Driver Skill Profiling using Machine Learning

Akhtar, Nadeem[1], Mohan, Mithun[2]

[1]Post Graduate Student, National Institute of Technology Karnataka
[2]Assistant Professor, National Institute of Technology Karnataka

Abstract. Road safety is a critical aspect of public safety, and driving skills are essential to ensuring safety on the road. An accurate understanding of one's driving abilities is crucial in promoting safe driving practices and reducing the risk of accidents. Overconfidence and underestimating road events can lead to a false sense of handling emergencies and may result in a higher risk of traffic offences and accidents. Young and novice drivers are particularly susceptible to these issues and may overestimate their abilities, leading to a higher risk tolerance. Machine learning is a viable approach that can compare perceived and actual skills to measure subjective driving skills accurately. A scoring system based on machine learning algorithms can quantify driver skills effectively and improve self-awareness, ultimately contributing to increased road safety. The proposed scoring system can give drivers an accurate assessment of their abilities, helping them take necessary corrective actions to work on their weaknesses. Driving style, encompassing violations, errors, and lapses, and driving skills, including perceptual motor skills and safety skills, are the two main components of the human factor in driving. Training sessions may be conducted based on the proposed scoring system using machine learning that can help improve drivers' self-awareness and reduce the risk of accidents.

Keywords: Road Safety, Driving Skills, Machine Learning.

1 Introduction

Road safety is a crucial element of public safety, and driving skills play a vital role in ensuring the safety of traffic operations on the road (Ahirwal et al., 2019). With the rising number of vehicles and drivers on the road, the likelihood of accidents and traffic offences also increases. Overconfidence and underestimating road events can create a false sense of handling emergencies, leading to a higher risk of traffic offences and accidents (Amado et al., 2014; Mohammadpour & Nassiri, 2021). Thus, it is essential to precisely understand one's driving abilities to promote safe driving practices and reduce accident risks.

Young and novice drivers are more susceptible to these problems and may overestimate their abilities, resulting in a higher risk tolerance (Martinussen et al., n.d.). Research shows that drivers aged 16-24 years are more likely to be involved in a car accident, with the highest rate of accidents occurring in the first year of driving (World

Health Organization, 2018). Therefore, it is crucial to improve young drivers' self-awareness and equip them with the necessary skills to handle potential road hazards.

Machine learning is a viable approach that can compare perceived and actual skills to accurately measure subjective driving skills (Malik et al., 2021). A scoring system based on machine learning algorithms can effectively quantify driver skills, improve self-awareness, and ultimately enhance road safety. The scoring system can give drivers an accurate assessment of their abilities, enabling them to take corrective action to address their weaknesses (Moharrer, 2011). The human factor in driving comprises driving style, including violations, errors, and lapses, and driving skills, including perceptual, motor and safety skills (Sundström, 2008).

Training sessions based on such scoring systems can improve drivers' self-awareness and reduce the risk of accidents (Xu et al., 2018). These training sessions can educate young and novice drivers on safe driving practices, road rules, and regulations. Additionally, the scoring system can identify areas where drivers need improvement and provide them with targeted feedback and training.

Thus, road safety is a critical aspect of public safety, and driving skills are vital for ensuring safety on the road. Accurate self-assessment of driving skills is crucial in promoting safe driving practices and reducing accident risks. This study proposes a scoring system for Indian drivers based on a machine learning algorithm which can effectively quantify driving skills, improve self-awareness, and enhance road safety. This system can be particularly useful for young and novice drivers who are more prone to overconfidence and underestimation of road events. Training sessions based on the proposed scoring system can help improve drivers' self-awareness and reduce the risk of accidents, making the roads safer for everyone.

This paper is organized into the following sections. The introduction section describes a brief outline of the need for the study and also summarizes excerpts from some of the previous studies. The section following this describes the procedure adopted in this work. The detailed procedure for data collection and observations drawn from it are respectively provided in sections 3 and 4. Section 5 explains the analysis performed, and the conclusions drawn from the study are summarized in section 6.

2 Methodology

The present research begins with the data collection based on a self-assessment questionnaire seeking inputs about the motor and safety skills of the respondents. The critical task that could differentiate drivers based on skill level was identified from these inputs. This was followed by field testing of the performance of drivers for this critical task. The performance of the drivers in this task was then analyzed, and by using a supervised machine-learning algorithm to predict the performance of new drivers. A scoring system was generated to score the new drivers based on the actual scores of the expert drivers. Further, an unsupervised machine learning approach was used to cluster the perceived data collected through the self-assessment questionnaire into different groups based on skill level. The actual scores and perceived scores of the four test drivers were compared to determine whether they overestimated or underestimated their skills.

3 Data Collection

3.1 Self-Assessment Questionnaire

The data collection process involved administering a self-assessment questionnaire to individuals living in Karnataka, both online and offline. The questionnaire was aimed to collect information about the participants' perceived driving skills and safety practices. It consisted of questions related to motor skills and safety skills, and participants were asked to rate themselves on a scale of 1 to 10, with 10 being the most expert driver. The questions presented to the participants are presented in Table 1. The questionnaire also included demographic questions, such as age, gender, level of education, job title, state of domicile, driving experience with cars and two-wheelers, and whether they possessed a driving license.

A total of 325 responses were collected, and only the responses from participants with driving license were used for further analysis, constituting 75% of the total responses. The data were collected through both online and offline modes to ensure maximum coverage of the target population. The data collection process aimed to provide a representative sample of individuals living in Karnataka. The collected data were then used to identify the most challenging driving skill and to develop a test to assess it on the ground.

Table 1. Self-Assessment Questionnaire

1	In the rainy situation, I can drive well on waterlogged roads
2	On turning left, I look for cyclist or a person who may come up on my side
3	I can slam on the brakes immediately to avoid a collision with front vehicle that stopped unexpectedly
4	I can judge the speed of oncoming vehicle when overtaking
5	I can perform reverse and parallel parking efficiently
6	I don’t lose my patience while driving behind a slow vehicle
7	I do not involve in unnecessary race competition with other vehicles
8	I can control my vehicle on slippery road without getting it skid
9	I am fluent in changing lanes in heavy traffic
10	I calmly tolerate other driver’s blunder in traffic

The data collected through the self-assessment questionnaire showed that the majority of the respondents were between the age group of 18-25 years (79%), followed by 26-35 years (15%) and 36-45 years (6%). In terms of gender, the majority were male (68%), and the remaining 32% were female. Regarding the level of education, only 2% of the respondents had completed their 12th grade, while 40% had completed their undergraduate degree. A significant percentage of respondents, 56%, had completed postgraduation or above. The remaining 1% fell under the "Other" category. In terms of driving experience with cars, 22% of the respondents had no experience, while 28% had less than 6 months of experience. A total of 9% of the respondents had 6-12 months of experience, while 10% had 1-2 years of experience. 18% of the respondents had 2-5 years of experience, 9% had 5-10 years of experience, and only 3% had more than 10 years of driving experience.

3.2 Field Data Collection

The responses from questionnaire were analyzed, and the average rating for each skill was calculated. For the field test, the lowest average rating of 6.46 was used, which was for the skill of efficiently performing reverse and parallel parking, which implies that this particular skill requires improvement among the respondents.

Parallel Parking and Reverse Parking Test

Set up before Field Test

Before conducting the field test for parallel and reverse parking, a pre-setup was carried out. As per the guidelines from the Indian Road Congress (IRC SP:12-2015), a rectangular parking lot was marked out using white marking powder on the college ground. The boundary of the lot (5.9 m in length and 2.5 m in width) was marked using lime and traffic cones. This is shown in Fig. 1. The car used for the experiment was a Maruti Suzuki Baleno, as shown in Fig. 2, which was used for all five days of the experiment. Piezoelectric sensors (shown in Fig. 3) were clamped to the clutch of the car and connected to the Arduino IDE through the ESP32 microcontroller using a laptop to measure the pressure applied to the clutch.

Fig. 1. Illustration of Parking Lot

Fig. 2. Trials conducted before experiment

During Field Test

During the experiment, acceleration values were measured using the 'Physics Toolbox' android application on a Samsung Galaxy S21 FE and a OnePlus 10T smartphone. In addition, the time taken to finish the parking maneuver was recorded, and a measuring tape was used to measure the offset of the parked car from the corner of the boundary to the wheels of the car. The angle of parking was then calculated using the offset data.

After Field Test

The data collected during the experiment were then analyzed using statistical software to determine the average time taken for each parking maneuver, the average angle of parking, and any other relevant metrics. Any anomalies or outliers in the data were identified and addressed, and the results were presented clearly and concisely in the final report. Overall, the pre-setup and post-setup steps were critical for ensuring the accuracy and reliability of the field test results.

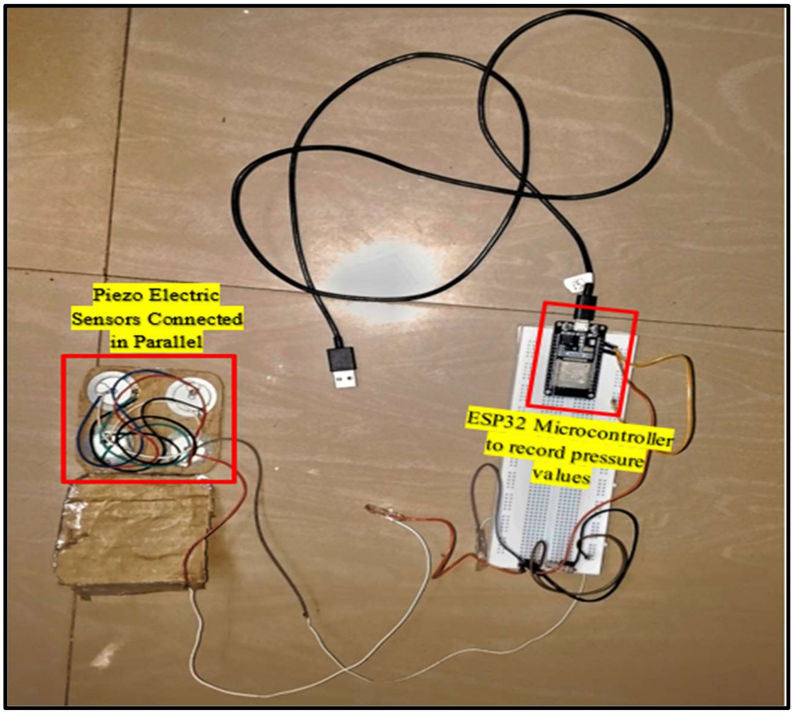

Fig. 3. Piezo Electric Sensor Setup

4 Observations

Table 2 represents the data collected during the parallel parking experiments. The experiments were conducted for 16 drivers, of which 12 were expert drivers. The remaining 4 were test drivers labelled A, B, C, and D. Similarly, Table 3 represents the data collected during the reverse parking experiments.

Table 2. Observations for Parallel Parking

Driver No.	Angle of Parallel Parking (degree)	Average of Offsets (m)	Time to finish the parallel parking (sec)	Average Acceleration during parallel parking (m/s^2)	Median Acceleration during parallel parking (m/s^2)	Std. Dev. Acceleration during parallel parking (m/s^2)	Average Pressure during parallel parking (Pa)	Minimum Pressure during parallel parking (Pa)	Maximum Pressure during parallel parking (Pa)	Std. Dev. Pressure during parallel parking (Pa)
1	0.36	1.63	60.00	0.36	0.33	0.22	75.87	4.00	268.91	68.43
2	9.68	1.60	50.00	0.80	0.78	0.34	62.31	8.47	180.98	53.44
3	1.29	1.75	65.45	0.73	0.53	0.59	58.41	0.31	227.29	43.76
4	9.70	1.65	67.00	0.52	0.44	0.54	74.20	2.67	252.33	65.16
5	8.23	1.60	66.00	0.69	0.63	0.34	77.19	4.49	187.18	41.53
6	9.99	1.75	44.96	0.51	0.50	0.51	63.63	8.47	180.98	43.76
7	1.29	1.78	39.76	0.92	0.78	0.46	40.55	5.74	145.76	29.93
8	4.17	1.60	44.00	0.74	0.62	0.44	45.93	4.10	151.94	33.49
9	3.54	1.80	36.00	0.78	0.46	0.85	34.57	0.31	247.37	47.50
10	9.13	1.73	44.94	0.75	0.55	0.52	55.33	3.72	188.20	37.31
11	0.64	1.73	58.42	0.78	0.69	0.56	70.38	8.47	221.89	59.97
12	0.97	1.65	49.94	0.67	0.62	0.43	63.29	2.16	180.98	55.03
A	9.99	1.88	80.25	0.22	0.62	0.86	65.67	0.31	267.29	46.43
B	21.65	1.65	90.00	0.27	0.21	0.24	78.32	1.07	258.37	59.58
C	9.13	1.65	76.00	0.52	0.45	0.42	60.49	1.48	232.04	54.35
D	12.47	1.63	127.00	0.74	0.64	0.55	70.32	0.31	267.29	49.77

Table 3. Observation Table for Reverse Parking

Driver No.	Angle of Reverse Parking	Average of Offsets	Time to finish the Reverse parking (sec)	Average Acceleration during Reverse parking (m/s^2)	Median Acceleration during Reverse parking (m/s^2)	Std. Dev. Acceleration during Reverse parking (m/s^2)	Average Pressure during Reverse parking (Pa)	Minimum Pressure during Reverse parking (Pa)	Maximum Pressure during Reverse parking (Pa)	Std. Dev. Pressure during Reverse parking (Pa)
1	4.89	1.68	40.00	0.51	0.39	0.38	73.54	1.17	293.52	52.17
2	8.99	1.68	26.23	0.62	0.78	0.34	63.44	4.84	185.39	46.18
3	8.04	1.65	35.67	0.80	0.53	0.59	67.35	17.82	163.81	37.29
4	6.04	1.65	18.24	0.83	0.44	0.54	44.79	0.44	141.57	44.50
5	2.49	1.58	42.90	0.66	0.63	0.34	84.43	2.24	210.20	52.76
6	0.94	1.60	28.76	4.84	0.50	0.51	58.41	3.46	174.20	50.19
7	2.57	1.63	22.00	0.42	0.78	0.46	39.54	6.09	129.05	36.74
8	3.53	1.60	21.19	0.81	0.62	0.44	38.04	0.47	138.94	32.16
9	7.40	1.68	24.00	0.65	0.46	0.85	38.16	1.70	128.27	35.09
10	4.81	1.75	23.58	0.80	0.55	0.52	40.11	1.63	101.67	26.78
11	3.85	1.68	44.95	0.85	0.69	0.56	51.23	0.01	185.39	44.12
12	11.62	1.70	44.79	0.54	0.62	0.43	50.57	0.08	174.20	45.22
A	5.49	1.60	45.00	0.62	0.62	1.86	52.73	0.01	185.39	47.11
B	2.55	1.60	32.80	0.39	0.35	0.28	73.04	1.17	224.42	49.12
C	3.00	1.70	23.24	0.47	0.45	0.42	45.26	5.55	117.88	33.81
D	2.65	1.68	40.45	0.55	0.64	0.55	87.05	39.37	228.05	38.59

5 Analysis and Result

In order to investigate the relationships between the target variables measuring drivers' skill (angle of parking, average of offsets, and time to finish parking) and other parameters (average acceleration, median acceleration, standard deviation of acceleration, average pressure, minimum pressure, maximum pressure), a Pearson correlation analysis was conducted. The analysis included data from both parallel and reverse parking experiments conducted by the expert drivers (Driver No. 1-12) and the test drivers (A, B, C, D). Tables 4 and 5 display the Pearson correlation coefficients for parameters in parallel parking and reverse parking, respectively, providing valuable insights into the relationships between variables and parking outcomes.

Table 4 presents a comprehensive overview of the correlation coefficients observed in parallel parking. There is a moderate negative correlation (-0.444) between average acceleration and the time taken to complete parallel parking, suggesting that higher average acceleration may lead to faster completion times. The correlation coefficient of 0.324 indicates a weak positive relationship between average acceleration and the final offset, implying that higher acceleration levels could result in larger deviations from the desired parking position. Additionally, the average pressure during parallel parking shows a strong positive correlation (0.833) with the average pressure recorded, indicating consistent pressure application throughout the maneuver.

Table 5 provides insights into the correlations observed in reverse parking. There is a weak negative correlation (-0.086) between average acceleration and the time taken to complete reverse parking, indicating that higher levels of acceleration do not necessarily lead to quicker parking times. The weak negative correlation of -0.464 suggests that average acceleration has minimal impact on the final angle of parking. Notably, the average pressure exerted during reverse parking demonstrates a moderate positive correlation (0.636) with the maximum pressure recorded, suggesting that higher average pressure tends to coincide with higher maximum pressure values.

These correlation coefficients provide valuable insights into the dynamics of parallel parking and reverse parking. They highlight the relationships between acceleration, pressure, and various parking outcomes, informing future research and potential strategies for enhancing parking skills.

Table 4. Pearson Correlation Coefficients For Parallel Parking

	Time	**Final Angle**	**Final Offset**
Average Acceleration	-0.444	-0.130	0.324
Median Acceleration	-0.243	0.005	-0.062
Standard Deviation Acceleration	-0.363	-0.042	0.754
Average Pressure	0.833	0.203	-0.558
Minimum Pressure	-0.094	0.292	-0.160
Maximum Pressure	0.473	-0.164	0.089
Standard Deviation Pressure	0.530	-0.108	-0.293

Table 5. Pearson Correlation Coefficients For Reverse Parking

	Time	Final Angle	Final Offset
Average Acceleration	-0.086	-0.464	-0.342
Median Acceleration	0.044	0.004	-0.116
Standard Deviation Acceleration	-0.272	0.119	0.239
Average Pressure	0.636	-0.047	-0.322
Minimum Pressure	0.002	0.152	-0.122
Maximum Pressure	0.655	-0.026	-0.165
Standard Deviation Pressure	0.557	-0.084	-0.387

5.1 Supervised Machine Learning

Random Forest is a popular machine learning algorithm that is widely used for classification and regression tasks. One of the main advantages of using Random Forest is that it can handle a large number of input variables and can be used for feature selection. Another advantage is that it is able to handle both linear and non-linear relationships between the input variables and the target variable. Additionally, Random Forest can help to reduce the risk of overfitting, which can be a common problem in machine learning. Therefore, this algorithm was chosen for analysis as it can handle the complexity of the dataset, help with feature selection, and reduce the risk of overfitting.

Scoring System

The scoring system was developed for the drivers using the predicted values obtained through the Random Forest machine learning algorithm and the observed values from on-ground tests. The scoring system was designed as follows:

Case 1: If the predicted value was less than the observed value, then the score was bound between 0 and 80 and was given by Eq. 1.

$$\text{Score} = 80 - \frac{Observed\ value - Predicted\ Value}{Observed\ Value} * 80 \tag{1}$$

Case 2: If the predicted value was more than the observed value, 1 point was given as a bonus for every 5% deviation from the predicted value, and Eq. 2 gave the score.

$$\text{Score} = 80 + \frac{Predicted\ Value - Observed\ value}{Predicted\ Val} * 100 * \frac{1}{5} \tag{2}$$

Parallel Parking

Random forest algorithm was used to predict the target variable values for test drivers A, B, C, and D. For parallel parking, two cases were considered, with two subcases each. In the first case, the first subcase involved all acceleration parameters as predictors, and the target variable was time to finish the parking. The second subcase included acceleration parameters as predictors, and the target variable was the average of offset. Angle was not included as a predictor variable in this subcase because the Pearson coefficient revealed that offset had a stronger correlation with acceleration.

In the second case, the first subcase considered all pressure parameters as predictors, and the target variable was time to finish the parking. The second subcase for this involved pressure parameters in combination with angle and offset as predictor variables.

Table 6. Final Score of drivers for Parallel Parking

Predictors			**Acceleration**
Target Variable			**Time to Finish Parking**
D. No.	**Observed Time**	**Predicted Time**	**Score**
A	80.25	52.32	52.16
B	90	57.89	51.46
C	76	54.16	57.01
D	127	54.24	34.17
Predictors			**Acceleration**
Target Variable			**Offset**
D. No.	**Observed Offset**	**Predicted Offset**	**Score**
A	1.88	1.74	74.04
B	1.65	1.63	79.03
C	1.65	1.64	79.52
D	1.62	1.72	81.16
Predictors			**Pressure**
Target Variable			**Time**
D. No.	**Observed Time**	**Predicted Time**	**Score**
A	80.25	59.61	59.42
B	90	64.26	57.12
C	76	59.65	62.79
D	127	63.11	39.75

Predictors			Pressure
Target Variable			**Angle + Offset**
D. No.	**Observed Result Value**	**Predicted Result Value**	**Score**
A	1.25	0.81	51.79
B	1.25	0.88	56.26
C	0.63	1.07	88.25
D	1.27	0.89	55.85

Reverse Parking

The same approach was used for reverse parking, with two cases considered, and two subcases in each. In the first case, the first subcase involved all acceleration parameters as predictors, and the target variable was time to finish the parking. The second subcase included acceleration parameters as predictors, and the target variable was a combination of offset and angle. This was because the Pearson coefficient revealed that offset and angle had a stronger correlation with acceleration.

In the second case, the first subcase included all pressure parameters as predictors, and the target variable was time to finish the parking, while the second subcase involved pressure parameters and offset as predictor variables.

Table 7. Final Score of drivers for Reverse Parking

Predictors			Acceleration
Target Variable			**Time to Finish Parking**
D. No.	**Observed Time**	**Predicted Time**	**Score**
A	45.00	31.36	55.75
B	32.80	37.63	82.57
C	23.24	35.31	86.84
D	40.45	32.95	65.17
Predictors			**Acceleration**
Target Variable			**Angle + Offset**
D. No.	**Observed Result Value**	**Predicted Result Value**	**Score**
A	0.64	0.98	86.85
B	1.39	1.11	63.80
C	0.25	1.26	96.05
D	0.22	0.37	88.02

Predictors			Pressure
Target Variable			**Time**
D. No.	**Observed Time**	**Predicted Time**	**Score**
A	45.00	42.94	76.34
B	32.80	40.54	83.82
C	23.24	23.09	79.48
D	40.45	34.52	68.27
Predictors			**Acceleration**
Target Variable			**Offsets**
D. No.	**Observed Offsets**	**Predicted Offsets**	**Score**
A	1.6	1.68	80.95
B	1.6	1.64	80.49
C	1.7	1.65	77.65
D	1.675	1.64	78.33

5.2 Unsupervised Machine Learning

Fuzzy C-means is an unsupervised machine learning algorithm adopted for clustering analysis. It is a clustering algorithm that aims to assign each data point to a certain cluster, where the cluster is defined by its centroid. Fuzzy C-means is a soft clustering algorithm, meaning that each data point is assigned to a certain cluster with a degree of membership rather than being assigned to only one cluster. The degree of membership is determined by the proximity of the data point to the centroid of the cluster. In this way, Fuzzy C-mean provides a more flexible approach to clustering analysis than traditional hard clustering algorithms.

Unsupervised machine learning was employed in this study, specifically Fuzzy C means clustering, to cluster participants based on their responses to a self-assessment questionnaire regarding their driving skills. Test drivers were also asked to complete the same questionnaire, allowing us to determine the clusters they belonged to. Three clusters were identified based on skill level, as shown in Fig.4, with Cluster 1 consisting of high-skilled drivers, Cluster 3 consisting of average-skilled drivers, and Cluster 2 consisting of low-skilled drivers. The skill level for each cluster was determined by calculating the centroid value of each cluster. All test drivers were then assigned to their respective clusters.

To assess the parking skills of the test drivers, the mean value of the parking skills for each cluster was compared to the perceived values of the driver's response, as shown in Table 8. The same scoring approach as before was employed but with one modification. The scoring system would be ineffective if the test drivers belonged to different clusters. To address this, weightages were assigned to each cluster. Cluster 1 was assigned a weightage of 1, Cluster 3 was assigned a weightage of 0.8, and Cluster 2 was assigned a weightage of 0.5. The overall score of each test driver was then multiplied by their respective cluster weightage to allow for comparison across all drivers. Eq. 3 was used for computing the weighted score.

$$\text{Modified Score} = \text{Score} * \text{Cluster Weightage} \tag{3}$$

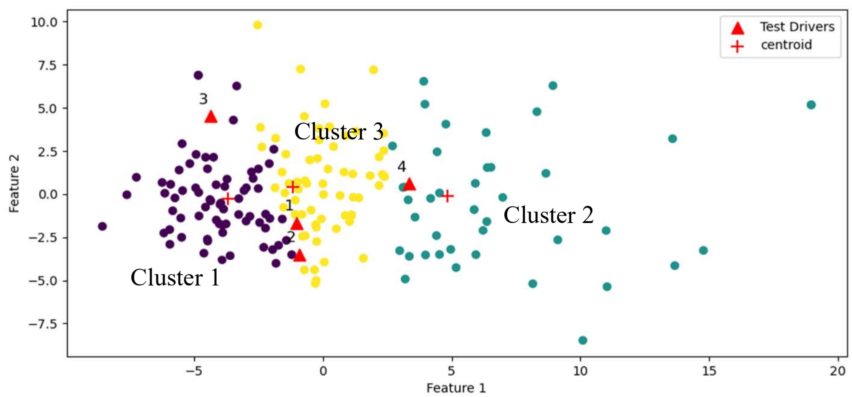

Fig. 2. Fuzzy C Means Clustering with 3 Clusters

Table 8. Centroid Feature Values of each Cluster and Test Driver's response

	Cluster 1	Cluster 2	Cluster 3	Driver 1	Driver 2	Driver 3	Driver 4
In the rainy situation, I can drive well on waterlogged roads	**7.81**	**4.75**	**6.99**	**7**	**8**	**5**	**7**
On turning left, I look for cyclist or a person who may come up on my side	**8.77**	**6.25**	**8.41**	**8**	**10**	**8**	**9**
I can slam on the brakes immediately to avoid a collision with front vehicle that stopped unexpectedly	**8.51**	**5.77**	**7.85**	**9**	**10**	**7**	**8**
I can judge the speed of oncoming vehicle when overtaking	**8.60**	**5.85**	**7.89**	**9**	**10**	**7**	**7**
I can perform reverse and parallel parking efficiently	**7.78**	**4.19**	**6.55**	**7**	**8**	**9**	**6**

I don't lose my patience while driving behind a slow vehicle	**7.88**	**5.59**	**6.84**	**5**	**10**	**5**	**7**
I do not involve in unnecessary race competition with other vehicles	**8.60**	**7.36**	**8.06**	**6**	**10**	**8**	**9**
I can control my vehicle on slippery road without getting it skid	**8.10**	**5.26**	**7.22**	**6**	**8**	**6**	**6**
I am fluent in changing lanes in heavy traffic	**8.21**	**4.84**	**7.21**	**10**	**9**	**5**	**7**
I calmly tolerate other driver's blunder in traffic	**7.40**	**5.60**	**6.68**	**8**	**10**	**7**	**8**

A combination of supervised and unsupervised machine learning techniques was used to evaluate the drivers' parking skills. The actual score was obtained by taking the average of all the scores predicted through supervised machine learning methods. In contrast, the perceived score was determined using unsupervised Fuzzy C means clustering to group drivers based on their self-assessment questionnaire responses.

Table 9 compares the perceived and actual scores for four test drivers: A, B, C, and D. Drivers A and B were found to have slightly underestimated their parking skills, as their actual scores were higher than their perceived scores. Conversely, Driver C overestimated their skills by a significant margin, with an actual score of 78.5 compared to a perceived score of 91. Driver D also performed poorly, with an actual score of 64 compared to a perceived score of 44. These results suggest that the combination of supervised and unsupervised machine learning methods can accurately evaluate drivers' parking skills while also identifying discrepancies between self-assessment and actual performance.

Table 9. Comparison of Perceived and Actual Driving Skill

	Perceived Score	**Actual Score**
Driver A	65	67
Driver B	67	69.5
Driver C	91	78.5
Driver D	44	64

6 Conclusion and Recommendation

In this research study, an evaluation of challenging driving skills identified from the questionnaire, parallel parking and reverse parking, was conducted among test drivers using various machine learning techniques. The aim was to assess and score the drivers based on their performance in parking scenarios.

To achieve this, supervised machine learning, specifically the Random Forest Algorithm, was employed to predict target variable values for the test drivers. This allowed for a comparison between the predicted values and the observed values obtained from on-ground tests. A scoring system was then developed to quantify the level of skill exhibited by each driver, considering both overestimation and underestimation of their abilities. Additionally, unsupervised machine learning in the form of the Fuzzy C Means clustering algorithm was utilized. helped group the drivers based on their responses to a self-assessment questionnaire, allowing for the identification of distinct skill levels.

Based on the results of our study, we recommend that further analysis should be done to incorporate additional skills from the self-assessment questionnaire to provide a more comprehensive evaluation of a driver's overall skill level. While this study focused specifically on parking skills, there are other skills that are likely to be important for safe and effective driving. Expanding the analysis's scope to include a wider range of skills can give a more accurate and holistic understanding of a driver's abilities. This information can be useful for identifying areas where drivers may need additional training or support to improve their skills and ultimately reduce the risk of accidents on the road.

References

1. Ahirwal, S., Gupta, A., & Gupta, S. (n.d.). The Indian Journal of Occupational Therapy Assessment of driving skills of indian adults on a driving simulator: A pilot study. https://www.ijotonweb.org/printarticle.asp?issn=0445-7706;year=2019;volume=51;issue=4;spage=125;epage=129;aulast=Ahirwal
2. Amado, S., Arikan, E., Kaça, G., Koyuncu, M., & Turkan, B. N. (2014). How accurately do drivers evaluate their own driving behavior? An on-road observational study. Accident Analysis and Prevention, 63, 65–73. https://doi.org/10.1016/j.aap.2013.10.022
3. Best, M. C. (2019). Real-time characterisation of driver steering behaviour. Vehicle System Dynamics, 57(1), 64–85. https://doi.org/10.1080/00423114.2018.1447678
4. Castignani, G., Derrmann, T., Frank, R., & Engel, T. (2015). Driver behavior profiling using smartphones: A low-cost platform for driver monitoring. IEEE Intelligent Transportation Systems Magazine, 7(1), 91–102. https://doi.org/10.1109/MITS.2014.2328673
5. Castignani, G., & Frank, R. (2014). SenseFleet: A smartphone-based driver profiling platform. 2014 11th Annual IEEE International Conference on Sensing, Communication, and Networking, SECON 2014, 144–145. https://doi.org/10.1109/SAHCN.2014.6990337
6. Eren, H., Makinist, S., Akin, E., & Yilmaz, A. (2012). Estimating driving behavior by a smartphone. IEEE Intelligent Vehicles Symposium, Proceedings, 234–239. https://doi.org/10.1109/IVS.2012.6232298
7. Malik, M., Nandal, R., Dalal, S., Jalglan, V., & Le, D. N. (2021). Driving pattern profiling and classification using deep learning. Intelligent Automation and Soft Computing, 28(3), 887–906. https://doi.org/10.32604/iasc.2021.016272
8. Martinussen, L. M., Møller, M., & Prato, C. G. (n.d.). 6th Road Safety on Four Continents Conference Driver Style And Driver Skill-Clustering Sub-Groups Of Drivers Differing In Their Potential Danger In Traffic.
9. Mohammadpour, S. I., & Nassiri, H. (2021). Aggressive driving: Do driving overconfidence and aggressive thoughts behind the wheel, drive professionals off the road? Transportation Research Part F: Traffic Psychology and Behaviour, 79, 170–184. https://doi.org/10.1016/j.trf.2021.04.008
10. Moharrer, M. (2011). Actual Skill Vs. Perceived Skill; A New Method For Assessing Overconfidence Among Drivers.
11. Sundström, A. (2008). Self-assessment of driving skill - A review from a measurement perspective. Transportation Research Part F: Traffic Psychology and Behaviour, 11(1), 1–9. https://doi.org/10.1016/j.trf.2007.05.002
12. Xu, J., Liu, J., Sun, X., Zhang, K., Qu, W., & Ge, Y. (2018). The relationship between driving skill and driving behavior: Psychometric adaptation of the Driver Skill Inventory in China. Accident Analysis and Prevention, 120, 92–100. https://doi.org/10.1016/j.aap.2018.07.032
13. Zhang, Y., Lin, W. C., & Chin, Y. K. S. (2010). A pattern-recognition approach for driving skill characterization. IEEE Transactions on Intelligent Transportation Systems, 11(4), 905–916. https://doi.org/10.1109/TITS.2010.2055239