International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Analysis of Time-Based Public Transport Demand Prediction Using OPTUNA Framework

R. Thiagarajan [1], Dr. S. Prakash kumar[2]
[1]Ph,D Research Scholar/PG and Research Department of Computer Science
Maruthu Pandiyar College
Vallam Post,Thanjavur-613 403.
(Affiliated to Bharathidasan University,Tiruchirappalli,Tamilnadu,India)
Email id: thiagarajan.nvr@gmail.com
[2]Assistant Professor/ PG and Research Department of Computer Science
Maruthu Pandiyar College
Vallam Post,Thanjavur-613 403.
(Affiliated to Bharathidasan University,Tiruchirappalli,Tamilnadu,India)
Email id: drsp1974@gmail.com

Abstract

Buses are the most popular and easy mode of transportation in all over the world. The state government operates bus service in all routes with low-cost fare. Traffic congestion has risen at an alarming rate due to an increase in the number of automobiles. Travel time have increased as a result, while accessibility and mobility have worsened. The primary challenge encountered by passengers is the absence of information regarding bus numbers that are accessible on a certain route and the approximate time of bus departure. The delay in bus operations, could have several reasons which are inclement weather, traffic jam, and breakdowns. Neither the arrival time of the bus nor the delay are known to the people waiting at the bus stop. In order to address this problem, encouraging the usage of public transportation seems to be a feasible way. Over the past ten years, prediction of bus arrival time has become a fascinating subject around the world. In the transportation sector, Machine Learning (ML) technologies have already shown great promise and have additionally shown to yield a larger return on investment than traditional methods. In this research work , authors propose and develop predictive models to predict public transport demand for passenger transit based on bus arrival time. The dataset shows the proportion of buses operated by the Rochester-Genesee Regional Transportation Authority (RGRTA) that arrive on time. Initially, lazy predict classifier is used for solving regression-based dataset for predicting the bus demand in On-time based passenger transit. Based on the examination of lazy classifiers, the Decision Tree Regressor (DTR) has been identified as the best model. It is assessed using the most advanced hyperparameter optimization framework (OPTUNA). The proposed OPTUNA based DTR which is utilized to identify On-time performance of bus services-based passenger transit. Using OPTUNA for search is an efficient and beneficial approach considering the search speed and the improvement in model accuracy. According to the experimental data, the proposed approach performs better where the R-squared score is 0.9878 with best hyperparameter to be optimized.

Keywords: Buses, public transport, prediction, demand, hyperparameter, optimization, OPTUNA, LazyPredict

Introduction

Traffic congestion has become a worldwide phenomenon and has grown at a rapid pace in recent years. The increase in population growth, motorization and urbanisation are blamed for this spike. Congestion damages the transportation network, lengthens travel time, emits more emissions and fuel also decreases travelling and easy access. Building more roads, highways, separate lanes, and other transportation infrastructure can help to mitigate this problem by increasing its capacity. Due to its own drawbacks, this option is not practical. Despite being the most popular mode of transportation, travellers are unaware of details like arrival time, precise routes, and buses that are available along the way. It is expected of passengers to remain immobile for an endless period of time without being informed regarding the bus where the people are waiting for or any other information about it. The challenging task for managing bus services are due to lack of information about route, location and traffic. In order to tackle the issues encountered by policy makers and bus managers, Intelligent Transportation System (ITS) has been suggested [1]. Without building new infrastructure, the goal of these transport services and technologies is to improve the effectiveness, safety, dependability, and environmental sustainability of existing transport networks. Promoting public transport to make it more desirable than private vehicles is a key component of ITS. From the viewpoint of the passenger, one of the most crucial markers of a dependable and appealing bus service is the provision of accurate journey and schedules for arrival [2]. A new age in transportation engineering and quality control has been brought about by the integration of growing technologies in public transportation, such as real-time tracking systems and new automatic data collecting. The challenge at hand is carrying out an extensive investigation to comprehend and measure the impact of punctuality on bus passengers. On-time performance refers to the ability of bus services to adhere to their published schedules, arriving and departing from stops as planned.

In the transportation sector, ML technologies have already shown great promise and have even shown to yield a larger return on investment than traditional methods. Still, there is room for improvement in the use and exploitation of ML techniques in transportation-related problems. The fundamental objectives of these solutions are to lessen traffic, enhance security and reduce human mistake, limit adverse environmental effects, optimum energy efficiency, and raise surface transportation productivity and efficiency. In feature engineering, data analytics can be enhanced by the capability of ML models and it transforms the unstructured data into features so that important information in the data can be highlighted. The prediction power of computational approaches can be increased while streamlining and accelerating data transformations by feature engineering to choose pertinent features or create unique features for both supervised and unsupervised learning. ML-based model might help bridge the gap between passengers and transport service. This research work has suggested a Lazy Predict (LP) model, which is simple and straightforward for anyone who is familiar with scikit-learn. In this study to provide prediction for create a Lazy Classifier (LC) instance which will get predictions from all of the models for each and every observation. The framework must fit the data for each model, then use metrics to determine which model has the highest accuracy for the current dataset, and finally pick the optimal model. The effectiveness of any classification algorithm is largely dependent on its ideal hyperparameters [3, 4]. The best set of hyperparameters can be chosen to increase the classification algorithm's accuracy.

To determine the ideal hyperparameter values for the ML model, an advanced hyperparameter optimisation framework (OPTUNA) [5] was used in this investigation. As a result, among the available hyperparameters, the best appropriate set was chosen for this investigation. There are several ways to achieve hyperparametric optimisation, including grid and random searches. The OPTUNA hyperparametric search is an additional technique. Conventional random and grid search methods waste a lot of time and are inefficient since they do not learn from the previous optimisation, which is mostly dependent on the amount of hyperparameters in the ML. As needed, the OPTUNA framework modifies the hyperparameters based on ongoing learning from prior optimisations. Consequently, OPTUNA was selected for hyperparameter optimisation in the present research.

The following is a list of this manuscript's primary contributions.

- First, the percentage of buses that depart from their starting point or reach their destination is included in a data collecting module that gathers information on on-time performance.
- To acquire data from open-source data repository and pre-process it using python libraries.
- To build ML models for the classification problem, and train them with the pre-processed dataset.
- To evaluate ML models based on various accuracy metrics to get the best fit model using LP classifier.
- To optimize the hyperparameters using optimization techniques/algorithms to fine tune the model.
- To compare the chosen models and finalize on the robust classification model to accurately predict the demand of bus required for a given period of time based on passenger transit.

Literature Review

A real-time bus management system that considers a variety of aspects, such as traffic congestion and the environment, is described by Shanthi et al. [6]. Passengers receive notification of the estimated bus arrival time at their endpoint through an HCI-based website and a mobile app. There are few factors such as weather and the flow of traffic in the bus's present location are utilised for the ETA prediction employing the Support Vector Regression technique. When tested, the model displayed an RMSE of 27 seconds. The goal of the Hossein Moosavi et al. [7] describes to identify the optimal tree-based ML algorithm and route creation approach for bus journey time prediction for high- and a low-frequency bus route. Furthermore, the precise, dependable, and useful "key stop-based" route creation method was first presented in this study. In the work by Faruk Serin et al. [8] discusses bus arrival time was predicted using ML techniques with a three-layer architecture. The data used in the case study came from Istanbul's public transport system, and it was processed using both classic and three-layer architecture to apply a variety of ML techniques. According to the experimental findings, the three-layer architecture produced effective outcomes with a MAPE of roughly 2.552. A model to simulate the actions of buses and forecast their delays is attempted to be built, as described by Palys et al. [9].

Shrivathsa et al [10] predicting travel time of bus from historical data is proposed using Artificial Neural Network (ANN). By analysing the data, calculate the time taken to reach the destination for every trip and every day. A new ANN method proves that, it is accurate and speed to predict the travel time. The public transportation system's success or

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

failure is largely dependent on its dependability. The predicting issue for transport travel systems was proposed by Dan Luo, Dong Zhao, et al. [11] discusses about bus transit system and passenger flow prediction using Deep Learning (DL). The real-time application proposed by Nagaraj et al. [12] aids in the detection of traveller movement at a separate site. Several elements in the dataset are taken into account for forecasting, including bus type, bus id, source, destination, number of passengers, slot number and income. After the parameters are processed and the clustered data is divided into regions and sent to the deep learning model for prediction. Next, the clustered data is moved to a Long Short-Term Memory (LSTM) model, which eliminates redundant data from the obtained data. Predicting bus passengers is made more accurate by these systems. To improve the precision of forecasting the demand for bike sharing, Yang et al. [13] discusses the several parameters where the suggested model has the potential to significantly increase forecast accuracy. Additionally, this paper examines the impact of various parameters on the predictions made by the model at different times. Bidirectional LSTM (Bi-LSTM) networks were employed by Collini et al. [14] to calculate the quantity of bikes and open bike slots at bike-sharing stations. Mehdizadeh Dastjerdi and Morency initially identified six communities in the bike-sharing network, and then this paper discusses about CNN-LSTM for predicting the demand for pickups in each neighbourhood [15]. The flow of passenger anticipating based on the UBTS was examined by Archana et al. [16]. Here, UBTS concerns are discussed, including prediction of delays, driver conduct, passenger flow, passenger comfort perception fleet size, vehicle failures and sound level. The strategies and solutions for urban transport have offered to address these problems are evaluated. In order to retain a respectable level of service across the whole bus network, Xubin Zhai and Yu Shen [17] discusses low-demand bus services are less likely to be disrupted by real-time cross-line bus dispatching that utilizes scarce vehicle resources, where the proposed model achieves 5% MAPE. The bus-passenger-flow data has erratic features, which Yulong Pei [18] addresses by WPD to break them down into smoother components. The Bi-LSTM model improves the model's capacity for analysing the passenger flow pattern due to the periodicity and nonlinearity of the passenger-flow data. A model for predicting traffic flow based on sparse regression was presented by Zheng et al. [19]. Extreme gradient boosting (XGBoost), was employed by Sun et al. [20] and Lu et al. [21] to predict the level of traffic on the highway. A DL model was developed by Chen et al. [22] for prediction of traffic flow in metropolitan road networks.

Dataset Description

Based on passenger travel, this statistic determines the proportion of RGRTA buses that arrive on time. The percentage of buses that depart from their starting point or reach their destination between 2:59 and 5:59 minutes early is known as the on-time performance. This dataset includes subsidiary, month, year, percent On-time and ridership. The total number of passengers, riders, or boardings is referred to as ridership. It contains a total of 1178 records, and it has 5 attributes described in figure 1.

	Subsidiary	Month	Year	Percent On-Time	Ridership
0	Regional Transit Service	4	2009	84.0	1377039
1	Regional Transit Service	5	2009	83.0	1483123
2	Regional Transit Service	6	2009	82.0	1434123
3	Regional Transit Service	7	2009	83.6	1221534
4	Regional Transit Service	8	2009	83.2	1115882

Figure.1 Recommended dataset based on randomizing the rows

Data Preprocessing

Along with data collection, a missing value check is performed, and the result is an imputed missing value. The dataset can be transformed to float values because the data cannot be strings. After missing imputation, the data is pre-processed using Robustscaler and label encoder to handle scaling the all-variable unit as unique. Scikit-Learn offers the Label Encoder class for this reason, which allows you to convert all string values to float values. It provides a unique numerical value for every category in a variable, making it easier for ML algorithms to examine and comprehend the data. After removing the median, RobustScaler scales the data using the quantile range. Figure 2 displays the pre-processed data.

	Subsidiary	Month	Year	Percent On-Time
0	Regional Transit Service	4	2009	84.00
1	Regional Transit Service	5	2009	83.00
2	Regional Transit Service	6	2009	82.00
3	Regional Transit Service	7	2009	83.60
4	Regional Transit Service	8	2009	83.20
...	...	...	...	...
1174	RTS Wyoming	8	2022	96.50
1175	RTS Wyoming	9	2022	97.00
1176	RTS Wyoming	10	2022	97.00
1177	RTS Wyoming	11	2022	95.10
1178	RTS Wyoming	12	2022	95.95

1178 rows × 4 columns

Figure.2 Pre-processed data

As a consequence, as shown in Figure 3, a heatmap is constructed utilizing a few data points from the dataset.

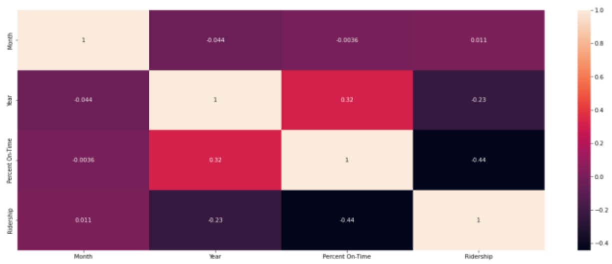

Figure 3 representation of correlation Heatmap

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Research Methodology

The performance of predictive support through data analysis and assist in providing prediction status to the user by proposed method which focus in identifying the best feature fits related to transportation demand. However, one critical factor that significantly influences the attractiveness of bus services to passengers is their on-time performance. The challenge at hand is carrying out an extensive investigation to comprehend and measure the impact of reliability on bus riders. This research work proposed OPTUNA framework has cogitates the best methodical approach due to it tuning parameters like learning rate, loss functions, etc. The Python language is used to develop and train this mode with the help of several libraries, including Matplotlib, NumPy, Scikit-Learn, and Pandas. After being cleaned, the real-time database's data is exported as a CSV file. The functions needed to carry out the hyperparameter optimizations are provided using Scikit-Learn module. The overall proposed model is as shown in figure.4

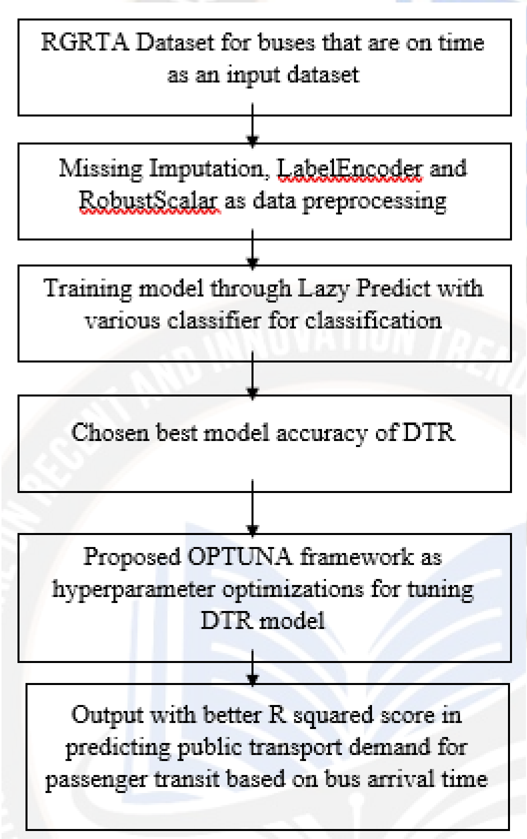

Figure.4 Proposed architecture for predicting transport demand for period of time

This study employs an ML classifier model that is created by grouping different classifiers into a single supervised LP library. The dataset is split in term of models and predictions with respect to train dataset and test dataset. The classification models are made to be evaluated using LP supervised LazyRegressor library. LC provides a convenient and efficient way to fit and evaluate multiple ML models, simplifying the model selection process and allowing to focus on building the best model for our data. LP aids in the development of multiple, distinct, fundamental ML models using specific code and helps identify which models may perform more accurately avoiding the need for parameter tuning. In this study, regression-based datasets are solved using a lazy regressor in classifying the problem with better prediction of demand in bus transportation required for a given period of time from dataset. In order to employing the best readily accessible ML model, the LP classifier can be optimized for accuracy. This can be achieved by optimizing the top model's hyperparameters, as suggested by the OPTUNA hyperparameter tuning. This research work DecisionTree (DT) Regressor, LGBMRegressor and KNeighbors(KN)Regressor has chosen the best fit model to predict the On-time performance of bus services. Further, this DT regressor of untuned model has chosen best fit models based on predicted value which can be improved using OPTUNA framework.

In ML, hyperparameter optimization is a crucial stage. OPTUNA was used in this proposed work to improve the untuned DTR model. This research work uses a unique define-by-run style that makes it relatively easy for users to adjust the ML algorithms of hyperparameters. According to OPTUNA, optimizing hyperparameters involves minimizing or maximizing an objective function that receives a collection of hyperparameters as input and outputs the function score. Without relying on externally supplied static constants, this function dynamically creates the search space of neural network architecture based on the number of layers and hidden units. A trial object is a special OPTUNA trace object that looks for the ideal value based on the hyperparameter's name and range. Through interaction with a trial object, OPTUNA gradually constructs the target function and, when the target function is being executed, uses the trial object to dynamically develop the search area. There are two types of sampling methods in OPTUNA: independent sampling and relational sampling. The correlations between parameters are exploited by relational sampling. The relationships between the parameters were not taken into consideration by independent sampling. Depending on the work and surroundings, both relational and independent sampling can be cost-effective.

The proposed work describes untuned lazy predict DT regressor model were optimized and the loss function is improved. The OPTUNA framework was utilized to improve the hyperparameters, and the primary loss function was utilized as the enhanced loss function. This research work, hyperparameter tuning of certain parameters are used to get the best hyperparameter model. The 13 hyperparameters, including ccp_alpha, min_impurity_decrease, criterion, max_leaf_nodes, min_weight_fraction_leaf, max_depth, max_features, min_samples_leaf, presort, min_samples_split, random_state, min_impurity_split and splitter were chosen for parameter optimization to get best hyperparameter. Although these 13 hyperparameters were chosen for optimization using OPTUNA as shown in table.1.

Table.1 DT regressor hyperparameter optimized using OPTUNA

S.No	hyperparameter name	Untuned parameter	Tuned parameter
1	ccp_alpha	0.0	0.0322
2	criterion	MSE	MSE
3	max_depth	None	12
4	max_features	None	Sqrt
5	max_leaf_nodes	None	None
6	min_impurity_decrease	0.0	0.01423
7	min_impurity_split	None	None
8	min_samples_leaf	1	1
9	min_samples_split	2	2
10	min_weight_fraction_leaf	0.0	9.57418
11	presort	Deprecated	Deprecated
12	random_state	None	None
13	splitter	best	best

The hyperparameters are described as follows:

ccp_alpha is at each stage of the pruning procedure, the cost difficulty pruning route which yields the real alphas and the related total leaf impurity. More of the tree is trimmed as alpha rises, growing the overall impurity of the leaves.

Criterion is a function that assesses a split's quality.

max_depth is the most levels that can be in each tree.

max_features is the most features that are taken into account while dividing a node.

min_samples_leaf represents samples that must be sorted at least once into a tree leaf.

min_samples_split represents minimum number of samples required in a node for node splitting to occur.

max_leaf_nodes represent hyperparameter controls the growth of the tree by imposing a condition on the node splitting.

min_impurity_decrease represents if splitting the node results in an impurity decrease higher than or equal to this value, the node will be divided.

min_weight_fraction_leaf is the percentage of input samples arrives in leaf node this addresses the issue of class imbalance. When presort is enabled, the problem will first be sorted and then permuted.

random_state is the component in charge of randomization during division. It may be an instance of RandomState or an int. The default value is None.

Results and Discussion

In this study, Jupiter IDE and Google Colab are used to help produce and distribute documents that can be explained using text, live code, and visualizations. After collection, the dataset is divided into 70% train and 30% test. In addition, the tunability hyperparameter made use of Seaborn, Sklearn and Pandas. All kind of regression techniques have been implemented by the given specified tools which assist in identifying the efficient untuned regression technique in prediction. In this research DT regressor, LGBM regressor and KN regressor is identified top three model are taken into consideration with an untuned model. Using 25% of the dataset, the accuracy of the trained model is evaluated. Regressions, in contrast to classification issues, yield a numeric value within a range rather than absolute binary values. An algorithm's errors can be measured using a variety of metrics.

Mean Squared Error (MSE): The average of the discrepancies between the actual and anticipated values. The outcome is never going to be negative. Better outcomes are those that are nearer zero.

$$\text{MSE} = \frac{\sum_{i=1}^{n} (at_i - pt_i)^2}{n}$$

Here at_i represents actual time pt_i represents predicted time, n represents number of predictions

Root Mean Square Error (RMSE): This formula calculates the total average magnitude of the error using a quadratic formula. In order to compute it, one must first add up all of the actual and anticipated values difference, squaring the difference value, divide the total by the number of predictions, and then take the square root of the result. The outcome is certain to be positive because the values are rooted and squared. The measured formula is given as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (at_i - pt_i)^2}{n}}$$

The MSE and RMSE methods are used to assess the accuracy of this trained prediction model.

R-Squared score: It explains how well the model performs. It explains the variance in the response or target variable that the data model's independent variables predict.

R square has a value in the interval [0,1].

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

SS_{res} represents the total squares of the data model's residual errors.

SS_{tot} represents the total sum of the errors.

Table 2 displays the outcomes that the models generated. The DT, LGBM, and KN regressor models were shown as graphs with actual values compared to expected values. The graph's linearity indicates how accurate the prediction was. The model is therefore 100% accurate if line y = x the graph is as linear. A scatter plot is the true by projected plot. For the ordinate, the experiential response (Y) is utilised. Another way to visually assess the likelihood of "lack of fit" is via the plot. An impartial prediction to yield anticipated values that, on average, correspond with the observed values. Figure.5, 6 and 7 show the predicted value versus true value plot for DT, LGBM and KN regressor model.

Table.2 Metrics to measure the errors for top three regressor ML model

Algorithm	R-Squared	MSE	RMSE
DT Regressor	0.9864	2361332682.94067	48593.54569
LGBM Regressor	0.9860	2438426254.05480	49380.423793
KN Regressor	0.9859	2456666405.6648	49564.7698

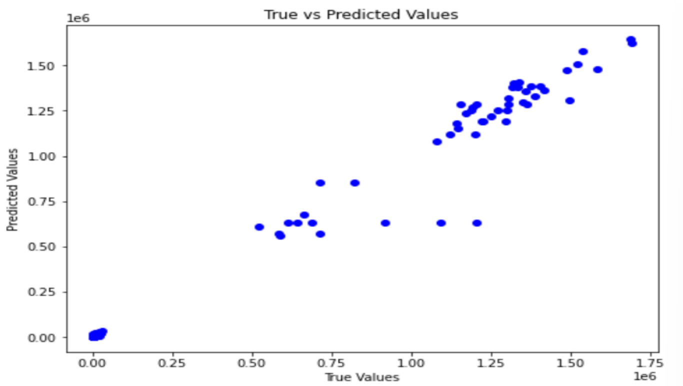

Figure.5 Predicted versus true value plot for DT regressor model

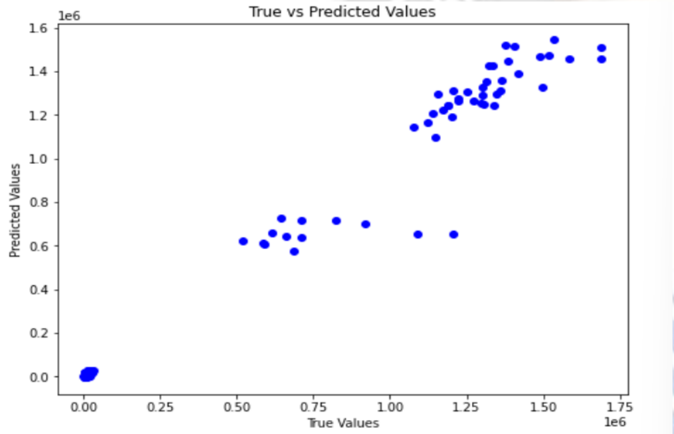

Figure.6 Predicted versus true value plot for LGBM regressor model

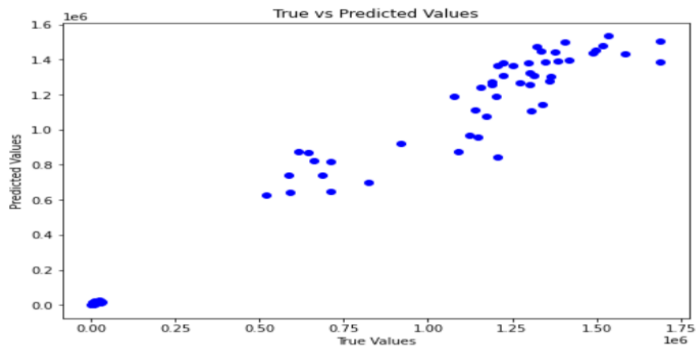

Figure.7 Predicted versus true value plot for KN regressor model

While the points in other models were a little more dispersed, the model with the DTR algorithm gives results from all the scattered point graphs of predicted values against true values that closely resemble the linear graph. Table 2 and Figure 5 demonstrate that, out of all the models considered, the DTR performs the best, with an accuracy of R-squared score of 0.9864. LGBM and KN regressor was performing with an accuracy of R-squared score is 0.9860 and 0.9859 which is nearly as good as but slightly lesser than DT regressor model. Similarly, the DT regressor model performs better prediction in RMSE and MSE whereas as in LGBM and KN regressor model performs higher error rate prediction than DT regressor model. The DTR model was therefore fitted with these hyperparameters, which were optimised using the proposed OPTUNA framework.

Table.3 Metrics to measure the errors for proposed model

Algorithm	R-Squared	MSE	RMSE
Proposed OPTUNA based DT Regressor Model	0.987	2120385948.23	46047.6486

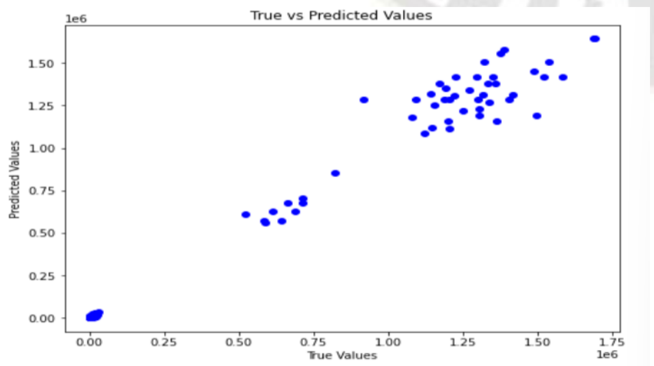

Figure.8 Predicted versus true value plot for proposed OPTUNA based DT Regressor Model

In comparison to existing classification methods, the experimental results of the proposed model show higher values for various evaluation metrics when the accomplished prediction model is checked for accuracy using R squared score. Table.3 and Figure.8 displays results of model with improved classification performance based on hyperparameter optimization to get best hyperparameter compared to other models was introduced. The proposed OPTUNA based DT Regressor model achieved a R-Squared of 0.987, a MSE of 2120385948.23, and an RMSE of 46047.6486.

Conclusion

This paper has discussed about understanding the relationship between on-time performance and bus ridership is crucial for urban planners, transportation authorities, and policymakers to make informed decisions that enhance the efficiency and appeal of bus services, ultimately contributing to more sustainable and accessible urban transportation systems. In this study, regression-based datasets are solved using the LP classifier for chosen the best model that generates high classification accuracy of R squared score for DT regressor is 0.9864 has been improved through optimizer and hyper parameter. The most sophisticated hyperparameter optimisation framework was used in this research to optimise the prediction model's hyperparameters, which allowed for an accurate estimation of the demand for bus transportation over a certain time period based on passenger transit. The model accuracy of R squared score is improved through OPTUNA with an increase of 0.987 for DT regressor model which is more likely to be suitable in the percentage of buses that depart from their starting point or arrive at a destination based on passenger transportation is one factor used to estimate on-time performance. It can facilitate the usage of public transportation and allow transportation authorities to effectively allocate their limited resources.

Reference

[1] V. E Sathishkumar and Y. Cho, “A rule-based model for Seoul Bike sharing demand prediction using weather data,” European Journal of Remote Sensing, vol. 53, no. 1, pp. 166–183, 2020.

[2] Tyndall J. Bus quality improvements and local commuter mode share. Transp Res Part A Policy Pract 2018;113:173–83.

[3] S. Nematzadeh, F. Kiani, M. Torkamanian-Afshar, and N. Aydin, ‘‘Tuning hyperparameters of machine learning algorithms and deep neural networks using metaheuristics: A bioinformatics study on biomedical and biological cases,’’ Comput. Biol. Chem., vol. 97, Apr. 2022, Art. no. 107619, doi: 10.1016/j.compbiolchem.2021.107619.

[4] M. Liang, B. An, K. Li, L. Du, T. Deng, S. Cao, Y. Du, L. Xu, X. Gao, L. Zhang, J. Li, and H. Gao, ‘‘Improving genomic prediction with machine learning incorporating TPE for hyperparameters optimization,’’ Biology, vol. 11, no. 11, p. 1647, Nov. 2022, doi: 10.3390/ biology11111647.

[5] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, ‘‘OPTUNA: A nextgeneration hyperparameter optimization framework,’’ in Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Anchorage, AK, USA, 2019, pp. 2623–2631.

[6] N. Shanthi, Sathishkumar V E , K. Upendra Babu, P. Karthikeyan, Sukumar Rajendran, and Shaikh Muhammad Allayear, “Analysis on the Bus Arrival Time Prediction Model for Human-Centric Services Using Data Mining Techniques”, Hindawi Computational Intelligence and Neuroscience Volume 2022, Article ID 7094654, 13 pages https://doi.org/10.1155/2022/7094654

[7] Seyed Mohammad Hossein Moosavi, Mahdi Aghaabbasi , Choon Wah Yuen , Danial Jahed Armaghani, “Evaluation of Applicability and Accuracy of Bus Travel Time Prediction in High and Low Frequency Bus Routes Using Tree Based ML Techniques”, Journal of Soft Computing in Civil Engineering, 2023;7(2):74–97. https://doi.org/10.22115/scce.2023.356348.1503.

[8] Faruk Serin, Yigit Alisan and Metin Erturkler, “Predicting bus travel time using machine learning methods with three-layer architecture”, Measurement, volume 198, July 2022,111403.

[9] Lukaz Palys, Maria Ganzha and Marcin Paprzycki, “Machine learning for bus travel prediction”,2022, ICCSCameraReadyVersion, https://dx.doi.org/10.1007/978-3-031-08754-7_72.

[10] Shamanth R Shrivathsa, Shailesh M R , Dr. Rajeswara Rao K V S, “Travel Time Prediction of Public Transport in A Selected Route Using Artificial Neural Networks”, 2019 JETIR February 2019, Volume 6, Issue 2.

[11] Luo D, Dong Z, Ke Q, You X, Liang L, Zhang D, Ma H (2021) Fine-grained service-level passenger flow prediction for bus transit systems based on multitask deep learning. IEEE Trans Intell Transp Syst 22:1–16.

[12] Nandini Nagaraj, Harinahalli Lokesh Gururaj, Beekanahalli Harish Swathi and Yu-Chen Hu, “Passenger flow prediction in bus transportation system using deep learning”, Multimedia Tools and Applications (2022) 81:12519–12542.

[13] Yang Yang , Xin Shao , Yuting Zhu , Enjian Yao ,Dongmei Liu, and Feng Zhao, “Short-Term Forecasting of Dockless Bike-Sharing Demand with the Built Environment and Weather”, Hindawi Journal of Advanced Transportation Volume 2023, Article ID 7407748, 13 pages https://doi.org/10.1155/2023/7407748

[14] E. Collini, P. Nesi, and G. Pantaleo, “Deep learning for short-term prediction of available bikes on bike-sharing stations,” IEEE Access, vol. 9, pp. 124337–124347, 2021.

[15] A. Mehdizadeh Dastjerdi and C. Morency, "Bike-sharing demand prediction at community level under COVID-19 using deep learning," Sensors, vol. 22, no. 3, p. 1060, 2022.

[16] Archana M. Nayak, Akhilesh Ladha, Nirbhay Kumar Chaubey, "A Comprehensive Comparative Analysis of Passenger Demand Prediction for Improving the Urban Bus Transportation System (UBTS) " *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 269-279, 2022. Crossref, https://doi.org/10.14445/22315381/IJETT-V70I9P227.

[17] : Zhai, X.; Shen, Y. Short-Term Bus Passenger Flow Prediction Based on Graph Diffusion Convolutional Recurrent Neural Network. Appl. Sci. 2023, 13, 4910. https://doi.org/10.3390/ app13084910.

[18] Pei, Y.; Ran, S.; Wang, W.; Dong, C. Bus-Passenger-Flow Prediction Model Based on WPD, Attention Mechanism, and Bi-LSTM. Sustainability 2023, 15, 14889. https://doi.org/10.3390/ su152014889

[19] Zheng, Z.; Shi, L.; Sun, L.; Du, J. Short-Term Traffic Flow Prediction Based on Sparse Regression and Spatio-Temporal Data Fusion. IEEE Access 2020, 8, 142111–142119

[20] Sun, B.; Sun, T.; Jiao, P. Spatio-Temporal Segmented Traffic Flow Prediction with ANPRS Data Based on Improved XGBoost. J. Adv. Transp. 2021, 2021, 5559562.

[21] Lu, X.; Chen, C.; Gao, R.; Xing, Z. Prediction of High-Speed Traffic Flow around City Based on BO-XGBoost Model. Symmetry 2023, 15, 1453.

[22] Chen, C.; Liu, Z.; Wan, S.; Luan, J.; Pei, Q. Traffic Flow Prediction Based on Deep Learning in Internet of Vehicles. IEEE Trans. Intell. Transp. Syst. 2021, 22, 3776–3789