

PENDEKATAN BERBASIS *DECISION TREE* DENGAN POLA WILDCARD UNTUK MENGIDENTIFIKASI PLAT NOMOR KENDARAAN PULAU JAWA

Chaesa Akmal Aryandra^{1*}, Diadora Sanggrahita², Safira Damayanti Kusumaningrum³, Iqbal Al Ahyubi⁴, Nur Iksan⁵.

^{1,2} Universitas Negeri Semarang; Sekaran, Kecamatan Gunung Pati, Kota Semarang, Jawa Tengah; chaemal09@students.unnes.ac.id

^{3,4} Universitas Negeri Semarang; Sekaran, Kecamatan Gunung Pati, Kota Semarang, Jawa Tengah; safiradamayanti04@students.unnes.ac.id

⁵ Universitas Negeri Semarang; Sekaran, Kecamatan Gunung Pati, Kota Semarang, Jawa Tengah; nur.iksan@mail.unnes.ac.id

Keywords:

DecisionTree;
Wildcard;
Kendaraan;
Jawa.

Correspondent Email:

chaemal09@students.unnes.ac.id

Abstrak. Pengembangan sistem klasifikasi wilayah berdasarkan kode plat nomor kendaraan di Pulau Jawa. Pentingnya penelitian ini terletak pada kebutuhan untuk mengidentifikasi asal wilayah kendaraan secara otomatis guna mendukung sistem informasi dan analisis data transportasi. Metode yang digunakan mencakup pengumpulan data plat nomor, preprocessing data dengan penghapusan karakter khusus dan pemisahan elemen, serta klasifikasi menggunakan pendekatan berbasis aturan wildcard dan algoritma pembelajaran mesin Decision Tree. Karakter wildcard () digunakan untuk membentuk pola deteksi huruf depan atau belakang plat nomor, yang dikonversi ke simbol REGEX untuk mempermudah pencocokan. Evaluasi dilakukan dengan skema train-test split 80:20, menghasilkan akurasi sebesar 58%. Hasil menunjukkan bahwa model lebih akurat pada kota/kabupaten dengan jumlah data lebih besar. Temuan ini menunjukkan bahwa kombinasi pendekatan wildcard dan Decision Tree cukup efektif dalam klasifikasi asal wilayah kendaraan, dan mengindikasikan bahwa keberagaman serta volume data sangat berpengaruh terhadap performa model klasifikasi.

Abstract. Development of a regional classification system based on vehicle license plate codes in Java Island. The importance of this research lies in the need to automatically identify the origin of vehicles to support information systems and transportation data analysis. The methods used include license plate data collection, data preprocessing by removing special characters and separating elements, and classification using rule-based wildcard patterns and the Decision Tree machine learning algorithm. Wildcard characters (*) are used to form detection patterns for the front or back letters of license plates, which are then converted into REGEX symbols to facilitate pattern matching. Evaluation is carried out using a train-test split scheme with an 80:20 ratio, resulting in an accuracy of 58%. The results show that the model performs better for cities/regencies with more abundant data. These findings indicate that the combination of wildcard and Decision Tree approaches is effective in classifying vehicle origin regions and highlight the significant

influence of data diversity and volume on the classification model's performance.

1. PENDAHULUAN

Pulau Jawa merupakan salah satu pulau terbesar di Indonesia. Pulau Jawa merupakan pulau terbesar kelima di Indonesia [1]. Di mana ini sebanding dengan banyaknya penduduk yang tinggal di pulau Jawa. Menurut Badan Pusat Statistik Aceh, provinsi DKI Jakarta sebanyak 10.684,9 ribu jiwa, Jawa Barat sebanyak 50.345,2 ribu jiwa, Banten sebanyak 12.431,4 ribu jiwa, Jawa Tengah sebanyak 37.892,3 ribu jiwa, DI Yogyakarta sebanyak 3.759,5 ribu jiwa dan Jawa Timur sebanyak 41.814,5 ribu jiwa. Dengan populasi lebih dari 150 juta jiwa, pulau jawa menjadi wilayah dengan tingkat kepadatan penduduk tertinggi di Indonesia.

Pulau Jawa merupakan pusat Industri dan Perdagangan di Indonesia [2]. Selain dari segi ekonomi, pulau Jawa merupakan pusat kegiatan politik, pendidikan dan budaya di Indonesia. Kepadatan tersebut berpengaruh pada mobilitas yang terjadi pad pulau Jawa. Kepadatan tersebut berpengaruh dalam meningkatnya volume kendaraan bermotor yang sangat tinggi. Fenomena ini menimbulkan tantangan dalam pengelolaan dan pengawasan kendaraan, terutama di kawasan perkotaan dengan lalu lintas yang padat. Oleh karena itu, pengelolaan data kendaraan yang efisien menjadi aspek penting dalam sistem transportasi modern.

Salah satu data yang berperan penting dalam manajemen kendaraan adalah informasi plat nomor. Plat nomor kendaraan di Indonesia disusun berdasarkan kode wilayah. Di mana data ini menyimpan data logistic wilayah yang berfungsi dalam pemetaan penyebaran penduduk. Plat nomor kendaraan merupakan identitas formal yang disebut sebagai Tanda Nomor Kendaraan Bermotor (TNKB) [3].

Dengan banyaknya data wilayah berdasarkan plat nomor kendaraan menuntut adanya efisiensi dalam klasifikasi wilayah. Dengan konteks inilah, teknologi berperan penting dalam efisiensi kinerja manusia untuk mengklasifikasi secara otomatis terhadap identifikasi asal plat nomor secara cepat, akurat,

dan adaptif. Teknologi adalah faktor terbesar yang mempengaruhi produktivitas tenaga kerja. Teknologi membantu dalam meningkatkan efisiensi dan efektivitas proses produksi sehingga *output* yang dihasilkan lebih besar [4].

Dalam konteks ini, teknologi *mechine learning* berperan penting aktif. Algoritma *Decision Tree* merupakan pendekatan yang digunakan dalam studi klasifikasi penelitian ini karena keandalannya dalam mengenali pola dari data. *Decision Tree* membentuk struktur pohon berbasis keputusan logis yang dipahami dan diinterpretasikan secara manusiawi.

Penelitian ini bertujuan menerapkan algoritma pendekatan *Decision Tree* dalam membangun sistem klasifikasi wilayah berdasarkan plat nomor kendaraan wilayah pulau Jawa. Dengan pendekatan ini, diharapkan sistem mampu dalam membangun keakuratan data dan efisiensi kinerja sistem sehingga dapat membantu manusia dalam memantau mobilitas manusia melalui sebuah plat nomor kendaraan.

2. TINJAUAN PUSTAKA

2.1 Text PreProcessing

Preprocessing merupakan tahap awal sebelum memasuki pola pengenalan dan klasifikasi. Di mana dalam suatu penelitian [5] *preprocessing* teks digunakan untuk meningkatkan kualitas data dan mempermudah proses analisis. Di mana *preprocessing* ini mencakup *cleaning* atau penghapusan atribut serta karakter yang tidak diperlukan. *Case folding* untuk mengubah seluruh teks menjadi huruf yang seragam. Dilanjutkan dengan *tokenizing* yang memecah teks menjadi kata-kata individual. Serta *stemming* yang mengubah kata menjadi bentuk dasar.

Berbagai penelitian sudah membahas mengenai pentingnya proses *preprocessing* ini. Di mana dalam sebuah penelitian, *preprocessing* digunakan untuk menghapus kata-kata atau karakter yang tidak digunakan dalam sebuah ulasan[6][7]. *Preprocessing* digunakan juga dalam sebuah penelitian, di mana digunakan sebagai pembersih label, pada

penelitiannya dataset IHSG memiliki 5 fitur menjadi 3 fitur yaitu *date*, *open*, dan *close* [8]. *Preprocessing* juga digunakan dalam sebuah penelitian mengenai peringkasan teks, di mana *preprocessing* ini digunakan untuk mengidentifikasi kata berdasarkan token dan melakukan *cleaning* teks [9].

2.2 Klasifikasi

Klasifikasi merupakan proses untuk mengidentifikasi suatu pola [5]. Klasifikasi merupakan suatu proses untuk sebuah kalimat dikelompokkan ke dalam satu atau lebih kategori [10]. Sebuah penelitian menggunakan klasifikasi dalam mendeteksi label berita hoaks kampanye pemilihan sehingga dapat mengetahui seberapa besar data yang tervalidasi hoaks [11].

2.3 Pola Wildcard

Pola wildcard merupakan jenis pola pencocokan teks yang menggunakan simbol khusus seperti bintang (*) untuk mewakili satu atau lebih karakter. Dalam sebuah penelitian oleh [12] pola wildcard digunakan dalam SQL untuk mewakili karakter atau sekumpulan karakter tertentu dalam melakukan pencarian data sehingga memudahkan dalam pencocokan pola yang tidak pasti. *Wildcard* tidak hanya berfokus pada karakter (*) namun juga dalam berbagai bentuk karakter unik. Karakter *Wildcard* merupakan karakter khusus yang digunakan sebagai pengganti karakter yang tidak diketahui oleh sebuah teks [13]. Dalam sebuah penelitian dibahas mengenai penambahan wildcard digunakan untuk melakukan pemblokiran terhadap sebuah situs website [14].

2.4 Decision Tree

Decision Tree merupakan algoritma yang bekerja dengan memecah dataset ke dalam subset-subset berdasarkan atribut tertentu. *Decision Tree* merupakan metode pemodelan prediktif dalam analisis data menggunakan struktur pohon [15]. *Decision Tree* ini berhasil dalam mengklasifikasi suatu *rating* film dari data Netflix [15]. Algoritma ini digunakan pada sebuah penelitian, di mana *Decision Tree* efektif dalam menganalisis data *kuesioner* sehingga mampu dalam meningkatkan strategi layanan dan kualitas produk suatu perusahaan [16]. Dalam sebuah penelitian [17] algoritma

ini menghasilkan akurasi 96.15% yang dimana algoritma ini sesuai untuk sebuah klasifikasi data dibandingkan dengan algoritma SVM yang memiliki akurasi 62.48%. Dalam konteks pengolahan data kendaraan, algoritma ini berpotensi sangat besar dalam mengklasifikasikan wilayah asal kendaraan berdasarkan kode plat secara akurat dan efisien. Hal ini sejalan dengan karakteristik *Decision Tree* yang mampu mengolah data simbolik seperti kode plat kendaraan yang menyajikan logika klasifikasi secara eksplisit sehingga mendukung kebutuhan sistem transportasi yang cerdas dengan berbasis data.

Dalam proses klasifikasi plat nomor kendaraan di pulau jawa, algoritma *Decision Tree* dapat digunakan untuk menyusun dalam pengelompokkan kendaraan berdasarkan kode wilayah, seperti huruf awal, deretan angka, dan huruf akhir. Bagian-bagian tersebut diambil dari plat nomor dan dijadikan sebagai fitur input dalam model. Kemudian algoritma *Decision Tree* membentuk pola-pola keputusan guna menentukan asal kendaraan, baik berdasarkan provinsi maupun kota di wilayah pulau jawa. pendekatan ini sangat cocok diterapkan dalam sistem pengenalan plat nomor otomatis, sistem tilang berbasis elektronik, hingga sistem parkir pintar, karena mampu memberikan proses klasifikasi yang efisien, jelas, dan mudah untuk diinterpretasikan.

Hubungan *Decision Tree* dengan penelitian ini terletak pada kemampuan dalam membentuk model klasifikasi yang didasari pada ciri-ciri simbolik dari plat nomor, seperti huruf awal, deretan angka, dan huruf akhir. dalam penelitian ini, *Decision Tree* dimanfaatkan untuk merancang struktur pengambilan keputusan bertingkat yang mengklasifikasikan kendaraan berdasarkan asal wilayah, mulai dari provinsi hingga kabupaten atau kota di pulau jawa. pendekatan ini memungkinkan sistem untuk menjalankan proses klasifikasi dengan cara yang sistematis, efisien, dan mudah dipahami. maka dari itu, penggunaan algoritma *Decision Tree* sangat tepat untuk mendukung tujuan utama dalam penelitian ini. Dalam penelitian [18] disebutkan mengenai kelebihan dan kekurangan dalam penggunaan *Decision Tree*.

a. Kelebihan *Decision Tree*

Decision Tree merupakan algoritma yang sangat fleksibel karena dapat

diterapkan untuk tugas klasifikasi maupun regresi. Dimana algoritma ini mampu mengelola berbagai tipe data, baik diskrit, kontinu, maupun kategorikal. Lain hal itu *Decision Tree* membantu dalam mengeksplorasi berbagai kemungkinan solusi terhadap suatu masalah. Dibandingkan dengan algoritma yang lain, metode ini umumnya menggunakan proses pembersihan data (*data cleaning*) yang lebih sedikit untuk mengatasi kesalahan atau ketidakkonsistenan yang ada pada data.

b. Kekurangan *Decision Tree*

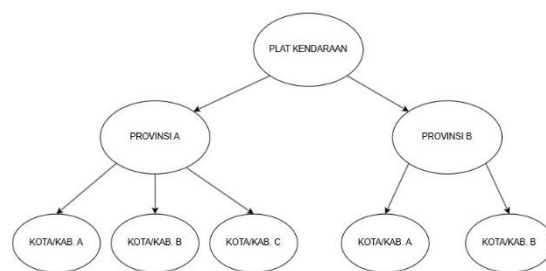
Meskipun efektif, *Decision Tree* mempunyai banyak lapisan yang membuat rumit fan beresiko mengalami *overfitting*, yaitu keadaan ketika model tidak mampu memberikan prediksi yang akurat pada data baru. Selain itu, dukungan pada *Decision Tree* di *Scikit-learn*, salah satu pustaka *machine learning* populer berbasis python masih terbatas. Algoritma ini juga kurang cocok untuk dataset yang berskala besar, karena proses pelatihannya yang membutuhkan waktu yang cukup lama, sumber daya yang lebih besar, serta kompleksitas model yang bisa meningkat secara signifikan. Tanpa disadari prinsip kerja *Decision Tree* sering diaplikasikan dalam kehidupan sehari-hari, seperti contohnya klasifikasi plat nomor kendaraan yang ada di wilayah pulau jawa.

Hal tersebut juga dikuatkan oleh [19] di mana *decision tree* memiliki kelebihan di felsibilitasnya sehingga membuat keputusan yang diambil dapat meningkat dan kekurangannya ada pada *overlap* jika data yang diuji memiliki begitu banyak kelas. Algoritma ini digunakan juga dalam sebuah penelitian keputusan latihan memanah yang memberikan rata-rata precision sebesar 88,89% dengan menggunakan algoritma tree [20].

3. METODE PENELITIAN

Metode ini menggunakan metode pengumpulan data melalui media-media penyedia kode plat kendaraan yaitu perpol No.7 tahun 2021 tentang regident di mana di sini terdapat kode-kode plat nomor kendaraan di

Indonesia, khususnya pulau Jawa. Selanjutnya diikuti dengan *preprocessing* untuk membuat data tersebut bersih dari karakter-karakter yang tidak diperlukan, labelisasi dan tokenisasi. Setelah itu, dilakukan transformasi transformasi teks menggunakan pembobotan kata TF-IDF. Klasifikasi dalam penelitian ini menggunakan klasifikasi *Decision Tree* dengan pola Wildcard.



Gambar 1 Alur Pemrosesan Input

Secara singkat proses validasi inputan data dengan dataset yang sudah diklasifikasiaskan seperti pada Gambar 1. Di mana ketika ada sebuah inputan baru, sistem mulai melihat kode plat kendaraan awal, lalu mengklasifikasikannya ke dalam provinsi mana. Selanjutnya ketika sudah diketahui masuk kedalam kelas provinsi mana maka sistem akan lanjut mengklasifikasikannya ke dalam kota/kabupaten mana berdasarkan kode plat belakang.

3.1 Pengumpulan Data

Penelitian ini menggunakan data kode plat nomor yang didapat melalui berbagai media yang selanjutnya diambil wilayah pulau Jawa. Label yang digunakan dalam dataset kode nomor kendaraan wilayah Jawa ini adalah label Provinsi, Kode Plat, Kota/Kabupaten dan Kode Belakang.

3.2 Text Preprocessing

Tahap ini bertujuan untuk meningkatkan kualitas teks dengan membersihkan dataset dari kode atau karakter yang tidak diperlukan. Selanjutnya dilakukan *Tokenizing* untuk memecah kalimat menjadi kata per kata. Dalam *preprocessing* ini untuk karakter * bintang tetap dipertahankan. Hal ini dikarenakan data memerlukan klasifikasi identifikasi secara manual dengan menggunakan pola wildcard

untuk mencocokkan kode belakang saat pencarian berbasis aturan.

3.3 Ekspansi dan Label Encoding

Dataset kemudian diubah kedalam format yang dapat terdeteksi oleh mesin. Label yang digunakan yaitu label Provinsi, Kode Plat, Kota/Kabupaten, dan Kode belakang. Setelah itu data di ekspansi menjadi format tabular yang menampilkan fitur-fitur penting seperti huruf belakang utama, pembuatan ekspresi reguler (regex) untuk mendukung pencocokan pola dinamis, dan pola posisi oleh *wildcard* yang kemudian dikodekan menggunakan Label Encoding.

3.4 Klasifikasi Decision Tree

Model klasifikasi dengan empat fitur utama yaitu kode plat, huruf belakang, posisi pola wildcard, dan provinsi sedangkan labelnya bertargetkan kota atau kabupaten asal kendaraan. Model dilatih menggunakan data yang telah dibagi menjadi rasio 80:20, serta dievaluasi menggunakan metrik akurasi, *confusion matrix*, dan *classification report*.

4. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan sebuah sistem klasifikasi wilayah yang berdasarkan kode plat nomor kendaraan dengan domisili pada pulau jawa. Sistem dirancang untuk dapat mendeteksi pola huruf belakang dan depan dari plat nomor kendaraan yang sesuai dengan kode dari plat tersebut berdasarkan kota/kabupaten dan provinsi dari hasil sistem penelitian. Sistem yang kami bangun memanfaatkan serangkaian tahapan, dimulai dari pengumpulan data kode plat, *teks preprocessing*, hingga penerapan algoritma klasifikasi dengan metode berbasis aturan (*wildcard*) dan metode berbasis pembelajaran mesin (*machine learning*).

Pada tahap *preprocessing*, yang dilakukan adalah penghapusan karakter tanda baca miring (/) yang kemudian di lanjutkan dengan pemisahan karakter sehingga menghasilkan list dari sebuah string. Selanjutnya digunakan fungsi *explode* untuk melonggarkan list di dalam kolom [Kode Belakang] menjadi baris serta menghapus serta menghapus spasi pada kolom data tersebut. Data yang semula berisi Provinsi Banten dengan Kode Plat A dan pada Kota Serang mempunyai Kode Belakang

A*/B*/C*/D* setelah melalui proses *preprocessing* akan berubah menjadi tabel baru.

Provinsi	Kode Plat	Kota/ Kabupaten	Kode Belakang
Banten	A	Kota Serang	A*
Banten	A	Kota Serang	B*
Banten	A	Kota Serang	C*
Banten	A	Kota Serang	D*

Tabel 1 Data setelah Preprocessing

Karakter bintang * tetap dipertahankan dikarenakan hal tersebut berguna dalam pola wildcard untuk mencari suatu pola tertentu. Dimana karakter bintang * digunakan sebagai indikator posisi huruf tersebut. Untuk mengaktifkan pola wildcard dibutuhkan REGEX. Dalam pemrosesan ini, karakter * diubah menjadi karakter \$ sebagai simbol dari posisi.

Provinsi	Kode Plat	Kota/Kabupaten	Kode Belakang	Regex Belakang
Banten	A	Kota Serang	A*	^A.*\$
Banten	A	Kota Serang	B*	^B.*\$
Banten	A	Kota Serang	C*	^C.*\$
Banten	A	Kota Serang	D*	^D.*\$
Banten	A	Kabupaten Serang	E*	^E.*\$

Gambar 2 REGEX

Untuk membuat sebuah klasifikasi dibutuhkan sebuah labeling, Di mana LabelEncoder ini berperan penting dalam mengubah nilai kategorikal menjadi angka *numeric* agar bisa diproses oleh algoritma *machine learning*. Dilanjutkan dengan klasifikasi menggunakan *DecisionTree*.

DecisionTree bekerja dengan cara pemikiran manusia, dimana ketika ada sebuah *input* akan dibaca dan dibandingkan cocok kedalam kategori kelas mana.

Hasil implementasi ini menunjukkan bahwa kedua algoritma yang digunakan, yaitu *wildcard* dan *decision tree*, yang baik dalam mengklasifikasi asal wilayah kendaraan berdasarkan kode plat nomor. Penggunaan *wildcard* dan *decision trees* terbukti sangat efisien dalam klasifikasi text.

Input Plat	Provinsi	Kota/Kabupaten	Pola Deteksi
H 3293 MG	Jawa Tengah	Semarang	*G
G 7860 MG	Jawa Tengah	Kabupaten Brebes	*G
B 2792 ST	DKI Jakarta	Jakarta Selatan	S*
F 4039 SJ	Jawa Barat	Sukabumi	S*

Tabel 2 Output Program

Berdasarkan hasil eksekusi dari penelitian ini, menunjukkan bahwa dapat melakukan pemodelan pembacaan plat nomor kendaraan secara presisi dengan pola belakang dari plat nomor kendaraan tersebut. Hal ini menandakan bahwa sistem mampu memberikan hasil yang akurat baik melalui pencocokan langsung maupun prediksi. Hal ini membuktikan bahwa sistem kerja dari pola *wildcard* dan *Decision Tree* bekerja dengan baik.

Proses evaluasi dilakukan menggunakan skema *train-test split*, di mana untuk data pelatihan dan data pengujian dengan rasio sebesar 80:20. Dengan menggunakan rasio ini untuk menguji kemampuan generalisasi model terhadap data yang tidak pernah dilihat sebelumnya. Dari hasil evaluasi penelitian yang telah dilakukan memiliki akurasi sebesar 0,58 atau sebesar 58%, yang menunjukkan bahwa struktur pohon keputusan berhasil menangkap pola relasi antara input plat kendaraan terhadap target keluaran provinsi dan kota maupun kabupaten.

Sementara itu, *classification report* memberikan hasil berupa matrix tambahan yang berisikan *precision*, *recall*, dan *F1-score* untuk masing-masing kelas kota/kabupaten dari dataset. Nilai *precision* merupakan hasil sejauh mana model yang digunakan dapat menghindari kesalahan dalam prediksi positif palsu, untuk *recall* merupakan kemampuan model dalam mengeksekusi semua contoh yang releban dari suatu kota/kabupaten, dan untuk *F1-score* merupakan rata-rata harmonik dari *precision* dan *recall* yang memberikan keseimbangan dalam melakukan proses pemodelan dalam hal akurasi dan sensitivitas.

Classification Report:

	precision	recall	f1-score
Kabupaten Bandung	0.00	0.00	0.00
Kabupaten Bandung Barat	0.00	0.00	0.00
Kabupaten Bangkalan	1.00	1.00	1.00
Kabupaten Banjarnegara	0.00	0.00	0.00
Kabupaten Bantul	0.00	0.00	0.00
Kabupaten Banyumas	0.00	0.00	0.00
Kabupaten Banyuwangi	0.00	0.00	0.00
Kabupaten Batang	0.00	0.00	0.00
Kabupaten Bekasi	0.00	0.00	0.00
Kabupaten Blitar	0.67	0.67	0.67
Kabupaten Blora	0.00	0.00	0.00
Kabupaten Bogor	1.00	1.00	1.00
Kabupaten Bojonegoro	1.00	1.00	1.00
Kabupaten Bondowoso	0.00	0.00	0.00
Kabupaten Boyolali	0.00	0.00	0.00
Kabupaten Brebes	0.00	0.00	0.00
Kabupaten Ciamis	0.00	0.00	0.00
Kabupaten Cianjur	0.00	0.00	0.00
Kabupaten Cilacap	0.00	0.00	0.00
Kabupaten Cirebon	1.00	0.83	0.91
Kabupaten Demak	0.00	0.00	0.00
Kabupaten Garut	1.00	0.50	0.67
Kabupaten Gresik	0.00	0.00	0.00
Kabupaten Grobogan	0.00	0.00	0.00
Kabupaten GunungKidul	0.00	0.00	0.00
Kabupaten Indramayu	0.75	0.75	0.75

Gambar 3 Klasifikasi Report Tiap Kota/Kabupaten

Pada hasil tersebut, pada Kabupaten Bandung memiliki *F1-score* lebih sedikit dari pada Kabupaten Blitar. Di mana *F1-score* dari Kabupaten Bandung sebesar 0,00. Sedangkan untuk *F1-score* dari Kabupaten Blitar sebesar 0,67. Hal ini dapat terjadi dikarenakan data pemodelan Kabupaten Blitar lebih banyak dari Kabupaten Bandung. Dari hasil tersebut, dapat disimpulkan bahwa semakin banyak data yang dibangun mampu mengidentifikasi asal kota/kabupaten kendaraan memiliki tingkat akurasi yang lebih tinggi dibandingkan kota/kabupaten yang memiliki data lebih sedikit.

Pembahasan ini mengindikasikan bahwa penerapan metode klasifikasi berbasis *decision tree* terhadap data identitas seperti plat nomor kendaraan memiliki potensi yang luas, khususnya dalam konteks pengawasan lalu lintas, pemetaan distribusi kendaraan, dan perumusan kebijakan transportasi berbasis data. Sistem klasifikasi ini tidak hanya mampu memproses informasi dalam jumlah besar secara efisien, tetapi juga dapat menjadi komponen penting dalam sistem transportasi cerdas (*intelligent transportation systems*).

5. KESIMPULAN

- Penggunaan *DecisionTree* untuk data jenis ini membantu dalam pemilihan kelas wilayah oleh kode plat nomor kendaraan. Di mana algoritma ini harus diselaraskan dengan pola *wildcard*. Hal ini dikarenakan menggunakan

- algoritma DecisionTree saja sistem belum mampu mengidentifikasi kode plat sesuai wilayah dengan baik dan benar, sehingga penggabungan metode ini efektif dalam mengklasifikan jenis plat nomor kendaraan pulau jawa dengan baik.
- b. Akurasi dalam penelitian ini terbilang cukup kecil yaitu 0.58. Namun, nilai ini dapat berkembang seiring dengan penggunaan metode tambahan yang mampu mengidentifikasi hal tersebut dengan baik. Penelitian selanjutnya dapat dilakukan penggabungan metode ini dengan citra camera atau citra gambar, sehingga keakuratan dalam pembacaan dapat berjalan dengan lebih baik
 - c. Tahapan preprocessing dan REGEX dengan pola wildcard memberikan kontribusi penuh dalam pengidentifikasian pola kodeplat yang unik. Preprocessing memberikan data bersih ke sistem sehingga sistem mampu mengidentifikasi kodeplat dengan lebih mudah dan baik.
 - d. Nilai score F1 sangat dipengaruhi oleh besarnya pemodelan tersebut, sedangkan semakin kecil pemodelan tersebut semakin kecil pula score F1.
 - e. Sistem ini memiliki potensi untuk dapat diterapkan dalam pengelolaan data kendaraan dan mendukung pengembangan sistem transportasi cerdas di wilayah padat penduduk.
- [2] D. Rudolf, C. Kuntadi, and R. L. Karunia, "FAKTOR-FAKTOR YANG MEMPENGARUHI BELANJA INFRASTRUKTUR JALAN DI PULAU JAWA: JUMLAH PANJANG JALAN, JUMLAH KENDARAAN BERMOTOR, DAN KONDISI KEMANTAPAN JALAN," *Jurnal Manajemen, Akuntansi dan Logistik (JUMATI)*, vol. 2, no. 3, 2024.
 - [3] S. Christanti, A. P. Sari, and M. M. Falihuddin, "IMPLEMENTASI SISTEM PENGENALAN PLAT NOMOR KENDARAAN LOKAL MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)," *Jurnal Teknik Informatika*, vol. 16, no. 3, pp. 120–126, 2024.
 - [4] S. R. Ningsih, "Pengaruh Teknologi Terhadap Produktivitas Tenaga Kerja di Indonesia," *Benefit: Journal of Bussiness, Economics, and Finance*, vol. 2, no. 1, pp. 1–9, Feb. 2024, doi: 10.37985/benefit.v2i1.341.
 - [5] S. F. Tahir and C. A. Sugianto, "OPTIMASI NAIVE BAYES MENGGUNAKAN ALGORITMA GENETIKA PADA KLASIFIKASI KOMENTAR CYBERBULLYING PADA MEDIA SOSIAL X," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4834.
 - [6] Abd. A. Syam, G. Hardy M, A. Salim, D. F. Surianto, and M. Fajar B, "ANALISIS TEKNIK PREPROCESSING PADA SENTIMEN MASYARAKAT TERKAIT KONFLIK ISRAEL-PALESTINA MENGGUNAKAN SUPPORT VECTOR MACHINE," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 3, pp. 1464–1472, Aug. 2024, doi: 10.29100/jipi.v9i3.5527.
 - [7] K. F. H. Holle, R. Alfianita, and H. M. Putri, "Evaluasi teknik preprocessing terhadap kinerja multinomial Naïve Bayes dalam klasifikasi pertanyaan insincere," *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, vol. 12, no. 4, pp. 707–715, 2024.
 - [8] D. E. Waluyo, H. W. Kinasih, C. Paramita, D. Pergiwati, R. Nohan, and F. A. Rafrastara, "Komparasi dan Implementasi Algoritma Regresi Machine Learning untuk Prediksi Indeks Harga Saham Gabungan," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 9, no. 1, pp. 12–17, Feb. 2024, doi: 10.30591/jpit.v9i1.6105.
 - [9] F. Raihanunnisa, M. Arhami, and R. Hidayat, "Pendekatan Hybrid Pada Sistem Peringkat Teks Artikel Berita Bahasa Inggris Menggunakan Natural Language Processing,"

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih kepada Bapak Nur Ikhsan, S.T., M.Kom. yang sudah membimbing kami sehingga dapat melakukan penelitian ini dengan lebih baik.

DAFTAR PUSTAKA

- [1] S. Marti'ah, "ANALISIS FAKTOR FAKTOR YANG MEMPENGARUHI KETIMPANGAN PENDAPATAN DI PULAU JAWA 2009-2021," *Juremi: Jurnal Riset Ekonomi*, vol. 4, no. 6, pp. 1275–1282, 2025.

- TELEMATIKA MKOM*, vol. 15, no. 2, pp. 86–92, 2023.
- [10] A. Septiana, G. Dwilestari, and A. Bahtiar, “PERBANDINGAN METODE KLASIFIKASI DENGAN MENERAPKAN ADABOOST DALAM ANALISIS SENTIMEN PENGGUNA TWITTER X TERHADAP PENERAPAN KURIKULUM MERDEKA,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 323–330, Feb. 2024, doi: 10.36040/jati.v8i1.8453.
- [11] S. Fernando, A. Voutama, and A. Andri Hendriadi, “KLASIFIKASI BERITA HOAKS KAMPANYE PEMILIHAN UMUM (PEMILU) 2024 MENGGUNAKAN ALGORITMA NAÏVE BAYES,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 2112–2115, Apr. 2024, doi: 10.36040/jati.v8i2.9400.
- [12] F. Adline, T. Tobing, A. Chandra, and R. Nainggolan, “PENERAPAN ALGORITMA KNUTH MORRIS PRATT (KMP) PADA PENCARIAN DATA DI SQL LIKE OPERATORS,” *Jurnal Widya*, vol. 3, no. 1, pp. 50–58, 2022, doi: 10.54593/awl.v3i1.82.
- [13] C. E. Suharyanto, P. Simanjuntak, and S. Adam, “Optimalisasi Sistem Keamanan Jaringan dan Manajemen Bandwidth pada Jaringan (Studi Kasus: CU Tunas Harapan),” *Jurnal Desain Dan Analisis Teknologi*, vol. 1, no. 1, pp. 13–18, 2022.
- [14] S. Wijanarko and A. Santoso, “Penerapan Fungsi Mid Dan Find pada Pembersihan Data Alamat,” *Jurnal Teknik Informatika*, vol. 10, no. 1, pp. 14–18, 2024.
- [15] D. A. Mukhsinin, M. Rafliansyah, S. A. Ibrahim, R. Rahmaddeni, and D. Wulandari, “Implementasi Algoritma Decision Tree untuk Rekomendasi Film dan Klasifikasi Rating pada Platform Netflix,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 570–579, Mar. 2024, doi: 10.57152/malcom.v4i2.1255.
- [16] N. Azwanti and N. E. Putra, “Analisis Kepuasan Customer pada Sdtechnology Computer dengan Algoritma Decision Tree,” *Jurnal Desain Dan Analisis Teknologi*, vol. 3, no. 2, pp. 137–148, Jul. 2024, doi: 10.58520/jddat.v3i2.62.
- [17] A. I. Putri, Y. Syarif, P. Jayadi, F. Arrazak, and F. N. Salisah, “Implementasi Algoritma Decision Tree dan Support Vector Machine (SVM) untuk Prediksi Risiko Stunting pada Keluarga,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 3, no. 2, pp. 349–357, Feb. 2024, doi: 10.57152/malcom.v3i2.1228.
- [18] Ubeitul Maltuf and Zaehol Fatah, “PENERAPAN ALGORITMA DECISION TREE UNTUK KLASIFIKASI KONSUMSI ENERGI LISTRIK RUMAH TANGGA DENGAN PENGGUNAAN RAPIDMINER,” *Jurnal Ilmiah Multidisiplin Ilmu*, vol. 2, no. 1, pp. 38–45, Feb. 2025, doi: 10.69714/0hmk8712.
- [19] A. P. Permana, K. Ainiyah, and K. F. H. Holle, “Analisis Perbandingan Algoritma Decision Tree, kNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 6, no. 3, pp. 178–188, 2021.
- [20] M. F. Hidayattullah, T. Abidin, I. N. Nururritzqi, and I. Khoeriyah, “Sistem Pendukung Keputusan Latihan Memanah Menggunakan Algoritma Decision Tree,” *Komputika : Jurnal Sistem Komputer*, vol. 10, no. 2, pp. 93–99, Aug. 2021, doi: 10.34010/komputika.v10i2.3774.