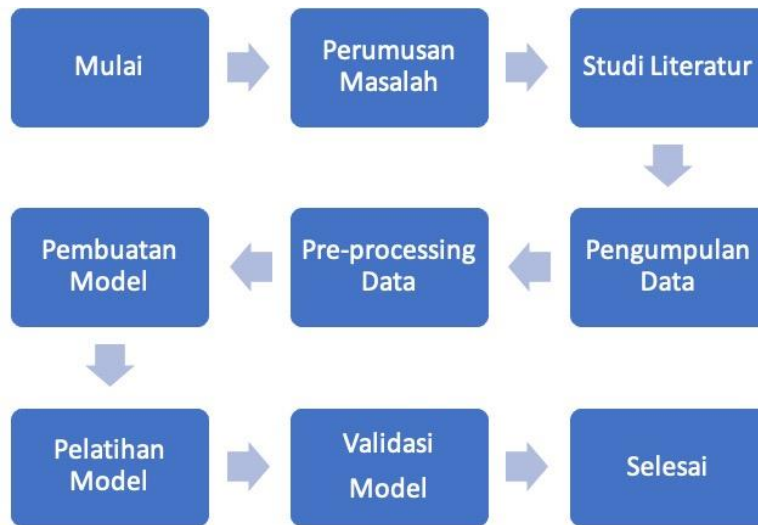


BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Tahapan penelitian untuk deteksi kerusakan jalan menggunakan metode Naïve Bayes dan Decision Tree ditunjukkan pada Gambar 3.1.



Gambar 3.1. Tahapan Penelitian

a) Perumusan Masalah

Pada tahapan ini dilakukan identifikasi masalah utama yang ingin diselesaikan, yaitu deteksi kerusakan jalan menggunakan metode Naïve Bayes dan Decision Tree.

b) Studi Literatur

Pada tahap ini dilakukan studi literatur terkait deteksi kerusakan jalan, studi tentang metode Naïve Bayes dan Decision Tree dalam bidang ini. Selain itu juga dilakukan analisis kelemahan dan kelebihan metode yang telah ada serta temuan terbaru dalam literatur.

c) Pengumpulan Data

Pada tahap ini dilakukan identifikasi jenis data yang diperlukan untuk pelatihan dan pengujian model Naïve Bayes dan Decision Tree, seperti data sensor, citra jalan, atau data historis kondisi jalan. Pada tahap ini dilakukan juga pengumpulan dan mempersiapkan kumpulan dataset yang relevan dengan tujuan penelitian.

d) Preprocessing Data

Pada tahap ini dilakukan pra-pemrosesan data untuk membersihkan, menormalkan, dan mempersiapkan data untuk analisis lebih lanjut. Pada tahap ini juga diterapkan teknik seperti penghilangan outlier, imputasi data yang hilang, dan pengkodean fitur jika diperlukan.

Proses pre-prosesing dalam konteks penerapan metode Naive Bayes dan Decision Tree untuk mendeteksi kerusakan jalan di Bandar Lampung akan melibatkan langkah-langkah berikut:

1. Pengumpulan Data: Data terkait kondisi jalan di Bandar Lampung harus dikumpulkan. Ini mungkin termasuk informasi tentang jenis kerusakan jalan, lokasi kerusakan, kondisi lalu lintas, jenis material jalan, dan faktor-faktor lain yang relevan.
2. Pembersihan Data: Data mentah yang dikumpulkan mungkin mengandung noise, missing values, atau outlier. Langkah pertama adalah membersihkan data dari masalah-masalah tersebut agar tidak memengaruhi hasil analisis. Ini bisa melibatkan penghapusan entri yang tidak lengkap, imputasi nilai yang hilang, atau deteksi dan penanganan outlier.
3. Transformasi Data: Beberapa fitur mungkin perlu diubah atau diubah bentuknya agar sesuai dengan kebutuhan algoritma. Misalnya, data spasial atau geografis mungkin perlu diubah menjadi format yang dapat diproses oleh algoritma.
4. Pemilihan Fitur: Pemilihan fitur melibatkan pemilihan subset dari semua fitur yang tersedia yang dianggap paling relevan untuk tujuan deteksi

kerusakan jalan. Fitur-fitur ini dapat dipilih berdasarkan pengetahuan domain, analisis korelasi, atau teknik pemilihan fitur lainnya.

5. Normalisasi: Normalisasi data mungkin diperlukan untuk memastikan bahwa semua fitur memiliki skala yang seragam. Ini bisa membantu mencegah fitur-fitur dengan rentang nilai yang besar mendominasi proses pembelajaran.
6. Pemrosesan Tambahan: Langkah-langkah pre-prosesing tambahan mungkin diperlukan tergantung pada karakteristik data dan kebutuhan spesifik dari algoritma yang akan digunakan.

Setelah proses pre-prosesing selesai, data siap untuk digunakan dalam penerapan metode Naive Bayes dan Decision Tree untuk mendeteksi kerusakan jalan di Bandar Lampung. Metode Naive Bayes dapat digunakan untuk klasifikasi kerusakan jalan berdasarkan fitur-fitur yang dipilih, sementara Decision Tree dapat memberikan pemahaman yang lebih baik tentang hubungan antara fitur-fitur tersebut dan kerusakan jalan yang terjadi. Proses ini akan membantu pihak terkait dalam pengambilan keputusan terkait perbaikan dan pemeliharaan jalan di Bandar Lampung.

e) Pembuatan Model

Pada tahap ini dilakukan implementasikan model Naïve Bayes dan Decision Tree untuk deteksi kerusakan jalan berdasarkan dataset yang telah dipersiapkan. Pada tahap ini juga dilakukan pengaturan parameter model yang paling sesuai dengan karakteristik data.

f) Pelatihan Model

Pada tahap ini dilakukan pembagian dataset menjadi subset pelatihan dan pengujian. Pada tahap ini juga dilakukan pelatihan model Naïve Bayes dan Decision Tree menggunakan subset pelatihan dan evaluasi kinerjanya menggunakan subset pengujian.

Saat pembuatan model untuk penerapan metode Naive Bayes dan Decision Tree dalam mendeteksi kerusakan jalan di Bandar Lampung, langkah-langkah yang mungkin dilakukan termasuk:

1. Pengumpulan Data: Tim peneliti akan mengumpulkan data terkait kondisi jalan di Bandar Lampung. Data ini mungkin mencakup informasi seperti jenis kerusakan jalan, lokasi kerusakan, intensitas lalu lintas, jenis permukaan jalan, dan faktor-faktor lain yang relevan.
2. Pre-Prosesing Data: Data yang dikumpulkan kemungkinan akan mengalami proses pre-prosesing, seperti membersihkan data yang tidak lengkap atau tidak konsisten, menghapus data yang tidak relevan, dan melakukan transformasi data jika diperlukan.
3. Pembuatan Model Naive Bayes: Langkah selanjutnya adalah mengembangkan model Naive Bayes menggunakan data yang telah diproses. Model Naive Bayes akan digunakan untuk mengklasifikasikan kondisi jalan berdasarkan atribut-atribut yang ada dalam dataset.
4. Pembuatan Model Decision Tree: Setelah model Naive Bayes dibuat, tim peneliti akan membuat model Decision Tree menggunakan data yang sama. Decision Tree akan menghasilkan aturan-aturan keputusan berdasarkan atribut-atribut dalam dataset yang dapat digunakan untuk mendeteksi kerusakan jalan.
5. Pelatihan Model: Model Naive Bayes dan Decision Tree akan dilatih menggunakan data yang telah diproses. Proses pelatihan ini akan membantu model untuk memahami pola-pola dalam data sehingga dapat melakukan klasifikasi dengan akurat.
6. Validasi Model: Setelah pelatihan selesai, kedua model akan divalidasi menggunakan set data uji yang tidak digunakan selama proses pelatihan. Validasi model akan membantu mengevaluasi kinerja model dalam mendeteksi kerusakan jalan dan mengukur tingkat akurasi serta keandalannya.

7. Optimasi dan Evaluasi: Setelah validasi, model-model tersebut mungkin akan dioptimasi dan dievaluasi ulang untuk memastikan bahwa mereka dapat memberikan prediksi yang akurat dan berguna dalam mendeteksi kerusakan jalan di Bandar Lampung.

Dengan langkah-langkah tersebut, diharapkan model-model Naive Bayes dan Decision Tree dapat memberikan kontribusi yang signifikan dalam upaya deteksi dan pemeliharaan jalan di Bandar Lampung.

g) Validasi Model

Pada tahap ini dilakukan validasi model menggunakan teknik seperti validasi silang atau validasi holdout. Sedangkan untuk evaluasi kinerja model diukur berdasarkan metrik yang relevan yaitu akurasi, presisi dan *recall*.

3.2 Pengumpulan Data

Dalam penyusunan penelitian ini, diperlukan data-data yang relatif lengkap sebagai bahan yang mendukung kebenaran materi pembahasan sehingga dilakukan pengumpulan data untuk mendapatkan informasi yang diperlukan. Adapun metode pengumpulan data dalam penelitian ini dilakukan sebagai berikut:

A. Studi Lapangan

Studi lapangan dilakukan menggunakan dua metode yaitu:

a. Wawancara

Wawancara dilakukan kepada pihak terkait dengan tema penelitian dalam hal ini dilakukan dengan memberikan pertanyaan-pertanyaan kepada pihak Dinas Perhubungan terkait di Kota Bandar Lampung.

b. Observasi

Observasi dilakukan untuk memperoleh data real, observasi dilakukan dengan cara langsung untuk memperoleh informasi dan data kerusakan jalan lalu lintas di Lampung.

B. Studi Pustaka

Metode ini digunakan untuk mencari sumber – sumber data yang diperlukan dalam penelitian yang biasa diperoleh dari membaca dan mengutip referensi serta jurnal penelitian yang pernah dilakukan sebelumnya yang mendukung dan berkaitan dengan penelitian. Adapun data yang di peroleh adalah data mengenai kerusakan jalan.

C. Lokasi dan Waktu Penelitian

1) Lokasi Penelitian

Penelitian ini dilakukan di Dinas Perhubungan Provinsi Lampung yang beralamat di Jl. Cut Mutia No.76, Pengajaran, Kec. Tlk. Betung Utara, Kota Bandar Lampung.

2) Waktu Penelitian

Rentang waktu penelitian yang dilakukan dimulai pada Bulan Oktober Tahun 2023 sampai dengan Bulan Januari Tahun 2024.

3.3 Pre-processing Data

a. Seleksi Atribut

Dalam tahap permulaan penelitian, seleksi atribut dilakukan untuk mereduksi jumlah atribut yang digunakan dalam penilaian kerusakan jalan. Pengurangan atribut dilakukan dengan mempertimbangkan nilai information gain dari setiap atribut terhadap entropi kelas targetnya. Semakin tinggi nilai information gain menunjukkan tingkat signifikansi atribut terhadap keputusan kelas targetnya. Proses penentuan atribut yang relevan dilakukan setelah melakukan wawancara dengan departemen pekerjaan terkait untuk menentukan atribut yang paling penting untuk evaluasi. Sebanyak 13 atribut telah ditetapkan sebagaimana ditunjukkan pada Tabel 3.1.

Tabel 3.1. Atribut Kerusakan Jalan

No	Nama Atribut
1.	Ukuran Kota
2	Tipe Lajur Jalan
3.	Lebar Jalan (M)
4.	Split Arah
5.	Hambatan Samping
6.	Lebar Bahu Jalan (m)
7.	Jarak Kerb (M)
8.	Faktor Koreksi Co
9.	Kapasitas
10.	Volume smp/jam(v)
11.	VCA
12.	Tingkat pelayanan jalan
13.	Tingkat Kerusakan

Pengurangan dataset dalam proses pengujian data dapat terjadi karena beberapa alasan:

1. Validasi Silang (Cross-Validation): Dalam beberapa kasus, dataset akan dibagi menjadi subset yang lebih kecil untuk validasi silang. Ini dilakukan untuk menghindari overfitting dan memastikan bahwa model dapat menggeneralisasi dengan baik pada data yang tidak terlihat.
2. Pembagian Data Training dan Testing: Dataset sering dibagi menjadi dua bagian: data training dan data testing. Bagian data testing direserve untuk menguji kinerja model setelah dilatih pada data training. Pengurangan dataset terjadi karena hanya sebagian kecil data digunakan untuk pengujian, sementara sebagian besar digunakan untuk melatih model.
3. Ketidakseimbangan Kelas: Dalam masalah klasifikasi, terkadang kelas-kelas dalam dataset tidak seimbang. Ini berarti jumlah sampel untuk setiap kelas tidak seimbang. Dalam hal ini, pengurangan dataset bisa terjadi karena hanya sebagian kecil dari kelas mayoritas yang digunakan untuk menjaga keseimbangan antara kelas mayoritas dan minoritas dalam data testing.

4. Efisiensi Komputasi: Terkadang, pengurangan dataset dilakukan untuk meningkatkan efisiensi komputasi. Menggunakan dataset yang lebih kecil dapat mempercepat waktu pelatihan dan pengujian model, terutama jika jumlah data sangat besar dan sumber daya komputasi terbatas.
5. Pemilihan Sampel yang Representatif: Dalam beberapa kasus, pengurangan dataset dilakukan untuk memilih sampel yang representatif dari populasi yang lebih besar. Ini bisa dilakukan dengan cara acak atau dengan menggunakan teknik pemilihan sampel yang cermat untuk memastikan bahwa sampel yang digunakan dalam pengujian mencerminkan karakteristik dari seluruh dataset.

Pengurangan dataset dapat bermanfaat untuk berbagai alasan praktis, termasuk meningkatkan efisiensi komputasi, menghindari overfitting, memastikan keseimbangan kelas, dan memilih sampel yang representatif. Namun, penting untuk memastikan bahwa proses pengurangan dataset dilakukan dengan hati-hati agar tidak mengorbankan representasi yang penting dari dataset asli.

b. Pengujian Algoritma Naïve Bayes

Analisis Pengujian Algoritma Naive Bayes untuk Kerusakan Jalan di Bandar Lampung:

1. Tujuan Pengujian: Tujuan dari pengujian algoritma Naive Bayes adalah untuk mengklasifikasikan kerusakan jalan di Bandar Lampung berdasarkan fitur-fitur yang diberikan dalam dataset.
2. Data Training: Data training terdiri dari berbagai fitur yang dapat mempengaruhi kerusakan jalan, seperti panjang jalan, lebar jalan, jenis lajur jalan, kondisi jalan, hambatan samping, dan faktor-faktor lainnya.
3. Pengolahan Data: Data yang dikumpulkan dari Bandar Lampung diproses dan disiapkan untuk digunakan sebagai data training algoritma Naive Bayes. Proses pengolahan data meliputi pemisahan data menjadi data training dan data tes, normalisasi data jika diperlukan, dan penanganan nilai yang hilang atau tidak valid.

4. Pelatihan Model Naive Bayes: Model Naive Bayes dilatih menggunakan data training yang telah disiapkan. Proses pelatihan melibatkan estimasi probabilitas dari setiap kelas berdasarkan fitur-fitur yang diamati dalam data training.
5. Validasi Model: Setelah pelatihan selesai, model Naive Bayes divalidasi menggunakan data tes yang terpisah. Performa model dievaluasi berdasarkan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score.
6. Analisis Hasil: Hasil pengujian memberikan gambaran tentang seberapa baik model Naive Bayes dapat mengklasifikasikan kerusakan jalan di Bandar Lampung. Performa model dievaluasi untuk memastikan keandalan prediksi dan kemampuan model dalam menggeneralisasi data yang belum pernah dilihat sebelumnya.
7. 7.Penggunaan Model: Setelah berhasil divalidasi, model Naive Bayes dapat digunakan untuk memprediksi kerusakan jalan di Bandar Lampung berdasarkan data baru yang diberikan. Prediksi yang dihasilkan dapat digunakan sebagai panduan untuk perencanaan dan pemeliharaan infrastruktur jalan di Bandar Lampung.
8. Peningkatan Model: Pengujian ini juga memberikan wawasan tentang area di mana model Naive Bayes mungkin perlu diperbaiki atau ditingkatkan. Berdasarkan analisis hasil, pengembang dapat mengidentifikasi fitur-fitur yang lebih penting atau memperbaiki masalah dalam pemrosesan data atau pelatihan model.

Pengujian algoritma Naive Bayes merupakan langkah penting dalam mengembangkan sistem yang dapat membantu pemerintah dan pihak terkait dalam memonitor dan mengelola infrastruktur jalan di Bandar Lampung secara lebih efektif dan efisien.

c. Pengujian Algoritma Decision Tree

Analisis Pengujian Algoritma C4.5 tentang Kerusakan Jalan di Bandar Lampung:

1. Konteks Penelitian:

Pengujian algoritma C4.5 dilakukan untuk menganalisis kerusakan jalan di Bandar Lampung. Kerusakan jalan merupakan masalah infrastruktur yang penting dan perlu dianalisis untuk meningkatkan kualitas transportasi dan keselamatan pengguna jalan.

2. Data Pengujian:

Data yang digunakan untuk pengujian algoritma C4.5 mencakup berbagai atribut terkait kondisi jalan di Bandar Lampung, seperti panjang jalan, lebar jalan, tipe lajur jalan, hambatan samping, lebar bahu jalan, dll. Data ini merupakan kumpulan informasi yang terkait dengan kondisi jalan di berbagai lokasi di Bandar Lampung.

3. Tujuan Pengujian:

Tujuan utama pengujian adalah untuk mengidentifikasi pola atau faktor apa yang mempengaruhi kerusakan jalan di Bandar Lampung. Algoritma C4.5 digunakan untuk membangun model keputusan yang dapat mengklasifikasikan kondisi jalan berdasarkan atribut-atribut yang diberikan.

4. Proses Pengujian:

Data yang telah dikumpulkan digunakan sebagai data latih untuk algoritma C4.5. Algoritma C4.5 kemudian menganalisis data latih untuk menentukan aturan-aturan keputusan yang dapat mengklasifikasikan kondisi jalan. Langkah-langkah algoritma, termasuk perhitungan gain informasi, pemilihan atribut terbaik, dan pembangunan pohon keputusan, diterapkan pada data.

5. Hasil dan Analisis:

Setelah proses pengujian selesai, algoritma C4.5 menghasilkan model keputusan yang dapat digunakan untuk memprediksi kerusakan jalan berdasarkan atribut-atribut yang diberikan..Analisis dilakukan terhadap struktur pohon keputusan yang dihasilkan untuk memahami faktor-faktor apa yang paling berpengaruh terhadap kerusakan jalan di Bandar Lampung. Variabel-variabel yang menjadi pembagi (splitter) yang paling signifikan dalam pohon keputusan dapat memberikan wawasan yang berharga tentang kondisi jalan di kota tersebut.

6. Kesimpulan:

Pengujian algoritma C4.5 membantu dalam pemahaman lebih lanjut tentang faktor-faktor yang mempengaruhi kerusakan jalan di Bandar Lampung.

Informasi yang diperoleh dari pengujian ini dapat digunakan untuk pengambilan keputusan yang lebih baik terkait perbaikan dan pemeliharaan infrastruktur jalan di kota tersebut.

Selain itu, model keputusan yang dihasilkan juga dapat menjadi dasar untuk pengembangan sistem pemantauan dan peramalan kerusakan jalan di masa depan.

Pengujian algoritma C4.5 merupakan langkah penting dalam memahami dan mengatasi masalah kerusakan jalan di Bandar Lampung, serta dapat memberikan kontribusi yang signifikan dalam perencanaan dan pengelolaan infrastruktur jalan yang lebih efisien dan efektif.