

Analisis keluhan publik terhadap Coretax melalui N-Gram pada komentar instagram kementerian keuangan

Ibrahim Kholilullah*

Magister Manajemen, Universitas Gajayana Malang, Indonesia

*) Korespondensi (e-mail: kholilullahibrahim@gmail.com)

Abstract

This study analyzes public complaints regarding the implementation of the coretax system based on Instagram comments from the Ministry of Finance of the Republic of Indonesia posted on October 16, 2025. This study addresses a research gap by applying an N-Gram approach that is exploratory, fast, and efficient in extracting patterns of public complaints specifically from Instagram, which has different interaction characteristics compared to other platforms such as Twitter or online forums. A total of 192 comments were collected through web scraping. Data processing was conducted through text preprocessing stages, including case folding, normalization, stopwords removal, tokenization, and stemming. The analysis employed the N-Gram method (unigram, bigram, and trigram) using Python to identify dominant complaint patterns. The results indicate that complaints are primarily related to technical issues such as system errors, access difficulties, and administrative constraints. In addition, the analysis identified increased user workload and dissatisfaction with tax services, particularly near reporting deadlines.

Keywords: Coretax, N-Gram, Comments, Instagram.

Abstrak

Penelitian ini menganalisis keluhan publik terhadap implementasi sistem coretax berdasarkan komentar Instagram dari Kementerian Keuangan Republik Indonesia pada 16 Oktober 2025. Penelitian ini mengisi kesenjangan penelitian dengan pendekatan N-Gram yang bersifat eksploratif, cepat dan efisien untuk mengekstraksi pola keluhan publik secara spesifik pada Instagram, yang memiliki karakteristik interaksi berbeda dibandingkan platform lain seperti Twitter atau forum berita. Sebanyak 192 komentar dikumpulkan melalui web scraping. Pengolahan data dilakukan melalui tahapan *text pre-processing* yang meliputi *case folding*, normalisasi, *stopword removal*, *tokenisasi*, dan *stemming*. Analisis menggunakan metode N-Gram (unigram, bigram, trigram) berbasis *Python* untuk mengidentifikasi pola keluhan dominan. Hasil menunjukkan bahwa keluhan didominasi oleh masalah teknis seperti sistem *error*, kesulitan akses, dan kendala administrasi. Selain itu, ditemukan peningkatan beban pengguna serta ketidakpuasan terhadap layanan pajak, terutama menjelang batas pelaporan.

Kata kunci: Coretax, N-Gram, Komentar, Instagram

How to cite: Kholilullah, I. (2026). Analisis keluhan publik terhadap Coretax melalui N-Gram pada komentar instagram kementerian keuangan. *Journal of Accounting and Digital Finance*, 6(2), 157–169. <https://doi.org/10.53088/jadfi.v6i2.3109>



1. Pendahuluan

Digitalisasi sistem perpajakan di Indonesia telah membawa perubahan signifikan, salah satunya adalah penggunaan coretax (Budiman et al., 2023). Coretax merupakan teknologi yang digunakan untuk mengelola seluruh administrasi perpajakan mulai dari pendaftaran, pembayaran, pengawasan dan pemeriksaan, yang tujuan utama dapat meningkatkan transparansi, efisiensi waktu, dan akurasi data, memungkinkan proses mempercepat penerimaan pajak. Penggunaan teknologi coretax memungkinkan semua proses administrasi perpajakan berjalan secara otomatis, seperti e-filing, e-payment, dan pengawasan, kewajiban pajak secara real-time (Korat & Munandar, 2025).

Coretax mulai beroperasi di Indonesia pada tahun 2024, penerapan coretax telah diatur dalam Peraturan Presiden No. 40/2018 tentang modernisasi sistem administrasi perpajakan (Dimethio et al., 2023). Pada pelaksanaannya, sistem coretax mengalami banyak kegagalan teknis yang menyebabkan banyak keluhan dari penggunaanya seperti, kesalahan dalam proses login, pembuatan faktur pajak, hingga ketidakmampuan sistem untuk menangani volume transaksi yang tinggi (Erstiawan, 2025).

Keluhan-keluhan ini banyak yang diungkapkan pada media sosial yang memperburuk citra dari lembaga terkait. Merespon hal tersebut Kementerian Keuangan Republik Indonesia mengeluarkan nomor kontak aduan mengenai pajak dan bea cukai yang dipublikasikan melalui akun Instagram kementerian resmi pada tanggal 16 Oktober 2025 (Gambar 1). Banyak keluhan mengenai coretax atau pajak mengisi kolom komentar postingan tersebut. Hal ini memungkinkan bisa dilakukan analisis teks untuk memahami masalah mayoritas yang dikeluhkan pengguna. Salah satu teknik sederhana namun efektif untuk ekstraksi pola keluhan adalah menggunakan analisis N-Gram.



Gambar 1. Postingan publikasi nomor kontak aduan mengenai pajak dan bea cukai di Instagram pada tanggal 16 Oktober 2025.

Sumber: <https://www.instagram.com/p/DP3QXbQE0vn/>

Analisis N-Gram merupakan analisis yang digunakan untuk memecah teks panjang menjadi teks yang lebih pendek, dengan salah satu tujuannya adalah untuk mencari rangkaian kata paling populer yang merupakan opini publik di media (Rofiq & Sujak, 2024). Jenis-jenis N-Gram yang umum digunakan dalam proses analisis bahasa untuk memahami hubungan antara kata-kata dalam teks antara lain; *unigram*, *bigram*, *trigram* (Nurhidayat & Dewi, 2023). *Unigram* terdiri dari satu kata, *bigram* terdiri dari dua kata dan *trigram* terdiri dari tiga kata (Ramadhan et al., 2021; Nurhidayat & Dewi, 2023).

Beberapa penelitian sebelumnya yang mengkaji keluhan coretax antara lain; Erstiawan (2025), menggunakan metode analisis konten (*Content Analysis*) dengan sumber data diperoleh dari artikel-artikel media massa online seperti CNBC Indonesia, Detik Finance, Bisnis Indonesia, DDTC News, Kompas.com, Tempo.com, Tribunnews, dan Diorama. Penelitian ini menjelaskan bahwa aplikasi coretax secara konsisten dilaporkan mengalami gangguan serius seperti *error*, kesulitan akses dan ketidakstabilan, yang berimplikasi langsung pada peningkatan beban administratif, hambatan operasional bagi pelaku usaha dan konsultan pajak, serta meluasnya keluhan publik. Aldo et al., (2025) menggunakan metode *netnografi* dengan sumber data melalui beragam platform di social media Instagram, Tiktok, Twitter, dan Facebook. Penelitian ini menjelaskan, coretax dinilai belum siap secara operasional, dengan keluhan utama meliputi sistem yang sering mengalami *error*, antarmuka yang tidak ramah pengguna, dan proses yang membingungkan.

Meskipun berbagai metode telah digunakan dalam analisis keluhan publik berbasis teks, seperti Support Vector Machine (SVM) dan analisis sentiment (Ahmad et al., 2018; Khan et al., 2024; Trisnawati & Wibowo, 2024; Rahmadina et al., 2025; Cahyani & Prasetyaningrum, 2026). Metode seperti SVM dan analisis sentimen umumnya berfokus pada klasifikasi polaritas opini (misalnya positif, negatif, atau netral), sehingga kurang mampu mengungkap konteks spesifik dari keluhan yang disampaikan pengguna. Sebaliknya, analisis N-Gram dapat mengekstraksi rangkaian kata (dua hingga tiga kata) yang lebih kontekstual (Schonlau et al., 2017), sehingga mampu merepresentasikan makna keluhan secara lebih utuh dan informatif. Analisis N-Gram menjadi relevan karena mampu mengungkap pola keluhan secara langsung tanpa memerlukan proses pelabelan data yang memakan waktu (Ojo et al., 2021; Rofiq & Sujak, 2024).

Urgensi penelitian ini juga diperkuat oleh dampak signifikan dari implementasi sistem perpajakan digital terhadap kepatuhan wajib pajak dan stabilitas penerimaan negara. Gangguan pada sistem seperti coretax tidak hanya berdampak pada aspek teknis, tetapi juga berpotensi menurunkan tingkat kepercayaan publik terhadap otoritas pajak dan menghambat efektivitas kebijakan fiskal (Ariyanto et al., 2024). Oleh karena itu, pemahaman yang cepat dan akurat terhadap keluhan publik menjadi krusial sebagai dasar evaluasi dan perbaikan sistem.

Kebaruan (*novelty*) penelitian ini terletak pada penggunaan analisis N-Gram sebagai pendekatan eksploratif yang cepat dan efisien untuk mengekstraksi pola

keluhan publik secara spesifik pada platform media sosial Instagram, yang memiliki karakteristik interaksi berbeda dibandingkan platform lain seperti Twitter atau forum berita. Selain itu, fokus pada data komentar Instagram resmi instansi pemerintah memberikan kontribusi empiris yang lebih kontekstual dalam memahami respons masyarakat terhadap kebijakan digitalisasi perpajakan. Tujuan penelitian ini adalah mengidentifikasi masalah paling sering dikeluhkan pengguna coretax dan pajak melalui komentar Instagram. Penelitian ini diharapkan dapat membantu pemerintah memahami keluhan publik secara lebih terstruktur. Secara akademis penelitian ini diharapkan juga dapat menambah literatur tentang text mining di sektor publik dan memberikan contoh penerapan N-Gram pada data media sosial.

2. Metode Penelitian

Pengumpulan data

Pengumpulan data dilakukan dengan metode *web scraping*. *Web scraping* merupakan teknik pengambilan atau ekstraksi data atau informasi secara otomatis langsung dari situs yang memudahkan dalam mencari data atau informasi yang dibutuhkan di dalam internet (Dzulkarnain et al., 2024). Sampel data merupakan ujaran pada Instagram Kementerian Keuangan Republik Indonesia @kemenkeuri. Ujaran yang diambil menjadi sampel, merupakan komentar-komentar yang ada dalam di postingan publikasi nomor kontak aduan mengenai pajak dan bea cukai pada tanggal 16 Oktober 2025. lebih jelasnya terlihat pada Gambar 1 dan pada alamat; <https://www.instagram.com/p/DP3QXbQE0vn/>. Proses *web scraping* dilakukan pada tanggal 16 Desember 2025 dengan menggunakan bahasa pemrograman *Python*. Data yang telah diperoleh disimpan dengan format .csv kemudian dilakukan penyaringan. Proses pengumpulan data melalui web scraping dilakukan menggunakan *library* yang terintegrasi dengan *Python* (misalnya “*requests*” dan “*BeautifulSoup*”) (Bhardwaj et al., 2026), kemudian data disimpan dan dikelola menggunakan “*pandas*” dan “*numpy*” (Harris et al., 2020).

Proses penyaringan dan klasifikasi data dilakukan secara manual dengan membaca seluruh isi komentar yang telah dikumpulkan. Komentar atau ujaran diklasifikasikan ke dalam tiga kategori utama, yaitu (1) keluhan terkait perpajakan, khususnya implementasi sistem coretax, (2) keluhan terkait layanan bea dan cukai, serta (3) dukungan dan penilaian terhadap menteri keuangan “Purbaya Yudhi Sadewa”. Hanya data ujaran mengenai pajak dan coretax saja yang digunakan dalam penelitian ini.

Justifikasi pemilihan sampel didasarkan pada relevansi konteks, di mana postingan yang dianalisis merupakan kanal resmi aduan publik sehingga komentar yang muncul mencerminkan respons langsung masyarakat terhadap layanan perpajakan dan bea cukai. Pemilihan metode penyaringan manual dilakukan karena karakteristik bahasa pada media sosial cenderung informal, ambigu, dan kontekstual, sehingga membutuhkan pemahaman semantik yang lebih mendalam yang sulit dicapai melalui pendekatan otomatis. Dengan demikian, pendekatan ini diharapkan mampu meningkatkan validitas dan ketepatan klasifikasi data dalam penelitian.

Pemilihan postingan tanggal 16 Oktober 2025 sebagai sumber data didasarkan pada beberapa pertimbangan. Pertama, postingan tersebut merupakan publikasi awal terkait nomor kontak aduan yang disampaikan melalui akun resmi Kementerian Keuangan Republik Indonesia, sehingga memiliki nilai penting sebagai titik awal respons publik terhadap kebijakan tersebut. Kedua, dibandingkan dengan postingan serupa lainnya, unggahan ini memiliki jumlah komentar tertinggi pada saat proses web scraping dilakukan pada 16 Desember 2025, sehingga dianggap lebih representatif dalam menggambarkan variasi dan intensitas keluhan publik. Ketiga, secara eksplisit postingan tersebut memuat informasi terkait aduan perpajakan dan bea cukai, sehingga relevan secara langsung dengan fokus penelitian. Berdasarkan pertimbangan tersebut, pemilihan sampel ini diharapkan mampu merepresentasikan dinamika keluhan publik secara lebih komprehensif dan kontekstual.

Metode Analisis

Sebelum melakukan analisis N-Gram, data yang terkumpul harus melewati tahap *text pre-processing* terlebih dahulu agar dapat di proses. Semua proses analisis dilakukan dengan menggunakan bahasa pemrograman *Python*.

Pre-processing

Pre-processing merupakan tahap awal untuk mengubah data menjadi sesuai dengan format yang dibutuhkan agar data siap untuk dianalisis (Mubaroroh et al., 2022). Proses *pre-processing* terdiri dari beberapa tahapan, yaitu *case folding*, normalisasi kata, *stopword removal*, *tokenizing*, *stemming* (Priyatno & Firmananda, 2022; Shevira et al., 2022).

Case folding, mengganti seluruh kata menjadi alfabet kecil (Priyatno & Firmananda, 2022). Normalisasi kata, mengubah kata-kata yang memiliki ejaan tidak standar atau tidak masuk dalam kamus KBBI dan singkatan menjadi bentuk yang lebih baku (Shevira et al., 2022). *Stopword removal*, menghapus kata-kata umum yang tidak memiliki arti signifikan dalam analisis seperti yang, dan, atau, di (Saputra & Wahyuni, 2024). Proses ini membantu model fokus pada kata-kata yang lebih bermakna. *Tokenizing*, memecah teks ulasan menjadi kata-kata tunggal dan *stemming*, mengubah kata-kata menjadi bentuk dasar mereka (Al-Husna et al., 2024). Dalam fungsi *stemming* juga membersihkan kata dari *hashtag*, mention dan emoticon. Contoh proses *pre-processing* terdapat pada Tabel 1.

Tabel 1. Contoh proses *pre-processing*

Ujaran	Paaak coretax gak bisa diakses, hr terakhir lapor SPT pot put nihi 🙄
<i>Case folding</i>	Paaak coretax gak bisa diakses, hr terakhir lapor spt pot put nihi 🙄
Normalisasi kata	pak coretax tidak bisa diakses, hari terakhir lapor spt pot put nih 🙄
<i>Stopword removal</i>	pak coretax tidak bisa diakses, hari terakhir lapor spt 🙄
<i>Tokenizing</i>	Pak, coretax, tidak, bisa, diakses, hari, terakhir, lapor, spt, 🙄
<i>Stemming</i>	coretax tidak bisa akses hari akhir lapor spt

Analisis N-Gram

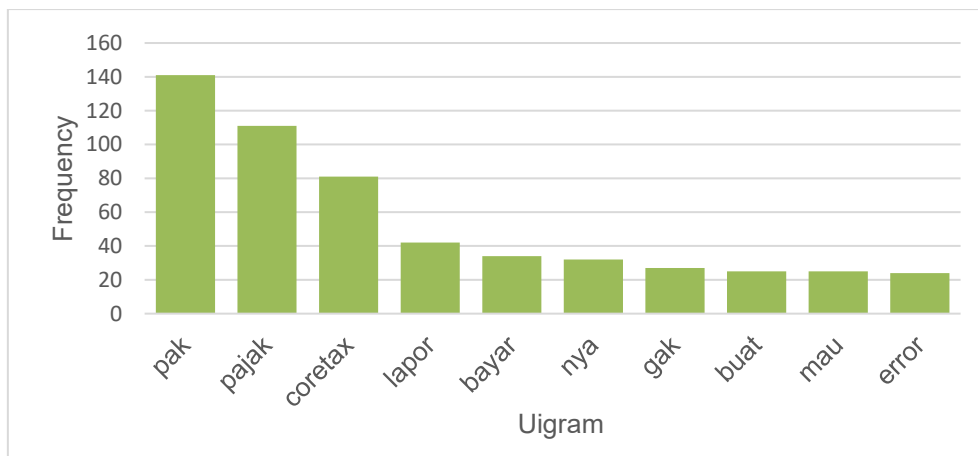
3. Hasil dan Pembahasan

[illegible]

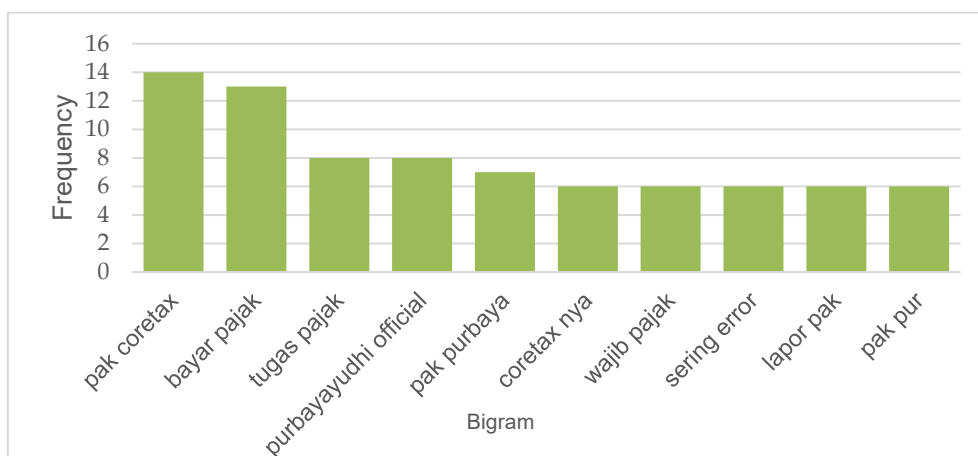
Visualisasi kata dari data yang tersedia menggunakan *WordCloud* dapat terlihat pada (Gambar 2). Dimana kata “coretax”, “pak”, “pajak” dan “lapor” menjadi kata-kata

yang sering muncul dalam data tersebut. *WordCloud* merupakan fungsi bahasa pemrograman yang sering di gunakan untuk visualisasikan kata yang sering muncul dalam sebuah data (Idris et al., 2023). Hasil *WordCloud* hanya berupa tulisan kata dengan ukuran yang semakin besar bila semakin sering muncul dalam sebuah data, tanpa memberitahu frekuensinya. Kekurangan ini dapat di sempurnakan dengan analisis N-Gram yang dapat menunjukkan jumlah frekuensi per kata (Atmadja et al., 2024).

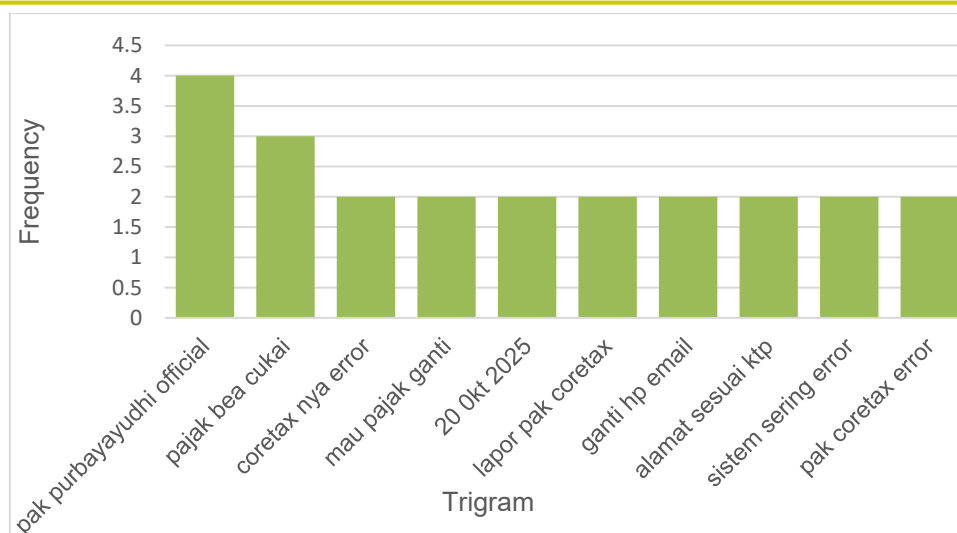
Sepuluh daftar kata *unigram* dengan frekuensinya bisa dilihat pada (Gambar 3). Daftar hasil *unigram* tertinggi ialah “pak” sebanyak 141, kemudian “pajak” sebanyak 111, dan “coretax” sebanyak 81. Beberapa kata seperti “lapor”, “bayar”, “gak”, “buat”, “mau” dan “error”, mengidentifikasi masih banyak kendala teknis dalam melakukan pembayar pajak coretex. Hal ini diperkuat dengan hasil *bigram* yang menunjukkan 2 kata dominan dan memiliki makna lebih utuh antara lain; “pak coretax”, “bayar pajak”, “sering error” (Gambar 4). Kendala teknis coretax juga tercermin dari hasil trigram yang menyebutkan “Ganti hp email”, “Alamat sesuai KTP” (Gambar 5). Hasil ini sejalan dengan penelitian Aldo et al., (2025) yang menyatakan, persoalan teknis menjadi isu utama yang banyak dibicarakan dan dikeluhkan oleh masyarakat digital di berbagai platform media sosial.



Gambar 3. Hasil *Unigram*



Gambar 4. Hasil *Bigram*

Gambar 5. Hasil *Trigram*

Salah satu hasil analisis bigram tertinggi ke tiga adalah kata “tugas pajak”. Hal ini mengidentifikasi ada keluhan kepada petugas pajak yang melayani. Nurhaeni et al., (2025) menjelaskan, beberapa hambatan yang signifikan saat menerapkan coretax seringkali, wajib pajak disarankan untuk melihat instruksi di YouTube, yang terdiri dari 55 video yang tidak teratur dan membingungkan pengguna.

Implementasi coretax juga berdampak buruk bagi para petugas pajak. Setiadi & Saluy (2025) menyatakan fokus yang berlebihan pada proses administrasi coretax menyebabkan peningkatan beban kerja, terutama karena wajib pajak harus memenuhi tenggat waktu pelaporan pajak. Ketidakpastian prosedur dan kompleksitas dalam penerapan awal sistem juga menambah tekanan kerja. Hal ini tercermin pada salah satu analisis *trigram* yang menghasilkan kata “20 okt 2025”. Tanggal 20 Oktober 2025 adalah Batas Akhir Pengunggahan Faktur Pajak (e-Faktur) bagi Pengusaha Kena Pajak (PKP). E-faktur yang diunggah setelah tanggal ini tidak akan disetujui dan dianggap bukan faktur pajak. Tanggal 20 Oktober 2025 juga merupakan batas akhir pelaporan Surat Pemberitahuan (SPT) Masa Pajak Penghasilan (PPh) untuk masa pajak September 2025.

Temuan mengenai dominasi keluhan terkait kendala teknis menunjukkan perlunya langkah strategis dari Kementerian Keuangan Republik Indonesia dalam meningkatkan kualitas sistem coretax. Salah satu solusi praktis yang dapat dilakukan adalah penguatan infrastruktur teknologi informasi, khususnya dalam meningkatkan kapasitas server dan stabilitas sistem untuk mengantisipasi lonjakan akses pengguna menjelang batas pelaporan pajak. Penelitian sebelumnya menunjukkan bahwa keandalan sistem (*system reliability*) dan kemudahan penggunaan (*perceived ease of use*) merupakan faktor utama yang mempengaruhi kepuasan dan kepatuhan pengguna dalam layanan digital pemerintah. Keandalan sistem dan kualitas layanan terbukti berpengaruh signifikan terhadap persepsi kemudahan penggunaan dan kepuasan pengguna (Adyaputra et al., 2025). Sementara dalam kerangka *Technology Acceptance Model* (TAM) atau model adopsi teknologi, kemudahan penggunaan dan manfaat yang dirasakan menjadi faktor utama dalam meningkatkan penerimaan,

kepercayaan, serta keberlanjutan penggunaan layanan *e-government* (Verástegui et al., 2025). Kualitas layanan digital pemerintah yang mencakup kemudahan interaksi dan keandalan sistem juga berkontribusi terhadap kepuasan dan loyalitas pengguna (Pham et al., 2023). Selain itu, pemerintah juga perlu mengembangkan mekanisme *real-time monitoring* dan *early warning system* untuk mendeteksi gangguan teknis secara cepat serta menyediakan layanan bantuan yang responsif, seperti *chatbot* atau layanan pelanggan berbasis digital (Alshehri et al., 2012; Lian & Wang, 2024; Rachman et al., 2025).

Dari sisi kebijakan, diperlukan evaluasi berkala terhadap implementasi sistem coretax melalui analisis umpan balik publik berbasis data, sehingga perbaikan sistem dapat dilakukan secara adaptif dan berkelanjutan. Pendekatan ini terbukti efektif dalam meningkatkan kualitas layanan digital dan kepercayaan masyarakat, di mana kualitas sistem dan layanan *e-government* berpengaruh signifikan terhadap tingkat kepercayaan dan penggunaan berkelanjutan (Nulhusna et al., 2017). Selain itu, partisipasi publik dan transparansi data dalam proses evaluasi layanan digital juga berkontribusi terhadap peningkatan legitimasi dan kepercayaan institusi pemerintah (Prasodjo, 2026). Dalam konteks transformasi digital, evaluasi berbasis pengalaman pengguna dan tata kelola adaptif menjadi faktor kunci dalam meningkatkan kualitas layanan publik dan memperkuat hubungan antara pemerintah dan masyarakat (Ramadhani et al., 2025; Sadiq et al., 2026; Suradi & Purnomo, 2026).

4. Kesimpulan

Hasil analisis N-Gram mengidentifikasi masih banyak kendala teknis dalam melakukan pembayaran pajak coretax, dari sistem error dan administrasi detail lainnya. Hasil analisis N-Gram juga mengidentifikasi meningkatnya beban kerja serta keluhan terhadap petugas pajak terutama mendekati tanggal tertentu.

Ucapan Terimakasih

Ucapan terimakasih diberikan kepada semua pihak yang telah memberikan dukungan serta kontribusi dalam proses pelaksanaan penelitian ini.

Referensi

- Adyaputra, R., Imtihan, K., & Saleh, M. (2025). Assessing User Satisfaction of Local Government Websites Through ISO 25010 and Technology Acceptance Model (TAM): A SmartPLS and IPMA-Based Study in Lombok Tengah. *Journal of Information and Organizational Sciences*, 49(2), 291–312. <https://doi.org/10.31341/jios.49.2.8>
- Ahmad, M., Aftab, S., Bashir, M. S., & Hameed, N. (2018). Sentiment analysis using SVM: a systematic literature review. *International Journal of Advanced Computer Science and Applications*, 9(2), 182–188. <https://doi.org/10.14569/IJACSA.2018.090226>
- Al-Husna, G. S., Asmarajati, D., Ihsannuddin, I. A., & Mahmudati, R. (2024). Perbandingan Metode Naïve Bayes Dan Support Vector Machine Untuk Analisis Sentimen Pada Ulasan Pengguna Aplikasi LinkedIn. *Storage: Jurnal Ilmiah Teknik*

- Dan Ilmu Komputer*, 3(2), 139–144. <https://doi.org/10.55123/storage.v3i2.3602>
- Aldo, Noviriani, E., & Roshani. (2025). Exploration of Netizen Perspectives on the Implementation of the Core Tax Administration System: (A Netnographic Study). *Jurnal Ilmu Ekonomi Manajemen Dan Akuntansi MH Tamrin*, 6(1), 131–150. <https://doi.org/10.37012/ileka.v6i1.2634>
- Alshehri, M., Drew, S., Alhussain, T., & Alghamdi, R. (2012). The effects of website quality on adoption of E-government service: An Empirical study applying UTAUT model using SEM. *23rd Australasian Conference On Information Systems*, 1–13. <https://aisel.aisnet.org/acis2012/24/>
- Ariyanto, D., Dewi, A. A., Paramadani, R. B., & Paramadina, A. A. (2024). Determinants of tax compliance and their impact on a sustainable information society: an investigation of MSMEs. *Cogent Business & Management*, 11(2414856), 1–17. <https://doi.org/10.1080/23311975.2024.2414856>
- Atmadja, A. R., Rahman, M. N., & Zulfikar, W. B. (2024). Permodelan Topik pada Layanan Akademik Perguruan Tinggi dengan Menggunakan N-Gram. *INTERNAL (Information System Journal)*, 7(2), 167–177. <https://doi.org/10.32627/internal.v7i2.1192>
- Bao, C., & Wang, Y. (2021). A survey of word cloud visualization. *Journal of Computer-Aided Design & Computer Graphics*, 33(4), 532–544. <https://doi.org/10.3724/SP.J.1089.2021.18811>
- Bhardwaj, A., Diwan, N., & Wang, G. (2026). *Beyond BeautifulSoup: Benchmarking LLM-Powered Web Scraping for Everyday Users*. <https://doi.org/10.48550/arXiv.2601.06301>
- Budiman, J., Limgestu, R., & Tri Sagianto, I. (2023). Transformasi Sistem Perpajakan Di Era Digital: Tantangan, Inovasi, Dan Kebijakan Adaptif. *Jurnal Ilmiah Akuntansi Dan Keuangan*, 5(9), 3518–3526. <https://doi.org/10.53866/jimi.v4i4.648>
- Cahyani, R. D., & Prasetyaningrum, P. T. (2026). Sentiment Analysis of User Reviews for AI Applications: Evaluating SVM, Logistic Regression, and Random Forest. *Journal of Information Systems and Informatics*, 8(1), 1–27. <https://doi.org/10.63158/journalisi.v8i1.1366>
- Dimetheo, G., Salsabila, A., & Izaak, N. C. A. (2023). Implementasi Core Tax Administration System sebagai upaya mendorong kepatuhan pajak di Indonesia. *Prosiding Seminar Nasional Ekonomi Dan Perpajakan*, 3(1), 10–25. <https://conference.um.ac.id/index.php/taxcenter/article/view/8557>
- Dzulkarnain, T., Ratnawati, D. E., & Rahayudi, B. (2024). Penggunaan Metode Naïve Bayes Classifier pada Analisis Sentimen Penilaian Masyarakat Terhadap Pelayanan Rumah Sakit di Malang. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(5), 993–1000. <https://doi.org/10.25126/jtiik.2024117979>
- Erstiawan, M. S. (2025). Analisis Tantangan dan Respon Pemangku Kepentingan terhadap Implementasi Sistem Coretax di Indonesia: Analisis Content Media. *GEMILANG: Jurnal Manajemen Dan Akuntansi*, 5(2), 304–324. <https://doi.org/10.56910/gemilang.v5i2.1969>
- Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., & Smith, N. J. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362.

<https://doi.org/10.1038/s41586-020-2649-2>

- Idris, I. S. K., Mustofa, Y. A., & Salihi, I. A. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM). *Jambura Journal of Electrical and Electronics Engineering*, 5(1), 32–35. <https://doi.org/10.37905/jjee.v5i1.16830>
- Khan, T. A., Sadiq, R., Shahid, Z., Alam, M. M., & Su'ud, M. B. M. (2024). Sentiment analysis using support vector machine and random forest. *Journal of Informatics and Web Engineering*, 3(1), 67–75. <https://doi.org/10.33093/jiwe.2024.3.1.5>
- Korat, C., & Munandar, A. (2025). Penerapan Core Tax Administration System (CTAS) Langkah Meningkatkan Kepatuhan Perpajakan di Indonesia. *Jurnal Riset Akuntansi Politala*, 8(1), 16–29. <https://doi.org/10.34128/jra.v8i1.453>
- Lian, Z., & Wang, F. (2024). Government chatbot: Empowering smart conversations with enhanced contextual understanding and reasoning. *Journal of Information Science*, 01655515241268863. <https://doi.org/10.1177/01655515241268863>
- Mubaroroh, H. H., Yasin, H., & Rusgiyono, A. (2022). Analisis Sentimen Data Ulasan Aplikasi Ruangguru Pada Situs Google Play Menggunakan Algoritma Naïve Bayes Classifier Dengan Normalisasi Kata Levenshtein Distance. *Jurnal Gaussian*, 11(2), 248–257. <https://doi.org/10.14710/j.gauss.v11i2.35472>
- Nulhusna, R., Sandhyaduhita, P. I., Hidayanto, A. N., & Phusavat, K. (2017). The relation of e-government quality on public trust and its impact on public participation. *Transforming Government: People, Process and Policy*, 11(3), 393–418. <https://doi.org/10.1108/TG-01-2017-0004>
- Nurhaeni, D. F., Masitoh, D., Shofurani, H., & Livtanta, N. K. (2025). Analisis Efektifitas dan Efisiensi Sistem CORETAX: Mengukur Kepercayaan Publik di Tengah Transisi Sistem Perpajakan 2025. *Jurnal Sosial Politika (JSP)*, 6(1), 21–37. <https://doi.org/10.54144/jsp.v6i1.103>
- Nurhidayat, R., & Dewi, K. E. (2023). Penerapan Algoritma K-Nearest Neighbor Dan Fitur Ekstraksi N-Gram Dalam Analisis Sentimen Berbasis Aspek. *Komputa: Jurnal Ilmiah Komputer Dan Informatika*, 12(1), 91–100. <https://doi.org/10.34010/komputa.v12i1.9458>
- Ojo, O. E., Gelbukh, A., Calvo, H., & Adebajji, O. O. (2021). Performance study of n-grams in the analysis of sentiments. *Journal of the Nigerian Society of Physical Sciences*, 3(2021), 477–483. <https://doi.org/10.46481/jnsps.2021.201>
- Pham, L., Limbu, Y. B., Le, M. T. T., & Nguyen, N. L. (2023). E-government service quality, perceived value, satisfaction, and loyalty: evidence from a newly emerging country. *Journal of Public Policy*, 43(4), 812–833. <https://doi.org/10.1017/S0143814X23000296>
- Prasodjo, T. (2026). Public Trust Reloaded: The Impact of Data Transparency and Digital Participation on Government Legitimacy in the Era of Open Government. *Golden Ratio of Social Science and Education*, 6(1), 53–67. <https://doi.org/10.52970/grsse.v6i1.1894>
- Pratyasto, A. (2026). Implementasi Metode Leksikon dengan Stemming Sastrawi pada Python untuk Ekstraksi Opini Pelanggan Warung Sate di Purbalingga. *Perwira Journal of Science & Engineering*, 6(1), 217–220. <https://doi.org/10.54199/cqbyak02>

- Priyatno, A. M., & Firmananda, F. I. (2022). N-Gram Feature for Comparison of Machine Learning Methods on Sentiment in Financial News Headlines. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 1(1), 1–6. <https://doi.org/10.31004/riggs.v1i1.4>
- Rachman, F., Setiawan, F. A., & Ginting, A. H. (2025). Development Of Digital Reporting System Using Whatsapp Chatbot Alert E-Pol (Police Service Frontline Integration System) To Accelerate Police Response To Public Complaints. *Open Society Conference*, 3(1), 164–170. <https://doi.org/10.33830/osc.v3i1.7046>
- Rahmadina, A. P., Setiawan, N. Y., & Rusydi, A. N. (2025). Analisis Sentimen Terhadap Kebijakan Kenaikan PPN 12% Menggunakan Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 9(9), 1–10. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/15322>
- Rahman, S. S. M. M., Biplob, K. B. M. B., Rahman, M. H., Sarker, K., & Islam, T. (2020). An investigation and evaluation of N-Gram, TF-IDF and ensemble methods in sentiment classification. *International Conference on Cyber Security and Computer Science*, 391–402. https://doi.org/10.1007/978-3-030-52856-0_31
- Ramadhan, R., Sari, Y. A., & Adikara, P. P. (2021). Perbandingan Pembobotan Term Frequency-Inverse Document Frequency dan Term Frequency-Relevance Frequency terhadap Fitur N-Gram pada Analisis Sentimen. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(11), 5075–5079. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10173>
- Ramadhani, D. P. S., Sulaiman, H. R., Anggraeni, A. W., & Aisyah, S. (2025). The effectiveness of e-government services in enhancing public trust: a comparative study across ASEAN countries. *Journal of Management and Informatics*, 4(1), 649–667. <https://doi.org/10.51903/jmi.v4i1.150>
- Rofiq, H. N., & Sujak, G. M. M. (2024). Analisis Persepsi Masyarakat Terhadap Lelang Indonesia Melalui Analisis n-gram dan Sentimen. *SINTECH (Science and Information Technology) Journal*, 7(3), 167–179. <https://doi.org/10.31598/sintechjournal.v7i3.1705>
- Sadiq, U., Alhassan, Y. M., & Issah, I. (2026). E-Government Services Quality and Trust in Government Organizations: The Mediating Roles of Perceived Transparency. *Cureus Journals*, 3(1). <https://doi.org/10.7759/s44404-025-00031-7>
- Saputra, M., & Wahyuni, S. (2024). Analisis sentimen pengguna pada aplikasi bank digital KROM dengan algoritma support vector machine. *Infotech Journal*, 10(2), 327–332. <https://doi.org/10.31949/infotech.v10i2.11801>
- Schonlau, M., Guenther, N., & Sucholutsky, I. (2017). Text mining with n-gram variables. *The Stata Journal*, 17(4), 866–881. <https://doi.org/10.1177/1536867X1801700406>
- Setiadi, T. B., & Saluy, A. B. (2025). The Influence Of Changes In The Coretax Taxation System On Stress Levels And Workload In The Tax Department. *Jurnal Fokus Manajemen*, 5(2), 145–150. <https://doi.org/10.37676/jfm.v5i2.8050>
- Shevira, S., Suarjaya, I., & Buana, P. W. (2022). Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia. *JITTER: Jurnal Ilmiah Teknologi Dan Komputer*, 3(2), 1074. <https://doi.org/10.24843/JTRTI.2022.v03.i02.p06>

- Suradi, A., Purnomo, H. D., & Hendry. (2026). A User-Centred E-Government Service Design Framework to Enhance Public Trust in Sustainable Policies. *Engineering, Technology & Applied Science Research*, 16(1), 31852–31862. <https://doi.org/10.48084/etasr.14851>
- Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, 57, 117–126. <https://doi.org/10.1016/j.eswa.2016.03.028>
- Trisnawati, W., & Wibowo, A. (2024). Sentiment analysis of ICT service user using Naive Bayes classifier and SVM methods with TF-IDF text weighting. *Jurnal Teknik Informatika (JUTIF)*, 5(3), 709–719. <https://doi.org/10.52436/1.jutif.2024.5.3.1784>
- Verástegui, M. D.-G., Medina-Quintero, J.-M., & Ortiz-Rodriguez, F. (2025). Perceived usefulness and ease of use of e-government to generate trust and intention to use by citizens. *Journal of Technology Management & Innovation*, 20(1), 49–60. <https://doi.org/10.4067/S0718-27242025000100049>