

Implementasi Metode *Support Vector Machine* untuk Klasifikasi Otomatis Pengaduan Publik di Kabupaten Trenggalek

¹Davin Zainur Robert, ²Ahmad Bagus Setiawan, ³Danang Wahyu Widodo

¹²³ Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ¹dappinro@gmail.com, ²ahmadbagus@unpkediri.ac.id,

³danangwahyuwidodo@unpkediri.ac.id

Penulis Korespondens : Davin Zainur Robert

Abstrak— Klasifikasi pengaduan masyarakat penting untuk meningkatkan efektivitas pelayanan publik daerah. Penelitian ini membangun model klasifikasi otomatis berdasarkan deskripsi aduan masyarakat di Kabupaten Trenggalek. Metode yang digunakan adalah *Support Vector Machine* (SVM) dengan *preprocessing* teks, mencakup tokenisasi, penghapusan tanda baca, *stopword removal*, dan *stemming*. Representasi teks dilakukan dengan *Term Frequency–Inverse Document Frequency* (TF-IDF). Hasil evaluasi menunjukkan bahwa model SVM mampu mengklasifikasikan pengaduan ke dalam kategori disposisi dinas dengan akurasi 77,4%, *precision* 79,6%, *recall* 77,4%, dan *F1-score* 77,5%. Pendekatan ini efektif untuk otomatisasi klasifikasi aduan masyarakat secara cepat dan akurat.

Kata Kunci— klasifikasi, pengaduan masyarakat, *support vector machine*, text mining, trenggalek

Abstract— *Public complaint classification plays a key role in improving local public service delivery. This study develops an automatic classification model based on complaint descriptions submitted to the Trenggalek District Government. The method used is Support Vector Machine (SVM), combined with text preprocessing including tokenization, punctuation and stopword removal, and stemming. Text representation is done using Term Frequency–Inverse Document Frequency (TF-IDF). Evaluation shows the SVM model classifies complaints into relevant agency categories with 77.4% accuracy, 79.6% precision, 77.4% recall, and 77.5% F1-score. The approach proves effective for fast and accurate automated complaint classification.*

Keywords— classification, public complaints, support vector machine, text mining, Trenggalek

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Pelayanan publik yang efektif dan responsif merupakan salah satu indikator utama kinerja pemerintah daerah dalam menjawab kebutuhan masyarakat [1]. Salah satu bentuk pelayanan tersebut adalah layanan pengaduan masyarakat, yang menjadi kanal penting untuk menyampaikan keluhan, laporan, maupun aspirasi warga terkait kondisi lingkungan, infrastruktur, atau layanan publik lainnya [2]. Di Kabupaten Trenggalek, sistem pengaduan masyarakat telah tersedia melalui kanal digital dan fisik, namun proses pengelolaan dan pengkategorian aduan masih dilakukan secara manual, sehingga mengakibatkan keterlambatan dalam respon dan tindak lanjut oleh instansi terkait.

Permasalahan umum yang dihadapi dalam pengelolaan laporan pengaduan adalah adanya variasi dalam struktur dan gaya penulisan yang digunakan oleh masyarakat. Banyak di antara laporan tersebut disusun dalam format teks bebas, menggunakan bahasa nonformal, bahkan dalam bahasa daerah, yang pada akhirnya mempersulit proses klasifikasi secara manual. Situasi ini menggarisbawahi pentingnya penerapan sistem klasifikasi otomatis yang mampu mengelompokkan jenis pengaduan dengan tingkat keakuratan dan kecepatan yang memadai. Salah satu solusi yang dinilai efektif untuk menjawab tantangan tersebut adalah pemanfaatan metode *Text Mining* yang dikombinasikan dengan algoritma klasifikasi seperti *Support Vector Machine (SVM)*, karena teknologi *Text Mining* telah menunjukkan efektivitasnya dalam mengidentifikasi dan mengolah informasi yang terkandung dalam data berbasis teks [3]. *Support Vector Machine (SVM)* merupakan salah satu metode klasifikasi yang bekerja dengan cara membangun hyperplane atau batas keputusan yang memisahkan data antar kelas secara optimal. *SVM* efektif untuk menangani data berdimensi tinggi seperti teks karena kemampuannya membentuk margin pemisahan maksimum antar kelas, bahkan ketika data tidak terdistribusi secara linier [4].

Sejumlah penelitian sebelumnya telah menunjukkan keberhasilan berbagai metode dalam klasifikasi data teks pengaduan masyarakat. Metode *Naive Bayes* dengan kombinasi seleksi fitur *Chi-Square* dan *Information Gain* telah diterapkan pada platform SAMBAT Online di Kota Malang dan menghasilkan akurasi sebesar 83,33% [5]. Algoritma *Naive Bayes Classifier* juga digunakan untuk klasifikasi teks pengaduan melalui layanan Call Center 110 dan memperoleh *precision* sebesar 90%, *recall* 93%, serta *f-measure* 92% [6]. Metode *Support Vector Machine (SVM)* diterapkan dalam analisis sentimen terhadap layanan Gojek di media sosial Twitter dan menghasilkan akurasi 79,19% menggunakan *kernel RBF* [7]. Sistem klasifikasi pengaduan masyarakat berbasis Organisasi Perangkat Daerah (OPD) di Kota Denpasar dikembangkan menggunakan pendekatan *Ensemble Naive Bayes* dan menghasilkan akurasi sebesar 91,82% [8]. Selain itu, metode *Maximum Entropy* digunakan untuk mengklasifikasikan teks pengaduan masyarakat pada platform Suara Warga Kabupaten Pasuruan dengan akurasi 89,27% [9].

Meskipun metode-metode tersebut telah banyak digunakan, belum terdapat penelitian yang secara khusus menerapkan *SVM* pada data pengaduan masyarakat di Kabupaten Trenggalek. Selain itu, masih terbatasnya penelitian yang mengimplementasikan tahapan *preprocessing teks* secara lengkap untuk bahasa Indonesia, seperti tokenisasi, penghapusan tanda baca, *stopword removal*, dan *stemming*, menjadi salah satu kesenjangan yang ingin dijawab dalam penelitian ini.

Mengacu pada permasalahan yang telah diuraikan, fokus dari penelitian ini adalah merancang sistem klasifikasi otomatis untuk data pengaduan masyarakat di Kabupaten Trenggalek dengan memanfaatkan algoritma *Support Vector Machine (SVM)*. Penelitian ini mencakup beberapa tujuan khusus, yaitu: (1) menerapkan tahapan *preprocessing teks* dalam bahasa Indonesia, seperti tokenisasi, penghapusan tanda baca, *stopword removal*, dan *stemming*, (2) membangun model klasifikasi berbasis SVM untuk data pengaduan masyarakat, serta (3) mengevaluasi performa model menggunakan metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score*. Dengan adanya sistem klasifikasi otomatis ini, diharapkan proses verifikasi dan penanganan pengaduan oleh dinas atau organisasi perangkat daerah terkait dapat dilakukan dengan lebih cepat, tepat, dan efisien.

II. METODE

Model klasifikasi otomatis terhadap pengaduan masyarakat berbasis teks dibangun menggunakan metode *Support Vector Machine (SVM)*. Penelitian dilakukan secara eksperimental berbasis data sekunder dari kanal aduan resmi Pemerintah Kabupaten Trenggalek. Subjek penelitian berupa teks aduan masyarakat yang memuat deskripsi masalah yang dilaporkan.

A. Materi dan Data Penelitian

Dataset yang digunakan dalam penelitian ini diperoleh secara langsung dari sistem pengaduan resmi milik Pemerintah Kabupaten Trenggalek. Pengumpulan data dilakukan dengan metode pengambilan acak berdasarkan laporan yang masuk dalam rentang waktu antara tahun 2024 hingga 2025. Data ini dikumpulkan dalam format *Microsoft Excel (.xlsx)* dan terdiri dari enam atribut, sebagaimana ditunjukkan pada Tabel 1.

Tabel 1. Struktur Dataset

Kolom	Tipe Data	Deskripsi
pengirim	Object	Nama atau identitas pengirim laporan pengaduan
tanggal	Object	Tanggal pengaduan dikirimkan (format tanggal dalam bentuk string)
judul	Object	Judul atau ringkasan singkat dari pengaduan
kategori	Object	Kategori aduan yang merujuk pada jenis layanan atau instansi terkait
deskripsi	Object	Uraian atau narasi detail dari isi pengaduan masyarakat
disposisi	Object	Label target berupa instansi atau dinas yang bertanggung jawab menangani

B. Desain Eksperimen dan Prosedur Penelitian

Rangkaian proses klasifikasi teks dilakukan secara sistematis dimulai dari preprocessing data teks, ekstraksi fitur menggunakan *TF-IDF*, pembagian data menjadi data latih dan uji, pelatihan model *SVM* dengan *kernel linear*, hingga evaluasi menggunakan metrik performa klasifikasi.

1. Preprocessing Teks

Preprocessing merupakan tahapan awal yang penting untuk menyiapkan data teks agar dapat dianalisis secara numerik. Beberapa langkah yang dilakukan dalam proses ini dapat dilihat pada Tabel 2.

Tabel 2. Tahapan pada Preprocessing Text

Tahapan	Struktur
Tokenisasi	Memecah kalimat aduan menjadi token (kata) menggunakan wordpunct_tokenize. Contoh: "jalan rusak parah" → ['jalan', 'rusak', 'parah']

Tahapan	Struktur
Penghapusan Tanda Baca	Menghapus karakter seperti titik, koma, atau tanda seru menggunakan modul string.punctuation.
Stopword Removal	Menghilangkan kata-kata umum seperti "yang", "dan", "di" yang tidak memiliki nilai klasifikasi tinggi, menggunakan kamus stopwords dari nltk.
Stemming	Mengubah kata ke bentuk dasarnya menggunakan PorterStemmer. Contoh: <i>berjalan</i> → <i>jalan</i>

2. Ekstraksi Fitur

Dokumen hasil *preprocessing* dikonversi ke dalam bentuk numerik menggunakan TF-IDF. Ekstraksi fitur dilakukan menggunakan metode TF-IDF untuk mengukur tingkat kepentingan kata dalam dokumen. *TF (Term Frequency)* mencerminkan seberapa sering suatu kata muncul dalam dokumen, sedangkan *IDF (Inverse Document Frequency)* menghitung sejauh mana kata tersebut unik di antara seluruh dokumen. Penerapan TF-IDF membantu menyaring kata-kata yang paling informatif dalam data teks, sekaligus mengurangi dominasi kata umum yang kurang bermakna dalam proses klasifikasi[10].

Rumus perhitungannya sebagai berikut:

$$tf - idf(t, d) = tf(t, d) \times \left(\frac{N}{df(t)} \right) \quad (1)$$

Dimana:

- a. $tf(t, d)$: frekuensi kemunculan term t dalam dokumen d
- b. N : jumlah seluruh dokumen
- c. $df(t)$: jumlah dokumen yang mengandung term t

3. Pembagian Data

Dataset dibagi menjadi 80% data latih dan 20% data uji menggunakan fungsi *train_test_split* dari pustaka *sklearn.model_selection*. Pembagian ini bertujuan agar model dapat dilatih menggunakan sebagian besar data, kemudian diuji kinerjanya pada data yang belum pernah dilihat sebelumnya.

Data latih merupakan bagian dari dataset yang dimanfaatkan untuk melatih model dalam mengenali serta memahami pola yang terdapat pada data[11]. Sementara itu, data uji digunakan untuk mengevaluasi kemampuan generalisasi model[12]. Fungsi *train_test_split* digunakan untuk membagi data secara otomatis dan acak sesuai dengan proporsi yang telah ditentukan dalam proses pelatihan dan pengujian model [13]. Modul *sklearn.model_selection* sendiri merupakan bagian dari pustaka *scikit-learn* yang menyediakan berbagai metode untuk pembagian data dan validasi model.

4. Pelatihan Model

Algoritma yang digunakan adalah *Support Vector Machine (SVM)* dengan *kernel linear*. Fungsi keputusan SVM dijelaskan dengan:

$$f(x) = \text{sign}(w \cdot x + b) \quad (2)$$

Dimana:

- a. w : vector bobot
- b. x : vektor fitur hasil TF-IDF
- c. b : bias

C. Evaluasi Model

Model klasifikasi yang telah dibangun dievaluasi menggunakan beberapa metrik performa yang umum digunakan dalam klasifikasi teks, yaitu:

1. Akurasi

Menunjukkan proporsi keseluruhan prediksi yang benar terhadap jumlah total data.

$$\text{Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Seluruh Prediksi}} \quad (3)$$

2. Precision

Mengukur tingkat ketepatan prediksi positif, yaitu seberapa banyak prediksi positif yang benar-benar relevan.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

3. Recall

Mengukur kemampuan model dalam menemukan semua kasus positif yang relevan dalam data.

$$\text{Precision} = \frac{TP}{TP + FN} \quad (5)$$

4. F1-score

Merupakan rata-rata harmonis dari *precision* dan *recall*, digunakan ketika diperlukan keseimbangan antara keduanya.

$$F1 - score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

III. HASIL DAN PEMBAHASAN

Metode *Support Vector Machine (SVM)* diimplementasikan untuk klasifikasi pengaduan masyarakat berdasarkan isi teks dalam kolom deskripsi masalah ke dalam label disposisi dinas. Dataset yang digunakan terdiri dari 1.153 data pengaduan, yang diperoleh dari sistem pengaduan resmi Pemerintah Kabupaten Trenggalek dalam rentang waktu tahun 2024 hingga 2025. Tahapan analisis diawali dengan eksplorasi struktur data, dilanjutkan dengan *preprocessing teks* sebagai input utama untuk model klasifikasi.

A. Struktur Dataset

Dataset yang digunakan terdiri dari tujuh atribut utama, yaitu: nomor, pengirim, tanggal pengaduan, judul laporan, kategori pengaduan, deskripsi masalah, dan disposisi dinas sebagai label target. Kolom Deskripsi merupakan fokus utama dalam proses klasifikasi karena memuat narasi masalah yang dilaporkan masyarakat. Sementara itu, kolom Disposisi berperan sebagai kelas atau label yang menunjukkan instansi atau dinas mana yang berwenang menangani laporan tersebut. Struktur dan contoh isi dari dataset ditunjukkan pada Tabel 3.

Tabel 3. Isi Dataset

No	Pengirim	Tanggal	Judul	Kategori	Deskripsi	Disposisi
1	Ibu Posyandu	17/06/2025	Tidak Ada Pemeriksaan Ibu Hamil Selama 2 Bulan	Layanan Kesehatan Masyarakat	Beberapa posyandu di desa-desa tidak mengadakan pemeriksaan ibu hamil rutin selama lebih dari 2 bulan.	DINAS KESEHATAN, PENGENDALIAN PENDUDUK DAN KELUARGA BERENCANA
...				
11	Anonim - 93FSQP	05/01/2024	Pohon Tumbang Menimpa Rumah dan Menutup Jalan Umum	Angin Kencang	Akibat hujan lebat dan angin kencang, pohon besar tumbang menimpa rumah dan menutup akses jalan warga di Desa Kendalrejo.	BADAN PENANGGULANGAN BENCANA DAERAH
53						

B. Preprocessing

Sebelum pelatihan model dilakukan, data teks pada kolom deskripsi diproses melalui tahapan *preprocessing* untuk menyiapkan data agar dapat dianalisis secara *numerik*. Proses ini bertujuan untuk menyederhanakan teks, mengurangi *noise*, dan memperkuat kualitas fitur yang akan diekstraksi. Contoh hasil sebelum dan sesudah preprocessing ditunjukkan pada Tabel 4.

Tabel 4. Preprocessing

No	Sebelum Preprosesing	Setelah Preprocessing
1	Beberapa posyandu di desa-desa tidak mengadakan pemeriksaan ibu hamil rutin selama lebih dari 2 bulan.	beberapa posyandu desa desa mengadakan pemeriksaan hamil rutin 2
...
1153	Akibat hujan lebat dan angin kencang, pohon besar tumbang menimpa rumah dan menutup akses jalan warga di Desa Kendalrejo.	akibat hujan lebat angin kencang pohon tumbang menimpa rumah menutup akses jalan warga desa kendalrejo

C. Ekstraksi Fitur *TF-IDF*

Teks yang telah melalui tahap *preprocessing* dikonversi ke dalam bentuk numerik menggunakan metode *Term Frequency–Inverse Document Frequency (TF-IDF)*. Proses ini bertujuan untuk merepresentasikan dokumen dalam bentuk vektor yang mencerminkan pentingnya suatu kata atau frasa dalam keseluruhan kumpulan data. Ekstraksi fitur dilakukan dengan mempertimbangkan kombinasi kata tunggal (*unigram*) dan pasangan kata berurutan (*bigram*). Jumlah fitur dibatasi hingga 20.000 fitur paling relevan, dengan kata yang terlalu umum tidak disertakan. Hasil representasi ditampilkan pada Tabel 5.

Tabel 5. Ekstraksi Fitur TF-IDF

No	<i>air</i>	<i>air bantuan</i>	<i>air bersih</i>	<i>akibat</i>	...	<i>warga menyebabkan</i>	<i>warga terdampak</i>	<i>wilayah</i>	<i>wilayah pegunungan</i>
1	0	0	0	0	...	0	0	0	0
2	0	0	0	0	...	0	0	0.208	0.300
3	0	0	0	0	...	0.245	0	0	0
...					...				
1151	0	0	0	0	...	0	0.221	0	0
1152	0.209	0.208	0.177	0	...	0	0	0	0
1153	0	0	0	0.154	...	0	0	0	0

D. Hasil Evaluasi

Evaluasi performa model dilakukan menggunakan empat metrik utama, yaitu akurasi, *precision*, *recall*, dan *F1-score*. Hasil evaluasi disajikan pada Tabel 6.

Tabel 6. Hasil Evaluasi

<i>Metrik</i>	<i>Evaluasi</i>
<i>Akurasi</i>	0.774
<i>Precision</i>	0.796
<i>Recall</i>	0.774
<i>F1-score</i>	0.775

Hasil ini menunjukkan bahwa model memiliki kinerja yang baik dalam mengklasifikasikan jenis aduan berdasarkan deskripsi teks. Nilai akurasi sebesar 77,4% menandakan bahwa sebagian besar data uji berhasil diprediksi dengan tepat. Selain itu, nilai *precision*, *recall*, dan *F1-score* yang seimbang menunjukkan bahwa model bersifat stabil, tidak bias terhadap kelas tertentu, dan mampu mempertahankan keseimbangan antara kemampuan mengenali aduan yang relevan dan menghindari kesalahan klasifikasi.

E. Pengujian Pengaduan Baru

Tabel 7. Hasil Pengujian Pengaduan Baru

No	Pengaduan Baru	Disposisi	Kesimpulan
1	trotoar di pusat kota rusak dan digunakan untuk parkir liar, menyulitkan pejalan kaki	DINAS PERHUBUNGAN	Tidak Valid
2	tidak ada event budaya tahunan seperti tahun lalu, padahal itu mendatangkan wisatawan	DINAS PARIWISATA DAN KEBUDAYAAN	Valid
3	banyak ASN sering datang terlambat dan pulang lebih awal tanpa sanksi tegas	BADAN KEPEGAWAIAN DAERAH	Valid
4	pipa air bersih bocor di jalan raya dan belum diperbaiki hingga seminggu	PDAM	Valid
5	alat tangkap bantuan untuk nelayan tidak sesuai spesifikasi dan cepat rusak	DINAS PERIKANAN	Valid
...
14	data kependudukan desa belum terhubung ke sistem online kecamatan	DINAS KOMUNIKASI DAN INFORMATIKA	Valid
15	perumahan baru tidak punya saluran limbah, menyebabkan bau tak sedap	DINAS PERUMAHAN, KAWASAN PERMUKIMAN DAN LINGKUNGAN HIDUP	Valid
16	tidak ada lagi sosialisasi KB di desa kami dan posyandu jarang aktif	DINAS KESEHATAN, PENGENDALIAN PENDUDUK DAN KELUARGA BERENCANA	Valid

17	Warung makan di area terminal melanggar aturan karena membuang limbah sembarangan ke selokan, mengganggu ketertiban umum dan kesehatan warga sekitar.	DINAS PERUMAHAN, KAWASAN PERMUKIMAN DAN LINGKUNGAN HIDUP	Tidak Valid
18	koleksi buku di perpustakaan desa sudah usang dan belum diperbarui	DINAS KEARSIPAN DAN PERPUSTAKAAN	Valid

Berdasarkan pengujian 18 pengaduan dalam Tabel 7, model klasifikasi disposisi menunjukkan performa yang baik dengan 16 dari 18 prediksi valid (rasio 8:9). Model mampu mengenali kata kunci sektoral dengan cukup tepat, namun masih mengalami kekeliruan pada dua pengaduan yang melibatkan tumpang tindih kewenangan antar dinas. Misalnya, pengaduan trotoar rusak diarahkan ke Dinas Perhubungan padahal seharusnya ke Dinas PUPR, dan pelanggaran pembuangan limbah oleh warung makan diarahkan ke Dinas Lingkungan Hidup padahal lebih tepat ke Satpol PP. Temuan ini menunjukkan model cukup andal namun masih perlu ditingkatkan dalam memahami konteks lintas sektor dan aspek pelanggaran.

IV. KESIMPULAN

Model klasifikasi otomatis terhadap data pengaduan masyarakat di Kabupaten Trenggalek menggunakan metode *Support Vector Machine (SVM)*. Proses *preprocessing* yang meliputi tokenisasi, penghapusan tanda baca, *stopword removal*, dan *stemming*, terbukti meningkatkan kualitas representasi teks secara numerik melalui metode *Term Frequency–Inverse Document Frequency (TF-IDF)*. Model SVM yang dibangun menunjukkan performa yang baik dengan nilai akurasi sebesar 77,4%, precision 79,6%, recall 77,4%, dan F1-score 77,5%. Hal ini menandakan bahwa model mampu melakukan klasifikasi dengan cukup seimbang dan efektif dalam mengenali pola dari deskripsi aduan masyarakat.

Penerapan sistem klasifikasi otomatis ini dapat memberikan kontribusi signifikan dalam mendukung proses pengelolaan pengaduan publik oleh instansi pemerintah daerah. Dengan adanya klasifikasi berbasis mesin, proses verifikasi aduan dapat dilakukan secara lebih cepat, efisien, dan akurat, sehingga mempercepat respons dan tindak lanjut dari laporan masyarakat. Ke depan, sistem ini dapat dikembangkan lebih lanjut dengan mengintegrasikannya ke dalam platform pengaduan digital yang telah tersedia, serta mengeksplorasi metode klasifikasi lain untuk meningkatkan akurasi dan stabilitas model.

DAFTAR PUSTAKA

- [1] Hamdillah, Hamdun. "Inovasi Pelayanan Publik dan Transformasi Birokrasi: Pendekatan Administrasi Publik dalam Meningkatkan Good Governance." *Resolusi: Jurnal Sosial Politik*, vol. 6, no. 2, 2023, doi: 10.32699/resolusi.v6i2.5672
- [2] Rahmia, Hilda Aulia, Telumab, Aurelius R.L., dan Hadi, Agus Purbathin. "Implementasi Komunikasi Pelayanan Publik Pemerintah Kota Mataram Melalui Aplikasi LAPOR!" *Tuturlogi*, vol. 1, no. 2, 2020. <http://dx.doi.org/10.21776/ub.tuturlogi.2020.001.02.4>

- [3] Kurnia, Y., Kusuma, E. D., Kusuma, L. W., Suwitno, & Apridius, W. (2024). *Perbandingan Naïve Bayes dan CNN yang Dioptimasi PSO pada Identifikasi Berita Hoax Politik Indonesia*. Buletin Teknologi, 6(3). <https://doi.org/10.32877/bt.v6i3.1225>
- [4] Tono, J. J., & Parjito, P. (2023). *Persepsi Publik terhadap Kepemimpinan Firli Bahuri di KPK: Pendekatan Sentimen Twitter dengan Naïve Bayes dan SVM*. Jurnal Ilmu Politik dan Ilmu Pemerintahan, 10(2). <https://doi.org/10.29100/jipi.v10i2.6181>
- [5] Idris, H. K., Fauzi, M. A., & Indriati, I. (2019). Klasifikasi Teks Bahasa Indonesia Pada Dokumen Pengaduan SAMBAT Online Menggunakan Metode Naive Bayes dan Kombinasi Seleksi Fitur. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(3), 2436–2442. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/4681>
- [6] Handayani, F., & Pribadi, F. S. (2018). Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110. Jurnal Teknik Elektro, 7(1). <https://journal.unnes.ac.id/nju/jte/article/view/8585>
- [7] Fitriyah, N., Warsito, B., & Maruddani, D. A. I. (2021). Analisis Sentimen Gojek pada Media Sosial Twitter dengan Klasifikasi Support Vector Machine (SVM). Departemen Statistika, FSM Universitas Diponegoro.
- [8] Triawan, A. Y., Wicaksono, S. A., & Purnomo, W. (2019). Sistem Klasifikasi Pengaduan Masyarakat Berdasarkan OPD Menggunakan Pendekatan Ensemble Naive Bayes. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(7), 7197–7204. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/5894>
- [9] Rini, M. P., Adikara, P. P., & Adinugroho, S. (2019). Klasifikasi Teks Pengaduan Suara Warga Kabupaten Pasuruan menggunakan Metode Maximum Entropy. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(8), 7774–7778. <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/5957>
- [10] Supardianto, Seriasih, dan L. Mutawalli, “Pemodelan Topik Isu Parwisata Pulau Lombok Menggunakan Focus Web Crawler dan Latent Dirichlet Allocation,” *Jurnal Ilmu Rekayasa Elektrikal*, vol. 7, no. 2, pp. 110–117, 2023. <https://doi.org/10.36595/jire.v7i2.1332>
- [11] A. Fatunnisa dan H. Marcos, “Prediksi Kelulusan Tepat Waktu Siswa SMK Teknik Komputer Menggunakan Algoritma Random Forest,” *Jurnal Manajemen Informatika (JAMIKA)*, vol. 14, no. 1, 2024. Tersedia: <https://doi.org/10.34010/jamika.v14i1.12114>
- [12] W. A. Naseer, S. Sarwido, dan B. B. Wahono, “Gradient Boosting Optimization with Pruning Technique for Prediction of BMT Al-Hikmah Permata Customer Data,” *Jurnal Inovasi dan Teknologi Sistem Informasi*, vol. 6, no. 3, 2024.. Tersedia: <https://doi.org/10.51401/jinteks.v6i3.4702>
- [13] E. J. Sudarman and S. Budi, "Pengembangan Model Kecerdasan Mesin Extreme Gradient Boosting untuk Prediksi Keberhasilan Studi Mahasiswa," *Strategi*, vol. 5, no. 2, Nov. 2023.
- [14] L. A. Susanto, “Komparasi Model Support Vector Machine dan K-Nearest Neighbor pada Analisis Sentimen Aplikasi POLRI Super App,” *Jurnal Ilmiah Teknik Elektro Terapan*, vol. 12, no. 2, pp. 85–91, 2023. <https://doi.org/10.23960/jitet.v12i2.4152>