merupakan banyaknya kemunculan sebuah kata dalam suatu dokumen dan jumlahnya sebanding dengan besaran bobot kata tersebut karena semakin sering muncul, maka bobotnya semakin besar. Sedangkan, pada konsep IDF, suatu kata yang sering muncul memiliki bobot kecil [17]. Rumus persamaan TF-IDF sebagai berikut [9]:

$$TF.IDF_{std}(t) = tf_d^t \times \log \frac{n}{df_t} \quad \dots\dots(1)$$

### 3.5. Pengujian

Untuk mengelompokkan perasaan dan ulasan pengguna aplikasi, kemudian dilanjutkan ke tahapan pengujian. Dalam tahapan pengujian terdapat beberapa tahap, yaitu klasifikasi Naive Bayes dan *Matriks konfusi (confusion matrix)*. Dalam proses klasifikasi Naive Bayes digunakan fungsi *Multinomial* Naive Bayes karena jenis klasifikasi ini mampu mengolah data berupa teks [18]. Pengklasifikasi *Naive Bayes* yang ditingkatkan menggunakan model *multinomial* dan komputasi paralel secara signifikan meningkatkan akurasi dan skalabilitas untuk klasifikasi teks skala besar [19].

Pengujian evaluasi performa menggunakan *Matriks konfusi (confusion matrix)* untuk mengukur akurasi, presisi, dan *recall*. *Matriks konfusi (confusion matrix)* adalah tabel yang digunakan untuk menilai kinerja model klasifikasi dalam masalah klasifikasi biner, membandingkan sampel aktual dan prediksi dalam masalah klasifikasi biner [20]. Perhitungan *Matriks konfusi (confusion matrix)* seperti rumus berikut:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad \dots\dots (2)$$

$$Precision = \frac{TP}{TP+FP} \quad \dots\dots (3)$$

$$Recall = \frac{TP}{TP+FN} \quad \dots\dots (4)$$

Keterangan:

TP: *True Positive*

TN: *True Negative*  
 FP: *False Positive*  
 FN: *False Negative*

#### 4. Hasil dan Pembahasan

##### 4.1. Scraping Data

Pada tahap ini peneliti menggunakan Teknik web *scraping* dengan bantuan *library* dalam bahasa pemrograman python dan data dikumpulkan dari halaman Google *Playstore*. Data yang sudah dikumpulkan di simpan dalam format file.csv. Jumlah ulasan yang didapatkan sebanyak 755 data. Data yang diperoleh adalah data pengguna yang memberikan ulasan yang tidak kosong dan memberikan nilai *rating* 1 – 5 terhadap aplikasi.

Rata-rata panjang karakter komentar pada ulasan adalah 54 karakter. Panjang karakter paling banyak yaitu 499 karakter dan paling sedikit 2 karakter. Data yang didapatkan dalam rentang waktu sejak tanggal 06 Oktober 2020 hingga tanggal 30 Oktober 2023. Seperti yang telah dijelaskan pada bab sebelumnya, nilai *rating* ulasan yang bernilai 3 tidak diproses. Maka jumlah data yang tersisa sebanyak 709 data yang akan dilanjutkan ke proses pelabelan data.

##### 4.2. Pelabelan Data

Dari jumlah data sebanyak 709 data. Diberikan label terhadap setiap data ulasan. Label negatif didefinisikan dengan nilai -1, sedangkan label positif didefinisikan dengan nilai 1. Sampel hasil pelabelan data seperti pada tabel berikut:

Tabel 2. Pelabelan data

No.	Comment	Value
1	Sebelum aplikasi eror cek biaya pajak sekian x...	-1
2	Aplikasi ini sangat bagus bisa buat ngecek paj...	-1
3	Sudah update ke versi terbaru (bapenda) tetap ...	-1
4	Ini gimana ya apk nya apa masih error udah lam...	-1
5	Tolong dong secepatnya diperbaiki, sebab ini a...	-1
...	...	...
705	Oke mantap	1
706	Sangat bagus	1
707	Apps oke	1
708	👍👍	1
709	👍👍👍👍👍	1

Rekapitulasi jumlah data negatif dan positif dapat dilihat pada tabel berikut:

Tabel 3. Jumlah ulasan berdasarkan label

No.	Label	Jumlah	%	Keterangan
1	Negatif	425	59,94%	-1
2	Positif	284	40,06%	1
<b>Total</b>		<b>709</b>		

##### 4.3. Pre-processing

Pada tahap ini dilakukan proses *Case folding*, *cleaning text*, *tokenization*, *stopwords* dan *stemming* sebelum dilakukan pengujian.

###### 4.3.1. Case Folding dan Cleaning Text

Tahapan *pre-processing* dengan mengubah teks menjadi huruf kecil (*lower case*) atau *case folding*. Membersihkan teks dari karakter-karakter dan kalimat yang tidak diperlukan, seperti tanda baca, simbol, url, *emoticon* dan lain-lain serta mengubah *emoticon* menjadi teks bermakna.

*Emoticon* yang diubah menjadi teks pada awalnya berbahasa inggris, kemudian diterjemahkan ke dalam bahasa Indonesia. Hal ini dilakukan agar pada proses *stemming* menjadi lebih optimal. Sampel hasil *case folding* dan *cleaning text* seperti tabel berikut:

Tabel 4. Case Folding dan Cleaning Text

No.	comment	value
1	sebelum aplikasi eror cek biaya pajak sekian x...	-1
2	aplikasi ini sangat bagus bisa buat ngecek paj...	-1
3	sudah update ke versi terbaru bapenda tetap aj...	-1
4	ini gimana ya apk nya apa masih error udah lam...	-1
5	tolong dong secepatnya diperbaiki sebab ini ap...	-1
...	...	...
705	oke mantap	1
706	sangat bagus	1
707	aplikasi oke	1
708	jempol ke atas jempol ke atas	1
709	jempol ke atas jempol ke atas jempol ke atas ...	1

#### 4.3.2. Tokenization

Proses yang kemudian dilakukan adalah *tokenization* atau memisahkan setiap kata dalam kalimat menggunakan tanda koma. Sampel hasil *tokenization* seperti pada tabel berikut:

Tabel 5. Tokenization

No.	Comment	value
1	[sebelum, aplikasi, eror, cek, biaya, pajak, s...	-1
2	[aplikasi, ini, sangat, bagus, bisa, buat, nge...	-1
3	[sudah, update, ke, versi, terbaru, bapenda, t...	-1
4	[ini, gimana, ya, apk, nya, apa, masih, error,...	-1
5	[tolong, dong, secepatnya, diperbaiki, sebab, ...	-1
...	...	...
705	[oke, mantap]	1
706	[sangat, bagus]	1
707	[aplikasi, oke]	1
708	[jempol, ke, atas, jempol, ke, atas]	1
709	[jempol, ke, atas, jempol, ke, atas, jempol, k...	1

#### 4.3.3. Stopwords

Tujuan dari stopwords adalah Menghapus kata sambung dan kata yang tidak efektif. Sampel hasil *stopwords* dapat dilihat pada tabel berikut:

Tabel 6. Hasil Stopwords

No.	comment	value	comment (stopwords)
1	[sebelum, aplikasi, eror, cek, biaya, pajak, s...	-1	[aplikasi, eror, cek, biaya, pajak, sekian, xx...
2	[aplikasi, ini, sangat, bagus, bisa, buat, nge...	-1	[aplikasi, bagus, ngecek, pajak, mobilmotor, u...
3	[sudah, update, ke, versi, terbaru, bapenda, t...	-1	[update, versi, terbaru, bapenda, aja, cek, pa...
4	[ini, gimana, ya, apk, nya, apa, masih, error,...	-1	[gimana, ya, apk, nya, error, udah, banget, er...
5	[tolong, dong, secepatnya, diperbaiki, sebab, ...	-1	[tolong, secepatnya, diperbaiki, apk, kalo, em...
...	...	...	...
705	[oke, mantap]	1	[oke, mantap]
706	[sangat, bagus]	1	[bagus]
707	[aplikasi, oke]	1	[aplikasi, oke]
708	[jempol, ke, atas, jempol, ke, atas]	1	[jempol, jempol]
709	[jempol, ke, atas, jempol, ke, atas, jempol, k...	1	[jempol, jempol, jempol, jempol, jempol]







sentimen positif menunjukkan aplikasi Info Pajak Kalsel masih perlu dievaluasi dan dikembangkan lebih lanjut.

## 5. Simpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa untuk analisis sentimen pada ulasan pengguna aplikasi Info Pajak Kalsel, model klasifikasi Naive Bayes dengan pembagian data 10:90 memberikan hasil yang cukup baik dalam hal akurasi, presisi, dan *recall*. Hasil evaluasi performa model menunjukkan akurasi sebesar 92%, presisi sebesar 92%, dan *recall* sebesar 90% dan *f1-score* 91%. Nilai-nilai tersebut menunjukkan bahwa model dapat mengklasifikasikan perasaan pengguna menjadi positif atau negatif dengan cukup baik.

Sebesar 59,94% pengguna aplikasi memberikan sentimen negatif terhadap aplikasi Info Pajak Kalses, sedangkan sentimen positif sebesar 40,06%. Berarti aplikasi Info Pajak Kalsel masih mendapat lebih banyak sentimen negatif oleh pengguna. Analisis sentimen dapat membantu Badan Pendapatan Daerah Provinsi Kalimantan Selatan memahami bagaimana pengguna melihat aplikasi mereka. Langkah-langkah perbaikan dapat diambil untuk memberikan pelayanan publik yang lebih baik berdasarkan tanggapan pengguna. Dengan fokus pada aplikasi berbasis Android, penelitian ini memberikan perspektif yang lebih luas tentang bagaimana teknologi dan pelayanan publik dapat terus berintegrasi dengan baik.

## Daftar Referensi

- [1] A. A. G. S. Utama, "The Implementation of e-Government in Indonesia," *International Journal Of Research In Business And Social Science*, vol. 9, no. 7, pp. 190-196, 2020.
- [2] W. Ariannor, D. R. Safitri dan R. Rahmi, "Analisis Usability Sistem Permohonan Legalisir Menggunakan Metode System Usability Scale," *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 2, pp. 415-424, 2023.
- [3] S. A. Saputra, D. Rosiyadi, W. Gata dan S. M. Husain, "Sentiment Analysis Analysis of E-Wallet Sentiments on Google Play Using the Naive Bayes Algorithm Based on Particle Swarm Optimization," *RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 3, no. 3, pp. 377-382, 2019.
- [4] D. M. K. Oshadi and S. Thelijjagoda, "AppGuider: Feature Comparison System using Neural Network with FastText and Aspect-based Sentiment Analysis on Play Store User Reviews," in *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1148 – 1155, Trichy, 2022.
- [5] A. Kovačević dan Ž. Kovačević, "Alati Za Analizu Sentimenta," *Polytechnic & Design*, vol. 9, no. 3, pp. 167-174, 2021.
- [6] B. Liu, *Sentiment Analysis and Opinion Mining*, Toronto: Springer Nature, 2022.
- [7] F. Kristanto, W.W. Winarno, & A. Nasiri, "Perbandingan Algoritme Naive Bayes dan Decision Tree Pada Analisis Sentimen Data Komentar Siswa Pada Aplikasi Digital Teacher Assessment. *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, vol. 12, no. 2, pp. 538-548, 2023.
- [8] S. Fathimah, B. Pambudhi, & D. Mulyani, "MODEL APLIKASI DIAGNOSIS PENYAKIT ANJING MENGGUNAKAN METODE NAIVE BAYES". In *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2023, Vol. 2, No. 2, pp. 1154-1163.
- [9] A. I. Tanggraeni and M. N. N. Sitokdana, "Analisis Sentimen Aplikasi E-Government Pada Google Play Menggunakan Algoritma Naive Bayes," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 2, pp. 785-795, 2022.
- [10] E. B. Susanto, P. A. Christianto, M. R. Maulana and S. W. Binabar, "Analisis kinerja algoritma naive bayes pada dataset sentimen masyarakat aplikasi NEWSAKPOLE samsat jawa tengah," *Jurnal Computer Science and Information Technology (CoSciTech)*, vol. 3, no. 3, pp. 234-231, 2022.
- [11] R. K. Dewi, B. Tantular, J. Suprijadi and A. A. Ptavitasari, "Analisis Sentimen Ulasan Pengguna Aplikasi E-Samsat Provinsi Jawa Barat Menggunakan Metode BiGRU," in *Seminar Nasional Statistika XI 2022*, pp. 1-8, Surabaya, 2022.

- [12] D. L. Kaka, G. K. Pati dan K. W. Rato, "Analisis Sentimen Komentar SIAKAD Menggunakan Metode Naive Bayes Classifier," *Jurnal Kridatama Sains dan Teknologi*, vol. 05, no. 2, pp. 266-277, 2023.
- [13] A. S. Bale, G. Ghorpade, N. Naveen, R. S, S. Kamalesh, R. R dan R. B. S, "Web Scraping Approaches and their Performance on Modern Websites," dalam *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 956-059, Coimbatore, 2022.
- [14] A. Alkabbabi dan M. Taileb, "Ensemble Learning Sentiment Classification for Un-labeled Arabic Text," dalam *ICC 2019: Advances in Data Science, Cyber Security and IT Applications*, pp. 203-210, Cham, 2019.
- [15] S. K. Johal dan R. Mohana, "Effectiveness of Normalization Over Processing of Textual Data Using Hybrid Approach Sentiment Analysis," *International Journal of Grid and High Performance Computing (IJGHPC)*, vol. 12, no. 3, pp. 43-56, 2020.
- [16] N. P. Ririanti dan A. Purwinarko, "Implementation of Support Vector Machine Algorithm with Correlation-Based Feature Selection and Term Frequency Inverse Document Frequency for Sentiment Analysis Review Hotel," *Scientific Journal of Informatics*, vol. 8, no. 2, pp. 297-303, 2021.
- [17] C. H. Yutika, A. Adiwijaya dan S. A. Faraby, "Analisis Sentimen Berbasis Aspek pada Review Female Daily Menggunakan TF-IDF dan Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 5, no. 2, pp. 422-430, 2021.
- [18] R. Noviana dan I. Rasal, "Penerapan Algoritma Naive Bayes dan SVM Untuk Analisis Sentimenboy Band BTS Pada Media Sosial Twitter," *Jurnal Teknik dan Science*, vol. 2, no. 2, pp. 51-60, 2023.
- [19] H. Chen dan D. Fu, "An Improved Naïve Bayes Classifier for Large Scale Text," dalam *Proceedings of the 2018 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2018)*, pp. 21-24, Chengdu, 2018.
- [20] M. M. Fahmy, "Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial," *Journal of Engineering Research (ERJ)*, vol. 6, no. 5, pp. T1-T10, 2022.