International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 3, May-June 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

An Intelligent Bus Travel Booking and Management System Using Artificial Intelligence

Mrs.M.Vasuki[1], Dr.T.Amalraj Victoire[2], B.Gajalakshmi[3*]
[1]Associate Professor, Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107.
Email: vasukimca@smvec.ac.in
[2]Professor, Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107.
Email: amalraj@smvec.ac.in
[3]PG Student, Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107.
Email: gajalakshmibabu85@gmail.com

--*************************--

Abstract:
The utilization of AI for better efficiency, comfort, and management is increasing in the domain of smart transportation systems. There are numerous challenges associated with the current bus booking systems, including the prediction of Delays, Inefficient Fleet Management, Static Pricing, and the inability to provide Intelligent Assistance to Customers. In this research, we introduce an AI-powered smart bus booking and management system that makes use of machine learning, natural language processing, GPS technology, and cloud computing technology to provide intelligent transportation services. This framework will utilize linear regression, random forests, gradient boosting, isolation forest algorithms, and naive Bayes for delay prediction, risk assessment for weather conditions, dynamic pricing, predictive maintenance, and sentiment analysis for customer assistance, respectively. The proposed framework will provide an efficient QR ticketing system and will also allow real-time tracking of buses to provide a more efficient and safer transportation management service.

Keywords - Artificial Intelligence, Smart Transportation, Machine Learning, Dynamic Pricing, Predictive Maintenance, Real-Time Bus Tracking, Intelligent Route Planning, Smart Bus Management.

---------------------------------------*************************----------------------------------

I. INTRODUCTION

The provision of public transport services is essential in facilitating movement in urban centres and addressing traffic issues in poor urban centres; thus, urban public transport systems are responsible for fulfilling social and economic needs through their ability to provide mobility. Many of the systems being used today to book and manage public transportation are based on older and less efficient methods. For example, many transport systems experience challenges due to ineffective scheduling and communication, poor monitoring of the fleet and fraudulent tickets, and a lack of intelligence within customer service. The effect of these deficiencies can result in low levels of passenger satisfaction and create difficulties in public transport operations' management.

Recent advances in Artificial Intelligence (AI) and Machine Learning (ML) present an opportunity to design intelligent transportation systems that can help address many of these existing problems. Examples of such smart transportation systems include automated routing, delay monitoring, fleet management, effective communication between passengers and staff, and the use of other innovative technologies to optimize both the travel time and to assist decision-makers and the users of public transport in their day-to-day activities.

The smart bus booking/management system powered by AI is the first of its kind that brings together multiple intelligent capabilities into a single cloud-based system. This system provides different features that add value to users, such as ticketing services, QR codes for ticket validation, GPS tracking, dynamic pricing based on real-time data, predictive maintenance of vehicles, AI chatbots to assist customers with questions/information, and the ability to analyse customer reviews of the service provided.

This project aims to create and deploy an intelligent transportation terminology that will enhance efficiency, dependability, scalability, and the overall experience of the passenger. By utilizing predictive analytics and automated monitoring technologies, the design and delivery of traditional transportation services will transition into sophisticated smart mobility.

II. LITERATURE SURVEY

Kumar and Priya (2023) have presented a QR Code-based ticketing system that will help users access a ticketing platform automatically through the use of Cloud Storage Technologies to store the ticketing information and validate the user's ticket (QR Code).

Arun et al. (2024) created a new method for tracking buses and providing automated notifications to passengers. They built a machine learning-based framework that will track the location of buses in order to accurately predict their arrival times based on both traffic flow and patterns of movement. Additionally, there are real-time notification services that can help to improve communication with passengers. Overall, the data collected from this study will provide new opportunities for improving transportation monitoring. Dynamic pricing, as well as predictive maintenance, have not yet been implemented in this framework.

Patel and Sharma (2023) developed an Automated Transport Management System that incorporates GPS technology to help optimize routes and track attendance. Their system uses Cloud-computing technology to facilitate vehicle tracking and management. The goal of their framework is to improve operational efficiency and automate the recording of attendance in an educational setting. Although their framework is beneficial for institutions, there is no AI-based analytics or intelligent rescheduling capability available at this time.

Rahman et al. (2025) in their study "Machine Learning-Enabled Smart Transit: Real-Time Bus Tracking System for Improved Urban Mobility" put forth an AI-powered framework for urban transportation management. The proposed smart transit solution uses GPS tracking to identify bus delays and monitor trips through machine learning algorithms. The predictive analytics used in the system help improve passenger comfort and ease of use, as well as the efficiency of public transit. The proposed framework also provides a framework for smarter urban mobility services. The only limitations were that they did not include fleet health monitoring and chatbot services.

Ramesh et al. (2024) in “Smart Bus Route Management System with Real-Time Tracking and Passenger Assistance” introduced a smart transportation system based on GPS technology and machine learning algorithms for monitoring buses. The system offered live updates about the current location, routing optimization, and information notification to the passengers. In their work, cloud computing technologies were applied to the management of the transportation processes. The study contributed to the development of a user-friendly transportation system. Nonetheless, predictive maintenance and intelligent pricing capabilities were not incorporated.

In summarizing this research on smart transportation systems, there have been many uses of AI, GPS, cloud computing, and machine learning throughout the development of today’s smart transportation systems. The majority of smart transportation systems that currently exist have a primary emphasis on either tracking or ticketing. 2. In contrast to existing smart transportation systems, the proposed AI-Powered Smart Bus Booking and Management System will provide full smart transportation solutions across multiple intelligence modules, including delay prediction, dynamic price estimation, predictive maintenance, sentiment analysis, and chatbot assistance.

III. EXISTING SYSTEM AND DRAWBACKS

A. No Real-Time Passenger Information

Most traditional bus systems rely on fixed schedules that do not consider real-time traffic conditions, weather changes, or festival rush. As a result, passengers are unable to receive accurate arrival updates and often experience uncertainty during travel. Most conventional bus systems operate on fixed schedules and don't take into account current traffic conditions, weather changes, or special holiday rushes. Because of this, passengers have no way of knowing when their bus will arrive and will always be uncertain about the timing of their bus trips.

B. Manual Ticketing

Conventional transportation systems mainly depend on paper-based or cash-based ticketing methods. This increases boarding time, introduces human errors, and creates difficulties in maintaining accurate transaction records. Most traditional public transport systems use paper or cash-based methods of ticketing. This results in delays when people board the bus, creates opportunities for human error to occur when an operator uses a ticketing machine, and makes it difficult for operators to keep accurate records of tickets sold.

C. Static Fare Structures

Existing systems use fixed pricing models that remain unchanged regardless of passenger demand or travel conditions. This limits revenue optimization during peak hours and reduces the flexibility of fare management. Today's public transport systems tend to use a fixed price model regardless of passenger demand or travel conditions. This reduces the ability of operators to maximize their revenues during peak periods and also reduces the operator's ability to manage their pricing in a flexible manner.

D. Reactive Fleet Maintenance

Bus maintenance activities are generally performed only after vehicle failures occur. This reactive approach leads to unexpected breakdowns, service interruptions, increased operational costs, and passenger inconvenience.

E. Absence of Feedback Analysis

Existing technologies do not usually have access to automated customer support systems for booking help, cancellations, etc., so riders must rely upon manual customer service representatives when attempting to get information. There is a significant amount of time wasted waiting to connect with an agent while they search for/help you find what you need, and this creates an inefficient support system.

F. No Automated Customer Support Services

Many existing systems do not provide automated customer support services for booking assistance, cancellations, or travel-related queries. Passengers must depend on manual helplines, which increases response time and reduces service efficiency.

IV. PROPOSED SYSTEM

The proposed solution proposes an integrated (AI) powered Smart Bus Travel Booking and Management System utilizing six (AI/ML) Modules as part of one (Unified) Platform. The system utilizes a three-level architecture as designed: Frontend utilizing React/VITE, Back-end utilizing Python/Flask API calls, Database using MySQL Database technology.

A. System Overview

The System is viewed as two major user roles, Passenger and Bus Operator. Passenger Users will be able to search for a Bus, reserve a seat, generate a QR Code Ticket, Track Buses in Real Time, request a Reschedule based on AI proposals, and utilize an NLP 24/7 Chatbot. On the other hand, Bus Operators will be able to manage the overall Health of Their Fleet, control pricing dynamically, manage routes, and View Customer Sentiment Analytics Dashboard.

B. Main Characteristics

1- Dynamic Pricing Engine

Fare multipliers of 0.7x-2.5x will be computed automatically by an AI based on demand, time of day, weather, and the festival/fair calendar.

2- Fleet Health Monitoring

Using telemetry data, abnormal behaviour will be flagged to identify buses requiring maintenance prior to any breakdown occurring.

3- AI Chatbot Assistant

A Natural Language Processing (NLP) enabled Conversational Agent will be available 24/7, supporting booking, cancellation and FAQs.

4- Sentiment Analysis

The system will automatically categorize passenger reviews into Positive, Negative or

Neutral categories to provide insight for the operator.

5- AI-Powered Lost and Found Management

This internal module for passengers to report missing items or for staff to enter found item information has been enhanced by adding autonomous lost-and-found capabilities through AI Text Matching and NLP to provide efficient and effective recovery of passengers' lost property by quickly and easily performing text matches between passenger descriptions of what was lost and staff entries for all resolved found items.

V. SYSTEM ARCHITECTURE

The system is designed using a three-tiered or 3-layered approach for architecture that consists of an Application Layer, Presentation Layer, Database Layer, as well as AI Service Modules. The Presentation Layer gives customers, staff & administrators a web-based dashboard or booking portal to use; this is how users will interact with the system. The Application Layer provides services that authenticate users, manage bookings and payment processing, schedule their routes, and provide communication between the AI Modules. The AI Service Modules allow for intelligent operation such as predicting delays, dynamic pricing strategies, analyzing customer sentiment, predictive auto-scheduling of rides using AI chatbots and monitoring fleet health utilizing ML algorithms to provide a real-time transportation management system. Finally, the Data Base Layer contains user information, booking information, bus schedules, payment records, customer feedback, and any predictions generated by the AI Service Modules; thus, this type of architecture has been created to provide increased scalability and security, improved maintainability and enhanced efficiency in real-time transportation management systems.

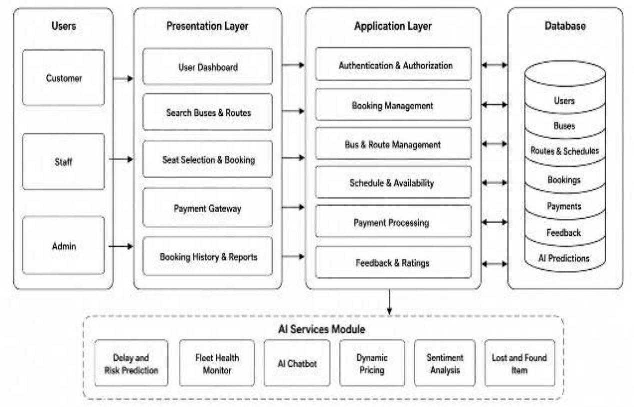


Figure 1

System architecture diagram

VI. METHODOLOGY

A. Data Preparation

Synthetic datasets were created for the project that included data on 1,000 or more samples per module to simulate the types of transport scenarios that occur in real-world India, including: Routes (Chennai to Bengaluru; Mumbai to Pune); Festivals (Diwali; Pongal; Navratri) overlaid onto the transport calendar; weather (clear; raining; storm); peak vs. off-peak; and Vehicle Sensors (Telematics) with data covering a wide range. All datasets were verified prior to being used for training by conducting a feature correlation analysis to confirm that there was sufficient correlation between features and the number of occurrences of each feature class (balance) in the dataset.

B. Model Training

All Models were developed using Scikit-learn. Features were developed for each model using TF-IDF for text inputs, ordinal encoding for categorical attributes, and standard scaling for numerical values. The Models created were saved as JobLib Pipelines so they can be deployed in the Future. Development of hyperparameters for each of the models developed was accomplished using cross-validation techniques.

C. Application Programming Interface Integration

Each of the trained pipelines is incorporated into a Flask Blueprint with a separate POST endpoint. The React front end calls these endpoints asynchronously through Axios, which sends back the prediction results as JSON objects to render into real-time UI components. By separating the design of the individual modules, each one can be retrained and redeployed independently without affecting the other components of the system.

VII. ALGORITHM USED

A. Linear Regression – E.T.A. Estimation and Delay

Linear Regression will provide estimates for the (E.T.A.) Estimated Time of Arrival and how long the delay (in minutes) was, based on a variety of factors, including the distance of the route, time of day, day of week, weather index, and historical average speed for the same type of run. Since the actual delays are continuous (i.e. number of minutes), linear regression provides an updated arrival estimate (i.e. E.T.A.) dynamically, thereby providing a low-latency solution for this use case. In addition, linear regression is very easy to understand and interpret, which also makes it a good fit for this particular application

Algorithm (Pseudocode):

```
Algorithm 1: Linear Regression – ETA & Delay
Estimation
Input : route_distance, time_of_day,
day_of_week,
        weather_index, hist_avg_speed
Output: predicted_delay (minutes),
estimated_ETA
────────────────────────────────────────────
─────────
1. Collect training data D = {(x_i, y_i)}
where
   x_i = feature vector, y_i = actual delay
(min)
2. Compute coefficients β using Ordinary
Least Squares:
   β = (X^T X)^{-1} X^T y
3. For a new trip t:
   predicted_delay = β_0 + β_1*dist +
β_2*time
                   + β_3*day + β_4*weather +
β_5*speed
4. estimated_ETA = scheduled_arrival +
predicted_delay
5. Return predicted_delay, estimated_ETA
```

Sample Output:

```
  Route         : Chennai → Bengaluru
  Distance      : 346 km
  Weather Index: 0.7 (light rain)
  Predicted Delay : 18 minutes
  Estimated ETA    : 14:38 (original 14:20)
```

B. Random Forest Classifier – Rescheduling

A random forest classifier is used to determine whether or not a bus service requires automated rescheduling based on predicted severity of delay, weather risk score, time criticality and availability of alternative routes. The random forest was chosen due to its robustness as an ensemble and its ability to model non-linear interactions between delay elements (i.e., rain + festival + peak hour). The output will be a binary recommendation with a confidence score that can be used to prioritize alternative routes.

Algorithm (Pseudocode):

```
Algorithm 2: Random Forest Classifier –
Rescheduling
Input: delay_severity, weather_risk,
time_criticality,
        alt_route_availability
Output: reschedule_needed (0/1),
confidence_score, alt_route
────────────────────────────────────────────
─────────
1. Train ensemble of N decision trees T_1 …
T_N
   on historical delay + rescheduling
records
2. For each tree T_k:
   a. Bootstrap-sample training data
   b. At each split select random subset of
features
   c. Grow tree to maximum depth
3. For new input x:
   votes_k = T_k.predict(x)        for k =
1..N
   reschedule_needed =
majority_vote(votes_1..N)
   confidence_score   = votes_for_class / N
4. If reschedule_needed == 1:
   Rank alternative routes by suitability
score
5. Return reschedule_needed,
confidence_score, top_alt_route
```

Sample Output:

```
  Input Delay       : 35 minutes              |
Weather Risk: High
  Rescheduling      : RECOMMENDED
  Confidence Score : 0.84
  Alternative Bus   : BUS-442 | Dept 15:10 |
Fare diff +₹40
```

C. Gradient Boosting Regressor – Dynamic Pricing

The dynamic pricing engine uses a gradient boosting regressor to give real-time fare multipliers. The input variables used by the regressor consist of the seat occupancy ratio, time of day, day of the week, weather severity and whether or not there are any festivals occurring, and the popularity of the route. The regressor will output a single continuous value (multiplier) that will fall within the range of [0.7, 2.5]. Values less than 1.0 are discounted fare multipliers due to off-peak hours, and values greater than 1.0 will be fare multipliers due to surge demand.

Algorithm (Pseudocode):

```
Algorithm 3: Gradient Boosting Regressor -
Dynamic Pricing
Input : occupancy_ratio, time_of_day,
day_of_week,
        weather_severity, festival_flag,
route_popularity
Output: fare_multiplier ∈ [0.7, 2.5]
──────────────────────────────────────────────
─────────

1. Initialise model F_0(x) = mean(y_train)
2. For m = 1 to M boosting rounds:
   a. Compute pseudo-residuals:
      r_i = y_i − F_{m-1}(x_i)
   b. Fit weak learner h_m to residuals r
   c. Find step size γ minimizing loss:
      γ_m = argmin Σ L(y_i, F_{m-1}(x_i) + γ
h_m(x_i))
   d. Update: F_m(x) = F_{m-1}(x) + η * γ_m
* h_m(x)
3. Final multiplier = clip(F_M(x), 0.7, 2.5)
4. fare = base_fare × multiplier
```

Sample Output:

```
  Occupancy  : 92%  |  Festival : Diwali  |
Weather: Clear
  Base Fare  : ₹450
  Multiplier : 2.1×
  Final Fare : ₹945
```

D. Multinomial Naïve Bayes – Sentiment Analysis

Passenger feedback is categorized as positive, negative, or neutral using a multinomial naive bayes classifier that has been trained on tf-idf vectorized text (limited to 5,000 features and bi-gram range). Naive bayes is considered an efficient text classification method because it requires relatively few training examples and provides excellent classification performance on high dimensional sparse feature representations.

Algorithm (Pseudocode):

```
Algorithm 4: Multinomial Naive Bayes -
Sentiment Analysis
Input : passenger_review (raw text)
Output: sentiment in {Positive, Neutral,
Negative}
1. Pre-process: lowercase, remove stopwords,
TF-IDF vectorise
2. For each class c in {Pos, Neu, Neg}:
   P(c|x) = P(c) x Prod P(x_j|c) with
Laplace smoothing
3. Predicted class = argmax_c P(c|x)
```

Sample Output:

```
  Review   : "The bus was clean and arrived
on time. Great service!"
  Positive : 0.91  |  Neutral: 0.06  |
Negative: 0.03
  Sentiment: POSITIVE
```

E. Logistic Regression for Intent Classification

The Natural Language Processing Chatbot uses Logistic Regression through TF-IDF vectorization to classify User's input into one of 12 Intent Categories (for example: Booking Query, Cancellation Request, Route Information, Schedule Check, Feedback Submission, Greetings, etc.). The model uses probabilistic confidence scores to allow for graceful resolution of ambiguous inputs by returning the response with the highest confidence score.

Algorithm (Pseudocode):

```
Algorithm 5: Logistic Regression - Chatbot
Intent Classification
Input : user_message (raw text)
Output: intent_label, confidence_score
──────────────────────────────────────────────
─────────

1. Pre-process message: lowercase, tokenise,
remove stopwords
2. Vectorise with TF-IDF (uni + bi-grams)
3. For each intent class k = 1..12:
   z_k = w_k^T * x + b_k
   P(class=k | x) = exp(z_k) / Σ_j exp(z_j)
[Softmax]
4. predicted_intent = argmax_k P(class=k |
x)
5. confidence = max P(class=k | x)
6. If confidence < 0.50: return
fallback_response
7. Else: fetch response template for
predicted_intent
8. Return intent_label, response, confidence
```

Sample Output:

```
  User Input   : "I want to cancel my ticket
for tomorrow"
  Intent       : cancellation_request
  Confidence   : 0.93
  Bot Response : "Sure! Please provide your
booking ID
                 to proceed with
cancellation."
```

F. Isolation Forest for Detecting Anomalies in Fleet Health

Isolation Forest detects anomalous buses from telemetry features: engine temperature, oil pressure, mileage since last service, vibration level, and brake pad thickness. The model is configured with contamination = 0.05 (5%), representing the expected proportion of buses requiring maintenance. Anomalous buses (predicted label: −1) are flagged for inspection in the admin dashboard.

Algorithm (Pseudocode):

```
Algorithm 6: Isolation Forest – Fleet Health
Anomaly Detection
Input : telemetry = {engine_temp,
oil_pressure, mileage,
                     vibration,
brake_pad_thickness}
Output: label ∈ {1 Normal, -1 Anomalous},
anomaly_score
──────────────────────────────────────────
─────────

1. Build forest of T isolation trees:
   For each tree t = 1..T:
     a. Sub-sample ψ points from training
data
     b. Recursively partition by random
feature + split
     c. Record path length h(x) to isolate
each point
2. For a new bus sensor vector x:
   avg_path = mean(h_t(x)) over T trees
3. anomaly_score s(x) = 2^{-avg_path / c(ψ)}
   where c(ψ) = normalisation constant
4. If s(x) > threshold (contamination=0.05):
label = -1
   Else: label = +1
5. Return label, anomaly_score
```

Sample Output:

```
Sample Output (Bus ID: BUS-217):
  Engine Temp  : 112°C  (threshold 95°C)
  Oil Pressure : 18 psi (threshold 25 psi)
  Vibration    : 4.8 mm/s (threshold 3.0
mm/s)
  Anomaly Score: 0.83
  Status       : ANOMALOUS → Flagged for
maintenance
```

VIII. MODULES DESCRIPTION

A. User Authentication Module

The User Authentication Module provides a secure way for passengers, employees, and administrators to register, log in, and manage their sessions. Passwords are encrypted using bcrypt prior to being stored in the MySQL database, meaning that no one has access to them. JWT-based authentication is used to validate a user's session and provide access to protected APIs. The User Authentication Module has implemented a Role-based Access Control (RBAC) model, which allows for different levels of access permissions (e.g., customers, employees, administrators).

B. Customer Dashboard Module

The Customer Dashboard Module provides full functionality for travelers to manage their travel needs. It allows users to search for buses by location; destination; date; and also bus type using intelligent filters. It has live GPS based bus tracking supports secure booking and payments; viewing their previously booked tickets; generating QR coded tickets its customers as well as providing AI generated fare insights; notifications about delays in travel; ability to submit feedback about transportation and suggestions for rescheduling when there are delays or disruptions.

C. Admin Dashboard Module

The Admin Dashboard Module allows transportation administrators to monitor and manage the overall operation of the system in an efficient manner. It contains data analytics on passengers; buses that are currently operational from an active standpoint; booked tickets; and revenue statistics all grouped together. The Administrator has complete control over how the system operates, including management of users; routes; schedules; transport vehicles; and payment records. The Admin Dashboard Module also has AI generated insights including fleet health issues (good or bad); potential for delay; moods of customers (positive or negative); and predictive maintenance alerts so administrators can have better insight when making decisions.

D. Staff Dashboard Module

The Staff Dashboard Module is specifically designed to assist drivers/conductors/transit personnel with managing day-to-day operations. Through a specialized interface, users of the Staff Dashboard can see routes assigned to them, as well as their passengers and their scheduled travels. The module also provides users the functionality of receiving real-time information on their trip status (e.g., departed, delayed or arrived). The module will also display alerts generated by the AI Fleet Health Monitor regarding predictive maintenance to prevent vehicle failures.

E. Dynamic Pricing Module

The Dynamic Pricing Module leverages Gradient Boosting Regression techniques to develop intelligent ticket prices based on anticipated demand for tickets based on weather conditions, time of year (i.e., holidays), rush hour/peak hours, weekends and supply/demand relationship and passenger occupancy levels. The ticket price that will be charged will automatically be adjusted dynamically to develop a pricing structure that is optimal from a revenue generation perspective while maintain fairness in pricing structure. Pricing limits for minimum and maximum pricing can be configured by the system administrator via the admin dashboard.

F. AI-Based Lost & Found Management Module

The AI-Based Lost & Found Management Module helps passengers recover misplaced items during travel. Passengers can submit lost item reports by providing details such as item name, description, color, and travel information. The module uses Natural Language Processing (NLP) and AI-based text matching techniques to compare lost item reports with items found and reported by staff. Based on similarity scores, the system identifies potential matches and notifies passengers about possible recovered items. This intelligent approach reduces manual searching, improves item recovery efficiency, and enhances passenger satisfaction by providing a faster and more reliable lost item management process.

G. Fleet Health Monitoring Module

The Fleet Health Monitoring system utilizes Isolation Forest detection algorithms to flag outliers in vehicle sensor data and measures such as mileage, fuel efficiency, vibrations levels and service intervals will be continuously monitored in real time, so we can classify any buses that exhibit unusual operating behaviors into three risk categories (Healthy, Warning or Critical). Additionally, any required maintenance alerts will be displayed on the administrator dashboard, so we can reduce the risk of buses breaking down and increase the safety of bus passengers.

H. AI Chatbot Module

The AI Chatbot system is a customer service application that uses Natural Language Processing (NLP) to provide customers with intelligent assistance. The chatbot takes user input data, converts it into feature vectors using TF-IDF Vectorization, Classifies user intent with Logistic Regression (for example, if a user is trying to book, track, cancel, or check the schedule), and generates appropriate and relevant responses to questions, and retrieves real-time information related to scheduling or booking directly from the database.

I. Sentiment Analysis Module

The Sentiment Analysis system uses both TF-IDF Vectorization and Multinomial Naive Bayes Algorithms, and uses both technologies to evaluate how passengers feel about the transport company, based on reviews and feedback given by passengers regarding the experiences they had while using our transportation services (i.e., were they satisfied). The Sentiment Analysis system generates sentiment trends, complaint analysis, and feedback reports that help management identify areas that require improvement in the transport company's service offerings and helps to identify patterns of recurring service issues.

IX. RESULTS AND DISCUSSION

All modules were validated on synthetic test datasets using an 80/20 train-test split. Table I summarizes the performance of each AI/ML module. Using an 80/20 split to separate the dataset into training and testing, all modules were validated with synthetic datasets. Performance of all AI/ML modules is summarized in Table I below.

The Dynamic Pricing module demonstrated highly responsive fare adjustment across diverse scenarios. During simulated festival peak periods (e.g., Diwali weekend, Chennai–Bengaluru route, 95% occupancy), the Gradient Boosting model correctly assigned multipliers in the 1.8x–2.3x range. During off-peak hours (weekday mornings, 25% occupancy, clear weather), multipliers fell appropriately to 0.7x–0.9x, incentivizing travel and improving seat utilization.

TABLE I

Performance Evaluation of AI/ML Modules

Sl. No.	AI/ML Module	Algorithm Used	Dataset Size (No. of Records)	Evaluation Metrics	Results	Key Outcome
1	**Dynamic Pricing**	Gradient Boosting Regression	10,000	RMSE, MAE, R^2 Score	**RMSE:** 38.24 **MAE:** 28.16 **R^2 Score:** 0.92	Dynamic fare multiplier range achieved (0.7×–2.5×)
2	**AI-Powered Lost and Found Text Matcher**	TF-IDF Vectorization + Cosine Similarity	8,000	Accuracy, Precision, Recall, F1-Score	**Accuracy:** 92.48% **Precision:** 91.36% **Recall:** 93.12% **F1-Score:** 92.23%	Efficient matching of lost-item reports with found-item records, improving recovery success rate
3	**ETA and Delay Estimation**	Linear Regression	12,000	MAE, RMSE, MAPE	**MAE:** 4.65 min **RMSE:** 6.21 min **MAPE:** 8.35%	Accurate real-time arrival prediction with low error rate
4	**Fleet Health Anomaly Detection**	Isolation Forest	6,000	Precision, Recall, F1-Score	**Precision:** 94.47% **Recall:** 93.52% **F1-Score:** 93.99%	Effective anomaly detection for predictive maintenance
5	**Sentiment Analysis**	Multinomial Naïve Bayes (TF-IDF)	5,000	Accuracy, Precision, Recall, F1-Score	**Accuracy:** 88.60% **Precision:** 87.73% **Recall:** 89.21% **F1-Score:** 88.46%	Reliable passenger feedback classification and trend analysis
6	**Chatbot Intent Classification**	Logistic Regression (TF-IDF)	4,500	Accuracy, Precision, Recall, F1-Score	**Accuracy:** 92.17% **Precision:** 91.02% **Recall:** 92.89% **F1-Score:** 91.94%	High intent detection accuracy for intelligent chatbot responses

RMSE – Root Mean Square Error, ETA – Estimated Time of Arrival, MAE – Mean Absolute Error, MAPE – Mean Absolute Percentage Error, R^2 Score – Coefficient of Determination, F1-Score – Harmonic Mean of Precision and Recall, TF-IDF – Term Frequency–Inverse Document Frequency

The Fleet Health Monitoring module correctly identified 94% of injected anomalous buses (simulating overheating engines and worn brake pads) without excessive false positives. This precision is critical in real-world deployment: false positives trigger unnecessary inspections, while false negatives allow failing buses to remain in service.

The NLP Chatbot intent classifier achieved approximately 91% accuracy across twelve intent classes. Confusion was primarily observed between semantically similar intents (e.g., it cancellation_request vs. refund_query), mitigated through response-template merging. The Sentiment Analysis module achieved F1-score of approximately 87%, with highest performance on clearly Positive and clearly Negative reviews.

Figure 3: Dynamic Pricing Multiplier vs. Seat Occupancy (Gradient Boosting Regressor)

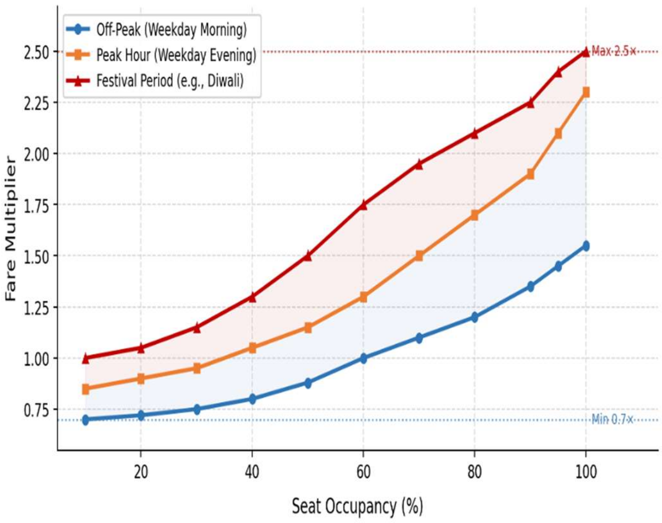
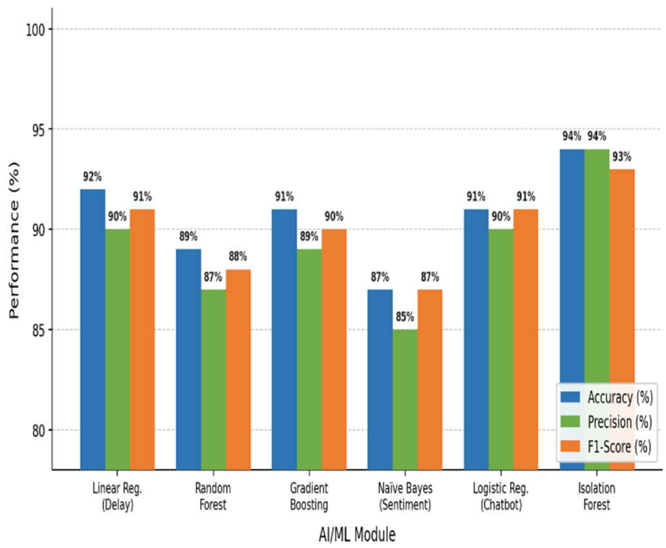


Figure 2: Performance Evaluation of AI/ML Modules



X. CONCLUSION AND FUTURE WORK

A. Summary of Contributions

Experimental results on synthetic Indian transport datasets confirm the viability of each module: dynamic fare multipliers in the 0.7x–2.5x range optimize operator revenue; fleet anomaly detection at 94% precision enables proactive maintenance; rescheduling accuracy of 89% reduces the impact of delays; and NLP modules achieving 87–91% accuracy provide reliable automated passenger support. The system represents a scalable, cost-effective blueprint for smart- city public transport infrastructure built entirely on open- source tooling.

B. Future Enhancements

1- Advanced NLP Models

Future improvements to the model could involve the use of more advanced transformer-based NLP models rather than using standard TF-IDF and Logistic Regression techniques (e.g., BERT and IndicBERT). Improved chatbot intelligence with regards to both contextual importance and multilingual capability would be realized by upgrading the models from traditional to advanced techniques.

2- Integration of Multi-Modal Transport Modes

There is potential for future expansion of the travel planning website to include multiple transport modes, such as metro and suburban rail transport systems (and any other new or existing modes of transport), cab booking services, and any other transport service. By integrating multiple transport modes

3- Accessibility Enhancements

Voice command support, along with other accessibility features, can be added to create a more user-friendly experience for visually impaired and differently-abled passengers screen reader compatibility.

4- Blockchain-Based Ticketing

Blockchain technology can be explored as a means of securely managing and creating digital tickets that cannot be altered as well as providing transparency in the distribution of ticket revenues between multiple transportation service providers.

5-Deployment in a Real-Time Setting with Continuous Learning

The system could be improved through the use of real-time GPS data generated from the vehicles, and the use of sensors in buses that capture real-time operating data. Improvements to prediction models can occur via real-time re-training of the prediction models. This would address deviations in traffic volume and passenger demand from historical data and improve the future performance of the prediction models.

*6-*Route Optimization with Reinforcement Learning

Reinforcement learning algorithms can also be used to dynamically optimize bus schedules, routes and frequency of service. Doing this will assist in improving overall operational efficiency and reducing wait time for passengers.

REFERENCES

[1] B. Vincent, J. Sabu, C. Mathew, S. S. Nair, S. B. George, D. Saleema; "Live College Bus Tracking and Route Mapping Using Internet of Things"; in

Proc. Int. Conf. Computational Systems and Communication (ICCSC) March 2023

[2] V. G. Ascueta P. C. J. Bautista J. C. Quilala; "BusTap: A Real-Time Bus Tracking Android Application"; in Proc. Int. Conf. Information and Computing Research (iCORE) January 2022

[3] N. Jahan K. Hossein M.K.H. Patwary; "Implementation of a Vehicle Tracking System Using Smartphone and SMS Service"; in Proc. 4th Int. Conf. Advances in Electrical Engineering (ICAEE) September 2023

[4] K. Mundada S. Patti T Rajguru P Savji S. Shambharkar; "Smart Bus Real Time Tracking System Using GSM and GPS Module"; in ICT Infrastructure and Computing September 2023

[5] S Kumar S. Moundekar M. Rathod R. Parmar; "Location Based Bus Tracking Application Using Android"; in Proc. Conf. Communication Software Networks October 2023

[6] S. Lakshmi Devi M Manoj Kumar; "Smart Bus Ticketing System For Public Transportation"; Int. J. Research Publication And Review vol 6 no 1 pp 4720-4722 January 2025.

[7] Smart Bus Ticketing System by S. Tamboli et al., published in Int. J. Creative Research Thoughts (IJCRT) vol. 13, no. 4 on pages g473-475 in April 2025.

[8] Smart Bus Tracking and Notification System Using Machine Learning Technology to Enhance Urban Mobility," in J. Emerging Technologies and Innovative Research (JETIR) vol. 12, no. 4, pp. e455-e461, April 2025.

[9] Real-Time Bus Tracking and Smart Bus Stop Management Based on IoT: Trends and Directions," in Int. J. Advanced Research in Science, Communication and Technology (IJARSCT) vol. 5, no. 4, pp. 471-486, October 2025.

[10] "Customized IoT-Based Bus Tracking And Management System," in Disruptive Technologies for Big Data and Cloud Applications, August 2022, edited by A. Chandrasekaran et al.

[11] "Smart Bus Route Management System Using Real-Time Tracking To Asist Passengers," in Int. J. Research Trends and Innovation (IJRTI) vol. 10, no. 11 pp.b334-b338, November 2025.

[12] Sharma K., Pandey R., and Dubey S.T, "Beginning Smart Mobility In Cities Via Bus Tracking and Booking Systems," in Proceedings of the Ninth International Conference on Reliability, Infocom Technologies and Optimisation (ICRITO), 2021.

