Towards Intelligent and Dynamic Road Speed Adaptation Model in Smart Cities

Pierre E. ABI-CHAR
College of Engineering and Technology
American University of the Middle East, Kuwait
Pierre.abichar@aum.edu.kw

Abduladhim ASHTAIWI
College of Engineering and Technology
American University of the Middle East, Kuwait
Abduladhim.ashtaiwi@aum.edu.kw

***Abstract*—Smart cities integrate data collection and communication technologies to operate more efficiently in order to provide better services to citizens, improved road traffic situation, better economic development, etc. However, safety still a main concern of road traffic as road crashes pose serious threats to road users causing death and injury all over the world. Road traffic safety is the process of implementing measures to reduce the number of crashes, death and severe injuries. In this paper, using Machine Learning (ML) algorithms, Off-line, we train and test the Intelligent and Dynamic Adaptation Model ($IDAM$) on the road traffic data collected for 14 years of car crashes in United Kingdom (UK). IDAM then on-line determines the optimal speed limit for each road segment based on many predictors such as weather, road type, light condition, and many others, stated later. We train and test IDAM using three different machine learning algorithms: Artificial Neural Network (ANN), Decision Tree (DT), and Linear Regression (LR) with Stochastic Gradient Descent (SGD). Using the testing dataset, IDAM achieves prediction accuracy of around 96%. Implemented in an operation algorithm, IDAM can continuously adapt the road speed limit to an optimal value as the predictors change on-line.**

***Index Terms*—Intelligent Speed Adaptation, Road Speed Management, Speed Limits, Smart Cities, Road Safety.**

I. INTRODUCTION

The rapid development of road transportation systems have resulted in a substantial increase in the number of fatal road crashes. Every year, many people are killed or injured in car accidents and collisions [1]. In 2016, more than 1.2 million people died in road crashes [2].

There are many reasons for road accidents including weather condition, visibility, road type, speed limits, road conditions, etc. Speeding has been identified as a key risk factor in road traffic crashes and has been demonstrated in several studies over the years. According to Finch et al. [3], the number of accidents resulting in injury can be decreased 3% when reducing the speed by 1 km/h. Nilsson [4], stated that decreasing the speed of a vehicle from 55 km/h to 50 km/h, can decrease the number of fatal and serious-injury accidents by almost 25%. Warner and Aberg [5] declared that almost one-fifth of road crash deaths would have survived if drivers did not exceed the speed limits.

Traditional road speed control systems such as, police enforcement and speed cameras, seek to prevent speeding by punishing violating drivers using fines or through license point system. However, these types of enforcement are not effective and tend to be quite inefficient. For example, drivers tend to reduce speed only near the enforcement area [6].

Recent and ongoing development systems provide possible opportunities in assisting drivers to adhere to speed limits. One system that is increasingly gaining attention is the Intelligent Speed Adaptation (ISA) system that can warn the driver when the vehicle exceeds the recommended speed limit. In addition, several advanced solutions relying on wireless communication standards have been proposed to provide Intelligent Transportation Systems (ITS) in smart cities. To the best of our knowledge, few of them focused on optimizing and solving the speed limits impact problem on roads and road segments.

Therefore, an advanced speed adaptation system where a speed limit of a road segment is set according to the current conditions such as weather, traffic, and road types is required. Such a system can improve safety and significantly, decrease road crashes levels.

In this paper, we propose an Intelligent Dynamic Adaptation Model ($IDAM$) that continuously assigns an optimal speed limit for road segments as response to the change of many car crash related attributes. Using Machine Learning (ML), IDAM is trained and tested, off-line, on real dataset collected by the Department of transportation (DoT) in United Kingdom (UK) which covers many Road Safety Data (RSD). DoT data collection covers 14 years of car crashes, for each car crash it records 33 information, in total, it contains 1.6 million records. To consolidate the prediction accuracy of IDAM, we use three different ML algorithms: Artificial Neural Network (ANN) regression, Decision Tree (DT) regression, and Linear Regression (LR) with Stochastic Gradient Descent (SGD). Using simple operation process, IDAM can, on-line, assigns an optimal speed limit for road sectors.

The rest of the paper is organized as follows. Section II summarizes related work. In section III, we introduce preliminaries that are used in the rest of this paper. In Section IV, we discuss and present the proposed speed adaptation model in details. Section V presents IDAM performance analysis with an on-line implementation example. Finally, conclusion and future work are presented in Section VI.

II. Related Work

Several speed management algorithms have been proposed for smart city applications to ensure road traffic safety. As our $IDAM$ method is used to find road speed limits based on the surrounding environmental conditions and the corresponding road segments, works related to these two problems are considered.

In 2010, Usman et al. [7], [8] have studied the important role of winter maintenance operations on road safety using a road surface condition index that depends on the level of applied winter operations instead of the commonly used method of friction measure. They found that road surface index plays an important role influencing both traffic volume and speeds on highways. In 2011, Yokota et al. [9] uses floating car data from 300 trucks running in Osaka, Japan and provides some information such as the average speeds for each route. Floating car data are very useful for predicting future road situations. However, authors did not consider the influence of bad weather in details. In 2012, Donaher et al. [10] developed two different models for speed on rural and urban highways in Ontario using wind speed, visibility, snow, and temperature. These two models were developed using linear regression approach. They found that operating speeds are also quite sensitive to adverse weather. In 2012, Lai et al. [11] present a study to predict the impacts of various forms of ISA and to assess whether ISA is viable in terms of benefit-to-cost ratio. They found that ISA is predicted to lead to savings of 30% in fatal crashes and 25% in serious crashes. However, several other factors need to be considered for precise speed control such as road surface, road type, weather conditions, etc. In 2015, Tanimura et al. [12] proposed a novel method to predict the speed of vehicles on each road segment in snowy cities by using multiple regression function. They classified winter urban road segments into busy road segments with frequent traffic jam, road segments with enough traffic amount, and road segments without traffic jam. However, this research provides only analysis on the vehicle speeds from 8:00 a.m. to 8:00 p.m. and they do not consider road type conditions. In 2017, Higashino [13] presented a method to predict the speed of vehicle in snowy urban roads for each road segment using the regression analysis technique. He considered weather information and vehicular traffic data as the dominant factors in the regression model equation. According to Higashino, the proposed method cannot predict the speed for roads with less than 10 floating cars, because the accurate average speed cannot be calculated.

III. Preliminaries

This section provides background details about technologies and techniques used in this work. Our model creation goes through three stages: data preprocessing, model building using machine learning, and lastly quantifying the prediction accuracy. This section gives more details on the aforementioned stage. The following are some defined notations used in this work. We use (X, Y) to refer to the used dataset. X is a matrix with dimension $(m \times n)$, where m is the number of samples and n is the number of features. Y is the class vector. We use X_c^r to refer to the c^{th} feature and the r^{th} sample. Y^r refers to the class of r^{th} sample.

A. Data Preprocessing

It is always known that real world data can contain duplicated data, missing target or class value, etc. Therefore, data preprocessing is a very important stage to improve the prediction accuracy. Next three preprocessing techniques are used in this study:

1) Scaling: Real world data created models come with different min-max ranges. Feeding data, as is, without scaling, cause the ML algorithms to create models mostly affected by big numbers assuming that they are the most important features. There are many scaling techniques used in machine learning, in this work a standard scaling is used, i.e., $sX_c = \dfrac{(X_c - \mu_c)}{\sigma_c^2}$ where sX_c is the scaled feature, X_c is the feature before scaling, μ_c is the mean of the feature and σ_c^2 is the standard deviation of the feature column.

2) Outliers: Outliers are defined as a data point that located far away from the mean. There are many mathematical techniques, in this work, we used quantile() function which brings outliers data points closer to data distribution.

3) Data Encoding: Real world data comes in different types such as categorical or numerical. ML algorithms works only on numerical data. Transforming data from categorical to numerical is called Encoding. There are many techniques that are used to transform data from categorical to numerical. In this work we use *Label Encoding* (LE) technique.

B. Machine Learning (ML)

To consolidate our findings for the proposed model, we use three different ML algorithms: Artificial Neural Network (ANN) regression, Decision Tree (DT) regression, and Linear Regression (LR) with Stochastic Gradient Descent (SGD). Next a background for each of these algorithms is given.

1) Artificial Neural Network (ANN) regression: ANN algorithm is a network of connected Processing Units (PUs) called Artificial Neurons (ANs). ANN goes through thousands of feed-forward and back-propagation operations to minimize the prediction error. Figure 1 shows the structure of ANN. The first and last columns of the neural network structure are called input and output layers, respectively. Columns in between are called hidden layers. ANN can have one or many hidden layers. The number of units of input layer is determined by number of features (n). However, the number of output units is determined by the problem domain under consideration, i.e., classification versus regression. As stated, the problem considered in this study is regression problem; hence, the output layer is composed of one unit and has identity function.

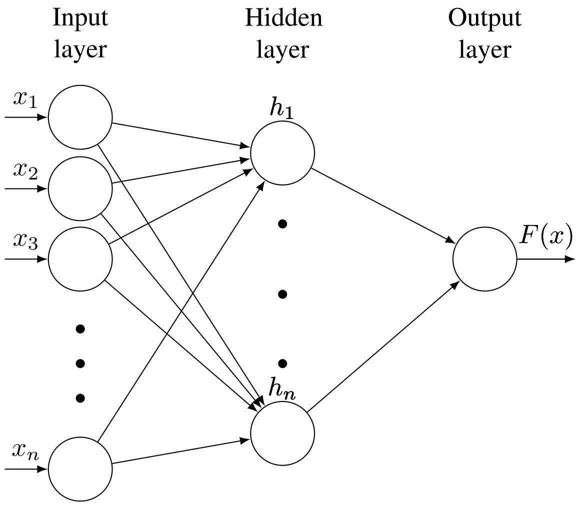


Fig. 1. Artificial Neural Networks (ANN) Regression

2) Decision Tree (DT) regression: Using the training dataset, the decision tree algorithm builds decision rules following top-down tree like structure. The most important feature is selected as the first branching node and is called the root-node. In DT regression, the node with highest Standard Deviation Reduction (SDV) value is selected as the root-node. SDV is calculated using Equation 1,

$$SDV(T, X_c) = S(T) - S(T, X_c) \tag{1}$$

where $S(T)$ is the standard deviation of the target attribute and $S(T, X_c)$ is equal to

$$S(T, X_c) = \sum_{g \in X_c} P(g)S(g) \tag{2}$$

Where, X_c is a specific feature in the dataset and $P(g)$ is the ratio of the sub-category with respect to the dataset, and $S(g)$ is the standard deviation with respect to the target/class. The feature with the highest $SDV(T, X_c)$ is selected as the root-node. Using the root-node branches, the whole dataset is splitted into sub-datasets. Similar to finding the root-node, Equation 1 and 2 are then applied to the sub-dataset, at each branch of the root-node, to find next branching feature. The process of splitting and finding most important feature is repeated until the whole tree structure is created.

3) Linear Regression (LR) using Stochastic Gradient Descent (SGD): Linear learning regression algorithms try to fit a linear model, with least cost error, on the data points. Using the learned model, then, the algorithm predicts the output given multiples predictors. A regression model involving multiple variables can be modelled as:

$$h_\theta(X^{r_i}) = \theta_0 + \theta_1 X_1^{r_i} + \ldots + \theta_n X_n^{r_i} \tag{3}$$

Equation 3 is called hyperplane model. Equation 4 models the cost function.

$$J(\theta_0, \theta_1, \ldots, \theta_n) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(X^{r_i}) - y^i \right)^2 \tag{4}$$

Given the hyperplane, Equation 3, and the cost function, Equation 4, the problem becomes minimizing $J(\theta_0, \theta_1, \ldots, \theta_n)$.

Linear model uses Gradient Descent (GD), Equation 5, to find the optimal model parameters that achieves the minimum cost function.

$$\theta_j := \theta_j = \alpha \frac{\delta}{\delta \theta_j} J(\theta_0, \theta_1, \ldots, \theta_n) \tag{5}$$

Since this model uses stochastic GD, the algorithm takes k random samples out of m samples as modelled in Equation 6.

$$\theta_j := \theta_j = \alpha \frac{\delta}{\delta \theta_j} \left(\frac{1}{2k} \sum_{i=1}^{k} \left(h_\theta(X^{r_i}) - y^i \right)^2 \right) \tag{6}$$

IV. Intelligent and Dynamic Adaptation Model (IDAM)

The proposed intelligent and dynamic adaptation model dynamically assigns an optimal speed limit for road-sectors based on many on-line car crash attributes. This section gives the details of all IDAM building stages.

A. Data set preparation

The data used to build IDAM is collected by the Department of Transportation (DoT) in United Kingdom (UK) which covers many Road Safety Data (RSD). DoT dataset covers 14 years of car crashes, for each car crash it records 33 information, in total, it contains 1.6 million car crash records. To know the whole list of captured car crash information please refer to DoT website [14]. Due to missing data, in this work, we only use 446804 car crash samples to create IDAM model. With regard to feature selection, we apply feature completeness, being informative or not, and whether has an impact or not on the speed limit as a selection criterion. Hence out of 33, 11 information/features, per accident, are selected. Table I lists the selected features. The column "*value*" of Table I lists the possible values of that feature. For instance, "*Accident Severity*" as a feature/attribute can be assigned a value as: "*Fatal*, *Serious*, or *Slight*". For shorthand writing, we refer to these features using codes, F_1, F_2,...F_{13}, as shown in Table I. Hence, $(446804, 11)$ is the dataset matrix used to build IDAM.

B. Target Creation

By observing the features listed in Table I, we can identify that F_1, F_2, and F_3, as an output or a result of an accident. In other words, these three features cannot be identified as the leading cause of an accident. Rather, it is obvious that each feature describes the accident output conditions, i.e., Accident Severity, Number of Vehicles, and Number of Casualties. Hence, in this work, these three features are used to create the target/class column. We refer to the target as *risk level* of the accident. Based on the collected data, the min-max

TABLE I
USED FEATURES/PREDICTORS

Code	Features	Values
F_1	Accident Severity	Fatal, Serious, or Slight
F_2	Number of Vehicles	Represents the number of vehicles involved in the accident
F_3	Number of Casualties	Number of deaths
F_4	Time	Time of the day when the accident happened
F_5	Speed Limit	The speed of road-sector where the accident happened
F_6	Day of Week	The day when the accident happened. Values: Sun, Mon,..., or Sat.
F_7	Urban or Rural Area	Urban, or Rural Area.
F_8	Road Type	Road-sector name. Values: Single carriageway (SC), One way street (OWS), 'Roundabout (RA), Dual carriageway (DC), or Slip road (SR).
F_9	Pedestrian Crossing-Human Control	None within 50 meters, Control by other authorized person, or Control by school crossing patrol.
F_{10}	Pedestrian Crossing-Physical Facilities	Pedestrian phase at traffic signal junction, No physical crossing within 50 meters, non-junction pedestrian crossing, Zebra crossing, Central refuge, or Footbridge or subway.
F_{11}	Light Conditions	Darkness: Street lights present and lit, Daylight: Street light present, Darkness: No street lighting, Darkness: Street lights present but unlit, Darkness: Street lighting unknown.
F_{12}	Weather Conditions	Fine without high winds (FWN), Raining without high winds (RWN), Snowing without high winds (SWN), Raining with high winds (RW), Fine with high winds (FW), Fog or mist (FM), or Snowing with high wind (SW).
F_{13}	Road Surface Conditions	Dry, Wet/Damp, Frost/Ice, Snow, or Flood (Over 3cm of water).

values of the three features range from small to large values. Where small represent less sever and large represents high sever. If linearly multiply these features with each other we still have min-max range, where min represents less sever and max represents high sever. Before we do the multiplication we fix the outlier problem by applying the technique explained in Section III-A. To fix the skew problem the target column after multiplication, the result column is further transformed to normal distribution by applying $log(-)$ function. Next, the result column is transformed to a percentage scale, i.e., from 0% to 100%, where 0% represents low car crash risk and 100% represents high risk.

C. Feature Encoding

As shown in table I, the remaining 10 features, i.e., features from F_4 to F_{13}, are considered model predictors. Features F_4 and F_5 are of type numerical data, while the remaining features, from F_6 to F_{13}, are of type nominal data. For numerical data we apply standard scaling as explained in Section III-A. For nominal data, we apply Label Encoding technique explained in Section III-A, to transform it to numerical data. Table II shows a randomly selected two rows after performing data encoding, scaling, and the respected risk level mapping.

TABLE II
TWO SAMPLES RANDOMLY SELECTED FROM THE PREPROCESSED DATASET

F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	Risk (%)
-1.262	-0.598	7.0	1.0	0	2	2	4	1	0	33.2
1.223	-0.598	5.0	1.0	3	2	2	2	1	4	19.8

D. IDAM Model Building

After preparing the data, the dataset is divided to 70% for training and 30% for testing the model prediction accuracy. The training dataset is used to train the three ML algorithms: ANN regression, DT regression, and LR with SGD. After many trials of algorithms parameters changing, the following algorithms parameters, shown in Table III, are proven to produce the highest prediction accuracy.

TABLE III
ALGORITHMS PARAMETERS SETTINGS

ML	Parameters values
ANN	Hidden-layer-sizes=(10,10,10), activation='relu', solver='adam', alpha=0.001,batch-size='auto', Learning-rate='constant', learning-rate-init=0.01, max-iter=1000, shuffle=True, Early-stopping=False.
DT	criterion='mse', splitter='best', max-depth=5, min-samples-split=3, and min-samples-leaf=1
LR	loss='squared-loss', fit-intercept=True, max-iter=1000, shuffle=True, early-stopping=False, and n-iter-no-change=5

E. Quality of Predictions Assessment

To assess the prediction accuracy of the created model, we used 30% of the dataset to be used for performance evaluation. Table IV shows the prediction accuracy for the three ML algorithms. As ANN slightly has a higher prediction accuracy compared with the other algorithms, we considered the model generated from ANN as our model for this work to control the speed limit over different road-sectors of the smart city.

TABLE IV
ACCURACY COMPARISON

Machine Learning Algorithm	Accuracy (%)
Artificial Neural Network (ANN)	96.95
Decision Tree (DT)	96.83
Linear Regression (LR) with SGD	96.88

V. IDAM PERFORMANCE ANALYSIS AND ON-LINE IMPLEMENTATION ALGORITHM

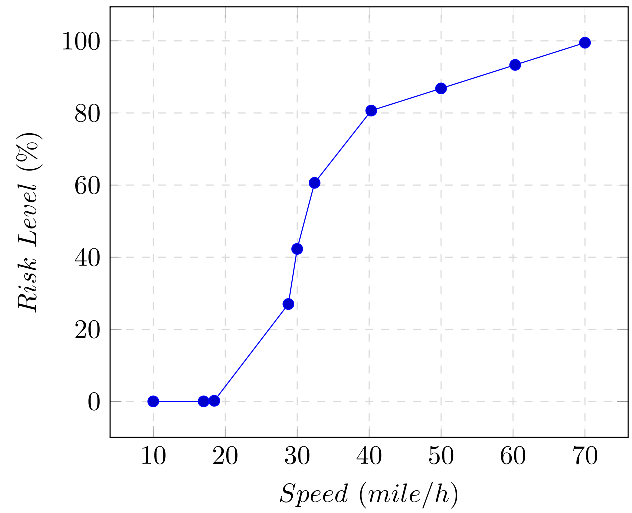
As explained, the model takes 10 predictors/features, features from F_4 to F_{13}, and predicts the risk percentage. Table V shows how different values of the predictors can have different risk level percentage values. For instance, the last row, in table V, has higher risk than the first row. That is because, from the learned knowledge during the training phase, the feature values found in the last row tend to produce higher risk level. Referring to Table I each feature can assume different values, for example, F_6, $Day\ of\ the\ week$, can be $Sunday$, $Monday$, etc. If we consider all possible values of all features in Table I, we can have 2.2×10^9 possible combination at the input of the model. For each combination we may have different risk level. Table V shows only 5 possible combinations and their respected risk levels.

TABLE V
DIFFERENT RANDOM VALUES OF PREDICTORS AND THEIR PREDICTED RISK LEVEL

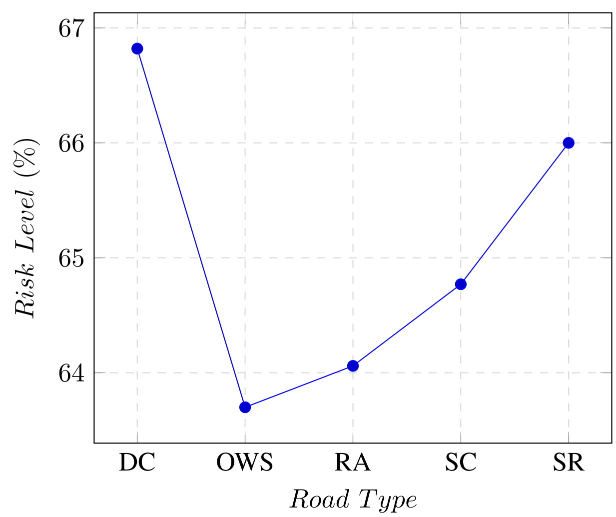
F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	Risk Level
0.35	-0.598	3	2	3	2	3	4	4	4	17.99368
1.854	-0.598	1	1	3	2	4	2	0	4	9.76121
-0.687	-0.598	1	1	3	2	5	4	1	0	0.000016
-0.073	-0.598	3	2	3	2	2	4	1	0	21.51778
0.063	1.572	6	2	3	2	2	4	4	4	46.24023

Further, to show the impact of individual feature on the risk level, we generated random values for the nine features (and we make them fixed during the prediction) and only change feature under examination. Figure 2 (A) shows how different speed limits impact the risk level. It is noticeable that higher speed limit bring higher risk levels. These risk level behavior characteristics may slightly change under any other different combinations of the nine features. Similarly, Figure 2 (B) shows the impact of different road types on the risk level percentage. Figure 3 (A) depicts the impact of different weather conditions on the risk level. Similarly, under any other combination of the other features, IDAM may produce different impact characteristics. Figure 3 (B) shows how the risk level is changing under different values of the minutes of the day. Due to the space limit we only showed the impact of only 4 features, however, other features, not shown in Figures 2 and 3, have their own impact on the risk level as well. Hence, the created model ties the impact of all features in one model, Equation. The Model then outputs the risk level based on any input values.

Figure 4 shows the IDAM implementation algorithm to find the optimal speed limit for different road sectors. As identified in the dataset used to build IDAM, roads classified under have 5 different road types: "Single carriageway", "One way street", "Roundabout", "Dual Carriageway", or "Slip Road". These road sector types, in general, can be considered as the constituent components of real world roads. The algorithm starts with a request to find an optimal speed limit for specific road sector (e.g., for "One way Street") is initiated. Next, the algorithm gets the current values for the features from F_4 to F_{13} except F_5 (the optimal speed limit) which is the output value in this case. Note that, current feature values mean current time, weather, day, light, and etc, for the road sector under consideration. Next, from speed limit set, the algorithm then assigns a speed limit value to F_5. The algorithm then calls IDAM to predict the risk level. If the risk level is within the accepted threshold then this speed limit is assigned to this road sector. If not, the algorithm iterates through the set of other speed limit values and for each speed limit value finds the risk



(a) Fig. 2 (A)



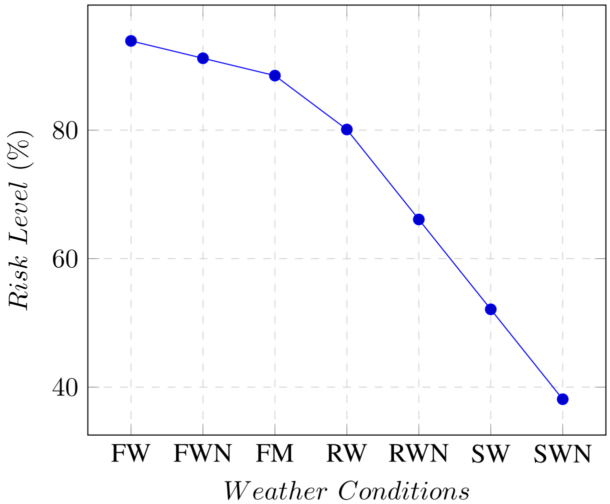
(b) Fig. 2 (B)

Fig. 2. Risk level (%) for different Speed limit values (mile/h) and at different road types

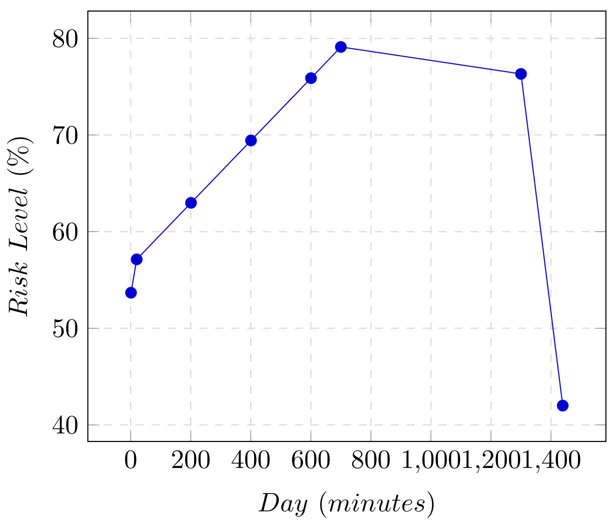
level. The algorithm stops iterating when the risk level is at the desired threshold. This speed limit value is then assigned to the road sector. The smart city continuously monitors any changes in the features and use IDAM accordingly to update the speed limit.

VI. CONCLUSION

In this paper, we proposed an Intelligent Dynamic Adaptation Model ($IDAM$) for smart cities. To consolidate the prediction accuracy of IDAM, we used three different M-L algorithms: Artificial Neural Network (ANN) regression, Decision Tree (DT) regression, and Linear Regression (LR) with Stochastic Gradient Descent (SGD). Using these ML algorithms, IDAM is trained and tested, off-line, on real dataset collected by the Department of transportation (DoT) in United



(a) Fig. 3 (A)



(b) Fig. 3 (B)

Fig. 3. Risk level (%) for different weather conditions and at different time (minutes)

Kingdom (UK). DoT dataset covers 14 years of car crashes. Evaluating on the testing dataset, IDAM achieved prediction accuracy of around 96%. Finally, an implementation process of IDAM is conducted. As a response to any on-line speed limit predictors' change, IDAM continuously computes the optimal speed limit for road segments.

References

[1] World Bank, World Health Organization & Global Road Safety Partnership, 2008, Speed management: a road safety manual for decision-makers and practitioners. Geneva: Global Road Safety Partnership.

[2] Research and Innovative Technology Administration, Bureau of Transportation Statistics, National Transportation Statistics, www.bts.gov/publications/national_transportation_statistics.

[3] D.J. Finch, P. Kompfner, C.R. Lockwood, G. Maycock Speed, Speed Limits and Accidents, Project Report 58, S211G/RB Transport Research Laboratory (TRL), Crowthorne, UK (1994).

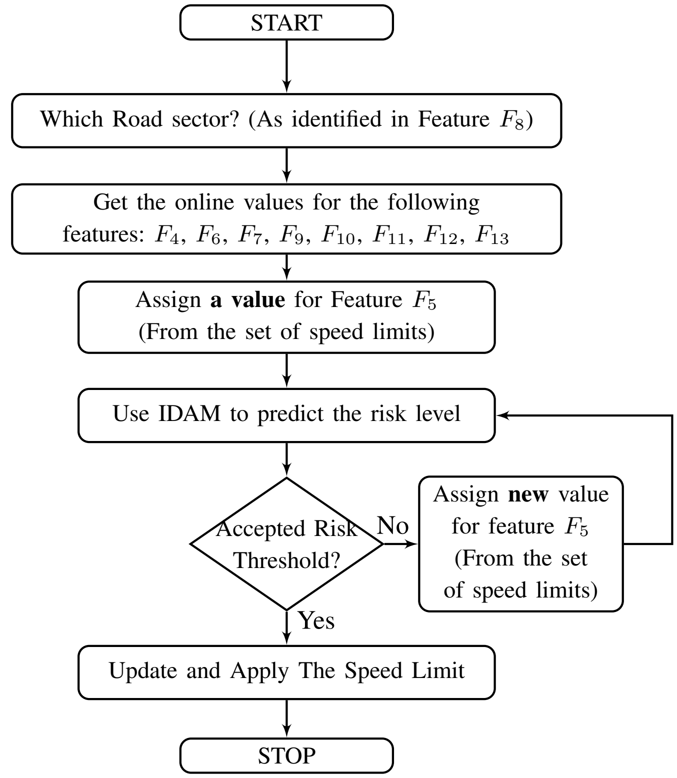


Fig. 4. Optimal Speed Limit Using IDAM

[4] G. Nilsson, Speed changes and traffic safety effects The Power Model, VTI notat 76-2000, Swedish National Road and Transport Research Institute (VTI), Linkoping (2000)

[5] H.W. Warner, L. Aberg, Drivers' beliefs about exceeding the speed limits Journal of Transportation Research Part F: Traffic Psychology and Behavior, 11 (5) (2008), pp. 376-389.

[6] N. Teed, A.K. Lund, R. Knoblauch, The duration of speed reductions attributable to radar detectors, In proceeding of the Journal of Accident Analysis and Prevention, 25 (2) (1993), pp. 131-137.

[7] Usman, T., L. Fu, and L. Miranda-Moreno. Quantifying safety benefit of winter road maintenance: Accident frequency modeling. Accident Analysis and Prevention, 2010, pp. 1878-1887.

[8] Usman, T., L. Fu, and L. Miranda-Moreno. Winter Road Safety: Effects of Weather, Maintenance Operations, and Road Characteristics. in the International Conference on Transportation Information and Safety, 2011, pp. 1152-1159.

[9] T.Yokota, D.Tamagawa,"Constructing Two-Layered Freight Traffic Network Model from Truck Probe Data, "International Journal of Intelligent Transportation Systems Research, vol. 9, no. 1, pp. 1-11, 2011.

[10] Donaher, G., L. Fu, and T. Usman. Quantifying the Mobility Effects of Winter Road Conditions. in the 9th International Transportation Specialty Conference, Edmonton, Alberta, Canada, 2012

[11] Lai, F.; Carsten, O.; Tate, F. How much benefit does Intelligent Speed Adaptation deliver: An analysis of its potential contribution to safety and environment. Accid. Anal. Prev. 2012, 48, 63-72.

[12] Tanimura R, Hiromori A, Yamaguchi H, Higashino T, Umedu T (2015) Prediction of deceleration amount of vehicle speed in snowy urban roads using weather information and traffic data. In: Proceedings of IEEE 18th international conference on intelligent transportation systems (ITSC2015), pp 2268-2273.

[13] T. Higashino, Edge computing for cooperative real-time controls using geospatial big data. In Proceedings of Smart Sensors and Systems, Springer, p. 441-466, 2017.

[14] https:Department for Transport and transport agencies, //data.gov.uk/dataset/, roadacc.stats@dft.gov.uk; 17 Dec., 2019.