International Journal of Computer Science and Engineering Development
Volume 1, Issue 5 | November - December 2026 | www.ijcsed.com
ISSN: 3139-

# Traffic Volume Prediction Based on Weather Conditions Using Random

Gayathri N #1, Nayana Rao S #2, Raju Bomma*3, Shamanth S*4, Shashank B*5, Y Surendra*6
#Assistant Professor, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India
*Undergraduate Scholar, Department of Computer Science and Engineering, City Engineering College, Bengaluru, Karnataka, India
1gayathrin@cityengineeringcollege.ac.in, 2nayana_rao@cityengineeringcollege.ac.in, 3rajubomma07@gmail.com, 4shamanthgooli2005@gmail.com,
5shashank9051yaragal@gmail.com, 6surendrasuri63633@gmail.com

## Abstract
Traffic volume forecasting allows for efficient traffic management, road planning, and reduced traffic congestion. Weather factors such as temperature, rainfall, and visibility can affect traffic patterns. This research provides an approach to traffic volume prediction using a machine learning method called Random Forest Regression. This system uses traffic and weather data. Once collected the data is preprocessed to extract the features. This model uses weather factors available, and other relevant traffic features to predict traffic volume. Random Forest Regression was chosen due to it combining many decision trees to make predictions. This can provide accurate complex predictions. The model will be evaluated using appropriate regression metrics. These include MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and R2 (coefficient of determination). The results from the experiments will show how accurate the model is, this will determine if the model could be useful. This approach aims to help manage traffic using data. The relationship between weather conditions and traffic volume will be investigated.

Keywords: Traffic Volume Prediction, Weather Conditions, Random Forest Regression, Machine Learning, Traffic Management, Regression Analysis.

## 1. INTRODUCTION
Traffic prediction is an important element within intelligent transportation systems and urban traffic. As cities become more populated and there are more vehicles on the road, traffic congestion has become an issue, increasing travel times, fuel usage, and causing pollution. Traffic volume can be affected by many different factors such as weather, time of day, day of the week, and location of roads. Factors such as rain, temperature, humidity, and visibility can affect how drivers behave whilst on the road. By taking into account historical data alongside weather information, we can make better traffic predictions.
Historically, traffic prediction would have been made based on previous traffic data. However, factors can influence traffic and it can be quite complex. Using machine learning techniques we can extract patterns from the historical data to predict future traffic volumes. Random Forest Regression is an ensemble learning algorithm which uses multiple decision trees to create a numerical output. This algorithm is a good choice for this project as it can predict non-linear relationships. However, its accuracy will depend on the quality of the dataset, input variables chosen, and how well the model has been trained.

## 2. PROBLEM STATEMENT
Accurately predicting traffic volume based on weather conditions is challenging because traffic patterns are influenced by multiple factors, including temperature, rainfall, humidity, visibility, wind speed, time of day, and road conditions. Traditional traffic prediction methods may not effectively capture the complex relationships between weather conditions and traffic volume, which can lead to inaccurate predictions. This paper examines a Random Forest-based machine learning approach that uses historical traffic and weather data to predict traffic volume. The intended goal is to provide preliminary traffic volume predictions to support traffic monitoring, congestion analysis, and better transportation.

OBJECTIVES
- To study the application of machine learning in traffic volume prediction based on weather conditions.
- To develop a Random Forest-based model for predicting traffic volume.
- To analyse the impact of weather conditions such as temperature, rainfall, humidity, and visibility on traffic volume.
- To evaluate the prediction model using suitable performance metrics such as MAE, RMSE, and $R^2$.
- To examine the advantages, limitations, and possible future improvements of the proposed approach.

## LITERATURE SURVEY

### 2.1 CNN-Based Skin Cancer Classification
Previous studies have investigated the application of machine learning techniques for predicting traffic volume using historical traffic data and environmental conditions. These studies highlight the importance of

factors such as temperature, rainfall, humidity, visibility, and time of day in understanding traffic patterns. Machine learning algorithms can identify relationships between these factors and traffic volume, helping improve prediction accuracy. Random Forest is particularly useful because it combines multiple decision trees to produce predictions and can capture complex relationships within traffic datasets.

### 2.2 Transfer Learning for Skin-Lesion Analysis

Several research works have explored the use of Random Forest and other machine learning algorithms for traffic volume forecasting. Random Forest can handle multiple input features, reduce overfitting through ensemble learning, and identify important variables influencing traffic conditions. Weather-related parameters, historical traffic volume, road information, and time-based features can be used as inputs to the prediction model. The effectiveness of these approaches depends on dataset quality, feature selection, preprocessing, and evaluation methods. Performance is commonly evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination ($R^2$).

## 3. DATASET DESCRIPTION

A traffic volume prediction system requires a dataset containing historical traffic records and corresponding weather information. The proposed study uses traffic and weather-related parameters to analyse traffic patterns and predict vehicle volume under different environmental conditions. The dataset may include features such as temperature, rainfall, humidity, visibility, wind speed, time of day, day of the week, road ID, and previous traffic volume.

## 4. PROPOSED METHODOLOGY

The proposed system uses a Random Forest machine learning algorithm to predict traffic volume based on weather conditions and other traffic-related parameters. The methodology involves collecting traffic and weather data, preprocessing the dataset, selecting relevant features, training the machine learning model, and evaluating its performance.

The proposed methodology consists of the following steps:

1. **Data Collection:** Collect historical traffic volume and weather-related data, including temperature, rainfall, humidity, visibility, and wind speed.

2. **Data Preprocessing:** Clean the dataset by handling missing values, removing duplicate records, and correcting inconsistent data.

3. **Feature Selection:** Identify important input features such as weather conditions, time of day, day of the week, road information, and previous traffic volume.

4. **Model Training:** Apply the Random Forest algorithm to the processed dataset to learn relationships between weather conditions and traffic volume.

5. **Traffic Volume Prediction:** Use the trained model to predict traffic volume based on user-provided weather and traffic parameters.

6. **Model Evaluation:** Evaluate the model using MAE, RMSE, and $R^2$ score to measure prediction performance.

7. **Web Application Integration:** Integrate the trained model with a web-based dashboard to display predicted traffic volume, traffic conditions, route information, and estimated travel time.

The overall methodology aims to provide reliable traffic volume predictions that can support traffic monitoring, congestion analysis, and transportation planning.

## 5. SYSTEM DESIGN

The proposed system is designed to predict traffic volume based on weather conditions using the Random Forest machine learning algorithm. The system consists of different modules that work together to collect input data, process the information, generate predictions, and display the results through a web-based dashboard.

The system follows a client-server architecture consisting of the following components:

1. **User Interface:** A React.js-based web interface allows users to enter travel details, weather conditions, and location information.

2. **Backend Server:** Node.js and Express.js handle user requests and provide communication between the frontend and machine learning model through REST APIs.

3. **Data Processing Module:** Python and Pandas are used to preprocess input data, handle missing values, and prepare features required by the prediction model.

4. **Machine Learning Model:** The Random Forest Regressor analyses the input parameters and predicts the expected traffic volume.

5. **Database Module:** MongoDB stores relevant application information and prediction records.

6. **Result Visualization:** The dashboard displays predicted traffic volume, traffic conditions, weather information, route distance, and estimated travel time.

## 6. PERFORMANCE EVALUATION

Random Forest model is evaluated by comparing predicted traffic volume with actual traffic volume. The performance is measured using the following metrics:

- **MAE:** Measures the average prediction error.
- **RMSE:** Measures prediction errors and penalizes larger errors.
- **$R^2$ Score:** Measures how well the model predicts traffic volume.
- **Feature Importance:** Identifies the most influential weather and traffic parameters.

## 7. RESULTS AND DISCUSSION

The following values are reproduced from the literature summary supplied for this paper. They are not results

ISSN: 3139-

from a new experiment.

| ML Model | Reported accuracy |
|---|---|
| Decision Tree | 85% |
| Linear Regression | 85% |
| K-Nearest Neighbors (KNN) | 87% |
| Random Forest | 90% |
| Gradient Boosting | 91% |

*Table 1. Accuracy values reported in the supplied literature summary.*

Among the selected machine learning models, Random Forest is the proposed algorithm for traffic volume prediction. Its performance will be compared with other models using the same dataset and evaluation metrics. The results may vary depending on dataset quality, feature selection, preprocessing, and training parameters. The final model performance will be determined using actual test-set evaluation results rather than assumed values.

## 8. ADVANTAGES

- CNNs can learn image features automatically from labelled examples.
- A trained model can apply a consistent classification procedure to new images.
- Transfer learning can reduce training time and data requirements compared with training from scratch.
- Multiple metrics can be used to examine different aspects of model performance.
- A user interface can make preliminary image analysis easier to access.

## 9. LIMITATIONS

- Predicts traffic volume using weather conditions and historical traffic data.
- Random Forest provides reliable predictions by combining multiple decision trees.
- Identifies important factors affecting traffic volume.
- Helps analyse and monitor traffic congestion.
- Supports better traffic management and transportation planning.
- Provides an interactive web-based dashboard for displaying prediction results.

## FUTURE SCOPE

The proposed system can be further improved by integrating real-time traffic and weather data to provide more accurate and updated traffic volume predictions. Advanced machine learning and deep learning algorithms can be explored to improve prediction performance. The system can also be extended to cover multiple roads, locations, and cities. In the future, a mobile application can be developed to provide easy access to traffic information. Integration with smart-city infrastructure, GPS, and real-time traffic monitoring systems can further support efficient traffic management, reduce congestion, and improve transportation planning.

## 10. CONCLUSION

The proposed system focuses on predicting traffic volume based on weather conditions using the Random Forest machine learning algorithm. By analysing factors such as temperature, rainfall, humidity, visibility, and historical traffic data, the system aims to provide useful traffic volume predictions. The integration of machine learning with a web-based dashboard helps users understand traffic patterns and monitor congestion. The system can support better traffic management and transportation planning. In the future, incorporating real-time data and advanced prediction techniques can further improve the effectiveness and reliability of the proposed approach.

## REFERENCES

[1] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324.

DOI

[2] “Traffic Volume Prediction Based on Weather Parameters,” *2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE)*, 2024. DOI: 10.1109/ICWITE59797.2024.10503349.

IEEE Xplore

[3] Scikit-learn, “RandomForestRegressor Documentation.” Available: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html.

scikit-learn 1.9.1 documentation

[4] Scikit-learn, “Machine Learning in Python.” Available: https://scikit-learn.org/

[5] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 201

ISSN: **3139-**