Bitlis Eren Üniversitesi Fen Bilimleri Dergisi
BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 13 NO: 1 PAGE: 1-14 YEAR: 2024
DOI:10.17798/bitlisfen.1292003

A Comparative Analysis of Passenger Flow Forecasting in Trams Using Machine Learning Algorithms

Beyhan ADANUR DEDETURK[1*], Bilge Kagan DEDETURK[2], Ayhan AKBAS[3]

[1]*Department of Computer Engineering, Abdullah Gul University, Kayseri, Türkiye*
[2]*Department of Software Engineering, Erciyes University, Kayseri, Türkiye*
[3]*Institute for Communication Systems, University of Surrey, Guildford, UK*
(ORCID: 0000-0003-4983-2417) (ORCID: 0000-0002-8026-5003) (ORCID: 0000-0002-6425-104X)

Keywords: Time Series Forecasting, Passenger Flow, Machine Learning, Deep Learning.

Abstract
Forecasting tram passenger flow is an important part of the intelligent transportation system since it helps with resource allocation, network design, and frequency setting. Due to varying destinations and departure times, it is difficult to notice large fluctuations, non-linearity, and periodicity in tram passenger flows. In this paper, the first-order difference technique is used to eliminate seasonal structure from the time series data, and the performance of different machine learning algorithms on passenger flow forecasting in trams is evaluated. Furthermore, the impact of the COVID-19 pandemic on forecasting success is examined. For this purpose, the tram data of Kayseri Transportation Inc. for the years 2018-2021 is used. Different estimation models, including Linear Regression, Support Vector Regression, Random Forest, Artificial Neural Network, Convolutional Neural Network, and Long Short-Term Memory are applied, and the time series forecasting performances of the models are evaluated with MAPE and R^2 metrics.

1. Introduction

With the continuous expansion of the corporate sector in major urban centers, traffic congestion has become increasingly prominent. Among its various challenges, overcrowding is particularly critical, presenting hidden threats to public safety and significant time wastage. A promising solution to this problem is the expansion of public transportation networks, focusing specifically on trams [1].

The rapid proliferation of tram systems and the development of sophisticated information management systems have led to a substantial generation of passenger trip data. This data surge has sparked significant interest in the scientific community, particularly in developing reliable methods to predict tram passenger flow. Accurate forecasting of passenger flows is crucial for efficient transportation management and plays a pivotal role in devising appropriate contingency plans for emergencies, thus enhancing the city's overall emergency response capabilities [2].

In essence, the growth of corporate activities in urban areas has exacerbated traffic congestion issues, especially overcrowding. This has underscored the need for solutions like expanding tram networks. Leveraging the wealth of data generated by these systems has become a key research focus. Accurate predictions of passenger flow can lead to better urban transportation management and improved emergency preparedness.

The aim of the article is to assess the efficiency of different machine learning algorithms in forecasting tram passenger flow. The study tackles the challenges in forecasting due to significant fluctuations, non-linearity, and periodicity in passenger numbers. Techniques such as log transform and first-order difference are utilized to preprocess the data, and the study investigates the impact of the COVID-19 pandemic on forecasting accuracy. Data

*Corresponding author: *beyhan.adanur@agu.edu.tr*
Received: 03.05.2023, Accepted: 02.02.2024

from Kayseri Transportation Inc. for 2018-2021 is analyzed using various models, including Linear Regression, Support Vector Regression, Random Forest, Artificial Neural Network, Convolutional Neural Network, and Long Short-Term Memory (LSTM). These models are evaluated using MAPE and R^2 metrics to establish effective methods for predicting tram passenger flows, which can enhance urban transportation management and emergency preparedness.

In the study, the log transform is applied to daily tram passenger number data from January 1, 2018, to July 1, 2021. Subsequently, the first-order difference of consecutive days' passenger numbers is calculated. Three datasets are prepared: pre-pandemic (799 data points), in-pandemic (445 data points), and the entire dataset (1244 data points). These are divided into training and test sets for the last 30 days' prediction using the previous 10 days' data. The aforementioned models are trained and tested, employing multi-step and multi-output forecasting techniques. The study compares the results and assesses the impact of the pandemic on forecasting success, aiming to contribute to the field with new insights and methodologies. The key contributions of this study can be summarized as follows:

- Creation and public sharing of the Kayseri tram passenger flow dataset.
- Investigation of the COVID-19 pandemic's effects on time series forecasting.
- Application of log transform and first-order difference techniques to reduce noise, linearize the data, and address seasonal patterns.
- Implementation and detailed comparative performance evaluations of various machine learning and deep learning methods.
- Comparison of the performance of multi-step and multi-output techniques.

We believe this will establish a hybrid method for future studies. The remainder of this paper is organized as follows: Section 2 reviews the related works. Section 3 describes the methodologies of LR, SVR, Random Forest, ANN, CNN, and LSTM. Section 4 discusses the experimental results. Finally, Section 5 concludes the paper, explaining the major results and limitations of the current study, their significance, and suggesting future research topics.

2. Related Work

Intelligent computing and machine learning technologies are increasingly being used in various forecasting application scenarios, yielding impressive results. This progress is largely due to advancements in artificial intelligence and the growth of big data [3]-[4]. Forecasting models can be categorized into three types: parametric, non-parametric, and hybrid [5]. Several parametric methods for forecasting transportation demand have been developed, including Box-Jenkins [6], smoothing techniques [7], autoregressive integrated moving average (ARIMA) [8], gray forecasting [9], and state space models [10]. Among these, the ARIMA model [8]-[11] is frequently used. It is a linear function of time-lagged variables and error terms; however, passenger flows are often characterized by high fluctuations, non-linearity, and periodicity. Therefore, traditional parametric models, which assume linear relationships between time-lagged variables, may not effectively represent the structure of non-linear flows.

Similarly, various non-parametric methods for forecasting transportation demand have been developed, including neural networks [14], k-nearest neighbors [15], Kalman filters [16], support vector regression (SVR) [17], and other methods [18]. Neural network models, such as Back Propagation (BP), stacked auto-encoders, and LSTM, often show good performance in trip mode analysis and flow prediction or similar issues [19]-[24]. However, they are susceptible to parameter selection and can be prone to local minima and overfitting [25]. SVM variants are also commonly used [26]-[28]. Unlike neural networks, SVR employs the structural risk minimization principle, aiming to reduce the generalization error upper bound rather than the training error [29], potentially overcoming some fundamental flaws of neural networks [30].

The use of hybrid models to enhance forecasting accuracy has become increasingly popular [31]. Each model constituting a hybrid model has its own set of advantages and disadvantages. The key idea behind hybrid modeling [32], [33] is to combine multiple models while retaining each's benefits. These models have shown promising results in addressing forecasting challenges.

Wang et al. [34] used an integrated model combining multivariate linear regression, K-nearest neighbor, XGBoost, and GRU as four submodels to accurately predict urban public transportation short-term passenger flows. They then integrated these models using a regression algorithm, demonstrating the integrated model's superiority over individual sub-models. Additionally, the popular hybrid forecasting model, the decomposition-integration method, decomposes the original data into several components, processes each component, and integrates them for final predictions [5], [35]-[36]. However, this method is rarely used for short-term bus passenger flow forecasting, and it typically involves only a single decomposition of the original data. Some components remain highly unstable after initial decomposition, which hinders accurate predictions. Therefore, additional noise reduction for unstable components post-decomposition is necessary.

Li et al. [37] developed a secondary decomposition-integration method for short-term bus route passenger flow prediction, integrating empirical modal decomposition, sample entropy, and kernel extreme learning machines. However, the superficial structure of traditional machine learning methods struggles with the complex nonlinearity of spatial and temporal travel demand patterns [38].

A recent study introduced a new model, ITS-Pro-Flow, for predicting short-term traffic flow in intelligent transportation systems (ITS) [49]. It builds upon the Pro-Energy model, utilizing historical data and current conditions for prediction. ITS-Pro-Flow improves upon Pro-Energy by dynamically adjusting past predictions and current observations, with extensive simulations showing its enhanced accuracy. The model incorporates a dynamic weighting factor and a thresholding strategy, improving adaptability and precision. The study also explores parameter variations for optimal prediction accuracy.

With the advent of the Internet of Things (IoT), numerous devices and software have been developed to assist in prediction tasks. Gao et al. [39] proposed a method to increase the efficacy of software-defined devices, while they also introduced a technology to transform business process execution language (BPEL) into timed automata for formal verification, bridging BPEL and IoT data in support of prediction tasks [40]. Huang et al. [41] optimized virtual machine allocation strategies for cloud data centers, and Ma et al. [42] proposed a real-time multiple workflow scheduling method in a cloud environment, enhancing the processing efficacy of large data sets for passenger flow source data analysis.

Despite the successes of these methods, current passenger flow forecasting faces challenges such as reliance on a single data source and insufficient analysis of influencing factors, leading to low accuracy in existing methodologies and impacting urban traffic management [43]. Often, the performance of a hybrid model is either compared with those in the literature or with a maximum of 2-3 methods.

3. Material and Method

In this section, we briefly talk about the data set and the preprocessing methods used. Following that, we discuss the machine learning models that were utilized, which include LR, SVR, Random Forest, ANN, CNN, and LSTM architectures. A flow diagram is presented in Figure 1 to demonstrate the research process.

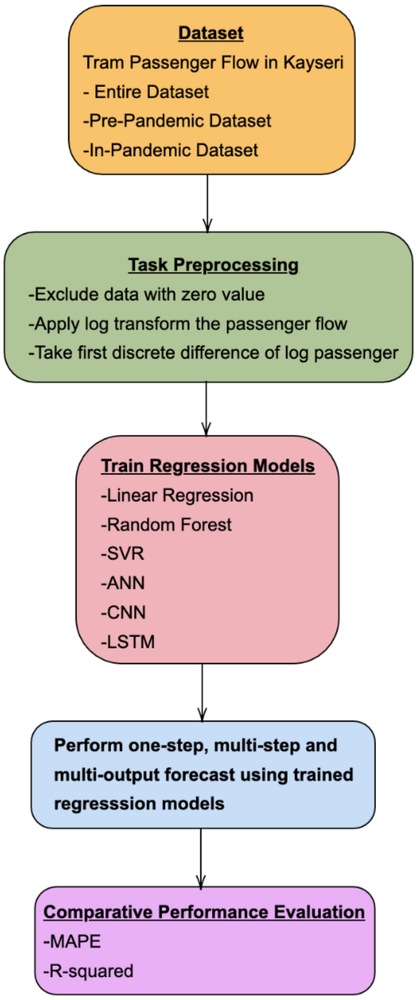


Figure 1. Overall methodology diagram.

3.1. Dataset

The dataset [44] consists of Kayseri tram daily passenger data between January 1, 2018 and July 1, 2021. The daily passenger flow data is shown in Figure 2. Because of eids such as Ramadan Eid, Eid of Sacrifice, and July 15 Democracy and National Unity Day, a free transit system is implemented in the city of Kayseri, and therefore, tram passenger numbers are not available. In addition, the number of passengers remained below fifty on the days when the curfew was performed during the COVID-19 pandemic. For this reason, the number of passengers on some days is seen as close to zero in Figure 2. These days were excluded from the dataset due to the absence of passenger flow data. Therefore, passenger data for 33 days out of 1277 was excluded from the dataset, leaving a total of 1244 passenger data points in the dataset. As seen from Figure 2, there has been a noticeable decrease in passenger numbers since mid-March 2020 due to the COVID-19 pandemic. The first detected COVID-19 case in Türkiye was announced by the Ministry of Health on March 11, 2020. The first death due to the virus in the country occurred on March 15, 2020. As of March 16, 2020, passenger flow in urban rail transportation in Kayseri started to decrease. In our study, in order to observe the effect of the COVID-19 pandemic, the entire dataset was divided into two parts: pre-pandemic (799 days) and during the pandemic (445 days), and three different datasets were obtained together with the whole dataset.

3.2. Preprocessing

Before being modeled with machine learning techniques, time series analysis often requires some data preparation. In time series forecasting, data transforms can be used to eliminate noise and improve the signal. One of the most frequently used data transformation techniques in time series data analysis is log transformation. By using log transformation, time series data with an exponential distribution can be turned into a linear trend, which is easy to model. Figure 3 shows the data obtained after applying the log transform to the passenger flow data, and, as can be seen, the data is squeezed into a smaller range. After obtaining log passenger data, the first-order difference technique was performed to make the time series data stationary. To simplify the prediction problem, differencing methods can be employed to eliminate trend and seasonal structure from the series. A difference transform is a simple approach to getting rid of a systematic structure in a time series. By subtracting the preceding value from each value in the series, a trend can be removed. The process is known as first-order differencing. After performing the first-order difference on the log transform data, the obtained data are shown in Figure 4.

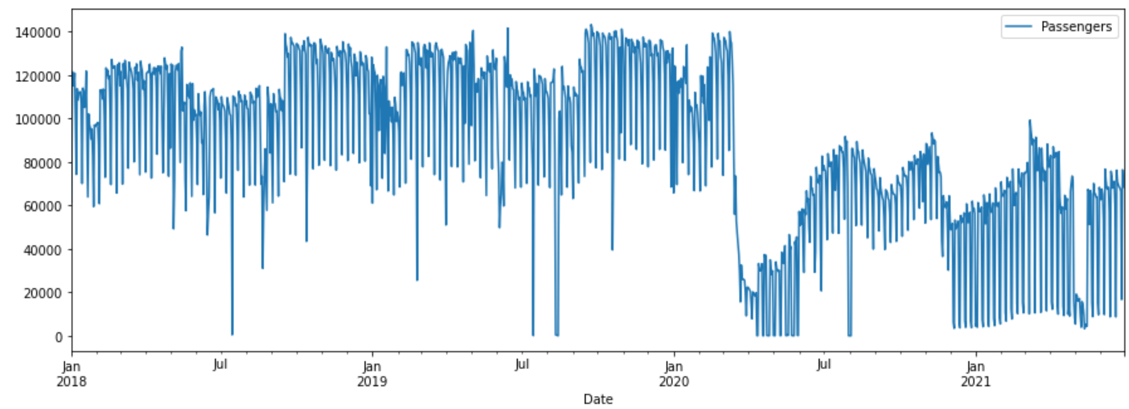


Figure 2. Tram passenger flow data in Kayseri.

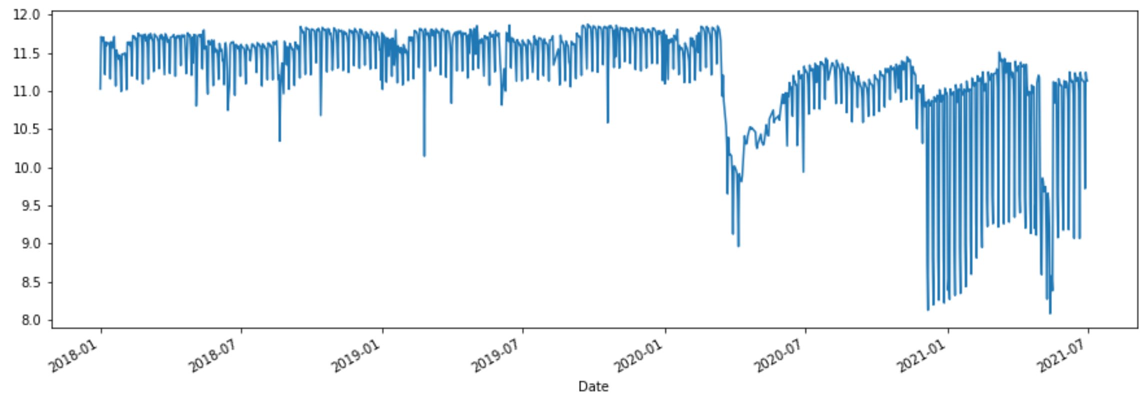


Figure 3. Tram passenger flow data after applying log transformation.

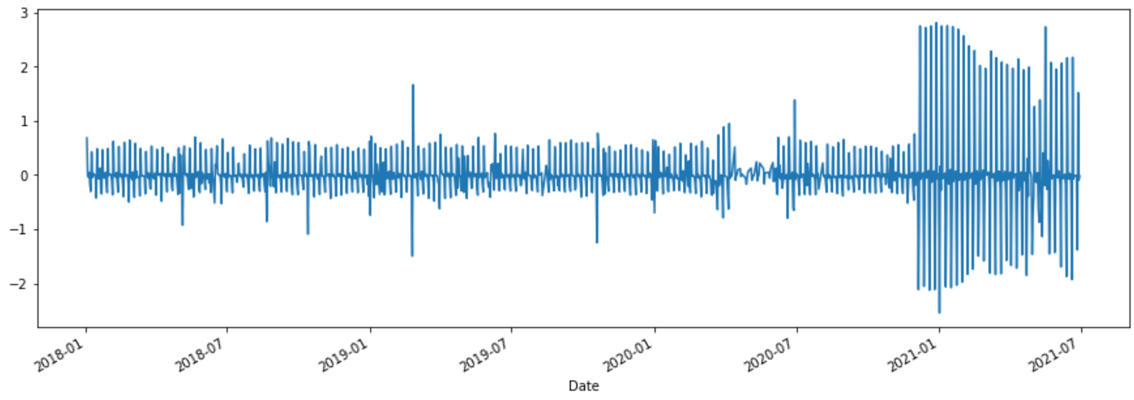


Figure 4. The data obtained after performing first order difference on log passenger data.

3.3. Linear Regression

Linear regression is a statistical technique used for predicting future values based on past data. It's a popular quantitative method for identifying underlying trends and determining if values have deviated significantly. A linear regression trendline uses the least squares method to plot a straight line across data, minimizing the gaps between the values and the trendline. Each data point's trendline value is plotted using this linear regression indicator [45].

In our experiments, two different LR models are trained. The first model, used for one-step forecasting, predicts the passenger flow of a day using first-order difference data from the previous 10 days, resulting in an input vector size of 10. The second model, designed for multi-output forecasting, predicts passenger flow for the next 30 days using the same input data. Both models use the same input layer shape and are trained and tested using the scikit-learn Linear Regression module with default settings.

3.4. Support Vector Regression

Support vector machines (SVM) are a supervised machine learning method based on the Vapnik-Chervonenkis (VC) theory, which identifies characteristics of machine learning conducive to accurate test data predictions. SVM is applicable to both classification and regression problems. Support vector regression (SVR) involves computing a linear regression function in a high-dimensional input space (feature space) where the input data is mapped through a nonlinear function. This transforms a nonlinear regression problem in low-dimensional input space into a linear one in high-dimensional space, where the solution is derived [46].

Two different SVR models are trained in the experiments. The first model, for one-step forecasting, and the second model, for multi-output forecasting, use the same inputs as the LR models. Both models are trained and tested using the Scikit-learn SVR module with default settings.

3.5. Random Forest

A random forest (RF) is an ensemble of several independent decision trees. Each tree generates a class prediction, and the model's forecast is the one with the most votes. Random forests are effective because they combine many generally uncorrelated trees, outperforming individual constituent models. RF is notable for its application to both regression and classification problems, faster training compared to other methods, higher estimation speed, fewer tuning parameters, and direct applicability to multidimensional problems [47].

Two different RF models are trained in the experiments. The first model is for one-step forecasting, while the second model is for multi-output forecasting. Both models use the same inputs as the previous models and are trained and tested using the Scikit-Learn Random Forest Regressor module with default settings.

3.6. ANN

The ANN (Artificial Neural Network) model is an intelligent system used for solving complex issues in various applications, including optimization and prediction. The ANN structure comprises an input layer for data collection, an output layer for computed information, and one or more hidden layers connecting the input and output. Each neuron, the fundamental processing unit of a neural network, performs two tasks: receiving inputs and generating output. Inputs are multiplied by connection weights, their products and biases are added, and then an activation function is used to generate output.

Two different ANN models are trained in the experiments. The first model, for one-step forecasting, uses first-order difference data from the previous 10 days, resulting in an input vector size of 10. It is enhanced by adding a dense layer of 24 units activated using the Rectified Linear Unit (ReLU) function and a 1-unit dense output layer for single-day prediction. The second model, for multi-output forecasting, uses the same input data. It includes a 24-unit dense layer and a 30-unit dense output layer, with the hidden layer activated using the ReLU function.

3.7. CNN

A standard CNN (Convolutional Neural Network) design includes an input layer, multiple hidden layers, and an output layer. The hidden layers consist of convolutional layers, an activation layer, pooling, and fully connected dense layers. The convolution layer, critical to the CNN model, accumulates discriminative features from the input using defined convolution filters. The activation layer introduces non-linearity with an activation function (e.g., ReLU, tanh, or sigmoid), helping to resolve the vanishing gradient problem during training. The pooling layer's main goal is to reduce the data representation size, the number of parameters, and the model's computational cost.

Two different CNN models are trained in the experiments. The first model, for one-step forecasting, uses first-order difference data from the previous 10 days with an input layer shape of 10x1. It includes a 16-filter convolutional layer with a 3-kernel and ReLU activation, followed by a max-pooling operation with a pool size of two. A second convolutional layer with a 32-filter and 3-kernel, along with ReLU activation, is added, followed by a GlobalMaxPooling1D operation and a 1-unit dense output layer. The second CNN model is similar but includes a 30-unit output layer for multi-output forecasting.

3.8. LSTM

LSTM (Long Short-Term Memory) is a type of RNN (Recurrent Neural Network) designed to address the problem of vanishing or exploding gradients. Its unique feature is the LSTM cell, which has specific gates: the input gate determines the relevance of incoming data, the forget gate decides what portions of the cell state to discard, and the output gate determines what information to forward to the next hidden state. Compared to traditional RNN models, LSTM can maintain long-term dependencies among input data items, helping to alleviate the vanishing gradient problem.

Two different LSTM models are trained in the experiments. The first model, for one-step forecasting, uses first-order difference data from the previous 10 days, with an input layer shape of 10x1 and a hidden state dimension of 24. It includes a dense layer with 1 unit. The second model, for multi-output forecasting, uses the same input data. An LSTM layer with 24 internal units is added, followed by a GlobalMaxPooling1D operation and a dense output layer of 30 units.

3.9. Evaluation Criteria

In this research, two popular scale invariant metrics named Mean Absolute Percentage Error (MAPE) and R-squared (R^2) are used to evaluate the performance of forecasting methods. The MAPE is the mean of absolute percentage errors. One disadvantage of this metric is that if there is an actual value that equals zero, then the MAPE value equals infinity, and this makes no sense with regard to percentage. However, the preprocessed data used in this research does not contain zero values.

$$\boldsymbol{MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - p_i}{y_i}\right|} \quad \textbf{(1)}$$

where N is the total number of data values, $\boldsymbol{p_i}$ is the predicted value, and $\boldsymbol{y_i}$ is the actual value for ith position. Because the MAPE value corresponds to the error in terms of percentages, the lower the MAPE value is, the better the forecast is. As indicated in Equation (2), the R^2 metric indicates how much variance is accounted for by the fitted model. It is a statistical measure that represents the proportion of the variance for a dependent variable that’s explained by an independent variable or variables in a regression model. The higher the R^2 value represents the better the prediction performance [4].

$$\boldsymbol{R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - p_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \quad \textbf{(2)}$$

where $\boldsymbol{y_i}$ and $\boldsymbol{p_i}$ correspond to the actual value and predicted value for i^{th} position, respectively. $\bar{y}$ is the mean of the actual values.

4. Results and Discussion

In this section, we will first evaluate the forecasting performances of time series forecasting methods on pre-pandemic, in-pandemic, and entire datasets, according to the visuals in Table 1, Table 2, and Table

3. Based on these tables, we will compare the performances of the methods with each other. Additionally, the best performances will be evaluated according to the MAPE and R^2 values presented in Table 4 and Table 5. Then, we will examine the impact of the COVID-19 pandemic on forecasting performance. Finally, the effects of multi-step and multi-output techniques on forecasting performance will be explored, and we will determine which technique is superior.

One-step forecasts are used to predict the next step's observation. In contrast, multi-step forecasts are utilized to predict a series of future values based on observed time series data. The multi-output technique involves creating a single model capable of predicting the entire forecast sequence in one attempt [48].

Table 1. Comparison of actual values and predicted values by LR and SVR.

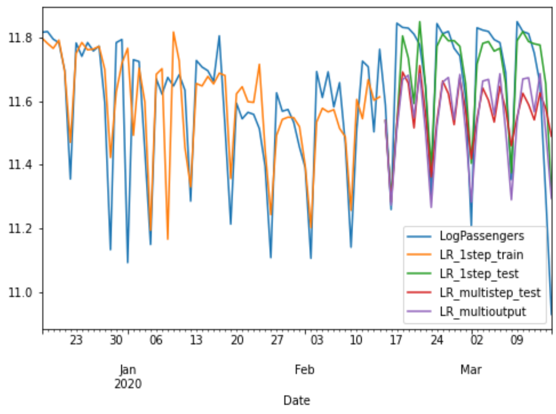
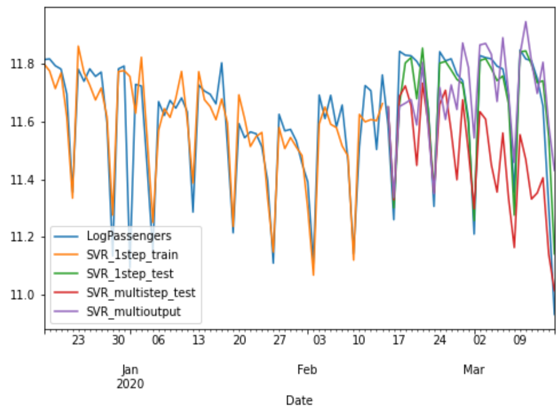
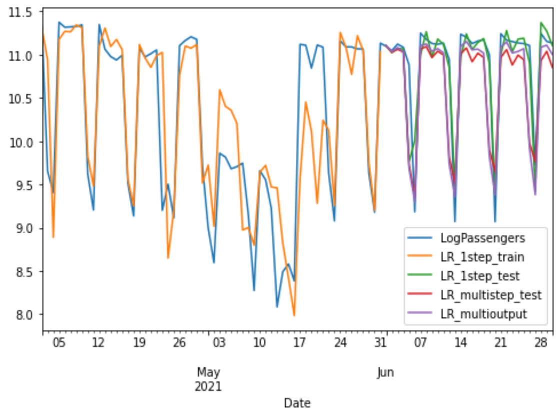
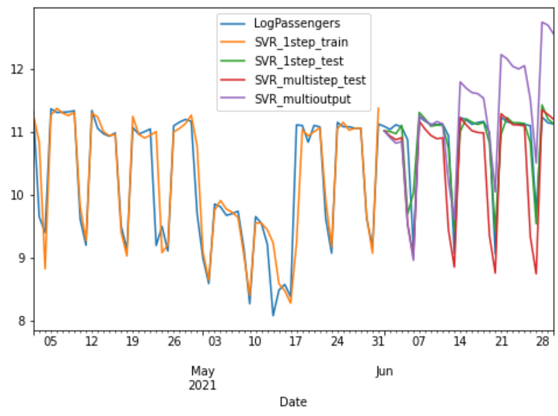
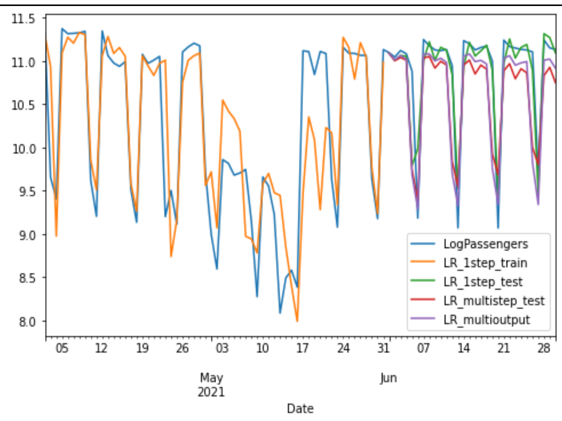
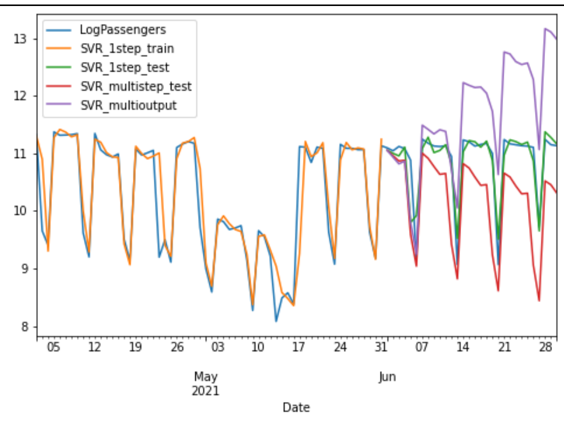
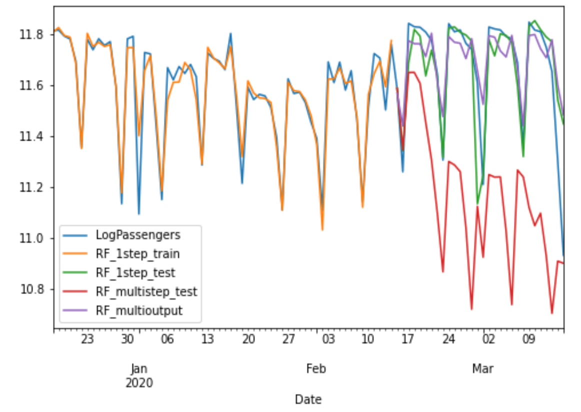
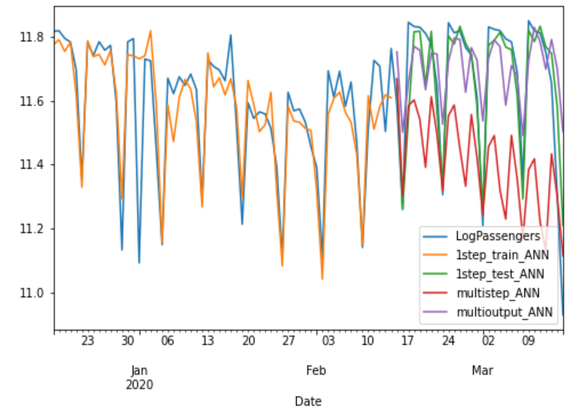
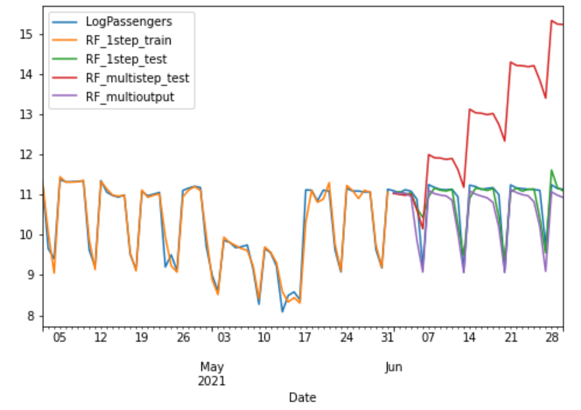
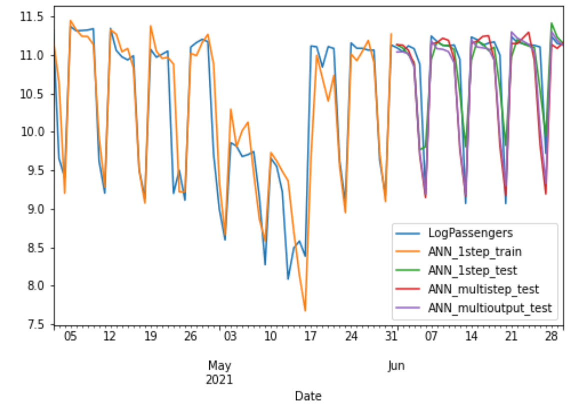
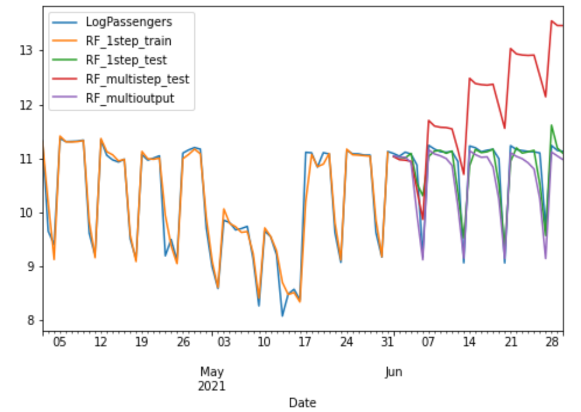
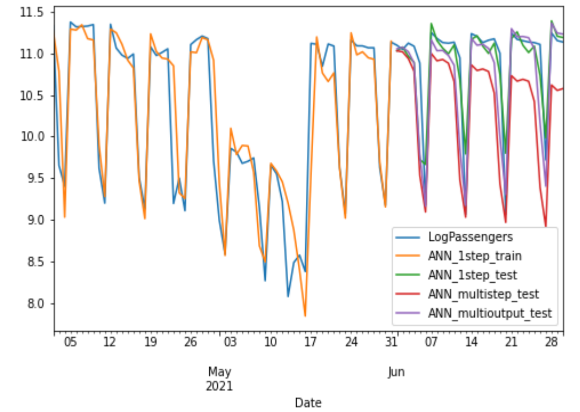
	LR	SVR
Pre-Pandemic		
In-Pandemic		
Entire Dataset		

Table 2. Comparison of actual values and predicted values by RF and ANN.

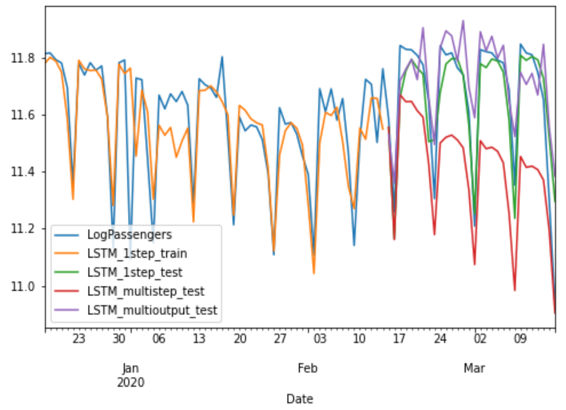
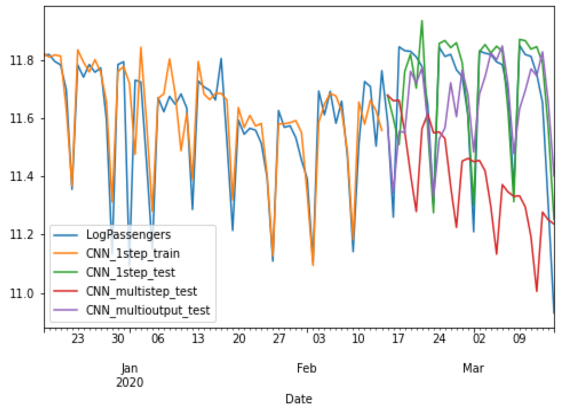
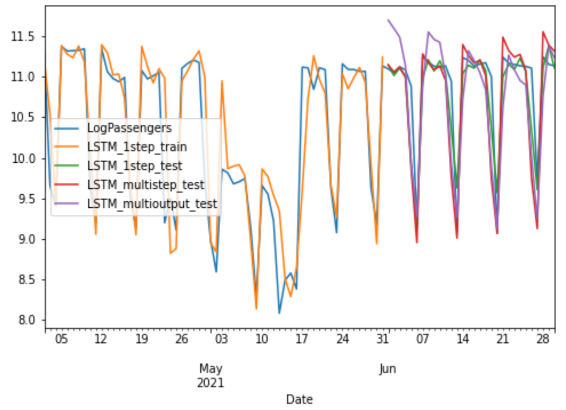
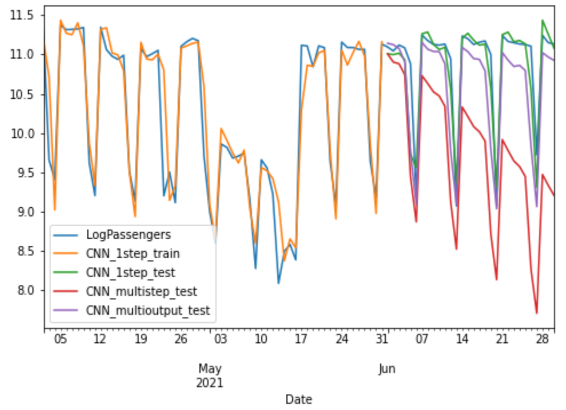
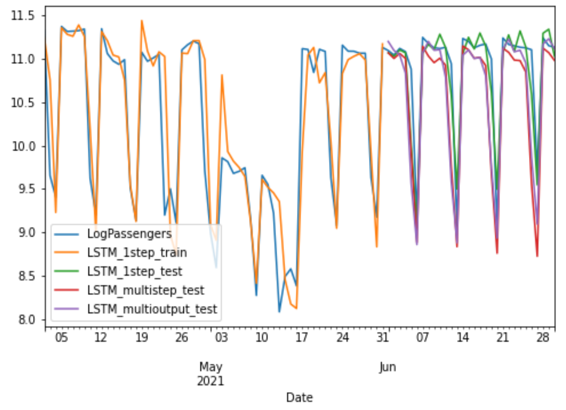
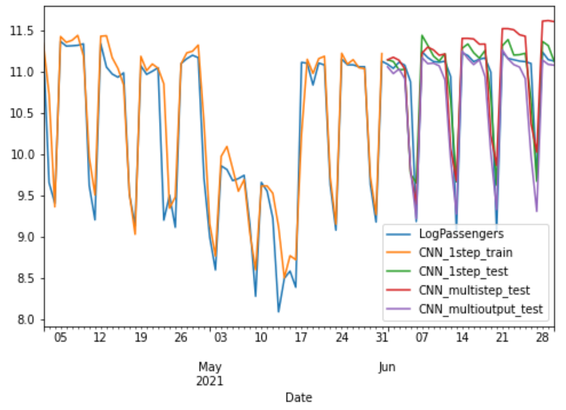
	RF	**ANN**
Pre-Pandemic		
In-Pandemic		
Entire Dataset		

In this study, both multi-step and multi-output forecasting techniques are applied to the models, as they are designed for 30-day time series forecasting.

Upon examining the figures in Table 1, Table 2, and Table 3 for the pre-pandemic dataset, it is evident that the multi-output performances of all models are superior to their multi-step counterparts. In all models except LR, a clear superiority of multi-output performances over multi-step performances is observed. When comparing the visuals in Table 1, Table 2, and Table 3 in terms of multi-output performance, it is apparent that RF provides the best performance, while LR exhibits the least impressive results.

When comparing the multi-step performances of the methods on the pre-pandemic dataset, it is evident that the LR model demonstrates the best performance, while the RF model exhibits the worst. As shown in Table 4 and Table 5, the LR model achieved the best MAPE (1.57) and the best R^2 (0.17) values for multi-step forecasting.

Table 3. Comparison of actual values and predicted values by LSTM and CNN.

<table>
<tr><th></th><th>LSTM</th><th>CNN</th></tr>
<tr><td>Pre-Pandemic</td><td>
</td><td>
</td></tr>
<tr><td>In-Pandemic</td><td>
</td><td>
</td></tr>
<tr><td>Entire Dataset</td><td>
</td><td>
</td></tr>
</table>

Conversely, the RF model's performance is the poorest, as indicated by its MAPE (4.30) and R^2 (-4.951) values.

Additionally, when considering the multi-output values in Table 4 and Table 5, it is observed that the RF model achieves the best MAPE (0.83) and R^2 (0.616) values, while the LR model shows the worst performance with a MAPE of 1.39 and an R^2 of 0.366.

Upon examining the visuals in Table 1, Table 2, and Table 3 for the in-pandemic dataset, it is noted that the multi-step performance of the SVR model is significantly better than its multi-output performance, whereas the multi-output performances of the RF and CNN models are notably better than their multi-step counterparts.

Analyzing the values in Table 4 and Table 5 reveals that the multi-output performance of the LR and ANN models is superior to their multi-step performance, while the LSTM model performs better in multi-step forecasting than in multi-output forecasting. For the in-pandemic dataset, the ANN model achieves the best multi-output performance with a MAPE value of 2.11, and the RF model attains an R^2 value of 0.644. The best multi-step performances are achieved by the ANN model with a MAPE of 2.37 and the LR model with an R^2 of 0.498. The SVR model exhibits the worst multi-output performance with a MAPE of 5.39 and an R^2 of -0.35, whereas the RF model shows the poorest multi-step performance with a MAPE of 17.44 and an R^2 of -11.547.

When analyzing the visuals in Table 1, Table 2, and Table 3 for the entire dataset, it is observed that the multi-output performances of all models surpass their multi-step counterparts, except for the SVR model. These tables also indicate that the RF model demonstrates the worst multi-step performance. Upon examining the MAPE and R^2 values in Table 4 and Table 5, it is seen that the LSTM model, with a MAPE of 2.99, and the CNN model, with an R^2 value of 0.55, achieve the best multi-step performance. Conversely, the RF model records the worst multi-step performance, with a MAPE of 10.86 and an R^2 value of -3.696.

It is observed that the CNN model, with a MAPE value of 2.10, and the RF model, with an R^2 value of 0.706, exhibit the best multi-output performance. Conversely, the SVR model shows the worst multi-output performance, with a MAPE value of 8.32 and an R^2 value of -1.816.

MAPE is a suitable metric for benchmarking performance across different datasets. Therefore, performance comparisons of pre-pandemic, in-pandemic, and entire datasets are made based on MAPE values. Table 4 indicates that the best MAPE value for the pre-pandemic dataset is 0.83 for the in-pandemic dataset, it is 2.11; and for the entire dataset, it is 2.10. Evaluating the best MAPE values from these three different datasets suggests that the COVID-19 pandemic has negatively impacted forecasting performance. Additionally, Tables 4 and 5 reveal that all the best results are achieved using the multi-output method. Although the multi-step performance of some models surpasses their multi-output performance, the multi-output method generally performs much better and has achieved the best results. Therefore, choosing the multi-output technique would be a more logical approach.

Table 4. MAPE results of different forecasting methods.

		LR(%)	SVR(%)	RF(%)	ANN(%)	LSTM(%)	CNN(%)
Pre-Pandemic	**Multi-step**	1.57	1.77	4.30	2.39	2.17	2.99
	Multi-output	1.39	1.09	0.83	1.14	0.95	1.13
In-Pandemic	**Multi-step**	2.82	3.19	17.44	2.37	2.45	10.53
	Multi-output	2.38	5.39	2.44	2.11	3.41	2.92
Entire Dataset	**Multi-step**	3.24	6.03	10.86	4.75	2.99	3.15
	Multi-output	2.75	8.32	2.19	2.17	2.80	2.10

Table 5. R^2 results of different forecasting methods.

		LR	SVR	RF	ANN	LSTM	CNN
Pre-Pandemic	**Multi-step**	0.170	-0.066	-4.951	-0.903	-0.375	-1.787
	Multi-output	0.366	0.518	0.616	0.419	0.597	0.434
In-Pandemic	**Multi-step**	0.498	0.099	-11.547	0.488	0.467	-3.249
	Multi-output	0.504	-0.350	0.644	0.577	0.430	0.439
Entire Dataset	**Multi-step**	0.459	-0.584	-3.696	-0.119	0.353	0.550
	Multi-output	0.449	-1.816	0.706	0.617	0.383	0.609

4.1. Benefits and Advantages of the Study

This study presents several significant benefits and advantages in the field of passenger flow forecasting on trams using machine learning algorithms. The key benefits of this research can be summarized as follows:

- **Enhanced Forecasting Accuracy:** By employing a range of machine learning algorithms, including LR, SVR, RF, ANN, CNN, and LSTM, this study demonstrates improved accuracy in forecasting tram passenger flow. The use of MAPE and R^2 metrics for performance evaluation further substantiates the reliability of the forecasts.

- **Innovative Data Preprocessing Techniques:** The application of log transform and first-order difference techniques for data preprocessing is a novel approach in this field. These methods effectively handle the challenges of large fluctuations, non-linearity, and periodicity in passenger numbers, leading to more accurate forecasting models.
- **Comprehensive Analysis of Pandemic Impact:** The study provides a thorough analysis of the impact of the COVID-19 pandemic on passenger flow forecasting. By dividing the dataset into pre-pandemic, in-pandemic, and entire dataset periods, the research offers valuable insights into how extraordinary events can affect passenger behavior and forecasting accuracy.
- **Practical Implications for Urban Transportation Management:** The findings of this study have practical implications for urban transportation planning and management. Accurate forecasting models can assist in efficient resource allocation, network design, and frequency setting of tram services, contributing to better urban transportation systems.
- **Methodological Contributions:** The study contributes methodologically to the field by comparing multi-step and multi-output forecasting techniques. This comparative analysis not only enhances the understanding of different forecasting methods but also guides future research in selecting appropriate techniques based on specific requirements.
- **Data Set Creation and Sharing:** The creation and public sharing of the Kayseri tram passenger flow dataset is a valuable contribution. It not only facilitates further research in this area but also promotes transparency and reproducibility in scientific studies.
- **Cross-Disciplinary Applicability:** While focused on tram passenger flow, the methodologies and findings of this study have potential applicability in other domains facing similar forecasting challenges, thereby extending its impact beyond the field of urban transportation.

This comprehensive approach to forecasting tram passenger flow using machine learning techniques underscores the study’s contribution to both the academic and practical realms of intelligent transportation systems. The advantages highlighted in this section demonstrate the study’s relevance, innovation, and applicability in addressing complex challenges in passenger flow forecasting.

5. Conclusion

In this study, daily passenger flow data of the Kayseri tram between January 1, 2018 and July 1, 2021 is used, and three different datasets are created using the pre-pandemic period, the pandemic period, and the entire dataset in order to observe the effects of the COVID-19 pandemic on the forecasting performance. Log transform and first-order difference techniques are applied to the datasets, respectively, in order to clear the noise in the data, obtain a more linear structure, and eliminate the seasonal structure from the data. LR, SVM, RF, ANN, CNN, and LSTM models are trained on three different datasets to estimate the passenger numbers for the next 30 days by looking at the passenger numbers of the previous 10 days. Multi-step and multi-output techniques are used to estimate the last 30 days.

From the results obtained, it is understood that the multi-output technique has superior performance to the multi-step technique and that the COVID-19 pandemic has a negative effect on forecasting performance. According to the R^2 metric, the RF model performs best on all datasets. According to the MAPE metric, RF on the pre-pandemic dataset, ANN on the in-pandemic dataset, and CNN on the entire dataset show the best performances. It is understood that the RF model, which uses the multi-output technique in the normal order, that is, in the pre-pandemic period, can predict the number of 30-day Kayseri passengers with quite high success.

In the future, efforts to develop machine learning methods combined with meta-heuristics and bio-inspired algorithms will be important in terms of both reducing the negative impact of the COVID-19 pandemic on forecasting performance and improving success.

Data availability. The dataset investigated in this study is available at

https://raw.githubusercontent.com/kagandedeturk/TimeSeries/main/Tramvay.csv .

Conflict of Interest Statement

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The study is complied with research and publication ethics.

References

[1] D. Li *et al.*, "Percolation transition in dynamical traffic network with evolving critical bottlenecks," *Proceedings of the National Academy of Sciences*, vol. 112, no. 3, pp. 669–672, 2014. doi:10.1073/pnas.1419185112.

[2] M. Ni, Q. He, and J. Gao, "Forecasting the subway passenger flow under event occurrences with social media," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2016. doi:10.1109/tits.2016.2611644.

[3] J. Yu, "Short-Term Airline Passenger Flow Prediction Based on the Attention Mechanism and Gated Recurrent Unit Model," *Cognitive Computation*, vol. 14, no. 2, pp. 693–701, 2022. doi:10.1007/s12559-021-09991-x .

[4] A. Kanavos, F. Kounelis, L. Iliadis, and C. Makris, "Deep Learning Models for Forecasting Aviation Demand Time Series," *Neural Computing and Applications*, vol. 33, no. 23, pp. 16329–16343, 2021. doi:10.1007/s00521-021-06232-y

[5] C. Li, X. Wang, Z. Cheng, and Y. Bai, "Forecasting Bus Passenger Flows by Using a Clustering-Based Support Vector Regression Approach," *IEEE Access*, vol. 8, pp. 19717–19725, 2020, doi: 10.1109/access.2020.2967867.

[6] S. Anvari, S. Tuna, M. Canci, and M. Turkay, "Automated Box–Jenkins Forecasting Tool With an Application for Passenger Demand in Urban Rail Systems," *Journal of Advanced Transportation*, vol. 50, no. 1, pp. 25–49, Sep. 2015, doi: 10.1002/atr.1332.

[7] A. Samagaio and M. Wolters, "Comparative Analysis of Government Forecasts for the Lisbon Airport," *Journal of Air Transport Management*, vol. 16, no. 4, pp. 213–217, Jul. 2010, doi: 10.1016/j.jairtraman.2009.09.002.

[8] M. Milenković, L. Švadlenka, V. Melichar, N. Bojović, and Z. Avramović, "Sarima Modelling Approach for Railway Passenger Flow Forecasting," *Transport*, pp. 1–8, Mar. 2016, doi: 10.3846/16484142.2016.1139623.

[9] X. Xiao, J. Yang, S. Mao, and J. Wen, "An Improved Seasonal Rolling Grey Forecasting Model Using a Cycle Truncation Accumulated Generating Operation for Traffic Flow," *Applied Mathematical Modelling*, vol. 51, pp. 386–404, Nov. 2017, doi: https://doi.org/10.1016/j.apm.2017.07.010.

[10] A. Stathopoulos and M. G. Karlaftis, "A Multivariate State Space Approach for Urban Traffic Flow Modeling and Prediction," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 2, pp. 121–135, Apr. 2003, doi: https://doi.org/10.1016/s0968-090x(03)00004-4.

[11] X. Tang and G. Deng, "Prediction of Civil Aviation Passenger Transportation Based on ARIMA Model," *Open Journal of Statistics*, vol. 06, no. 05, pp. 824–834, 2016, doi: https://doi.org/10.4236/ojs.2016.65068.

[12] N. K. Chauhan and K. Singh, "A Review on Conventional Machine Learning vs Deep Learning," in 2018 *International Conference on Computing, Power and Communication Technologies (GUCON)*, Sep. 2018, doi: https://doi.org/10.1109/gucon.2018.8675097.

[13] Y. Park, Y. Choi, K. Kim, and J. K. Yoo, "Machine Learning Approach for Study on Subway Passenger Flow," *Scientific Reports*, vol. 12, no. 1, p. 2754, Feb. 2022, doi: https://doi.org/10.1038/s41598-022-06767-7.

[14] F. Moretti, S. Pizzuti, S. Panzieri, and M. Annunziato, "Urban Traffic Flow Forecasting Through Statistical and Neural Network Bagging Ensemble Hybrid Modeling," *Neurocomputing*, vol. 167, pp. 3–7, Nov. 2015, doi: https://doi.org/10.1016/j.neucom.2014.08.100.

[15] B. Sun, W. Cheng, P. Goswami, and G. Bai, "Short-Term Traffic Forecasting using Self-Adjusting K-Nearest Neighbours," *IET Intelligent Transport Systems*, vol. 12, no. 1, pp. 41–48, Feb. 2018, doi: https://doi.org/10.1049/iet-its.2016.0263.

[16] J. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman Filter Approach for Stochastic Short-Term Traffic Flow Rate Prediction and Uncertainty Quantification," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 50–64, Jun. 2014, doi: https://doi.org/10.1016/j.trc.2014.02.006.

[17] Y. Sun, B. Leng, and W. Guan, “A Novel Wavelet-Svm Short-Time Passenger Flow Prediction in Beijing Subway System,” *Neurocomputing*, vol. 166, pp. 109–121, Oct. 2015, doi: https://doi.org/10.1016/j.neucom.2015.03.085.

[18] L. Li, X. Chen, and L. Zhang, “Multimodel Ensemble for Freeway Traffic State Estimations,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 3, pp. 1323–1336, Jun. 2014, doi: https://doi.org/10.1109/tits.2014.2299542.

[19] J. Zhang, F. Chen, Z. Cui, Y. Guo, and Y. Zhu, “Deep Learning Architecture for Short-Term Passenger Flow Forecasting in Urban Rail Transit,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7004–7014, Nov. 2021, doi: https://doi.org/10.1109/tits.2020.3000761.

[20] X. Feng, T. Gan, and X. Wang, “Feedback Analysis of Interaction between Urban Densities and Travel Mode Split,” *International Journal of Simulation Modelling*, vol. 14, no. 2, pp. 349–358, Jun. 2015, doi: 10.2507/ijsimm14(2)co9.

[21] Y. Wei and M.-C. Chen, “Forecasting The Short-Term Metro Passenger Flow with Empirical Mode Decomposition and Neural Networks,” *Transportation Research Part C: Emerging Technologies*, vol. 21, no. 1, pp. 148–162, Apr. 2012, doi: https://doi.org/10.1016/j.trc.2011.06.009.

[22] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, “Long Short-Term Memory Neural Network for Traffic Speed Prediction Using Remote Microwave Sensor Data,” *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, May 2015, doi: https://doi.org/10.1016/j.trc.2015.03.014.

[23] P. He, “Optimization and Simulation of Remanufacturing Production Scheduling under Uncertainties,” *International Journal of Simulation Modelling*, vol. 17, no. 4, pp. 734–743, Dec. 2018, doi: https://doi.org/10.2507/ijsimm17(4)co20.

[24] M. Castro-Neto, Y.-S. Jeong, M.-K. Jeong, and L. D. Han, “Online-Svr for Short-Term Traffic Flow Prediction under Typical and Atypical Traffic Conditions,” *Expert Systems with Applications*, vol. 36, no. 3, pp. 6164–6173, Apr. 2009, doi: https://doi.org/10.1016/j.eswa.2008.07.069.

[25] G. Xie, S. Wang, Y. Zhao, and K. K. Lai, “Hybrid Approaches Based on Lssvr Model for Container Throughput Forecasting: A Comparative Study,” *Applied Soft Computing*, vol. 13, no. 5, pp. 2232–2241, May 2013, doi: https://doi.org/10.1016/j.asoc.2013.02.002.

[26] X. Wang, N. Zhang, Y. Zhang, and Z. Shi, “Forecasting of Short-Term Metro Ridership with Support Vector Machine Online Model,” *Journal of Advanced Transportation*, vol. 2018, pp. 1–13, Jun. 2018, doi: https://doi.org/10.1155/2018/3189238.

[27] A. Singhal, C. Kamga, and A. Yazici, “Impact of Weather on Urban Transit Ridership,” *Transportation Research Part A: Policy and Practice*, vol. 69, pp. 379–391, Nov. 2014, doi: https://doi.org/10.1016/j.tra.2014.09.008.

[28] A. Koesdwiady, R. Soua, and F. Karray, “Improving Traffic Flow Prediction With Weather Information in Connected Cars: A Deep Learning Approach,” *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9508–9517, Dec. 2016, doi: https://doi.org/10.1109/tvt.2016.2585575.

[29] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, “Multimodal Deep Support Vector Classification with Homologous Features and Its Application to Gearbox Fault Diagnosis,” *Neurocomputing*, vol. 168, pp. 119–127, Nov. 2015, doi: https://doi.org/10.1016/j.neucom.2015.06.008

[30] M.-L. Huang, “Intersection Traffic Flow Forecasting Based on N-Gsvr with a New Hybrid Evolutionary Algorithm,” *Neurocomputing*, vol. 147, pp. 343–349, Jan. 2015, doi: https://doi.org/10.1016/j.neucom.2014.06.054

[31] Y. Zhang and Y. Liu, “Analysis of Peak and Non-Peak Traffic Forecasts Using Combined Models,” *Journal of Advanced Transportation*, vol. 45, no. 1, pp. 21–37, Sep. 2010, doi: https://doi.org/10.1002/atr.128.

[32] M. Khashei and M. Bijari, “A Novel Hybridization of Artificial Neural Networks and Arıma Models for Time Series Forecasting,” *Applied Soft Computing*, vol. 11, no. 2, pp. 2664–2675, Mar. 2011, doi: https://doi.org/10.1016/j.asoc.2010.10.015.

[33] T. Pant, C. Han, and H. Wang, “Examination of Errors of Table Integration in Flamelet/Progress Variable Modeling of a Turbulent Non-Premixed Jet Flame,” *Applied Mathematical Modelling*, vol. 72, pp. 369–384, Aug. 2019, doi: https://doi.org/10.1016/j.apm.2019.03.016.

[34] X. Wang, L. Huang, H. Huang, B. Li, Z. Xia, and J. Li, "An Ensemble Learning Model for Short-Term Passenger Flow Prediction," *Complexity*, vol. 2020, pp. 1–13, Dec. 2020, doi: https://doi.org/10.1155/2020/6694186.

[35] H. Li, J. Bai, and Y. Li, "A Novel Secondary Decomposition Learning Paradigm with Kernel Extreme Learning Machine for Multi-Step Forecasting of Container Throughput," *Physica A: Statistical Mechanics and its Applications*, vol. 534, p. 122025, Nov. 2019, doi: https://doi.org/10.1016/j.physa.2019.122025.

[36] M. Niu, Y. Hu, S. Sun, and Y. Liu, "A Novel Hybrid Decomposition-Ensemble Model Based on Vmd and Hgwo for Container Throughput Forecasting," *Applied Mathematical Modelling*, vol. 57, pp. 163–178, May 2018, doi: https://doi.org/10.1016/j.apm.2018.01.014.

[37] Y. Li and C. Ma, "Short-Time Bus Route Passenger Flow Prediction Based on a Secondary Decomposition Integration Method," *Journal of Transportation Engineering, Part A: Systems*, vol. 149, no. 2, Feb. 2023, doi: https://doi.org/10.1061/jtepbs.teeng-7496.

[38] H. Nguyen, L. Kieu, T. Wen, and C. Cai, "Deep Learning Methods in Transportation Domain: A Review," *IET Intelligent Transport Systems*, vol. 12, no. 9, pp. 998–1004, Jul. 2018, doi: https://doi.org/10.1049/iet-its.2018.0064.

[39] H. Gao, X. Qin, R. J. D. Barroso, W. Hussain, Y. Xu, and Y. Yin, "Collaborative Learning-Based Industrial IoT API Recommendation for Software-Defined Devices: The Implicit Knowledge Discovery Perspective," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 1, pp. 66–76, Feb. 2022, doi: https://doi.org/10.1109/TETCI.2020.3023155.

[40] H. Gao, Y. Zhang, H. Miao, R. J. D. Barroso, and X. Yang, "SDTIOA: Modeling the Timed Privacy Requirements of IoT Service Composition: A User Interaction Perspective for Automatic Transformation from BPEL to Timed Automata," *Mobile Networks and Applications*, vol. 26, no. 6, pp. 2272–2297, Nov. 2021, doi: https://doi.org/10.1007/s11036-021-01846-x.

[41] A. A. Amer, I. E. Talkhan, R. Ahmed, and T. Ismail, "An Optimized Collaborative Scheduling Algorithm for Prioritized Tasks with Shared Resources in Mobile-Edge and Cloud Computing Systems," *Mobile Networks and Applications*, Apr. 2022, doi: https://doi.org/10.1007/s11036-022-01974-y.

[42] X. Ma, H. Xu, H. Gao, and M. Bian, "Real-time Multiple-Workflow Scheduling in Cloud Environment," *Research Square*, Feb. 2021, Published, doi: 10.21203/rs.3.rs-170491/v1.

[43] K. Peng, W. Bai, and L. WU, "Passenger Flow Forecast of Railway Station Based on Improved Lstm," 2020 *2nd International Conference on Advances in Computer Technology, Information Science and Communications (CTISC)*, Mar. 2020, doi: https://doi.org/10.1109/ctisc49998.2020.00033.

[44] Dedeturk, B.K.: Dataset. https://raw.githubusercontent.com/kagandedeturk/TimeSeries/main/Tramvay.csv (accessed May 12, 2023).

[45] W. Xu, H. Peng, X. Zeng, F. Zhou, X. Tian, and X. Peng, "A Hybrid Modelling Method for Time Series Forecasting Based nn a Linear Regression Model and Deep Learning," *Applied Intelligence*, vol. 49, no. 8, pp. 3002–3015, Feb. 2019, doi: https://doi.org/10.1007/s10489-019-01426-3.

[46] Y. Bai et al., "A Comparison of Dimension Reduction Techniques for Support Vector Machine Modeling of Multi-Parameter Manufacturing Quality Prediction," *Journal of Intelligent Manufacturing*, vol. 30, no. 5, pp. 2245–2256, Jan. 2018, doi: https://doi.org/10.1007/s10845-017-1388-1.

[47] X. Qiu, L. Zhang, P. Nagaratnam Suganthan, and G. A. J. Amaratunga, "Oblique Random Forest Ensemble via Least Square Estimation for Time Series Forecasting," Information Sciences, vol. 420, pp. 249–262, Dec. 2017, doi: https://doi.org/10.1016/j.ins.2017.08.060.

[48] T. Xiong, Y. Bao, and Z. Hu, "Beyond One-Step-Ahead Forecasting: Evaluation of Alternative Multi-Step-Ahead Forecasting Models for Crude Oil Prices," *Energy Economics*, vol. 40, pp. 405–415, Nov. 2013, doi: https://doi.org/10.1016/j.eneco.2013.07.028.

[49] H. I. Kazıcı, S. Kosunalp, and M. Arucu, "Its-Pro-Flow: A New Enhanced Short-Term Traffic Flow Prediction For Intelligent Transportation Systems," *Scientific Journal of Silesian University of Technology. Series Transport*, vol. 120, pp. 117–136, Sep. 2023, doi: 10.20858/sjsutst.2023.120.8.