e-ISSN: 2147-8228

INTERNATIONAL JOURNAL OF APPLIED MATHEMATICS ELECTRONICS AND COMPUTERS

www.dergipark.org.tr/ijamec

International Open Access

Volume 10
Issue 03

September, 2022

Research Article

Detection of accident situation by machine learning methods using traffic announcements: the case of metropol Istanbul

Eren Dağlı [a], ***Mustafa Büber*** [a,*], ***Yavuz Selim Taşpınar*** [a]

[a]*Doganhisar Vocational School, Selcuk University, Konya, TÜRKİYE*

ARTICLE INFO

Article history:
Received 19 July 2022
Accepted 3 August 2022

Keywords:
Accident Status
Classification
Detection
Machine Learning
Traffic Announcement

ABSTRACT

Information about the reality of the traffic accident, the clearness of the roads and the status of the accident can be obtained from the traffic accident announcements. By using the words in the radio or telephone announcements, you can be informed about the status of the accident. Inferences can be made with machine learning methods using a large number of data. In this study, the accident situation was classified using three different machine learning methods using radio and telephone announcements in Istanbul in Türkiye. The dataset contains 156.856 announcement data. Classifications were performed using Artificial Neural Network (ANN), k-Nearest Neighbor (kNN) and Decision Tree (DT) machine learning methods. Classification accuracy was 92.1% in the classification made with the ANN model, 91% in the classification made with the kNN model, and 89.8% in the classification made with the DT model. Classification performances of the models were also analyzed with precision, recall, F-1 Score and specificity metrics. In addition, the estimation abilities of the models with ROC curves and AUC values were analyzed. In addition, the training and testing times of the models were also analyzed. It will be possible to use the suggested models to automatically detect the accident situation from the announcements. In this way, it is thought that the most accurate direction can be made by obtaining information about crew orientation, traffic jams and the size of the accident.

1. Introduction

An accident is an event that occurs at any time owing to human omission and wrongful behavior, resulting in negative effects such as loss of life and property [1]. On the other hand traffic accidents are defined as the unforeseen events that occur on the road, in which the vehicle, driver and pedestrian are mixed separately or together and cause various losses as a result [2]. In addition to deaths and injuries, traffic accidents cause economic losses and psychological problems for both society and individuals. In order to eliminate these negative effects, it is important to determine the accuracy of emergency interventions and create an effective traffic culture and eliminate the effects that cause traffic accidents.

The timely and correct response to victims is a determining factor in survival rates after an accident [3]. By eliminating the time elapsed between the moment the accident occurs and the intervention of first aid teams, mortality rates can be reduced by 6% [4]. In their study, [5] emphasized the importance of in-vehicle systems that detect the occurrence of the accident and notify the emergency personnel, noting that rapid emergency response increases the survival rate of the victims. They suggested that this process should be carried out with smartphones without additional cost, especially in old type vehicles in which there are no in-vehicle accident detection and emergency call systems.

The most important stage to eliminate traffic accidents and the negative effects of these accidents is to determine the causes of traffic accidents correctly and eliminate them. In this direction, many pieces of researches have been carried out in order to determine and eliminate the factors that cause traffic accidents from past to present and to predict traffic accidents. [6] designed a three-dimensional model-based tracking system and followed the vehicles in their study.

* Corresponding author. E-mail address: *mbuber@selcuk.edu.tr*
DOI: 10.18100/ijamec.1145293

They analyzed the trajectory and maneuvering movements of the vehicles by modeling and determined the probability of an accident with algorithms that predict vehicle efficiency and statistical methods. [7] mentioned that they work with limited data in the studies in the literature in traffic accidents predictions. Based on this point, in their study, they made traffic accident predictions using deep learning methods such as traffic volume, road condition, weather and environmental characteristics emphasized that the results improved the prediction accuracy compared to previous methods [8]. drew attention to the heterogeneity in the estimation analysis made on the available data. In order to reach accurate predictions, they processed large data efficiently and performed clustering and classification analyzes. They aimed to increase the accuracy of prediction by supporting location and road information and traffic data. [9] examined whether there is a relationship between population, number of vehicles and death, with data collected from different countries. [10] stated that the estimation models cannot be generalized and cannot be applied to all countries due to regional traffic patterns.

According to the [11] report, 93.6% of passenger transport and 88.3% of freight transport in Türkiye are done by road. Considering these rates, it is crucial for Türkiye that road transportation continues without interruption. In addition, population growth and the increasing the number of motor vehicles increase traffic accidents [12]. The increase in population and socio-economic level together with the possible increase in traffic accidents pose a risk. Researchers, local administrations and relevant public institutions carry out various studies provide a transportation infrastructure that meets the needs at an optimum level. Thanks to the concept of the smart city which aims to use information and communication technologies in an integrated manner with the city, it is aimed by all stakeholders to establish the systematic functioning infrastructure of the city. Intelligent transportation systems, one of the smart city applications, appear as innovative systems that aim at the safety and efficiency of traffic elements in the field of transportation of information communication technology opportunities [13]. With intelligent transportation systems, it is aimed to provide a safer, more environmentally friendly, more efficient road system.

As of March 2022 in Türkiye, 4,714,138 (approximately 18.5%) of 25,478,989 motor vehicles registered in traffic are registered in the province of Istanbul, which has been chosen as the pilot city [14]. Istanbul Metropolitan Municipality Transportation Management Center operates in order to manage the traffic of the city in line with the visual and numerical data provided by various traffic observation and measurement systems. In addition to data provider systems, the center activates emergency response scenarios in line with instantaneous user notifications. In the meantime, it directs the drivers with various warnings to prevent the density, queue formation, increase in delay times and confusion that may occur in the region. With the variable message system and variable traffic signs, drivers are informed about closed lanes and areas where traffic density is observed. In this way, it is tried to prevent possible confusion and accidents that may occur because of confusion. However, there may be errors in momentary user information entries, and the current situation cannot be evaluated objectively when the same case is reported by more than one user, as well as exaggerated and malicious notifications. It is aimed to prevent erroneous redirects by evaluating the accuracy of the notifications coming to the system with machine learning methods. Thus, unnecessary emergency response operations will be eliminated and faster interventions will be possible when necessary. In addition, directing the drivers with more notifications will prevent confusion.

When the literature is examined, there are few studies on accident detection from traffic announcements. The contributions of this study can be listed as follows;

- 136,964 traffic announcements in Istanbul city of Türkiye were used. This number is quite high and is a sufficient number of training data for the models to be created.
- The machine learning models used in the study have not been used for accident detection from traffic announcements before.
- With the proposed methods, accident detection can be done automatically from the announcements.
- The classification accuracy of the proposed models is sufficient for the models to be used in different applications.

The rest of the study is organized as follows: In the second section, the dataset, methods and performance metrics used in the study are explained. In third section, the results obtained from the experiments carried out in the study are given. In fourth section, the results obtained from the study are explained.

2. Material and Methods

In this section, the dataset used in the study, methods and methods used in performance evaluation are explained.

2.1. Dataset

The dataset used in the study was created by the Istanbul Metropolitan Municipality Data Center by using the traffic announcements records in Istanbul, Türkiye. The dataset has been published as open access on the website of this institution [15]. The dataset contains a total of 136,964 data. There are 12 features in the dataset. In the study, the Accident Status feature was estimated using 11 features. The features in the dataset and the features of these features are shown in Table 1.

Table 1. Features in the traffic announcements dataset

FEATURES	DATA TYPE	LABEL	DESCRIPTION
ANNOUNCEMENT_STARTING_DATETIME	timestamp	Announcement Start Date	It is the field that contains date, hour, minute and second information.
ANNOUNCEMENT_ENDING_DATETIME	timestamp	Announcement End Date	It is the field that contains date, hour, minute and second information.
ANNOUNCEMENT_TITLE	text	Announcement Title	It is the field where the title information of the announcement is located.
ANNOUNCEMENT_TEXT	text	Announcement Text	It is the field that contains the text information of the announcement.
ANNOUNCEMENT_TYPE	numeric	Announcement Type	It is the field that contains the type information of the announcement.
ANNOUNCEMENT_TYPE_DESC	text	Announcement Description	It is the field that contains the explanation information of the announcement type.
INTERVENTION_DATETIME	timestamp	Intervention Date	It is the field that contains date, hour, minute and second information.
ACCIDENT_STATUS	numeric	Accident Status	If the announcement is made for a traffic accident, it is the field containing the accident situation.
ACCIDENT_DESCRIPTION	text	Accident Description	If the announcement is made for a traffic accident, it is the field that contains an explanation for the accident situation.
CLOSED_LANE	numeric	Indoor Strip	It is the numerical field that shows how many lanes are out of use if the lanes are not used as a result of an accident.
LONGITUDE	numeric	Longitude	It is the field that contains longitude information.
LATITUDE	numeric	Latitude	It is the field that contains latitude information.

2.2. Artificial Neural Network (ANN)

Artificial neural networks (ANN) are a method that mimics the human brain and tries to recognize the main relationships in the data set used by pretending to be it [16]. In general, the structure of an ANN consists of an input layer, several hidden layers, and an output layer. Figure 1 shows the structure of a simple network of neurons. While the neurons of each layer are connected to each other, the neurons in the layers are not connected to each other. [17, 18]. The data received from the input layer passes through a transfer or activation function and is transmitted to the next neuron. The weight parameters between each neuron are correlated between neurons and have a constant value. Even if there is no non-linear relationship between input and output variables, ANN calculates by considering all possibilities with its self-learning ability [19]. The optimal number of neurons located in the hidden layer is found by trial and error. In this way, it optimizes parameters such as weights and bias values in the training process [20]. Accidents have direct and indirect effects on society and human life. It is accepted that ANNs are a powerful and effective approach to dealing with the accident dilemma [21].

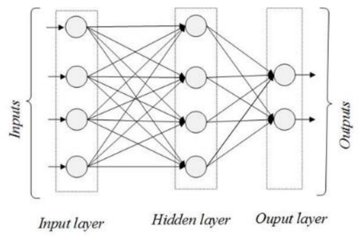


Figure 1. Artificial Neural Network (ANN) Layers

2.3. k-Nearest Neighbor (kNN)

kNN is one of the widely used classification and estimation methods. The conclusion process takes place as follows;

- Calculation of the distance between the target and other known values
- Detection of the k nearest neighbors of the calculated distance to the target
- Obtaining the value of k neighbors
- Estimated value of target

gives successful results with processes [22, 23]. The result of the target is the mean of its K neighbors. This method is widely used to predict possible classes of data in a given dataset with a simple classifier technique [24]. The kNN method is better understood by examining Figure 2 visually. In the estimation of the target, a small number of neighbors gives the estimation result low, while too many neighbors can give very satisfactory results. For this, it is necessary to determine the appropriate K value [25, 26].

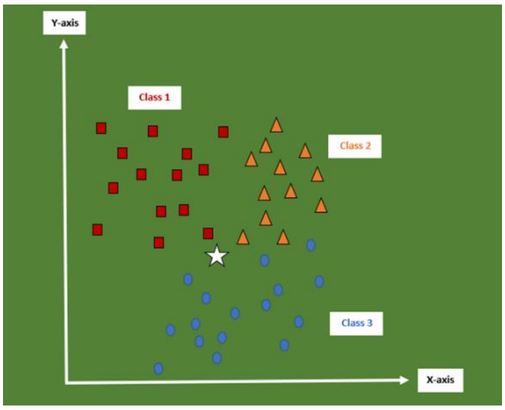


Figure 2. k-Nearest Neighbor (kNN)

2.4. Decision Tree (DT)

The main purpose of the decision tree method is based on the idea of separating the input data into smaller groups using a clustering algorithm. When Figure 3 is examined, the clustering process continues until all elements in the input data have class labels. There are three steps to perform this process in general. These;

- Determination of target and attribute variables
- The dataset is divided into sub-nodes using various partitioning algorithms (such as ID3, C4.5) according to the selected attribute variables.
- Each child node is considered a parent node if it is further divided according to the desired attribute [27, 28].

The algorithm used in the division of the dataset is to ensure that each sub-node is formed homogeneously. In this study, the binary tree algorithm was chosen for the division of the root. This chosen algorithm decides by calculating entropy and information gain values to see if further division of the node is required. Entropy represents uncertainty, and if the entropy value is high, the uncertainty is high. The fact that the entropy is close to zero indicates that the data set is completely homogeneously distributed [29, 30].

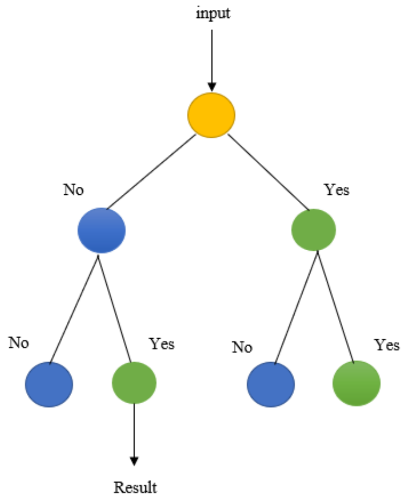


Figure 3. Decision Tree (DT)

2.5. k-fold Cross validation

How much of the dataset will be used as the training set and how much as the test set varies according to the nature of the studies [31]. In this study, k-fold cross validation method was used to separate the dataset into training and test sets. The biggest advantage of the method is that it allows all observations to be used at least once for both training and testing. For this reason, all the data in the dataset are subject to both training and testing, thus positively affecting the classification performance. In the k-fold expression, the letter k represents a numerical value and indicates how many parts the data set will be divided into. The number k is usually chosen with 5 or 10. The process is based on dividing this k number of datasets and using k-1 training data as test data if it is 1 piece. Then, each part divided in turn is used as test data. As a result, the performance of the method used is expressed as the arithmetic average of the sum of the performances obtained by using each part as test data. In this study, the number k was chosen as 10 [29, 30].

2.6. Confusion matrix

One of the techniques used to evaluate the prediction performance of training and test data is the complexity matrix [32]. The values in the matrix are used to evaluate the results of classification problems and to compare their performance [33, 34]. When a two-class confusion matrix given in Figure 4 is examined, the values of TP, TN, FP and FN given in the matrix mean the following;

		True Class	
		Positive (P)	Negative (N)
Predicted	True (T)	TP	TN
	False (F)	FP	FN

Figure 4. A two-class confusion matrix

- TP: True Positive. The real class of the data and the predicted class are the same
- TN: True Negative. Correct guessing that the data belongs to another class
- FP: False Positive. Estimating the data in the real class even though it belongs to another class
- FN: False Negative. Predicting data in another class when it should belong to the real class.

2.7. Performance Metrics

Different performance metrics are used to evaluate machine learning methods. In classification problems, various metrics are calculated from the values obtained from the confusion matrix [35]. Accuracy, precision, recall, F-1 score, specificity, AUC (Area Under Curve) and ROC (Receiver Operating Characteristic) performance metrics were used in this study.

Accuracy: It is found by dividing the number of correct guesses by the number of all guesses.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} x100$$

Precision: The positively predicted (TP + FP) is the ratio of how many of the samples were predicted correctly.

$$Precision = \frac{TP}{TP + FP} x100$$

Recall: It is the ratio of how many of the samples that need to be predicted positively (TP + FN) are predicted correctly.

$$Precision = \frac{TP}{TP + FN} x100$$

F1 score: Combines the precision and recall values to a single number.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} x100$$

The ROC curve provides a visualization of the accuracy of the methods used and a visual comparison of the differences between the classification models. In the coordinate system from which the ROC curve will be obtained, it shows the true positive value (sensitivity) of the dataset on the Y axis and the false positive value (1-specificity) on the X axis. The ROC curve is obtained by combining the points corresponding to true positive and false positive at each cut point. The area under the ROC curve (AUC) is a measure of how well a parameter can distinguish between two groups [30, 31].

3. EXPERIMENTAL RESULSTS

A dataset containing announcements about traffic accidents in the city of Istanbul, Türkiye, was used. The dataset contains 136,964 data. Using these data, classification processes were carried out with three different machine learning methods. The block diagram of the study is shown in Figure 5.

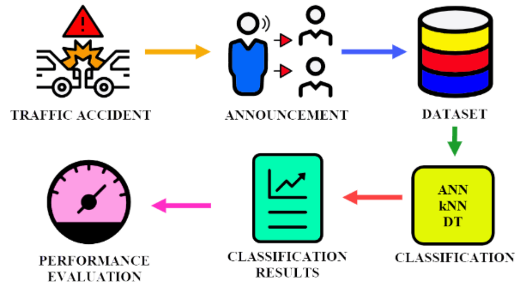


Figure 5. Block diagram of traffic accident status classification

The computer features on which the experiments were performed are shown in Table 2, and the training parameters of each classification method are shown in Table 3.

Table 2. Computer features used in experiments

Processor	Intel i5 10200H @2.4Ghz
RAM	24 GB
Graphical Card	Nvidia GeForce 1650Ti 4GB
Operating System	Windows 10

Table 3. Training parameters of machine learning methods

ANN	kNN	DT
Hidden layer neurons: 100 Activation: ReLu Solver: Adam Iterations: 100	Number of neighbors: 5 Metric: Euclidean Weight: Uniform	Number of instances in leaves: 2 Split subsets: 5 Maximal tree depth: 100

The cross-validation method was used in the training and testing stages of machine learning models. In the cross-validation method, the k value was determined as 10. As a result of the classifications, a confusion matrix was obtained for each machine learning model. The confusion matrix of the ANN model is shown in Figure 6, the confusion matrix of the kNN model is shown in Figure 7 and the confusion matrix of the DT model is shown in Figure 8.

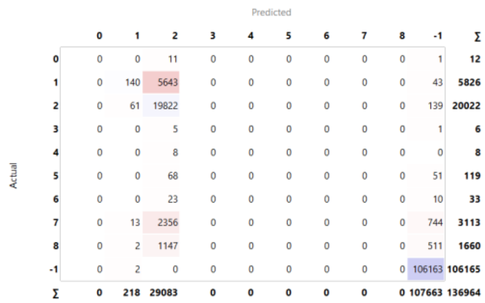
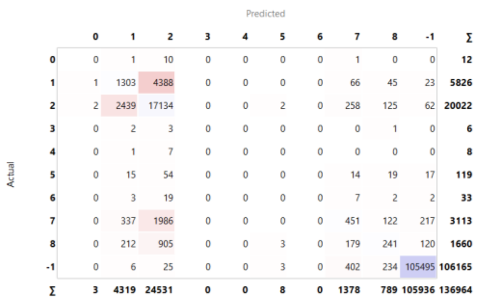
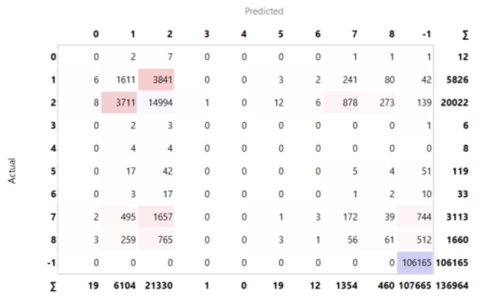


Figure 6. Confusion matrix of ANN model



Actual \ Predicted	0	1	2	3	4	5	6	7	8	-1	Σ
0	0	1	10	0	0	0	0	1	0	0	12
1	1	1303	4388	0	0	0	0	66	45	23	5826
2	2	2439	17134	0	0	2	0	258	125	62	20022
3	0	2	3	0	0	0	0	0	1	0	6
4	0	1	7	0	0	0	0	0	0	0	8
5	0	15	54	0	0	0	0	14	19	17	119
6	0	3	19	0	0	0	0	7	2	2	33
7	0	337	1986	0	0	0	0	451	122	217	3113
8	0	212	905	0	0	3	0	179	241	120	1660
-1	0	6	25	0	0	3	0	402	234	105495	106165
Σ	3	4319	24531	0	0	8	0	1378	789	105936	136964

Figure 7. Confusion matrix of kNN model



Actual \ Predicted	0	1	2	3	4	5	6	7	8	-1	Σ
0	0	2	7	0	0	0	0	1	1	1	12
1	6	1611	3841	0	0	3	2	241	80	42	5826
2	8	3711	14994	1	0	12	6	878	273	139	20022
3	0	2	3	0	0	0	0	0	0	1	6
4	0	4	4	0	0	0	0	0	0	0	8
5	0	17	42	0	0	0	0	5	4	51	119
6	0	3	17	0	0	0	0	1	2	10	33
7	2	495	1657	0	0	1	3	172	39	744	3113
8	3	259	765	0	0	3	1	56	61	512	1660
-1	0	0	0	0	0	0	0	0	0	106165	106165
Σ	19	6104	21330	1	0	19	12	1354	460	107665	136964

Figure 8. Confusion matrix of DT model

When the confusion matrices of the machine learning models in Figure 6, Figure 7 and Figure 8 are examined, it is seen that the data belonging to the same classes are classified correctly and incorrectly in close numbers. Due to the small number of data in some classes, the training of the models was insufficient for the classes with a small number of data. For this reason, classification difficulties were encountered during the testing phase. If the data numbers in the classes are close, the training of the models can be carried out exactly. However, it can be understood from the data on the confusion matrices that the data in the most important classes are correctly classified. The performance metrics of the models were calculated using the data on the confusion matrices. Obtained performance metrics are shown in Table 4.

Table 4. Performance metrics of all machine learning models (%)

	Accuracy	Precision	Recall	F-1 Score	Specificity	AUC
ANN	92.1	89.1	92.1	89	95.1	1.035
kNN	91	89.8	91	90.1	97.8	1.167
DT	89.8	88.3	89.8	89	95.3	1.267

According to Table 4, it is seen that the model with the highest classification accuracy is the ANN model. The highest precision value belongs to the kNN model. The highest recall value belongs to the ANN model, as in the classification accuracy. F-1 Score and specificity values were obtained from the highest kNN model. The highest AUC value was obtained from the DT model. Another parameter that shows the training and test performances of the models is the ROC curves. ROC curves for all models are shown in Figure 9.

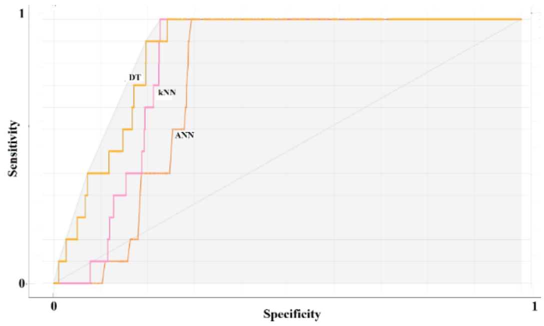


Figure 9. ROC curves of all models

When Figure 9 is examined, it can be said that the DT model is the best learning model. The ANN model is seen as the worst learning model. However, the opposite situation was observed in classification accuracy. The highest classification accuracy belongs to the ANN model. The lowest classification accuracy belongs to the DT model. Classification times of models are also an important factor for their usability in applications. Training and testing times for all models are shown in Table 5.

Table 5. Training and testing time of all models (sec)

	Training Time	Testing Time
ANN	216.240	0.747
kNN	13.142	13.073
DT	80.471	0.027

When Table 5 is examined, it is seen that the highest education period belongs to the ANN model. The lowest training time belongs to the kNN model. The highest test time belongs to the kNN model and the lowest test time belongs to the DT model.

4. Conclusions

Detection of traffic accidents through announcements is an extremely important issue in terms of time and cost. Based on this problem, accident detection models from traffic announcements were proposed by using 136,964 traffic announcements made in Istanbul. Three different machine learning methods were used to classify the data. ANN, kNN and DT methods, which are frequently used in the literature, were used to classify the data. As a result of the classifications, the highest classification accuracy was obtained from the ANN model, 92.1%. The lowest classification accuracy was obtained from the DT model, 89.8%. When the training and testing times of the models are compared, the model that completes the training in the least time is the kNN model with 13.142 seconds. The least test time belongs to the DT model with 0.027 seconds.

According to the results obtained from the classification models, it is thought that the proposed models can be used in the prediction of accidents from traffic announcements. In this way, positive benefits will be provided in terms of time and cost. By creating a decision support system with the models obtained, it is possible to automatically detect the accident and direct the teams.

References

[1] Kıran, S., S. Şemin, and A. Ergör, *Kazalar ve toplum sağlığı yönünden önemi.* Sürekli Tıp Eğitimi Dergisi, 2001. **10**(2): p. 50-1.

[2] Uyurca, Ö. and İ. Atılgan, *Ankara ilinde meydana gelen trafik kazalarının incelenmesi.* Kent Akademisi, 2018. **11**(4): p. 618-626.

[3] Champion, H.R., J. Augenstein, A.J. Blatt, B. Cushing, K. Digges, J.H. Siegel, and M.C. Flanigan, *Automatic crash notification and the Urgency algorithm: Its history, value, and use.* Advanced Emergency Nursing Journal, 2004. **26**(2): p. 143-156.

[4] Rauscher, S., G. Messner, P. Baur, J. Augenstein, K. Digges, E. Perdeck, G. Bahouth, and O. Pieske. *Enhanced automatic collision notification system-improved rescue care due to injury prediction-first field experience*. in *The 21st International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV)-International Congress Center Stuttgart, Germany*. 2009.

[5] White, J., C. Thompson, H. Turner, B. Dougherty, and D.C. Schmidt, *Wreckwatch: Automatic traffic accident detection and notification with smartphones.* Mobile Networks and Applications, 2011. **16**(3): p. 285-303.

[6] Weiming, H., X. Xuejuan, D. Xie, T. Tieniu, and S. Maybank, *Traffic accident prediction using 3-D model-based vehicle tracking.* IEEE Transactions on Vehicular Technology, Vehicular Technology, IEEE Transactions on, IEEE Trans. Veh. Technol., 2004. **53**(3): p. 677-694.

[7] Yuan, Z., X. Zhou, and T. Yang. *Hetero-convlstm: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data*. in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018.

[8] Park, S.-h., S.-m. Kim, and Y.-g. Ha, *Highway traffic accident prediction using VDS big data analysis.* The Journal of Supercomputing: An International Journal of High-Performance Computer Design, Analysis, and Use, 2016. **72**(7): p. 2815-2831.

[9] Smeed, R.J., *Some statistical aspects of road safety research.* Journal of the Royal Statistical Society. Series A (General), 1949. **112**(1): p. 1-34.

[10] Andreassen, D.C., *Linking deaths with vehicles and population.* Traffic Engineering and Control, 1985. **26**(11): p. 547-549.

[11] UAB. *Ulastirma ve Altyapi Bakanligi*. 2021 [cited 2022 1 May]; Available from: https://www.uab.gov.tr/uploads/pages/bakanlik-yayinlari/ulasan-ve-erisen-turkiye-2021.pdf.

[12] Akgüngör, A. and E. Doğan, *Smeed ve Andreassen kaza modellerinin Türkiye uygulaması: Farklı senaryo analizleri.* Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 2008. **23**(4).

[13] Shaheen, S. and R. Finson, *Intelligent transportation systems.* Reference module in earth systems and environmental sciences, 2013.

[14] TUIK. *Türkiye Istatistik Kurumu*. 2022 [cited 2022 1 May]; Available from: https://data.tuik.gov.tr/Bulten/Index?p=Motorlu-Kara-Tasitlari-Mart-2022-45706

[15] *Traffic Announcements Data*. 2022 [cited 2022 14 June]; Available from: https://data.ibb.gov.tr/en/dataset/ulasim-yonetim-merkezi-trafik-duyuru-verisi/resource/1c043914-8a76-4793-bae9-c60a68c7d389.

[16] Koklu, M., I. Cinar, and Y.S. Taspinar, *Classification of rice varieties with deep learning methods.* Computers and electronics in agriculture, 2021. **187**: p. 106285.

[17] Koklu, M., R. Kursun, Y.S. Taspinar, and I. Cinar, *Classification of Date Fruits into Genetic Varieties Using Image Analysis.* Mathematical Problems in Engineering, 2021. **2021**: p. 4793293.

[18] Elsheikh, A.H., S.W. Sharshir, M. Abd Elaziz, A.E. Kabeel, W. Guilan, and Z. Haiou, *Modeling of solar energy systems using artificial neural network: A comprehensive review.* Solar Energy, 2019. **180**: p. 622-639.

[19] Taspinar, Y.S., I. Cinar, and M. Koklu, *Classification by a stacking model using CNN features for COVID-19 infection diagnosis.* Journal of X-ray science and technology, 2022(Preprint): p. 1-16.

[20] Chakraborty, A., D. Mukherjee, and S. Mitra, *Development of pedestrian crash prediction model for a developing country using artificial neural network.* International Journal of Injury Control and Safety Promotion, 2019. **26**(3): p. 283-293.

[21] Najafi Moghaddam Gilani, V., S.M. Hosseinian, M. Ghasedi, and M. Nikookar, *Data-Driven Urban Traffic Accident Analysis and Prediction Using Logit and Machine Learning-Based Pattern Recognition Models.* Mathematical Problems in Engineering, 2021. **2021**: p. 9974219.

[22] Tang, J., L. Zheng, C. Han, W. Yin, Y. Zhang, Y. Zou, and H. Huang, *Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review.* Analytic Methods in Accident Research, 2020. **27**: p. 100123.

[23] Koklu, M. and Y.S. Taspinar, *Determining the Extinguishing Status of Fuel Flames With Sound Wave by Machine Learning Methods.* IEEE Access, 2021. **9**: p. 86207-86216.

[24] Qin, Z., A.T. Wang, C. Zhang, and S. Zhang. *Cost-Sensitive Classification with k-Nearest Neighbors*. in *Knowledge Science, Engineering and Management*. 2013. Berlin, Heidelberg: Springer Berlin Heidelberg.

[25] Fiorentini, N. and M. Losa, *Handling Imbalanced Data in Road Crash Severity Prediction by Machine Learning Algorithms.* Infrastructures, 2020. **5**(7).

[26] Al-Doori, S.K.S., Y.S. Taspinar, and M. Koklu, *Distracted Driving Detection with Machine Learning Methods by CNN Based Feature Extraction.* International Journal of Applied Mathematics Electronics and Computers, 2021. **9**(4): p. 116-121.

[27] Zhou, X., P. Lu, Z. Zheng, D. Tolliver, and A. Keramati, *Accident Prediction Accuracy Assessment for Highway-Rail Grade Crossings Using Random Forest Algorithm Compared with Decision Tree.* Reliability Engineering & System Safety, 2020. **200**: p. 106931.

[28] Taspinar, Y.S., M. Koklu, and M. Altin, *Classification of flame extinction based on acoustic oscillations using artificial intelligence methods.* Case Studies in Thermal Engineering, 2021. **28**: p. 101561.

[29] Sarkar, S., R. Raj, S. Vinay, J. Maiti, and D.K. Pratihar, *An optimization-based decision tree approach for predicting slip-trip-fall accidents at work.* Safety Science, 2019. **118**: p. 57-69.

[30] Gutierrez-Osorio, C. and C. Pedraza, *Modern data sources and techniques for analysis and forecast of road accidents: A review.* Journal of Traffic and Transportation Engineering (English Edition), 2020. **7**(4): p. 432-446.

[31] Taspinar, Y.S. and M. Selek, *Object recognition with hybrid deep learning methods and testing on embedded systems.* International Journal of Intelligent Systems and Applications in Engineering, 2020. **8**(2): p. 71-77.

[32] TAŞPINAR, Y.S., I. Cinar, and M. Koklu, *Improvement of Football Match Score Prediction by Selecting Effective Features for Italy Serie A League.* MANAS Journal of Engineering, 2021. **9**(1): p. 1-9.

[33] Singh, D., Y.S. Taspinar, R. Kursun, I. Cinar, M. Koklu, I.A. Ozkan, and H.-N. Lee, *Classification and Analysis of Pistachio Species with Pre-Trained Deep Learning Models.* Electronics, 2022. **11**(7).

[34] Taspinar, Y.S., I. Cinar, and M. Koklu, *Prediction of Computer Type Using Benchmark Scores of Hardware Units.* Selcuk University Journal of Engineering Sciences, 2021. **20**(1): p. 11-17.

[35] Ropelewska, E., X. Cai, Z. Zhang, K. Sabanci, and M.F. Aslan, *Benchmarking Machine Learning Approaches to Evaluate the Cultivar Differentiation of Plum (Prunus domestica L.) Kernels.* Agriculture, 2022. **12**(2).