

ANALISIS SENTIMEN MASYARAKAT INDONESIA DI TWITTER TERHADAP SISTEM PERPAJAKAN 'CORETAX' MENGGUNAKAN METODE NAÏVE BAYES

Fathoni, Audia Faradhisa Ansori, Indira Nailah Ramadhani,
Cahya Rahmi Anissa, Shinta Amelia Putri
Sistem Informasi, Universitas Sriwijaya
Jl. Raya Palembang – Prabumulih No. KM. 32, Indralaya Indah,
Kabupaten Organ Ilir, Sumatera Selatan
fathoni@unsri.ac.id

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia terhadap sistem perpajakan digital Coretax melalui media sosial Twitter dengan menggunakan metode Naive Bayes. Data dikumpulkan dari aplikasi X sebanyak 1009 tweet, yang setelah melalui proses pembersihan menghasilkan 899 data bersih. Sentimen dalam tweet diklasifikasikan ke dalam tiga kategori: positif, negatif, dan netral. Proses analisis melibatkan tahapan crawling data, preprocessing teks (tokenizing, stopword removal, dan stemming), serta pembobotan menggunakan TF-IDF sebelum diklasifikasikan oleh algoritma Naive Bayes. Hasil pengujian menunjukkan bahwa model mampu mengklasifikasikan sentimen dengan akurasi sebesar 77%, dengan performa lebih baik pada sentimen negatif dan netral dibandingkan sentimen positif. Ketidakseimbangan jumlah data pada masing-masing kategori menjadi salah satu faktor penyebab ketidaktepatan klasifikasi. Penelitian ini memberikan gambaran awal terhadap persepsi publik mengenai implementasi Coretax, serta menunjukkan potensi penggunaan analisis sentimen sebagai alat bantu evaluasi sistem layanan digital pemerintah. Temuan ini diharapkan dapat menjadi bahan pertimbangan bagi pengembang dan instansi perpajakan untuk meningkatkan kualitas layanan, khususnya dalam aspek pengalaman pengguna, aksesibilitas, dan efisiensi sistem Coretax ke depannya.

Kata Kunci : *Coretax*, Pajak, Naive Bayes, Sentimen, *Twitter*

1. PENDAHULUAN

Pajak menjadi salah satu pilar utama dalam pemasukan negara yang dimanfaatkan untuk mendanai berbagai sektor pembangunan, termasuk bidang pendidikan, kesehatan, hingga infrastruktur. Pengelolaan pajak yang optimal akan berdampak positif terhadap kesejahteraan masyarakat serta mendorong laju pertumbuhan ekonomi nasional [1]. Untuk memastikan pendapatan negara yang tetap optimal, maka diperlukan masyarakat yang patuh dalam membayar pajak melalui sistem perpajakan yang mudah, transparan, dan efisien. Namun, sistem perpajakan di Indonesia masih menghadapi berbagai tantangan, seperti rendahnya tingkat kepatuhan wajib pajak, rumitnya administrasi, serta kurangnya pemahaman masyarakat terhadap kewajiban pajak. Proses pelaporan pajak yang masih sangat kompleks dan banyak memakan waktu membuat masyarakat enggan untuk memenuhi kewajiban pajaknya dengan baik [2]. Selain itu, kurangnya transparansi dalam sistem perpajakan dapat menimbulkan potensi kesalahan dalam pengelolaan pajak, baik di pihak wajib pajak maupun lembaga otoritas yang mengelola perpajakan.

Sistem perpajakan merupakan salah satu elemen penting dalam strategi peningkatan pendapatan negara. Efektivitas dan efisiensi sistem perpajakan sangat menentukan keberhasilan dalam mencapai tujuan bisnis tersebut. Oleh karena itu, pemerintah Indonesia memperkenalkan Coretax atau Coretax Administration System sebagai upaya modernisasi

administrasi perpajakan [3]. Sistem ini dirancang untuk mengintegrasikan berbagai aspek administrasi perpajakan, termasuk pendaftaran, pelaporan, penegakan hukum, dan pembayaran. Dengan implementasi Coretax, diharapkan pengelolaan pajak menjadi lebih efisien serta mengurangi risiko kesalahan dalam administrasi perpajakan. Modernisasi sistem ini juga dinilai mampu meningkatkan kepatuhan wajib pajak melalui pemanfaatan teknologi informasi dan komunikasi yang lebih terstruktur dan terintegrasi.

Meskipun Coretax memiliki berbagai keunggulan dalam otomatisasi pengelolaan pajak, masih terdapat sejumlah kendala yang dialami oleh pengguna. Beberapa permasalahan yang sering dikeluhkan meliputi antarmuka yang kurang responsif serta waktu respon yang cenderung lambat. Masalah-masalah ini memicu diskusi di kalangan masyarakat dan menjadi perhatian utama bagi pengembang sistem. Untuk memperoleh pemahaman yang lebih menyeluruh terkait pengalaman pengguna serta hal-hal yang masi perlu ditingkatkan, diperlukan analisis sentimen terhadap ulasan pengguna terhadap Coretax.

Naive Bayes merupakan salah satu metode yang umum dimanfaatkan dalam proses analisis sentimen. Metode ini memiliki kemampuan dalam mengelola data teks dengan efisiensi yang tinggi, menjadikannya pilihan yang tepat dalam mengklasifikasikan sentimen pengguna terhadap Coretax. Namun, penerapan metode ini masih menghadapi tantangan, terutama dalam meningkatkan ketepatan model. Proses pre-

processing yang kompleks serta kebutuhan untuk menyiapkan data yang optimal sebelum dianalisis menjadi faktor yang perlu diperhatikan dalam penelitian ini (Faisol A, 2024). Keunggulan utama dari Naive Bayes adalah kemampuannya dalam menangani data yang besar serta efisiensi dalam perhitungan, sehingga banyak digunakan dalam berbagai aplikasi pemrosesan bahasa alami. Namun, penerapan metode ini masih menghadapi tantangan, terutama dalam meningkatkan ketepatan model dalam mengklasifikasikan sentimen secara tepat. Sehubungan dengan hal tersebut, riset ini diarahkan untuk mengimplementasikan metode Naive Bayes dengan pendekatan yang lebih akurat dan relevan guna memahami sentimen pengguna terhadap Coretax.

Penelitian berfokus pada pemanfaatan metode Naive Bayes dalam mengklasifikasikan sentimen pengguna Coretax berdasarkan data ulasan yang dikumpulkan melalui Aplikasi X. Aplikasi ini memiliki berbagai fitur seperti otomatisasi penghitungan pajak, integrasi data, serta panduan langkah demi langkah yang bertujuan untuk mengurangi risiko kesalahan dalam pelaporan pajak (Purnomo T, et al., 2025). Melalui penerapan algoritma Naive Bayes, penelitian ini mengklasifikasikan ulasan pengguna menjadi tiga kategori utama, yaitu positif, negatif, dan netral (Anwar, 2022). Hasil analisis ini akan digunakan untuk mengevaluasi kualitas Coretax serta memberikan rekomendasi yang dapat membantu meningkatkan sistem ini. Studi ini bertujuan untuk memperoleh dan menganalisis tanggapan pengguna terhadap aplikasi Coretax melalui data yang diambil dari platform X, menerapkan metode Naive Bayes yang telah dioptimasi untuk mengklasifikasikan sentimen ulasan menjadi positif, negatif, atau netral, mengevaluasi kinerja model dengan metrik seperti akurasi dan presisi, serta memberikan rekomendasi berdasarkan hasil analisis untuk peningkatan kualitas aplikasi Coretax. Dengan penelitian ini, diharapkan sistem Coretax dapat mengalami peningkatan yang signifikan dalam hal pengalaman pengguna, sehingga tujuan modernisasi administrasi perpajakan di Indonesia dapat tercapai secara lebih optimal.

2. TINJAUAN PUSTAKA

Pada bagian ini, penulis menyoroti beberapa penelitian yang sudah dilakukan sebelumnya. Penelitian tersebut berkaitan dengan analisis sentimen melalui aplikasi twitter yang diolah dengan cara menerapkan metode Naive Bayes. Pada penelitian oleh Salim & Solichin, dilakukan penelitian mengenai analisis sentimen pada Twitter yang ditujukan untuk mengevaluasi pelayanan Dina Kependudukan dan Pencatatan Sipil dengan menggunakan metode Naive Bayes. Penelitian berikutnya dilakukan oleh Rahman mengenai analisis sentimen ulasan pengguna aplikasi samsat digital nasional (SIGNAL) dengan menggunakan metode Naive Bayes. Adapun pada penelitian lain yang dilakukan oleh Maulana yang

membahas mengenai analisis sentimen mengenai korupsi bansos beras masa pandemi yang ada pada aplikasi twitter. Pada penelitian tersebut, mereka menggunakan metode SVM (*Support Vector Machine*) dan juga Naive Bayes. Data yang dianalisis didapat melalui *crawling twitter*. Penelitian yang telah dilakukan tersebut dikelompokkan menjadi 2, yaitu positif dan negatif.

Penelitian terdahulu yang menerapkan metode Naive Bayes menunjukkan bahwa, metode tersebut pertama kali dilakukan dengan pengambilan data (*crawling data*). Pengambilan data tersebut didapatkan melalui unggahan *tweet* di platform daring yang dikumpulkan berdasarkan kata kunci tersebut. Setelah berhasil, akan menampilkan notifikasi jumlah data *tweet*. Hasil dari proses pengumpulan data akan menghasilkan file dalam format *Microsoft Excel*. Dataset yang telah diperoleh selanjutnya akan melalui tahap pembersihan data. Setelah itu, dilanjutkan dengan pemberian label *tweet* yang diuji. Sedangkan, pada penelitian sebelumnya yang menggunakan metode SVM (*Support Vector Machine*) memiliki proses yang hampir sama dengan Naive Bayes. Perbedaannya terdapat pada tahapan mengklasifikasi sentimen *tweet* yang diuji. Metode SVM bekerja dengan menentukan *hyperlane* optimal yang mampu membedakan dua kategori data secara maksimal, guna memperoleh margin selebar mungkin demi memperoleh tingkat akurasi klasifikasi yang lebih tinggi.

Berdasarkan pendekatan pada studi terdahulu oleh Salim dan Solichin metode Naive Bayes menghasilkan akurasi sebesar 89%, dengan 43,6% ulasan bersentimen negatif 36,4 positif, dan 20% netral pada telaan opini publik di Twitter mengenai layanan administrasi kependudukan dan pencatatan sipil. Sementara itu, dalam penelitian oleh Rahman, dilakukan dilakukan penilaian sentimen terhadap ulasan pengguna aplikasi Samsat Digital Nasional (SIGNAL) menggunakan pendekatan algoritma Naive Bayes, tetapi jurnal ini tidak mencantumkan persentase detail untuk masing-masing kategori sentimen, hanya menyajikan evaluasi model berdasarkan akurasi dan performa klasifikasi. Selanjutnya, pada penelitian oleh Maulana et al., (2024), dilakukan analisis sentimen terhadap kasus korupsi bansos beras melalui platform *Twitter* dengan pendekatan Naive Bayes dan Support Vector Machine (SVM), dengan hasil 58,79% *tweet* memiliki sentimen negatif, 28,41% positif, dan 12,8% netral, serta akurasi 66,67% untuk SVM dan 60,61 untuk Naive Bayes, menunjukkan bahwa SVM lebih unggul dalam menangani data teks kompleks.

Berdasarkan penelitian-penelitian sebelumnya, metode Naive Bayes telah terbukti banyak digunakan dalam analisis sentimen di berbagai bidang, seperti pelayanan publik, aplikasi digital, dan isu sosial. Penelitian yang dilakukan oleh Salim & Solichin serta Rahman mengindikasikan bahwa pendekatan ini memiliki kemampuan yang layak dalam

mengidentifikasi sentimen secara akurat. Meskipun dalam penelitian Maulana pendekatan SVM menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan metode Naive Bayes, tetapi metode Naive Bayes tetap menjadi pilihan populer karena sederhana, cepat, dan efisien dalam mengelola data besar. Selain itu, Naive Bayes masih bisa bekerja dengan baik meskipun data yang dianalisis mengandung banyak *noise* seperti *typo*, *slang* dan emoji yang sering ditemukan dalam *tweet*.

Dalam penelitian ini menggunakan algoritma Naive Bayes karena dinilai efisien dalam mengolah data *Twitter* yang berjumlah Besar. Selain itu pendekatan ini relatif lebih sederhana dalam implementasinya dibandingkan metode klasifikasi lainnya, sehingga memudahkan dalam memahami pola sentimen yang muncul. Meskipun ada metode lain yang mungkin memiliki akurasi lebih tinggi, metode Naive Bayes tetap menjadi pilihan yang tepat karena lebih fleksibel dan tidak terlalu bergantung pada banyak fitur rumit. Dengan mempertimbangkan keterbatasan data serta tujuan penelitian ini, penggunaan Naive Bayes diharapkan mampu menghasilkan output yang konsisten serta memiliki tingkat keandalan yang tinggi.

Penelitian ini memiliki *novelty* karena fokus pada analisis sentimen masyarakat Indonesia terhadap sistem perpajakan digital *Coretax* melalui *Twitter*, yang belum banyak dikaji dalam studi sebelumnya. Sejauh ini, penelitian lebih banyak membahas layanan publik lain seperti Samsat Digital Nasional (SIGNAL), sedangkan studi khusus mengenai *Coretax* masih sangat sedikit. Analisis ini dapat memberikan wawasan baru bagi pemerintah dan pengembang sistem perpajakan digital dalam meningkatkan layanan serta memahami persepsi publik secara lebih mendalam. Selain itu, penelitian ini juga mengeksplorasi efektivitas metode Naive Bayes dalam bidang perpajakan digital yang masih jarang dieksplorasi, sehingga dapat berpotensi menjadi acuan bagi penelitian-penelitian mendatang dalam ranah yang sejenis.

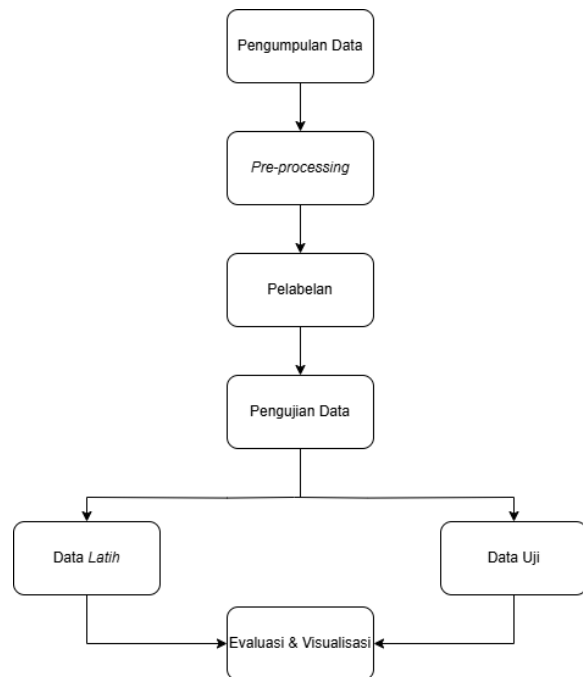
3. METODE PENELITIAN

Tahapan metode berupa mencakup langkah-langkah untuk melakukan analisis sentimen masyarakat Indonesia di *Twitter* terhadap sistem perpajakan yang bernama *Coretax*. Pendekatan yang dilakukan yaitu kuantitatif dengan menggunakan metode *data mining* yang berbasis klasifikasi teks.

3.1. Pengumpulan Data

Tahap pertama yang perlu dilakukan adalah dengan mengumpulkan data. Pengumpulan data didapatkan dari *tweet* pada aplikasi X dengan menggunakan kata kunci maupun *hashtag* yang berkaitan dengan *Coretax* kemudian dilakukan proses *crawling* untuk mendapatkan *dataset*. Proses *crawling* ini memanfaatkan API *Twitter* dengan alat bantu *Tweet Harvest* yang berbasis *javascript* untuk

mengumpulkan data dengan cara mengakses data *Twitter* menggunakan API token akun. Lalu, data tersebut diproses dan disajikan dalam format tabel menggunakan *Python*.



Gambar 1. Langkah-langkah Penelitian

3.2. Pre-processing

Pada langkah ini, dilakukan penyaringan seperti pembuangan dan perbaikan kata agar teks yang dihasilkan lebih terstruktur untuk digunakan dalam tahap selanjutnya. Proses ini meliputi pembersihan teks (*cleaning text*) dari simbol, angka, URL, *retweet*, *hashtag*, spasi berlebih, dan elemen tidak penting lainnya [9].

3.3. Pelabelan

Langkah pelabelan ini dilakukan setelah proses penyaringan, perbaikan dan pembuangan kata-kata yang terdapat unsur tidak pantas pada proses sebelumnya. Pelabelan dilakukan dengan cara membaca satu per satu *tweet* yang sudah didapat sebelumnya kemudian ditentukan label sentimen berupa positif, negatif, ataupun netral.

3.4. Pengujian Data

Pada langkah pengujian data ini, data dibagi menjadi 2, yaitu data latih dan data uji. Pada penelitian ini, jumlah untuk data latih sebanyak 719 data dan jumlah data uji sebanyak 167 data. Pengujian data ini menggunakan perangkat lunak Google Colab sebagai *tools* untuk mengukur akurasi data.

3.5. Evaluasi & Visualisasi

Pada tahapan terakhir ini, dilakukan evaluasi dan juga visualisasi hasil dari tahapan pengujian data. Evaluasi merupakan tahapan yang dilalui untuk dapat mengetahui bagaimana kinerja dari model yang

diusulkan [10]. Tahapan evaluasi dan visualisasi ini memberikan hasil berupa persen akurasi yang didapatkan dan gambaran berupa *confusion matrix*.

Pada tahapan terakhir ini, dilakukan evaluasi dan juga visualisasi hasil dari tahapan pengujian data. Evaluasi merupakan tahapan yang dilalui untuk dapat mengetahui bagaimana kinerja dari model yang diusulkan [10]. Tahapan evaluasi dan visualisasi ini memberikan hasil berupa persen akurasi yang didapatkan dan gambaran berupa *confusion matrix*.

4. HASIL DAN PEMBAHASAN

Dalam analisis sentimen, terdapat beberapa tahapan yang penulis lalui. Berdasarkan metode penelitian yang sudah dirancang, tahapan pertama yang dilakukan adalah pengumpulan data. Data yang penulis gunakan dalam penelitian ini dikumpulkan melalui aplikasi X sebanyak 1009 data. Pengumpulan data dilakukan kurang lebih lima kali sampai akhirnya penulis mendapatkan jumlah data yang ingin dikumpulkan. Data tersebut dikumpulkan dari kurun waktu 1 Januari 2025 sampai 13 April 2025. Setelah dikumpulkan, tahapan berikutnya ada pembersihan data menggunakan tools Google Colab. Setelah data dibersihkan dari adanya *hashtag*, huruf kapital, angka, tanda baca. Data yang melalui pembersihan jadi hanya sebanyak 899 data.

```
[ ] #@title Twitter Auth Token

twitter_auth_token = '6b97c4bc92d3e2aa588f9d2b872027228a94a9b9' # change this auth token
```

Gambar 2. Kode Penyimpanan Token Autentikasi Twitter

[illegible]

Gambar 3. Instalasi Pandas dan Node

```
# Crawl Data

filename = 'coretaxnew.csv'
search_keyword = 'coretax lang:it'
limit = 1000

[mpex -v tweet-harvester@2.6.1 -o "{{filename}}" -s "{{search keyword}}" --tab "A1A3B" -l {{limit}} --token {{twitter auth token}}
```

Gambar 4. Perintah Crawling Tweet

```
#Cek jumlah yang didapatkan

num_tweets = len(df)
print(f"Jumlah tweet dalam dataframe adalah {num_tweets}.")
```

Gambar 5. Menampilkan Data yang Berhasil Diambil

```

Akurasi: 0.7784431137724551

Classification Report:
              precision    recall  f1-score   support

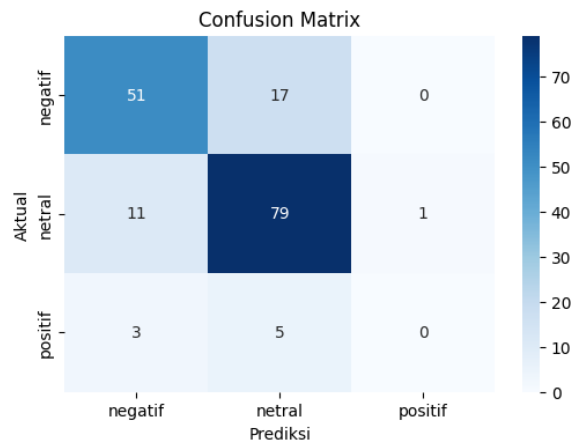
negatif      0.78         0.75         0.77         68
netral       0.78         0.87         0.82         91
positif      0.00         0.00         0.00          8

accuracy                    0.78         167
macro avg      0.52         0.54         0.53         167
weighted avg   0.75         0.78         0.76         167

```

Gambar 6. Hasil akurasi

Berdasarkan hasil akurasi pada gambar 6, dapat dilihat bahwa data yang digunakan memiliki nilai keakuratan sebesar 0,77 atau 77%. Dengan total nilai *support* negatif sebesar 68, *support* netral dan *support* positif sebesar 8. Setelah dilakukan analisis terkait data yang didapatkan melalui twitter dan diolah menggunakan metode naive bayes, data tersebut mengandung banyak pernyataan netral.



Gambar 7. Hasil *Confusion Matrix*

Gambar di atas merupakan hasil dari *confusion matrix* dari pengujian yang sudah dilakukan. Berdasarkan matrix tersebut ssdapat diperoleh informasi mengenai performa model dalam mengelompokkan data uji ke dalam tiga kategori sentimen, yaitu negatif, netral, dan positif. Dari matriks tersebut dapat dilihat bahwa terdapat 79 data aktual netral yang berhasil diprediksi dengan benar, sementara sisanya adalah salah dalam pengelompokan kategori negatif yang berjumlah 51 data dan positif yang berjumlah 3 data.

Untuk kategori sentimen negatif, model menunjukkan performa yang cukup baik dengan 51 data berhasil diprediksi dengan benar. Tetapi, terdapat 17 data negatif yang salah dikelompokkan sebagai netral dan tidak ada data negatif yang diprediksi sebagai positif. Sementara itu, pada kategori sentimen positif, model belum mampu melakukan prediksi dengan baik. Dari total data uji yang termasuk dalam kategori positif, tidak ada satupun yang berhasil diprediksi dengan benar. Sebanyak 3 data positif dikelompokkan sebagai negatif dan 5 data sebagai netral. Hal ini menunjukkan bahwa model masih kesulitan dalam mengenali karakteristik data dengan sentimen positif. Kondisi ini dapat terjadi karena jumlah data positif dalam dataset relatif lebih sedikit dibandingkan dengan kategori lainnya, sehingga model tidak memiliki cukup informasi atau pola untuk belajar secara optimal dalam mendeteksi sentimen positif.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menganalisis sentimen masyarakat Indonesia terhadap sistem perpajakan digital *Coretax* melalui media sosial *Twitter* dengan menggunakan metode *Naive Bayes Classifier*. Berdasarkan hasil pengumpulan dan pengolahan data, diketahui bahwa metode ini mampu mengklasifikasikan ulasan masyarakat ke dalam tiga kategori sentimen, positif, negatif dan netral. Dari total data yang dikumpulkan 1009 *tweet*, dilakukan tahap pra-pemrosesan sehingga diperoleh 899 data bersih yang kemudian digunakan untuk pelatihan dan pengujian model. Hasil pengujian menunjukkan bahwa akurasi model *Naive Bayes* mencapai 77%, yang menunjukkan performa cukup baik dalam memprediksi sentimen terhadap *Coretax*. Model ini menunjukkan kemampuan klasifikasi yang cukup kuat pada sentimen negatif dan netral, tetapi masih lemah dalam mengklasifikasikan sentimen positif.

DAFTAR PUSTAKA

- [1] A. Purwitasari, B. Mutafarida, and Yuliani, "Urgensi Pajak dalam Mendorong Pembangunan Infrastruktur dan Pertumbuhan Ekonomi di Indonesia," *Jurnal Ilmiah Ekonomi dan Manajemen*, vol. 2, no. 6, pp. 656–666, Jun. 2024.
- [2] R. S. Aliyudin, E. F. Ahmad, and N. Nizhan, "PENGARUH SISTEM PERPAJAKAN, DISKRIMINASI, TEKNOLOGIDAN INFORMASI PERPAJAKAN TERHADAP PERSEPSI WAJIB PAJAK MENGENAI PENGELAPAN PAJAK (Studi Pada Wajib Pajak Orang Pribadi di Universitas Majalengka)," *Jurnal Akuntansi Keuangan dan Sistem Informasi*, vol. 2, no. 2, pp. 182–200, Feb. 2021.
- [3] M. Panjaitan and Yuna, "Pengaruh *Coretax* Terhadap Transparansi dan Akuntabilitas Sistem Perpajakan," *Jurnal Riset Akuntansi*, vol. 2, no. 4, pp. 51–60, 2024.
- [4] M. I. Zuhdi, D. Suryadi, and Yuniati, "Pengaruh Modernisasi Sistem Administrasi Dan Sanksi Perpajakan Terhadap Kepatuhan Wajib Pajak Orang Pribadi Pada KPP Pratama Bandung X," *Jurnal Ilmiah Manajemen Ekonomi dan Akuntansi*, vol. 3, no. 1, pp. 116–135, 2020.
- [5] K. Anwar, "Analisa Sentimen Pengguna Instagram di Indonesia Pada Review Smartphone Menggunakan *Naive Bayes*," *Kajian Ilmiah Informatika dan Komputer*, vol. 2, no. 4, pp. 148–155, 2022.
- [6] E. Salim and A. Solichin, "ANALISIS SENTIMEN PADA MEDIA SOSIAL TWITTER TERHADAP PELAYANAN DINAS KEPENDUDUKAN DAN PENCATATAN SIPIL MENGGUNAKAN ALGORITMA NAIVE BAYES," *Indonesia Journal Information System (IDEALIS)*, vol. 5, no. 2, pp. 79–86, Jul. 2022.
- [7] I. F. Rahman, A. N. Hasanah, and N. Heryana, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI SAMSAT DIGITAL NASIONAL (SIGNAL) DENGAN MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, pp. 963–969, Apr. 2024.
- [8] B. A. Maulana, R. A. Fauzi, R. I. Agustin, S. A. Azhaar, and T. Rohana, "KOMPARASI ALGORITMA NAIVE BAYES DAN SVM UNTUK ANALISIS SENTIMEN TWITTER KORUPSI BANSOS BERAS MASA PANDEMI," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, pp. 912–918, Apr. 2024.
- [9] S. Shevira, I. M. Suarjaya, and P. Buana, "Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia," *JITTER- Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, 2022.
- [10] M. I. Fikri, T. S. Sabrila, and Y. Azhar, "Perbandingan Metode *Naive Bayes* dan Support Vector Machine pada Analisis Sentimen Twitter," *Jurnal Stiki Informatika*, vol. 10, no. 02, pp. 71–76, 2020.