

Bab I

Pendahuluan

1.1 Latar Belakang

Pada lalu lintas kendaraan di jalanan, kemacetan merupakan salah satu masalah yang paling krusial. Kemacetan lalu lintas adalah kondisi dimana volume arus lalu lintas pada ruas jalan tertentu melebihi dari rencana kapasitas jalan tersebut sebenarnya. Kondisi tersebut menyebabkan menurunnya kecepatan arus bebas hingga mendekati 0 km/jam [5]. Masalah ini bukanlah hal yang asing bagi lalu lintas kendaraan pada umumnya, terutama pada kota Bandung. Kota Bandung merupakan salah satu dari beberapa kota dengan luas terbesar serta pertumbuhan populasi yang tinggi di negara Indonesia. [8]. Jumlah penduduk di Kota Bandung mencapai angka 2.510.103 di tahun 2020 saja, bertambah 6395 orang dari tahun 2018 yang semula berjumlah 2.503.708 [3]. Dengan jumlah penduduk yang bertingkat setiap tahunnya, penggunaan jalanan pun akan meningkat. Baik itu dengan tujuan transportasi, keagamaan, olahraga, dan sebagainya. Bahkan, jumlah kendaraan bermotor pribadi pada tahun 2020 mencapai angka 1.538.788 [4]. Hal ini dapat menyebabkan meningkatnya resiko volume arus lalu lintas kendaraan hingga melebihi kapasitas.

Di era digital ini, terdapat berbagai metode dalam mengumpulkan informasi untuk menyelesaikan sebuah masalah. Metode konvensional seperti survei atau wawancara akan sangat sulit dilakukan untuk menemukan solusi masalah dengan volume besar seperti kemacetan lalu lintas kendaraan. Perlu penggunaan metode konvensional yang diiringi metode digital untuk mendapatkan hasil yang cocok. Metode digital seperti pembelajaran mesin adalah salah satu metode paling cocok untuk menemukan hasil dalam masalah ini. Pembelajaran mesin adalah studi akan algoritma dan model statistik yang digunakan komputer untuk melakukan suatu pekerjaan tanpa diminta secara eksplisit [10]. Dengan menggunakan algoritma pembelajaran mesin, dapat dilakukan pengelompokan data yang sejenis, ataupun prediksi dari kumpulan informasi yang ada. Dalam pembelajaran mesin, perlu dilakukan suatu pengelompokan sekumpulan data dengan label yang dapat dikenali

secara unik. Proses pengelompokan ini dinamakan Klasifikasi, dan program menggunakan metode pembelajaran mesin perlu dilatih mengenai pengelompokan ini. Dengan mengelompokkan data yang sudah ada, program akan mengenali pola dan karakteristik data baru yang dimasukkan sehingga dapat mengelompokkan kembali data terbaru. Untuk menemukan solusi terbaik untuk permasalahan kemacetan di Kota Bandung, perlu dilakukan prediksi dan klasifikasi menggunakan metode pembelajaran mesin.

Beragam penelitian telah dilakukan demi menemukan metode terbaik untuk menemukan solusi kemacetan di Kota Bandung. Pada studi prediksi kemacetan lalu lintas kendaraan yang dilakukan oleh Fazry, ditemukan bahwa akurasi dari classifier SVM mencapai 93%, dan akurasi dari Naive bayes mencapai 90% dengan dataset daerah Bandung yang disediakan oleh ATCS Bandung [14]. Penelitian ini menggunakan atribut Hari, Waktu, dan okupansi lalu melabelinya berdasarkan nilai kategorikal menjadi numerikal. Atribut Hari yang memiliki nilai kategorikal seperti Senin, Selasa, Rabu, Kamis, Jumat, Sabtu, dan Minggu diubah menjadi nilai numerik seperti 0, 1, 2, 3, 4, 5, dan 6. Sementara atribut Waktu yang awalnya memiliki nilai kategorikal Pagi, Siang, dan Sore, diubah menjadi 0, 1, dan 2. Atribut okupansi dibagi menjadi empat level dan dilabeli sesuai dari nilainya masing masing. 60% kebawah dilabeli Free Flow atau 0. Diantara 60% dan 71% dilabeli Steady Flow atau 1. Diantara 70% dan 80% dilabeli Stable Controlled atau 2. Lalu, diantara 80% dan 90% dilabeli Unstable atau 3 [14].

Sementara pada studi yang dilakukan oleh Arabiat , ditemukan hasil yang cukup berbeda. Penelitian ini menggunakan metode Random Forest, Logistic Regression, Support Vector Machine, dan K-Nearest Neighbour [1]. Selain itu, penelitian ini menggunakan dua program pembelajaran mesin tambahan, yaitu WEKA yang dikembangkan oleh Waikato University, dan Orange tool yang dikembangkan oleh laboratorium bioinformatika di Universitas Ljubljana. Menggunakan Orange Tool, ditemukan bahwa akurasi dari classifier Random Forest dan Logistic Regression mencapai 100%, sementara KNN dan SVM masing masing mendapatkan akurasi 99.8% dan 99.1%. Akan tetapi, saat menggunakan WEKA, didapatkan bahwa SVM mencapai akurasi 99.7%. Sementara K-Nearest Neighbours, Logistic Regression, dan Random Forest mendapatkan masing masing 98.7%, 97.6%, dan 96.2%. Penelitian ini menggunakan dataset daerah Amman pada jalan King Abdullah. [1]

Di tahun 2022, Bokaba melakukan studi komparatif dari model ensemble untuk memprediksi kemacetan lalu lintas kendaraan [13]. Penelitian ini menggunakan dataset dari salah satu jalan raya di provinsi Gauteng, Afrika Selatan. Hasil

dari penelitian ini didapatkan dengan menggunakan metode pembelajaran mesin tradisional (seperti Random Forest, Decision Tree, dan Support Vector Machine), metode Bagging, AdaBoost, dan Logistic Regression sebagai metode stacking untuk mengkombinasi Random Forest, Decision Tree, dan Support Vector Machine. Variabel yang menggunakan waktu perjalanan, volume lalu lintas, dan rata rata kecepatan membantu prediksi agar berjalan lebih baik. Diantara semua metode, metode yang menggunakan Decision Tree menghasilkan akurasi paling tinggi. Dengan angka akurasi metode tradisional 98.3%, Bagging 98.2%, AdaBoost 99.7%, dan Logistic Regression pada angka 99.7% [13].

Berdasarkan penelitian penelitian diatas, metode metode pembelajaran mesin pada permasalahan ini sudah dieksplorasi dengan cukup luas. Akan tetapi, terdapat perbedaan hasil akurasi yang cukup nyata dari penerapan algoritma pembelajaran mesin pada permasalahan ini. Fazry mendapatkan akurasi 93% saat menggunakan metode Support Vector Machine tanpa melakukan penelitian menggunakan Random Forest [14]. Sementara Arabiat mendapat akurasi 99.7% menggunakan Support Vector Machine, dengan akurasi Random Forest yang lebih dari itu yaitu 100% [1]. Maka dari itu, penelitian ini akan memeriksa efektivitas dari algoritma pembelajaran mesin dengan menggunakan Random Forest dalam melakukan prediksi dan klasifikasi, untuk memberikan solusi yang lebih efektif dengan harapan dapat berguna untuk menentukan kebijakan terbaik bagi Bandung. Penelitian ini akan menggunakan atribut yang memengaruhi kemacetan lalu lintas kendaraan seperti lebar ruas jalan, volume lalu lintas, waktu, dan hari. Diharapkan bahwa peta dari interpolasi kriging akan membantu pemerintah dalam mengetahui titik titik krusial pada masalah kemacetan lalu lintas kendaraan.

1.2 Perumusan Masalah

Masalah utama yang akan diteliti pada penelitian ini adalah

- Bagaimana pengembangan metode algoritma Random Forest dan K-Nearest Neighbours dalam permasalahan prediksi kemacetan lalu lintas kendaraan di Bandung?
- Bagaimana bentuk model Random Forest dan K-Nearest Neighbours dalam melakukan prediksi dan klasifikasi kemacetan lalu lintas kendaraan di Bandung berdasarkan waktu?
- Bagaimana performansi metode algoritma Random Forest dan K-Nearest Neighbours dalam permasalahan prediksi kemacetan lalu lintas kendaraan di Bandung?

1.3 Tujuan

Berdasarkan penjelasan rumusan masalah, tujuan dari penelitian ini yaitu:

- Mengembangkan metode algoritma Random Forest dan K-Nearest Neighbours dalam permasalahan prediksi kemacetan lalu lintas kendaraan di Bandung.
- Membentuk model Random Forest dan K-Nearest Neighbours dalam melakukan prediksi dan klasifikasi kemacetan lalu lintas kendaraan di Bandung berdasarkan waktu.
- Menemukan performansi metode algoritma Random Forest dan K-Nearest Neighbours dalam permasalahan prediksi kemacetan lalu lintas kendaraan di Bandung.

1.4 Rencana Kegiatan

Rencana kegiatan untuk proposal tugas akhir dengan topik "Prediksi dan Klasifikasi Kemacetan Lalu Lintas Kendaraan di Kota Bandung Menggunakan Algoritma Random Forest dan K-Nearest Neighbours" adalah sebagai berikut :

1. Studi Literatur

Mencari literatur dan informasi terkait dengan topik yang dipilih, seperti metode interpolasi kriging, penerapan Random Forest dan KNN, serta penerapan klasifikasi dan prediksi. Penelitian ini akan menggunakan referensi berupa artikel, jurnal, serta situs yang akurat.

2. Pengumpulan Data

Pengumpulan data dilakukan dengan metode kuantitatif. Metode ini memerhatikan faktor faktor atau atribut terkait dengan kemacetan lalu lintas kendaraan yang sudah ditentukan sebelumnya.

3. Analisis dan Perancangan Sistem

Melakukan analisis dan merancang sistem yang akan digunakan. Berdasarkan data yang sudah dikumpulkan, dilakukan analisis untuk menentukan langkah yang tepat. Setelah dilakukan analisis, sistem akan dirancang sesuai dengan kesimpulannya.

4. Implementasi Sistem

Dari hasil analisa dan perancangan, sistem akan diimplementasikan dengan matang dan sesuai rencana.

5. Analisa Hasil Implementasi

Setelah sistem berhasil diimplementasikan, hasilnya akan dianalisa kembali untuk menemukan kesimpulan terbaik dari implementasi yang dilakukan.

6. Penulisan Laporan

Hasil dari implementasi akan dicatat dan direkam di laporan tahap ini. Seluruh tahapan dan luaran dari implementasi dan hasil sistem ini akan didokumentasikan untuk pembelajaran kedepannya.

1.5 Jadwal Kegiatan

Tabel 1.1: Jadwal kegiatan proposal tugas akhir

No	Kegiatan	Bulan ke-					
		1	2	3	4	5	6
1	Studi Literatur	■	■	■	■	■	■
2	Pengumpulan Data	■	■	■			
3	Analisis dan Perancangan Sistem		■	■	■		
4	Implementasi Sistem			■	■	■	
5	Analisa Hasil Implementasi				■	■	■
6	Penulisan Laporan		■	■	■	■	■