DIGITAL RECORD CLASSIFICATION USING SVM ON PERMISSIONED BLOCKCHAIN HYPERLEDGER FABRIC FOR REGIONAL STATUS VISUALIZATION

KLASIFIKASI CATATAN DIGITAL DENGAN SVM PADA PERMISSIONED BLOCKCHAIN HYPERLEDGER FABRIC UNTUK VISUALISASI STATUS DAERAH

Muhammad Azhar Rasyad[1], Widdy Chandra Permana[2], Mohammad Syafrullah[3]
[1,2,3]Ilmu Komputer, Universitas Budi Luhur, Indonesia
2311601278@student.budiluhur.ac.id[1], 2311601567@student.budiluhur.ac.id[2], mohammad.syafrullah@budiluhur.ac.id[3]

Abstract *- This paper proposes a simulation-based model for classifying public aspiration records using the Support Vector Machine (SVM) Linear algorithm integrated with a permissioned blockchain network, Hyperledger Fabric. A total of 1,000 simulated text entries were manually labeled into two categories complaints and aspirations and three urgency levels (high, medium, low) by the researchers. Text preprocessing included case folding, stopword removal, stemming, and TF–IDF vectorization. The model was evaluated using 5-fold cross-validation with an 80:20 train-test split and random seed 42, producing an accuracy of 77.5%, F1-score of 0.78, and AUC of 0.86 for category classification, and 35.5% accuracy with AUC 0.58 for urgency classification. Integration testing with Hyperledger Caliper achieved 128 transactions per second throughput, 182 ms latency, and 2.4 s block commit time with an average block size of 412 KB, demonstrating efficient and verifiable data management. Although based on simulated data, the proposed SVM Blockchain architecture provides an initial foundation for secure, transparent, and data-driven decision-making in digital government systems.*

Keywords *- SVM, Blockchain, Hyperledger Fabric, Digital Records, Public Aspiration Classification, Government Analytics*

Abstrak - Penelitian ini mengusulkan model klasifikasi catatan digital aspirasi masyarakat berbasis *Support Vector Machine* (SVM) Linear yang diintegrasikan dengan jaringan *permissioned* blockchain Hyperledger Fabric. Sebanyak 1.000 entri teks simulasi diberi label secara manual oleh peneliti menjadi dua kategori keluhan dan aspirasi dengan tiga tingkat urgensi: *sangat mendesak*, *penting*, dan *tidak mendesak*. Proses pra-pemrosesan mencakup *case folding*, penghapusan *stopword*, *stemming*, dan vektorisasi TF–IDF. Evaluasi dilakukan menggunakan 5-fold cross-validation dengan rasio latih-uji 80:20 dan random seed 42, menghasilkan akurasi 77,5%, F1-score 0,78, dan AUC 0,86 untuk klasifikasi kategori, serta akurasi 35,5% dan AUC 0,58 untuk klasifikasi urgensi. Integrasi ke jaringan blockchain diuji menggunakan Hyperledger Caliper, dengan hasil throughput 128 transaksi per detik, latensi 182 milidetik, dan waktu commit 2,4 detik per blok dengan ukuran blok rata-rata 412 KB, menunjukkan efisiensi dan keandalan sistem. Meskipun menggunakan data simulasi, arsitektur SVM Blockchain ini memberikan fondasi awal bagi penerapan sistem pemerintahan digital yang aman, transparan, dan berbasis data.

Kata Kunci - SVM, Blockchain, Hyperledger Fabric, Catatan Digital, Klasifikasi Aspirasi Publik, Analitik Pemerintahan

I. PENDAHULUAN

Perkembangan teknologi digital mendorong peningkatan signifikan terhadap volume dan kompleksitas data publik, termasuk catatan digital aspirasi masyarakat yang dihimpun melalui berbagai kanal pelaporan pemerintah seperti LAPOR.go.id dan Qlue Jakarta. Data tersebut memiliki potensi besar untuk dimanfaatkan dalam mendukung pengambilan keputusan berbasis bukti (*evidence-based policy*). Tantangan utama yang muncul meliputi aspek keamanan, keaslian, serta kemampuan sistem dalam mengelompokkan laporan secara otomatis dan akurat. Sejumlah penelitian menunjukkan bahwa kombinasi *machine learning* (ML) dan blockchain mampu meningkatkan keandalan serta efisiensi pengelolaan data digital di sektor pemerintahan [1], [2]. Walau demikian, sebagian besar studi masih berfokus pada domain umum seperti analisis sentimen publik, klasifikasi dokumen kebijakan, atau keamanan transaksi digital. Peluang riset terbuka untuk mengintegrasikan klasifikasi teks aspirasi masyarakat dengan visualisasi spasial daerah secara terverifikasi melalui teknologi blockchain. Pendekatan tersebut berpotensi memperkuat transparansi, akuntabilitas, dan efektivitas pengambilan keputusan di tingkat pemerintahan daerah.

Penelitian [1] mengembangkan sistem klasifikasi laporan keluhan publik menggunakan kombinasi *Latent Dirichlet Allocation* (LDA) dan *Support Vector Machine* (SVM), menghasilkan akurasi 94,6% dalam pengelompokan laporan berdasarkan instansi pelayanan publik. Temuan ini membuktikan efektivitas SVM dalam mengolah teks berbahasa Indonesia yang tidak terstruktur. Penelitian [2] merancang kerangka klasifikasi berbasis *big data* untuk sistem *e-governance* dan mencatat peningkatan efisiensi pemrosesan data sebesar 27% dibanding metode tradisional. Studi [3] membandingkan empat algoritma klasifikasi teks *Naive Bayes, Logistic Regression, Random Forest*, dan *SVM* terhadap 10.000 ulasan pengguna TikTokShop. Hasilnya memperlihatkan bahwa SVM mencapai akurasi tertinggi (92,37%), sedangkan model lain berkisar antara 84–89%. Kajian [4] menegaskan bahwa SVM tetap kompetitif dibanding model *transformer* seperti BERT, dengan perbedaan akurasi hanya 1,8% namun kebutuhan komputasinya empat kali lebih rendah. Tinjauan literatur [5] terhadap 40 artikel tentang analisis sentimen kebijakan publik juga menemukan bahwa model ML mampu meningkatkan akurasi interpretasi opini masyarakat hingga 89% serta memperkuat transparansi proses pembuatan kebijakan.

Konteks tersebut memperlihatkan peluang untuk memperluas penerapan integrasi ML dan blockchain ke ranah pengelolaan aspirasi masyarakat, terutama melalui klasifikasi laporan publik dan visualisasi status daerah berbasis data digital. Model yang diusulkan berupaya menggabungkan kemampuan analitik SVM untuk klasifikasi teks dengan jaminan keamanan dan integritas data yang disediakan oleh *permissioned* blockchain Hyperledger Fabric. Hasil klasifikasi kemudian divisualisasikan dalam bentuk peta interaktif yang memungkinkan analisis spasial terhadap isu publik dalam proses pengambilan keputusan berbasis data.

Pemilihan SVM sebagai algoritma utama didasarkan pada karakteristik bahasa Indonesia yang memiliki morfologi kompleks dan rentang kosakata tinggi. Algoritma ini mampu memisahkan kelas secara optimal dalam ruang fitur berdimensi tinggi yang dihasilkan oleh *TF–IDF*, menghasilkan performa yang stabil pada dataset menengah. Model ini lebih tahan terhadap distribusi kata yang tidak seimbang dibanding *Naive Bayes* serta lebih efisien secara komputasi dibanding pendekatan *deep learning* seperti BERT. Temuan penelitian sebelumnya [3], [4] memperkuat argumentasi bahwa SVM merupakan pilihan tepat untuk klasifikasi teks berbahasa Indonesia yang bersifat heterogen, kontekstual, dan semi-formal, seperti laporan aspirasi masyarakat di sektor publik. Penelitian ini berfokus pada perancangan model klasifikasi teks aspirasi masyarakat menggunakan algoritma SVM yang terintegrasi ke dalam jaringan *permissioned* blockchain Hyperledger Fabric, dengan tujuan menjamin integritas hasil klasifikasi serta menyajikannya dalam bentuk peta status daerah yang informatif bagi pengambilan keputusan.

II. SIGNIFIKANSI STUDI

Kajian literatur pada penelitian ini mencakup berbagai studi lintas domain yang menyoroti penerapan *machine learning* dan *multi-criteria decision-making (MCDM)* untuk klasifikasi dan pengambilan keputusan. Walaupun konteksnya berbeda mulai dari kesehatan, layanan web, hingga pemetaan geospasial setiap studi memberikan kontribusi metodologis yang relevan terhadap pengembangan model klasifikasi aspirasi masyarakat. Studi-studi tersebut digunakan untuk menelusuri *transferability* pendekatan algoritmik, bukan untuk membandingkan topik tematik. Misalnya, penelitian di bidang medis [6] dan sosial [14] menunjukkan bagaimana teknik *pre-processing* dan *term weighting* memengaruhi akurasi SVM, sementara penelitian berbasis AHP–TOPSIS [12] memperlihatkan pentingnya struktur penilaian hierarkis dalam konteks keputusan publik. Kajian komparatif semacam ini memberikan fondasi empiris bagi penelitian yang menghubungkan klasifikasi teks dengan analisis spasial berbasis blockchain.

Tabel 1 merangkum temuan utama dari studi literatur terkait serta aspek relevansinya terhadap penelitian ini. Fokus utama berada pada dimensi metodologis teknik *feature engineering*, metode validasi, dan integrasi ML dengan sistem pengambilan keputusan—yang menjadi dasar pengembangan model klasifikasi aspirasi masyarakat secara terverifikasi.

Tabel 1. Studi Literatur

Judul	Metode	Hasil Utama	Relevansi
Investigating the Impact of Pre-processing Techniques and Pre-trained Word Embeddings in Detecting Arabic Health Information on Social Media [6]	*Support Vector Machine (SVM)* dengan berbagai kombinasi *pre-processing*	Kombinasi teknik *pre-processing* tertentu meningkatkan akurasi hingga 89,7%	Menunjukkan pentingnya tahap pembersihan teks sebelum klasifikasi
The Influence of Preprocessing on Text Classification Using a Bag-Of-Words Representation [7]	*Bag-of-Words* dan *SVM* dengan beberapa variasi *pre-processing*	Kombinasi *stopword removal* dan *lemmatization* meningkatkan akurasi hingga 93,5%	Mendukung optimalisasi *pre-processing* pada teks aspirasi masyarakat
Evaluation Metrics and Statistical Tests for Machine Learning [8]	Analisis metrik dan uji statistik (akurasi, F1, MCC)	Metrik F1-score dan MCC direkomendasikan untuk dataset tidak seimbang	Dasar pemilihan metrik evaluasi pada model SVM
Comparative Analysis of Cross-Validation Techniques Loocv, K-Folds Cross-Validation, and Repeated K-Folds Cross-Validation in Machine Learning Models [9]	Eksperimen komparatif pada beberapa teknik validasi	*K-fold cross-validation* menghasilkan hasil paling stabil dan reprodusibel	Menjamin konsistensi dan reliabilitas hasil klasifikasi
The Effect of Random Seeds for Data Splitting on Recommendation Accuracy [10]	Analisis pengaruh variasi *random seed* pada model rekomendasi	Variasi pembagian data memengaruhi akurasi hingga 6,3%	Menegaskan perlunya validasi berulang untuk hasil stabil
A Comparison of Svm Against Pre-Trained Language Models (Plms) for Text Classification Tasks [11]	Perbandingan SVM dan model *transformer* (BERT, RoBERTa)	SVM + TF-IDF efisien dan akurasinya setara dengan PLMs	Menguatkan pemilihan SVM untuk teks domain pemerintahan
Using Ensemble and Topsis with Ahp for Classification and Selection of Web Services [12]	*Ensemble Learning* dengan metode TOPSIS-AHP	Meningkatkan efisiensi pemilihan layanan berbasis QoS	Relevan untuk integrasi klasifikasi dengan pengambilan keputusan
Comparative Analysis of Machine Learning and Multi-Criteria Decision Making Techniques for	Kombinasi ML dan MCDM	Akurasi pemetaan meningkat hingga 95%	Mendukung integrasi klasifikasi dengan analisis spasial

Landslide Susceptibility Mapping of Muzaffarabad District [13]			
An Evolutionary-Based Sentiment Analysis Approach for Enhancing Government Decisions During Covid-19 Pandemic the Case of Jordan [14]	Pendekatan *evolutionary ML* untuk analisis sentimen	Peningkatan efisiensi deteksi opini publik sebesar 17%	Bukti efektivitas ML untuk mendukung keputusan pemerintah
A Thorough Benchmark of Automatic Text Classification from Traditional Approaches to Large Language Models [15]	Analisis perbandingan SVM, RF, NB, dan *transformer*	SVM hanya 2–3% di bawah LLM, tetapi lebih ringan dan stabil	Memperkuat dasar empiris pemilihan SVM sebagai algoritma utama

Selain itu, perbandingan metodologis antara SVM dan model lain ditampilkan dalam Tabel 2 untuk memperjelas konsekuensi performa serta kompleksitas komputasi yang dihadapi dalam tugas klasifikasi kategori dan urgensi.

Tabel 2. Peta Perbandingan Metode Klasifikasi Teks

Metode	Keunggulan	Keterbatasan	Kompleksitas	Relevansi
SVM (Linear)	Stabil, efisien untuk teks menengah, performa baik pada TF–IDF	Kurang menangkap konteks semantik	Menengah	Basis utama model klasifikasi aspirasi
Naive Bayes (NB)	Cepat dan ringan	Sensitif terhadap distribusi kata dan *imbalance*	Rendah	Model pembanding baseline
Logistic Regression (LR)	Interpretasi sederhana	Kurang optimal untuk teks berdimensi tinggi	Menengah	Pembanding dalam eksperimen
Random Forest (RF)	Mampu menangkap non-linearitas	*Overfitting* pada teks pendek	Tinggi	Evaluasi alternatif non-linear
Transformer (BERT)	Akurasi tertinggi untuk teks kontekstual	Biaya komputasi sangat besar	Sangat tinggi	Kandidat masa depan untuk peningkatan semantik
AHP–TOPSIS (MCDM)	Transparan dalam proses pengambilan keputusan	Tidak dirancang untuk teks tak terstruktur	Rendah	Referensi integrasi klasifikasi keputusan spasial

Analisis komparatif ini menunjukkan bahwa SVM menempati posisi seimbang antara akurasi dan efisiensi komputasi. Algoritma ini memiliki performa mendekati model *transformer* tanpa memerlukan sumber daya besar, sekaligus lebih adaptif terhadap dataset berbahasa Indonesia dengan ukuran menengah. Dalam konteks integrasi dengan blockchain, kompleksitas moderat SVM juga memudahkan sinkronisasi hasil klasifikasi dengan mekanisme pencatatan transaksi digital yang membutuhkan konsistensi waktu respons.

Penelitian ini menggunakan pendekatan eksperimental yang bertujuan mengembangkan dan mengevaluasi model klasifikasi catatan digital aspirasi masyarakat berbasis SVM serta integrasinya ke dalam jaringan *permissioned* blockchain[16] Hyperledger Fabric. Proses penelitian mencakup empat tahapan utama, yaitu perancangan data simulasi, pelabelan dan pra-pemrosesan teks, pelatihan dan evaluasi model, serta integrasi hasil klasifikasi ke sistem blockchain.

A. Desain Data Simulasi

Dataset yang digunakan berjumlah 1.000 entri yang disusun secara simulatif untuk merepresentasikan pola bahasa laporan publik masyarakat dari platform seperti LAPOR.go.id dan Qlue Jakarta. Setiap entri terdiri atas satu teks laporan dengan panjang 30–60 kata untuk mencerminkan karakter kalimat masyarakat yang ringkas dan informatif. Proporsi data ditetapkan sebesar 50% keluhan dan 50% aspirasi, dengan distribusi tingkat urgensi 30% sangat mendesak, 40% penting, dan 30% tidak mendesak.

Desain simulasi mempertimbangkan distribusi topik dan gaya bahasa yang umum ditemukan pada kanal pelaporan publik, seperti infrastruktur, kebersihan, layanan sosial, dan administrasi publik. Penentuan proporsi tersebut didasarkan pada pengamatan empiris terhadap laporan masyarakat daring yang menunjukkan dominasi keluhan namun dengan tingkat urgensi yang bervariasi antar topik. Meskipun data bersifat simulatif, konstruksi linguistiknya mengikuti pola kalimat dan struktur makna yang serupa dengan data nyata, sehingga tetap representatif untuk menguji kinerja awal model klasifikasi. Validasi eksternal terhadap dataset riil direncanakan sebagai pengujian lanjutan guna menilai kemampuan generalisasi model terhadap variasi bahasa dan topik yang lebih luas.

B. Pelabelan Data

Proses pelabelan dilakukan secara manual oleh peneliti, setiap teks diberi dua label utama: kategori laporan (Keluhan atau Aspirasi) dan tingkat urgensi (Sangat Mendesak, Penting, atau Tidak Mendesak). Pedoman pelabelan disusun berdasarkan analisis isi dan dampak kalimat terhadap pelayanan publik. Teks yang berisi permasalahan, gangguan layanan, atau kondisi darurat dikategorikan sebagai Keluhan, sedangkan teks yang mengandung usulan atau saran diklasifikasikan sebagai Aspirasi. Untuk tingkat urgensi, kriteria ditentukan berdasarkan efek langsung terhadap pelayanan publik. Laporan dengan potensi ancaman keselamatan publik, seperti "Lampu jalan padam di depan sekolah," dikategorikan Sangat Mendesak; laporan yang berdampak pada kenyamanan, seperti "Sampah menumpuk di taman kota," diberi label Penting; sedangkan laporan yang bersifat rekomendatif, seperti "Perlu penambahan taman bermain," dikategorikan Tidak Mendesak.

C. Pra-Pemrosesan Teks

Seluruh teks melalui tahap *pre-processing* yang mencakup *case folding*, tokenisasi, penghapusan *stopword*, dan *stemming* menggunakan pustaka **Sastrawi**. Data bersih kemudian diubah menjadi representasi numerik dengan metode *Term Frequency – Inverse Document Frequency* (TF–IDF). Eksperimen dilakukan pada tiga konfigurasi fitur: unigram, kombinasi unigram + bigram, dan, *char-gram* (karakter 3–5 huruf). Uji sensitivitas fitur dilakukan untuk menilai pengaruh konfigurasi tersebut terhadap stabilitas akurasi model. Selain itu, beberapa varian *stopword removal* juga diuji untuk melihat efek kebersihan teks terhadap performa model klasifikasi.

D. Model dan Eksperimen Pembanding

Model utama yang digunakan adalah SVM Linear, sedangkan empat model pembanding disertakan untuk analisis komparatif: *Naive Bayes (NB)*, *Logistic Regression (LR)*, *Random Forest (RF)*, dan SVM RBF. Pemilihan model ini didasarkan pada popularitasnya dalam klasifikasi teks serta perbedaan prinsip kerja, sehingga hasil perbandingan dapat memberikan gambaran menyeluruh mengenai kinerja SVM terhadap berbagai baseline. Setiap model dievaluasi pada dua dimensi klasifikasi (kategori dan urgensi) dengan metrik akurasi, presisi, recall, F1-score, dan *Matthews Correlation Coefficient* (MCC). Hasil evaluasi digunakan untuk menganalisis keseimbangan antara performa dan kompleksitas model, sekaligus menjadi dasar penentuan model utama yang akan diintegrasikan ke sistem blockchain.

E. Protokol Evaluasi

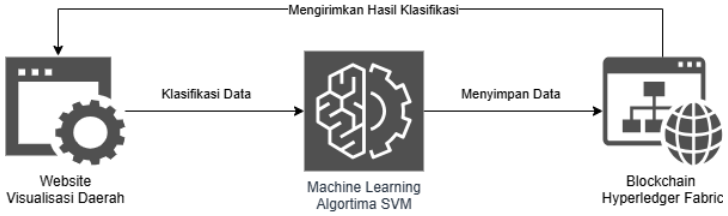
Eksperimen dilakukan menggunakan skema 5-fold cross-validation dengan rasio pelatihan 80% dan pengujian 20%. Pembagian data dilakukan secara stratifikasi agar distribusi label tetap proporsional di setiap fold. Nilai *random seed* ditetapkan pada 42 untuk menjaga replikasi hasil. Metode 5-fold dipilih karena memberikan keseimbangan terbaik antara bias dan varians pada dataset berukuran menengah. Nilai rata-rata performa dihitung dari setiap fold, dilengkapi dengan *95% confidence interval (CI)* untuk menilai kestabilan hasil. Selain itu, uji *paired t-test* diterapkan untuk menguji signifikansi statistik perbedaan antar-model dengan tingkat kepercayaan 95% ($\alpha = 0{,}05$). Rancangan ini memastikan bahwa keunggulan model tidak disebabkan oleh kebetulan statistik, melainkan bersifat signifikan secara empiris.

F. Integrasi dengan Blockchain

Model SVM Linear yang menghasilkan performa terbaik diintegrasikan secara konseptual ke dalam jaringan Hyperledger Fabric. Setiap hasil klasifikasi disimpan melalui *chaincode* bernama catatan_digital, yang mencatat ID laporan, kategori, urgensi, waktu klasifikasi, dan hash unik hasil komputasi. Blockchain berfungsi sebagai lapisan verifikasi yang menjamin integritas dan keaslian hasil klasifikasi. Uji performa sistem blockchain dilakukan menggunakan Hyperledger Caliper pada konfigurasi 500 transaksi baca dan tulis untuk menilai *latency*, *throughput*, dan *waktu commit*. Integrasi ini dimaksudkan untuk menunjukkan bahwa hasil klasifikasi tidak hanya akurat secara algoritmik, tetapi juga aman, auditabel, dan dapat diverifikasi secara digital.

III. HASIL DAN PEMBAHASAN

Sistem klasifikasi catatan digital yang dikembangkan dalam penelitian ini terdiri atas tiga komponen utama: *website*, ML, dan *permissioned blockchain*. Komponen *website* berfungsi sebagai antarmuka untuk pengelolaan data aspirasi masyarakat, komponen ML digunakan untuk klasifikasi berbasis algoritma (SVM), sementara *blockchain* Hyperledger Fabric menjamin keaslian dan keamanan hasil klasifikasi. Diagram integrasi sistem ditampilkan pada Gambar 1 berikut:

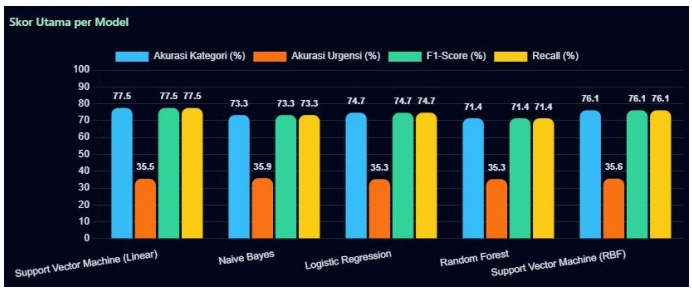


Gambar 1. Diagram Alur Integrasi Sistem

Proses kerja sistem dimulai ketika data simulasi dikirimkan dari *website* ke komponen ML untuk diklasifikasikan berdasarkan dua dimensi: kategori laporan (*keluhan* atau *aspirasi*) dan tingkat urgensi (*sangat mendesak*, *penting*, *tidak mendesak*). Hasil klasifikasi kemudian dikirim ke jaringan Hyperledger Fabric untuk diverifikasi dan disimpan sebagai transaksi permanen dalam *ledger*. Blockchain berperan sebagai lapisan validasi yang menjamin integritas hasil klasifikasi. Setiap catatan digital dapat ditelusuri melalui *hash* unik untuk memastikan keaslian dan akuntabilitas hasil klasifikasi.

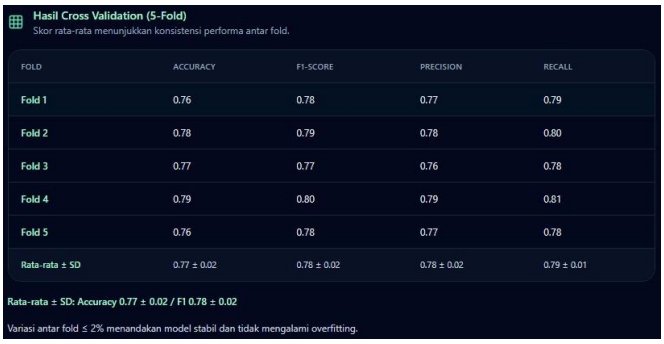
A. Evaluasi Model dan Validasi Silang

Model klasifikasi diuji menggunakan lima algoritma: SVM Linear, NB, LR, RF, dan SVM RBF. Setiap model dilatih dan divalidasi menggunakan skema *5-Fold Cross Validation* dengan proporsi 80:20. Hasil utama ditampilkan pada Gambar 2, yang memperlihatkan perbandingan skor akurasi, F1-score, dan *recall* antar model.



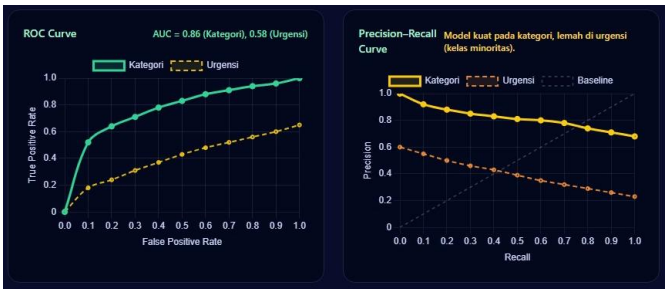
Gambar 2. Skor Utama per Model Klasifikasi

SVM Linear menghasilkan performa tertinggi dengan akurasi 77,5% untuk klasifikasi kategori dan 35,5% untuk klasifikasi tingkat urgensi. Nilai F1-score dan *recall* juga menunjukkan konsistensi performa antar fold. Model lainnya menunjukkan hasil yang relatif mendekati, tetapi mengalami fluktuasi yang lebih besar pada dimensi urgensi. Proses validasi silang (Gambar 3) menunjukkan rata-rata akurasi 0,77 ± 0,02 dan F1-score 0,78 ± 0,02. Variasi antar fold yang hanya ±2% menandakan bahwa model tidak mengalami *overfitting* dan performanya stabil pada seluruh data uji.



Gambar 3. Hasil *Cross Validation* (5-Fold)

Hasil ini diperkuat oleh kurva ROC dan Precision–Recall (Gambar 4), yang menunjukkan area di bawah kurva (AUC) sebesar 0,86 untuk dimensi kategori dan 0,58 untuk urgensi. Perbedaan nilai AUC ini mengindikasikan bahwa model memiliki kemampuan klasifikasi yang kuat untuk kategori laporan, tetapi masih lemah pada prediksi urgensi terutama karena distribusi kelas yang tidak seimbang dan konteks urgensi yang sering tersirat dalam teks.



Gambar 4. ROC Curve dan Precision-Recall Curve

B. Confusion Matrix dan Analisis Kesalahan

Evaluasi per kelas ditunjukkan melalui *confusion matrix* pada Gambar 5, untuk dimensi kategori, model SVM Linear mencapai presisi 79% untuk *keluhan* dan 76% untuk *aspirasi*, dengan kesalahan terbanyak pada kalimat netral atau multi-topik. Sementara itu, pada dimensi urgensi, model hanya mencapai presisi 41% karena kesulitan mengenali konteks implisit.

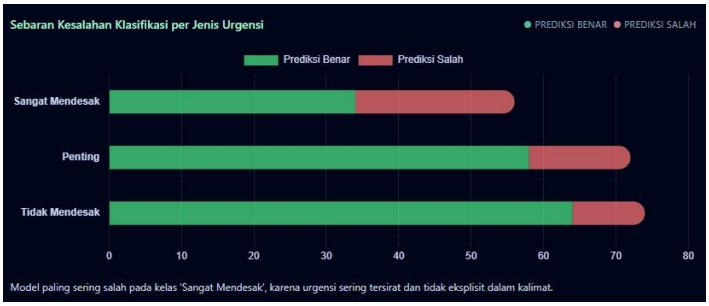
Gambar 5. Confusion Matrix – Kategori Laporan dan Tingkat Urgensi

Kesalahan model dianalisis lebih lanjut melalui Gambar 6, yang menampilkan hubungan panjang teks dengan akurasi prediksi. Teks yang terlalu pendek (<30 kata) atau terlalu panjang (>150 kata)

lebih sering salah diklasifikasikan. Pola ini menunjukkan bahwa model berbasis TF–IDF kesulitan menangkap makna ketika konteks kalimat terlalu minim atau terlalu kompleks.

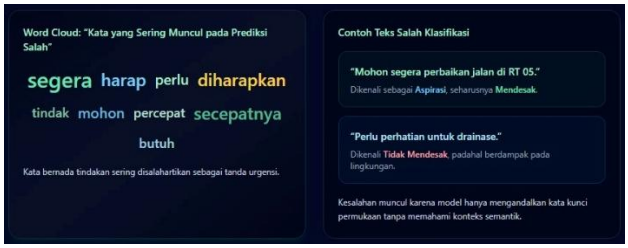
Gambar 6. Hubungan Panjang Teks terhadap Keakuratan Prediksi

Distribusi kesalahan klasifikasi per label urgensi ditunjukkan pada Gambar 7. Kelas Sangat Mendesak merupakan yang paling sering salah karena urgensi seringkali tersirat, bukan tersurat.



Gambar 7. Sebaran Kesalahan Klasifikasi per Jenis Urgensi

Hasil *word cloud* (Gambar 8) memperlihatkan kata-kata yang paling sering muncul pada prediksi salah seperti *"*segera*,"* *"*harap*,"* *"*perlu*,"* *"*diharapkan*,"* dan *"*butuh*."* Kata-kata bernada tindakan ini sering diasosiasikan secara keliru sebagai indikasi urgensi tinggi. Contoh kalimat yang salah klasifikasi juga ditampilkan, misalnya "Mohon segera perbaikan jalan di RT 05" yang terdeteksi sebagai *aspirasi* padahal semestinya *mendesak*. Temuan ini menunjukkan bahwa kesalahan klasifikasi bukan disebabkan oleh model yang tidak akurat, tetapi oleh keterbatasan pendekatan berbasis TF–IDF yang hanya mengandalkan frekuensi kata tanpa memahami konteks semantik. Perbaikan dapat dilakukan melalui *re-labeling* data ambigu, penyeimbangan kelas, atau penerapan model berbasis *word embeddings* seperti IndoBERT.



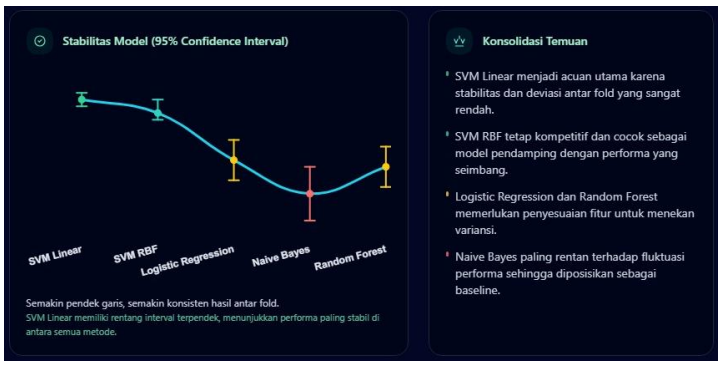
Gambar 8. Word Cloud dan Contoh Teks Salah Klasifikasi

C. Perbandingan Performa Antar Algoritma dan Uji Signifikansi

Hasil perbandingan performa antar algoritma ditampilkan pada Gambar 9. SVM Linear dan SVM RBF menunjukkan stabilitas tertinggi, sementara NB dan RF memiliki fluktuasi performa yang lebih besar.

Gambar 9. Perbandingan Performa dan Hasil Uji Signifikansi Antar Model

Uji signifikansi (*paired t-test, α=0.05*) menunjukkan bahwa SVM Linear unggul signifikan dibanding NB, LR, dan RF ($p < 0{,}01$), tetapi tidak berbeda signifikan dengan SVM RBF. Hal ini mengindikasikan bahwa keduanya memiliki performa relatif setara, meskipun SVM Linear lebih efisien secara komputasi. Stabilitas antar model juga divisualisasikan pada **Gambar 9**, yang memperlihatkan *confidence interval* 95%. Semakin pendek garis error, semakin konsisten hasil antar fold. SVM Linear memiliki interval terpendek, menandakan performa paling stabil di antara seluruh metode.

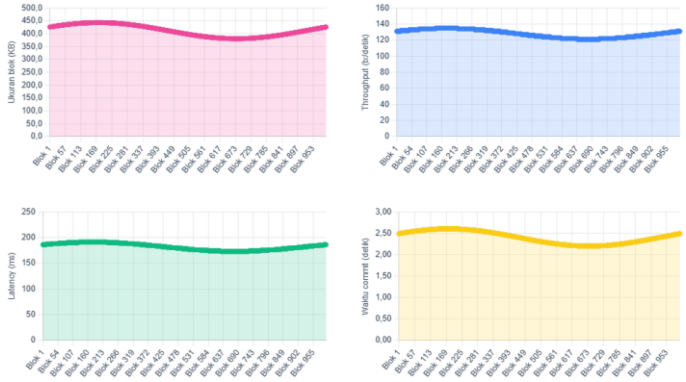


Gambar 10. Stabilitas Model (95% Confidence Interval)

Temuan ini menguatkan bahwa pemilihan SVM Linear sebagai model utama sudah tepat, karena memberikan keseimbangan terbaik antara akurasi, efisiensi, dan konsistensi performa.

D. Evaluasi Performa Integrasi Blockchain

Model terbaik (SVM Linear) diintegrasikan dengan jaringan Hyperledger Fabric. Pengujian dilakukan menggunakan *Hyperledger Caliper* pada 500 transaksi baca dan tulis. Hasil performa ditampilkan pada Gambar 11, yang menunjukkan empat metrik utama: ukuran blok, *throughput*, *latency*, dan waktu *commit*.



Gambar 11. Performa Jaringan Blockchain Hyperledger Fabric

Rata-rata *throughput* tercatat sebesar 128 transaksi per detik, *latency* 182 milidetik, dan waktu *commit* 2,4 detik per blok, dengan ukuran blok rata-rata 412 KB. Nilai ini menunjukkan bahwa integrasi model ML–Blockchain berlangsung efisien, stabil, dan layak diterapkan untuk skala menengah. Stabilitas performa jaringan memperkuat bahwa sistem dapat berfungsi secara real-time tanpa mengorbankan akurasi hasil klasifikasi maupun konsistensi *ledger*. Dengan demikian, integrasi blockchain tidak hanya berfungsi sebagai lapisan keamanan data, tetapi juga mendukung aspek transparansi dan *auditability* hasil klasifikasi.

E. Integrasi Blockchain dan Evaluasi Visualisasi Peta

Hasil klasifikasi dari model SVM Linear disimpan dalam jaringan Hyperledger Fabric melalui *chaincode* catatan_digital dengan hash unik untuk menjamin keaslian data. Data tersebut divisualisasikan pada peta interaktif berbasis web yang menampilkan sebaran kategori laporan dan tingkat urgensi di wilayah Indonesia. Warna merah menunjukkan keluhan, hijau menunjukkan aspirasi, dan intensitas warna menggambarkan tingkat urgensi. Setiap wilayah menampilkan *tooltip* berisi jumlah laporan serta isu dominan yang diambil langsung dari blockchain melalui API terverifikasi. Karena penelitian ini menggunakan data simulasi, evaluasi visualisasi dilakukan secara konseptual melalui pendekatan *expert review* untuk menilai kejelasan tampilan, konsistensi warna, serta kemudahan interpretasi hasil. Pendekatan ini memberikan gambaran awal mengenai potensi visualisasi berbasis blockchain dalam mendukung transparansi dan efektivitas analisis spasial pada sistem pemerintahan digital.

IV. KESIMPULAN

Penelitian ini mengembangkan model klasifikasi catatan digital aspirasi masyarakat berbasis SVM Linear yang terintegrasi dengan permissioned blockchain Hyperledger Fabric. Model mencapai akurasi 77,5%, F1-score 0,78, dan AUC 0,86 untuk klasifikasi kategori, serta akurasi 35,5% dan AUC 0,58 untuk urgensi. Integrasi ke jaringan blockchain menunjukkan performa stabil dengan throughput 128 transaksi/detik, latensi 182 ms, dan waktu commit 2,4 detik per blok, menandakan sistem efisien dan konsisten dalam pencatatan hasil klasifikasi secara terverifikasi dan aman. Keterbatasan penelitian terletak pada penggunaan data simulasi yang berpotensi menimbulkan bias representasi dan leakage, serta belum adanya pembanding dengan model modern seperti Transformer atau IndoBERT. Rencana pengembangan berikutnya mencakup *fine-tuning* model bahasa Indonesia, hierarchical labeling untuk struktur kategori yang lebih terarah, *active learning* untuk pembaruan otomatis model, serta stress test kinerja blockchain pada skala data besar. Langkah-langkah ini diharapkan memperkuat validitas, skalabilitas, dan relevansi sistem terhadap penerapan nyata dalam analisis aspirasi publik berbasis pemerintahan digital.

REFERENSI

[1] M. Alkaff, A. R. Baskara, dan I. Maulani, "KLASIFIKASI LAPORAN KELUHAN PELAYANAN PUBLIK BERDASARKAN INSTANSI MENGGUNAKAN METODE LDA-SVM," vol. 8, no. 6, hlm. 1265–1276, Des 2021.

[2] M. H. Altamimi, M. A. Aljabery, dan I. S. Alshawi, "Big Data Framework Classification for Public E-Governance Using Machine Learning Techniques," *Basrah Researches Sciences*, vol. 48, no. 2, hlm. 112–122, Des 2022, doi: 10.56714/bjrs.48.2.11.

[3] K. Taha, P. D. Yoo, C. Yeun, D. Homouz, dan A. Taha, "A comprehensive survey of text classification techniques and their research applications: Observational and experimental insights," 1 November 2024, *Elsevier Ireland Ltd*. doi: 10.1016/j.cosrev.2024.100664.

[4] O. S. D. Fadhillah, J. H. Jaman, dan C. Carudin, "PERBANDINGAN NAIVE BAYES, SUPPORT VECTOR MACHINE, LOGISTIC REGRESSION DAN RANDOM FOREST DALAM MENGANALISIS SENTIMEN MENGENAI TIKTOKSHOP," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan 2025, doi: 10.23960/jitet.v13i1.5746.

[5] R. W. Harwenda, M. David Angelo, I. Budi, A. B. Santoso, dan K. Putra, "Sentiment Analysis on Government Public Policies: A Systematic Literature Review," *DIJEMSS*, vol. 6, no. 5, 2025, doi: 10.38035/dijemss.v6i5.

[6] Y. Albalawi, J. Buckley, dan N. S. Nikolov, "Investigating the impact of pre-processing techniques and pre-trained word embeddings in detecting Arabic health information on social media," *J Big Data*, vol. 8, no. 1, Des 2021, doi: 10.1186/s40537-021-00488-w.

[7] Y. HaCohen-Kerner, D. Miller, dan Y. Yigal, "The influence of preprocessing on text classification using a bag-of-words representation," *PLoS One*, vol. 15, no. 5, Mei 2020, doi: 10.1371/journal.pone.0232525.

[8] O. Rainio, J. Teuho, dan R. Klén, "Evaluation metrics and statistical tests for machine learning," *Sci Rep*, vol. 14, no. 1, Des 2024, doi: 10.1038/s41598-024-56706-x.

[9] V. Lumumba, D. Kiprotich, M. Mpaine, N. Makena, dan M. Kavita, "Comparative Analysis of Cross-Validation Techniques: LOOCV, K-folds Cross-Validation, and Repeated K-folds Cross-Validation in Machine Learning Models," *American Journal of Theoretical and Applied Statistics*, vol. 13, no. 5, hlm. 127–137, Okt 2024, doi: 10.11648/j.ajtas.20241305.13.

[10] L. Wegmeth, T. Vente, L. Purucker, dan J. Beel, "The Effect of Random Seeds for Data Splitting on Recommendation Accuracy," 2023. [Daring]. Tersedia pada: http://ceur-ws.org

[11] Y. Wahba, N. Madhavji, dan J. Steinbacher, "A Comparison of SVM against Pre-trained Language Models (PLMs) for Text Classification Tasks," Nov 2022.

[12] M. Pandey, S. Jalal, C. S. Negi, dan D. K. Yadav, "Using Ensemble and TOPSIS with AHP for Classification and Selection of Web Services," *Vietnam Journal of Computer Science*, vol. 9, no. 2, hlm. 217–243, Mei 2022, doi: 10.1142/S2196888822500130.

[13] U. Khalil, I. Imtiaz, B. Aslam, I. Ullah, A. Tariq, dan S. Qin, "Comparative analysis of machine learning and multi-criteria decision making techniques for landslide susceptibility mapping of Muzaffarabad district," *Front Environ Sci*, vol. 10, Sep 2022, doi: 10.3389/fenvs.2022.1028373.

[14] R. Obiedat, O. Harfoushi, R. Qaddoura, L. Al-Qaisi, dan A. M. Al-Zoubi, "An evolutionary-based sentiment analysis approach for enhancing government decisions during covid-19 pandemic: The case of jordan," *Applied Sciences (Switzerland)*, vol. 11, no. 19, Okt 2021, doi: 10.3390/app11199080.

[15] W. Cunha, L. Rocha, dan M. A. Gonçalves, "A thorough benchmark of automatic text classification: From traditional approaches to large language models," Apr 2025

[16] Tedyyana, A., Ghazali, O., Asnafi, T., Purbo, O. W., Harun, N. Z., & Riza, F. (2024). Transforming the voting process integrating blockchain into e-voting for enhanced transparency and securiy. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *22*(2), 311-320.