

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemacetan di Jakarta terjadi akibat laju pertumbuhan kendaraan yang jauh lebih cepat dibandingkan pembangunan infrastruktur jalan yang tidak seimbang. Hal ini terjadi karena dalam proses pembangunan atau pelebaran jalan, Sebagian kapasitas jalan justru berkurang, sementara jumlah kendaraan terus meningkat. Selain itu, kurangnya informasi mengenai keadaan lalu lintas yang mampu memberikan rute alternatif juga memperparah situasi.

Masalah kemacetan tidak hanya menyebabkan pemborosan waktu, tetapi juga berdampak pada tingginya konsumsi bahan bakar, meningkatnya polusi udara, hingga turunnya produktivitas masyarakat. Oleh sebab itu, diperlukan analisis dan prediksi keadaan lalu lintas secara akurat yang diperlukan sebagai bahan pertimbangan bagi pemerintah atau instansi terkait dalam menyusun kebijakan pengelolaan lalu lintas.

Dengan tersedianya data volume kendaraan harian yang dicatat secara berkala, penelitian ini berfokus pada pemanfaatan data historis untuk membangun model prediksi keadaan lalu lintas. Model ini diharapkan dapat mengklasifikasikan keadaan lalu lintas menjadi tiga kategori, yaitu: lancar, sedang, dan padat. Kekurangan informasi lalu lintas yang lengkap sering membuat pengguna jalan terjebak kemacetan yang kemudian memunculkan berbagai permasalahan lanjutan.

Penelitian ini memanfaatkan data historis lalu lintas yang berasal dari “*Traffic Prediction Dataset*” yang diperoleh melalui situs *Electronic Registration Identification* (ERI) Korlantas Polri. Dataset tersebut mencakup fitur-fitur input seperti jenis kendaraan (mobil penumpang, bus, mobil barang, sepeda motor), waktu, lokasi, dan cuaca yang mewakili kondisi lalu lintas di wilayah Jakarta Selatan.

Untuk mencapai tujuan tersebut, algoritma *K-Nearest Neighbors* (K-NN) digunakan karena memiliki karakteristik yang sederhana namun efektif untuk masalah klasifikasi. K-NN memanfaatkan jarak terdekat antar data untuk memetakan pola lalu lintas berdasarkan parameter seperti hari, jam, volume kendaraan, dan kondisi cuaca.

Hasil akhir penelitian ini berupa model prediksi lalu lintas berdasarkan data historis, yang mampu mengelompokkan keadaan lalu lintas menjadi kategori seperti lancar, sedang, macet. Meskipun tidak langsung menghasilkan aplikasi, model ini diharapkan menjadi landasan untuk pengembangan sistem informasi atau aplikasi pintar berbasis prediksi lalu lintas di masa mendatang. Selain itu, hasil penelitian ini diharapkan menjadi referensi bagi instansi terkait dalam merumuskan kebijakan serta strategi pengelolaan lalu lintas di Jalan Ciledug secara lebih efektif. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi di bidang akademik, tetapi juga solusi nyata untuk mengurangi masalah transportasi di kota besar.

Oleh karena itu, penulis berencana melakukan penelitian untuk memprediksi arus lalu lintas yang akan dijelaskan lebih lanjut dalam skripsi yang berjudul **“PREDIKSI KEADAAN LALU LINTAS DI JAKARTA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR”**.

1.2 Rumusan Masalah

Bagaimana membangun model yang dapat memprediksi keadaan lalu lintas (lancar, sedang, padat) di Jalan Ciledug, Jakarta Selatan menggunakan algoritma K-Nearest Neighbors (K-NN) berbasis data historis, seperti volume kendaraan, hari, jam, dan cuaca?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan fokus, maka terdapat beberapa batasan masalah yang ditetapkan, yaitu:

1. Penelitian ini hanya berfokus pada prediksi lalu lintas berdasarkan data historis dan variabel tertentu seperti hari, jam, volume kendaraan, dan cuaca
2. Model yang dikembangkan hanya diuji pada data lalu lintas di wilayah tertentu dan tidak mencakup semua daerah.
3. Teknik prediksi yang digunakan dalam penelitian ini terbatas pada metode kecerdasan buatan seperti machine learning.
4. Penelitian ini tidak membahas aspek kebijakan transportasi yang berkaitan dengan pengelolaan lalu lintas secara langsung.

5. Pembatasan data selama tiga bulan dilakukan karena pertimbangan keterbatasan waktu penelitian dan proses pemrosesan data. Meski periode tersebut relatif singkat, data yang diperoleh tetap mencakup berbagai variasi pola lalu lintas, seperti hari kerja, akhir pekan, serta kondisi cuaca yang berbeda, sehingga tetap relevan untuk membangun dan menguji model prediktif secara representatif.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Membangun model prediksi lalu lintas yang akurat menggunakan metode kecerdasan buatan.
2. Menganalisis faktor-faktor yang memengaruhi prediksi lalu lintas.
3. Mengevaluasi efektivitas model prediksi dalam membantu pengelolaan lalu lintas.

1.4.2 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

1. Bagi Penulis:

Penulis memperoleh pemahaman dan wawasan baru terkait permasalahan kemacetan lalu lintas, khususnya yang terjadi di ruas Jalan Ciledug Raya, Kelurahan Cipulir, Jakarta Selatan.

2. Bagi Masyarakat:

Memberikan informasi yang lebih jelas dan akurat kepada masyarakat mengenai kondisi kemacetan yang sering terjadi di wilayah Jalan Ciledug Raya, sehingga dapat meningkatkan kesadaran dan pemahaman dalam menghadapi masalah lalu lintas.

3. Bagi Pemerintah:

Hasil penelitian ini dapat dijadikan sebagai masukan dalam menyusun strategi perbaikan infrastruktur transportasi serta penataan ruang kota. Dengan begitu, diharapkan dapat meningkatkan efisiensi penggunaan jalan dan menekan tingkat kemacetan di kawasan perkotaan.

1.4.3 Kerangka Penelitian

BAB I: PENDAHULUAN

1. Latar Belakang

- a. Penjelasan mengenai kemacetan lalu lintas di Jakarta, terutama di Jalan Ciledug Raya, dan dampaknya terhadap waktu perjalanan, konsumsi bahan bakar, serta polusi udara.
- b. Tujuan penelitian untuk membangun model prediksi keadaan lalu lintas berbasis data historis dengan menggunakan algoritma K-Nearest Neighbors (K-NN).
- c. Pentingnya prediksi keadaan lalu lintas untuk pengelolaan arus lalu lintas yang lebih efisien.

2. Rumusan Masalah

Bagaimana membangun model yang dapat memprediksi keadaan lalu lintas (lancar, sedang, padat) di Jalan Ciledug, Jakarta Selatan menggunakan algoritma K-Nearest Neighbors (K-NN)?

3. Batasan Masalah

- a. Penelitian ini hanya akan menggunakan data historis lalu lintas yang diambil dari Traffic Prediction Dataset ERI Korlantas Polri dengan variabel seperti hari, jam, volume kendaraan, dan cuaca.
- b. Penelitian terbatas pada penggunaan algoritma K-Nearest Neighbors (K-NN) dalam mengklasifikasikan keadaan lalu lintas.

4. Rumusan Masalah

Bagaimana membangun model yang dapat memprediksi keadaan lalu lintas (lancar, sedang, padat) di Jalan Ciledug, Jakarta Selatan menggunakan algoritma K-Nearest Neighbors (K-NN)?

5. Batasan Masalah

- a. Penelitian ini hanya akan menggunakan data historis lalu lintas yang diambil dari Traffic Prediction Dataset ERI Korlantas Polri dengan variabel seperti hari, jam, volume kendaraan, dan cuaca.
- b. Penelitian terbatas pada penggunaan algoritma K-Nearest Neighbors (K-NN) dalam mengklasifikasikan keadaan lalu lintas.

6. Tujuan Penelitian

- a. Membangun model prediksi keadaan lalu lintas yang akurat menggunakan algoritma K-NN.
- b. Menganalisis faktor-faktor yang mempengaruhi keadaan lalu lintas seperti volume kendaraan, waktu, dan cuaca.
- c. Mengevaluasi efektivitas model prediksi dalam mendukung pengelolaan lalu lintas secara lebih efisien.

7. Manfaat Penelitian

- a. Bagi Penulis: Menambah wawasan dalam mengembangkan model prediksi menggunakan algoritma K-NN dalam konteks lalu lintas.
- b. Bagi Masyarakat: Memberikan informasi yang lebih akurat mengenai kondisi lalu lintas yang dapat membantu pengguna jalan.
- c. Bagi Pemerintah: Menyediakan model yang dapat digunakan untuk merumuskan kebijakan pengelolaan lalu lintas yang lebih efektif di Jakarta.

BAB II: TINJAUAN PUSTAKA

1. Teori Dasar

- a. Prediksi Lalu Lintas: Penjelasan mengenai definisi prediksi lalu lintas dan teknik-teknik yang digunakan untuk prediksi lalu lintas.
- b. Algoritma K-Nearest Neighbors (K-NN): Prinsip dasar K-NN, cara kerjanya, dan penerapannya dalam prediksi lalu lintas.
- c. Cross-Validation: Penjelasan tentang teknik k-fold cross-validation untuk mengevaluasi kinerja model.

2. Penelitian Terkait

- a. Tinjauan tentang penelitian sebelumnya yang menggunakan K-NN untuk prediksi lalu lintas dan evaluasi menggunakan cross-validation.
- b. Keunggulan dan perbedaan penelitian ini dengan penelitian sebelumnya.

BAB III: METODOLOGI PENELITIAN

1. Sumber Data

- a. Traffic Prediction Dataset yang diperoleh dari ERI Korlantas Polri mencakup variabel seperti hari, jam, volume kendaraan, dan cuaca.
- b. Data historis yang mencakup 3 bulan pengamatan untuk wilayah Jakarta Selatan.

2. Metode Penelitian

- a. Klasifikasi Keadaan Lalu Lintas: Menggunakan total kendaraan untuk mengklasifikasikan keadaan lalu lintas ke dalam tiga kategori: lancar, sedang, dan padat.
- b. Algoritma K-Nearest Neighbors (K-NN): Penggunaan K-NN untuk mengklasifikasikan keadaan lalu lintas berdasarkan data historis.

3. Tahapan Penelitian

- a. Pembersihan Data (Data Cleaning): Mengatasi missing values, menghapus duplikat, dan memastikan format data konsisten.
- b. Rekayasa Fitur (Feature Engineering): Mengubah data waktu ke format desimal, normalisasi data, dan encoding variabel kategorik.
- c. Pembagian Data: Membagi dataset menjadi 80% data latih dan 20% data uji untuk pengujian model.

4. Evaluasi Model

- a.k-Fold Cross-Validation: Menggunakan teknik 5-fold cross-validation untuk mengevaluasi kinerja model K-NN.
- b.Metrik Kinerja: Pengukuran akurasi, presisi, recall, dan F1-Score pada setiap fold.
- c.Pemilihan Nilai kkk Optimal: Menggunakan grid search untuk memilih nilai kkk terbaik dalam K-NN.

BAB IV: ANALISA HASIL DAN PEMBAHASAN

1. Gambaran Umum Data
 - a. Deskripsi tentang dataset, termasuk distribusi kondisi lalu lintas (lancar, sedang, padat).
 - b. Visualisasi data untuk menggambarkan pola lalu lintas.
2. Hasil Evaluasi Model
 - a. Hasil dari 5-fold cross-validation yang mencakup akurasi dan metrik lainnya seperti presisi, recall, dan F1-Score.
 - b. Pembahasan mengenai kinerja model pada setiap fold dan rata-rata hasilnya.
3. Pembahasan
 - a. Analisis mengapa model K-NN dapat memberikan hasil yang baik pada dataset ini dan alasan pemilihan nilai kkk optimal.
 - b. Diskusi mengenai potensi perbaikan dan aplikasi model untuk pengelolaan lalu lintas di Jakarta.

BAB V: PENUTUP

1. Kesimpulan

Merangkum hasil penelitian, yaitu keberhasilan model K-NN dalam memprediksi keadaan lalu lintas dan efektivitas k-fold cross-validation dalam mengevaluasi model.

2. Saran

Saran untuk penelitian lanjutan, misalnya memperbaiki data yang lebih besar, memperkenalkan fitur tambahan, atau mencoba teknik prediksi lainnya.

1.5 Alasan Pemilihan Topik dan Urgensi

Pemilihan topik prediksi keadaan lalu lintas di Jakarta menggunakan algoritma K-Nearest Neighbors (K-NN) didasarkan pada tiga pertimbangan utama.

1. Pertama, urgensi problem kemacetan di Jakarta yang berdampak pada pemborosan waktu, kenaikan konsumsi BBM, polusi udara, dan penurunan produktivitas—sehingga diperlukan dukungan informasi prediktif untuk pengelolaan lalu lintas yang lebih efektif.
2. Kedua, ketersediaan dataset historis ERI Korlantas Polri yang memuat variabel hari, jam, volume kendaraan, dan cuaca, memungkinkan pengembangan model klasifikasi kondisi lancar/sedang/padat yang berbasis data.
3. Ketiga, dari sisi metodologis, K-NN dipilih karena karakteristiknya yang sederhana namun efektif untuk pola nonlinier dan mudah ditafsirkan melalui kedekatan antar data, sehingga sesuai untuk prototipe awal sistem prediksi lalulintas.

Secara praktis, model yang dihasilkan diharapkan menjadi landasan pengembangan sistem informasi berbasis kecerdasan buatan serta referensi bagi instansi terkait dalam perumusan strategi pengelolaan lalu lintas pada koridor studi. Selain itu, penelitian ini menutup sebagian celah yang teridentifikasi pada keterbatasan ruang-waktu dan evaluasi model, sekaligus menyiapkan dasar untuk validasi eksternal dan perluasan fitur pada studi lanjutan (lihat Batasan dan Saran).

