INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 31-05-2022

Accepted: 01-09-2022

Published: 29-10-2022

Citation: Jenifer J, Priyadarsini RJ (2022) Improved Mayfly Optimization and LightGBM Classifier for Smart City Traffic Prediction. Indian Journal of Science and Technology 15(40): 2085-2092. https://doi.org/10.17485/IJST/v15i40.1155

* **Corresponding author**.

jenifer.john3@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Improved Mayfly Optimization and LightGBM Classifier for Smart City Traffic Prediction

J Jenifer[1]*, R Jemima Priyadarsini[2]

1 Research Scholar, Department of Computer Science, Bishop Heber College, Affiliated with Bharathidasan University, Trichy-17, Tamilnadu, India
2 Associate Professor, Department of Computer Science, Bishop Heber College, Affiliated with Bharathidasan University, Trichy-17, Tamilnadu, India

Abstract

Objectives: This research work focuses on predicting traffic for the Smart City. **Methods:** Current research methods for traffic prediction are based on machine learning (ML) model. This article presents two contributions related to it. First, it provides feature engineering that includes feature extraction and a nature inspired optimization algorithm for selecting the best features. The mayfly optimization algorithm is improved by using the mode-based ranking method to select the best feature. Second, it uses the light-weight boosting method to train the datasets for better accuracy.**Findings:** The proposed Improved MayFly Optimization with LightGBM (IMFO-LGBM) is experimented with popular smart city datasets which is available in Kaggle website. IMFO-LGBM shows an improvement in the prediction accuracy when compared with the baseline methods. It shows 2% of increase in the overall accuracy. **Novelty:**Nature inspired Mayfly optimization is improved and used to find the best feature for prediction. The selected features are then trained with the light weight boosting algorithm (i.e., Light Gradient Boosting Model). The hybrid of improved mayfly optimization and light GBM outperformed well.

Keywords: IoT; Smartcity; Mayfly optimization; Machine learning and LightGBM

1 Introduction

A new congestion-based traffic prediction model (CTPM)[1] is proposed to predict traffic more accurately. The proposed CTPM used the XGBoost decision tree classifier model for traffic congestion prediction. The work was compared with Long Short-Term Memory (LSTM) neural network and the refinement results showed an improvement of 9.76% accuracy.

In addition to the traffic congestion prediction, a genetic algorithm based[2] optimization model was proposed to reduce traffic in Amman, a capital city of Jordan. The research work used Human Community based Genetic Algorithm (HCBGA) with several parameters such as speed performance index, speed reduction index, road congestion index, and congestion period to forecast the traffic in the city. Machine learning plays a key role in forecasting traffic with the historical data in the recent research article[3], deep Extreme Learning Machine (ELM) was proposed to predict the traffic.

Hybrid Multimodal Deep Learning System (HMDL)[4] was proposedtoward estimating traffic flow. This model integrates CNN with GRU dives and detects deep non-linear connection credits, combining cadet repeat units (GRUs) and single layer CNN to accomplish the highlights of the relationship between floats and any modular transport information level. Multimodal input information. The beginning of the completion of multimedia collaborations was the introduction of the Traffic Continuous Information Handling System with locally located components, long-lasting reliability features and spatial-global connections. CNN-GRU solves traffic stream prediction problems by detecting long-term transient conditions and locating components of spatial-quick connection to determine the time-to-time correlation of speed stream excursions. The number of vehicles at the advertising point is expected to be re-evaluated. Determining time series acceleration, coordinate coordinates, and spatial and transient dependence of multimedia input information are used to advantage. Gathering strong traffic information in the short term is a barrier. Data collected from the interstate areas of the UK include traffic flow, speed and traffic casualties and when they are fully confronted in extreme weather conditions.

The goal of authors was to develop a radical model for predicting traffic flow in light of the elements that make up secret fragments of knowledge in automobiles[5]. The creator of Help Vector Relapse uses mapping information using non-linear planning. The average square error approach evaluates the exhibition by measuring the squares of the regular error. Raises the straight regeneration model in measuring the correlation between scalar reaction and free factors. Selection tree learning calculation calculates entropy or data gain. The biggest advantage is that it works effectively in differentiating the best model for accessible open information. This model is exclusive; No longer is it tiring to constantly evacuate traffic jams in conjunction with existing offices.

Bang and Lee[6] coordinate the revolts and the expected situation at the head of each movement or prevent accidental access between vehicles. The vector-based portability expectation model in the TDMA-based VANET prevents downtime by allocating time allocations and displaying the versatility of public vehicles using control table opening, vehicle ID, and home data of the vehicle combination trend. Hopping data, and the scope and longitude of a vehicle. Addition of the calculation function, the street around the street where the rigidity of the traffic is high is modified and the movements recreate the bearings for more compact and continuous movement. Access and integration accidents due to vehicle composition design and level of traffic.

Wei et al.[7] focus on improving forecast accuracy as traffic progresses. Auto-encode accepts the internal contract of the stream in the Rush Hour Gridlock by deleting the attributes of the stream information in the Rush Hour Gridlock stream, using the Auto-encode Long Short-Term Memory (AE-LSTM) approach. The LSDM Network separates executable trademark information and predictable and verifiable information from the luxurious live information in the Rush Hour gridlock stream. This approach is challenging to implement and has a level-headed objective. It additionally adorns the exhibition appreciated in forecasting traffic flow. Whatever it is, the strike considers only time examples and direct spatial examples.

Adaptive Traffic management system (ATM)[8] based on ML was proposed to reduce the road accidents. The data was collected from the IoT devices embedded in the automatic vehicles. Vehicle, infrastructure and events are the main parameters used to design the ATM. Clustering based machine learning was used to detect accidental anomaly. In a recent study, Random Forest was used to improve the LightGBM model[9] to forecast and fault classification of smart meters. Artificial Neural Network (ANN) and Support Vector Machine (SVM) based traffic congestion system[10] was proposed for the vehicular networks. The proposed work used ANN and SVM to build a fuzzy logic for traffic prediction. Various optimization methods such as ant colony optimization, genetic algorithm, particle swarm optimization, artificial bee colony optimization and differential evolution are experimented with the machine learning models to forecast the traffic prediction[11].In a research work[12] , researchers haveproposed aConvolutional Neural Network (CNN) technique for traffic monitoring system. The authors have collected datasets using drones. The captured images are pre-processed and the proposed system produced 91.67%.

From the above study, the application of machine learning has attracted more attention to predict traffic in smart cities, with less attention focused on the use of nature inspired optimization schemes. To help with this, the present research work provides hybrid of nature optimization method with the machine learning model to forecast the traffic in smart-city. The interest of this paper lies in proposing a more weight-less approach while being exceptionally accurate at present.

Key contributions of this studyis listed as follows:

- To apply nature inspired optimization algorithm to reduce the high-dimensions in the time-series data
- To improve the may-fly optimization for attaining better accuracy than the baseline methods
- To hybrid nature inspired may-fly optimization algorithm with Light weight Gradient Boosting machine learning model to predict the traffic in smart-cities.

2 Materials and Methods

2.1 IMFO-LGBM Architecture:

This section discusses the proposed Improved Mayfly Optimization Algorithm with Light Gradient Boosting Machine (IMFO-LGBM) for IoT data prediction. The proposed IMFO-LGBM research work has the following components as depicted in the Figure 1. The collected IoT data is load to the preprocessing phase. The preprocessing phase contains a 'feature engineering' module to generate meaningful data. As the data is timeseries in nature, this work uses the 'feature extraction' (i.e., a submodule in this proposed system) technique to extract a greater number of features. The dataset [13] contains 4 features namely date, count of vehicles, instance ID and junction ID. The date feature is then extracted into 11 features. More features may cause computational sparsity. Feature selection is a best method to reduce the dimension of datasets and produce better accuracy with help of relevant features. In the proposed system, a nature inspired optimization algorithm namely mayfly optimization is used to select the appropriate features. Improved mayfly optimization is used for better results. The selected information is passed to the classifier model which uses the Light Gradient Boosting Machine (LGBM). The following section describes the proposed algorithm.

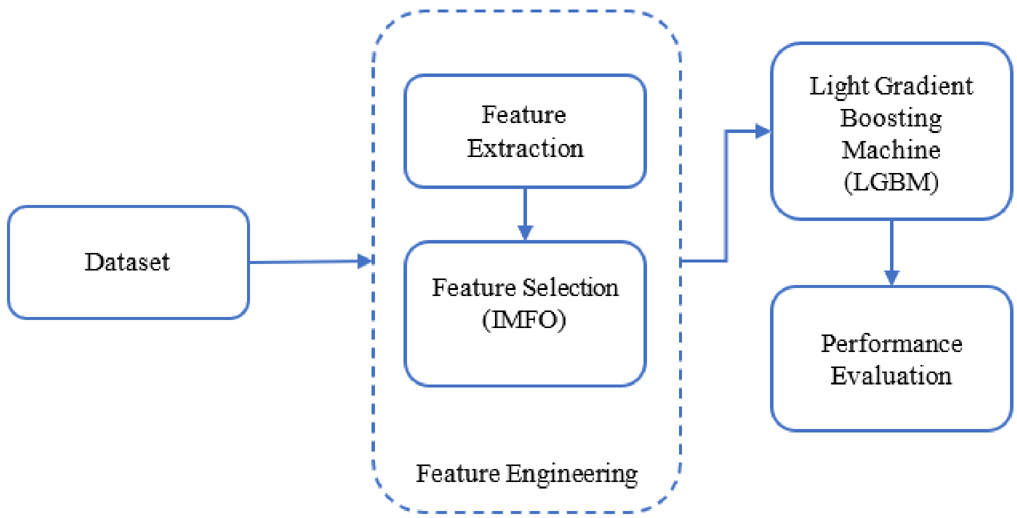

Fig 1. Research Methodology

2.2 Improved Mayfly Optimization:

In the year 2020, Zervoudakis and Tsafarakis [14] proposed the Mayfly Optimization Algorithm (MFOA) based on the social behaviour of the mayfly mating process. In general, female mayfly (FMF) chooses the best male mayfly (MMF) using the potential of best position in the search space. This concept produces the best solution. In recent years, MFOA has played a major role in the feature selection process of high dimensionality data. Among several intelligent optimization algorithms, the basic MFOA produced good accuracy when compared with the nature inspired optimization algorithms such as swarm intelligence algorithms and Ant Colony Optimization (ACO). Even though, MFOA faced lagging in the search process. To overcome the above-mentioned issue, Guo et. al. [15] proposed an improved mayfly optimization algorithm (IMFOA) using the median position parameter and the proposed IMFOA outperformed well than the following nature inspired algorithms: Particle Swarm Optimization (PSO), Salvia Swarm Algorithm (SSA), Whale Optimization Algorithm (WOA), Genetic Algorithm (GA), and Firefly Algorithm (FA).

As discussed earlier, the mayfly algorithm has a lot of similarities like the PSO algorithm in terms of updating the position and velocity. Hence, the movement of mayfly is based on the social sharing is as same as the biological information to create the advantage. The social sharing contains the cognitive measures and memory of individual mayfly. To obtain the best solution of the MFO algorithm, the individual memory of mayfly provides the best position whereas the cognitive measure produces the best position of the mayfly in the total population.

To obtain the individual best position of the male mayfly, the current position of each male mayfly is calculated. The position is calculated based on the velocity of the male mayfly. Equation 1 gives the current position of mayfly:

$$mmf_i^{t+1} = mmf_i^t + vel_i^{t+1} \quad (1)$$

where mmf denotes male mayfly, t is search space at time, i is the position of mmf and vel is the velocity of the iteration.

The velocity of mmf is updated based on the speed. Usually, the mmf dances to attract the female mayfly (fmf) above the surface of the water. While dancing, the velocity of mmf will increase constantly based on the movement. The velocity will be updated by using the following equation (2)

$$vel_{ij}^{t+1} = vel_{ij}^t + a_1 e^{-\beta r_{pos}^2}\left(posBest_{ij} - mmf_{ij}^t\right) + a_2 e^{-\beta r_{gbl}^2}\left(gblBest_j - mmf_{ij}^t\right) \quad (2)$$

where vel_{ij}^t and mmf_{ij}^t denotes the velocity and position of i^{th}mmf in j^{th} dimension at the time t respectively. In the equation (2), e is the exponential constant with $'\beta'$ coefficient of mmf. This 'e' controls the perceptibility of mmf. The r_{pos} and r_{gbl} is the Cartesian distance between the individual position and global best position using the equation (3)

$$(mmf_i - MMF_i| = \sqrt{\sum_{j=1}^{n}\left(mmf_{ij} - MMF_{ij}^2\right)} \quad (3)$$

where MMF_i corresponds to the best position of mmf_i or global best position of mmf_i

The female mayflies are not having congregation behaviour as like the male mayfly. The best female mayfly attracts the best male mayfly to reproduce. As like equation (1), the velocity of female mayfly will be calculated using equation (4)

$$fmf_i^{t+1} = fmf_i^t + vel_i^{t+1} \quad (4)$$

where fmf denotes female mayfly, t is search space at time, i is the position of fmf and vel is the velocity of the iteration.

The velocity of fmf is updated based on the speed that attracts the male mayfly. The best female mayfly is attracted by the best male mayfly with equal rank (i.e., first best fmf is paired with first best mmf. The updated velocity is calculated using the equation (5).

$$vel_i^{t+1} = \left\{vel_{ij}^t + a_2 e^{-\beta r_{mmf}^2}\left(mmf_{ij}^t - fmf_{ij}^t\right), \text{if } f(fmf_i) > f(mmf_i)\, vel_{ij}^t + fw.r\,,\, if\ f(fmf_i) \le f(mmf_i)\ and -1 \le\ r \le 1 \right. \quad (5)$$

where vel_{ij}^t and fmf_{ij}^t denotes the velocity and position of i^{th}fmf in j^{th} dimension at the time t respectively. fw denotes the random walk coefficient. The cartesian distance of mmf and fmf is calculated using the equation 3.

The fmf and mmf mate and produce a pair of children for each one of them. The offspring is calculated using the equation 6 and 7 given below.

$$\mathrm{OS}_1 = L * \mathrm{mmf} + (1-L) * \mathrm{fmf} \quad (6)$$

$$\mathrm{OS}_2 = L * \mathrm{fmf} + (1-L) * \mathrm{mmf} \quad (7)$$

where, L is the Gaussian distribution based random number.

From the above discussion, identifying the best position of mayflies is a difficult task and it plays a major role in providing the best solution. Searching the best position needs more attention. This research work uses mode based ranking methods to identify the best position of the mayfly. The authors have improved the performance of the MFO algorithm using linear gravity and median position of the group. The proposed IMOA uses mode-based ranking method to identify the best position of male mayfly. IoT data is continuous in nature. The smart-city dataset is continuous and grouped data. Mode is one of the well-known methods used to identify the ranking in grouped or ungrouped data.

The mode is working based on the following steps:

1: Identify the modal class

2: Calculate the mode using equation 8

$$M = L + \left(\frac{fq_c - fq_1}{2fq_c - f_1 - f_2}\right] * w \quad (8)$$

where M is the Mode, Lis lower limit of modal class, fq_c is the frequency of modal class, f_1 is the frequency of class preceding modal class, f_2 is the frequency of class succeeding modal class and w is the weight of the class.

The best position of mayflies is calculated using the multiclass modal method that is given in the above equation 8. The mode position of the mayfly of a grouped data is given in the following equation (9)

$$pmode = \{Calculate\ M\ using\ eq.\ (8),\ \ if\ data\ is\ multiclassvel\ with\ max.\ frequency,\ if\ data\ is\ ungrouped \quad (9)$$

The updated velocity will be calculated using the following equation (10). The new constraint equation (9) is added to the equation (2) and gives the velocity of mmf.

$$vel_{ij}^{t+1} = vel_{ij}^{t} + a_1 e^{-\beta r_{pos}^2}\left(posBest_{ij} - mmf_{ij}^{t}\right) + a_2 e^{-\beta r_{gbl}^2}\left(gblBest_{j} - mmf_{ij}^{t}\right) + a_3 e^{-\beta r_{mode}^2}\left(pmode_{j} - mmf_{ij}^{t}\right) \quad (10)$$

where a_3 is the positive attraction-coefficient based on social behaviour. $r_{mode}$is the cartesian distance between the mode and current position of mmf. The cartesian product is calculated by equation 3.

In the optimization algorithm fitness function plays a major role to retain and remove the individuals in the total population. For selecting the best position value, this research paper uses the following fitness function given in the equation 11,

$$f\left(best\right) = \frac{\sum_{i=1}^{N} posBest_i}{N} \quad (11)$$

where, the fitness value is the arithmetic mean of the i^{th} mayfly global best position in the total population N.

2.3 LGBM:

The Light Gradient Boosting Machine (LGBM) developed by Microsoft is an integrated learning model (i.e. ensemble learning). There are different types of boosting methods based on the various loss functions that are as follows: AdaBoosting, Gradient boosting decision tree (GBDT), Gradient-boost AR ensemble learning algorithm, eXtreme gradient boosting decision tree, XGBoost, Categorical boosting, CatBoost, Deep decision tree transfer boosting, DTrBoost[(16)] .

LGBM is a lightweight computing algorithm developed using the Gradient Boosting (GB) system. This is currently the most developed and mature technique for machine learning problems. The following are the main advantages of using LGBM over XGBoost:

- lower memory
- faster output
- parallel training
- adopt multiple loss functions
- higher accuracy

Furthermore, the leaf-wise growth method based on best-first search helps to achieve the minimal loss than other level-wise algorithms. For a minimal number of datasets, the leaf-wise machine learning model may cause overfitting problems. To overcome this issue, LGBM uses max_depth parameter to limit the depth of the tree growing.

2.4 IMFO-LGBM Algorithm

Figure 2 represented the flowchart of the proposed IMFO-LGBM algorithm. LGBM algorithm has various parameters. Due to the minimum size of the datasets, the maximum depth of the parameter is set by using the IMFO algorithm. The proposed IMFO algorithm finds the best features and passes it into the LGBM model. The optimal combination of the IMFO and LGBM increases the accuracy and reduces the computational sparsity when compared with the baseline methods.

Based on the optimization process of IFMO, the steps of the IMFO-LGBM algorithm are given as follows:

1: Load the dataset with 80% of training and 20% for testing
2: Set the population size (mmf, fmf and OS) to the LGBM for the 80% training dataset
3: Set the other parameters (learning factors, visibility and dancing coefficient)
4: Initialize position and velocity of mayflies
5: Evaluate the fitness function
6: Find the fitness function values of OS, update the fitness function, posbest and gblbest
7: Run the step 1 to 5 till the maximum criterion reaches the end
8: LGBM trained with the 80% of training dataset
9: Evaluate the trained IMFO-LGBM with 20% of testing dataset
10: Stop

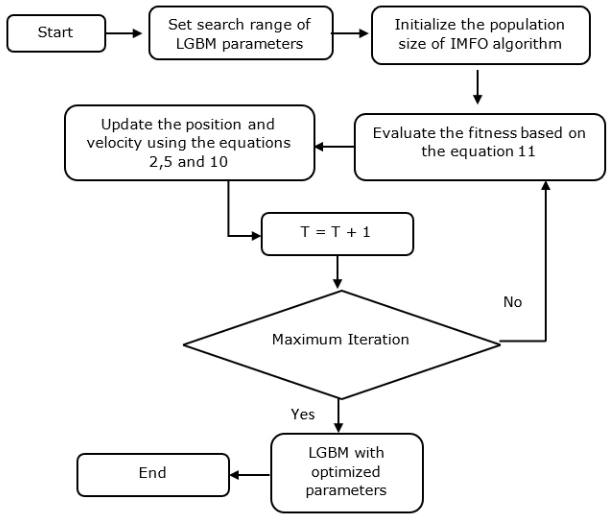

Fig 2. Process flow of IMFO-LGBM

3 Results and Discussions

The proposed IMFO-LGBM algorithm is experimented on the publicly available dataset[13]. The dataset was published in a Kaggle competition for building an effective smart city. The major problem faced by the government is to predict the traffic and provide efficient transportation for the public. The dataset contains 48,121 train records and 11,809 test records. The prepared dataset is split into training and test sets as 80% and 20% respectively.

The proposed IMFO is a nature inspired based optimization algorithm applied along with the lightweight boosting model. The dataset is evaluated with the baseline LGBM, Random Forest and XGBoost models. It is known that random forest can forecast the results accurately when compared with the traditional linear regression. The bagging technique analyses the values of historical data and produces the results based on average. The boosting-based method predicts the next value based on the previous values. It learns sequentially and adaptively to improve the prediction accuracy. The proposed IMFO algorithm helps to identify the best features for the prediction. In this research work, the mayfly optimization is improved by adding a new parameter mode ranking to select the best position of the instance. This helps to improve the accuracy when compared to the baseline models. In this research work, the competence of the proposed work with the baseline methods is evaluated with the following metrics: precision, recall, f-score, accuracy, mean absolute error (MAE) and root mean square error (RMSE).

For experimenting the dataset, this research work used python 3.9, numpy library and scikit learn machine learning libraries. The baseline methods LGBM, random forest and XGBoost produced 89.1%, 85.1% and 86.7 % of accuracy respectively. The mode ranking based mayfly optimization showed an improvement when compared with the baseline LGBM, random forest model. It showed 2 % of increase in accuracy when compared with the LGBM. The boosting-based model LGBM showed an improvement of 3 % in the accuracy when compared with the XGBoost. Among both boosting and bagging, LGBM model outperformed well. Figure 3 represents the chart outcomes of the proposed IMFO-LGBM model along with LGBM, random forest and XGBoost models.

MAE and RMSE are the two important error evaluation metrics for the time-series data. The performance of MAE and RMSE is evaluated based on the time intervals. The time interval is calculated with 5 minutes. Figure 4 shows the improvement of the proposed IMFO-LGBM algorithm in MAE error reduction when compared with the RF, LGBM and XGBoost models. The proposed IMFO-LGBM records the lowest MAE error with 2% which is lower than the RF, XGBoost and LGBM models. The RF models produced high error rate with 8% of accuracy. The RMSE error is reduced in the proposed work. It produces 3.5% or error which lower than the existing models for 5 min interval of time and 10% of error for 25 mins interval of time. The result in the Figure 5 depicts that the increase in the error rate is directly proportional to the increase in the interval of time. The proposed IMFO-LGBM model produces lower error rate in the taken intervals of time. This shows the efficiency of the proposed model. Proposed IMFO-LGBM model produces less incorrect classification when compared with the existing models.

The proposed work IMFO-LGBM provides less improvement when compared with the existing work done using CNN technique. The feature selection model shows an increase of 0.23% of improvement when compared with the existing technique.

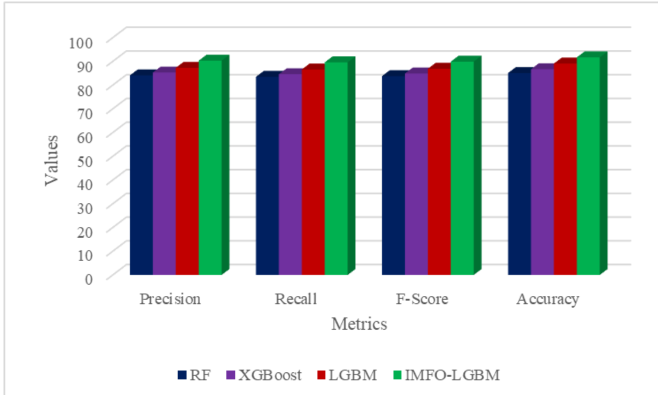

Fig 3. Comparative Results of Smart-City Prediction

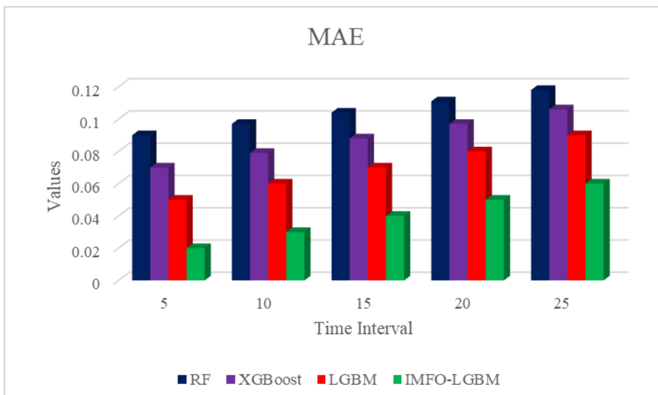

Fig 4. Comparative Results of MAE with Different Time Intervals

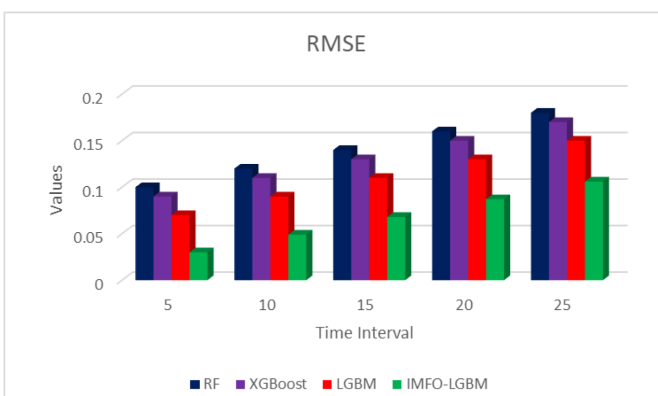

Fig 5. Comparative Results of RMSE with Different Time Intervals

The existing work was experimented using the image datasets captured with drone technologies.Table 1 depicts the comparative analysis of proposed IMFO-LGBM model with existing technique [12].Feature selection plays an important role to show improvement in accuracy.

Table 1. Comparison Results of Proposed IMFO-LGBM with literature

S. No.	Techniques	Accuracy (%)
1.	CNN technique [12]	91.67
2.	Proposed IMFO-LGBM	91.9

4 Conclusion

Newly added modal rank parameter in the May Fly algorithm fine tunes the feature selection process. It helps to train the model quickly and improve the accuracy results. The proposed IMFO algorithm with LGBM produced 2 % of improvement in accuracy. The proposed work has shown better accuracy for the evaluation metrics MAE and RMSE error. The LGBM works faster than the XGBoost model and occupies less memory space.Theproposed IMFO-LGBM model shows an improvement of 0.23% compared with the existing literature.In the future, the proposed model is planned to be evaluated with real-time data collection by unmanned aerial vehicles (UAV).

References

1) Attila M, Nagy V, Simon. Improving traffic prediction using congestion propagation patterns in smart cities. *Advanced Engineering Informatics50101343 Available from*. 2021. Available from: https://doi.org/10.1016/j.aei.2021.101343.
2) Nagham A, Al-Madi AA, Hnaif. Optimizing Traffic Signals in Smart Cities Based on Genetic Algorithm. *Computer Systems Science & Engineering*;2022(1). Available from: https://doi.org/10.32604/csse.2022.016730.
3) Hassan M, Kanwal A, Jarrah M, Pradhan M, Hussain A, Mago B. Smart City Intelligent Traffic Control for Connected Road Junction Congestion Awareness with Deep Extreme Learning Machine. *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*. 2022;p. 1–4. Available from: https://doi:10.1109/ICBATS54253.2022.9759073.
4) Du S, Li T, Gong X, Shi-Jinn, Horng. A hybrid method for traffic flow forecasting using multimodal deep learning. *International journal of computational intelligence systems*;2020(13):85–97. Available from: https://doi.org/10.48550/arXiv.1803.02099.
5) Moses A. Vehicular Traffic analysis and prediction using Machine learning algorithms. *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*. 2020;2020:1–4. Available from: https://doi.org/10.1109/ic-ETITE47903.2020.279.
6) Bang JH, Lee JR. Collision Avoidance Method Using Vector-Based Mobility Model in TDMA-Based Vehicular Ad Hoc Networks. *Applied Sciences*;10(12):4181–4181. Available from: https://doi.org/10.3390/app10124181.
7) Wei W, Wu H, Ma H. An AutoEncoder and LSTM-Based Traffic Flow Prediction Method. *Sensors*. 2019;19(13):2946–2946. Available from: https://doi.org/10.3390/s19132946.
8) Lilhore UK, Imoize AL, Li CT, Simaiya S, Pani SK, Goyal N, et al. Design and Implementation of an ML and IoT Based Adaptive Traffic-Management System for Smart Cities. *Sensors*. 2908;22(8):2908–2908. Available from: https://doi.org/10.3390/s22082908.
9) Tao P, Shen H, Zhang Y, Ren P, Zhao J, Jia Y. Status Forecast and Fault Classification of Smart Meters Using LightGBM Algorithm Improved by Random Forest. *Wireless Communications and Mobile Computing*. 2022;2022:1–11. Available from: https://doi.org/10.1155/2022/3846637.
10) Saleem M, Abbas S, Ghazal TM, Khan MA, Sahawneh N, Ahmad M. Smart cities: Fusion-based intelligent traffic congestion control system for vehicular networks using machine learning techniques. *Egyptian Informatics Journal*. 2022;23(3):417–426. Available from: https://doi.org/10.1016/j.eij.2022.03.003.
11) Syed AS, Sierra-Sosa D, Kumar A, Elmaghraby A. Making Cities Smarter-Optimization Problems for the IoT Enabled Smart City Development: A Mapping of Applications, Objectives. *Constraints Sensors*. 2022;22:4380–4380. Available from: https://doi.org/10.3390/s22124380.
12) Jian L, Li Z, Yang X, Wu W, Ahmad A, Jeon G. Combining Unmanned Aerial Vehicles With Artificial-Intelligence Technology for Traffic-Congestion Recognition: Electronic Eyes in the Skies to Spot Clogged Roads. *IEEE Consumer Electronics Magazine*. 2019;8(3):81–86. Available from: https://doi:10.1109/mce.2019.2892286.
13) Vetrivel. . Available from: https://www.kaggle.com/vetrirah/ml-iot.
14) Zervoudakis K, Tsafarakis S. A mayfly optimization algorithm. *Computers & Industrial Engineering*;2020(145):106559–106559. Available from: https://doi.org/10.1016/j.cie.2020.106559.
15) Lei G, Chang X, Tianhang Y, Tuerxun W. An Improved Mayfly Optimization Algorithm Based on Median Position and Its Application in the Optimization of PID Parameters of Hydro-Turbine Governor. *IEEE Access*;2022(10):36335–36349. Available from: https://doi:10.1109/ACCESS.2022.3160714.
16) Jiang S, Mao H, Ding Z, Fu Y. Deep Decision Tree Transfer Boosting. *IEEE Transactions on Neural Networks and Learning Systems*. 2020;31(2):383–395. Available from: https://doi.org/10.1109/TNNLS.2019.2901273.