BALTIC JOURNAL OF LAW & POLITICS
A Journal of Vytautas Magnus University
VOLUME 15, NUMBER 4 (2022)
ISSN 2029-0454

Cite: *Baltic Journal of Law & Politics* 15:4 (2022): 168-174
DOI: 10.2478/bjlp-2022-004017

Drowsiness Detection of Driver using Novel Random Forest Classifier and Logistic Regression Classifier with Improved Accuracy

D. Nikitha
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

S. Kalaiarasi
Project Guide, Corresponding Author, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105

Received: August 8, 2022; reviews: 2; accepted: November 29, 2022.

ABSTRACT

Aim: The aim is to detect the driver drowsiness using Logistic Regression Classifier method in Comparison with Novel Random Forest Classifier. **Materials and Methods:** Two groups such as logistic regression Classifier and the novel Random Forest Classifier are applied. Each group has a sample size of 10. SPSS was used for predicting significance value of data set considering G-power value as 80%. Using different sample sizes, their accuracies are compared to each other. **Result:** Novel Random forest classifier provides a higher accuracy of 63.20% when compared to Logistic Regression Classifier with accuracy 60.00% in predicting stock driver drowsiness detection. There is a significant difference between two groups with a significance value of 0.001 ($p<0.001$). **Conclusion:** The results show that the Novel Random Forest classifier detects driver drowsiness better than the logistic Regression Classifier algorithm. It can also be thought of as a superior choice for classifying driver drowsiness.

Keywords

Machine Learning, Road Accidents, Drowsiness Detection, Logistic Regression Classifier, Novel Random Forest Classifier, Driver drowsiness.

INTRODUCTION

Driver drowsiness is a leading cause of road accidents all around the world, according to extensive statistical evidence (Roshini et al. 2021). If you don't take breaks while driving for long periods of time, you can end up in an accident. According to the World Health Organization, South Africa has the highest rate of road traffic accident fatalities in Africa, with 26.6 percent per 100 000 people. Machine learning Novel Random Forest Classifier stock driver drowsiness detection Using logistic Regression Classifier to predict drowsiness detection. Furthermore, 1,700 individuals perished on South African roads during the festive season of 2016, up 5% from the previous year. The South African transport ministry released a study on the statistics for the 2014-2015 fiscal year, revealing that 80

percent of road accidents involve adult males aged 19 to 34 (Shukla et al. 2020). Furthermore, the minister stated that women are more likely to die as passengers in car accidents, particularly on public transportation. Furthermore, studies reveal that distracted drivers (for example, a motorist on a phone conversation), speeding, and driving under the influence of alcohol are the top three causes of road accidents in South Africa. These tragedies have prompted researchers all across the world to look at methods for detecting and warning drowsiness early on. Furthermore, several countries and government leaders are focusing on putting solutions in place to increase driving safety (Xue and Bai 2016; Ecer et al. 2020). Drowsiness, often known as sleepiness, is a biological state in which the body is transitioning from an awake to a sleeping state. A motorist may lose focus at this point and be unable to perform steps such as avoiding head-on crashes or braking in a timely manner. There are several telltale symptoms that a motorist is drowsy.

Previously our team has a rich experience in working on various research projects across multiple disciplines (Venu and Appavu 2021; Gudipaneni et al. 2020; Sivasamy, Venugopal, and Espinoza-González 2020; Sathish et al. 2020; Reddy et al. 2020; Sathish and Karthick 2020; Benin et al. 2020; Nalini, Selvaraj, and Kumar 2020).The drawbacks of this algorithm is to find out the locations where the road accidents are held and identify the reason, because there are many ways to get road accidents like lack of sleep, tideness, eye closure, yawning, etc (Swain, Pattnaik, and Gupta 2020; Khanna et al. 2020; Kumar and Mozar 2020; *Drowsy Driving and Automobile Crashes* 1998; Snášel et al. 2013). Devised a method for determining a person's breathing and, as a result, drowsiness based simply on heart rate monitoring. The approach showed to be a good predictor of respiration and as a result. The aim is to prove the accuracy rate using an enhanced Novel Random Forest Classifier in comparison with the Logistic Regression classifier for driver drowsiness detection .

MATERIALS AND METHODS

This study setting was done in Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences. The number of required samples in research are two in which group 1 is Novel Random Forest Classifier with group 2 of Logistic Regression Classifier. The sample size was estimated using G power and was set at 10 per group with a pre test power of 80%, threshold 0.001% and CI 95%. From the Kaggle repository, a dataset containing a collection of stocks was downloaded.

The data was collected from the IEEE-dataport.org open source website. The database consists of 29 columns and 210 rows, which was utilized to estimate software effort using the Novel Random Forest and logistic regression classifier. TrialID, ObsNum, and IsAlert were found in 180 samples from three different species. The Computer Vision of a jupyter notebook software is used to evaluate this research effort for display. An Intel Core i7 processor with 6GB of RAM was used in the hardware configurations. The system's software setup is 64-bit Windows OS, 64-bit processor, and 1TB HDD.

Novel Random Forest Classifier

The procedure for dividing a node in a random forest takes into account only a random subset of the features. This Novel Random Forest Classifier is widely used in Machine Learning to make predictions. Instead of searching for the greatest possible thresholds, you may make trees even more random by employing random thresholds for each feature (like a normal decision tree does). It has a big effect on drowsiness detection. So, the program predicts the driver's detection.

Pseudocode for Random Forest Classifier

Input: K is the training dataset.
Output:A class of testing dataset.
Step1: Import packages.
Step2: Create an input dataset.

Step3: Analyze the size of the taken input data.
Step4: Split the datasets for testing and training the dataset.
Step5: Apply Novel Random Forest Classifier.
Step6: Predict the results.

Logistic Regression Classifier
The statistical analysis approach of logistic regression Classifier is used to forecast a data value based on prior observations of a data set. Logistic regression Classifier comes under the machine learning A logistic regression Classifier model analyses the relationship between one or more existing independent variables to predict a dependent data variable.

Pseudocode for Logistic Regression Classifier
Input: K is the training dataset.
Output:A class of testing dataset.
Step1: Import packages.
Step2: Create an input dataset.
Step3: Analyze the size of the taken input data.
Step4: Split the datasets for testing and training the dataset.
Step5: Apply Logistic Regression Classifier.
Step 6: Predict the results.

Recall that the testing setup includes both hardware and software configuration choices. An Intel Core i7 7th generation processor with 16GB of RAM, an x86-based processor, a 64-bit operating system, and a hard drive are included in the laptop. The software is now running on Windows 10 and is written in Python. The accuracy value will appear once the programme is completed. Procedure: Laptop is connected to the internet via Wi-Fi. Collaborative search from Chrome to Google Python should be used to write the code. Execute the code. To save the file, put it on the CD and put it in a folder. Use the ID from the mail to log in. To get the accuracy and graph, run the code.

Statistical Analysis
SPSS is a statistical analysis software program. The independent variable is facial tracking and other attributes such as bar graphs are dependent variables taken for this work. For each group, the proposed system used ten iterations, with expected accuracy logged and analyzed. The significance between two groups was determined using an independent sample t-test.

RESULTS

Table 1 shows the accuracy value of iteration of Novel Random Forest Classifier and Logistic Regression Classifier. Table 2 represents the Group statistics results which depicts the Novel Random Forest Classifier with mean accuracy of 63.20% and standard deviation is 3.2, Logistic Regression Classifier has a mean accuracy of 60.00% and standard deviation is 0.8. Proposed Novel Random Forest Classifier provides better performance when compared to Logistic Regression Classifier. Table 3 shows the independent samples T-test value for Novel Random Forest Classifier and Logistic Regression Classifier with Mean difference as 3.2, std Error Difference as 0.81 Significance value is observed as 0.001 ($p<0.001$).
Figure 1 shows the bar graph comparison of mean of accuracy on Novel Random Forest Classifier and Logistic Regression Classifier. Mean accuracy of the Novel Random Forest Classifier is 63.20% and Logistic Regression Classifier is 60.00%. Compare with both Logistic Regression and Novel Random Forest Classifier. The Novel Random Forest Classifier has more accuracy.

DISCUSSION

In this investigation, Novel Random Forest Classifier had a much greater accuracy of 63.20% in predicting driver sleepiness detection than Logistic Regression Classifier with 60.00%. Novel Random Forest Classifier appears to give outcomes that are more consistent and have a lower standard deviation (Zainal, Khan, and Abdullah 2014).

The similar findings of the paper had an accuracy of 63% with RF which was used to predict Driver Detection. The proposed work of Mkhuseli Ngxande, Jules-Raymond Tapamo, Michael Burke reported LGhas 60% accuracy which is used to predict Driver Detection (Kumar and Mozar 2020). The work proposed by Michael Burke shows the RF has a better accuracy of 63%. LG is a parameter to measure Driver Detection which is used in both traditional and modern methods as per their research it opposes 63% has highest accuracy and LG will get least accuracy compared to other machine learning techniques which ranges between 60% when compared to other machine learning algorithms will get more accuracy than this (Roshini et al. 2021). By using LG for forecasting Driver Detection it will have key issues to pretend, Jules-Raymond Tapamo, in this paper shows LG has least accuracy of 60% Increasing the dataset's value only tends to get desired accuracy (Majaranta and Päivi 2011). Random Forest algorithm performs better with a combination of other machine learning algorithms (Breiman 2017).
The limitation of this research is that it cannot give appropriate results for smaller data. In this model it is not possible to consider all given feature variable parameters for training. The future scope of proposed work will be prediction of stock price based on classification using class labels for lesser time complexity (Sabet et al. 2012).

CONCLUSION

In this study, driver drowsiness detection using the Novel Random Forest Classifier has accuracy 63.20% and Logistic Regression has accuracy with 60.00%. By comparing both, the Novel Random Forest Classifier has more accuracy. The discussion of the research paper also proves that the Novel Random Forest Classifier method has better accuracy than the Logistic Regression and accuracy.

DECLARATION

Conflict of Interests
No conflict of interests in this manuscript

Authors Contribution
Author DN was involved in data collection, data analysis, and manuscript writing. Author DN, SKA was involved in conceptualization, data validation, and critical review of 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 the necessary Infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.
1. HKM Info Systems pvt.ltd
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering

REFERENCES

Benin, S. R., S. Kannan, Renjin J. Bright, and A. Jacob Moses. 2020. "A Review on Mechanical Characterization of Polymer Matrix Composites & Its Effects Reinforced with Various Natural Fibres." *Materials Today: Proceedings* 33 (January): 798–805.

Breiman, Leo. 2017. *Classification and Regression Trees*. Routledge.

Drowsy Driving and Automobile Crashes. 1998.

Ecer, Fatih, Sina Ardabili, Shahab S. Band, and Amir Mosavi. 2020. "Training Multilayer Perceptron with Genetic Algorithms and Particle Swarm Optimization for Modeling Stock Price Index Prediction." *Entropy* 22 (11). https://doi.org/10.3390/e22111239.

Gudipaneni, Ravi Kumar, Mohammad Khursheed Alam, Santosh R. Patil, and Mohmed Isaqali Karobari. 2020. "Measurement of the Maximum Occlusal Bite Force and Its Relation to the Caries Spectrum of First Permanent Molars in Early Permanent Dentition." *The Journal of Clinical Pediatric Dentistry* 44 (6): 423–28.

Khanna, Ashish, Deepak Gupta, Zdzisław Pólkowski, Siddhartha Bhattacharyya, and Oscar Castillo. 2020. *Data Analytics and Management: Proceedings of ICDAM*. Springer Nature.

Kumar, Amit, and Stefan Mozar. 2020. *ICCCE 2020: Proceedings of the 3rd International Conference on Communications and Cyber Physical Engineering*. Springer Nature.

Majaranta, and Päivi. 2011. *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies: Advances in Assistive Technologies*. IGI Global.

Nalini, Devarajan, Jayaraman Selvaraj, and Ganesan Senthil Kumar. 2020. "Herbal Nutraceuticals: Safe and Potent Therapeutics to Battle Tumor Hypoxia." *Journal of Cancer Research and Clinical Oncology* 146 (1): 1–18.

Reddy, Poornima, Jogikalmat Krithikadatta, Valarmathi Srinivasan, Sandhya Raghu, and Natanasabapathy Velumurugan. 2020. "Dental Caries Profile and Associated Risk Factors Among Adolescent School Children in an Urban South-Indian City." *Oral Health & Preventive Dentistry* 18 (1): 379–86.

Roshini, G., Y. Kavya, R. Hareesh, M. Suma, and N. Sunny. 2021. "Driver Distraction and Drowsiness Detection System." *2021 International Conference on Intelligent Technologies (CONIT)*. https://doi.org/10.1109/conit51480.2021.9498348.

Sabet, Mehrdad, Reza A. Zoroofi, Khosro Sadeghniiat-Haghighi, and Maryam Sabbaghian. 2012. "A New System for Driver Drowsiness and Distraction Detection." *20th Iranian Conference on Electrical Engineering (ICEE2012)*. https://doi.org/10.1109/iraniancee.2012.6292547.

Sathish, T., and S. Karthick. 2020. "Gravity Die Casting Based Analysis of Aluminum Alloy with AC4B Nano-Composite." *Materials Today: Proceedings* 33 (January): 2555–58.

Sathish, T., D. Bala Subramanian, R. Saravanan, and V. Dhinakaran. 2020. "Experimental Investigation of Temperature Variation on Flat Plate Collector by Using Silicon Carbide as a Nanofluid." In *PROCEEDINGS OF INTERNATIONAL CONFERENCE ON RECENT TRENDS IN MECHANICAL AND MATERIALS ENGINEERING: ICRTMME 2019*. AIP Publishing. https://doi.org/10.1063/5.0024965.

Shukla, Usha, Urvashi Yadav, Jay Brijesh Singh Yadav, and Sanket Agrawal. 2020. "Comparison of End-Tidal Anesthetic Gas Concentration versus Bispectral Index-Guided Protocol as Directing Tool on Time to Tracheal Extubation for Sevoflurane-Based General Anesthesia." *Anesthesia, Essays and Researches* 14 (4): 600–604.

Sivasamy, Ramesh, Potu Venugopal, and Rodrigo Espinoza-González. 2020. "Structure, Electronic Structure, Optical and Magnetic Studies of Double Perovskite Gd2MnFeO6 Nanoparticles: First Principle and Experimental Studies." *Materials Today Communications* 25 (December): 101603.

Snášel, Václav, Pavel Krömer, Mario Köppen, and Gerald Schaefer. 2013. *Soft Computing in Industrial Applications: Proceedings of the 17th Online World Conference on Soft Computing in Industrial Applications*. Springer Science & Business Media.

Swain, Debabala, Prasant Kumar Pattnaik, and Pradeep K. Gupta. 2020. *Machine Learning*

and Information Processing: Proceedings of ICMLIP 2019. Springer Nature.
Venu, Harish, and Prabhu Appavu. 2021. "Experimental Studies on the Influence of Zirconium Nanoparticle on Biodiesel–diesel Fuel Blend in CI Engine." *International Journal of Ambient Energy* 42 (14): 1588–94.
Xue, Hongxin, and Yanping Bai. 2016. "Research on Prediction of Stock Index Based on PCA and SVM." *Applied Mechanics, Mechatronics and Intelligent Systems*. https://doi.org/10.1142/9789814733878_0124.
Zainal, Mohd Shamian Bin, Ijaz Khan, and Hadi Abdullah. 2014. "Efficient Drowsiness Detection by Facial Features Monitoring." *Research Journal of Applied Sciences, Engineering and Technology* 7 (11): 2376–80.

TABLES AND FIGURES

Table 1. Accuracy Values for RF and LG

S.NO	RF	LG
1	65	63
2	64	62
3	63	61
4	63	63
5	62	60
6	64	59
7	61	63
8	65	58
9	63	57
10	60	62

Table 2. Group Statistics Results-RF has an mean accuracy (63.20%), std.deviation (1.61), whereas for LG has mean accuracy (60.00%), std.deviation (2.00).

Group Statistics					
	Groups	**N**	**Mean**	**Std deviation**	**Std. Error Mean**
Accuracy	RF	10	63.20	1.61	0.51
	LG	10	60.00	2.00	0.63

Table 3. Independent Samples T-test - RF seems to be significantly better than LG(p=0.001)

Accuracy	Independent Samples Test								
	Levene's Test for Equality of Variances		T-test for Equality of Means						
	F	Sig	t	df	Sig(2-tailed)	Mean Difference	Std.Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	0.250	0.623	3.932	18	0.001	3.200	0.814	1.490	4.910
Equal variances not assumed			3.932	17.253	0.001	3.200	0.814	1.485	4.915

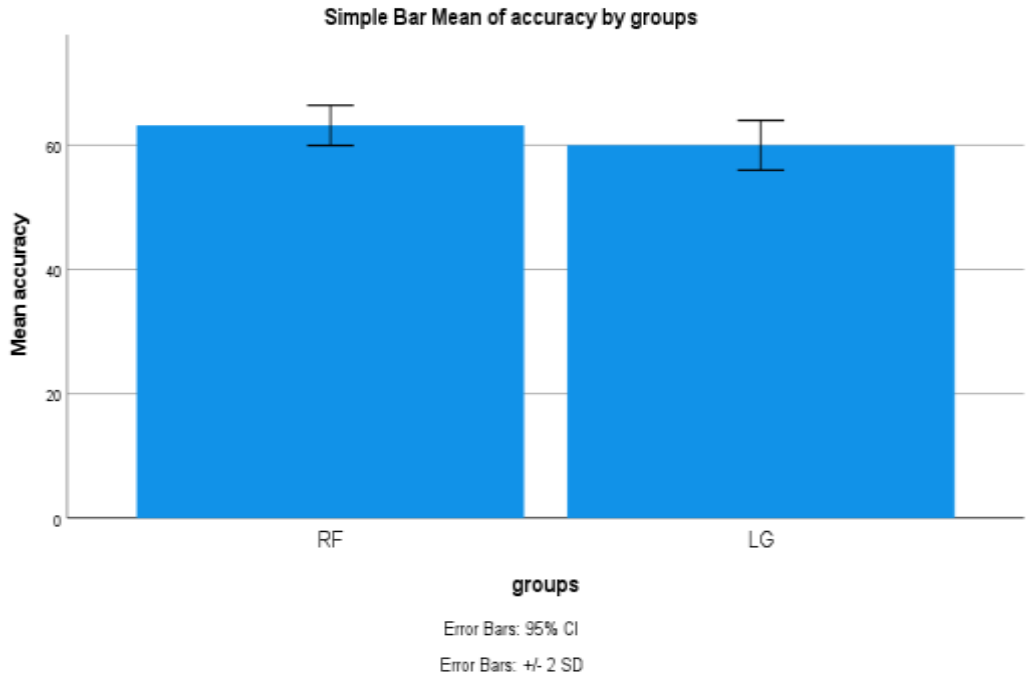

Fig. 1. Bar Graph Comparison on mean accuracy of RF(63.20%) and LG (60.00%). X-axis:RF,LG, Y-axis: Mean Accuracy with ±1 SD.