www.revistageintec.net
ISSN: 2237-0722
revista GEINTEC
Gestão, Inovação e Tecnologias

An Efficient Analysis of Road Accidents Severity Using Novel Support Vector Machines Over Artificial Neural Networks with Improved Accuracy Rate

U. Sravan Kumar[1]; Dr.R. Beaulah Jeyavathana[2*]

[1]Research Scholar, Department of Computer Science Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

[1]sravanroyal1999@gmail.com

[2*]Project Guide, Department of Computer Science Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India.

[2*]Beaulahjayavathanar.sse@saveetha.com

Abstract

Aim: *To perform an efficient analysis of road accidents severity prediction using Support Vector Machine and Artificial Neural Network algorithms.* ***Materials and Methods:*** *Road accident severity prediction is performed by Support Vector Machine algorithm (N=10) and Artificial Neural Network algorithm (N=10) using Traffic-crashes dataset.* ***Results and Discussion:*** *The road accident severity prediction using Support Vector Machine algorithm (Accuracy 83.6%) and with Artificial Neural Network algorithm (Accuracy 73.3%). There is no significant difference between the groups.* ***Conclusion:*** *Within the limits of this study, SVM algorithm has significantly better accuracy than ANN algorithm.*

Key-words: Machine Learning, Computer Vision, Novel SVM, Road Accident Severity Prediction, Artificial Neural Networks.

1. Introduction

Road accident severity prediction is the process of predicting the severity measure of accidents. In recent years, the road accident has become a major problem and marked as ninth prominent cause of deaths in the world (Labib et al. 2019). It is estimated from a study ("Road Traffic Injuries" n.d.) that 1.3 million people died and a large number of people got injured in 2019 around the world. An efficient and accurate accident severity prediction can help to provide vital information for emergency service providers to evaluate the severity level of accidents, estimate the potential impacts and proceed with the efficient accident management.

Road traffic accident severity prediction has been carried out by researchers and 20 related research articles in IEEE Digital Xplore and 15 articles are published in Research gate. (Bharti Sharma, Katiyar, and Kranti Kumar 2016) analyzed the urban traffic accidents using the Support Vector Machine algorithm. (Alkheder, Taamneh, and Taamneh 2017) used Artificial Neural Networks to predict the severity of around 6000 accidents that occurred in AbuDhabi during 2008 to 2013. (Jacobé de Naurois et al. 2019) used Artificial Neural Networks to detect driver drowsiness level and predict when the state of the driver will be impaired. (Zhang et al. 2018) used Decision trees, Random forest, KNN and ANN algorithms to predict accident severity. The ANN and Random forest produced better results with accuracy rates of 52.9% and 53.9% respectively. (García de Soto and Bumbacher 2018) used SVM models to predict accident severity and they have given clear explanations about accident severity prediction with more accuracy.

Previously our team has a rich experience in working on various research projects across multiple disciplines (Sathish and Karthick 2020; Varghese, Ramesh, and Veeraiyan 2019; S. R. Samuel, Acharya, and Rao 2020; Venu, Raju, and Subramani 2019; M. S. Samuel et al. 2019; Venu, Subramani, and Raju 2019; Mehta et al. 2019; Sharma et al. 2019; Malli Sureshbabu et al. 2019; Krishnaswamy et al. 2020; Muthukrishnan et al. 2020; Gheena and Ezhilarasan 2019; Vignesh et al. 2019; Ke et al. 2019; Vijayakumar Jain et al. 2019; Jose, Ajitha, and Subbaiyan 2020). Now the growing trend in this area motivated us to pursue this project.

In the existing system, the data considered to predict the accident severity does not contain functional features like weather and lighting conditions which tend to increase the severity of accidents in major cases. It is important to include the features like weather and lighting conditions. So, this research focuses on considering these conditions along with other factors that help in predicting accident severity using Support Vector Machine algorithm with better accuracy.

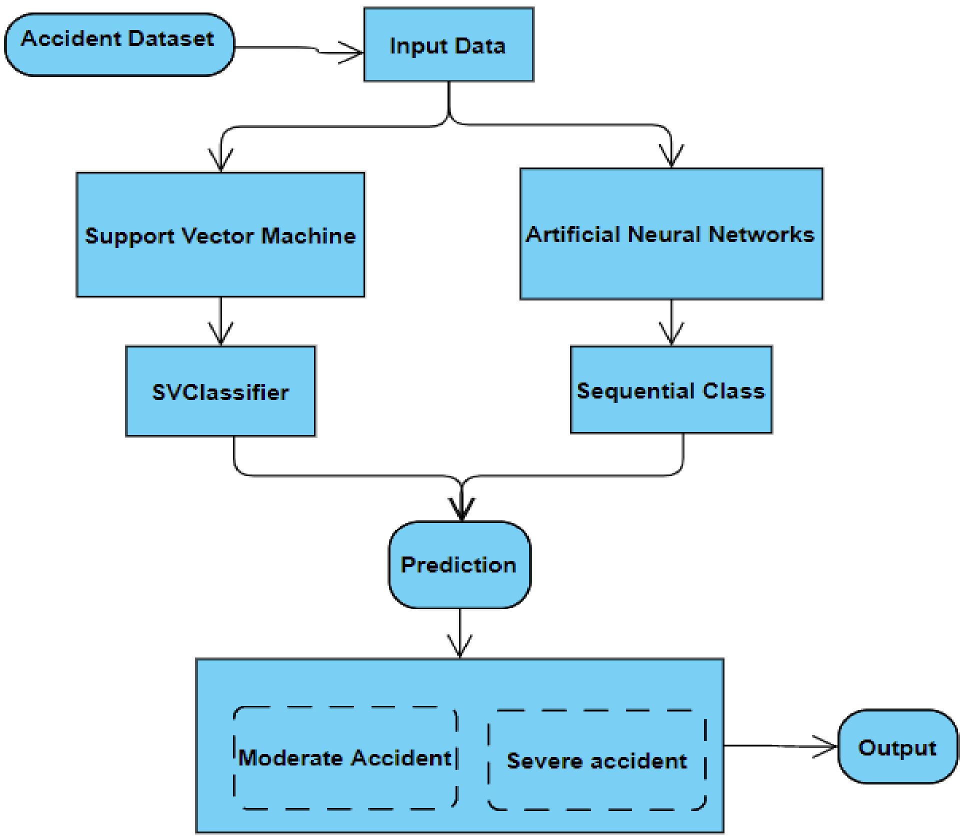
2. Materials and Methods

The study setting of the proposed work is done in Saveetha School of Engineering, SIMATS. The number of groups identified for the study is two. The group 1 is Artificial Neural Networks and group 2 is Support Vector Machines. Using G power 10 sample sizes and totally 20 sample sizes have been carried out for our study, 95% confidence and pretest power 80% (Chong, Abraham, and Paprzycki 2005)**.**

The dataset named ‘Traffic-Crashes’ is downloaded from a public domain**.** The data in the dataset is about the traffic accidents that occurred in the LasVegas region in the year 2016. The data was preprocessed and all the null values and unwanted columns have been removed from the dataset. The dataset includes information about the condition of the driver, weather(rainy, cloudy, clear) and lighting condition(dark, day, dull). The dataset was splitted into two parts namely the training part and the testing part. 90% of the data was used for training and the remaining 10% was used for testing. The algorithms were implemented by evaluating the train and test sets.

Fig 1 represents the framework of predicting the severity of an accident step by step. The input is given from the records that are present in the dataset by reading the dataset into the program. Using Support Vector Machines and Artificial Neural Networks, the input is processed and the data is divided into training and testing parts. In this case, 90% of the data is used for training and the remaining 10% is used for testing.

Fig. 1 - Accident Severity Prediction Framework



SVM classifier is used to train the data and the Sequential class of ANN is used to add hidden layers that are connected to the input layer and train the data. The function predict() which is present in Sequential and SV Classifier classes is used to predict the output of the testing data. The predicted output is the accident severity and it is displayed as either Moderate Accident or Severe Accident.

Artificial Neural Network Algorithm - Group-1

Pseudocode for ANN algorithm is given in Table-1. Artificial Neural Network algorithm is a deep learning based algorithm. As the name suggests, this algorithm is inspired by human brains and it learns as the human brain learns. Neural network consists of an input layer, hidden layers and an output layer to predict and store the output.

Table 1 - Pseudocode for ANN Algorithm

// I : Input dataset records
1. Import the required packages.
2. Convert the string values in the dataset to numerical values.
3. Assign the data to X_train, y_train, X_test, y_test variables.
4. Using train_test_split() function, pass the training and testing variables and give test_size and random_state as the parameters.
5. Import the Sequential() class.
6. Using Sequential() class, add input layer, hidden layers, output layer.
7. Predict the output using.predict() function.
8. Calculate accuracy and other measures of the model from the confusion matrix.
Output **// Accident Severity**

The activation function known as Transfer Function used here in the input and the output layers is 'relu' which is a Rectified Linear Unit. The optimizer is set as 'adam'. It is the mostly used and best stochastic gradient descent function. The hidden layer is interconnected to the input layer and connected to the output layer where the output is stored.

SVM Algorithm - Group 2

Pseudocode for SVM Algorithm is given in Table-2. Support Vector Machine Algorithm is a supervised machine learning algorithm. It is a classification algorithm that performs data classification by forming a hyperplane between the data points existing in the plane. The hyperplane will try to be as far away from both the data points. The kernel function that is used in the SVM Classifier here is 'linear' as our data can be linearly related.

Table 2 - Pseudocode for SVM Algorithm

// I : Input dataset records
1. Import the required packages.
2. Convert the string values in the dataset to numerical values.
3. Assign the data to X_train, y_train, X_test, y_test variables.
4. Using train_test_split() function, pass the training and testing variables and give test_size and the random_state as parameters.
5. Import the SVClassifier from sklearn library.
6. Using SVClassifier, predict the output of the testing data.
7. Calculate the accuracy and other metrics of the model from the confusion matrix.
Output **// Accident severity**

The software tool used to evaluate ANN and SVM algorithms is Kaggle in Python programming language. The hardware configuration includes an intel i3 processor with a RAM size of 4GB. The system used was 64-bit Windows 10 Operating System.

Statistical Analysis

For statistical implementation, the software tool used here is IBM SPSS V26.0. Statistical Package for Social Sciences is used for calculating the statistical calculations such as mean, standard deviation, significance and also to plot the graphs etc.,. The independent variables are weather_condition, driver_condition and lighting and the dependent variable is 'result' (which describes the severity of accidents in numerical form). In SPSS, the datasets are prepared using 10 as sample size for each group and Accuracy is given as the testing variable.

3. Results

The accuracy and F1_Score measured for different samples using Artificial Neural Network algorithm is given in Table 3. Table 4 shows the accuracy and F1_Score measured for different samples using Support Vector Machine algorithm. Table 5 shows various statistical measures like Accuracy, Precision, Recall and F1- Score calculated using the confusion matrix in the program implementation. The confusion matrix is a 2x2 matrix containing True Positive(tp), True Negative(tn), False Positive(fp), False Negative(fn).

Table 3 - Accuracy of Road Accident Severity Prediction Using ANN Algorithm
(Mean Accuracy = 73.30, mean F1_Score = 84.50)

Test	**Accuracy**	**F1_Score**
Test 1	71.00	83.00
Test 2	71.00	83.00
Test 3	70.00	82.00
Test 4	81.00	89.52
Test 5	69.00	81.65
Test 6	70.00	82.35
Test 7	76.00	86.36
Test 8	74.00	85.05
Test 9	75.00	85.71
Test 10	76.00	86.36

Table 4 - Accuracy of Road Accident Severity Prediction Using SVM Algorithm
(Mean Accuracy = 83.60, Mean F1_Score = 88.76)

Test	**Accuracy**	**F1_Score**
Test 1	79.00	86.27
Test 2	86.00	90.66
Test 3	83.00	88.59
Test 4	85.00	91.12
Test 5	81.00	86.71
Test 6	83.00	88.88
Test 7	81.00	88.05
Test 8	83.00	89.03
Test 9	82.00	88.88
Test 10	83.00	89.44

Table 5 - Calculation of Performance Measures Using Confusion Matrix
(Accuracy of SVM 84% is More Compared to ANN 74%)

Measure	**Support Vector Machine (SVM)**	**Artificial Neural Network (ANN)**
Accuracy	84.00	74.00
Precision	89.18	100.0
Recall	86.84	74.00
F1-Score	88.00	85.05

Table 6 shows group statistics results. For Independent Samples tests we take equality means for every group and parameters. F-score and Significance is calculated for Levene‘s Test, Equality of Variances are taken. Whereas for t-test equality means are calculated. This includes Significance, Mean Difference for each group, standard deviation for variance assumed and not assumed values. Confidence interval of the difference as lower and upper values range as shown in Table 7.

ISSN: 2237-0722
Vol. 11 No. 2 (2021)
Received: 10.03.2021 – Accepted: 12.04.2021

Table 6 - Group Statistics Results (Mean of SVM 83.60 is more Compared to ANN 73.30 and Standard Error mean for SVM is 0.63 and ANN is 1.19)

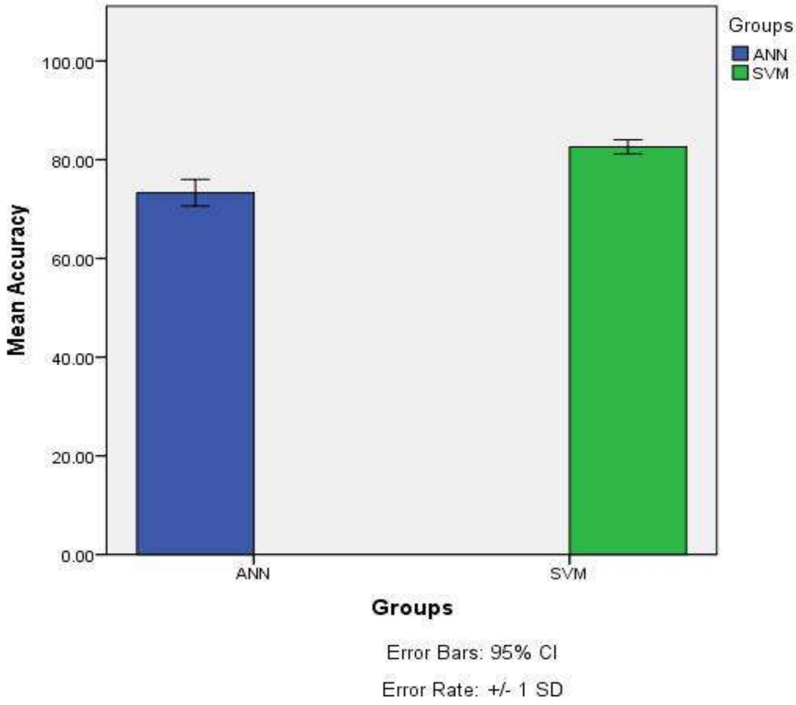
	Groups	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	ANN	10	73.3000	3.77271	1.19304
	SVM	10	83.6000	2..01108	.63596
F1_score	ANN	10	84.5000	2.52501	.79848
	SVM	10	88.7630	1.51581	.47934

Table 7 - An Independent Samples t Test of Accuracy for Predicting Accident Severity Using SVM and ANN algorithms. SVM Algorithm Appears to Perform Significantly Better than ANN Algorithm. (p=0.001)

		F	Sig	t	df	Sig(2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Accuracy	Equal Variances assumed	5.080	.037	-6.879	18	.001	-9.30	1.35195	-12.14035	-6.45
	Equal Variances not assumed			-6.879	13.733	.001	-9.30	1.35195	-12.14035	-6.35
F1_Score	Equal variances assumed	4.173	.051	-4.577	18	.001	-4.263	.93131	-6.21961	-2.30
	Equal variances not assumed			-4.577	14.741	.001	-4.2630	.93131	-6.21961	-2.27

Figure 2 is a bar graph that is plotted by selecting Mean Accuracy on Y-axis and the Groups on X-axis. From the graph, it is clear that SVM has significantly higher accuracy than ANN. The error bars are shown in the graph and the error rate is less for SVM compared to ANN.

Fig. 2 - Comparison of ANN Algorithm and SVM Algorithm in Terms of Mean Accuracy. The mean accuracy of SVM is better than ANN and the standard deviation of SVM is slightly better than ANN. X Axis: SVM vs ANN Y Axis: Mean accuracy of detection ± 1 SD



The sample data is tested statistically using the SPSS tool with group ID-1 is ANN algorithm and group ID-2 is SVM algorithm. In this study, it is observed that the SVM algorithm proved with better significant results and improved accuracy than the ANN algorithm.

4. Discussion

In this study, we observed that Novel SVM algorithm has better significant accident severity prediction accuracy and error difference than ANN algorithm ($p<0.005$, Independent Sample t test). The improved accuracy and reduced Loss for SVM (mean accuracy = 83.60%, mean Loss=5.47%) than ANN (mean Accuracy = 73.30%, mean Loss=10.39%).

Various machine learning algorithms are used to predict accident severity. ((Chong, Abraham, and Paprzycki 2005) used SVM and hybrid tree-networks to predict accident severity. Out of which, the hybrid decision-tree neural network produced better and accurate results compared to the other algorithms. (García de Soto and Bumbacher 2018) used neural networks to predict accident severity. This model is compared with Decision trees and Multilayer perceptron (MLP). (Pradhan and Ibrahim Sameen 2020) used neural networks and support vector machines to determine the factors that greatly affect the severity of driver injuries caused by traffic accidents. (Mokhtarimousavi et al. 2019)

ISSN: 2237-0722
Vol. 11 No. 2 (2021)
Received: 10.03.2021 – Accepted: 12.04.2021

investigates the prediction of work zone crash severity and the contributing factors by using support vector machine algorithm. (Z. Li et al. 2012) developed an SVM model (48.8%) for analysing crash injury severity and it is compared with ordered probit model (44.0%). SVM and Negative Binomial regression models are used to predict motor vehicle crashes (X. Li et al. 2008) and it is found that SVM models effectively and accurately predicted crash data.

Our institution is passionate about high quality evidence based research and has excelled in various fields ((Vijayashree Priyadharsini 2019; Ezhilarasan, Apoorva, and Ashok Vardhan 2019; Ramesh et al. 2018; Mathew et al. 2020; Sridharan et al. 2019; Pc, Marimuthu, and Devadoss 2018; Ramadurai et al. 2019). We hope this study adds to this rich legacy.

The limitation in our study is that this proposed model takes longer execution time for preprocessing. In the future, this limitation can be overcome by optimizing this algorithm to work faster.

5. Conclusion

Support Vector Machine algorithm has significantly greater accuracy than the Artificial Neural Networks. The accuracy which we achieved using the algorithm is a good score in predicting accident severity.

Declarations

Conflict of interests

No conflict of interest in this manuscript.

Authors Contributions

Author U. Sravan Kumar was involved in data collection, data analysis, manuscript writing. Author R. Beaulah Jeyavathana was involved in conceptualization, data validation, and critical review of manuscript.

Acknowledgements

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 the opportunities and facilities to carry out this work successfully.

Funding

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

1. Technosoft labs Pvt Ltd., Visakhapatnam.
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of engineering.

References

Alkheder, S., Taamneh, M., & Taamneh, S. (2017). Severity prediction of traffic accident using an artificial neural network. *Journal of Forecasting*, *36*(1), 100-108.

Jeyavathana, R.B., Kanimozhi, K.V., & Kumar, P. (2021). Comparative Analysis of Unstructured Data Using FWS, FIMDO, WFUPAC Algorithm. *European Journal of Molecular & Clinical Medicine*, *7*(5), 1734-1738.

Sharma, B., Katiyar, V.K., & Kumar, K. (2016). Traffic accident prediction model using support vector machines with Gaussian kernel. In *Proceedings of fifth international conference on soft computing for problem solving*, 1-10.

Chong, M., Abraham, A., & Paprzycki, M. (2005). Traffic accident analysis using machine learning paradigms. *Informatica*, *29*(1), 89-98.

http://www.informatica.si/index.php/informatica/article/download/21/15.

Ezhilarasan, D., Apoorva, V.S., & Ashok Vardhan, N. (2019). Syzygium cumini extract induced reactive oxygen species-mediated apoptosis in human oral squamous carcinoma cells. *Journal of Oral Pathology & Medicine*, *48*(2), 115-121.

García De Soto, B., & Bumbacher, A. (2018). Predicting Road Traffic Accidents Using Artificial Neural Network Models. *Infrastructure Asset*, *5*(4), 132-144.

https://www.icevirtuallibrary.com/doi/abs/10.1680/jinam.17.00028.

Gheena, S., & Ezhilarasan, D. (2019). Syringic acid triggers reactive oxygen species–mediated cytotoxicity in HepG2 cells. *Human & experimental toxicology*, *38*(6), 694-702

De Naurois, C.J., Bourdin, C., Stratulat, A., Diaz, E., & Vercher, J.L. (2019). Detection and prediction of driver drowsiness using artificial neural network models. *Accident Analysis & Prevention*, *126*, 95-104.

Jose, J., & Subbaiyan, H. (2020). Different treatment modalities followed by dental practitioners for ellis class 2 fracture–A questionnaire-based survey. *The Open Dentistry Journal*, *14*(1) 59–65.

Ke, Y., Al Aboody, M.S., Alturaiki, W., Alsagaby, S.A., Alfaiz, F.A., Veeraraghavan, V.P., & Mickymaray, S. (2019). Photosynthesized gold nanoparticles from Catharanthus roseus induces caspase-mediated apoptosis in cervical cancer cells (HeLa). *Artificial cells, nanomedicine, and biotechnology*, *47*(1), 1938-1946.

Krishnaswamy, H., Muthukrishnan, S., Thanikodi, S., Antony, G. A., & Venkatraman, V. (2020). Investigation of air conditioning temperature variation by modifying the structure of passenger car using computational fluid dynamics. *T hermal Science*, *24*(1 Part B), 495-498.

Labib, M.F., Rifat, A.S., Hossain, M.M., Das, A.K., & Nawrine, F. (2019). Road accident analysis and prediction of accident severity by using machine learning in Bangladesh. *In 7th International Conference on Smart Computing & Communications (ICSCC)*, 1-5.

Li, X., Lord, D., Zhang, Y., & Xie, Y. (2008). Predicting motor vehicle crashes using support vector machine models. *Accident Analysis & Prevention*, *40*(4), 1611-1618.

Li, Z., Liu, P., Wang, W., & Xu, C. (2012). Using support vector machine models for crash injury severity analysis. *Accident Analysis & Prevention*, *45*, 478-486.

Malli Sureshbabu, N., Selvarasu, K., Nandakumar, M., & Selvam, D. (2019). Concentrated growth factors as an ingenious biomaterial in regeneration of bony defects after periapical surgery: A report of two cases. *Case reports in dentistry*, 2019.

Mathew, M.G., Samuel, S.R., Soni, A.J., & Roopa, K.B. (2020). Evaluation of adhesion of Streptococcus mutans, plaque accumulation on zirconia and stainless steel crowns, and surrounding gingival inflammation in primary molars: Randomized controlled trial. *Clinical oral investigations*, *24*(9), 3275-3280.

Mehta, M., Tewari, D., Gupta, G., Awasthi, R., Singh, H., Pandey, P., & Satija, S. (2019). Oligonucleotide therapy: an emerging focus area for drug delivery in chronic inflammatory respiratory diseases. *Chemico-biological interactions*, *308*, 206-215.

Mokhtarimousavi, S., Anderson, J.C., Azizinamini, A., & Hadi, M. (2019). Improved support vector machine models for work zone crash injury severity prediction and analysis. *Transportation research record*, *2673*(11), 680-692.

Muthukrishnan, S., Krishnaswamy, H., Thanikodi, S., Sundaresan, D., & Venkatraman, V. (2020). Support vector machine for modelling and simulation of Heat exchangers. *Thermal Science*, *24* (1 Part B), 499-503

PC, J., Marimuthu, T., Devadoss, P., & Kumar, S.M. (2018). Prevalence and measurement of anterior loop of the mandibular canal using CBCT: A cross sectional study. *Clinical implant dentistry and related research*, *20*(4), 531-534. https://europepmc.org/article/med/29624863.

Pradhan, B., & Sameen, M.I. (2020). Modeling traffic accident severity using neural networks and support vector machines. *In Laser Scanning Systems in Highway and Safety Assessment*, 111-117.

Ramadurai, N., Gurunathan, D., Samuel, A.V., Subramanian, E., & Rodrigues, S.J. (2019). Effectiveness of 2% Articaine as an anesthetic agent in children: randomized controlled trial. *Clinical oral investigations*, *23*(9), 3543-3550

Ramesh, A., Varghese, S., Jayakumar, N.D., & Malaiappan, S. (2018). Comparative estimation of sulfiredoxin levels between chronic periodontitis and healthy patients–A case-control study. *Journal of periodontology*, *89*(10), 1241-1248.

“Road Traffic Injuries.” https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Samuel, M.S., Bhattacharya, J., Raj, S., Santhanam, N., Singh, H., & Singh, N.P. (2019). Efficient removal of Chromium (VI) from aqueous solution using chitosan grafted graphene oxide (CS-GO) nanocomposite. *International journal of biological macromolecules*, *121*, 285-292.

Samuel, S.R., Acharya, S., & Rao, J.C. (2020). School Interventions–based Prevention of Early-Childhood Caries among 3–5-year-old children from very low socioeconomic status: Two-year randomized trial. *Journal of public health dentistry*, *80*(1), 51-60.

Sathish, T., & Karthick, S. (2020). Wear behaviour analysis on aluminium alloy 7050 with reinforced SiC through taguchi approach. *Journal of Materials Research and Technology*, *9*(3), 3481-3487.

Sharma, P., Mehta, M., Dhanjal, D.S., Kaur, S., Gupta, G., Singh, H., & Satija, S. (2019). Emerging trends in the novel drug delivery approaches for the treatment of lung cancer. *Chemico-biological interactions*, *309*, 108720.

Sridharan, G., Ramani, P., Patankar, S., & Vijayaraghavan, R. (2019). Evaluation of salivary metabolomics in oral leukoplakia and oral squamous cell carcinoma. *Journal of Oral Pathology & Medicine*, *48*(4), 299-306.

Varghese, S.S., Ramesh, A., & Veeraiyan, D.N. (2019). Blended Module-Based Teaching in Biostatistics and Research Methodology: A Retrospective Study with Postgraduate Dental Students. *Journal of dental education*, *83*(4), 445-450.

Venu, H., Raju, V.D., & Subramani, L. (2019). Combined effect of influence of nano additives, combustion chamber geometry and injection timing in a DI diesel engine fuelled with ternary (diesel-biodiesel-ethanol) blends. *Energy*, *174*, 386-406

Venu, H., Subramani, L., & Raju, V.D. (2019). Emission reduction in a DI diesel engine using exhaust gas recirculation (EGR) of palm biodiesel blended with TiO2 nano additives. *Renewable Energy*, *140*, 245-263.

Vignesh, R., Ditto Sharmin, C., Annamalai, S., & Baghkomeh, P.N. (2019). Management of complicated crown-root fracture by extra-oral fragment reattachment and intentional reimplantation with 2 years review. *Contemporary clinical dentistry*, *10*(2), 397-401.

Jain, S.V., Muthusekhar, M.R., Baig, M.F., Senthilnathan, P., Loganathan, S., Wahab, P.A., & Vohra, Y. (2019). Evaluation of three-dimensional changes in pharyngeal airway following isolated lefort one osteotomy for the correction of vertical maxillary excess: a prospective study. *Journal of maxillofacial and oral surgery*, *18*(1), 139-146.

Vijayashree Priyadharsini, J. (2019). In silico validation of the non-antibiotic drugs acetaminophen and ibuprofen as antibacterial agents against red complex pathogens. *Journal of periodontology*, *90*(12), 1441-1448

Zhang, J., Li, Z., Pu, Z., & Xu, C. (2018). Comparing prediction performance for crash injury severity among various machine learning and statistical methods. *IEEE Access, 6*, 60079-60087.