DOI:10.25092/baunfbed. 937333 *J. BAUN Inst. Sci. Technol., 24(1), 300-312, (2022)*

Prediction of travel time for railway traffic management by using the AdaBoost algorithm

Mehmet Taciddin AKÇAY[1*], Abdurrahim AKGUNDOĞDU[2], Hasan TİRYAKİ[2]

[1]*Department of Electrical-Elektronics Engineering, Faculty of Engineering, Halic University, Istanbul, Turkey.*
[2]*Department of Electrical-Elektronics Engineering, Faculty of Engineering, Istanbul-Cerrahpasa University, Istanbul, Turkey.*

Geliş Tarihi (Received Date): 14.05.2021
Kabul Tarihi (Accepted Date): 18.11.2021

Abstract

While determining the travel time between stations, a number of design parameters such as waiting time, motion resistance, slope, curve, traction force, maximum speed, vehicle mass, and distance between two stations are taken into consideration. These parameters form the infrastructure of the system definition of the motion of the vehicle. Furthermore, while creating the speed profile, special attention should be paid to the travel time in order to ensure the defined headway for the line. In this study, the travel time value between stations for intracity metro stations was predicted using the adaptive boosting method, which is one of the machine learning methods, and compared with various well-known methods. The data used were applied to the proposed model with the cross-validation and random sampling hold-out methods, and the values of the coefficient of determination (R^2) were calculated.

Keywords: *AdaBoost, machine learning, railway, travel time, signalization*

AdaBoost algoritmasını kullanarak demiryolu trafik yönetimi için seyir süresinin tahmini

Öz

İstasyonlar arasında geçen seyir süresi belirlenirken bekleme süresi, hareket direnci, eğim, kurp, cer kuvveti, maximum hız, aracın kütlesi ve iki istasyon arası mesafe gibi bir takım tasarım parametreleri göz önünde bulundurulmaktadır. Bu parametreler

[*]Mehmet Taciddin AKÇAY, mehmettaciddinakcay@halic.edu.tr, https://orcid.org/0000-0002-1050-4566
Abdurrahim AKGUNDOGDU, akgundogdu@iuc.edu.tr, https://orcid.org/0000-0001-8113-0277
Hasan TİRYAKİ, hasan.tiryaki@iuc.edu.tr, https://orcid.org/0000-0001-9175-0269

aracın hareketine ait sistemin tanımının alt yapısını oluşturmaktadır. Ayrıca, hız profili oluşturulurken hat için tanımlanmış sefer sıklığının sağlanabilmesi için seyir süresine özellikle dikkat edilmelidir. Bu çalışmada şehiriçi metro sistemlerine ait istasyonlar arası seyir süresi değerinin makine öğrenmesi yöntemlerinden Adaptive Boosting yöntemi ile tahmini gerçekleştirilmiş ve iyi bilinen çeşitli yöntemler ile karşılaştırılmıştır. Kullanılan veriler çapraz doğrulama ve rastgele örnekleme yöntemleri ile önerilen modele uygulanmış ve belirleme katsayısı (R^2) değerleri hesaplanmıştır.

Anahtar kelimeler: *AdaBoost, demiryolu, makine öğrenmesi, seyir süresi, sinyalizasyon.*

1. Introduction

Rail systems are complex systems that include many sub-branches in construction and electromechanics. While conducting transportation surveys and feasibility studies, rail system lines are planned according to population density and passenger movement. Performance criteria related to the designed line emerge as a result of these studies. The number of vehicles planned for the line, capacity, and other features are determined in this way. It is aimed to carry out the operation effectively and efficiently in order to realize the passenger movement, which is the main duty of the line, in the most appropriate way. It is the subject of the signalization system to transport the passenger load safely on railways in the desired time and at the desired headway. The main goal of this system is to reach the desired passenger capacity target. The signalization system is designed to comply with the highest safety requirements, such as SIL4 (Safety Integrity Level), taking into account the European standard CENELEC. It is carried out with an uninterrupted exchange of information between the on-board systems and equipment in order to guide operations for the company. When designing the signalization system, the minimum lifetime of the system is determined to be 30 years. The signalization system is used to control train movement and to ensure train safety. In this way, necessary arrangements are made for the operation traffic. The most important parameter for the signalization system is defined as the headway. The headway is the time interval measured from front bumper to front bumper between two vehicles passing in one direction of a line. The headway determines the passage possibilities of vehicles at the intersection or ramp points. Speed and headway information constitutes important information for traffic engineering and is required for designing high-performance transportation systems [1-7]. The speed characteristic of the vehicle also constitutes travel time. In order to simulate the signalization system, travel time is calculated by creating vehicle speed profiles. While creating speed profiles, speed restrictions in the related part, switching layouts, and other parameters are taken into consideration. Since the slope and curve placements of the line affect vehicle motion, various limitations can occur in the speed profile. After the vehicle speed profiles are created, the operation is planned according to the headway value between vehicles. In Figure 1, the formation of the travel time, which is one of the most important parameters of the signalization system, is expressed.

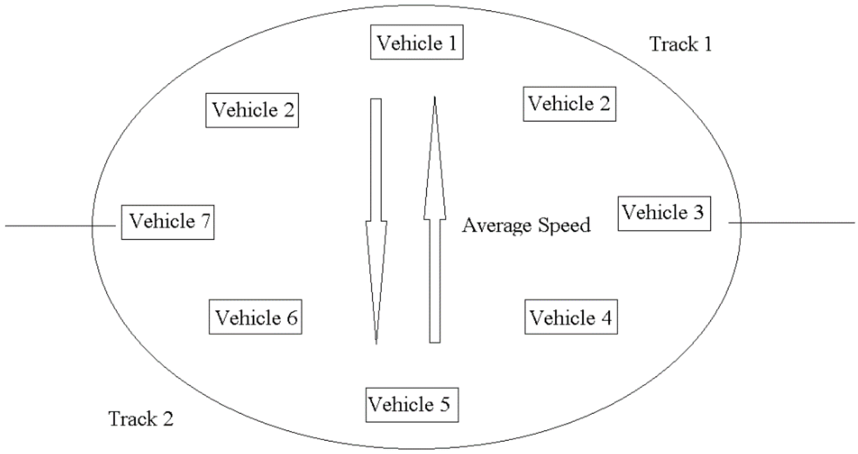

Figure 1. Travel time

The motion of vehicles over a period is expressed in a circular manner in Figure 1. While Line 1 and Line 2 are shown in two hemispheres of the circle, the vehicle placement is made within these limits. Necessary infrastructures for the electrification system should be provided in order to verify the targeted headway of the operation traffic. Other subsystems in the system are designed to be integrated into the whole system. Operational simulations and signalization simulations are performed with the models created by using vehicle-driving algorithms created in the computer environment, and these tests are carried out before the application.

The signalization system manages the driving characteristics of vehicles, and even vehicle motions are performed according to the speed profile defined for the vehicles used. The vehicle signalization system includes speed limitations according to the slope, curve, switching zones, and entrance to the depot area. The graph of the formation of the speed-time and position-time characteristics of vehicles is given in Figure 2.

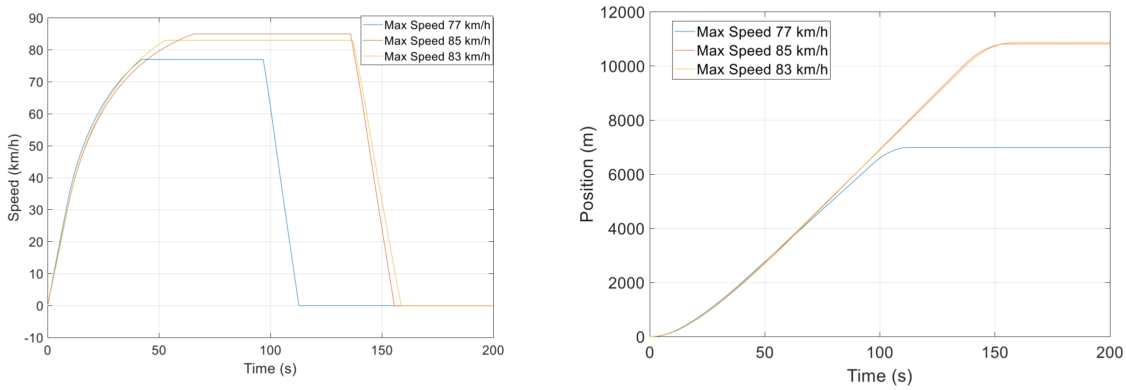

Figure 2. Graphs of speed-time and position-time characteristics

In this case, there are three different speed-position characteristics. They were designed by considering many parameters such as maximum speed that the vehicle can reach, travel time at a constant speed, initial braking moment, and acceleration. In metro systems, the maximum speed varies between 70 km/h and 90 km/h, and this value is determined according to the condition and features of the line.

There are various studies in the literature on the prediction of travel time. In the study [8], the travel time was predicted by taking into account accident situations. In [9], the prediction of the travel time for dynamic route determination was investigated by using the ant colony method. In [10], a work was conducted on the prediction of the travel time using the advanced artificial neural networks method. In [11], the prediction of the travel time was carried out in a long-term and efficient manner with the help of fuzzy logic. In the study [12], the prediction of the travel time was carried out using speed predictions. In [13], the prediction of the travel time for travel sections in planned transportation was studied. In [14], the prediction of the travel time was investigated with the help of GPS-based traveling vehicles. In [15], the models proposed to increase the accuracy of travel time predictions were described with the emphasis on speed predictions. In [16], the prediction of the travel time of Warsaw vehicles was examined. In study [17], the prediction of the travel time on highways was investigated. In [18], the prediction of the travel time was made by using mobile phone GPS data. In [19], real-time travel time prediction was investigated with the help of the Multi-level k-Nearest Neighbor Algorithm and Data Fusion Method. In [20], online travel time prediction was made by using advanced machine learning models. The analysis of the operation traffic is an important factor that determines the characteristics of the traffic flow [21]. The behavior of the traffic flow depends on a number of parameters, such as the flow of water [22]. Since there are various speed applications in the operation, the harmonization of the traffic flow by reducing this diversity is among the subjects studied separately for operational safety [23].

As is known, the travel time includes the waiting times and defines the time elapsed for a full turn of movement between stations. Operation management is carried out by creating a traffic schedule in control centers, where the operational vehicle traffic is managed on rail system lines. The train schedule represents a plan that shows which vehicle will be at which location in which time. Through this plan, information is provided on operational data, including the number of vehicles in operation, headway, travel time, and station locations. Before the signalization system is applied to the line, operational simulation is performed, and the relevant performance details are calculated. Operational simulation is a model in which the operational behavior of the rail system line is tested by obtaining other operational data along with vehicle traffic. Tests for many systems and subsystems such as headway, operation time, number of vehicles, electrification system adequacy, and line parameters are performed in this way. In this way, an integrated operation is ensured by integrating vehicle traffic and all systems. The travel time information to be used in the operation in these tests is of great importance for the optimization of the system designed. For this reason, it is necessary to accurately predict the travel time information to be used in the operation during the simulation stage. In line with this prediction, the system can be designed and implemented in accordance with the schedule. In this work, which was carried out in the light of this information, the travel time value between the stations of intracity metro systems was determined using the adaptive boosting method, which is one of the machine learning methods.

In the second section of this article, the proposed model for the prediction of travel time, the inputs-outputs used to create the structure of the model, and the proposed method were explained. In the third section, the simulation results obtained were presented and compared with various well-known methods. In the conclusion section, the evaluation of the article was examined in detail.

2. Material and method

2.1. Simulation model and dataset

Some equations are used in the mathematical expression of the traction force of the vehicle given with (1).

$$\mathbf{F_{traction}} = \mathbf{F_{motion}} + \mathbf{F_{slope}} + \mathbf{F_{curve}} + \mathbf{ma} \quad (1).$$

$F_{traction}$ expresses the traction force, while acceleration is expressed by the a value. The forces acting on the vehicle are calculated with equation (2):

$$\sum_{i=1}^{n} F_i = m' \times a \quad (2).$$

Within the scope of this work, a dataset consisting of 500 samples, each with a different value, obtained using different operational data was used. From these data, the data such as Waiting Time, Motion Resistance, Slope, Curve, Traction Force, Maximum Speed, Vehicle Mass, and Distance Between Two Stations constitute the input of the proposed method. The travel time parameter was selected as the output of the system. The statistical distribution of these data is shown in Figure 3. Here, the dark blue vertical line shows the mean and standard deviation. The blue area represents the values between the first and third quarters.

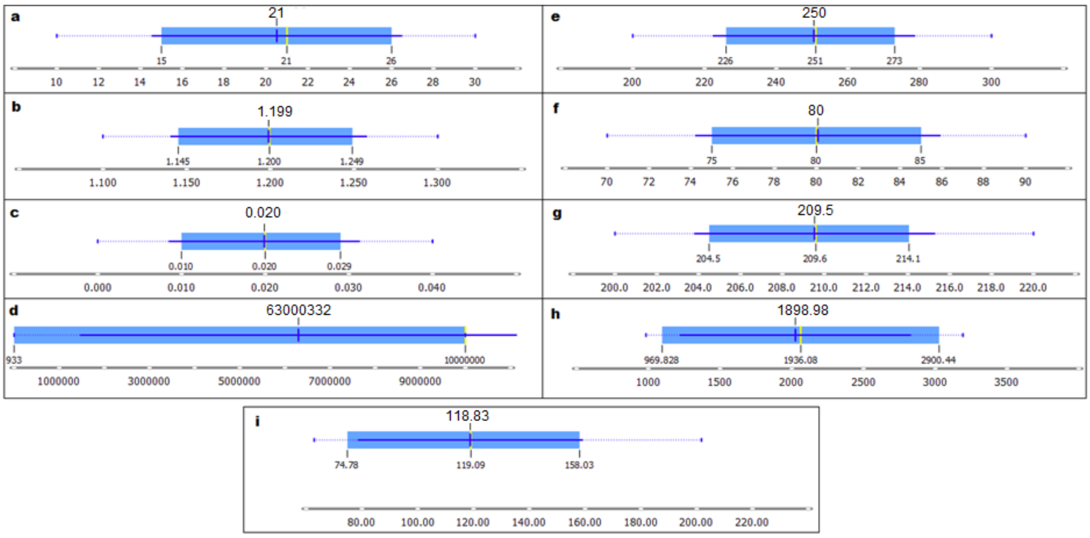

Figure 3. Distribution of input and output parameters that form the model. Input parameters: (a) Waiting Time (b) Motion Resistance, (c) Slope, (d) Curve, (e) Traction Force, (f) Maximum Speed, (g) Vehicle Mass, and (h) Distance Between Two Stations. Output parameter: (i) Travel time.

Random variables were used for all the data used in the simulation. The simulation screen is shown in Figure 4.

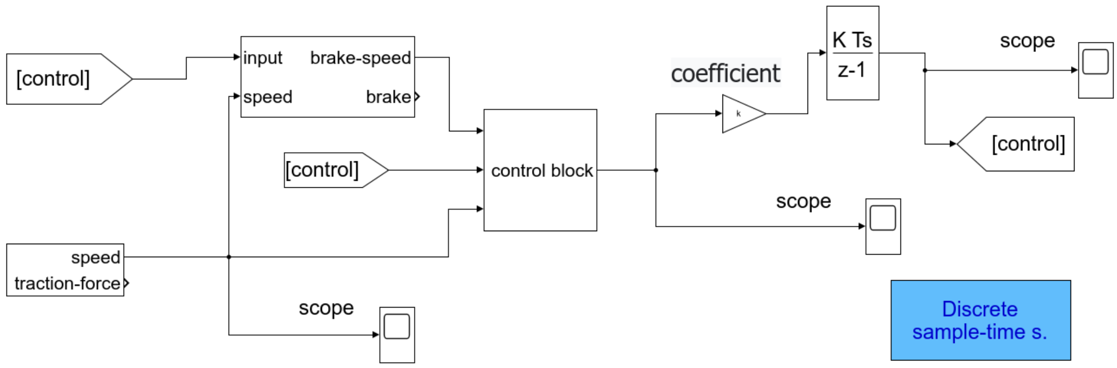

Figure 4. Simulation screen

The sampling time was chosen from 1 to 3 seconds to obtain the precision required for the simulation. In the simulation, the vehicle maximum speed range was created with values ranging from 70 km/h to 90 km/h. Due to the diversity of parameters in simulation and variability in operation conditions, the use of machine learning methods is suitable for the solution of this problem.

2.2 Adaptive boosting

The Adaptive Boosting (AdaBoost) algorithm, developed by Freund and Schapire in 1995, solved most of the practical difficulties of previously used boosting algorithms. The AdaBoost steps are given in Algorithm 1. The algorithm takes as an input a training set $(x_1, y_1), \ldots, (x_m, y_m)$ in which each xi belongs to a domain or sample area (X) and each yi tag is in a tag set $(Y = \{-1, +1\})$. AdaBoost repeatedly calls a given weak or base learning algorithm in a series of $t = 1, \ldots, T$ turns. One of the main ideas of the algorithm is to provide a distribution or set of weights on the training set. The weight of this distribution on the training sample i in the turn t is indicated as $D_t(i)$. Initially, all weights are set equally, but with each turn, the weights of misclassified samples are increased so that the weak learner has to focus on more difficult samples in the training set [24-25].

Algorithm 1: The boosting algorithm of AdaBoost.

Step 1 Given $(x_1, y_1), \ldots, (x_m, y_m)$ where $x_i \in X, y_i \in Y = \{-1, +1\}$

Step 2 Initialize $D_1(i) = 1/m$

For $t = 1, \ldots, T$:

Step 3 Train weak learner using distribution D_t

Step 4 Get weak hypothesis $h_t: X \rightarrow \{-1, +1\}$ with error $\epsilon_t = P_{r_i \sim D_t}[h_t(x_i) \neq y_i]$

Step 5 Choose $\propto_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$

Step 6 Update: $D_{t+1} = \frac{D_t(i)}{Z_t}\begin{cases} e^{-\propto_t} \; if \; h_t(x_i) = y_i \\ e^{\propto_t} \; if \; h_t(x_i) \neq y_i \end{cases} = \frac{D_t(i)\exp\left(-\propto_t y_i h_t(x_i)\right)}{Z_t}$

where Z_t is a normalization factor (chosen so that D_{t+1} will be a distribution)

Step 7 Output the final hypothesis: $$H(x) = sign\left(\sum_{t=1}^{T} \propto_t h_t(x)\right)$$

The task of the weak learner is to find a weak hypothesis ($h_t : X \rightarrow \{-1, +1\}$) with a distribution of D_t. How good is a weak hypothesis is measured by its error.

$$\epsilon_t = P_{r_i \sim D_t}[h_t(x_i) \neq y_i] = \sum_{i:h_t(x_i) \neq y_i} D_t(i) \qquad (3).$$

The error is measured according to the distribution D_t in which the weak learner is trained. In practice, the weak learner can be an algorithm that can use weights D_t in training samples. Alternatively, when this is not possible, a subset of training samples can be sampled according to D_t, and these (weightless) resampled samples can be used to train the weak learner.

When the weak hypothesis h_t is taken, it selects a $\propto_t$ parameter, as seen in the AdaBoost algorithm. Intuitively, $\propto_t$ measures the importance assigned to h_t. If $\epsilon_t \leq 1/2$, $\propto_t \geq 0$ (can be assumed without the loss of generalization) and $\propto_t$ increases as ϵ_t decreases.

The distribution D_t is then updated using the rule shown in Algorithm 1. The effect of this rule is to increase the weight of the samples classified incorrectly with h_t and to reduce the weight of the samples classified correctly. Thus, the weight tends to concentrate on "more difficult" samples.

The final hypothesis H is the weighted majority vote of the weak hypotheses T. Here, $\propto_t$ is the weight assigned to h_t.

2.3. Performance criteria

In this article, the Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Correlation Coefficient (R^2) values were calculated as performance criteria. The formulae of these criteria are given in Table 1.

Table 1: Performance measurements.

MSE	**MAE**	**RMSE**	**R²**
$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$	$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y}_{avg})^2}$

Here, y_i, $\hat{y}_i$, and $\hat{y}_{avg}$ are the average of the desired output i, the predicted output, and the desired output, respectively. n represents each sample in the dataset.

3. Results and discussion

The dataset of railways was studied, and a prediction method with eight inputs and one output was proposed. In the literature, Decision Trees, kNN, Random Forest (RF), Neural Network (NN), SVM, and the AdaBoost methods recommended in this study were applied using the Orange machine learning program, and the results obtained were examined. The knowledge flow of these methods applied can be observed in Figure 5.

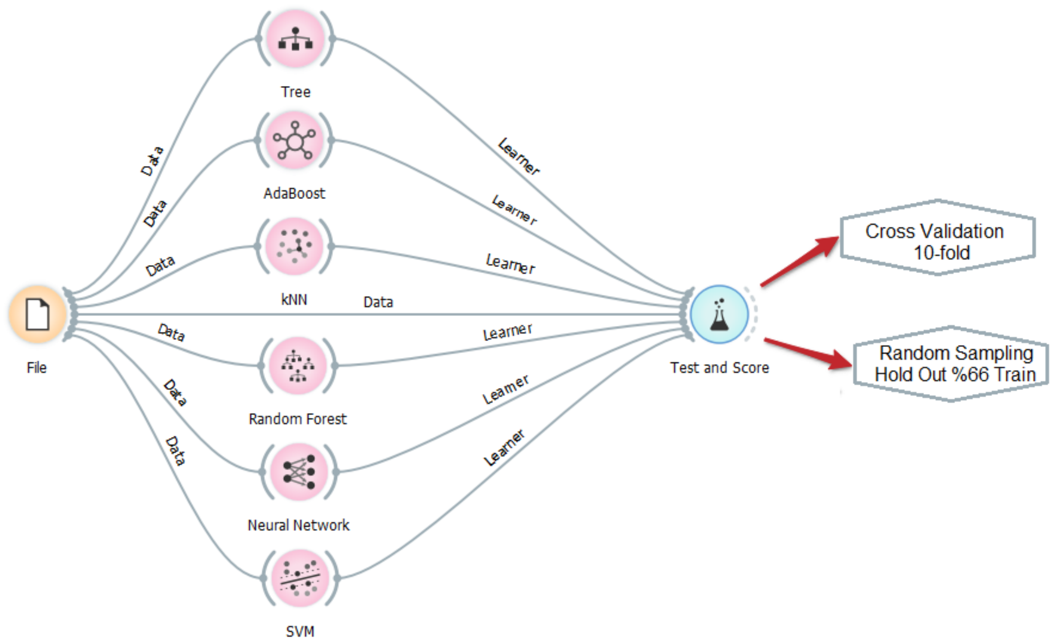

Figure 5. Knowledge flow chart of the compared methods.

Here, the dataset consisting of 500 samples was evaluated in two different ways to avoid the overfitting of the system during the learning process. In the first evaluation, the cross-validation method was used (CV), and in the second evaluation, the random sampling hold-out (RSHO) method was used. In the CV method, the data were divided into ten groups, and nine were used in training, while the remaining group was used in the test process, which was repeated ten times. In the second method, 66% (330 samples) of the total data were used in training, and the remaining data (170 samples) were used in the test process with the RSHO method to calculate the success rate of the system. Figure 6 demonstrates how the CV and RSHO methods were performed.

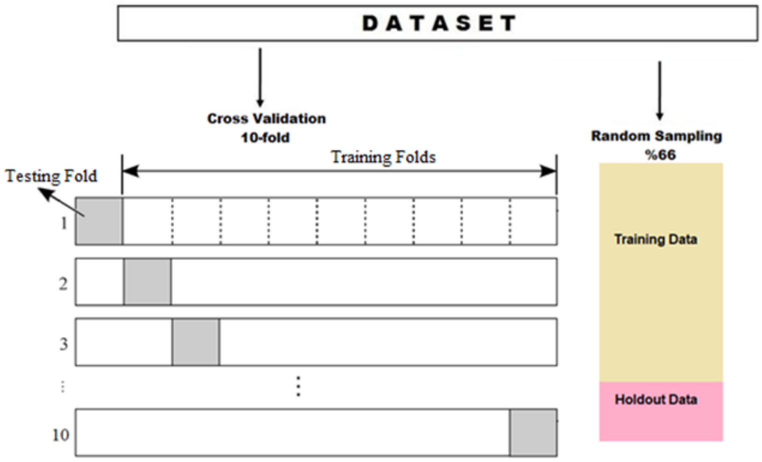

Figure 6. Block diagram of the CV and RSHO methods

In Figure 7, the results obtained for the 170 samples used in the test process performed after the AdaBoost algorithm following RSHO, the real values, and the error rates that occur between them are shown on the graph.

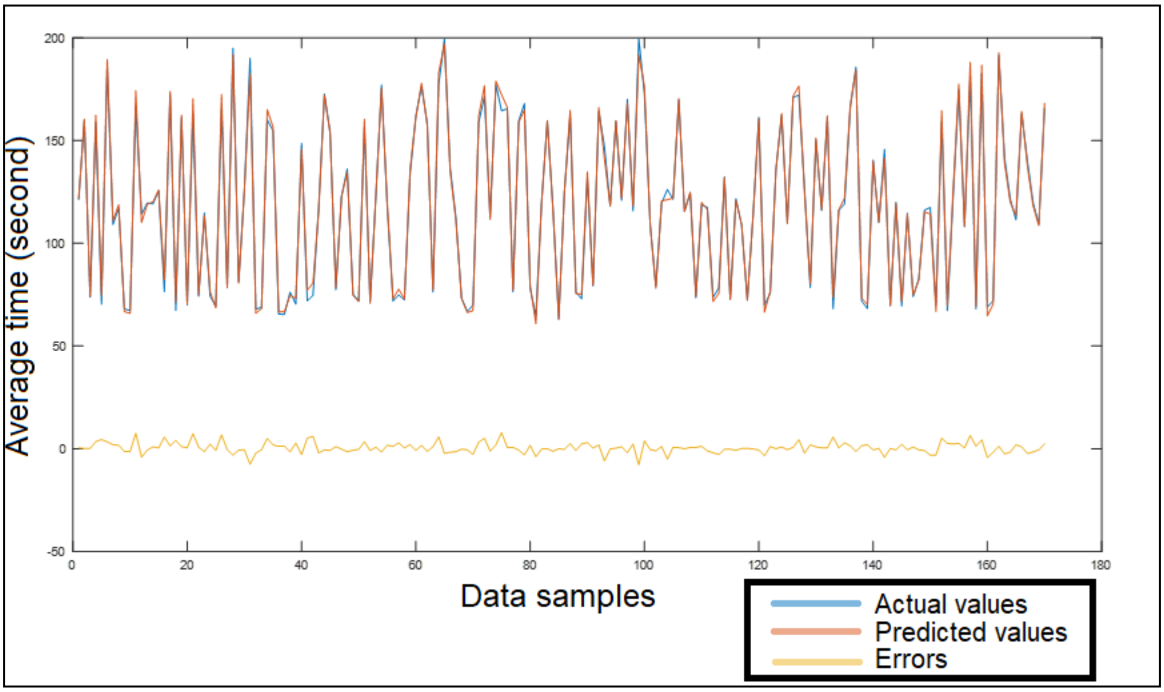

Figure 7. Real values, predicted values, and values of error changes after the RSHO method

Figure 8 shows the regression curves demonstrating the results and real values obtained with the Adaboost, the method proposed.

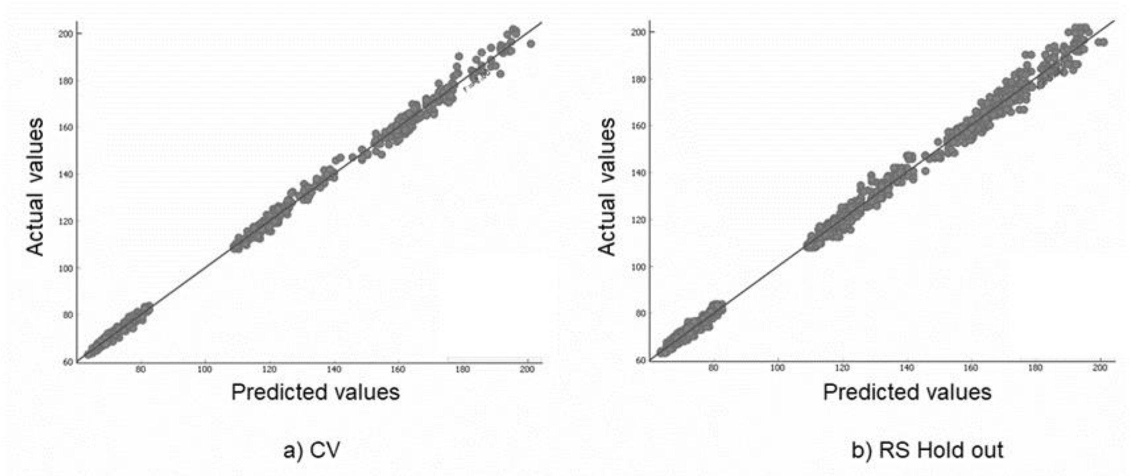

Figure 8. Regression graphs of the results obtained with AdaBoost as a result of the CV (a) and RS Hold out methods (b)

The comparison of the results obtained after different methods is given in Table 2 and Table 3, respectively, after CV and RSHO. The performance values of the Random Forest, Neural Network, kNN, Decision Trees, and SVM methods frequently used in the literature, which are recommended and with which the most successful results are obtained, can be observed in these tables, respectively.

Table 2: Comparison of the predicted performance of the models used with the 10-fold CV method

Model	MSE	RMSE	MAE	R^2
AdaBoost	4.5239935999	2.126968170894	1.5520399999999	0.9971890922375
Random Forest	5.2427362190	2.289702211872	1.6963276072372	0.9967425135317
Neural Network	5.6990668610	2.387271844810	1.8001799961579	0.9964589801191
kNN	8.4055911777	2.899239758588	2.1754466666666	0.9947773264998
Tree	27.872001595	5.279394055744	2.5462483333333	0.9826821979500
SVM	52.867139701	7.270979280781	5.3954551600592	0.9671518869159

With the CV method, the best performance values are obtained with the AdaBoost method. The MSE value is 4.53, the RMSE value is 2.13, the MAE value is 1.55, and the R^2 value is 0.997. When comparing the performance values, the most successful results were obtained with the Random Forest, Neural Network, kNN, Tree, and SVM methods, respectively.

Table 3: Performance of the compared models after the RS 66% hold-out method

Model	MSE	RMSE	MAE	R2
AdaBoost	5.7148984117	2.390585370105	1.7449823529411	0.9964304818858
Neural Network	5.9831307616	2.446043900189	1.8454250369695	0.9962629443090
Random Forest	7.2639750941	2.695176264029	1.9369419953739	0.9954629306050
kNN	11.493498679	3.390206288670	2.4522784313725	0.9928211756752
Tree	39.227793943	6.263209556067	2.9207691176470	0.9754983709301
SVM	40.717286562	6.381009838732	4.5802160669577	0.9745680357779

When calculations are made with the RSHO method, the best performance values are obtained with the AdaBoost method. While the R^2 value is 0.9964, the MSE and RMSE values are 5.71 and 2.39, respectively. The MAE value was determined to be 1.74. It was stated that there were ± 3 minutes between the data produced in the work [8] and the realized data. It was explained that a 53% improvement was achieved in the route arrangement with the ABC algorithm developed in [9]. In [10], the mean error values of the prediction results are 3% for training data and 7% for test data. In [11], an error range of 5-22% was obtained for predictions produced regarding 90% travel time data. In [12], the PRD (percent root mean distortion) value of the prediction values is in the range of 0.1-0.001%. In [13], the R^2 value for the predictions produced was determined to be 0.35. In [14], it was stated that 42.86% success was achieved in the travel time predictions. In [15], the RMSE value varies in the range of 0.01-0.04 in the prediction results produced. In the study indicated with [16], it was explained that there was a 2% error amount related to the values realized with the prediction results. In the study given in [17], the results obtained were produced with an error in the range of 50-300 seconds. In the data produced in [18], the amount of MAE belonging to the travel time was found to be in the range of 120-150 seconds for the travel situation. In the study indicated with [19], the difference between the prediction results and actual values was found to be less than 2 minutes. In the study specified in [20], the MAPE (mean absolute percentage error) value of the prediction results produced is in the range of 14-26%.

4. Conclusion

Travel time significantly affects the efficiency of the railway operation and the performance of the signalization system. The model created for travel time calculations has the feature to be applied for each line regardless of the specific features of lines. In this study, the prediction of the travel time between stations in intracity metro systems was carried out with machine learning methods. With this work, the most important output for reaching the desired destination in the target period, which is the aim of the signalization system, the travel time was predicted with high accuracy by the proposed method. In this study, the machine learning methods compared were applied to the created dataset using the CV and RSHO techniques. In addition to the AdaBoost method recommended in this article and giving the most successful results in the travel time prediction studies found by using experimental data, the NN, RF, kNN, Tree, and SVM machine learning methods were used and compared. Here, besides the R^2 values, the error values obtained (MSE, RMSE, and MAE) were also taken into consideration while performing the performance evaluation. For both CV and RHSO techniques, when the AdaBoost method proposed in this study and the methods discussed are examined in detail, it is seen that R^2 values are very close to each other. However, it has been observed that the proposed method outperforms the closest method by 15,67% in MSE value, 7.51% in RMSE value and 9.68% in MAE value compared to the closest method.

Thanks to the proposed method, travel time can be predicted with high accuracy in accordance with the schedule determining the availability of the system, even under changing conditions. Thus, it can be easily foreseen whether the system can adapt to flexible working conditions easily and quickly. This work is important for the dissemination and promotion of machine learning applications for the vehicle traffic signalization system used/to be used in railway operations.

Acknowledgments

We would like to thank Istanbul Metropolitan Municipality, Department of Rail System for its support during the realization of this study.

This study was funded by the Scientific Research Projects Coordination Unit of Istanbul University - Cerrahpasa. Project numbers: 23444 and 23446

References

[1] Riccardo, R., Massimiliano, G., "An empirical analysis of vehicle time headways on rural two-lane two-way roads", **Procedia - Social and Behavioral Sciences**, 2012, 54: 865 – 874.

[2] Suweda, I., W., "Time Headway Analysis to Determine the Road Capacity", **Jurnal Spektran**, 2016, 4 (2): 71-75.

[3] Nakamura, H., "Analysis of minimum train headway on a moving block system by genetic algorithm", **Transactions on the Built Environment**, 1998, 34: 1014-1022.

[4] Jang, J., Park, C, Kim, B., Choi, N., "Modeling of Time Headway Distribution on Suburban Arterial: Case Study from South Korea", ", **Procedia - Social and Behavioral Sciences**, 2011, 16: 240 – 247.

[5] Maurya, A., K., Das, S., Dey, S., Nama, S., "Study on Speed and Time-headway Distributions on Two-lane Bidirectional Road in Heterogeneous Traffic Condition", **Transportation Research Procedia**, 2016, 17: 428 – 437.

[6] Maurya, A., K., Dey, S., Das, S., "Speed and Time Headway Distribution under Mixed Traffic Condition", **Journal of the Eastern Asia Society for Transportation Studies**, 2015, 11: 1774-1792.

[7] Minh, C., C., Sano, K., Matsumoto, S., "The Speed, Flow and Headway Analyses of Motorcycle Traffic", **Journal of the Eastern Asia Society for Transportation Studies**, 2005, 6: 1496 – 1508.

[8] Domenichini, L., Salerno, G., Fanfani, F., Bacchi, M., Giaccherini, A., Costalli, L., Baroncelli, C., "Travel time in case of accident prediction model", **Procedia - Social and Behavioral Sciences** 53, 2012, 1079 – 1088.

[9] Tatomir, B., Rothkrantz, L., J., M., Suson, A., C., Travel time prediction for dynamic routing using Ant Based Control. In: Proceedings of the 2009 **Winter Simulation Conference** (WSC), 2009, 1069–1078.

[10] Kisgyörgy, L. and L.R. Rilett, Travel time prediction by advanced neural network. **Periodica Polytechnica Ser. Civ. Eng**, 2002, 46(1), 15–32.

[11] Li R, Rose G, Chen H, and Shen J. Effective Long Term Travel Time Prediction with Fuzzy Rules for Tollway, **Neural Computing and Applications**, 2017, 30, 2921–2933.

[12] A. Narayanan, N. Mitrovic, M. T. Asif, J. Dauwels, and P. Jaillet, "Travel time estimation using speed predictions," in Intelligent Transportation Systems (ITSC), 2015 IEEE **18th International Conference on. IEEE**, 2015, 2256–2261.

[13] A. Gal, A. Mandelbaum, F. Schnitzler, and A. Senderovich, "Traveling time prediction in scheduled transportation with journey segments," **Inform. Syst.**, 2017, 64: 266-280.

[14] Li, Y., Gunopulos, D., Lu, C. and Guibas, L. 2017. UrbanTravel Time Prediction using a Small Number of GPS Floating Cars. InSIGSPA-TIAL GIS. ACM, 2017, 3: 1-10.

[15] Bauer, D., & Tulic, M. Travel time predictions: should one modelspeeds or travel times?**European Transport Research Review**, 2018, 10(46): 1-12.

[16] Zychowski, A.; Junosza-Szaniawski, K.; Kosicki, A. Travel Time Prediction for Trams in Warsaw.In Proceedings of the 10th International Conference on Computer Recognition Systems (CORES),Polanica Zdroj, Poland, 22–24 May 2017; Kurzynski, M., Wozniak, M., Burduk, R., Eds.; **Springer InternationalPublishing**: Cham, Germany, 2017, 53–61

[17] J. Rupnik, J. Davies, B. Fortuna, A. Duke, and S. S. Clarke, “Travel timeprediction on highways,” in **International Conference on Computer andInformation Technology**, 2015, 1435–144.

[18] Woodard, D.; Nogin, G.; Koch, P.; Racz, D.; Goldszmidt, M.; Horvitz, E. Predicting travel time reliabilityusing mobile phone GPS data.**Transp. Res. Part** C, 2017,75, 30–44.

[19] S. hak, S. Kim, K. Jang, et al. Real-hime hravel hime Prediction Using Multi-Level k-Nearest Neighbor Algorithm and Data Fusion Method. 2014 **International Conference on Computing in Civil and Building Engineering**, 2014.

[20] Yusuf, A., “Advanced Machine Learning Models for Online Travel-time Prediction on Freeways”, PhD. Thesis, Georgia Institute of Technology.

[21] Faheem, H., Hashim, I., H., “Analysis of Traffic Characteristics at Multi-lane Divided Highways, Case Study from Cairo-Aswan Agriculture Highway”, **International Refereed Journal of Engineering and Science** (IRJES), 2014, 3 (1): 58-65.

[22] Mathew, T., V., Rao, K., V., K.,” Fundamental Parameters of Traffic Flow”, **Introduction to Transportation Engineering**, NPTEL, May 3, 2007.

[23] Geistefeldt, J., “Capacity effects of variable speed limits on German freeways”, Procedia - **Social and Behavioral Sciences**, 2011, 16: 48 – 56.

[24] Freund, Yoav, Robert Schapire, and Naoki Abe. "A short introduction to boosting." **Journal-Japanese Society For Artificial Intelligence**, 1999, 14 (5), 771-780.

[25] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. **Journal of Computer and System Sciences**, 1997, 55(1): 119-139.