SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
VOL. 5, NO. 1, APRIL 2022
DOI: 10.35377/saucis.05.01.932969

Research Article

Calculation of Driving Parameters for GOA4 Signaling System using Machine Learning Methods

Mehmet Taciddin AKÇAY[1], Abdurrahim AKGUNDOGDU[2]
[1]Corresponding Author; Department of Electrical-Electronics Engineering, Faculty of Engineering, Halic University, Istanbul, Turkey; mehmettaciddinakcay@halic.edu.tr
[2]Department of Electrical-Electronics Engineering, Faculty of Engineering, Istanbul University Cerrahpasa, Istanbul, Turkey; akgundogdu@iuc.edu.tr

Received 5 May 2021; Revised 4 March 2022; Accepted 10 March 2022; Published online 30 April 2022

Abstract

Among the electromechanical components of the rail system, the rail system vehicle is one of the most important units that carrying the passenger load. In terms of the efficiency of the signalization system, it is very critical to create the optimum vehicle driving profile. While many parameters of the vehicle are important during the design of the driving profile, determining the acceleration and braking accelerations directly affects this characteristic. The use of programmable devices and software instead of human factors is becoming more and more widespread day by day with the developed technology in rail transportation systems. Among the software used, artificial intelligence and machine learning applications constitute a large share in the general distribution. Especially, in the driverless (GOA4) signaling systems, these softwares become more important. In this study, the estimation of Vehicle Acceleration and Braking Acceleration with travel time has been done by using Machine Learning Methods. The ideal results obtained were given comparatively and interpreted on the graphics.

Keywords: Acceleration, braking, driverless, machine learning, vehicle.

1. Introduction

Today, the rail system technology is constantly renewed and updated in the best way to respond to the needs of the age at an adequate level. Especially the high development in the field of electrical-electronics is of interest to the systems used in this field, and new systems are replaced by old systems. With the high performance achieved in systems where communication and data systems used in rail system technology in our age are integrated, the systems leave themselves to automatic software and control methods. Especially with the efficient management of data with big data technology, great advantages have been provided to the rail system operation in controlling and controlling the systems. High success is achieved in machine learning applications, artificial intelligence algorithms, and applications with the wide data networks obtained from the systems. These studies also help companies in optimization of operational traffic by increasing their management skills. It is the most important subsystem signaling system that closely concerns electromechanical issues, operating performance, and RAMS competencies in rail systems. While designing the signalization system, compliance with EN 50126, EN 50128, and EN 50129 standards is the most basic principle. The rating of the signaling system according to the automation level (GOA) is specified in the IEC 62290 standard. Among these ratings, GOA 4 has the highest performance in terms of the highest automation level and features it provides. At this level, the vehicle is operated completely without a driver, while the system is managed with a full automation level. It is important to minimize the problems experienced for the signalization system, obtain the targeted RAMS values, and increase the performance. Figure 1 gives the graphic expressing the situations related to the malfunctions that occur**.**

Actions are developed against the problems experienced in the business according to the decrease, the same progress, and increase of the malfunctions. Therefore, possible errors in the signaling system are prevented with the highest automation level following SIL 4 criteria. At the GO4 signaling level, the acceleration and braking (driving characteristic) of the vehicle are automatically activated.

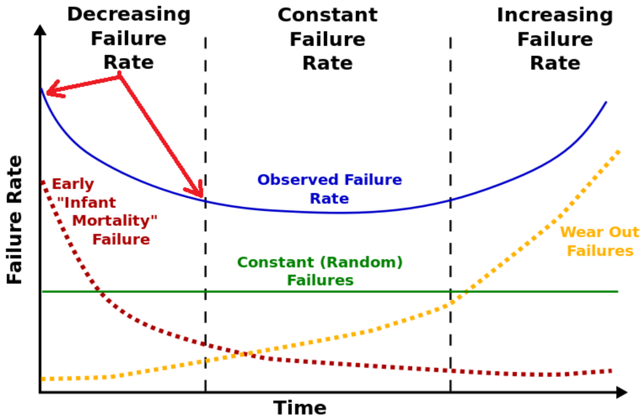

Figure 1 Failure Rate Situation

In this study, the estimation of Vehicle Acceleration and Braking Acceleration has been carried out using Machine Learning Methods. In the second part of this article, the model proposed for the estimation of vehicle acceleration and braking acceleration, and the input-outputs used to create the structure of the model and the proposed method are explained. By presenting the simulation results obtained in the third section, a comparison has been made with other similar methods. In the last part, the evaluation of the article is given in detail. Various studies have been carried out in the literature on this subject. While estimating the arrival times of vehicles with vehicle speed measurement done in the study [1], the method for calculating vehicle motion resistance with the help of a vehicle data collection device is explained with [2]. In [3], the prediction of critical speed for the vehicle on soft ground is studied, and in [4] a low-speed estimation has been made with the time signal-based warping algorithm. In [5], the simulation of the vehicle movement has been done by the element increment method. While traffic speed estimation is made with the deep learning method in [6], a study has been done on speed prediction models with [7] and hybrid systems have been investigated. In [8], the estimation of vehicle motion resistance with the help of a vehicle onboard system was examined. While the estimation of the cruise speed is made using learning methods in (9), in [10] the situation of estimating the vehicle speed according to the energy consumption has been investigated. While estimating the average vehicle speed according to weather conditions and traffic characteristics with [11], in [12] the vehicle travel speed was estimated by using machine learning methods. While real-time arrival estimates were made in light rail systems with [13] and [14], vehicle speed was calculated with ground vibrations in [15]. In this study, the performance outputs used by the signaling system and effectiveness in the vehicle speed profile were estimated by machine learning methods.

2. Material and Method

2.1 Experimental Study and Simulation

In this study, the distance between two stations with the setup set up, the acceleration, braking acceleration, and travel time parameters of the signaling system was used for the structure of the design. Each row from the data sets represents the data sets obtained separately from each other for all data sections. While the system was in operation, records were taken and the characteristics of the speed profiles were reached. Figure 2 shows the experimental setup.

After the driving algorithm is applied, speed position profiles are formed and the data for the specific acceleration curve for the vehicle with the following query code sequence below is taken from the recorded part. Figure 2 shows how the driving profile is created with the characteristic driving parameters of the vehicle. This curve is limited according to the operating speed limits, and starting and braking situations are created according to the conditions determined in the operation with the control blocks.

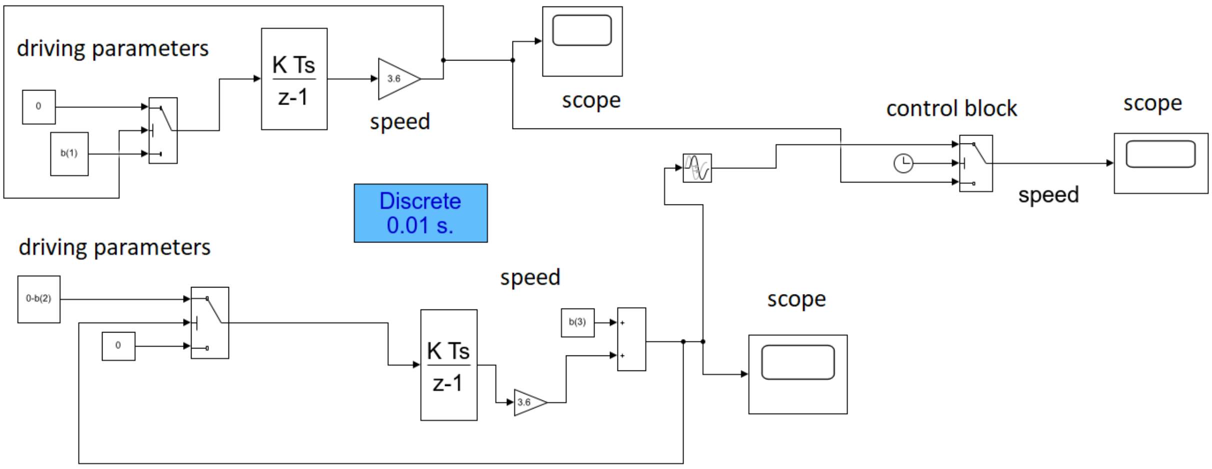

Figure 2 Experimental setup

Code 1 Query

```
for datas=1:1:length(speed.signals.values)
if speed.signals.values(datas)>=max operation speed
acceleration(n)=speed.time(datas);
 n=n+1;
end
if   speed.signals.values(datas)<0
total(m)=zaman.time(datas);
m=m+1;
end
end
```

2.2 Data Set

Within the scope of the study, 300 data arrays were used, while the distance between stations and maximum operating speed were used as input parameters, vehicle acceleration, deceleration, and travel speed were used as output parameters. The distribution of some of the data is given in Figure 3.

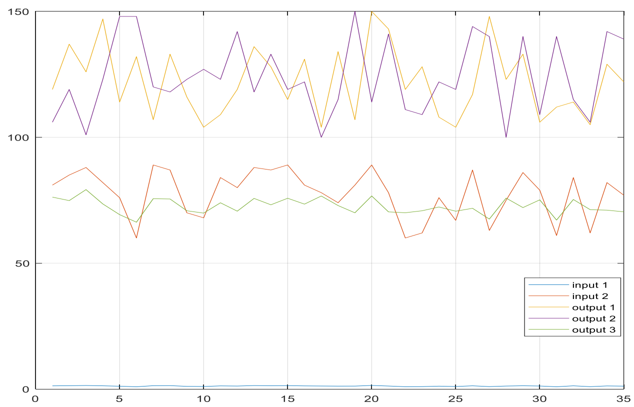

Figure 3 Graph of a section of data

The number of data was chosen according to the success of the methods, and the data realized under operating conditions were preferred. Values ranging from 60 to 90 km/h were used for the maximum operating speed. As the output, the acceleration values of the vehicle in the starting and braking states were used, and this data varies between 1 and 1.5 m/s2. The distance between the stations was taken between 1 and 2 km since this design was made for the metro stations.

2.3 Machine Learning Methods

Machine learning is the computer modeling of systems that make predictions by making inferences from operations on data using mathematics and statistics. The model is created with the current data set and the algorithm used. Machine learning is used to get the maximum performance from models. In this study, Linear Regression, Random Forest, Multi-Layer Perceptron, and K-Neighbors Regressor models, which are among the supervised learning methods, were applied to the data.

2.3.1 Linear Regression

Linear regression is a method used to model the relationship between one or more independent variables and another dependent variable. The purpose of linear regression; is to find the values of β_i using given x and y. Once the β_i values are found, y's of unknown value can be predicted from the x values [16,17]. Linear regression can be formulated as follows:

$$y_i=\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+...+\beta_p x_{ip}+\epsilon i \tag{1}$$

In this equation;

yi: dependent variable
xi: explanatory variables
β_0: y-intercept (constant term)
βi: is the regression coefficient
β_p: slope coefficients for each explanatory variable
ϵ = the model’s error term.

2.3.2 Random Forest

Random Forest is a supervised learning algorithm. The Random Forest creates multiple decision trees and combines them with the bagging method to obtain a more accurate and stable forecast. Figure 4 shows a simple Random Forest model.

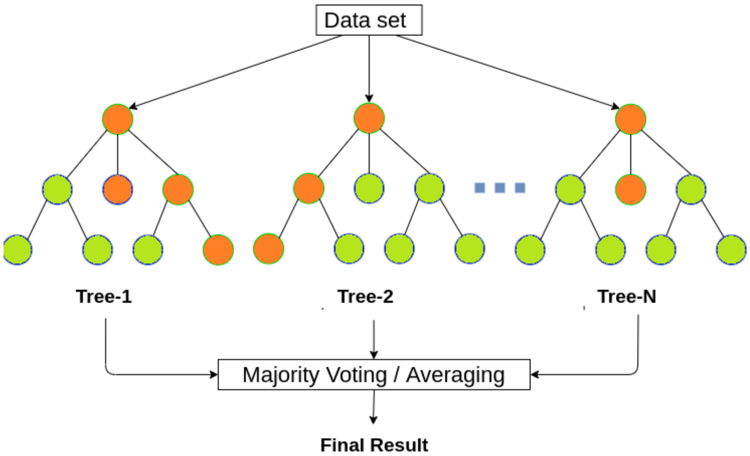

Figure 4 An example of the random forest with N decision trees [18]

The main advantage of the random forest is that it can be used for both classification and regression problems that make up most of the current machine learning systems [19, 20]

2.3.3 Multi-Layer Perceptron (MLP)

MLP is the simplest in artificial neural networks. It consists of three layers: input, hidden, and output. The input layer is entered into the system and is the layer where it is processed towards the hidden layer. In these networks, error learning is performed by distributing backward in each transaction cycle [21]. A nonlinear activation function transmits the sum of the weighted input signals. The actual observation results are compared with the results of the network and the error of the network is calculated. Then the calculated network error is propagated back by the system and the weights of the coefficients are updated [22].

MLP generally contains one or more hidden neuronal layers. After these neurons are the output layer of the neurons. The network learns the linear and nonlinear relationships between input and output vectors through transfer functional neuron layers. An example MLP model is shown in Figure 5.

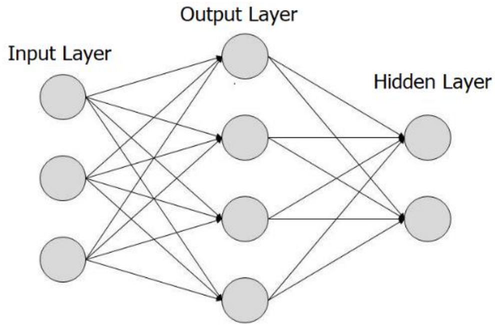

Figure 5 Example of an MLP model with three inputs and two outputs

2.3.4 K Neighbors Regressor

K-NN, which is sensitive to the distance function, is a non-parametric learning algorithm. The main idea of the algorithm is that if the most similar training data in the feature space belongs to a cluster, it includes the training data in this cluster. To determine to which cluster the training data in the feature, space belongs to, the distance between training data is determined by distance functions such as Euclidean, Manhattan, Minkowski, and Kullback-Leibler [23, 24].

2.4 Performance Calculations

2.4.1 Statistical Performance Validation

In this study, Root Mean Square Error (RMSE) and coefficient of determination (R^2) values were calculated. These performance values can be formulated as follows.

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(T_i - Y_i)^2}{N}} \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(T_i - Y_i)^2}{\sum_{i=1}^{N}(T_i - \bar{T})^2} \tag{3}$$

where T_i is the targets, $\bar{T}$ is the mean of all target values, Y_i is the neural network outputs and N is the number of samples. These performance results were used as a reference for the comparisons about the all methods [25].

3. Results and Discussion

In this study, acceleration braking acceleration and travel times are estimated simultaneously with three different machine learning methods. While performing regression, the data were subjected to training and testing processes in 10 groups with the cross-validation method. The regression curves obtained after each method are shown in Figure 6, Figure 7, Figure 8, and Figure 9 respectively. RMSE values and R^2 values obtained after the estimates are shown in Table I. When Table I is examined, it is seen that the most successful method is the MLP.

Table 1 Performance evaluations of different models

Model	**RMSE**	**R^2**
Random Forest	9.79	0.30
kNN	8.88	0.42
Linear Regression	8.04	0.52
MLP	7.69	0.56

Figure 6 shows the regression distribution related to the estimation results of the kNN method, and a partial success has been achieved in this case.

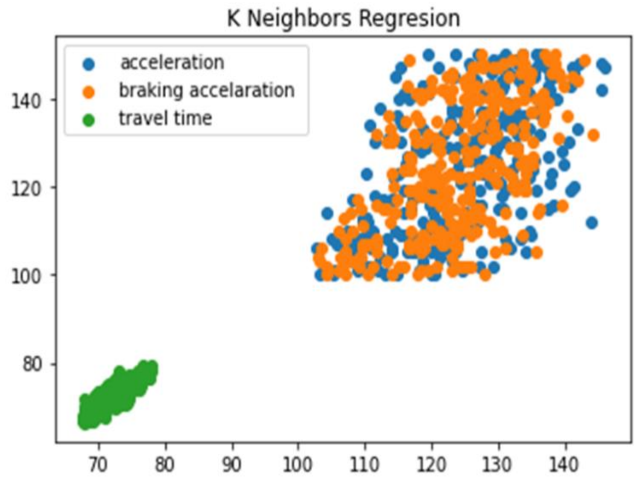

Figure 6 Regression curves of models (kNN)

The regression distribution of the Random Forest method is given in Figure 7, and in this case, it can be understood from the rate of deviation at the 45 degree slope line as it can be seen that less successful results are obtained compared to the kNN method.

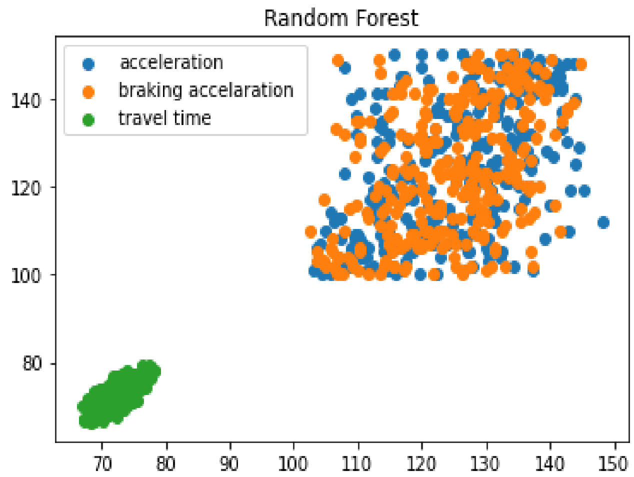

Figure 7 Regression curves of models (Random Forest)

Figure 8 shows the regression distribution related to the estimation results of the Linear Regression method, and in this case, the most successful results were obtained compared to the previous two methods.

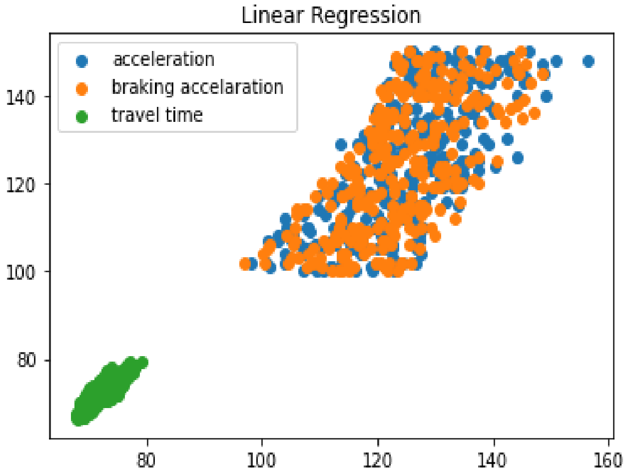

Figure 8 Regression curves of models (Linear Regression)

While Figure 9 shows the regression distribution related to the estimation results of the MLP method, in this case, the most successful estimation results compared to the previous methods were obtained. In this case, while the R2 value is obtained as 0.56, it is understood that the results achieved under these conditions are close to ideal and success, and at the same time, the results seem to be at a usable level.

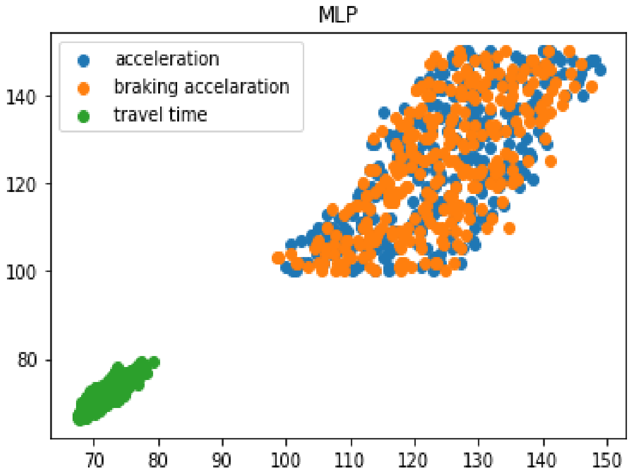

Figure 9 Regression curves of models (MLP)

The results obtained for different methods are given in Table 2. The SGD method proposed in this study and the most successful results are obtained, and the performance values of AdaBoost, Random Forest, Neural Network, kNN, Decision Trees and SVM methods, which are frequently used in the literature, can be seen in these tables, respectively.

4. Conclusion

In this study, the driving parameters of the driverless signaling system with GOA4 technology used in subway lines were estimated using machine learning methods. Speed profiles for the vehicle were supported by calculating vehicle acceleration, braking acceleration, and travel times. The performance values of the results obtained with the study using Random Forest, kNN, Linear Regression, and MLP methods are given on the table and interpreted. RMSE and R^2 values were obtained in calculations as performance criteria. The study aims to integrate new technologies used in the design of automatic driverless systems with machine learning and artificial intelligence applications and to increase operational performance. Analyzes can be diversified with the help of new machine learning methods, artificial intelligence applications, and algorithms by using different data series and types for the estimation of signaling system parameters with future studies.

Acknowledgments

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

References

[1] Y. Chen and L. R. Rilett, “A train speed measurement and arrival time prediction system for highway-rail grade crossings”, *Transportation Research Record Journal of the Transportation Research Board*, vol. 2608, no.1., pp. 96-104, 2017.

[2] T. Ogawa, S. Manabe, G. Yoshikawa, Y. Imamura, and M. Kageyama, “Method of Calculating Running Resistance by the Use of the Train Data Collection Device”, *Quarterly Report of RTRI*, no. 58, pp. 21-27, 2017.

[3] K. N. Cosgriff, E. G. Berggren, A. M. Kaynia, N. N. Dam, and N. A. Mortensen, “new method for estimation of critical speed for railway tracks on soft ground”, *International Journal of Rail Transportation*, vol. 6, no. 4, pp. 203-217, 2018.

[4] S. Hensel, M. Marinov, “Time Signal Based Warping Algorithms for Low Speed Velocity Estimation of Rail Vehicles”, *Annual Journal of Electronics*, no. 8., pp. 177-180, 2014.

[5] G. Xu, F. Li, J. Long, and D. Han, “Train movement simulation by element increment method”, *Journal of advanced transportation*, no. 50, pp. 2060–2076, 2017.

[6] Z. Lv, J. Xu, K. Zheng, H. Yin, P. Zhao, and X. Zhou, “LC-RNN: A Deep Learning Model for Traffic Speed Prediction”, in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, 2018, pp. 3470-3476.

[7] A. Dhamaniya and S. Chandra, “Speed Prediction Models for Urban Arterials under Mixed Traffic Conditions”, *Procedia-Social and Behavioral Sciences*, no. 104, pp. 342 – 351, 2013.

[8] S. Aradi, T. Becsi, and P. Gaspar, “Estimation of running resistance of electric trains based on on-board telematics system”, *International Journal of Heavy Vehicle Systems*, no. 22, pp. 277-291, 2015.

[9] M. Gmira, M. Gendreau, A. Lodi, and J. Potvin, “Travel Speed Prediction Based on Learning Methods For Home Delivery”, *Interuniversity Research Center On Business Networks, logistics and transport*, pp. 1-34, 2018.

[10] M. Bysveen, *Vehicle speed prediction models for consideration of energy demand within road design*, Norwegian University of Science and Technology, Civil and Environmental Engineering, Master’s Thesis, 2017.

[11] B. Mirbaha, M. Saffarzadeh, S. A. Beheshty, M. Aniran, M. Yazdani, and B. Shirini, “Predicting Average Vehicle Speed in Two Lane Highways Considering Weather Condition and Traffic Characteristics”, *IOP Conference Series: Materials Science and Engineering*, pp. 1-7, 2017.

[12] M. Gmira, M. Gendreau, A. Lodi, and J. Potvin, “Travel speed prediction using machine learning techniques”, *ITS World Congress*, pp. 1-10, 2017.

[13] E. Naye, *Real-time arrival prediction models for light rail train systems*, Royal Insistute of Technology, Department of Engineering, Master’s Thesis, 2014.

[14] O. Cats, “Real-Time Predictions for Light Rail Train Systems”, *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1-10, 2014.

[15] G. Kouroussis, D. P. Connolly, M. Forde, and O. Verlinden, “Train speed calculation using ground vibrations”, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 229, no. 5, pp. 466-483, 2015.

[16] D. Freedman, *Statistical Models: Theory and Practice*, Cambridge: Cambridge University Press, pp.25, 2005.

[17] A. C. Rencher, and W. F. Christensen, “Chapter 10, Multivariate regression – Section 10.1, Introduction”, Methods of Multivariate Analysis, *Wiley Series in Probability and Statistics, 709 (3rd ed.), John Wiley & Sons*, pp. 19, 2012.

[18] Available: https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/

[19] T. K. Ho, “Random Decision Forest”, *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, 14-16 August, pp. 278-282, 1995.
[20] L. Breiman, “Random Forests”, *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[21] J. A. Freeman and D. M. Skapura, “Neural Networks Algorithms”, *Applications and Programming Techniques*, New York, USA: Addison-Wesley Publishing Company, 1991.
[22] G. J. Garret and C. Wu, “Knowledge-based modeling of material behavior with neural networks”, *Journal of Engineering Mechanics*, vol. 117, no.1, pp. 132-153, 1991.
[23] T. M. Cover, and P. E. Hart, “Nearest neighbor pattern classification”, *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.
[24] J. Walters-Williams, Y. Li, “Comparative study of distance functions for nearest neighbors”, *In Advanced Techniques in Computing Sciences and Software Engineering, Springer, Dordrecht*, pp. 79-84, 2010.
[25] M. Akçay, “Estimation of Constant Speed Time for Railway Vehicles by Stochastic Gradient Descent Algorithm”, *Sakarya University Journal of Computer and Information Sciences*, vol. 3, no. 3, pp. 355-365, 2020.