JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DRIVING BEHAVIOR STUDY USING HYBRID MACHINE LEARNING FRAMEWORK WITH RECOMMENDATION SYSTEM

[1]Meghana Prakash, [2]Ravi V

[1, 2]Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumakuru, India

[1]meghanaprakash77@gmail.com , [2]ravi@sit.ac.in

Abstract : In today's modern world, everyone is in a rush to accomplish their objectives (to meet targets/destinations) as quickly as possible. People drive on the road and drive faster in order to get there, which might result in traffic accidents and severe/serious injuries. It's also a general belief that closely monitoring a person's actions would significantly reduce accidents while also enabling safe driving. A driver behaviour analysis and prediction system has been proposed as a solution to this problem, which analyzes the vehicle's speed mean while collects the essential information about reckless driving, which is a frequent occurrence among drivers. For each ride taken by the driver, a real-time data is obtained, and the data is then uploaded (updated) in the Cab Management System (CMS) for interpretation. Drivers can be employed, rewarded (by incentives), and promoted using CMS data. The strategic recommendation aids in the recruitment and advancement of employees, as well as the reduction of road accidents.

***IndexTerms* - Cab Management System, Traffic Accidents, Rash Driving, Machine Learning, Naive Bayes Algorithm, Random Forest Algorithm.**

I. INTRODUCTION

The number of vehicles on the road now is increasing every year, in lockstep with economic expansion in every country. Simultaneously, the majority of non drivers are rapidly increasing. Because most inexperienced drivers are untrained, unfamiliar with automotive conditions, and unaware of traffic regulatory standards, behavioural variables are becoming the leading causes of traffic accidents. There seems to be a lot of research done in this space beforehand, but researchers required specialised hardware installed inside the automobile or on the wayside, which is intensive and risky upkeep; as a result, past studies were mostly aimed at examining either driving patterns/behavior or road conditions. This suggested technique computes the speed and position of vehicle using a dataset containing the driver ID, latitude, and longitude (location) of the vehicle. The space will be calculated using latitude and longitude (distance). It also generates a driving pattern depending on distance travelled among distinct points by the vehicle. This will find things simpler to figure out the driver's driving patterns. The system employs a dataset to generate results.

The information extracted from a sample dataset can then be used to gather insight using machine learning techniques [9]. This study employed the transportation dataset, which incorporates driving data processed by a vehicle's accelerometer, gyroscope, and proximity sensor. A segmentation or classification technique will be applied on data that has been labelled [3]. In previous studies [6], [10], and [11], the researchers looked at the drivers' driving patterns. The bulk of the researchers used gyroscope sensors in their trials, but some used SVM, decision trees, fuzzy logic, and other techniques.

There are two class labels in the example dataset: "class 0" for a safe drive/trip and "class 1" for the dangerous drive/trip, according to data analysis, which uses statistical science to derive the dataset's properties from the technique. Safe travel data outnumbers hazardous travel data in the sample data label by a 3:1 ratio. Classifiers seem to be more likely to uncover information into predominant classes when a dataset is uneven, posing a problem in the classification process [12]. The classifier was prone to misclassifying harmful trips as secure (safe) due to the unbalanced dataset utilised in this work, resulting in a low sensitivity value for dangerous trips. Because of the low sensitivity, classifier performance deteriorates, and numerous errors arise in the recognition of transportation information.

Figure 1: Reckless Driving.

Cab service connects cab drivers with passengers, consumers, and folks looking to rent a cab. Customers can book a cab using the service via the internet. This enables customers to discover registered travel businesses, cab operators, and cab owners. This proposed solution recommends the Cab Management System (CMS) to monitor drivers' incentives and promotions, based on the aforementioned database and public (passengers) inputs. They also recommend providing employment for the drivers based on their preliminary/early observations.

II. LITERATURE SURVEY

A literature review has been conducted at this step to investigate the classification techniques employed in prior studies. Relying on the dataset presented in the case, there are a variety of approaches to recognize driving behaviour. The following are several previous researches on the identification of driving activities using pertinent research sensor readings (data).

2.1. In vehicle-driver-environment systems, driver behaviour is a significant factor. Real-time driving monitoring technologies play an important role in improving driver safety. Fatigue, interruption, expertise, climatic circumstances, vehicle condition, and other aspects can influence the driver's conduct. Inappropriate driving behaviour will increase the likelihood of an accident occurring. As a result, detecting driver behaviour is a fascinating study topic. The driving monitoring system warns the driver about his unsafe driving by recognizing and differentiating among regular and reckless driving. The driver can improve his driving technique and reduce the likelihood of an accident occurring. The accuracy of the driving monitoring system is a critical component. Depending on the system's goals, different methodologies have been employed to identify driver behaviour. Driving monitoring systems rely heavily on sensors. This paper addresses different approaches for detecting driver behaviour that have also been presented, as well as the benefits and drawbacks of each method.

2.2. The software collects data from accelerometers, GPS, and a microphone to record audio, which is then combined and analysed to find risky/rash driving patterns. Many patterns, such as speed breaker, lane shift and turn (left/right), rapid halting, and abrupt acceleration, were assessed and verified using 'Ground Truth.' Audio and accelerometer data are linked to reveal new patterns. For illustration, a lane change without an indicating signal denotes a hasty driving action. The study's use of machine learning techniques to define driving types is a flaw.

2.3. A one-of-a-kind programme that assesses driving style using a smart phone that is connected into a vehicle. Using the three-axis accelerometer of an Android-based Smartphone, researchers identify and analyse a variety of driver characteristics or patterns and external road circumstances that could be hazardous/harmful to the driver's wellbeing. They measured the driver's direct control of the vehicle as they steer, accelerate, and brake using accelerometer data from the x and y axes. The maximum safe acceleration or deceleration is 0.3 g, however rapid acceleration or deceleration can reach 0.5 g. It is simple to measure the difference among safe and abrupt acceleration and deceleration using this comparison. Secure right/left lanes have a g force of less than 0.1 g, while unsafe or sudden right/left lanes have just average g force of more than 0.5 g. It was discovered that completing a safe lane change took 75 percent longer than completing a spontaneous lane change. A device’s (phone) location in a vehicle was also observed, and location 1, the centre console, offered the best relevant data with the least amount of engine input. The study's drawback is that the phone's substantial performance for forecasting driving patterns was achieved when it was installed on a canter dashboard; however, because the phone can be mounted anywhere in a car, a system to digitally re-orient the accelerometer is required.

III. PROBLEM STATEMENT

Driver behaviour may play a vital role in road accidents. Machine Learning techniques must be employed to provide driving behaviour analysis and prediction models to overcome this issue. To assign drivers to certain trips/drives and manage driver promotions and incentives, a Driver Recommendation System and a Cab Management System were required. Cab agencies’ long-term objective is to eliminate uncertainty/risk. Cab services/companies and agencies can improve their customer services by examining data and getting insights (information) from sensors and earlier track records. It can assist clients in identifying risk areas and implementing customised solutions to help maintain profitability or prevent losses.

IV. PROPOSED SYSTEM

We can enhance the accuracy/correctness of the driving behavior analysis formed on the sample dataset in this approach, because there is absence of optimised pattern analysis in the driving styles (existing system) and existing approaches may not deliver the accurate analysis of the drivers' driving style because all potential/possible combinations of patterns are not evaluated while analysing. As a result, the pattern analysis as well as the activities (actions) performed on the basis of that analysis, will be incorrect. Therefore, a solution is required which enhances the correctness of driving behavior analysis using input data.

Incorporating Naive Bayes and Random Forest techniques (machine learning methods), the proposed methodology examines the drivers' driving behaviours). By using these two algorithms, accuracy will be checked. Then based on the accuracy predictions or forecasts further process will be carried out by the analysis of prior data, current pattern or style, as well as certain private information. It will assist cab agencies/companies in risk management. Driving strategies might encourage individuals to drive more cautiously, so boosting road safety. The analysis will give prediction whether a particular driver is rash/not in peak/non peak hours. Once prediction is done based on some constraints imposed on the dataset, these results will be used in recommendation engines like Driver Recommendation for the drive/trip and Cab Management System (CMS) in hiring drivers and managing their incentive & promotions, where customers' review and past track records play a vital role.

V. METHODOLOGY

5.1 Data Extraction

The dataset may be read with the java POI API, which returns all of the dataset's records. For reading and writing Excel, Apache contains a well-known library called POI. This library supports both the XLS and XLSX formats for Excel files. An Excel file can be read using the Java IO method. This will necessitate the use of an Apache POI Jar. POI interfaces are in a variety of shapes and sizes, including Workbook, Sheet, Row, and Cell. Apache POI has provided rudimentary text extraction for the entire projects supported file formats. In addition to the (plain) text, they allow access to the information associated with a certain file.

5.2 Data Clustering

The goal of this phase is to collect all (or almost all) of the data and organise it into clusters. Clustering is the division of a population or set of data points into different groups so that data points from one group are more similar to data points from other groups. To put it another way, the goal is to organise people into clusters based on their shared characteristics. Hard partitioning is a sort of grouping that allows an object to not be a member of a cluster or strictly belong to it. Soft partitioning, on the other hand, states that each item is in some way a part of a cluster. More complicated divisions are possible, such as things belonging to many clusters, compelling an object to operate in only one cluster, or even creating hierarchical trees based on group relationships.

K-means Clustering: In contrast to supervised learning, the unsupervised learning technique K-Means clustering does not use labelled data to group data. K-Means classifies things based on how similar and dissimilar they are to those in other groups. The letter 'K' stands for a number. You must inform the system of the number of clusters you require. K = 2 denotes two clusters, for example. There is a method for determining the best or optimal or optimum value of K for a given set of data.

5.3 Association Rule Mining Algorithm

The basic premise of association rule-mining is that driving behaviour has a pattern that may be used to avoid collisions. It is a rule-based machine learning technique for uncovering significant associations between variables in large databases. Its purpose is to find strong rules in databases using some interestingness measurements. This rule-based technique produces new rules as more data is analysed. The ultimate goal, assuming a large enough dataset, is to let a machine imitate the human brain's feature extraction and abstract association abilities from new unsorted input.

VI. METHODOLOGY

6.1 Naive Bayes Algorithm

It's a probabilistic machine learning method for classifying a wide variety of objects. Spam Filtering, document categorization, and predicting or forecasting sentiment are all typical applications. The name is derived from the works of Rev. Thomas Bayes. "Naive" was the term given to it for a reason?

The naive assumption is that the design's features are irrelevant to one another. In other words, changing the value of attribute has no bearing/effect on the value of algorithm's remaining (other) attributes/features.
Using P(c), P(x), and P(x|c), the Bayes theorem may be used to calculate the posterior probability P(c|x).

$$P(X/Y) = P(Y/X) * P(X) / P(Y)$$

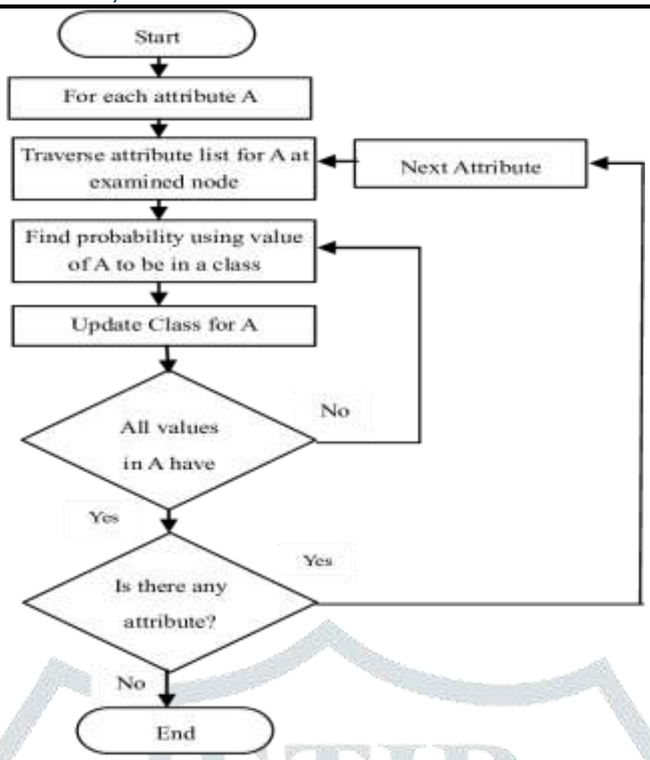

Figure 2: Flow diagram of Naive Bayes.

How does the above algorithm work?

Step 1: From the gathered information, design a frequency table.
Step 2: Construct a Likelihood chart using the probabilities.
Step 3: Calculate the posterior probability for each class using the Naive Bayesian equation.
The outcome of prediction is the class with the highest posterior probability.

6.2 Random Forest Algorithm

The supervised learning method is used by Random Forest, a well-known machine learning algorithm. It can be used for both/either classification and regression problems in machine learning. It's based on ensemble learning, which is a method of combining multiple classifiers to solve complicated issues and enhances the model's performance. According to the term, "Random Forest (RF) is a classifier typically incorporates a number of decision trees on various subsets of a training sample and chooses the mean to boost the dataset's projected accuracy."Rather than relying on a single decision tree, it collects forecasts out of each tree and anticipates or predicts the final output based on the majority votes of projections/forecasts. The more trees in the forest, the more precise it becomes, avoiding the issue of over-fitting.

How does the above algorithm work?

The random forest is developed in two stages: the first includes integrating N decision trees to form the random forest, and the next involves making predictions for every tree formed in the first phase. To demonstrate how well the process works, use the steps below:

Step 1: Select K data points from the training set at random.
Step 2: For the data points you've picked, make decision trees (Subsets).
Step 3: Choose N for the number of decision trees you want to construct.
Step 4: Repeat the above steps 1 and 2 if required.
Step 5: Locate each decision tree's forecasts for latest data points, and assign the latest data points to the category/classification with the highest votes.

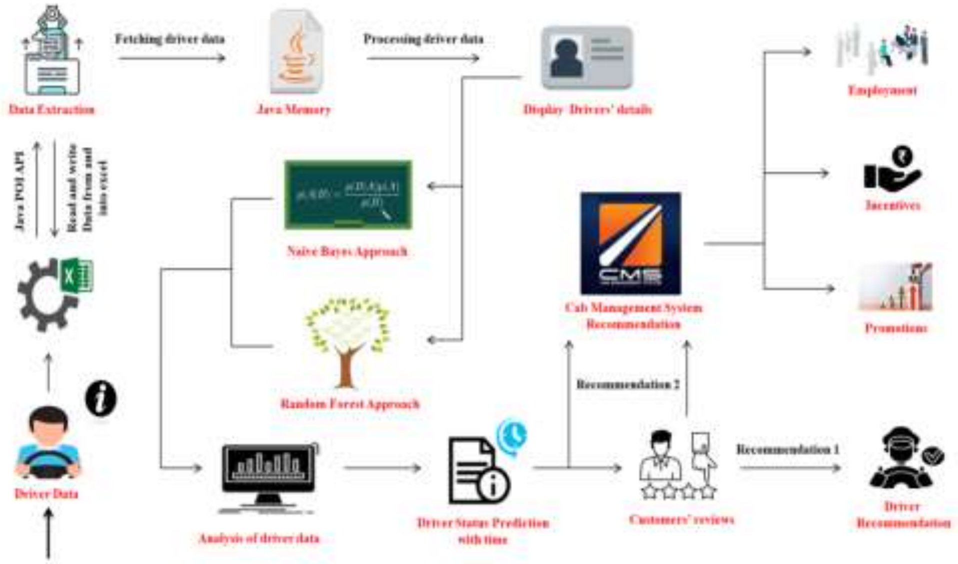

Figure 3: System Architecture.

VII. IMPLEMENTATION

The proposed system is basically implemented to analyse whether the driver is rash or not with respect to time (peak/non peak hours) on the basics of drivers' past records and by some other parameters, then prediction and analysis will be done based on mentioned machine learning techniques.
Cab management admin should login to the page with correct credentials.

Figure 4: Login Page.

Once the admin has gained access, they will proceed to the prediction page. The admin can select the driver ID from a dropdown list. The data of the chosen driver will be maintained in a database that has been developed using statistics from the cab agency and other sources.

1	Customer	Claim Amount	Coverage	Gender	Policy	Vehicle Size	no of accident record	no of rule breaking records	no of times threshold crossed	active status	inactive days	age	date
2	BU79786	276.3519279	Basic	F	Corporate L3	Medsize	6	15	104.0	yes	0	51	01-05-2021
3	QZ44356	697.9535903	Extended	F	Personal L3	Medsize	4	6	159.0	yes	0	28	01-05-2021
4	AI49188	1288.743165	Premium	F	Personal L3	Medsize	6	15	115.0	yes	0	51	01-05-2021
5	WW63253	764.5861827	Basic	M	Corporate L2	Medsize	6	4	126.0	no	45	45	01-05-2021
6	HB64268	281.3692575	Basic	M	Personal L1	Medsize	30	34	90.0	yes	0	23	01-05-2021
7	OC83172	825.62978	Basic	F	Personal L3	Medsize	15	25	74.0	yes	0	58	01-05-2021
8	XZ87318	538.0898636	Basic	F	Corporate L3	Medsize	4	7	116.0	yes	0	56	01-05-2021
9	CF85061	721.6100311	Premium	M	Corporate L3	Medsize	7	20	150.0	yes	0	34	01-05-2021
10	DY87989	2412.750402	Basic	M	Corporate L3	Medsize	12	21	95.8	no	60	21	01-05-2021
11	BQ94931	738.8178085	Extended	F	Special L2	Medsize	6	18	59.0	yes	0	32	01-05-2021
12	SX51350	473.8992022	Basic	M	Personal L3	Small	3	5	13.0	yes	0	28	01-05-2021
13	VQ65197	819.7197078	Basic	F	Personal L3	Medsize	4	6	183.0	yes	0	28	01-05-2021
14	DP39365	879.8797003	Premium	M	Corporate L1	Medsize	9	30	57.0	yes	0	38	01-05-2021
15	SJ95423	881.9018934	Basic	M	Corporate L3	Medsize	0	2	32.0	yes	0	40	01-05-2021
16	IL66569	538.4431665	Basic	M	Corporate L3	Medsize	0	1	12.0	yes	0	33	01-05-2021
17	BW63560	746.3139377	Basic	F	Corporate L2	Medsize	3	5	24.0	yes	0	53	01-05-2021
18	FV94802	256.8867823	Basic	M	Personal L3	Medsize	4	10	26.0	yes	0	60	01-05-2021
19	OE15005	394.5241604	Basic	M	Personal L2	Medsize	0	0	66.0	yes	0	20	01-05-2021
20	WC83389	571.0333115	Basic	M	Personal L2	Medsize	2	8	24.0	no	80	45	01-05-2021
21	FL50705	816.2617053	Premium	F	Corporate L2	Small	9	33	75.0	yes	0	29	01-05-2021
22	AB34567	342.566	Premium	F	Personal L3	Medium	3	12	22.0	yes	0	30	01-05-2021
23	BU79786	276.3519279	Basic	F	Corporate L3	Medsize	1	2	31.0	yes	0	50	15/2/2021
24	QZ44356	697.9535903	Extended	F	Personal L3	Medsize	4	8	65.0	yes	0	28	15/2/2021
25	AI49188	1288.743165	Premium	F	Personal L3	Medsize	1	2	31.0	yes	0	50	15/2/2021
26	WW63253	764.5861827	Basic	M	Corporate L2	Medsize	6	4	94.0	no	45	45	15/2/2021
27	HB64268	281.3692575	Basic	M	Personal L1	Medsize	30	34	61.0	yes	0	23	15/2/2021
28	OC83172	825.62978	Basic	F	Personal L3	Medsize	15	25	53.0	yes	0	58	15/2/2021
29	XZ87318	538.0898636	Basic	F	Corporate L3	Medsize	4	7	88.0	yes	0	56	15/2/2021
30	CF85061	721.6100311	Premium	M	Corporate L3	Medsize	7	20	119.0	yes	0	34	15/2/2021
31	DY87989	2412.750402	Basic	M	Corporate L3	Medsize	12	21	99.8	no	60	21	15/2/2021
32	[illegible]	[illegible]	[illegible]	F	[illegible]	[illegible]	6	18	[illegible]	[illegible]	0	[illegible]	[illegible]

Figure 5: Drivers' prior accomplishments, present conduct, and personal information.

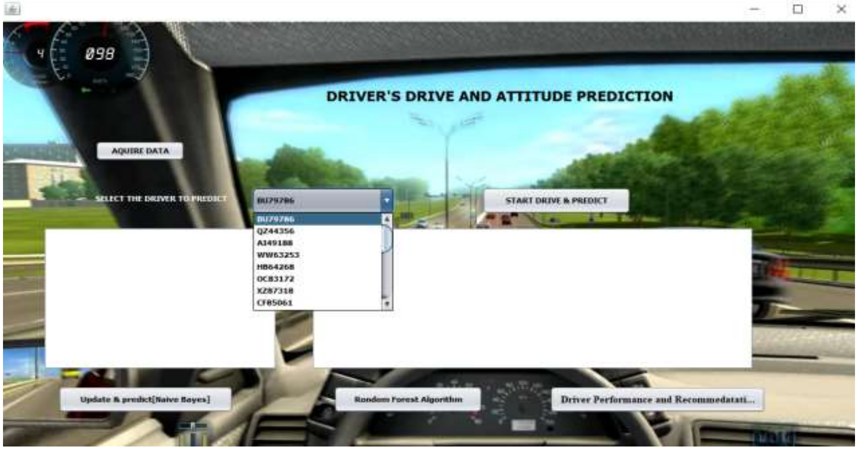

Figure 6: Prediction Page (Selecting driver ID for driver behavior prediction).

1. Acquire Data button: will use the Java POI API to fetch the driver details.

2. Select driver to predict button: this option allows the administrator to choose a specific driver ID. This is used to predict if the selected driver is driving recklessly or rashly during peak as well as non-peak hours.

3. Following the processes outlined above, data will be successfully configured; the driver's forecast/prediction will then implemented based on the database information. The information gathered includes the maximum speed permitted in specific places and the actual speed of the driver driving the vehicle to that location. This is accomplished through the use of the "K means clustering" and "Association rule algorithm."

4. Start drive and predict button: based on prior track records and personal information, the prediction is made using two techniques (Naive Bayes and Random Forest).

Table 1: Driving limits of various locations.

Longitude	Latitude	Speed limits
12.947967	77.69907	70
12.948981	77.699338	70
12.949525	77.699424	70
12.949933	77.69951	70
0	0	0
12.950065	77.700202	70
12.949897	77.700921	60
12.949772	77.701404	60
12.94972	77.701629	60
0	0	0
12.949416	77.701554	50
12.949155	77.701468	50
12.948894	77.701565	50
12.948789	77.701887	50
0	0	0
12.949082	77.702037	40
12.949385	77.702187	40
12.949312	77.702435	40

Step 1: Naive Bayes Approach

To know whether the driver is rash or not we are implementing machine learning (Naive Bayes) technique on driver data.

1. Update and predict button: will give the implementations of Naive Bayes algorithm by predicting whether the driver is rash or not and parameters which are considered to analyse this are: age, gender, number of accidents, number of rules broken, vehicle size, etc…(datas collected by driver database).

2. Based on the above attributes all the conditions will be monitored. If any of the conditions crosses the threshold limits with corresponding time (peak or non peak hours) then the driver is considered as a rash driver.

3. The final prediction will be done by analyzing drivers past records and by the obtained results.

Figure 7: Naive Bayes Approach.

Step 2: Random Forest Approach

We follow the same procedure as in Naive Bayes approach to carry out predictions.
The implementation of hybrid models will help us to know which algorithm best suits for predicting drivers' behaviour.

Figure 8: Random Forest Approach.

Step 3: Performance Analysis of the driver and Recommendation System

Figure 9: UI of the Cab Management Analytics.

Recommendation 1: Driver Recommendation for drive/trip

1. Customer's Reviews regarding the previous drives/trips are stored in the database.

2. If a customer books a cab, then the cab management will search for the type of trip (short/medium/long) and based on the reviews of customers and previous track records, the driver will be allotted for a particular drive/trip.

3. If the driver data meets the proposed constraints, then he/she will be considered for long drives/trips. It is to ensure the safety of the customers as well as the cab.

Figure 10: Recommends Cab Management about which driver has to be assigned for particular trip/drive.

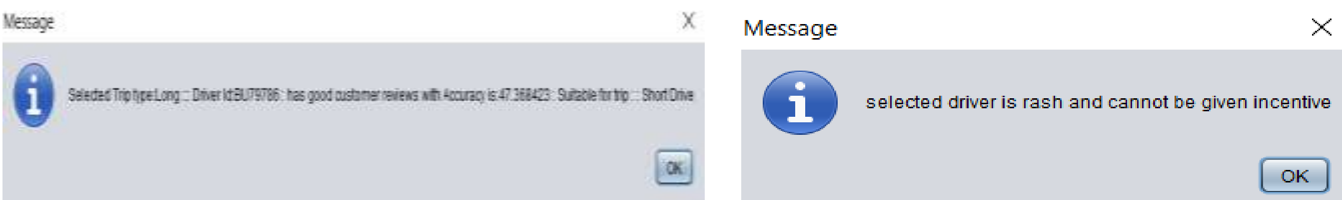

Figure 11: Recommendation of driver BU79786 based on customer's reviews.

Recommendation 2: Cab Management System

Figure 12: Recommendations for CMS (Driver ID: VQ65197)

This project is beneficial to cab companies/agencies' since it advises them in providing employment and to maintain drivers' incentives and promotions. These are thoroughly explained.

Employment: is based on prior track records, current documentation, and customer feedback (if any), and candidates are only selected if they have a solid track record, in order to ensure the safety of the vehicle, customers, and driver. It also lowers maintenance expenses because the car is protected from wear and tear costs caused by aggressive driving.

A. Outcome 1 of CMS - Employment (Driver ID: BU79786)

Incentives: The second feature of CMS is that it suggests whether or not a driver should be given an incentive. This is based on a three-month track record with favourable customer feedback. The chosen driver will be rewarded if he/she has a good track record and has received positive comments. The pay scale will be increased by ₹1000. There will be a reduction of ₹1000 if personnel do not meet the above-mentioned criteria. In moderate circumstances, there will be no modifications to the pay scale.

B. Outcome 2 of CMS - Incentive (Driver ID: BU79786)

Promotion: This is the most important aspect of every employee's life, and achieving it is certainly not easy. To advance in one's career, one must work diligently and patiently. In this application, the driver is promoted over a year of excellent performance and, of course, with positive comments.

C. Outcome 3 of CMS - Promotion (Driver ID: BU79786)

VIII. RESULTS

This project was developed using the Netbeans IDE and Java library packages. The Random Forest Approach and Naive Bayes had been used to develop this. Using these Machine Learning approaches, it was trivial to predict and analyse driver behaviour. The results of the procedures listed above are extracted and stored for use in recommendation engines (Driver Recommendation and Cab Management System).

The predictions were acquired by monitoring drivers' prior track records, current behaviour, and some personal information, and were implemented using Naive Bayes (figure 13) and Random Forest (figure 14) techniques. Some constraints were imposed on the acquired data in order to ensure accurate prediction and analysis. The conduct of drivers are predicted and analysed based on their gender; the results are given in Figure 16. Figure 17 shows the attributes that were taken into account for clustering and the association rule. If a driver's conduct is within the threshold limit, both during peak and non-peak hours, he is not a rash driver.

Figure 13: Naive Bayes Approach.

Figure 14: Random Forest Approach.

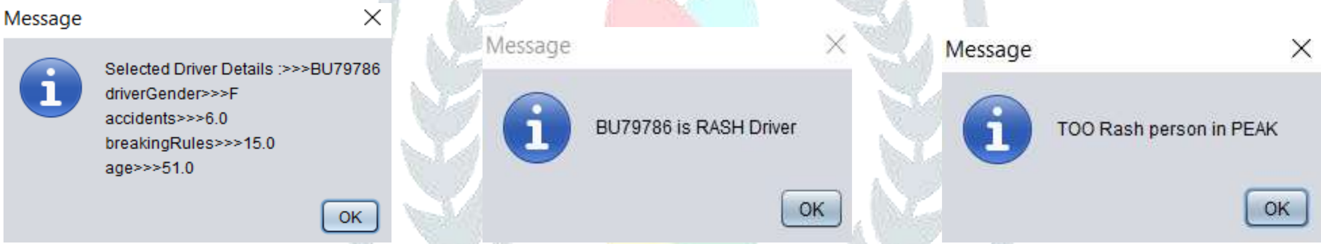

Figure 15: Driving Behavior Prediction (Driver ID: BU79786)

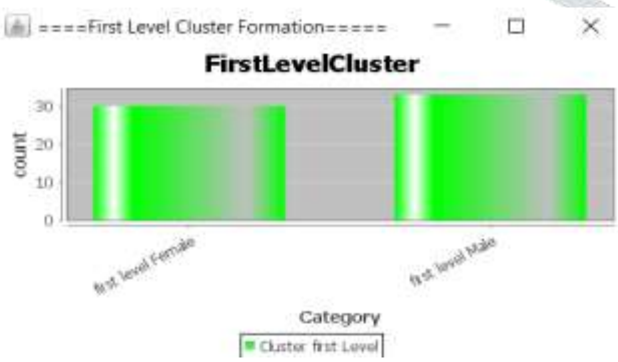

Figure 16: Cluster of drivers based on Gender.

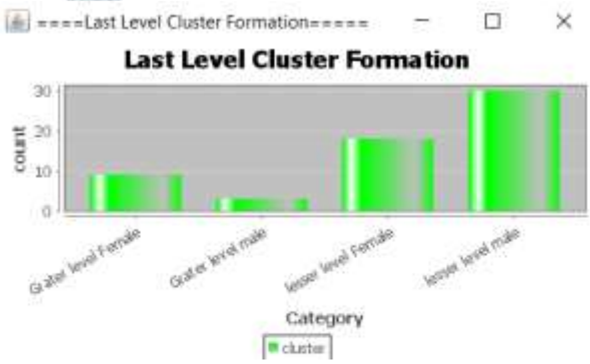

Figure 17: Cluster of drivers based on Gender and Age (above and below 50)

The data is then incorporated into recommendation systems. The driver will be assigned to the drive/trip based on client feedback (Figures 18 and 19). Likewise, these reviews, together with statistics (documentation), assist CMS in managing incentives and promotions for drivers as well as in providing employment.

1. Analyzing the reviews collected by the customers.

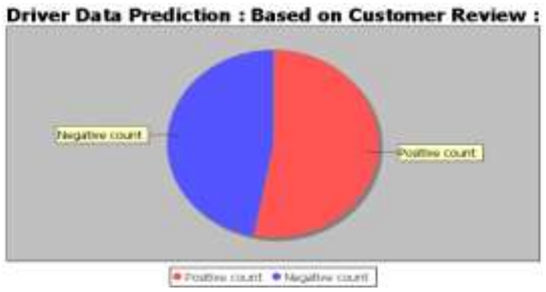

Figure 18 (a)

2. Driver Performance Analysis by customer reviews and past records

Figure 18 (b)

Figure 18 (a) shows the Customers review analysis of the driver (BU79786) in graph and Figure 18 (b) illustrates the Driver (BU79786) Performance Analysis by customer reviews and track records.

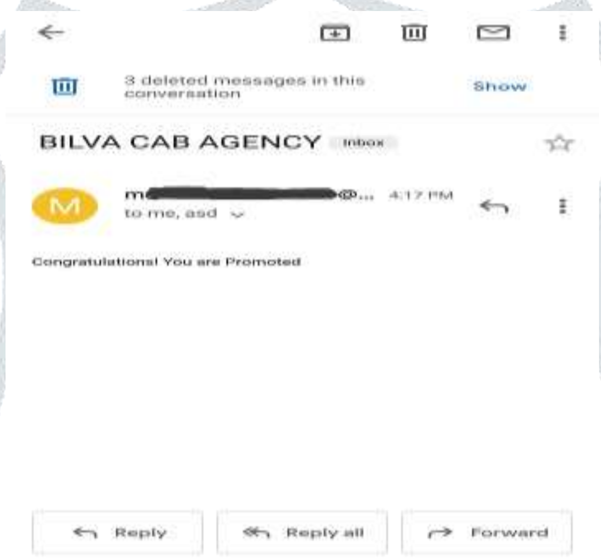

Figure 20: Snapshot of email showing the promotion details of the driver.

IX. Conclusion And Future Enhancement

This proposed system gives optimized driver pattern analysis with accurate outputs of the driver and will give details about the users' driving styles and patterns. It provides improved driving pattern analysis. The proposed system finds out the driving pattern and behaviour of the driver by applying the Naive Bayes and Random Forest algorithms on the dataset. These algorithms find the fitness of each of the solutions. If the fitness is less than the mean fitness, it discards and replace with the new solutions. And if the solution is more than mean fitness, pass to the next iteration and selects the solution which has the max fitness and use it as the driver pattern. These algorithms automate the development, acquisition, processing, extraction, and interpretation of data and information, removing the humanistic flaws associated with data acquisition. Based on the input dataset, this proposed system develops solutions that would enhance the accuracy of driving behavior analysis. This system collects the information from the dataset of the vehicle; it gives information about the speed, acceleration and driver's ID etc. This information is used to derive the pattern of driving. An outlier will be excluded from the sample; outliers can be caused by measurement variability or an experimental error, the latter of which will be excluded from the given data, dataset is used to find the pattern of driving and generate the solution for the driving pattern and also finds the fitness of each solution. One can analyse and compare both the results obtained from Random Forest and Naive Bayes approach and then predict the best result for driver data prediction. Based on the predictions and the reviews collected by the customers the proposed system recommends drivers for the drives/trips. It also recommends Cab Management System (CMS) in hiring drivers and managing their incentives and promotions.

Future Enhancement

This software should also promote cab companies to provide rental cars depending on prior driving records and rash status. This application must be capable of selecting a driver based on the vehicle's size (which is selected by the customer). This should not just apply to four-wheelers, but also to two-wheeler taxi services such as Rapido, UberMOTO, and others.

REFERENCES

[1] Nima Khairdoost, Mohsen Shirpour, Michael A. Bauer, Steven S. Beauchemin, Real-Time Driver Maneuver Prediction Using LSTM, IEEE Transactions (2020)

[2] Nasim Arbabzadeh, M. Jafari, A Data-Driven Approach for Driving Safety Risk Prediction Using Driver Behavior and Roadway Information Data, IEEE Transactions (2018)

[3] Nidhi Kalra and Divya Bansal, Analyzing Driver Behavior using Smartphone Sensors: A Survey, International Journal of Electronic and Electrical Engineering. ISSN 0974-2174 Volume 7, Number 7 (2014)

[4] Singh, P., Juneja, N., Kapoor, S., Using mobile phone sensors to detect driving behaviour, Proceedings of the 3rd ACM Symposium on Computing for Development, ACM (2013)

[5] Pushpendra Singh, Nikita Juneja, S. Kapoor, Using Mobile Phone Sensors to Detect Driving Behaviour, IEEE Transactions (2013)

[6] Fazeen, M., Gozick, B., Dantu, R., Bhukhiya, M., Gonzalez, M.C., Safe Driving Using Mobile Phones, IEEE Transactions on Intelligent Transportation Systems (2012)

[7] Chigurupa, S., Polavarapu, S., Kancherla,Y., Nikhath, K.A., Integrated Computing System for measuring Driver Safety Index, International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 2 (2012)

[8] Johnson, D.A., Trivedi, M.M., Driving Style Recognition using a Smartphone as a sensor platform, IEEE 14th International Conference on Intelligent Transportation system(2011)

[9] Amardeep Sathyanarayana, Pinar Boyraz, John H.L. Hansen, Driver behaviour analysis and route recognition by hidden Markov models, IEEE Transactions (2008)

[10] Hickman, J.S., Geller, E.S., Self-management to increase safe driving among short haul truck drivers, Journal of Organizational Behavior Management (2005)

[11] https://en.wikipedia.org/wiki/Genetic_algorithm

[12] Driving Conditions in India, http://www.indiandrivingschools.com/index.php

[13] http://www.ripublication.com/irph/ijeee_spl/ijeeev7n7_07.pdf

AUTHORS PROFILE

Meghana Prakash is currently pursuing Master of Technology degree program in Computer Science and Engineering, Siddaganga Institute of Technology (VTU, India).

Ravi V is an Assistant Professor in Department of Computer Science and Engineering, Siddaganga Institute of Technology (VTU, India).