DETERMINING THE ROAD ACCIDENTS BY USING A SUPPORT VECTOR MACHINE ALGORITHM IN COMPARISON WITH THE LOGISTIC REGRESSION ALGORITHM FOR OBTAINING ACCURACY

I. Krishna Teja[1], S. Ashokkumar[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The aim of the study is to determine road accidents using support vector machine algorithms and logistic regression algorithms.
Materials and Methods: Novel support vector machines compared to logistic regression algorithms are used to determine road accidents of time series. Sample size is determined using the G Power calculator and found to be 10 per group.Totally 20 samples are used. Pretest power is 80% with a CI of 95%.
Results and Discussion: Based on analysis support, vector machine algorithms have significantly better accuracy (92%) compared to logistic regression algorithms (87.13%). The statistical significance difference value p=0.01 ($p<0.05$, Independent sample T-test) states that traffic flow and the results in study are significant.
Conclusion: Within the limits of this study, the Support vector machine offers better accuracy than logistic regression algorithm to determine road accidents.

Keywords: Novel Support Vector Machine, Logistic Regression Algorithm, Road accident Prediction, Machine Learning, Time Series, Traffic Flow.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu. India. Pincode: 602105.
[2*]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu. India. Pincode: 602105.

1. Introduction

The purpose of this study is to determine road accidents using support vector machines and logistic regression algorithm.Every year road accidents count increases drastically. The main reasons for road accidents are vehicle conditions and road conditions (Wang et al. 2021). The World Health Organization declares that every year due to road accidents 1.5 million casualties are recorded all over the world (Wang et al. 2021; Jha, Chatterjee, and Tiwari 2021). The main factors that affect diver negligence and road conditions. Some of the drivers' rash driving like overtaking vehicles, drinking and driving leads to a high death rate. Due to road accidents humans life and property of government damaged in seconds traffic flow of (Davoodabadi et al. 2021; Mehdizadeh et al. 2021).The death rates are heavy in capital cities of countries like Mumbai, New York,Tokyo etc,,,.Road transport authority (RTA) has to take Novel Support Vector Machine certain measures to determine accidents (Vankov and Schroeter 2021; Taamneh and Taamneh 2021). The outcome of this study could help road transport authorities to determine roads accidentally accurately. Real world applications of Support Vector Machine are used in Bioinformatics, Generalised Predictive Control, Face detection etc.

There are 156 research articles published on road accident prediction using support vector machine and logistic regression algorithms in science direct and 462 research articles on Google Scholar and 39 research articles were published in IEEE xplore for road accident prediction. The data for road accidents is vast. According to the World Health Organization, twenty thousand people die in 24 hrs of time. So, outliers and null values are removed in data preprocessing. Sometimes huge data helps to get a lot of insight about problems and assists to solve challenges in different ways of Novel Support Vector Machines (Rustagi et al. 2021). At the sametime huge data leads to a time taken process to read data (Jha, Chatterjee, and Tiwari 2021; Taamneh and Taamneh 2021). There are several softwares and gadgets to determine the road accidents, but machine learning algorithms give more accuracy than others (Jha, Chatterjee, and Tiwari 2021).Our team has extensive knowledge and research experience that has translated into high quality publications(Pandiyan et al. 2022; Yaashikaa, Devi, and Kumar 2022; Venu et al. 2022; Kumar et al. 2022; Nagaraju et al. 2022; Karpagam et al. 2022; Baraneedharan et al. 2022; Whangchai et al. 2022; Nagarajan et al. 2022; Deena et al. 2022)

Based on the literature survey, Functional features like weather and lighting conditions, which tend to increase the severity of accidents in major cases, weren't included in the data to predict accident severity. It is critical to know features like weather and lighting conditions.of time series. The aim of these conditions, along with others, are the focus of research. traffic flow and other factors that help in predicting accident severity using the Support Vector Machine algorithm.

2. Materials and Methods

The proposed work's topic is carried out in DBMS Laboratory, department of computer science and engineering, Saveetha School of Engineering. The number of groups identified for the study are two. The group-1 is a support vector machine algorithm algorithm and group-2 is a logistic regression algorithm. Sample size for each Machine Learning.group was calculated by using previous study results by keeping the g power as 80 %, threshold 0.05 and confidence interval as 95% (Taamneh and Taamneh 2021). According to that, the sample size of the support vector machine algorithm (N=10) and logistic regression algorithms (N=10) were calculated.

The dataset contains 117537 rows and 37 columns. The dataset contains which traffic flows different data about road accidents like locations, latitude, longitude, speed, police force, time etc..The data samples splitted for training is 107036 and for testing is 26775. The dataset was collected from the UK government open data.

The software tool used to evaluate support vector machine algorithms and logistic regression algorithms was a Google collaboration with python programming language . The hardware configuration was an intel core i5 processor with a RAM size of 12GB. The system type used was a 64-bit, OS, X64 based processor with an HDD of 917 GB. The software configuration includes Windows 10 operating system.

In the proposed model, train the dataset and implement a classification algorithm based on the dataset. After collecting the dataset, Novel Support Vector Machine outliers and null values were removed .By data preprocessing was done. After data preprocessing the dataset is split into two parts, one for training Machine Learning. and another for testing. In the dataset 50% is split for training and the remaining 50% given to the testing process. By evaluating algorithms with train and test sets accuracy percentage was predicted.

Support Vector Machine

Support vector machine is a supervised learning algorithm which can be used for both classification and regression processes. In a support vector machine there are data points in traffic flow

between data points there is a line which is called a hyperplane. This hyper plane divides the data points into two segments,one is margin and other one is a large margin.(Wang et al. 2021)

Logistic Regression Algorithm

Logistic regression is a supervised learning model. It takes the only discrete values for a given set of features, and uses sigmoid function Machine Learning to get output. Logistic regression is a regression model. It predicts output of a categorical dependent variable. It is similar to linear regression. The ROC curve approach, which is a rationale (Boztepe and Usul 2019)**.**

Statistical Analysis

The analysis was done using IBM SPSS version 21 . It is a statistical software tool used for data analysis. For both proposed and existing algorithms 10 iterations were done with a maximum of 10-20 samples and for each iteration predicted accuracy was noted for analysing accuracy.Independent T-Test is carried out for analysis.Independent.variables are dataset and dependent variable is accuracy (Boo and Choi 2021).

3. Results

In this study ,Machine Learning algorithms are used for determining road accidents. We test for the performance of these algorithms how precisely a technique can predict the accidents.Two algorithms are selected and tested for which algorithm produces the highest rate of accuracy.

The graph explains the comparison of accuracy between the algorithms Support vector machine and Logistic regression.Where the accuracy of the SVM is 90.5% and LR is 83.5%.The data gives clarifications of different factors that affect road accidents.The accuracy comparison of support vector machine algorithm and logistic regression algorithm Fig. 1 results of a support vector machine to determine the road accidents Fig. 2 results of logistic regression algorithms to determine road accidents Fig. 3 results clearly show that support vector machine algorithms got better Machine Learning . significance than logistic regression.The support vector machine model achieved precision 87%, recall 88%, accuracy 92% and 88% F1-score. Finally, the proposed classifier achieved an accuracy of 92%. Pseudo code for support vector machine Table 1. Pseudocode for logistic regression algorithm Table 2. Table 3 shows Sample Data of the Road Accidents.Table 4 Consists of SPSS inputs which are taken to calculate accuracy. Table 5 Consists of SPSS inputs which are taken to calculate accuracy. Table 5 Predicted accuracy of support vector machine and random forest algorithm (Support vector machine accuracy is 92% and random forest accuracy of 87.%). Table 6 shows Statistical analysis of mean, standard deviation and standard error of Accuracy for Support vector machine and random forest algorithm. There is a statistically significant difference between groups. Support vector machines have a higher mean (92) than random forest (87). Table 7 shows Independent sample T-test result is done with confidence interval as 95% and significance level as p=0.01 (support vector machine appears to perform significantly better than logistic regression with a value of $p<0.05$).Bar graph compares mean accuracy of support vector machine algorithm based road accident prediction and determination and logistic regression based road accident determination and prediction shown in Fig. 4.

4. Discussion

Support vector machine algorithm-based road accident prediction and determination have better accuracy compared to logistic regression based road accident prediction and determination gives lots of insights and information about road accidents to understand much better. Head injury caused by road accidents (Wang et al. 2021). (Taamneh and Taamneh 2021) et al used machine learning algorithms like support vector machine and logistic regression algorithm to predict real world road user impact accidents (Taamneh and Taamneh 2021); (Lin 1994; Robin 2014; Riyapan et al. 2021).

The factors that affect road accident prediction are computational cost, dataset size, number of data in dataset and null values in data. The identification ability of the model is completely dependent on data and its attributes; a small size of datasets with less null values and outliers performs better convergence. Aim of this research is to develop simple networks to reduce computational cost (Tsukahara, Haseyama, and Kitajima 1999; Davoodabadi et al. 2021), these networks produce good results against large datasets. Some simple pre-trained neural networks have found difficulty in learning one class successfully with high accuracy (Taamneh and Taamneh 2021) have 3GB of data and 265030 images. This research used SVM and LRA classifiers to achieve accuracy of 93% (Rowlands and DeCrescenzo 1990; Jha, Chatterjee, and Tiwari 2021). The digital data used in this dataset are collected from various UK locations. The obtained original data may not have the same parameters

and might be different data. Therefore, the time series collected data should be similar, and maintaining consistency is important in terms of making efficient analysis and consistency (E. and K. 2015; Kaur et al. 2012). There is no opposite finding related to this proposed algorithm.

Due to limitations such as threshold, precision and recall, fraud data used in this dataset is collected from various sources. The evaluation of accuracy cannot provide a better outcome on larger data sets. Moreover, in logistic regression algorithms, mean error appears to be higher than support vector machine algorithms. It would be better if mean error can be reduced to a considerable extent. However, work can be enhanced by applying optimization algorithm techniques to automate road accident data, time series to achieve better accuracy and less mean error. Feature selection algorithms can be used before classification to improve classification accuracy of classifiers. The feature selection algorithm can be used to reduce computation time and improve classification accuracy of classifiers.

5. Conclusion

Based on the obtained results, the support vector algorithm got a better significance value as compared to the logistic regression algorithm. The time series and accuracy of road accident prediction is done by an algorithm support vector machine with better accuracy 92% as compared to logistic regression with accuracy 87.13%.

Declaration

Conflict of Interest

No conflicts of interest in this manuscript.

Authors Contribution

Author IKT was involved in data collection, data analysis, and manuscript writing. Author SAK was involved in conceptualization, data validation, and critical review of a manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing necessary infrastructure to carry out this work successfully.

Funding:

We thank the following organisations for providing financial support that enabled us to complete the study.

1.Soft Square Pvt.Ltd.

2.Saveetha University.

3.Saveetha Institute of Medical and Technical Sciences.

4.Saveetha School of Engineering.

6. References

Baraneedharan, P., Sethumathavan Vadivel, C. A. Anil, S. Beer Mohamed, and Saravanan Rajendran. 2022. "Advances in Preparation, Mechanism and Applications of Various Carbon Materials in Environmental Applications: A Review." Chemosphere. https://doi.org/10.1016/j.chemosphere.2022.134596.

Boo, Yookyung, and Youngjin Choi. 2021. "Comparison of Prediction Models for Mortality Related to Injuries from Road Traffic Accidents after Correcting for Undersampling." International Journal of Environmental Research and Public Health 18 (11). https://doi.org/10.3390/ijerph18115604.

Boztepe, Engin, and Hayrettin Usul. 2019. "Using the Analysis of Logistic Regression Model in Auditing and Detection of Frauds." Khazar Journal of Humanities and Social Sciences. https://doi.org/10.5782/2223-2621.2019.22.3.5.

Davoodabadi, Abdoulhosein, Esmail Abdorrahim Kashi, Mahdi Mohammadzadeh, Noushin Mousavi, Shima Shafagh, Leila Ghafoor, Mojtaba Sehat, Shahrzad Ale Mohammad, and Abbas Hajian. 2021. "Predicting Factors and Incidence of Preventable Trauma Induced Mortality." Annals of Medicine and Surgery (2012) 68 (August): 102609.

Deena, Santhana Raj, A. S. Vickram, S. Manikandan, R. Subbaiya, N. Karmegam, Balasubramani Ravindran, Soon Woong Chang, and Mukesh Kumar Awasthi. 2022. "Enhanced Biogas Production from Food Waste and Activated Sludge Using Advanced Techniques – A Review." Bioresource Technology. https://doi.org/10.1016/j.biortech.2022.127234.

E., Bhagyashree, and Tanuja K. 2015. "Phishing URL Detection: A Machine Learning and Web Mining-Based Approach." International Journal of Computer Applications. https://doi.org/10.5120/ijca2015905665.

Jha, Alok Nikhil, Niladri Chatterjee, and Geetam Tiwari. 2021. "A Performance Analysis of Prediction Techniques for Impacting Vehicles in Hit-and-Run Road Accidents." Accident; Analysis and Prevention 157 (July): 106164.

Karpagam, M., R. Beaulah Jeyavathana, Sathiya Kumar Chinnappan, K. V. Kanimozhi, and M. Sambath. 2022. "A Novel Face Recognition Model for Fighting against Human Trafficking in Surveillance Videos

and Rescuing Victims." Soft Computing. https://doi.org/10.1007/s00500-022-06931-1.

Kaur, Jaspreet, Student Masters of Technology, Shri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, and Rajneet Kaur. 2012. "Speckle Noise Reduction in Biomedical Images Using Haar Wavelets with Wiener Filter." International Journal of Scientific Research. https://doi.org/10.15373/22778179/may2013/43.

Kumar, P. Ganesh, P. Ganesh Kumar, Rajendran Prabakaran, D. Sakthivadivel, P. Somasundaram, V. S. Vigneswaran, and Sung Chul Kim. 2022. "Ultrasonication Time Optimization for Multi-Walled Carbon Nanotube Based Therminol-55 Nanofluid: An Experimental Investigation." Journal of Thermal Analysis and Calorimetry. https://doi.org/10.1007/s10973-022-11298-4.

Lin, Tzuoo-Ding. 1994. Exploration of Motor Carrier Accident Occurrence Using Time Dependent Logistic Regression.

Mehdizadeh, Amir, Mohammad Ali Alamdar Yazdi, Miao Cai, Qiong Hu, Alexander Vinel, Steven E. Rigdon, Karen Davis, and Fadel M. Megahed. 2021. "Predicting Unsafe Driving Risk among Commercial Truck Drivers Using Machine Learning: Lessons Learned from the Surveillance of 20 Million Driving Miles." Accident; Analysis and Prevention 159 (September): 106285.

Nagarajan, Karthik, Arul Rajagopalan, S. Angalaeswari, L. Natrayan, and Wubishet Degife Mammo. 2022. "Combined Economic Emission Dispatch of Microgrid with the Incorporation of Renewable Energy Sources Using Improved Mayfly Optimization Algorithm." Computational Intelligence and Neuroscience 2022 (April): 6461690.

Nagaraju, V., B. R. Tapas Bapu, P. Bhuvaneswari, R. Anita, P. G. Kuppusamy, and S. Usha. 2022. "Role of Silicon Carbide Nanoparticle on Electromagnetic Interference Shielding Behavior of Carbon Fibre Epoxy Nanocomposites in 3-18GHz Frequency Bands." Silicon. https://doi.org/10.1007/s12633-022-01825-1.

Pandiyan, P., R. Sitharthan, S. Saravanan, Natarajan Prabaharan, M. Ramji Tiwari, T. Chinnadurai, T. Yuvaraj, and K. R. Devabalaji. 2022. "A Comprehensive Review of the Prospects for Rural Electrification Using Stand-Alone and Hybrid Energy Technologies." Sustainable Energy Technologies and Assessments. https://doi.org/10.1016/j.seta.2022.102155.

Riyapan, Sattha, Jirayu Chantanakomes, Bongkot Somboonkul, Sang Do Shin, Wen-Chu Chiang, and PATOS Clinical Research Network. 2021. "Effect of Nighttime on Prehospital Care and Outcomes of Road Traffic Injuries in Asia: A Cross-Sectional Study of Data from the Pan-Asian Trauma Outcomes Study (PATOS)." Prehospital Emergency Care: Official Journal of the National Association of EMS Physicians and the National Association of State EMS Directors, August, 1–11.

Robin, Paul. 2014. Use on Multinomial Logistic Regression in Work Zone Crash Analysis for Missouri Work Zones.

Rowlands, J. A., and G. DeCrescenzo. 1990. "Wiener Noise Power Spectra of Radiological Television Systems Using a Digital Oscilloscope." Medical Physics. https://doi.org/10.1118/1.596528.

Rustagi, Neeti, Abhishek Jaiswal, Dervla Kelly, Naveen Dutt, Arvind Sinha, and Pankaja Raghav. 2021. "Measuring Medical Graduate Behavioral Intention for Administering on-Site Care to Road Traffic Accident Victims: Development and Validation of a Questionnaire." Indian Journal of Public Health 65 (1): 39–44.

Taamneh, Salah, and Madhar M. Taamneh. 2021. "A Machine Learning Approach for Building an Adaptive, Real-Time Decision Support System for Emergency Response to Road Traffic Injuries." International Journal of Injury Control and Safety Promotion 28 (2): 222–32.

Tsukahara, Masato, Miki Haseyama, and Hideo Kitajima. 1999. "Removal of Additive White Noise Using an Adaptive Wiener Filter with Edge Retention." Systems and Computers in Japan. https://doi.org/10.1002/(sici)1520-684x(199901)30:1<81::aid-scj8>3.0.co;2-2.

Vankov, Daniel, and Ronald Schroeter. 2021. "Driving under the Influence of Drugs or Alcohol: Predicting the Intentions of Young Drivers." Traffic Injury Prevention 22 (2): 97–101.

Venu, Harish, Ibham Veza, Lokesh Selvam, Prabhu Appavu, V. Dhana Raju, Lingesan Subramani, and Jayashri N. Nair. 2022. "Analysis of Particle Size Diameter (PSD), Mass Fraction Burnt (MFB) and Particulate Number (PN) Emissions in a Diesel Engine Powered by Diesel/biodiesel/n-Amyl Alcohol Blends." Energy. https://doi.org/10.1016/j.energy.2022.123806.

Wang, Fang, Zhen Wang, Lin Hu, Hongzhen Xu, Chao Yu, and Fan Li. 2021. "Evaluation of Head Injury Criteria for Injury Prediction Effectiveness: Computational Reconstruction

of Real-World Vulnerable Road User Impact Accidents." Frontiers in Bioengineering and Biotechnology 9 (June): 677982.

Whangchai, Niwooti, Daovieng Yaibouathong, Pattranan Junluthin, Deepanraj Balakrishnan, Yuwalee Unpaprom, Rameshprabu Ramaraj, and Tipsukhon Pimpimol. 2022. "Effect of Biogas Sludge Meal Supplement in Feed on Growth Performance Molting Period and Production Cost of Giant Freshwater Prawn Culture." Chemosphere 301 (August): 134638.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Advances in the Application of Immobilized Enzyme for the Remediation of Hazardous Pollutant: A Review." Chemosphere 299 (July): 134390.

TABLES AND FIGURES

Table 1. Pseudo Code for the Support Vector Machine

1.Find the optimal values for a tuning parameters of the SVM model;
2.Train the SVM model;
3.pp*;
4.while p> 2 do
5.SVMP + SVM with a optimised tuning parameters for p variables and
6.observations in Data;
7.Wp + calculate weight vector of a SVM, (Wpl, ... , Wpp);
8.rank.criteria + (wpl, ..., wmp);
9.min.rank.criteria + variable with lowest value in rank.criteria vector;
10.Remove min.rank.criteria from Data;
11.Rankp + min.rank.criteria;
12.ptp-1;
13.end
14.Rankı + variable in Data & (Rank2, ..., Rankpop);
15.return (Rankı,..., Rankp)

Table 2. Pseudo Code for Random Forest Algorithm

1.#Feature Scaling
2.from sklearn.preprocessing import StandardScaler
3.sc_X = StandardScaler()
4.X_train =sc_X.fit_transform(X_train)
5.X_test = sc_X.transform(X_test)
6.#Fitting Logistic Regression to dataset
7.from sklearn.linear_model import LogisticRegression
8.classifier = LogisticRegression()
9.classifier.fit(X_train, y_train)
10.#Predicting the test set result
11.y_pred = classifier.predict(X_test)
12.#Making the confusion matrix
13.from sklearn.metrics import confusion_matrix
14.cm = confusion_matrix(y_test, y_pred)

Table 3. Sample Data of the Road Accidents

Accident_ Index	**Location_Easting_OSGR**	**Location_Northing_OSGR**	**Longitude**	**Latitude**	**Police_ Force**	**Accident_Severity**	**Number_of_ Vehicles**
2.02E+12	528218	180407	-0.15384	51.50806	1	3	2
2.02E+12	530219	172463	-0.12795	51.43621	1	3	2
2.02E+12	530222	182543	-0.12419	51.5268	1	3	2
2.02E+12	525531	184605	-0.19104	51.54639	1	2	1
2.02E+12	524920	184004	-0.20006	51.54112	1	3	2
2.02E+12	540188	185266	0.020461	51.54888	1	3	2

2.02E+12	532424	164886	-0.09907	51.36761	1	3	1
2.02E+12	532773	178460	-0.08898	51.48951	1	3	3
2.02E+12	548535	188113	0.141957	51.57233	1	3	2

Table 4. Consists of SPSS inputs which are taken to calculate accuracy.

group_id	Svm	Lr
1	95	80
1	94.5	83
1	95.3	77
1	92	78
1	97	82
2	96	80
2	95	77.5
2	94.7	82
2	95.5	81
2	91	79

Table 5. Consists of SPSS inputs which are taken to calculate accuracy.Table-5: Predicted accuracy of support vector machine and random forest algorithm (Support vector machine accuracy is 92% and random forest accuracy of 87.%)

Algorithm	Accuracy	F1 score	Recall	Precision
Support vector machine	92%	88%	88%	87%
Logistic regression	87%	0.93%	100%	87%

Table 6. Statistical analysis of mean, standard deviation and standard error of Accuracy for Support vector machine and random forest algorithm. There is a statistically significant difference between groups. Support vector machines have a higher mean (92) than random forest (87).

Group		**N**	**Mean**	**Std.Deviation**	**St Std.Error mean**
AAccuracy	**Support vector mMachine**	10	92.5278	1.16725	.36912
	Logistic Regression	10	87.9700	1.67294	.21280
Precision	**Su Support vector M Machine**	10	87.0644	1 1.63325	.20025
	Logistic Regression	10	8 87.5755	1.78327	.24769

Table 7. The significance value p=0.01 (p<0.05) shows that two groups are statistically significant.

		Levene's Test for Equality of Variance		t-test for Equality of Means						
		F	Sig.	t	df	Sig.(2-tailed}	Mean difference	Std.Error difference	95% confidence interval of the differences	
									Lower	Upper
Accuracy	Equal variance assumed	3.153	0.01	11.085	18	<0.01	4.723	.42607	3.8278	5.6181

	Equal variance not assumed			11.085	14.388	<0.06	4.723	.42607	3.8114	5.6345

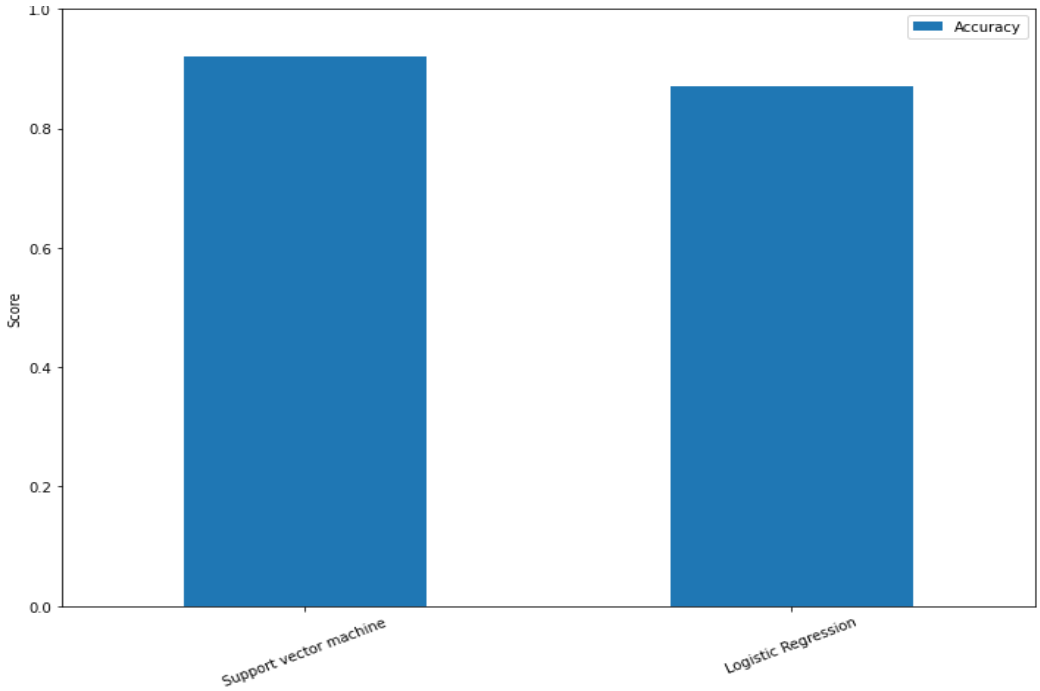

Fig. 1. Data visualisation of comparison of accuracy analysis of support vector machine and logistic regression algorithm

Accuracy for support vector machine is : 92.5%

S.NO	**Precision**	**recall**	**F1-score**	**support**
1	0.00	0.00	0.00	0
2	0.14	0.13	0.13	125
3	0.87	0.88	0.88	862
Accuracy			0.92	987
Macro avg	0.34	0.34	0.34	987
Weighted avg	0.78	0.79	0.78	987

Fig. 2. Results of support vector machine shows that accuracy of 92% and Accuracy for Logistic Regression is:87.132%

S.NO	**Precision**	**recall**	**F1-score**	**support**
1	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	127

S.NO	Precision	recall	F1-score	support
3	0.87	1.00	0.93	860
Accuracy			0.87	987
Macro avg	0.44	0.50	0.47	987
Weighted avg	0.76	0.87	0.81	987

Fig. 3. Results of logistic regression algorithm shows that accuracy of 87.13%

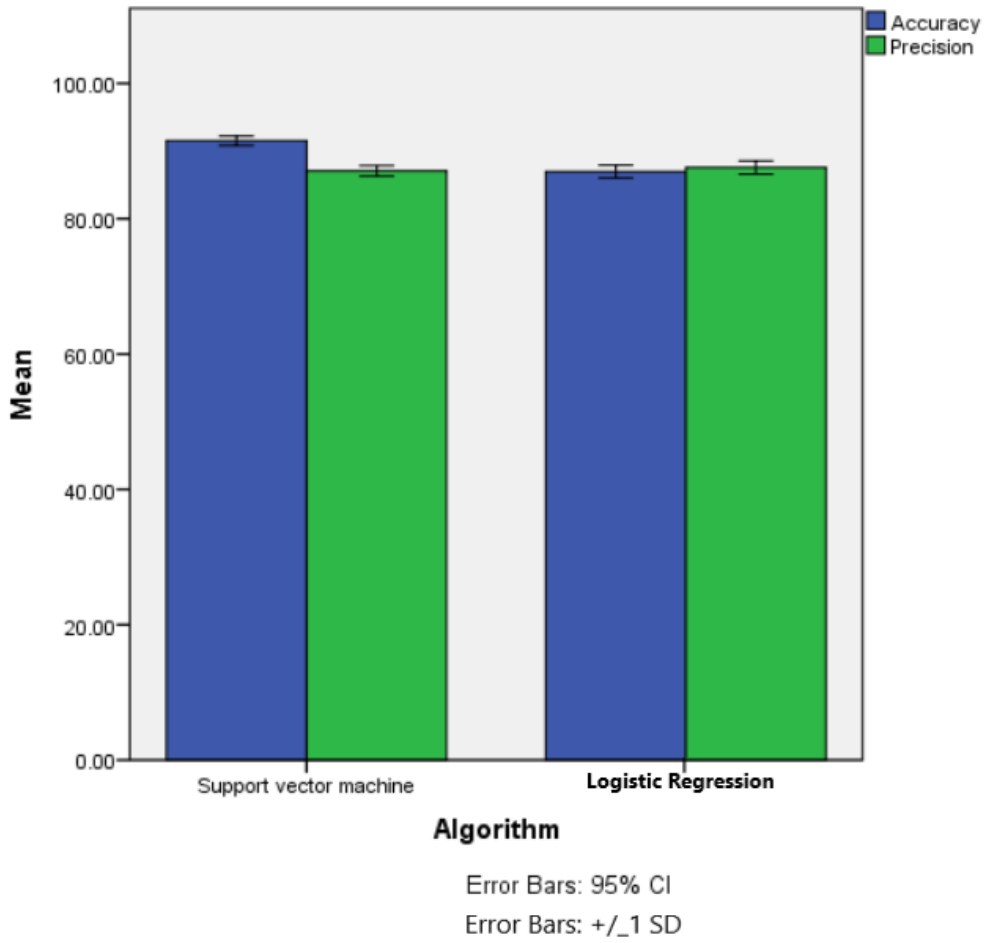

Fig. 4. Comparison of support vector machine algorithm and random forest algorithm in terms of mean accuracy and precision. Mean accuracy and precision of the support vector machine (92%) is better than logistic regression (87.13%). Standard deviation of a support vector machine is slightly better than logistic regression. X Axis: SVM vs RFA. Y Axis: Mean accuracy of detection ± 1 SD.