

**ANALISIS MODEL PREDIKSI  
UNTUK LAYANAN BUS SEKOLAH DI JAKARTA  
MENGUNAKAN PENDEKATAN *MACHINE LEARNING***

**Berdasarkan database open data Jakarta  
Tahun 2017-2019**

**SKRIPSI**

**Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer**



**Oleh:**

**SRIWAHYUNI**

**09011282025045**

**JURUSAN SISTEM KOMPUTER  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA**

**2024**

**LEMBAR PENGESAHAN**

**ANALISIS MODEL PREDIKSI  
UNTUK LAYANAN BUS SEKOLAH DI JAKARTA  
MENGUNAKAN PENDEKATAN *MACHINE LEARNING***

Berdasarkan Database Open Data Jakarta  
Tahun 2017-2019

**SKRIPSI**

Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Serjana Komputer

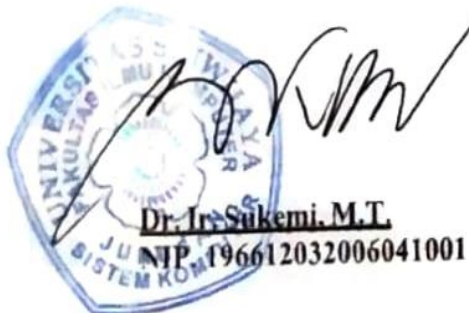
Oleh:

**SRIWAHYUNI  
09011282025045**

Indralaya, 17 Juli 2024

Mengetahui,

Ketua Jurusan Sistem Komputer, Pembimbing Tugas Akhir,

  
**Dr. Ir. Sukemi, M.T.**  
NIP. 196612032006041001


  
**Dr. Rossi Passarella, M.Eng.**  
NIP. 197806112010121004

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada  
Hari : Kamis  
Tanggal : 20 Juni 2024

Tim Penguji

1. Ketua : Kemahyanto Exaudi, M.T



2. Sekretaris : Abdurahman, S.Kom., M.Han.



3. Penguji : Huda Ubaya, M. T.



4. Pembimbing : Dr. Rossi Passarella, M.Eng

Mengetahui, <sup>12/6/24</sup>  
Ketua Jurusan Sistem Komputer



Dr. Ir. Sukemi, M. T

NIP. 196612032006041001

## HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini

Nama : Sriwahyuni

Nim : 09011282025045

Judul : Analisis Model Prediksi Untuk Layanan Bus Sekolah di Jakarta  
Menggunakan Pendekatan *Machine Learning* Berdasarkan Database  
Open Data Jakarta Tahun 2017-2019

### Hasil Pengecekan Plagiat / Turnitin : 6%

Menyatakan bahwa laporan tugas akhir ini adalah hasil karya saya sendiri dan tidak mengandung unsur penjiplakan atau plagiat. Saya sepenuhnya menyadari bahwa jika terbukti adanya penjiplakan atau plagiat dalam laporan tugas akhir ini, saya siap menerima sanksi akademik dari Universitas Sriwijaya. Demikian, pernyataan ini saya buat dengan kesadaran penuh dan tanpa adanya unsur paksaan dari pihak manapun.



Indralaya, Juli 2024  
Penulis.



Sriwahyuni  
NIM. 09011282025045

## KATA PENGANTAR

Assalamualaikum Warahmatullahi Wabarakatuh

Dengan menyebut nama Allah SWT yang Maha Pengasih lagi Maha Penyayang, Puji syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan segala rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan penulisan Proposal Tugas Akhir yang berjudul “**Analisis Model Prediksi Untuk Layanan Bus Sekolah di Jakarta Menggunakan Pendekatan *Machine Learning***”. Proposal ini disusun sebagai salah satu syarat untuk menyelesaikan Program Studi Sistem Komputer di Universitas Sriwijaya.

Dalam laporan ini penulis menjelaskan mengenai model prediksi untuk layanan bus sekolah di Jakarta menggunakan pendekatan machine learning dengan disertai data-data yang diperoleh penulis saat melakukan penelitian dan pengujian data. Penulis berharap agar tulisan ini dapat bermanfaat bagi orang banyak.

Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada beberapa pihak atas ide dan saran serta bantuannya dalam menyelesaikan penulisan Proposal Tugas Akhir ini. Oleh karena itu, penulis ingin menyampaikan rasa syukur kepada Allah SWT dan terima kasih kepada:

1. Kedua orang tua penulis serta keluarga besar atas doa dan dukungan moral dan materil kepada penulis.
2. Bapak Dr. Ir. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Universitas Sriwijaya.
3. Bapak Huda Ubaya, M.T, selaku Sekretaris Jurusan Sistem Komputer Universitas Sriwijaya.
4. Ibu Prof. Dr. Ir. Siti Nurmaini, M.T., selaku Dosen Pembimbing Akademik Jurusan Sistem Komputer, Universitas Sriwijaya.
5. Bapak Dr. Rossi Passarella, M.Eng. selaku Pembimbing Tugas Akhir yang selalu memberikan ilmu serta masukan dalam penulisan tugas akhir.
6. Mbak Sari dan Kak Angga selaku admin Jurusan Sistem Komputer yang telah membantu mengurus seluruh berkas.
7. Semua pihak yang telah membantu.

Penulis menyadari dalam pembuatan proposal ini masih terdapat kekurangan serta sangat jauh dari kata sempurna. Oleh karena itu dibutuhkan kritik dan saran yang bersifat membangun. Harapan dengan adanya Proposal Tugas Akhir ini bermanfaat dan berguna bagi semuanya.

Wassalamualaikum Warahmatullahi Wabarakatuh

Indralaya, Juli 2024

Penulis,

Sriwahyuni

NIM. 09011282025045

**ANALISIS MODEL PREDIKSI  
UNTUK LAYANAN BUS SEKOLAH DI JAKARTA  
MENGUNAKAN PENDEKATAN *MACHINE LEARNING***

**Berdasarkan Database Open Data Jakarta  
Tahun 2017-2019**

**SRIWAHYUNI (09011282025045)**

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

Email: [sw9723011@gmail.com](mailto:sw9723011@gmail.com)

**ABSTRAK**

Penelitian ini bertujuan memprediksi jenis bus sekolah di Jakarta menggunakan metode *machine learning*. Data dari tahun 2017 hingga 2019 mencakup jumlah penumpang, jumlah sekolah, dan jenis bus. Analisis data eksplorasi mengidentifikasi pola dan tren, dengan *feature engineering* menghasilkan tiga variabel utama. Tujuh model machine learning diuji, termasuk SVM, *Logistic Regression*, KNN, *Gaussian Naive Bayes*, *Decision Tree*, *AdaBoost*, dan *Gradient Boosting*, dengan fokus pada f1-score untuk menangani ketidakseimbangan data. Evaluasi menunjukkan bahwa Gradient Boosting memiliki performa terbaik dengan akurasi, precision, recall, dan f1-score tertinggi. Hasil penelitian memberikan wawasan tentang faktor-faktor yang mempengaruhi jenis bus sekolah dan menawarkan model prediktif yang efektif untuk mendukung pengambilan keputusan dalam pengelolaan transportasi sekolah di Jakarta. *Gradient Boosting* terbukti paling andal dalam memprediksi jenis bus sekolah, sehingga model ini bisa menjadi dasar untuk strategi peningkatan keselamatan dan efisiensi transportasi sekolah.

**Kata kunci :** Bus sekolah, Machine learning, Prediksi, Gradient Boosting, Transportasi sekolah

**PREDICTIVE MODEL ANALYSIS  
FOR SCHOOL BUS SERVICES IN JAKARTA  
USING A MACHINE LEARNING APPROACH**

*Based On Database Open Data Jakarta  
For 2017-2019*

**SRIWAHYUNI (09011282025045)**

*Department of Computer Systems, Faculty of Computer Science, Sriwijaya  
University*

Email: [sw9723011@gmail.com](mailto:sw9723011@gmail.com)

**ABSTRACT**

*This research aims to predict the types of school buses in Jakarta using machine learning methods. Data from 2017 to 2019 includes the number of passengers, the number of schools, and bus types. Exploratory data analysis identified patterns and trends, with feature engineering generating three main variables. We tested seven machine learning models, including SVM, Logistic Regression, KNN, Gaussian Naive Bayes, Decision Tree, AdaBoost, and Gradient Boosting, with a focus on f1-score to handle data imbalance. The evaluation shows that gradient boosting has the best performance with the highest accuracy, precision, recall, and f1-score. The results provide insights into the factors that influence school bus types and offer an effective predictive model to support decision-making in school transportation management in Jakarta. Gradient boosting proved to be the most reliable in predicting school bus types. Therefore, this model can serve as a basis for strategies to improve the safety and efficiency of school transportation.*

**Keywords:** *School bus, Machine learning, Prediction, Gradient Boosting, School transportation*



## DAFTAR ISI

LEMBAR PENGESAHAN .....	<b>Error! Bookmark not defined.</b>
HALAMAN PERSETUJUAN.....	<b>Error! Bookmark not defined.</b>
HALAMAN PERNYATAAN .....	iii
KATA PENGANTAR .....	iv
ABSTRAK .....	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR .....	x
DAFTAR TABEL.....	xi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Metodologi Penelitian .....	5
1.7 Sistematis penulisan .....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 Data <i>Science</i> dan Analisis .....	7
2.2 Analisis Data Eksplorasi.....	8
2.3 Pengolahan Data.....	9
2.4 Model Prediksi .....	10
2.4.1 <i>Support Vector Machine (SVM)</i> .....	12
2.4.2 <i>Logistic Regression (LR)</i> .....	12
2.4.3 <i>K-Nearest Neighbors (KNN)</i> .....	13
2.4.4 <i>Gaussian Naive Bayes (GNB)</i> .....	14
2.4.5 <i>Decision Tree (DT)</i> .....	15
2.4.6 <i>Adaptive Boost (AB)</i> .....	16
2.4.7 <i>Gradient Boost (GB)</i> .....	17
2.5 Studi Terkait .....	17
2.6 Landasan Teori .....	20

2.6.1 Transportasi .....	20
2.6.2 Rute Bus .....	22
2.6.3 Pembagian Rute Berdasarkan <i>Type</i> Operasi Bus di Jakarta.....	22
<b>BAB 3 METODOLOGI PENELITIAN.....</b>	<b>25</b>
3.1 Data .....	25
3.2 Kerangka Kerja Penelitian.....	25
3.2.1 Pengumpulan Data .....	27
3.2.2 Analisis Data Eksplorasi.....	27
3.2.3 Pengolahan Data.....	30
3.2.4 Tahap Pemodelan .....	31
3.2.4 Evaluasi Model.....	31
3.2.5 Kesimpulan.....	32
3.3 Spesifikasi Perangkat .....	32
3.3.1 <i>Hardware</i> (Perangkat Keras).....	32
3.3.2 <i>Software</i> (Perangkat Lunak) .....	33
<b>BAB 4 HASIL DAN PEMBAHASAN.....</b>	<b>34</b>
4.1 Analisis Data Eksplorasi.....	34
4.1.1 Diagram Batang Variabel yang digunakan.....	34
4.1.2 <i>Missing Value</i> Objek .....	38
4.1.3 Matriks korelasi .....	39
4.2 Pengolahan Data.....	40
4.2.1 <i>Feature Engineering</i> .....	40
4.2.2 <i>Feature Selection</i> .....	41
4.3 Tahap Pemodelan .....	43
4.4 Evaluasi Model.....	44
4.4.1 Model Menggunakan Data Training .....	44
4.4.2 Model Menggunakan Data <i>Testing</i> .....	47
4.5 Hasil Pengujian.....	49
<b>BAB 5 KESIMPULAN DAN SARAN .....</b>	<b>51</b>
5.1 Kesimpulan.....	51
5.2 Saran.....	51
<b>DAFTAR PUSTAKA .....</b>	<b>52</b>

## DAFTAR GAMBAR

Gambar 2. 1 Bus sedang.....	21
Gambar 2. 2 Bus kecil.....	22
Gambar 3. 1 Kerangka Kerja Penelitian .....	26
Gambar 3. 2 <i>Flowchart</i> Tahapan Penelitian .....	26
Gambar 3. 3 Data mentah.....	28
Gambar 3. 4 Dataset.....	29
Gambar 4. 1 Jumlah sekolah tahun 2017-2019.....	35
Gambar 4. 2 Jumlah Penumpang tahun 2017-2019 .....	36
Gambar 4. 3 Jenis bus tahun 2017-2019 .....	37
Gambar 4. 4 Program Mengecek nilai yang hilang.....	38
Gambar 4. 5 Program matriks korelasi .....	39
Gambar 4. 6 Matriks Korelasi.....	39
Gambar 4. 7 <i>Train</i> dan <i>Test Split</i> .....	43
Gambar 4. 8 Program untuk melihat jumlah nilai setiap variabel.....	43
Gambar 4. 9 Hasil klasifikasi data <i>training</i> dan data <i>testing</i> model GB.....	50

## DAFTAR TABEL

Tabel 2. 1 Perbedaan Model Analisis .....	8
Tabel 2. 2 Lintasan Rute Regular Bus Sekolah Jakarta .....	23
Tabel 2. 3 Lintasan Rute Zonasi Bus Sekolah Jakarta .....	24
Tabel 3. 1 variabel yang dipilih.....	29
Tabel 3. 2 Spesifikasi Hardware .....	32
Tabel 3. 3 Spesifikasi Software.....	33
Tabel 4. 1 <i>Missing Value</i> Objek.....	38
Tabel 4. 2 Statistik variabel dataset.....	40
Tabel 4. 3 Variabel <i>predictor</i> .....	42
Tabel 4. 4 Distribusi dataset variabel target.....	43
Tabel 4. 5 Perbandingan matrix evaluasi data training .....	44
Tabel 4. 6 Peforma f1-score 2 Model terbaik data <i>training</i> .....	46
Tabel 4. 7 Perbandingan matrix evaluasi data <i>testing</i> .....	47
Tabel 4. 8 Peforma f1-score 2 Model terbaik data testing .....	49
Tabel 4. 9 Hasil peforma model data training dan data testing.....	49

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pertumbuhan populasi yang pesat di kota-kota besar tidak hanya berarti bertambahnya jumlah manusia yang tinggal dan bekerja di kota tersebut, tetapi berimplikasi juga terhadap pertumbuhan jumlah permintaan perjalanan. Pertumbuhan permintaan perjalanan ini bukan hanya dari sisi jumlahnya saja, tetapi juga dalam panjang perjalanan karena areal kota bertambah luas seiring dengan meningkatnya kebutuhan ruang bagi kegiatan kota. Pada dasarnya, permintaan akan transportasi dipicu oleh beberapa faktor yaitu kebutuhan manusia untuk berpindah dari satu lokasi ke lokasi lainnya dengan tujuan berpartisipasi dalam berbagai aktivitas, seperti bekerja, berbelanja, pergi ke sekolah, dan lain sebagainya; serta kebutuhan akan transportasi barang yang memungkinkan barang-barang tersebut dapat diangkut dan digunakan atau dikonsumsi di lokasi yang berbeda.

Peningkatan permintaan perjalanan ini tidak hanya berdampak pada jumlah perjalanan yang lebih tinggi, tetapi juga pada panjang perjalanan yang ditempuh. Salah satu konsekuensi dari tingginya penggunaan kendaraan pribadi sebagai sarana transportasi adalah polusi udara yang dihasilkan oleh emisi kendaraan. Istilah “polusi udara” mengacu pada keberadaan gas atau zat *aerosol* di *atmosfer* yang terakumulasi cukup banyak sehingga berdampak pada kesehatan manusia atau kesehatan ekosistem. Ini mungkin hasil dari aktivitas manusia (termasuk pertanian, industri, dan transportasi) atau peristiwa alam (termasuk kebakaran hutan). Menurut laporan dari *World Health Organization (WHO)*, dalam skala global, polusi udara semakin marak [1]. Saat ini, hanya 9% dari semua orang tinggal di daerah di mana pedoman kualitas udara WHO terpenuhi. Menurut EPA, empat sumber utama pencemaran udara adalah sumber tidak bergerak, sumber bergerak, kebakaran, dan biogenik. Transportasi, yang tercakup dalam sumber bergerak, merupakan salah satu sumber polusi udara terbesar, terutama di daerah perkotaan. Di negara-negara *Organization for Economic Co-operation and Development (OECD)*, misalnya, transportasi bertanggung jawab atas setengah dari semua emisi partikel [2].

Untuk mengatasi masalah polusi udara dan mengurangi tingkat emisi kendaraan bermotor, perlu adanya langkah-langkah pengendalian polusi yang berkelanjutan. Salah satu solusi yang dapat diimplementasikan adalah mempromosikan penggunaan kendaraan umum, seperti bus sekolah bagi para pelajar. Dengan mengoptimalkan ketersediaan dan efisiensi rute bus sekolah di Jakarta, diharapkan dapat mengurangi jumlah kendaraan pribadi yang beroperasi di jalan-jalan kota. Penggunaan kendaraan umum juga dapat membantu dalam mengurangi polusi udara karena bus sekolah biasanya beroperasi dengan kapasitas penumpang yang lebih besar dibandingkan kendaraan pribadi. Selain mengurangi tingkat polusi di udara, kendaraan pribadi juga bisa menyebabkan kemacetan di lalu lintas. Salah satu faktor utama penyebab kemacetan lalu lintas adalah volume kendaraan di jalan, yang dipengaruhi oleh beberapa faktor antara lain aksesibilitas ke berbagai tujuan dan moda transportasi. Pada penelitian [3], menyatakan bahwa terdapat lokasi-lokasi di mana orang melakukan banyak lalu lintas dan melakukan berbagai aktivitas meskipun terjadi kemacetan, yang cenderung terpusat, Kawasan terbangun dengan tingkat aksesibilitas yang lebih tinggi. Dengan meningkatkan akses terhadap transportasi umum, Masyarakat akan lebih cenderung menggunakan transportasi umum dibandingkan mengemudi sendiri, sehingga akan membantu mengurangi jumlah kendaraan di jalan dan mengurangi kemacetan.

Transportasi sekolah memainkan peran vital dalam mendukung mobilitas dan akses pendidikan bagi siswa di berbagai daerah. Di kota metropolitan seperti Jakarta, transportasi sekolah menjadi komponen penting untuk memastikan siswa tiba di sekolah dengan aman dan tepat waktu. Namun, meningkatnya jumlah siswa dan kompleksitas jaringan transportasi perkotaan telah menimbulkan tantangan baru dalam mengelola sistem transportasi sekolah yang efektif. Minat pelajar untuk menggunakan transportasi umum sebagai moda transportasi pilihan yang digunakan untuk ke sekolah masih rendah. Hal ini disebabkan oleh beberapa faktor, termasuk *headway* yang terlalu lama antara kedatangan bus [4]. Peningkatan jumlah penumpang dan kepadatan siswa dalam bus sekolah dapat mengindikasikan tingkat keberhasilan suatu rute dan tingkat pemenuhan layanan. Di sisi lain, ketidakseimbangan antara jumlah penumpang dan kapasitas bus dapat menyebabkan ketidaknyamanan dan ketidak efisienan dalam perjalanan.

Berdasarkan penelitian [5], Beberapa temuan utama mengenai kepentingan fitur dan plot dalam hubungan *non-linier* adalah sebagai berikut: Pertama, indikator aksesibilitas menunjukkan sekitar 5 hingga 10% kepentingan fitur kecuali Mart (sekitar 50%). Kedua, aksesibilitas yang lebih baik terhadap infrastruktur transportasi umum, seperti halte bus dan stasiun transit, dikaitkan dengan lalu lintas harian rata-rata tahunan yang lebih tinggi, khususnya di wilayah metropolitan. Ketiga, akses terhadap pasar berskala besar mungkin mempunyai dampak yang tidak diinginkan terhadap volume lalu lintas kendaraan dan mobil. Keempat, terlihat bahwa tingkat lalu lintas rata-rata tahunan yang lebih rendah berhubungan dengan aksesibilitas yang lebih tinggi ke sekolah dasar untuk ketiga moda transportasi tersebut.

Ketersediaan rute bus sekolah yang memadai sangat penting untuk memastikan aksesibilitas siswa ke sekolah. Namun, di Jakarta, ketersediaan rute bus sekolah masih belum memadai [6]. Salah satu faktor yang mempengaruhi ketersediaan transportasi sekolah adalah jumlah penumpang dan kepadatan siswa. Jumlah penumpang yang terlalu banyak dan kepadatan siswa yang tinggi dapat menyebabkan keterlambatan dan ketidaknyamanan dalam perjalanan ke sekolah [7].

Penelitian yang dilakukan oleh [8], terkait penentuan rute dan jumlah kendaraan pada kasus *school bus routing problem*. Hasil penelitian ini memberikan analisis data tentang pengetahuan dibidang penentuan rute, baik dari pengembangan model maupun pengembangan algoritma untuk penyelesaian masalah, namun masih ada beberapa pengembangan yang harus dilakukan. seperti, penentuan nilai parameter pada travel yang lebih sesuai, atau mempertimbangkan jumlah sekolah yang dilintasi penumpang ataupun wilayah apa saja yang dilintasi pada rute bus.

Oleh sebab itu, penelitian ini dilakukan untuk mengetahui lebih dalam terkait masalah tersebut, serta mengetahui bagaimana kinerja dari model prediksi menggunakan pendekatan *machine learning* untuk memprediksi jenis bus yang optimal untuk digunakan dalam layanan bus sekolah. Sehingga, diharapkan dapat meningkatkan efisiensi, kualitas, dan kepuasan layanan bus sekolah di Jakarta melalui penggunaan jenis bus yang tepat sesuai dengan kebutuhan dan permintaan yang diprediksi.

## 1.2 Rumusan Masalah

Berdasarkan data bus sekolah yang telah dikumpulkan dari *website* OpenData Jakarta, akan digunakan untuk mengetahui bagaimana hubungan jumlah penumpang dan jumlah sekolah dalam memprediksi jenis bus yang optimal untuk digunakan dalam layanan bus sekolah. Sehingga, diharapkan dapat meningkatkan efisiensi, kualitas, dan kepuasan layanan bus sekolah di Jakarta melalui penggunaan jenis bus yang tepat sesuai dengan kebutuhan dan permintaan yang diprediksi.

## 1.3 Batasan Masalah

Batasan masalah yang menjadi acuan dalam penelitian ini adalah:

1. Raw data yang digunakan yaitu jenis bus sekolah yang telah dikumpulkan dari *website* OpenData Jakarta.
2. Dalam melakukan analisis model prediksi untuk layanan bus sekolah di Jakarta menggunakan pendekatan *Machune Learning* yang menggunakan perbandingan 7 model yaitu *SupportVerctorMachine* (SVM), *LogisticRegression* (LR), *K-Nearest Neighbors* (KNN), *GaussianNaiveBayes* (GNB), *DecisionTreeClassifier* (DT), *Adative Boost* (AB), dan *Gradient Boost* (GB).
3. Data rute bus yang diambil mulai dari bulan januari 2017 hingga bulan desember 2019.
4. Analisis hanya akan mempertimbangkan jenis bus sekolah yang beroperasi secara regular dan tidak akan mempertimbangkan jenis bus yang bersifat khusus atau tidak tetap.

## 1.4 Tujuan Penelitian

Penelitian ini juga memiliki tujuan yang akan dicapai yaitu:

1. Untuk memprediksi jenis bus yang optimal untuk digunakan dalam layanan bus sekolah Jakarta.
2. Untuk mengevaluasi dan membandingkan kinerja tujuh model machine learning dalam memprediksi jenis bus.

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini dapat dilihat sebagai berikut:

1. Dengan mengetahui jenis bus yang paling sesuai untuk kebutuhan transportasi



sekolah, sekolah atau lembaga pendidikan dapat mengalokasikan sumber daya mereka dengan lebih efisien.

2. Dengan menggunakan jenis bus yang sesuai, risiko kecelakaan atau insiden lainnya dapat dikurangi, meningkatkan keamanan siswa. Selain itu, penggunaan bus yang sesuai juga dapat meningkatkan kenyamanan dan kepuasan siswa dalam perjalanan.
3. Mendukung pengambilan keputusan yang lebih efektif dalam perencanaan transportasi sekolah. Dengan menggunakan pendekatan *machine learning*, penelitian ini akan memberikan pandangan yang komprehensif tentang layanan bus sekolah.

## **1.6 Metodologi Penelitian**

Metode yang digunakan dalam penelitian ini yaitu sebagai berikut:

### **1. Tahap Studi Pustaka/Literatur**

Tahap pertama yang dilakukan penulis yaitu mengidentifikasi permasalahan yang berkaitan dengan penelitian yang akan dilakukan, kemudian mencari sumber dari buku, jurnal, paper dan melalui internet yang berhubungan dalam penelitian ini.

### **2. Tahap Perancangan**

Tahap perancangan ini penulis akan merancang permasalahan dengan menentukan metode atau teknik yang akan digunakan untuk melakukan analisis data yang akan digunakan pada penelitian ini.

### **3. Tahap Pengujian**

Tahap pengujian ini dilakukan untuk mendapatkan hasil dari metode atau teknik yang digunakan dalam melakukan analisis data dari penelitian ini.

### **4. Tahap Analisa dan Pembahasan**

Pada tahapan ini penulis melakukan analisis dari hasil pengujian metode yang digunakan sebelumnya untuk menemukan hasil dari permasalahan pada penelitian yang dilakukan.

### **5. Tahap Kesimpulan dan Saran**

Tahapan terakhir yang dilakukan kesimpulan, penulis menarik kesimpulan dari hasil analisa dan pembahasan yang sudah dijelaskan pada tahap sebelumnya, dan juga memberikan saran dari hasil penelitian yang telah dilakukan.

### **1.7 Sistematis penulisan**

Penelitian ini memiliki sistematika penulisan dari setiap bab akan dibuat secara sistematika penulisan, yaitu sebagai berikut:

#### **BAB I: PENDAHULUAN**

Bab ini berisi pemaparan mengenai topik penelitian yang mencakup latar belakang, tujuan penelitian, manfaat penelitian, rumusan masalah, batasan masalah dan metodologi penelitian secara sistematis.

#### **BAB II: TINJAUAN PUSTAKA**

Bab ini berisi tentang penjabaran mengenai dasar teori yang relevan dengan topik penelitian yang akan dilakukan.

#### **BAB III: METODOLOGI PENELITIAN**

Bab ini menjelaskan bagaimana kerangka kerja dari penelitian secara rinci dan proses penelitian yang dilakukan hingga selesai.

#### **BAB IV: HASIL DAN ANALISA**

Bab ini menjelaskan bagaimana hasil dari metode yang digunakan dalam proses penelitian ini.

#### **BAB V: KESIMPULAN**

Bab ini berisi kumpulan tentang penelitian yang dilakukan, dan menjawab tujuan dari pencapaian pada BAB I Pendahuluan. Serta memberikan saran untuk penelitian yang akan dilakukan selanjutnya.

## DAFTAR PUSTAKA

- [1] P. Sicard, E. Agathokleous, A. De Marco, E. Paoletti, and V. Calatayud, “Urban population exposure to air pollution in Europe over the last decades,” *Environ. Sci. Eur.*, vol. 33, no. 1, p. 28, 2021, doi: 10.1186/s12302-020-00450-2.
- [2] D. L. Mendoza *et al.*, “Idle-Free Campaign Survey Results and Idling Reductions in an Elementary School,” *Vehicles*, vol. 4, no. 3, pp. 865–902, 2022, doi: 10.3390/vehicles4030048.
- [3] A. Mondschein and B. Taylor, “Is traffic congestion overrated? Examining the highly variable effects of congestion on travel and accessibility,” *J. Transp. Geogr.*, vol. 64, pp. 65–76, Oct. 2017, doi: 10.1016/j.jtrangeo.2017.08.007.
- [4] B. B. PERDANA, “Perencanaan Angkutan Sekolah Di Kabupaten Bantul,” 2022, [Online]. Available: [http://digilib.ptdisttd.net/2416/%0Ahttp://digilib.ptdisttd.net/2416/1/Draft KKW Final %28Bhima Perdana%29.pdf](http://digilib.ptdisttd.net/2416/%0Ahttp://digilib.ptdisttd.net/2416/1/Draft%20KKW%20Final%20Bhima%20Perdana%29.pdf)
- [5] S. Lee, J. Yang, and K. Cho, “The Influence of Transportation Accessibility on Traffic Volumes in South Korea: An Extreme Gradient Boosting Approach,” 2023.
- [6] F. L. Nabilah, S. Anggada, S. K. Koto, and R. P. Naskah, “Perencanaan angkutan sekolah di kabupaten landak,” no. 5, pp. 1–15, 2022.
- [7] N. A. K. Sakti and S. T. Pamungkas, “Kondisi Fisik Jalur Pejalan Kaki pada Koridor Jalan Veteran Malang dalam Konteks Kenyamanan Spasial,” *J. Mhs. Jur. Arsit. Univ. Brawijaya*, vol. 7, no. 3, 2019.
- [8] F. Ramadhan and A. Imran, “Penentuan Rute Dan Jumlah Kendaraan Pada Kasus School Bus Routing Problem: Penerapan Algoritma Record-to-record Travel,” *J. Rekayasa Sist. Ind.*, vol. 5, no. 01, p. 15, 2018, doi: 10.25124/jrsi.v5i01.307.
- [9] K. Smedt, D. Koureas, and P. Wittenburg, *FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units*. 2020. doi:

10.20944/preprints202003.0073.v1.

- [10] S. Bibri, “Smart Sustainable Urbanism: Paradigmatic, Scientific, Scholarly, Epistemic, and Discursive Shifts in Light of Big Data Science and Analytics,” 2019, pp. 131–181. doi: 10.1007/978-3-030-17312-8\_6.
- [11] F. Emmert-Streib and M. Dehmer, “Defining Data Science by a Data-Driven Quantification of the Community,” *Mach. Learn. Knowl. Extr.*, vol. 1, no. 1, pp. 235–251, 2019, doi: 10.3390/make1010015.
- [12] F. Emmert-Streib and M. Dehmer, “Understanding Statistical Hypothesis Testing: The Logic of Statistical Inference,” *Mach. Learn. Knowl. Extr.*, vol. 1, pp. 945–961, Aug. 2019, doi: 10.3390/make1030054.
- [13] V. Yellapu, “Descriptive statistics,” *Int. J. Acad. Med.*, vol. 4, p. 60, Apr. 2018, doi: 10.4103/IJAM.IJAM\_7\_18.
- [14] V. Yellapu, “Full Text Introduction,” *Descr. Stat.*, vol. 4, no. 1, pp. 60–63, 2018, doi: 10.4103/IJAM.IJAM.
- [15] T. Shukla, *Data Processing*. 2018. doi: 10.13140/RG.2.2.35660.10889.
- [16] S. Peng, J. Zhu, Z. Liu, B. Hu, M. Wang, and S. Pu, “Prediction of Ammonia Concentration in a Pig House Based on Machine Learning Models and Environmental Parameters,” *Animals*, vol. 13, no. 1, pp. 1–14, 2023, doi: 10.3390/ani13010165.
- [17] T. D. Synthesis, “HT-Fed-GAN: Federated Generative Model for Decentralized,” pp. 1–22, 2022.
- [18] C. M. Q. Ramos, P. J. S. Cardoso, H. C. L. Fernandes, and J. M. F. Rodrigues, “A Decision-Support System to Analyse Customer Satisfaction Applied to a Tourism Transport Service,” *Multimodal Technol. Interact.*, vol. 7, no. 1, 2023, doi: 10.3390/mti7010005.
- [19] L. Kolsi, S. Al-Dahidi, S. Kamel, W. Aich, S. Boubaker, and N. Ben Khedher, “Prediction of Solar Energy Yield Based on Artificial Intelligence Techniques for the Ha’il Region, Saudi Arabia,” *Sustain.*, vol. 15, no. 1, pp. 1–15, 2023, doi: 10.3390/su15010774.
- [20] S. Isak-Zatega, A. Lipovac, and V. Lipovac, “Logistic regression based in-service assessment of mobile web browsing service quality acceptability,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2020, no. 1, 2020, doi:

10.1186/s13638-020-01708-2.

- [21] Y. Lin *et al.*, “Application of Logistic Regression and Artificial Intelligence in the Risk Prediction of Acute Aortic Dissection Rupture,” *J. Clin. Med.*, vol. 12, no. 1, 2023, doi: 10.3390/jcm12010179.
- [22] N. Shabbir, K. Vassiljeva, H. Nourollahi, O. Husev, E. Petlenkov, and J. Belikov, “Comparative Analysis of Machine Learning Techniques for Non-Intrusive Load Monitoring,” *Electron.*, vol. 14, 2024.
- [23] B. Poggi, C. Babatoude, E. Vittori, and T. Antoine-Santoni, “Efficient WSN Node Placement by Coupling KNN Machine Learning for Signal Estimations and I-HBIA Metaheuristic Algorithm for Node Position Optimization,” *Sensors*, vol. 22, no. 24, 2022, doi: 10.3390/s22249927.
- [24] S. Park, D. Jung, H. Nguyen, and Y. Choi, “Diagnosis of problems in truck ore transport operations in underground mines using various machine learning models and data collected by internet of things systems,” *Minerals*, vol. 11, no. 10, 2021, doi: 10.3390/min11101128.
- [25] X. Dai *et al.*, “A Decision-Tree Approach to Identifying Paddy Rice Lodging with Multiple Pieces of Polarization Information Derived from Sentinel-1,” *Remote Sens.*, vol. 15, no. 1, 2023, doi: 10.3390/rs15010240.
- [26] M. Balfaqih, S. A. Alharbi, M. Alzain, F. Alqurashi, and S. Almilad, “An accident detection and classification system using internet of things and machine learning towards smart city,” *Sustain.*, vol. 14, no. 1, pp. 1–13, 2022, doi: 10.3390/su14010210.
- [27] C. Eyupoglu, “Novel CAD Diagnosis Method Based on Search , PCA , and AdaBoostM1 Techniques,” pp. 1–17, 2024.
- [28] G. Guo, S. Li, Y. Liu, Z. Cao, and Y. Deng, “Prediction of Cavity Length Using an Interpretable Ensemble Learning Approach,” *Int. J. Environ. Res. Public Health*, vol. 20, no. 1, 2023, doi: 10.3390/ijerph20010702.
- [29] D. Bertsimas, A. Delarue, and S. Martin, “Optimizing schools’ start time and bus routes,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 13, pp. 5943–5948, 2019, doi: 10.1073/pnas.1811462116.
- [30] H. I. Calvete, C. Galé, J. A. Iranzo, and P. Toth, “A partial allocation local search matheuristic for solving the school bus routing problem with bus stop

- selection,” *Mathematics*, vol. 8, no. 8, 2020, doi: 10.3390/MATH8081214.
- [31] Z. Liu, L. Gang, B. Yu, and H. Zhang, “The routing problem for school buses considering accessibility and equity,” *Transp. Res. Part D Transp. Environ.*, vol. 107, no. January, p. 103299, 2022, doi: 10.1016/j.trd.2022.103299.
- [32] M. Li and J. Y. J. Chow, “School bus routing problem with a mixed ride, mixed load, and heterogeneous fleet,” *Transp. Res. Rec.*, vol. 2675, no. 7, pp. 467–479, 2021, doi: 10.1177/03611981211016860.
- [33] S. Sakphrom, K. Suwannarat, R. Haiges, and K. Funsian, “A simplified and high accuracy algorithm of rssi-based localization zoning for children tracking in-out the school buses using bluetooth low energy beacon,” *Informatics*, vol. 8, no. 4, 2021, doi: 10.3390/informatics8040065.
- [34] Z. Yin and B. Zhang, “Bus Travel Time Prediction Based on the Similarity in Drivers’ Driving Styles,” *Futur. Internet*, vol. 15, no. 7, pp. 1–18, 2023, doi: 10.3390/fi15070222.
- [35] M. H. Suh and M. Jeong, “Development of Bus Routes Reorganization Support Software Using the Naïve Bayes Classification Method,” *Sustain.*, vol. 14, no. 8, 2022, doi: 10.3390/su14084400.
- [36] S. Deb and X. Z. Gao, “Prediction of Charging Demand of Electric City Buses of Helsinki, Finland by Random Forest,” *Energies*, vol. 15, no. 10, 2022, doi: 10.3390/en15103679.
- [37] S. Sarjana, “Urban Public Transportation Perspective in Meta-Analysis Study,” *Media Komun. Tek. Sipil*, vol. 27, no. 2, pp. 277–287, 2021, doi: 10.14710/mkts.v27i2.40635.
- [38] O. Shirokorad, S. Arkhipov, P. Volodkin, A. I. Fadeev, and S. Alhuseini, “Improving the Bus Stops Capacity Depending on Capacity of Bus Lane and Regulated Intersection Capacity Improving the Bus Stops Capacity Depending on Capacity of Bus Lane and Regulated Intersection Capacity,” 2023, doi: 10.1088/1755-1315/1232/1/012050.