INTERNATIONAL JOURNAL OF INNOVATIVE STUDIES
A MONTHLY, INDEXED, REFEREED AND PEER REVIEWED OPEN
ACCESS INTERNATIONAL JOURNAL
VOL.6. ISSUE.1, 2026 (JANUARY)

RESEARCH ARTICLE

ISSN: 3027-2173

ACCIDENT SEVERITY PREDICTION FOR ROAD TRANSPORTATION SYSTEM USING MACHINE LEARNING

UMEJURU DANIEL[1], IGBUDU KINGSLEY EZEBUNWO[2]
[1]Department of Computer Science, University of Port Harcourt, Choba, Nigeria
[2]Department of Computer Science, Rivers State University, Port Harcourt Nigeria
E-mails: daniel_umejuru@uniport.edu.ng, kingsley.igbudu@ust.edu.ng
D.O.I.: 10.5281/zenodo.18458959

ARTICLE INFORMATION

Received: 1st December, 2025

Accepted: 30th December 2025

Published: 15th January 2026

KEYWORDS: Classification, Prediction, Traffic, Accident Severity, Machine Learning

Publisher: Empirical Studies and Communication – (A Research Center)
Website: www.cescd.com.ng

ABSTRACT

Road accidents constitute some of the most horrifying experience one can have and can sometimes leave an indelible mark on victims for the rest of their life. The constant occurrence of accidents involving cars and collisions is an immediate consequence of road conditions, related environmental factors, and the carelessness of drivers involved. Professionals and agencies can utilize accident severity prediction models to obtain insight into the factors influencing road traffic incidents, and this can help predict the degree of severity of traffic accidents. Machine learning algorithms can be used to determine accident trends and predict cases of fatalities, major injuries, or minor injuries. The aim is to provide a method for predicting the degree of severity caused by a traffic accident using a ML algorithm. We created a prediction model using support vector machine (SVM). Road traffic accident (RTA) datasets obtained from the Kaggle website served as the data in the experiment. The artificial neural network was designed with the decay concept in mind, which penalizes the model to acquire knowledge from larger weights and forces it to learn from smaller weights, leading to a simpler architecture. The fine-tuned hyper-parameter variables enhanced the classification results and assisted in minimizing the model's complexity with the aid of the penalty term. The SVM yielded a recommendable result of at 89.0% efficiency which shows that the model performed excellently as expected.

1.0 Introduction

Reducing the number and complexities of road accidents requires the development of an effective predictor of accident severity. We can notify nearby drivers or persuade them to take the least dangerous path by estimating the chance of a traffic collision in a particular region. Nevertheless, because there are so many interrelated factors that could contribute to a traffic accident, it can be challenging to predict with accuracy the likelihood of one. In this regard, the frequency of collisions while driving varies greatly across the nation. Weather-related visibility and traffic efficiency reductions increase the chance of an accident. There is likely a variation in the frequency of incidents on the roads based on the drivers' tiredness levels throughout the day. It has been challenging to reliably estimate the possibility of traffic accidents in real time although risk indicators for traffic accidents have received a lot of attention. ML approaches can be used by computer models to learn and determine the importance of features for probabilities. The arrays of numbers in a space with multiple dimensions that are used in machine learning are termed features. (Aassi et. al 2020)

The intricate interplay involving one or more of the four main factors—person, vehicle, road, and environment in order to effectively determine the severity of traffic accidents; it has been determined that the human element is the most significant but also the most difficult to modify (Fu et al.,2022). Accidents on the roads cause substantial losses for both the concerned parties and the nation in terms of infrastructural damage, lost production, and payouts from road accident funds. Achieving global environmental objectives and making traffic safety and accident prevention a priority in the management of transportation are contingent upon reducing or preventing these numerous losses (Yan and Shen 2022). Every road tragedy results in the collection of an accident report, which includes various accident characteristics and can be utilized to look into the incident's potential cause at that specific segment of the road. Nevertheless it comes to the accessibility of trustworthy accident data, the majority of developing and underdeveloped nations are falling beyond the rest of the globe (Mukherjee and Mitra 2022). Stretches of road with a high frequency of accidents can also be identified from collision reports, and these areas are then the topic of an accident study that may entail the expert reconstruction of accident scenes. Replicating the real-world behaviors of the drivers and the technical functioning of the vehicle that caused the accident can be costly and difficult when reconstructing a scene from an accident. Therefore, for an existing road or a newly constructed road network, the utilization of accident traffic data and analytical tools could be helpful in anticipating and preventing a future traffic accident.

Accidents on the roads account for a sizable share of the reported major injuries each year(Islam et al 2022). Nevertheless, it is difficult to pinpoint the precise circumstances that trigger such an incident, making it problematic for the road authority to effectively address the frequency and seriousness of traffic accidents (Tarlochan et al.,2022). Additionally, studies have demonstrated the critical roles that ecological, vehicle, road, and human variables play (Jeong et al.,2022). Human variables include driving experience, gender, and age (Pervez et al.,2021). Conversely, the engine capacity, make and model, carrying capability, and articulation are all considered vehicle variables (Yang et al.2022). Road features include the kind of road, its state, its class, its geometry, and its maximum speed limit (Rabbani et al., 2022). In addition, environmental variables comprise the day of the week, the climate, and the lighting (Chen et al., 2019). Human natures have been well studied and numerous countermeasures have been implemented (Bonnet et al.,2018). Further factors must be investigated, because it is challenging to focus on just one aspect, the following concerns about traffic accidents continue to arise: What are the direct and indirect contributing reasons to traffic accident fatalities, and (ii) what are the tactics employed to prevent such accidents in the future? To address these issues and provide particular insights into how the risk score affects the likelihood that a driver will be involved in a fatal or serious accident based just on information obtained from personal and vehicle data.

In remote areas, the likelihood of just one car accident occurring is higher than that of two collisions between cars (Weng et al.,2021). This is achieved through the use of negative values for collisions that happen on the road and favorable parameters for flipping accidents.

There is a greater chance of a severe accident when the collision happens off the road. Nonetheless, in metropolitan regions, collisions involving a single vehicle and many vehicles are noteworthy, while collisions involving animals have no bearing on the extent of the incident. The following criteria are used to classify road accidents according to the degree of seriousness based on the majority casualties involved, whether they were slight, serious, or fatal:

(a). **Slight injury**: a minor injury that needs to be treated at the scene, such as a sprain, bruise, cut, or laceration that is not considered to be severe.

(b). **Serious injury**: an injury that requires hospitalization as a "in-patient," or any of the following kinds of injuries, regardless of whether detention is necessary: fractures, concussions, internal injuries, crushing, severe cuts and lacerations, severe general shock necessitating medical attention, injuries that result in death 30 days or longer after the accident

(c). **Fatal**: death as a result of wounds received over a thirty-day period following the incident.

(d). **Damage only**: There was damage to the car. It excludes injuries and fatalities.

2.0 Conceptual Framework

One of the contributing reasons to traffic accidents is speeding; at greater rates of speed, a driver's time to react to an unforeseen situation is shortened, and the magnitude of an impact is increased (Yuan, et al.,2021). The subsequent text, which was released by the Organization of British Drivers, illustrates the significance of slower speeds: The only thing that almost all traffic accidents have when it comes common is that they could have all been prevented if everyone involved had known for sure, even only a few seconds beforehand, that something unfortunate was going to happen. Reduced speeds consequently give you that extra few seconds of time. A scalar variable with an importance component, speed is used to assess the efficiency of traffic flow. It can be described as the rate at which something moves expressed as the distance traveled in a given amount of time, usually expressed in kilometres every hour. Velocity is the vector component that corresponds to speed (Yu, et al.,2019). While speed and velocity are measured using the same physical units, speed lacks the direction components that velocity possesses.

According to a report by the Transportation Research Laboratory, collisions occur more frequently depending on how quickly traffic moves on aggregate. The study also came to the conclusion that drivers traveling faster are linked to noticeably higher crash participation than drivers traveling slower, and that the incidence of crashes rises roughly with the square of normal road speed. It is also observed that accidents decrease by 2–7% for every 1 mph decrease in the mean speed. More precisely, the percentage of accidents that occur on low-average-speed urban roads is 6%; on medium-speed urban roads and lower-class rural major highways, it is 4%; and on high-speed urban roads and rural main roads, it is 3%. There are different speeds limits on the road, and they vary depending on the location. The variation in the speed restriction that was posted in that region was what determined this discrepancy. The 85th percentile speed was determined in order to reach the posted speed restriction.

2.1 Road intersection accident prediction models

A great deal of the models put forth in the scientific literature are for signalized junctions, despite the fact that substantial study has been done on the prediction of accidents at junctions(Hauer et al.2019). Nicholson and Turner (2020) employed a smaller number of models for unsignalized junctions as they were unable to incorporate categorical variables. There are quite a few distinct types of APMs that employ traffic flow as their sole input; a few typical model forms are presented in equations 2.6 and 2.8(Nelder and Wedderburn 2019):

$$E\{k\} = \boldsymbol{\alpha}(F_1 + F_2)^{\beta} \quad 2.1$$

$$E\{k\} = \boldsymbol{\alpha}(F_1 F_2)^{\beta} \quad 2.2$$

$$E\{k\} = \boldsymbol{\alpha}(F_1 + F_2)^{\beta_2} (\frac{F_2}{F_1})^{\beta_2} \quad 2.3$$

Where E{k} is the expected number of road accidents in a specific interval of time

F_1, F_2 as the minor and major traffic flows and $\boldsymbol{\alpha}$, β_1, β_2 are the estimated coefficients

There are differences in the accuracy of the computational forms used to predict accidents at crossings between equations 2.4 and 2.6. The choice of model form frequently depends on the analyst's judgment and data features. According to (Mountain et al., 2020) the following model form is the most appropriate for road intersection modeling:

$$E\{k\}=\boldsymbol{\alpha} F_1^{\beta_1} F_2^{\beta_2} \quad 2.4$$

The variant of equation 2.7 that is most frequently encountered in safety publications is the model form. Additionally, Equation 2.7 offers two key benefits and the model operates on the assumption that "no collisions, no traffic flows" (Green 2019). If traffic is nonexistent, it would be reasonable to assume that accidents would not happen. Furthermore, the model is predicated on the idea that there isn't a straight line connecting traffic flow and incidents. This association proved to be well-researched, verified, and has a long history (Miaou 2020). Equation 2.7 is capable of being used for conflicting movements that depends on the desired result, and the overall approach can flow in predicting the number of accidents for an intersection. The parameters F_1 and F_2 are typically defined as the sum of the two approaches flows in any one direction in order to anticipate the overall number of incidents at an intersection. These models typically have estimated coefficients (P) ranging from 0.2 to 1. Numerous variables affect the value of these predictions, including the intersection (urban versus rural) and the features of the neighborhood population. Depending on the description provided by different researchers, there are anywhere from 15 to 25 separate kinds of collisions at junctions. As predicted by the model mentioned in the preceding paragraph, the overall number of fatalities for every component of a junction should, in theory, equal the sum of the estimated accidents in the various categories. The primary disadvantage of this method is the immense amount of data required to create these models. Every vehicle's maneuver engaged in an accident needs to be recorded. When dealing with a small number of intersections, this might not be an issue, but if the junction and accident collections are very big, it might become an issue. APMs may not always be able to be developed for categories with insufficient accident data. APMs are constructed based on the kind of accident that happened at the intersection, such as accidents between straight-through and left-turning cars, or right-angled incidents, in order to anticipate accidents for particular divergent motions (Hasan et al 2022).

2.1.1 Models for predicting accidents on major roadways

Very few APMs are utilized for predicting accidents on parts of metropolitan main roads. Studies that are now available show that a large variety of ML techniques are employed to evaluate the safety of different road features, including the existence of a median, two-way left turns, speed restrictions, the quantity of points of intersection, and pedestrian density. It is difficult some times to identify which geometric property might account for the accidents that occur along urban road segments. These models have numerous explanatory variables as proposed by (Diggle et. al 2020).

$$E\{k\} = \boldsymbol{\alpha} LY F^{\beta_1} e^{[x^{\beta}, x_i]} \quad 2.5$$

Where E{k} is the expected number of road accidents in a specified period of time

L is the length of road segment, Y is the number of years, F is the traffic flow on road segments, xi is the series of independent features for the ranging j=2 to m and $\boldsymbol{\alpha}$, β_1, β_i are the coefficients variables to be estimated which I ranges from 2 to m. A model was developed by (Liang and Zeger 2021) for predicting property loss and injuries from traffic accidents. Eight explanatory variables total, including numerous dummy variables (such as 1 or O), made up the best-fitting model that was produced. They employed two distinct models: one for intersections and one for mid-block sections—instead of utilizing a single model to forecast accidents for the full road segment. They had a model form that resembled equation 2.7. They proposed isolating incidents that transpired among junctions from those that occurred at intersections. They suggested two APMs: a single for mid-block segments and another for intersections. The form of the accident prediction model for mid-block road segments was as follows:

$$E\{R\} = \boldsymbol{\alpha} e^{\left(x_i^{(\beta_i, x_i)}\right)} \quad 2.6$$

Where

E{R}is the expected road accident rate measured in 10^8 vkms

X_i is a series of independent variable and i variable range from 1 to 5

X_1 denotes residential development when its equal to 1 and zero(0) as otherwise

X_2 is the flush median when its equal to zero(0)

X_3 =1 is the solid median and x3=0 otherwise

X_4 =1 is for 50 kilometer per hour, and zero otherwise

X_5 is the number of intersections per kilometer

$\boldsymbol{\alpha}$, β_i is the coefficients to be estimated with I ranging from 1 to 5.

The intersection length and being exposed are both included in the formula provided in equation 2.9. The creation of APMs for highway route segments with minor intersections when traffic counts for intermediate approaches are unavailable(Mountain et al.,2020). The system they created was designed for isolated highways, and it continues to remain in use today because conventional methods are unable to predict accidents on certain road sections where traffic counts are not always conducted at all junctions along a highway corridor. The following forms appeared in their proposed road accident prediction model.

$E\{K\} = \alpha L^{\beta_1} F^{\beta_2} e^{(\beta_3 N)}$ and $E\{K\} = \alpha L^{\beta_1} F^{\beta_2}$ 2.7

Where E{k} is the total number of expected road accidents on intersection per unit of time

L = Length of road intersection, F= the traffic flow on the road segment

α, β_1, β_2, β_3 are the coefficients to be estimated.

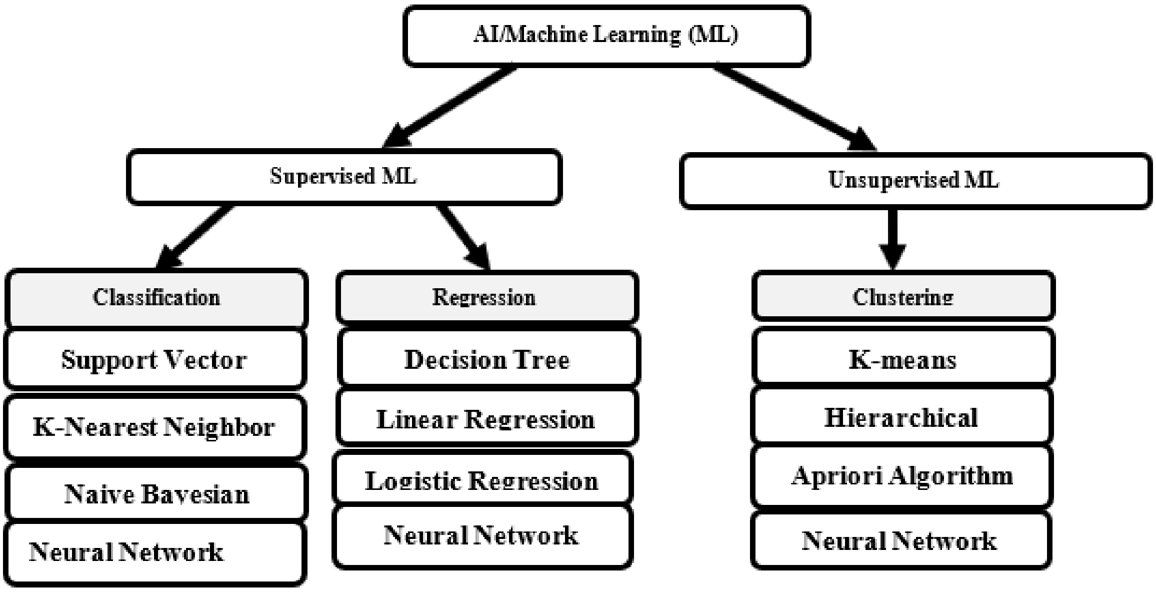


Fig 2.1: Machine learning Algorithms (Source: Theiler et al. 2020)

3.0 Methodology

This research work will be achieved following the Object-Oriented Analysis and Design Methodology (OOADM). This is aimed at viewing, modeling and implementing the new system as a collection of interacting classes and objects. OOADM is adopted because it is more effective, efficient, reliable, reusable and a faster way of developing systems.

Data Preprocessing: Preprocessing is contained in a collection of functions that can filter attribute values or database entities. The data preparation phase was utilized in order to find and eliminate false and missing values from the suggested system dataset. The numerical fields are preprocessed into an appropriate format prior to training. The training dataset was padded up and the replace missing value function was used during the preprocessing phase before model creation, training and accident prediction.

We defined a SVM and created a class to locate the SVM candidate with the closest pair from the support vectors classification. If there are any violating point, add them to the candidate points; otherwise, exclude points. Check for points that are less than zero as a result of addition violations and then select that candidate as the SV and repeat this step to prune all data points.

Training and Testing Data Split: A procedure known as "train and test split" function is used to validate model and simulate how the model function with testing data. About 80% of the entire dataset was put aside for training and 20% for testing with a random state parameter set to zero. We ensure that the structure in which our data is organized is suitable for the train and test split. This involves splitting our whole data set into "Features" and "Target".

Fitting SVM Algorithm: The extent to which a ML model generalizes with testing data is comparable to the training set as determined by its model fitting performance. More accurate results are obtained from a well-fitting model.. The fit approach is utilized to train the existing SVM model using dataset that contain both severe and non-severe road accident cases. Fitting the model to the data involves first taking in a dataset that is usually represented as a 2D array or matrix and a set of labels.

3.1 High level Model of the Proposed System.

High-level model of the proposed system provides an easily understandable overview of the key concepts and principles of an organization. A high level model is utilized as an abstract representation of organizational business data to help technical and functional team members communicate with one another. Every team member has a unique background and experience level. The purpose is to facilitate communication within the team in the organization. High level models facilitate data exchange by being utilized in the design of models in the databases and information systems. It facilitates effective communication between team members with varying backgrounds and degrees of experience. Terms are the regulations that have a single possible interpretation. This will help us identify stake holder, available resources, The high-level model illustrates the link between the proposed model's features and demonstrates dependence with respect to various degrees of abstraction. We enhanced our suggested model with abstraction linkages using use-case diagram, flowchart, database design tools, processing subsystems and class diagrams as contained in this research.

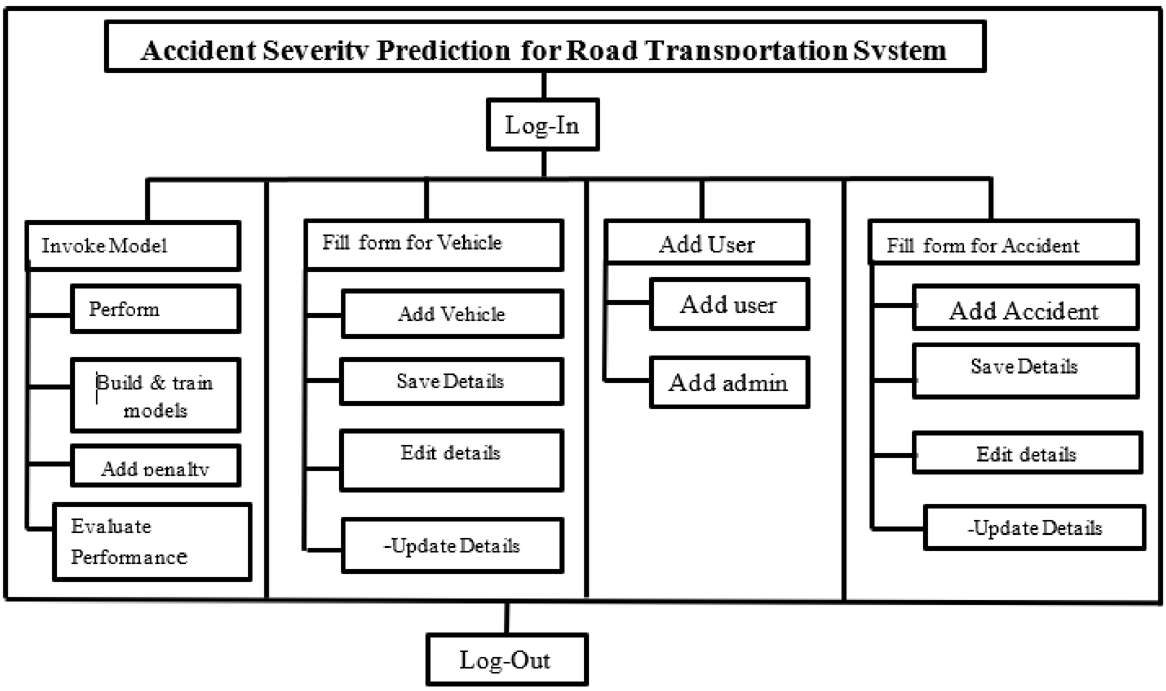


Figure 3.1: High Level Model of the Proposed System

4.0 Results and Discussion

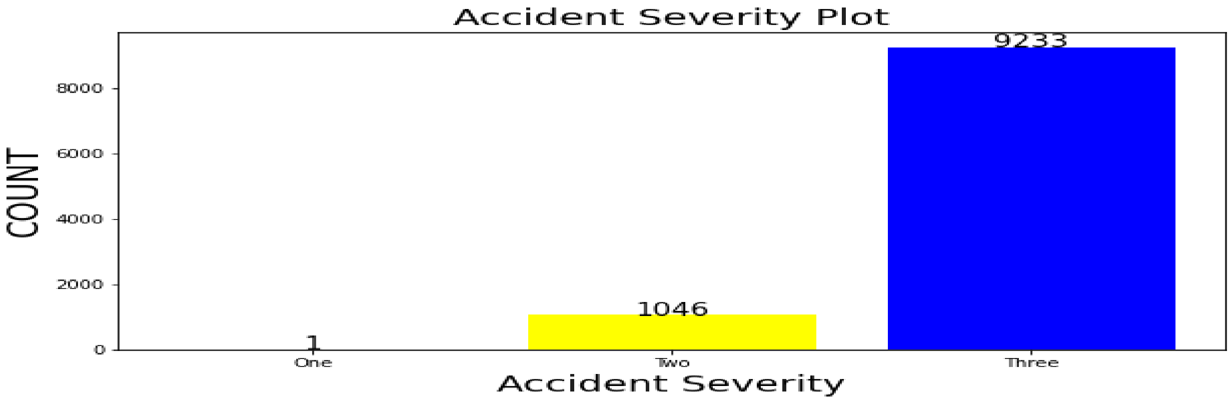


Figure 4.1: Accident severity plot

The road transport system dataset, which produced a total of 10280 items for training and testing the model, is shown in Figure 4.1 with the count of accident severity level. According to the dataset, the number of cases for accident severity level 1 was 1, level 2 produced 1046 cases, and level 3 produced 9233 cases. The blue bar chart representing accident severity level-3 showed the greatest count, while the yellow bar chart representing level-2 and level-1 showed the lowest counts, respectively.

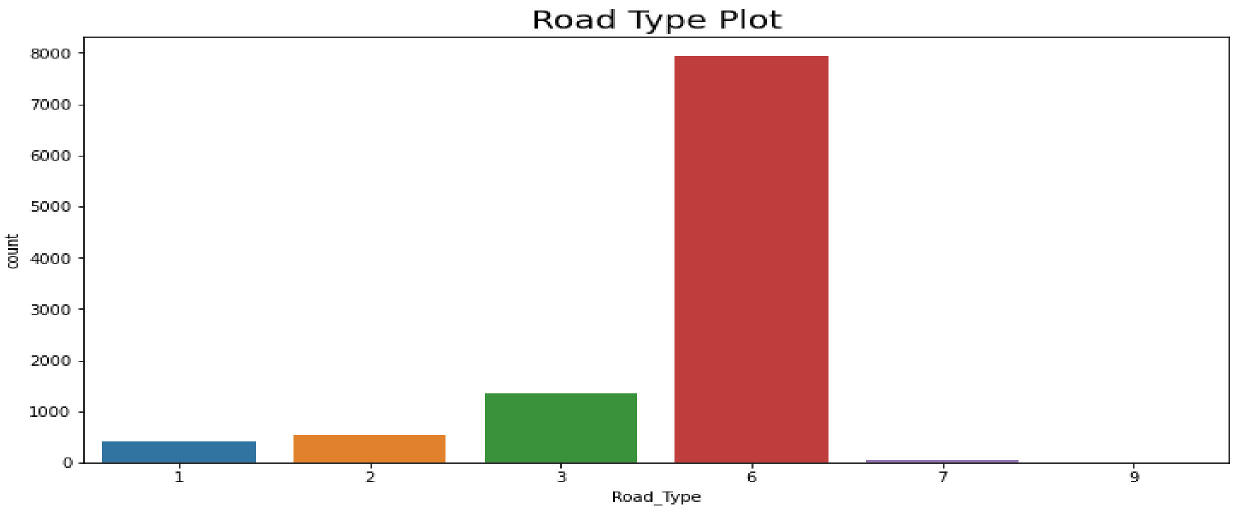


Figure 4.2: Road types

Figure 4.2 shows the various road types are coded using 1, 2, 3, and 9 as shown in Figure 4.2. The numbers 1, 2, 3, 4, and 9 stand for several kinds of roads. It appears that single carriageways are where most accidents happen.

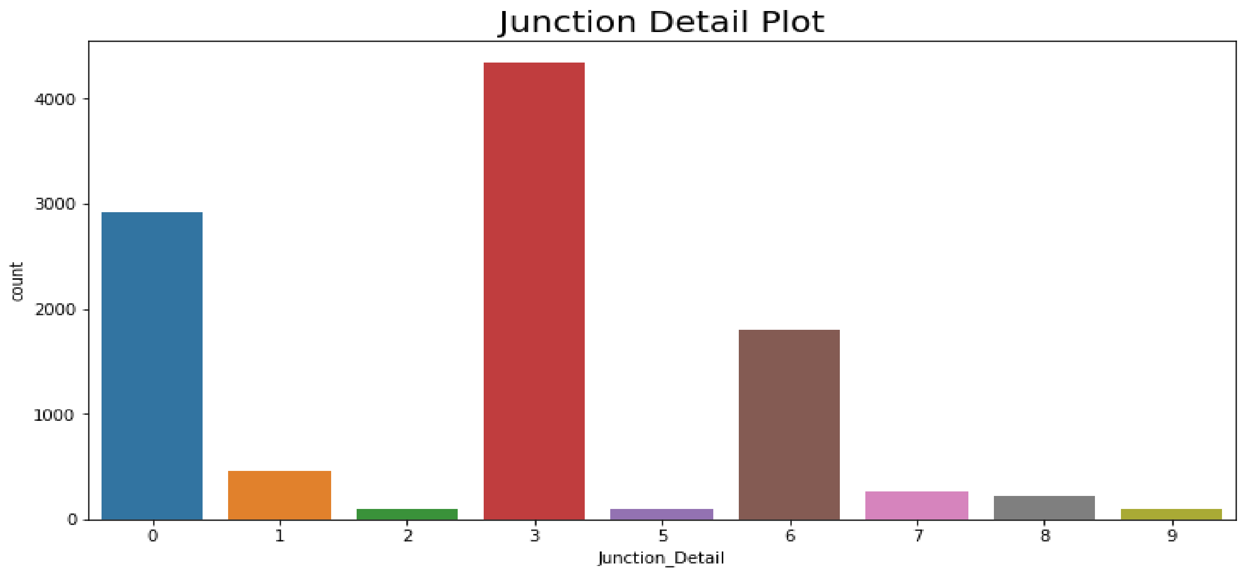


Figure 4.3: Plot of Road Junction details

Figure 4.3 depicts road transport junction details. The numbers 1, 2, 3, 4, and 9 stand for no junction, single, T-junction, and other combinations. Figure 4.3 shows that the majority of accidents occur in areas without intersections, at crossroads spaced 20 meters or less apart, or at staggered junctions.

.

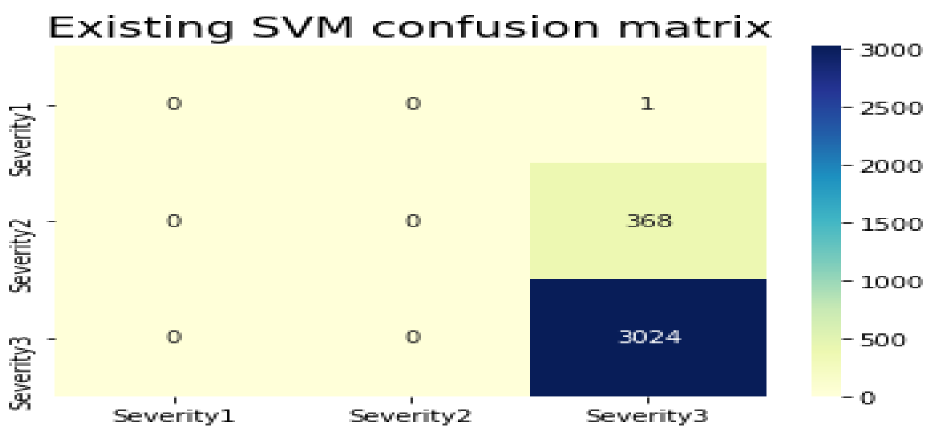


Figure 4.4: Confusion matrix of SVM

Figure 4.4 shows the confusion matrix, which displays a table structure of the various SVM predicted outcomes of a binary-classification task to aid in visualizing its results. Cell values above and below the main diagonal or off-diagonal elements showing the incorrectly predicted values, show the total number of correctly predicted values that are equal to the actual or true values. The greater the diagonal value, the more accurate the predicted EV Battery charging duration. According to the confusion matrix, accident severity level 1 had 573 incorrectly predicted cases with zero (no) correct predictions. Level 2 produced 368 incorrectly misclassified cases and level 2 provided 3024 with zero incorrectly predicted values having zero(0) true positive class prediction. 3024 accident severity levels were accurately predicted by the SVM, while 369 accident cases were incorrectly classified.

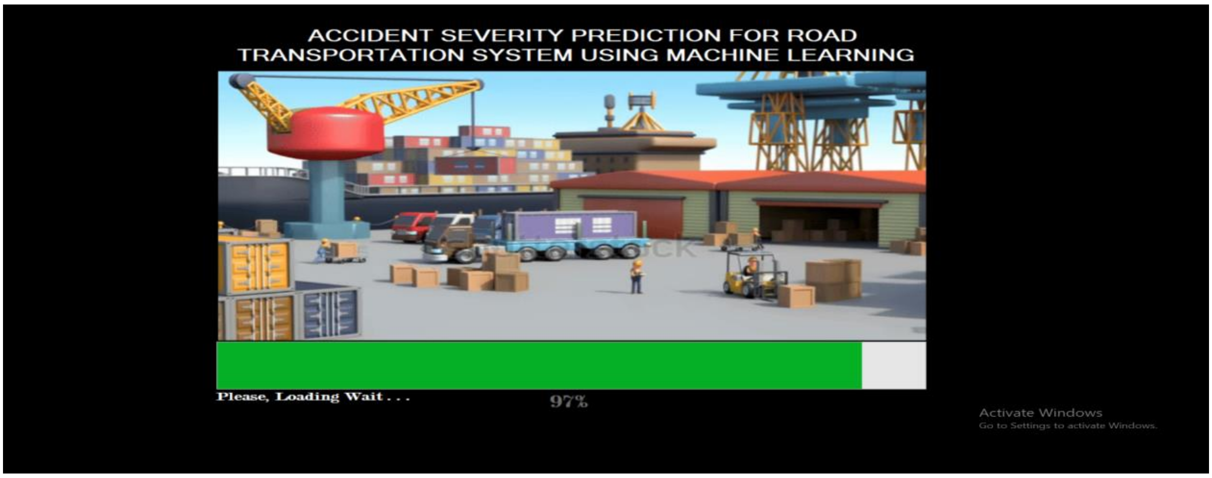


Figure 4.5: Splash screen of Road Accident API

in the progressive bar. Figure 4.5 depicts the program splash screen for accident severity prediction system software interface with a progressive bar that shows the percentage that loads from 1 to 100%. The steps vary from 1% to 100%, and the user will be prompted to wait until the user log-in screen appears after it reaches 100%. This is the splash screen that appears at the very first time a program is launched to inform the user that the program is loaded. The loading progresses as shown

Figure 4.6: Log-in GUI

Figure 4.6 depicts the user log-in menu for users (road users and transport companies) that prompts a user to supply user name and password; if they match, access is granted; otherwise, access to the program main menu is denied. If the user enters the wrong user name or password, the application display labels with the phrase forgot your user password and forgot user name respectively; but if password matches grants user an authorized access.

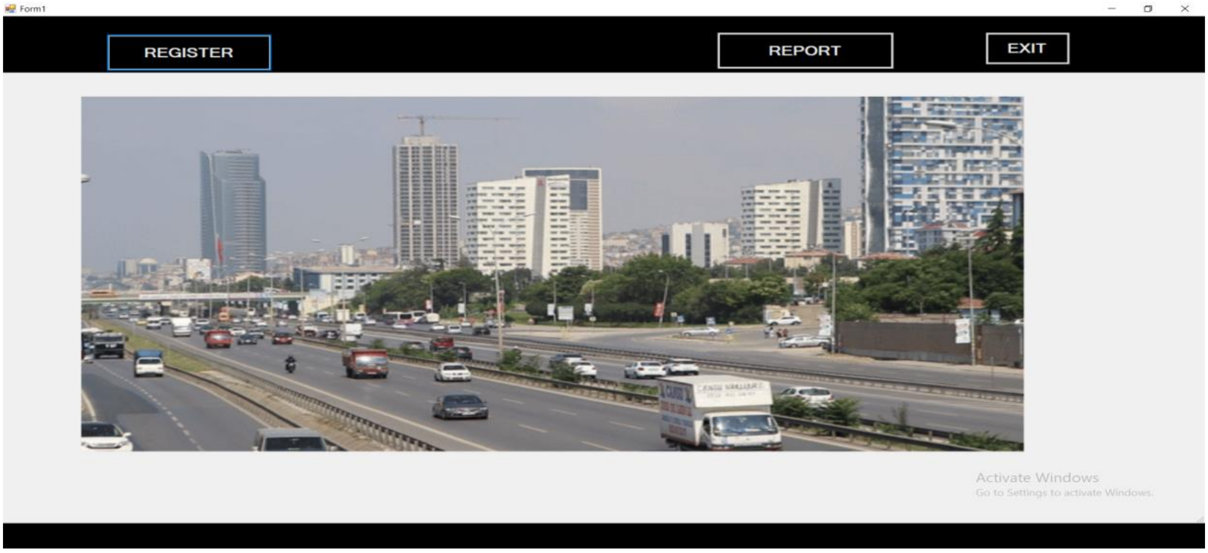


Figure 4.7: Program Main menu

Figure 4.7 shows the program main menu, which has three main icons: the register menu, report and exit icon. User registration, Accident registration, and prediction icons are found in the pull-down menu of registration panel control. The registration pull-down menu simply includes user registration and accident registration while the report pull-down menu includes User registration, prediction, and statistics of accident severity levels.

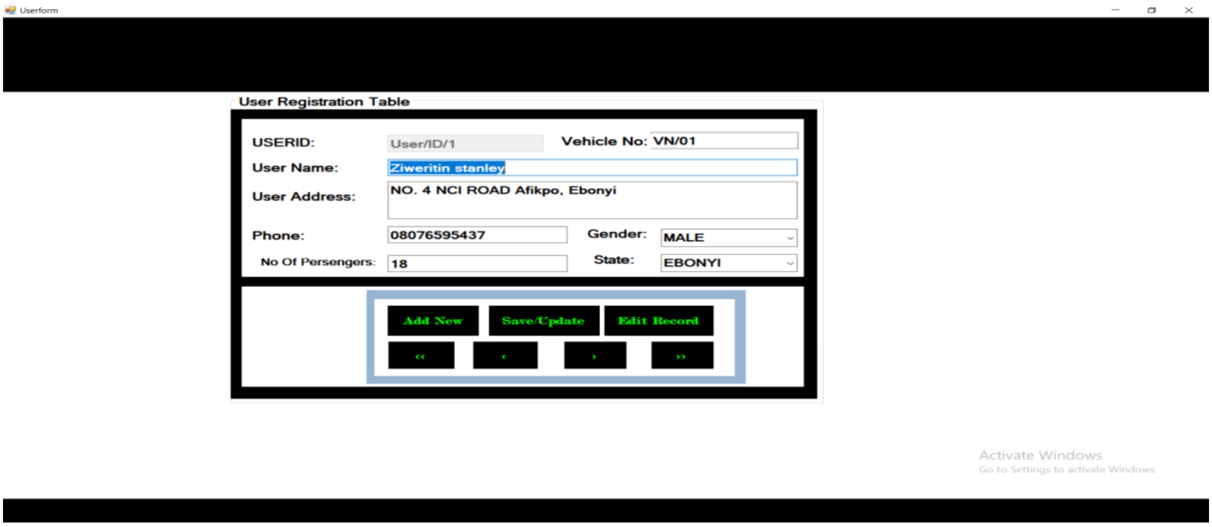


Figure 4.8: User registration menu

The user registration form acquired from the registration module is shown in Figure 4.8 The registration form includes fields like User ID, Vehicle number, user name, user address, number of passengers, gender and status. The user has been granted some administrative rights, including the ability to save records to a database, modify, update, add new items, search for record as well as move back, forward, between records and to last record.

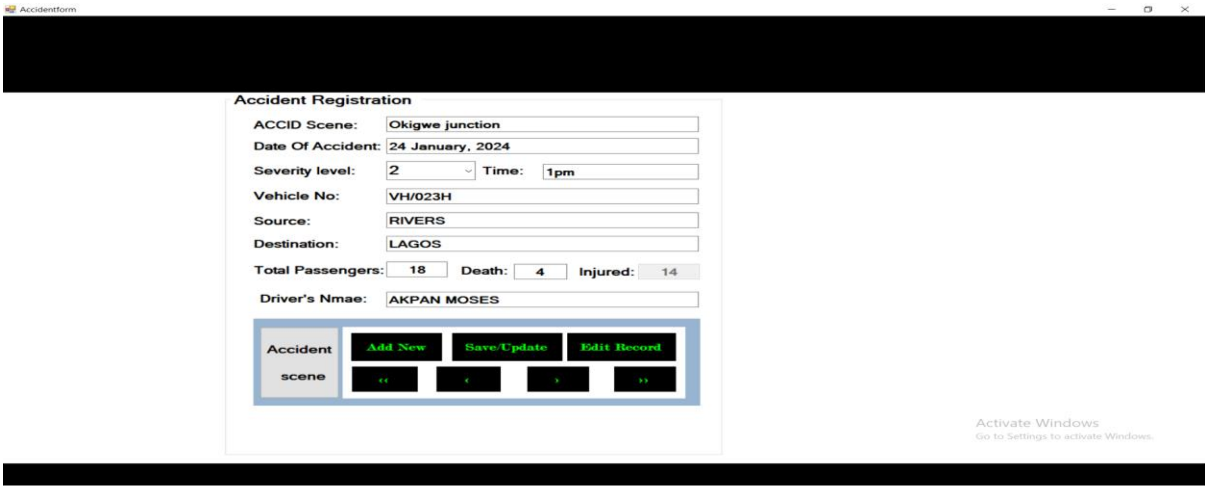


Figure 4.9: Accident registration menu

Figure 4.9 depicts the accident registration form as it requires the user to supply details relating to road traffic accident severity rates. The accident registration form includes fields like accident ID, accident scene, date of accident, severity level, vehicle number, source, destination, total passengers and driver's name, death cases injured cases. The user has been granted some administrative rights, having the ability to save records to a database, modify, update, add new records, search for record as well as move back, forward, between records and to last record.

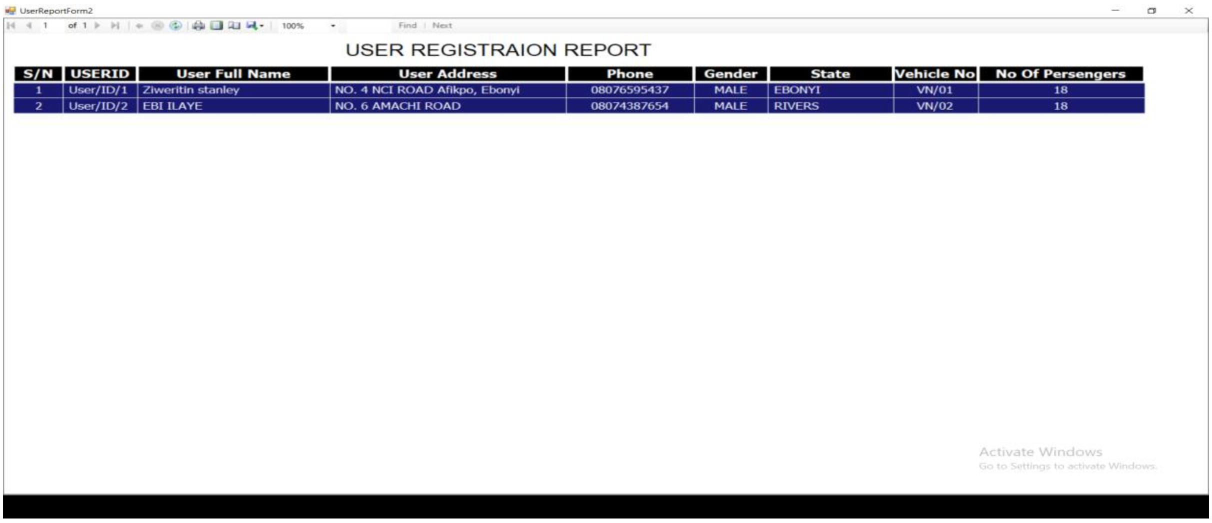
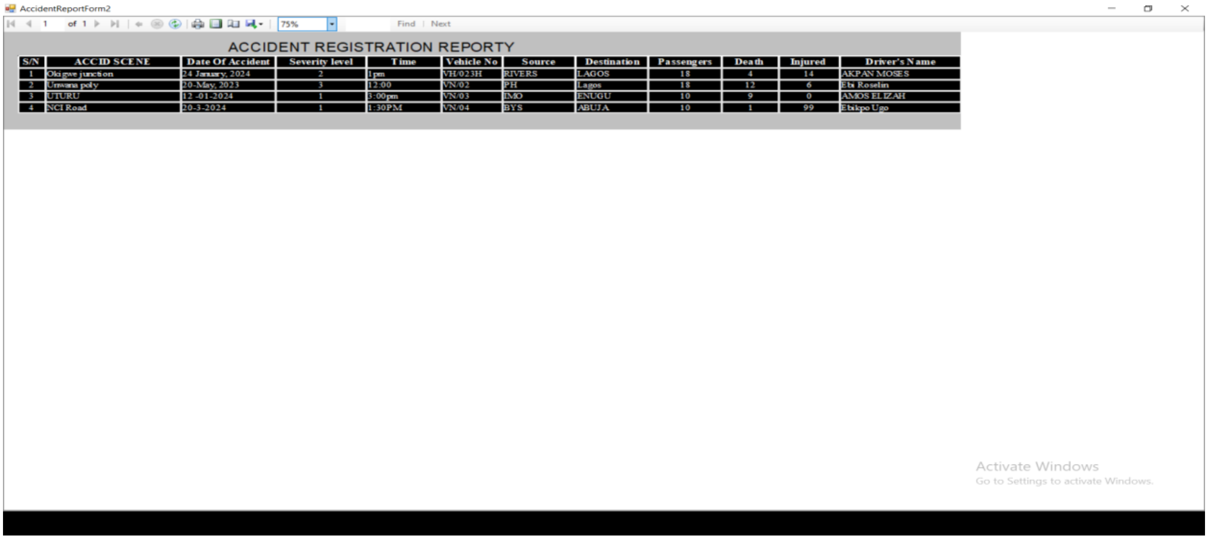


Figure 4.10: User registration report

Figure 4.10 above shows user registration report for road traffic accident prediction system with the attributes: Username, address, phone number, gender, State of origin, vehicle number and number of passengers as displayed using VB.NET report viewer control tool. The user has been granted some administrative rights like converting of reports to PDF format, excel or other database supported formats, with an interactive report viewer tools to zoom in/out.



ACCIDENT REGISTRATION REPORTY

S/N	ACCID SCENE	Date Of Accident	Severity level	Time	Vehicle No	Source	Destination	Passengers	Death	Injured	Driver's Name
1	Okigwe junction	24 January, 2024	2	1pm	VH/023H	RIVERS	LAGOS	18	4	14	AKPAN MOSES
2	Umuna poly	20-May, 2023	3	12:00	VN/02	PH	Lagos	18	12	6	Ebi Roselin
3	UTURU	12 -01-2024	1	3:00pm	VN/03	IMO	ENUGU	10	9	0	AMOS ELIZAH
4	NCI Road	20-3-2024	1	1:30PM	VN/04	BYS	ABUJA	10	1	99	Ebikpo Ugo

Figure 4.11: Accident registration report

Figure 4.11 above shows the data-based table report for accident registration with the attributes: Accident id, date of accident, accident severity level, time, vehicle number, death cases, passengers, injured cases and etc. The user has been allowed certain rights as an administrator, such as the ability to convert reports to PDF, Excel, or other database-supported formats and to zoom in and out using an interactive report viewer tool.

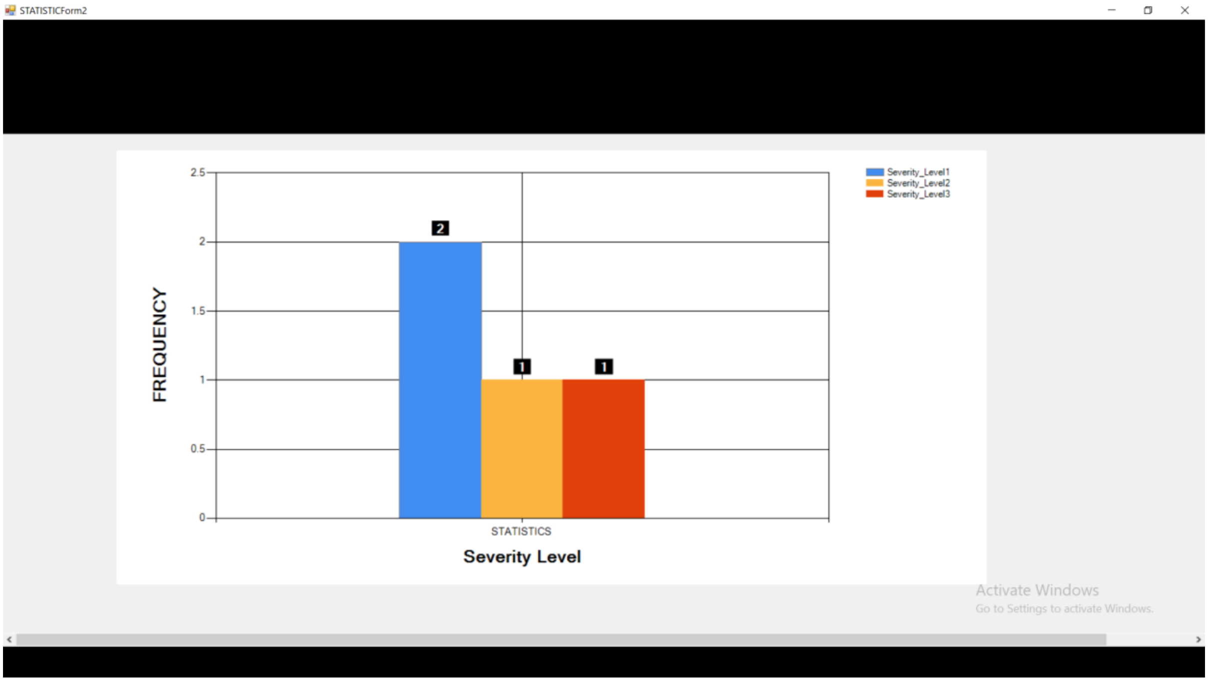


Figure 4.12: Statistics of accident severity rate

The accident severity level is depicted in Figure 4.12 by the bars which depicts mild, average, and severe cases respectively. As observed, severity levels 1 and 1 are values that are identical, with level 2 producing the highest value.

5.0 Conclusion

Adoption of the proposed support vector machine (SVM) demonstrated a notable breakthrough. Experts can use it to raise the prediction accuracy of deep learning systems since the penalty term improved detection accuracy for all adjusted hyper-parameter values and solved the model complexity problem. The SVM exhibited excellent generalization with testing set, ignore level-1 occurrences, and correctly classified all level 2 accident instances. The primary risk factor causing driving-related accidents was quickly and effectively determined by the suggested prediction model. The level-1, level-2, and level 3 accident predictions for the suggested system were significantly higher which sets off some hyper-parameters for high performance.

The penalty term gave the user control over the layer's output and assisted the SVM outputs to be smaller, converge faster, and possess higher accuracy. This was done with the Python Jupyter Notebook independent development environment (IDE) using the Tensorflow, Keras, and Graphviz deep learning tools. Furthermore, a VB.NET graphical user interface platform was created that enables users to input accident details and receive an automatic prediction of the accident severity level, which is categorized into 1, 2, and 3 targeted classes by the model.

REFERENCES

Aassi, K., Rahman, S.M., Mansoor, U. and Ratrout, N.(2020) Predicting Crash Injury Severity with Machine Learning Algorithm Synergized with Clustering Technique: A Promising Protocol. *Int. J. Environ. Res. Public Health*, *17*, 5497.

Bonnet, E., Lechat, L., Ridde, V.(2018) What interventions are required to reduce road traffic injuries in Africa? A scoping review of the literature. *PLoS ONE*, *13*, e0208195

Chen, T.Y. and Jou, R.C.(2019) Using HLM to investigate the relationship between traffic accident risk of private vehicles and public transportation. *Transp. Res. Part A Policy Pract.*, *119*, 148–161.

Diggle, P.J., Liang, K. Y. and S.L. Zeger, S. L.(2020) Analysis of Longitudinal Data. Clarendon Press, Oxford, U.K, 1-20.

Fu, X., Meng, H., Wang, X., Yang, H. and Wang, J.(2022) A hybrid neural network for driving behavior risk prediction based on distracted driving behavior data. *PLoS ONE*, *17*, e0263030.

Green, P.(2019) literative Reweighted Least Squares for Maximum Likelihood Estimation and Some Robust and Resistant Alternative (with discussion). Journal of the Royal Statistical Society. Series 6(46), 1 49-1 62.

Hasan, R.A., Irshaid, H., Alhomaidat, F., Lee, S. and Oh, J.S.(2022) Transportation Mode Detection by Using Smartphones and Smartwatches with Machine Learning. KSCE J. Civ. Eng., 26, 3578–3589

Hauer, E., J., Ng, C. N. and Lovell, J.(2019) Estimation of Safety at Signalized Intersections. In Transportation Research Record 1185, TRB, National Research Council, Washington, D.C., 48-61.

Jeong, H., Kim, I., Han, K., Kim, J.(2022) Comprehensive Analysis of Traffic Accidents in Seoul: Major Factors and Types Affecting Injury Severity. *Appl. Sci.*, *12*, 1790.

Liang, K.Y., and Zeger, S. L.(2021) Longitudinal Data Analysis Using Generalized Linear Models. Biometrika, Vol. 73, 13-22.

Miaou, S. P.(2020) MeasuMg the Goodness-of-Fit of Accident Prediction Models. Federal Highway Administration, Publication No. FHWA-RD-96-040, Virginia, U.S., 1-20

Mountain, L., Maher, M. J. and Fawaz, B.(2020) The Influence of Trend on estimates of Accidents at Junctions, Accident Analysis & Prevention, Vol. 30, No. 5, 641 49.

Mukherjee, D. and Mitra, S.(2022) Pedestrian safety analysis of urban intersections in Kolkata, India using a combined proactive and reactive approach. *J. Transp. Saf. Secur*, *14*, 754–795. Müller, A.C., and Guido, S., (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists 1st ed., United States of America, O'Reilly Media

Nelder, LA., and Wedderburn, R. W. N.(2019) Generalized Linear Models. Journal of the Royal Statistical Society, Series A, 53, 370-384.

Nicholson, A. and Turner, S.(2020) Estimating Accidents in a Road Network. ln Proceedings of Roads 96 Conference, Part 5, New Zealand, 5346.

Pervez, A., Huang, H., Lee, J., Han, C., Li, Y.; Zhai, X.(2021) Factors affecting injury severity of crashes in freeway tunnel groups: A random parameter approach. *J. Transp. Eng. Part A Syst.*, *48*, 04022006.

Rabbani, M.B.A., Musarat, M.A., Alaloul, W.S., Ayub, S., Bukhari, H., Altaf, M.(2022) Road Accident Data Collection Systems in Developing and Developed Countries: A Review. *Int. J. Integr. Eng.*, *14*, 336–352.

Tarlochan, F., Ibrahim, M.I.M., Gaben, B.(2022) Understanding Traffic Accidents among Young Drivers in Qatar. *Int. J. Environ. Res. Public Health*, *19*, 514. Weng, J.X.; Gan, X.F. and Zhang, Z.Y.(2021) A quantitative risk assessment model for evaluating hazmat transportation accident risk. Saf. Sci., 137, 11

Yu, R.J.; Zheng, Y.; Qin, Y. and Peng, Y.C.(2019) Utilizing Partial Least-Squares Path Modeling to Analyze Crash Risk Contributing Factors for Shanghai Urban Expressway System. ASCE-ASME J. Risk. Uncertain. Eng. Syst. Part A. Civ. Eng., 5, 9

Yuan, Y.L.; Yang, M.; Guo, Y.Y.; Rasouli, S.; Gan, Z.X. and Ren, Y.F.(2021) Risk factors associated with truck-involved fatal crash severity: Analyzing their impact for different groups of truck drivers. J. Saf. Res., 76, 154–165

Yan, M., and Shen, Y.(2022) Traffic Accident Severity Prediction Based on Random Forest. *Sustainability*, *14*, 1729.

Yang, Z., Zhang, W., Feng, J.(2022) Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework. *Saf. Sci.*, *146*, 105522