



# Machine Learning-Based Intelligent Taxi System Analysis for the Tourism Sector

K Naresh<sup>1</sup>, Reddypalli Charitha Sai<sup>2</sup>

<sup>1</sup>Assistant Professor, Department of MCA, Annamacharya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India.

<sup>2</sup>Postgraduate, Department of MCA, Annamacharya Institute of Technology and Sciences, Tirupati, Andhra Pradesh, India.

## How to Cite this Article:

Sai, R. C. (2026). Machine Learning-Based Intelligent Taxi System Analysis for the Tourism Sector. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04).

<https://doi.org/10.55041/ijcope.v2i4.068>

## License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i4.068>

## Abstract

A machine learning-based analytical system for researching cab transportation services in the Indian tourism industry is presented in this paper. Taxi services are essential for promoting tourism since they give passengers easy and adaptable transportation. However, the amount of transportation data created in urban contexts makes it difficult to analyse taxi usage trends, service effectiveness, and transportation demand. This paper suggests a data-driven approach that combines machine learning methods with an online application for examining transportation trends and taxi system performance in order to solve this problem.

In order to find trends in transportation utilisation, demand fluctuations, and operational efficiency, the suggested system gathers and analyses taxi-related data and uses machine learning algorithms. The web framework used in the system's implementation allows administrators and users to engage with the platform for analytical tasks and taxi service monitoring. Before the machine learning model is trained, the dataset is cleaned and arranged using data preprocessing techniques. In order to improve taxi services in tourist locations, the trained model aids in the analysis of transportation behaviour and produces valuable insights.

The suggested method can successfully analyse taxi traffic patterns and assist in decision-making for tourism transportation management, according to

experimental assessment. The outcomes show how machine learning approaches can enhance transportation analysis and maximise taxi service management. The created framework can help tourism planners and transportation authorities improve service quality and operational efficiency by offering a scalable and intelligent method for analysing taxi systems.

## Keywords

Machine Learning, Taxi Transportation System, Tourism Analytics, Data Analysis, Transportation Management, Intelligent Transport Systems



## I. Introduction

One of the key industries supporting the economic growth of many nations, including India, is tourism. By making it simple for visitors to move between tourist attractions, lodging facilities, airports, and other sites of interest, efficient transportation networks are essential to promoting tourism-related activities. Taxis are regarded as one of the most practical and adaptable forms of transportation for travellers in urban and tourist areas. By offering on-demand mobility, taxi services let passengers to go where they're going without depending on set routes or schedules. In many Indian cities, the need for dependable taxi services has grown dramatically due to the country's growing tourism and urban population.

Traditional taxi services have evolved into more sophisticated and intelligent transportation systems in recent years due to the quick growth of digital technologies. The efficiency and accessibility of taxi services have been enhanced by digital payment methods, mobile applications, and online booking platforms. However, the increasing amount of data on taxi utilisation poses significant difficulties for comprehending service demand, transportation trends, and operational effectiveness. It is challenging and time-consuming to manually analyse this vast volume of transportation data. To gain valuable insights from taxi service data and enhance transportation management in the tourism industry, sophisticated data analysis techniques are therefore necessary.

In order to analyse big datasets and find hidden patterns in intricate systems, machine learning has become a potent tool. Data from taxi services may be analysed, transportation trends can be predicted, and system performance can be assessed by using machine learning techniques. Large amounts of travel data may be processed by machine learning models, which can then offer insightful information to help tourism planners and transport authorities make wise decisions. Peak travel periods, taxi demand forecasting, service distribution optimisation, and overall transportation efficiency can all be enhanced with the use of these technologies. The combination of machine learning and taxi system analysis can be quite advantageous when it comes to tourism transportation.

It enables researchers and system developers to better manage taxi services in well-known tourist locations, comprehend tourist travel patterns, and examine transportation behaviour. Furthermore, intelligent

analytical systems can facilitate improved tourism infrastructure planning, lower transit delays, and increase passenger satisfaction. The goal of this project is to create an analytical system based on machine learning for researching cab transportation systems utilised in the tourism industry. The goal of the suggested system is to examine data pertaining to taxis and spot trends that can help improve transportation administration. The machine learning model is integrated with an interactive interface on a web-based platform that enables consumers to analyse data and track the effectiveness of taxi services.

The system produces analytical insights that can help enhance cab transportation services by utilising machine learning algorithms and data preparation techniques. This study's primary objective is to show how machine learning methods may be successfully used to examine taxi transportation networks in tourist settings. The suggested strategy advances the creation of intelligent transportation analytic tools that can enhance service effectiveness and facilitate data-driven decision-making in the management of tourism transportation.

## II. Objectives of the Study

Developing an intelligent analytical system for researching cab transportation services utilised in the tourism industry is the primary goal of this project. The demand for effective solutions that can evaluate transportation data and offer helpful insights for enhancing taxi services is rising due to the quick expansion of tourist and urban transportation. The suggested system analyzes cab transportation data and facilitates improved transportation management by combining machine learning techniques with a web-based platform. Analysing taxi service data to find trends in transportation demand, usage patterns, and operational efficiency is one of the study's main goals.

The technology uses machine learning techniques to extract useful information from transportation datasets that can help comprehend how tourists and travellers use taxi services. Designing and implementing a web-based system that enables administrators and users to communicate with the taxi analysis platform is another goal of this project. The system offers features for processing transportation data, managing taxi-related information, and carrying out machine learning-based analysis. The analytical process is made more accessible and effective by combining machine learning with an intuitive web interface.





procedures like tokenisation and the elimination of frequent stop words are carried out. The machine learning model's precision and effectiveness are enhanced by these preprocessing procedures.

### Feature Extraction

The cleaned text must next be transformed into numerical features that machine learning algorithms can understand once the textual input has been preprocessed. Feature extraction techniques are used to represent the text input in numerical form because machine learning models are unable to interpret raw text directly. The review text is converted into feature vectors in this system using vectorisation techniques. The model is able to examine the frequency and significance of words that appear in the reviews thanks to this procedure. The algorithm can successfully recognise patterns linked to both positive and negative attitudes by transforming the text into structured numerical data.

### Machine Learning Model

The suggested solution uses machine learning methods to categorise reviews of taxi services according to their emotion. The model is trained using labelled data, which shows whether a review is good or negative, using a supervised learning approach. The system discovers patterns and connections between the sentiment labels and the retrieved characteristics during the training phase. After the training process is finished, the trained model is able to categorise fresh reviews and identify whether the input represents a favourable or unfavourable customer experience. This categorisation aids in comprehending the general level of client satisfaction with taxi services.

### Model Training and Evaluation

The dataset is split into training and testing sets once the machine learning model has been chosen. While the testing dataset is used to assess the model's performance, the training dataset teaches the model how to identify trends in customer evaluations. The model's ability to predict the sentiment of taxi service reviews is evaluated using performance indicators including accuracy and classification outcomes. A thorough assessment guarantees that the model can accurately analyse consumer feedback and produce significant outcomes.

### V. System Implementation

The suggested approach is put into practice as a web application that combines an intuitive user interface with machine learning analysis. The system's backend

processes the input data, preprocesses the text, and uses the machine learning model that has been developed to forecast the review's sentiment. Users can submit reviews of taxi services using the frontend interface and receive the anticipated sentiment outcome. Users can interact with the system and analyse client comments on taxi transportation services with ease thanks to the integration of machine learning with a web platform.

### VI. Experimental Results and Performance Analysis

Several machine learning techniques were used and evaluated in order to assess the efficacy of the suggested taxi review sentiment analysis system. Logistic Regression, Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and Stochastic Gradient Descent (SGD) are among the models employed in this investigation. The preprocessed taxi review dataset was used to train these algorithms, and their performance was evaluated using a different testing dataset.

DETAILED CLASSIFICATION REPORT			
LOGISTIC REGRESSION	ACCURACY	PRECISION	RECALL
RANDOM FOREST	ACCURACY	PRECISION	RECALL
DECISION TREE	ACCURACY	PRECISION	RECALL
NAIVE BAYES	ACCURACY	PRECISION	RECALL
SVM	ACCURACY	PRECISION	RECALL
KNN	ACCURACY	PRECISION	RECALL
STOCHASTIC GRADIENT DESCENT (SGD)	ACCURACY	PRECISION	RECALL

Fig: Results comparison

Standard classification criteria including accuracy, precision, recall, and F1-score were used to assess each model. While recall gauges the model's capacity to find all pertinent positive occurrences, precision quantifies the number of projected positive instances that are really accurate. The F1-score offers a fair metric that incorporates recall and precision. The model's total accuracy in predicting the emotion of taxi service reviews is known as accuracy.

The experimental findings show that when applied to the taxi review dataset, several machine learning models yield differing performance levels. In terms of precision and recall, logistic regression performs steadily and evenly. Decision Tree provides moderate classification capability but may suffer from overfitting in certain cases. By integrating several decision trees, Random Forest increases prediction reliability and produces more solid forecasts. Naïve Bayes's probabilistic methodology makes it effective for text categorisation jobs. By determining the best judgement boundaries between



positive and negative attitudes, Support Vector Machine (SVM) also offers dependable classification.

Models like SVM and Logistic Regression provide relatively superior classification performance for the specified dataset among the assessed algorithms. The findings show that machine learning methods may successfully analyse reviews of taxi services and spot trends in customer sentiment. The best algorithm for sentiment analysis of taxi transportation reviews in tourism-related contexts is found through this comparison investigation. According to the experimental research, sentiment categorisation based on machine learning can be a useful technique for assessing consumer satisfaction in cab transportation systems. The suggested system can offer useful insights that assist tourism planners and taxi service providers in addressing consumer complaints and improving service quality by evaluating passenger reviews.

## VII. Conclusion

This study offered a machine learning-based method for examining Indian taxi service user feedback. The suggested approach divides reviews of taxi services into positive and negative categories using sentiment analysis techniques. To assess their effectiveness, a number of machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, Naïve Bayes, SVM, KNN, and SGD, were put into practice. The outcomes of the experiment demonstrated that machine learning models are capable of efficiently analysing textual feedback and spotting trends in customer satisfaction. Users may quickly analyse taxi service evaluations and get sentiment forecasts thanks to the built web-based system. All things considered, the suggested approach shows how machine learning methods can help with improved analysis and enhancement of cab transportation services in the travel industry.

## References

- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113.
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.
- Joachims, T. (1998). Text categorization with Support Vector Machines: Learning with many relevant features. In *Proceedings of the European Conference on Machine Learning*, 137–142.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.