Behavioral Analysis and Risk Assessment of Two-Wheeler Drivers

Submitted in partial fulfillment of the requirements of the degree

BACHELOR OF ENGINEERING IN COMPUTER ENGINEERING

By

Simran Ahuja 02

Jesica Bijju 10

Sejal Datir 14

Sania Khan 36

Name of the Mentor

Prof. Dr. Mrs. Nupur Giri

Since 1962

Vivekanand Education Society's Institute of Technology,

An Autonomous Institute affiliated to University of Mumbai

HAMC, Collector's Colony, Chembur,

Mumbai-400074

University of Mumbai (AY 2023-24)

CERTIFICATE

This is to certify that the Mini Project entitled **“ Behavioral Analysis and Risk Assessment of Two-Wheeler Drivers”** is a bonafide work of **Simran Ahuja(02), Jesica Bijju(10), Sejal Datir(14) and Sania Khan(36)** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **“Bachelor of Engineering”** in **“Computer Engineering” .**

(Prof.__________________)

Mentor

(Prof.__________________)

Head of Department

(Prof.__________________)

Principal

Mini Project Approval

This Mini Project entitled "Behavioral Analysis and Risk Assessment of Two-Wheeler Drivers " by **Simran Ahuja(02), Jesica Bijju(10), Sejal Datir(14) and Sania Khan(36)** is approved for the degree of **Bachelor of Engineering** in **Computer Engineering.**

Examiners

1………………………………………

(Internal Examiner Name & Sign)

2………………………………………

(External Examiner name & Sign)

Date:

Place:

Contents

Abstract

With the growing popularity of two-wheelers in urban and rural areas, it is imperative to enhance road safety and reduce the alarming rate of accidents involving two-wheeler riders. This study combines behavioral analysis, data collection, and risk assessment to provide insights that can be leveraged for developing effective road safety strategies.The "Behavioral Analysis of Two-Wheeler Drivers" project aims understand the behavior of driving patterns of individuals operating two-wheeled vehicles, such as motorcycles and scooters.

It involves the collection of extensive time - series data related to two wheelers followed by behavioral analysis to gain insights into the decision-making processes and habits of two-wheeler drivers. It considers parameters like acceleration, orientation, magnetometer and gyroscope readings, motion pitch, etc. to build a robust machine learning model.

Further, analyzing the collected data, patterns and correlations will be identified to establish a risk profile for different types of two-wheeler riders. This risk assessment will help in developing targeted intervention strategies to improve road safety.

The ultimate goal of this project is to provide valuable insights that can be used by policymakers, traffic authorities, and safety advocates to develop and implement evidence-based strategies for reducing the number of accidents involving two-wheeler riders. The results of this study can potentially lead to the formulation of more effective road safety campaigns, improved traffic regulations, and the promotion of safe driving behaviors among two-wheeler users.

Acknowledgments

We extend our heartfelt gratitude to Vivekanand Education Society's Institute of Technology for their unwavering support and assistance throughout our project. In particular, we are deeply appreciative of Professor Dr. Mrs. Nupur Giri for her invaluable guidance and advice during the project's inception and development.

We are indebted to Dr. (Mrs.) Nupur Giri, the Head of the Computer Department, and Dr. (Mrs.) J.M. Nair, our Principal, for granting us this valuable opportunity to undertake this project. Their trust in us has been instrumental in our success.

Our sincere thanks go to all the teaching and non-teaching staff whose constant encouragement and selfless assistance were instrumental in completing both the project synopsis and the project review successfully. We also acknowledge the Department of Computer Engineering for their support and guidance.

We are profoundly thankful to everyone who contributed to our project by providing information and resources. Our families also deserve recognition for their unwavering moral support and encouragement throughout this journey.

Chapter 1 : Introduction

1.1 Introduction

Alarming statistics reveal that two-wheeler riders account for a significant 38% of road traffic fatalities and 51% of nonfatal road traffic injuries in the nation, and 75% of fatalities and 82% of non-fatal injuries occur in the age group between 15 and 44 years, while 50% of the accidents involve commuters between 20 and 30 years of age.

These concerning figures highlight the urgent need to address the road safety challenges faced by two-wheeler riders. To confront this pressing public health issue and promote safer roads, we embark on a comprehensive research project. Our project seeks to address this pressing issue by delving into the behavior analysis of driving patterns. It is widely acknowledged that the behavior of road users significantly impacts their safety and the safety of others on the road. By closely examining the actions and habits of two-wheeler drivers, we can better comprehend the underlying causes of accidents and develop strategies to mitigate the associated risks.Alarming statistics reveal that two-wheeler riders account for a significant 38% of road traffic fatalities and 51% of nonfatal road traffic injuries in the nation, and 75% of fatalities and 82% of non-fatal injuries occur in the age group between 15 and 44 years, while 50% of the accidents involve commuters between 20 and 30 years of age.

These concerning figures highlight the urgent need to address the road safety challenges faced by two-wheeler riders. To confront this pressing public health issue and promote safer roads, we embark on a comprehensive research project. Our project seeks to address this pressing issue by delving into the behavior analysis of driving patterns. It is widely acknowledged that the behavior of road users significantly impacts their safety and the safety of others on the road. By closely examining the actions and habits of two-wheeler drivers, we can better comprehend the underlying causes of accidents and develop strategies to mitigate the associated risks.

1.2 Motivation

With the growth of various start-ups promising 10 minute deliveries the strain of these targets is obviously borne by the delivery partners, who often drive hastily to meet the demands. The primary transportation used by these delivery partners is 2 wheelers. In 2019, the Mumbai Traffic Police had to summon the executives of delivery partners like Swiggy, Zomato, Uber-Eats to address the issue of rash and negligent driving by the delivery drivers who are perpetually in a tearing hurry to meet their targets. Moreover, maintenance of such huge fleets of vehicles is also a challenge. Thus, to address these concerns behavior analysis of these driving patterns is the need of the hour with primary focus on 2 wheelers.

1.3 Problem Statement

The project aims to develop a comprehensive behavioral analysis and risk assessment AI-model using time-series data for 2-wheeler motorbikes that would incorporate parameters to identify abnormal driving behaviors. Hence, help improve road safety, and create a safer environment for all road users in India.

Chapter 2 : Literature Survey

SR NO.	Title Of The Paper	Abstract
1	Driving-Pattern Identification and Event Detection Based on an Unsupervised Learning Framework: Case of a Motorcycle-Riding Simulator	This paper suggests a method to study how people drive in a detailed way. They have a process with several steps. The main step uses computer programs to figure out how people usually drive and to find unusual or risky driving moments. When these risky moments are found, they are analyzed further by looking at the most important aspects using graphs.
2	(PDF) Detection of Two wheeler Driver Safety Using Machine Learning	In this paper a method for spotting people who break the law by not wearing a helmet while riding a bike. This new tool can also help the police find these rule-breakers even in tough weather conditions. These experiments show how well it can recognize the bikers and identify the violations. A system was also proposed that could adapt to different situations with a few adjustments if needed.
3	Safety of motorised two wheelers in mixed traffic conditions: Literature review of risk factors - ScienceDirect	This research looks at other studies that have examined the things that make motorized two-wheelers (MTWs) less safe, especially when they're riding in diverse traffic without clear lanes. The goal of this paper is to gather the findings from these studies and emphasize the latest insights. The paper also talks about what we still don't know and need to research more about to improve the safety of MTWs in mixed traffic.
4	D 3 : Abnormal driving behaviors detection and identification using smartphone sensors	This paper focuses on enhancing driving safety by monitoring and identifying specific abnormal driving behaviors using smartphone sensors. It can pinpoint actions like weaving, swerving, and sudden braking. The system, called "D3,"

		extracts unique patterns from smartphone sensor data and employs machine learning to identify these behaviors with an impressive 95.36% accuracy, as shown in real-world tests. It aims to improve drivers' awareness and help prevent accidents.
5	https://www.academia.edu/23985750/Detecting_Powered_Two_Wheeler_incidents_from_high_resolution_naturalistic_data	This paper focuses on how motorcycle riders (Powered-Two-Wheeler or PTW drivers) change their behavior during risky situations. The paper suggests a method to identify these changes by looking at detailed driving data. It finds irregular behavior by spotting values that stand out in the data. These irregularities can be seen as critical moments (incidents) that relate to common driving situations.
6	Driver Behavior Classification System Analysis Using Machine Learning Methods	This paper is used to classify driver behavior based on data from lane detection and traffic conditions. It tested various models and found that gradient boosting works best. It also analyzes important features for classification and suggests future improvements, considering factors like speed limits and mental workload, and a hybrid classification system.
7	Real-time detection of abnormal driving behavior based on long short-term memory network and regression residuals - ScienceDirect	This paper introduces LSTM-R, a real-time algorithm for detecting abnormal driving behavior. It's tested using smartphone-collected vehicle data and outperforms other methods, even with limited abnormal driving data. This approach is cost-effective and can improve road safety by assessing driving risk and behavior.
8	CNN-LSTM Driving Style Classification Model Based on Driver Operation Time Series Data	This paper presents a method for classifying driving styles using data from a simulator environment. A CNN+LSTM network is trained to detect driving styles, which shows high generalization and cost-efficiency when tested with real car data. The system can signal drivers and surrounding vehicles to improve driving route planning.

2.1 Survey of Existing System/SRS

In the landscape of driver assistance systems, a noticeable gap exists when it comes to providing comprehensive and proactive solutions tailored specifically for two-wheeler riders. Many existing systems are primarily designed for cars, leaving a void in holistic support for two-wheeler safety. These systems, while proficient in their respective tasks, often function reactively rather than proactively. Instead of assisting riders in avoiding potential risks, they typically alert them after detecting hazards. Furthermore, these systems may offer features such as navigation or vehicle diagnostics, but they often lack the integration of multimodal datasets, which could combine camera data with information from other sensors to enhance safety.

AI technology, although prominent in contemporary systems, is not fully harnessed for two-wheeler applications, especially concerning comprehensive diagnostics, risk assessment, and behavioral analysis. Furthermore, these systems rarely provide specific alerts tailored to two-wheelers regarding hazardous road conditions, such as wet or slippery surfaces. Unlike car-centric solutions, which include advanced fleet management tools, two-wheeler-specific solutions are less developed and may lack advanced functionalities. Maintenance tracking, typically offered as a standalone feature, is often disconnected from real-time diagnostics and AI-driven risk analysis.

Chapter 3: Proposed System

3.1 Introduction

By monitoring the rider's driving patterns and comparing them to a database of safe driving behaviors, the system can assess the risk associated with the rider's style. Over time, it can identify tendencies that might lead to accidents or violations and offer corrective feedback. For example, if a rider frequently brakes hard or takes turns aggressively, the system might provide tips or tutorials to encourage safer habits.

3.2 Architectural Framework / Conceptual Design

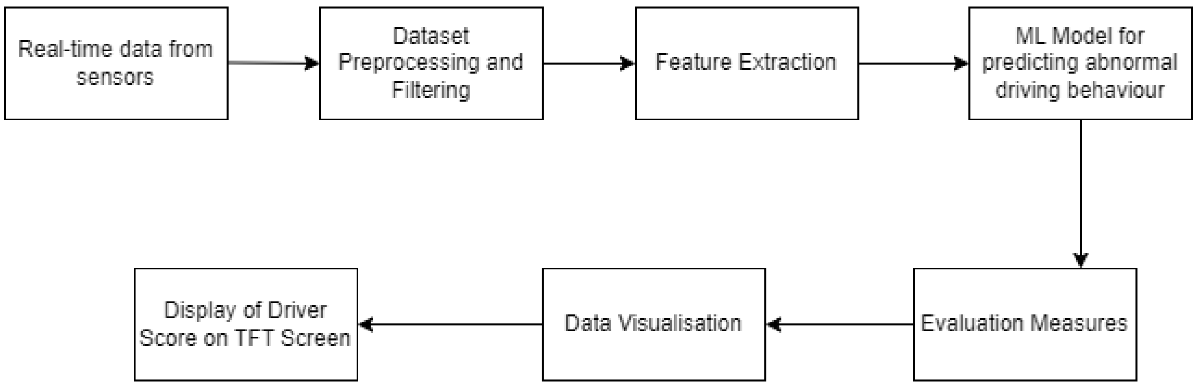

Fig. 1: Architectural framework

3.4 Methodology Applied

The aim of this project is to analyze the behavior of two-wheeler riders and make predictions regarding the normalcy or abnormality of their driving patterns. This section outlines the methodologies and procedures employed to achieve this objective.

Data Collection:

1. Data Source: The data for this study was collected from riders navigating a predefined route -"Bhakti Bhavan route."

2. Data Recording: A variety of sensors were employed to capture the rider's behavior. These included:

- Accelerometer (x, y, z)

- Gyrometer (x, y, z)
- Magnetometer (x, y, z)
- Location Speed
- Motion Pitch
- Motion Yaw

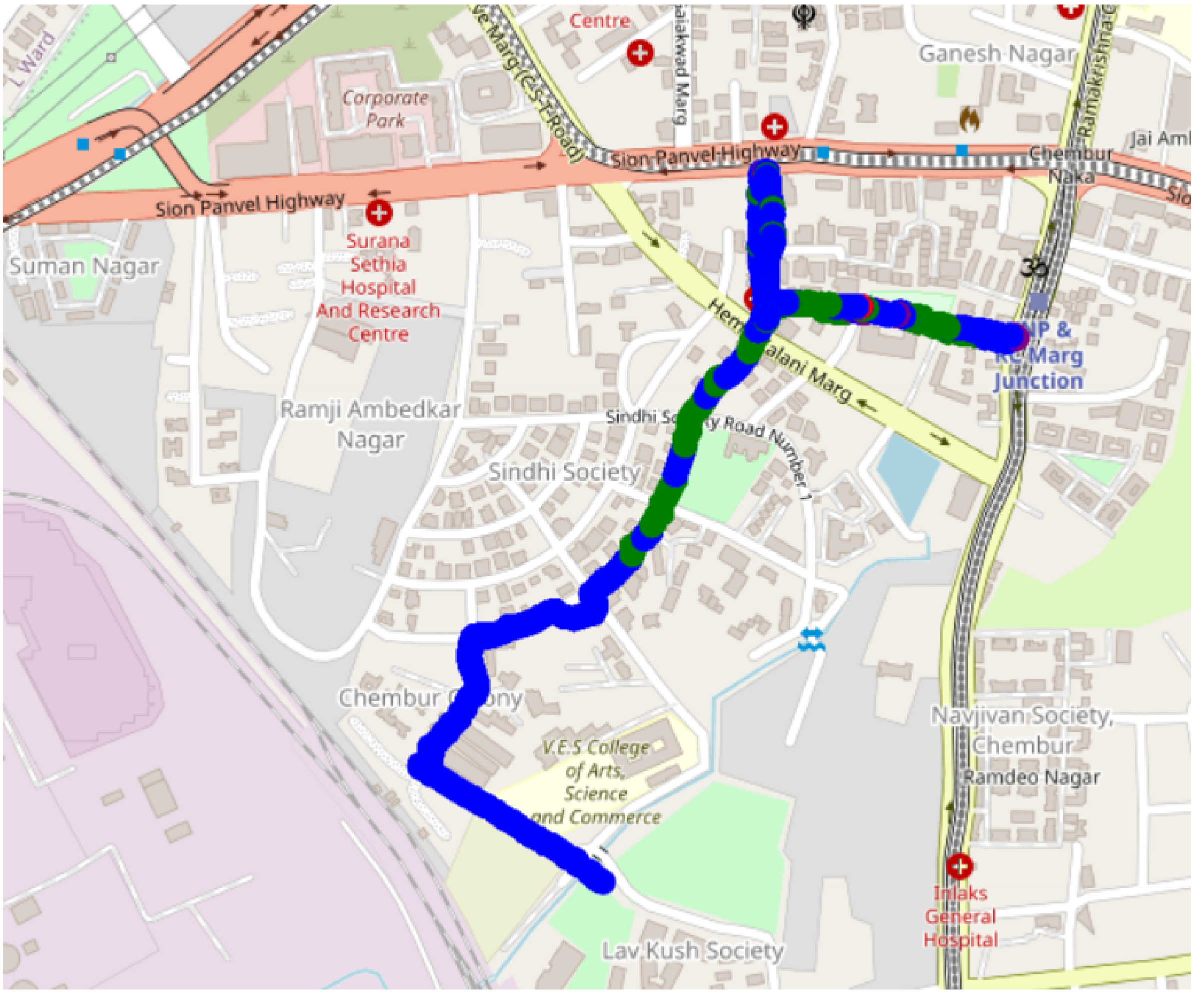

fig. 2: Test Route (blue: normal driving, green: sharp acceleration, red: sharp braking, purple: sharp turning)

3. Labeling: To categorize the data, we introduced four labels:
- 0: Normal Driving
- 1: Sharp Acceleration
- 2: Sharp Braking
- 3: Sharp Turning

These labels were manually assigned by an observer seated behind the rider using the SensorLog iOS app. This labeling process aimed to differentiate between normal and abnormal driving patterns.

Data Preprocessing:

1. Feature Selection: Data preprocessing began with the identification of relevant features. We pruned the parameters by examining correlations between the features to eliminate redundancies and enhance model efficiency.

2. Data Splitting: The dataset was divided into training and testing subsets to facilitate model training and evaluation.

Model Selection:

We experimented with a variety of machine learning and deep learning models to ascertain the most effective approach for predicting driving behavior. The models considered in this study were:

1. **Long Short-Term Memory (LSTM-R)**
We trained an LSTM model on the time series data, considering each parameter separately.
For each parameter (e.g., accelerometer, gyrometer, magnetometer), we performed LSTM predictions.
We calculated the residuals using the formula: residuals = np.sqrt(y_test - y_pred.flatten()).
Note: In the future, we plan to use these residuals for further classification.

2. **Convolutional Neural Network (CNN)**
Convolutional Neural Networks (CNNs) were leveraged as a powerful tool for analyzing and classifying two-wheeler driving behavior based on a dataset containing sensor data. CNNs, typically known for their effectiveness in image recognition, were adapted to analyze sequences of sensor readings. The model architecture included convolutional and pooling layers to capture spatial patterns within the sensor data, followed by fully connected layers for classification. The CNN model was trained and tested on a dataset that included parameters such as speed, accelerometer data, gyroscope readings, and magnetometer data,

along with corresponding driving behavior labels. The results demonstrated a high accuracy of 79% in classifying different driving behaviors, which showcased the adaptability of CNNs in the context of multi-class classification for two-wheeler driving. CNNs, with their ability to extract meaningful features from sequential data, offered a valuable approach to behavioral analysis and risk assessment in this project.

3. Random Forest

An ensemble learning technique known for its robustness and ability to handle a variety of data types. We leveraged the power of the Random Forest ensemble learning technique to build a classification model. We used it to predict a target variable (label) based on a set of input features from a real-world dataset. After training the Random Forest classifier on the training data, we evaluated its performance on a separate test set. The model achieved impressive results, with an accuracy of 89%, precision of 88%, recall of 89%, and an F1 score of 88%. These metrics showcase the Random Forest's robustness and effectiveness in handling a variety of data types.

4. XG Boost

We employed the versatile ensemble learning technique, XGBoost, renowned for its robustness and adaptability to diverse data types. Our objective was to construct a classification model capable of predicting a target variable (label) based on a set of input features drawn from a real-world dataset. Following the training of the XGBoost classifier on the training dataset, we assessed its performance on a separate test set. The model delivered remarkable results, exhibiting an exceptional accuracy of 91.2%. This high level of accuracy underlines XGBoost's strength in effectively handling various data types and its capacity to yield outstanding results in classification tasks.

Model Evaluation:

MAE (Mean Absolute Error)

MAE measures the average absolute difference between the predicted values and the actual values. A lower MAE indicates that the model's predictions are closer to the actual values.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y}_i|$$

RMSE (Root Mean Square Error)

RMSE is a measure of the average magnitude of the errors between predicted and actual values. It's similar to MAE but gives more weight to larger errors. A lower RMSE is better.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}$$

R2 Score

The R-squared (R2) score measures the proportion of the variance in the dependent variable that is predictable from the independent variable(s). A higher R2 score indicates a better fit of the model to the data, with a value of 1 being a perfect fit.

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

The LSTM model has an MAE of 0.0227, RMSE of 0.0294, R2 Score of 0.964, MSE of 0.00086, and RMS of 0.0294 for predicting Accelerometer data. These values indicate that the model's predictions for Accelerometer data are quite accurate, with a strong correlation (R2 Score) and low errors (MAE and RMSE).

The model performs even better when predicting Location Speed, with an MAE of 0.0599, RMSE of 0.074, R2 Score of 0.9989, MSE of 0.00584, and RMS of 0.0764. These values indicate very accurate predictions for Location Speed data.

The model's performance for Magnetometer is decent but not as accurate as Accelerometer and Location Speed. It has an MAE of 0.6620, RMSE of 0.918, R2 Score of 0.9718, MSE of 0.00086, and RMS of 0.294. The high MAE and RMSE suggest larger errors and less accurate predictions for Magnetometer data.

The LSTM model performs well when predicting Gyro Rotation, with an MAE of 0.0138, RMSE of 0.035, R2 Score of 0.9981, MSE of 0.001241, and RMS of 0.03523. These values indicate very accurate predictions for Gyro Rotation data.

The model's performance for Motion Pitch is also strong, with an MAE of 0.01584, RMSE of 0.0216, R2 Score of 0.9803, MSE of 0.000467, and RMS of 0.021627. These values suggest highly accurate predictions for Motion Pitch data.

The performance of each model was assessed using standard evaluation metrics. These metrics include accuracy, precision, recall, and F1-score.

Precision

Precision is a key performance metric used in classification tasks, particularly when dealing with imbalanced datasets. It measures the accuracy of positive predictions made by a model.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall

Recall, often referred to as sensitivity or true positive rate, measures a model's ability to identify all the relevant instances within a dataset. It quantifies the percentage of true positive predictions (correctly identified positive instances) out of all actual positive instances (true positives and false negatives combined).

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1 Score

F1 Score is a metric that combines both precision and recall into a single value. It is particularly useful when there is a trade-off between precision and recall, and we want to balance these two factors.

$$\text{F1 - score} = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

The following results were achieved:

- LSTM: Accuracy of 73.87%
- CNN: Accuracy of 79%
- Random Forest: Accuracy of 89%
- XG Boost: Accuracy of 91.2%

The methodology employed in this project involved data collection, labeling, preprocessing, and the application of various machine learning models. The results indicate that the Random Forest and XGBoost models outperformed others in predicting driving behavior, providing valuable insights for behavioral analysis and risk assessment in the context of two-wheeler riders.

3.6 Hardware & Software Specifications

Hardware:

- Sensors for capturing vehicle data.

Software:

- Python (Tensorflow, Keras, Sklearn)
- Google Colab

3.7 Result Analysis and Discussion

For non-labeled data:

Actual vs Predicted LSTM fitted results with regression residuals for individual parameters

```
MAE: 0.033543894506554504
RMSE: 0.04578088012025078
R2: 0.9293290928859587
```

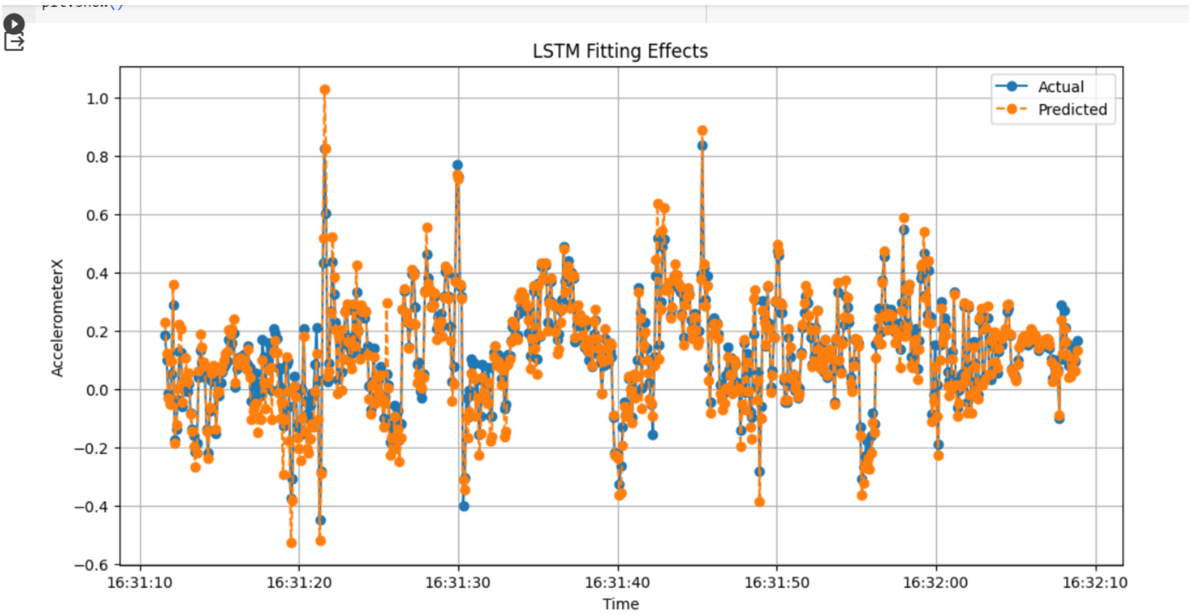

fig. 3: Actual vs Predicted values of AccelerometerX

```
0.0024885972424101865
0.04988584210384933
```

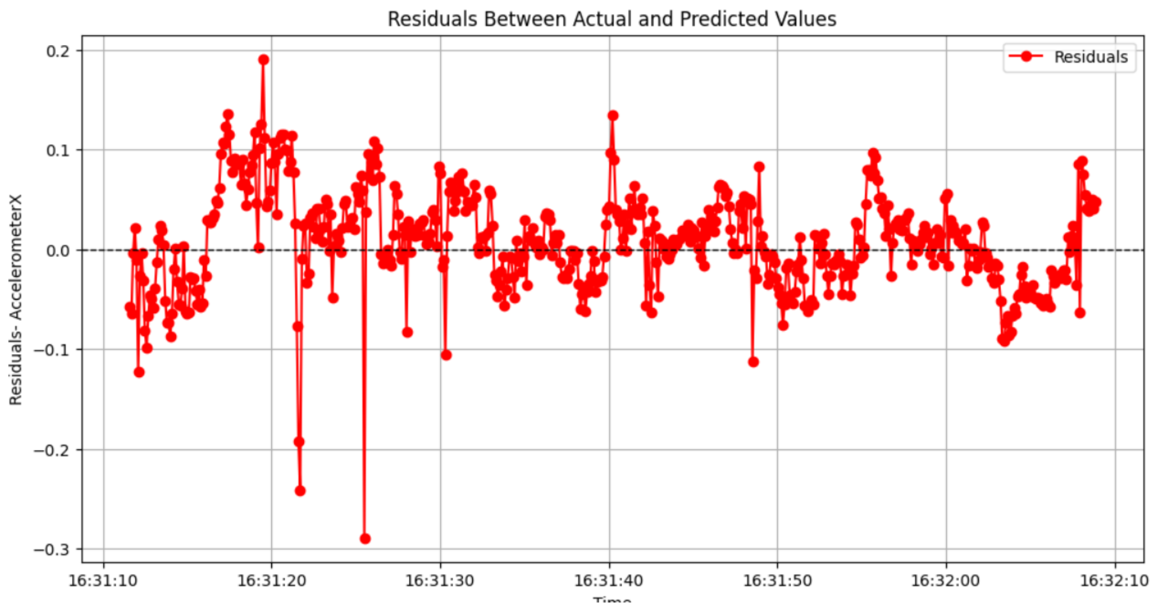

fig. 4: Residuals of AccelerometerX

```
MAE: 0.04129394100910706
RMSE: 0.05540431845702206
R2: 0.9180873270302313
```

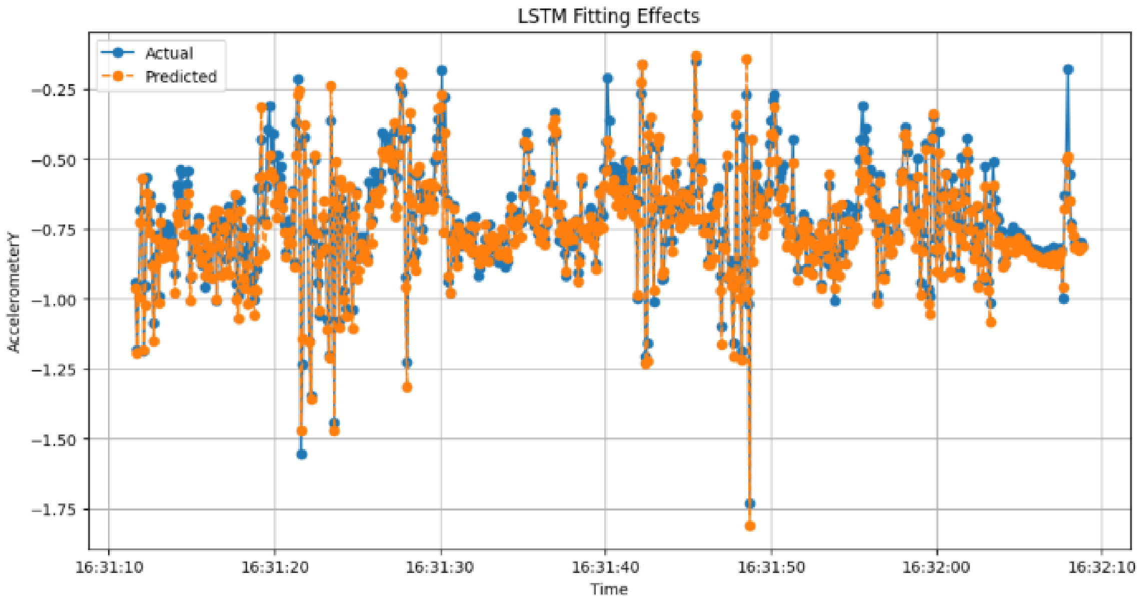

fig. 5: Actual vs Predicted values of AccelerometerY

```
0.0030696385036871154
0.05540431845702206
```

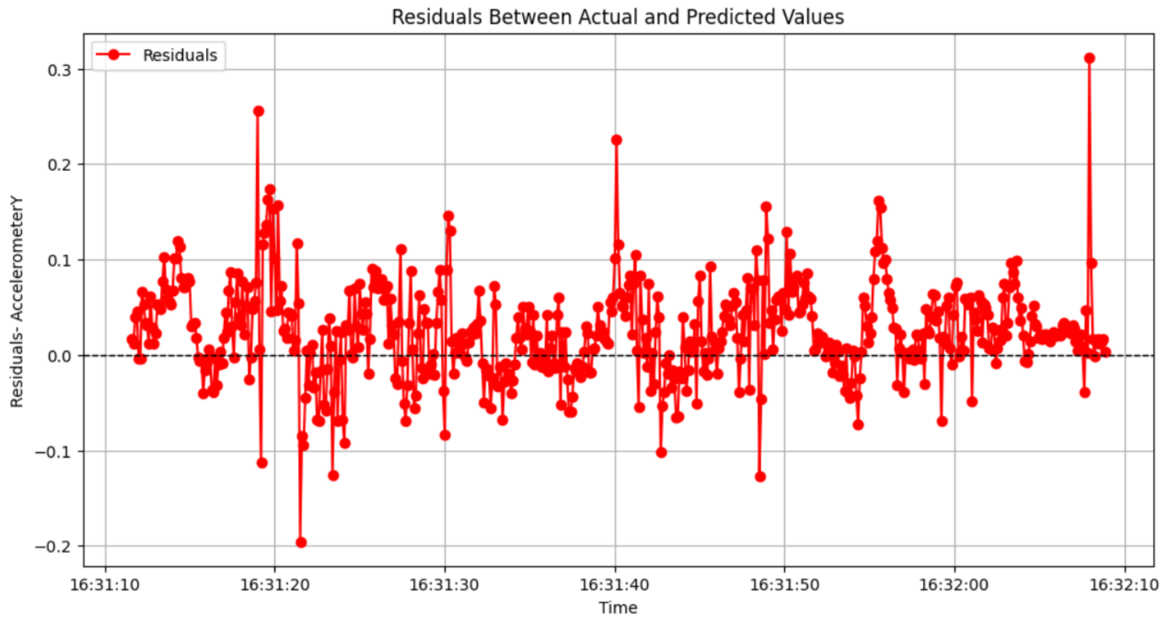

fig. 6: Residuals of AccelerometerY

```
MAE: 0.05296301519829222
RMSE: 0.07607557094486156
R2: 0.7985960980871547
```

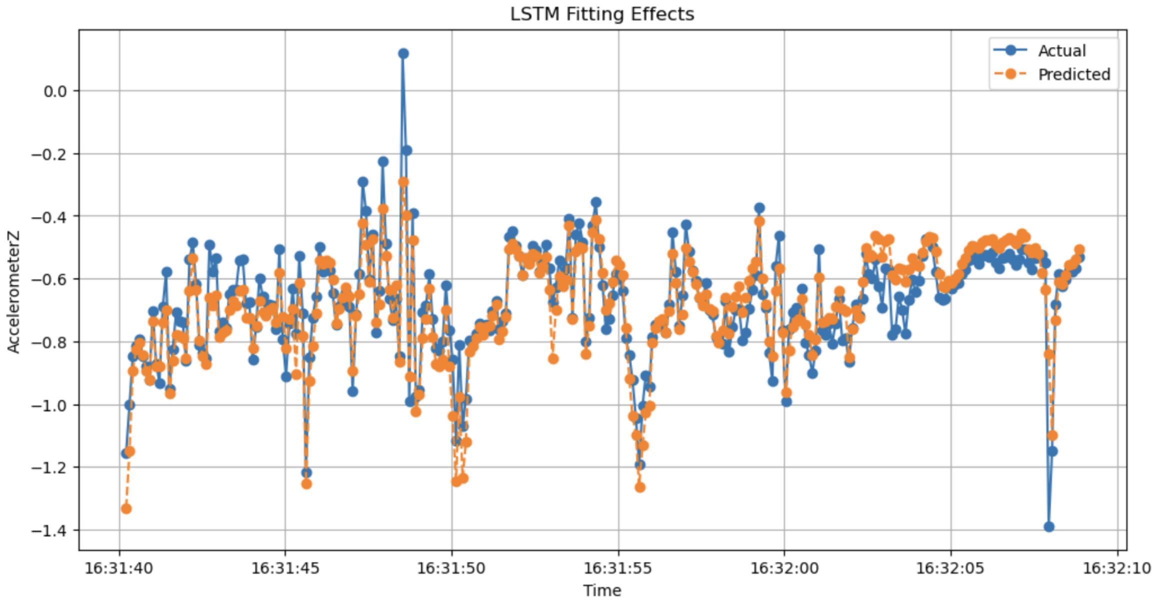

fig.7: Actual vs Predicted values of AccelerometerZ

```
0.005787492494586663
0.07607557094486156
```

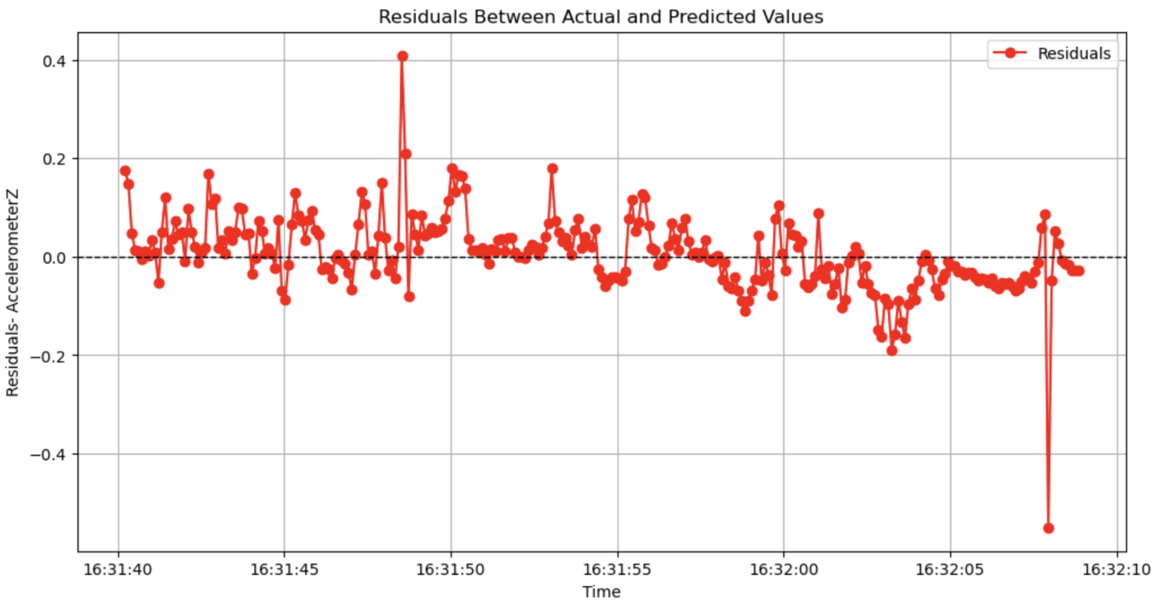

fig. 8: Residuals AccelerometerZ

MAE: 0.7692779154885712
RMSE: 0.9970465881222829
R2: 0.9775695363340825

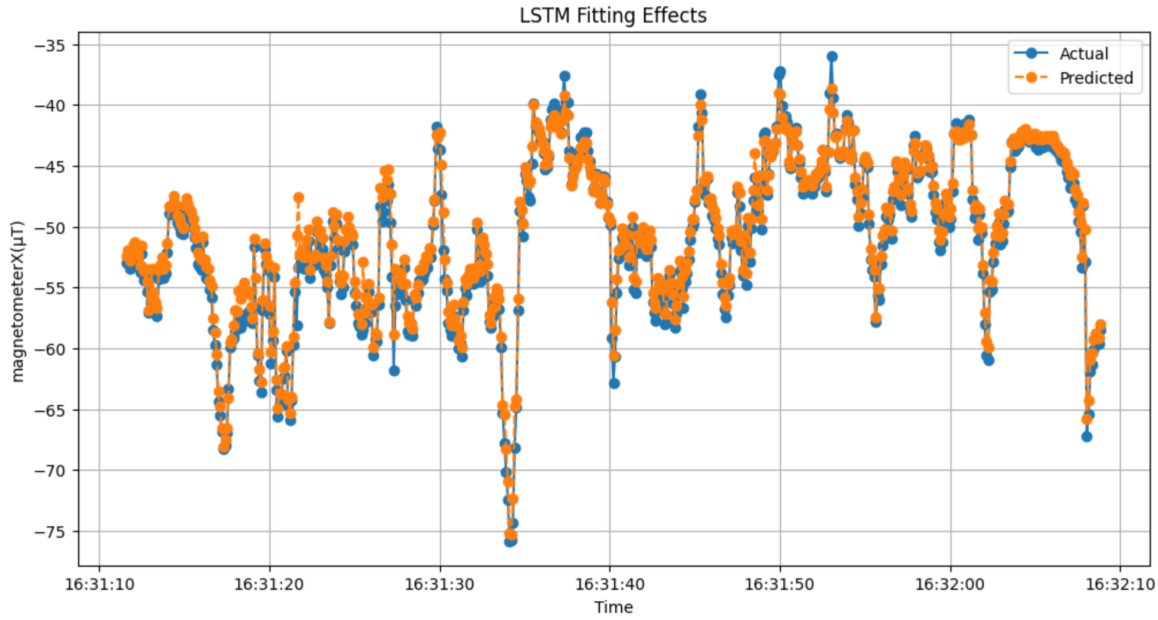

fig. 9: Actual vs Predicted values of MagnetometerX

mse 0.9941018988862851
rms 0.9970465881222829

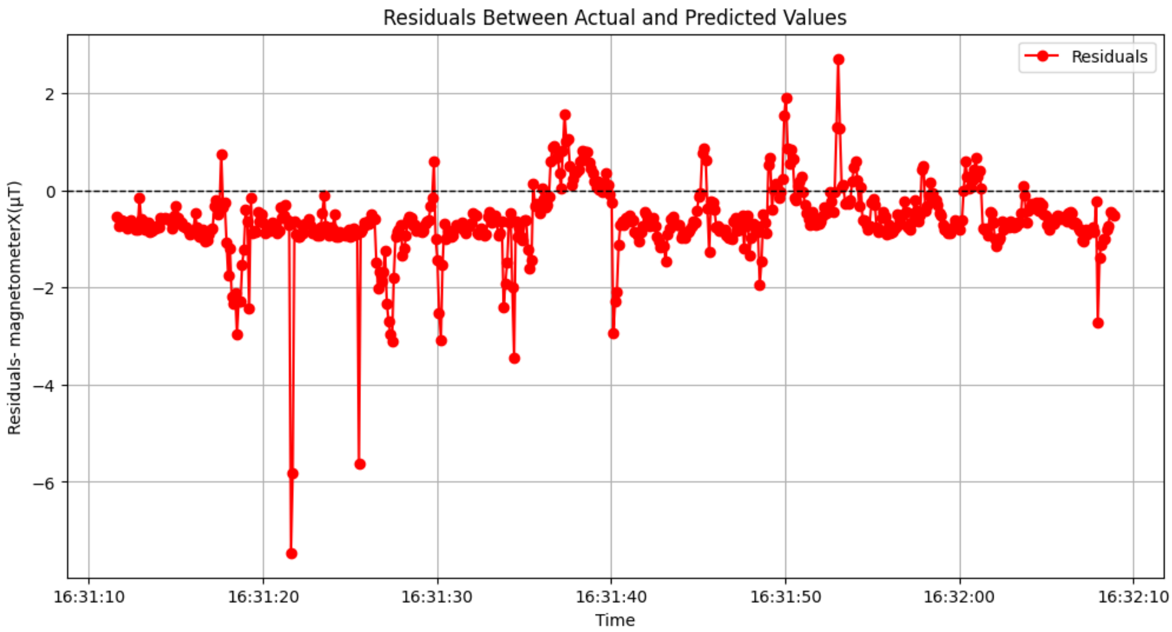

fig. 10: Residuals-Magnetometer(X)

```
MAE: 0.6036308381777996
RMSE: 0.8606629362391726
R2: 0.9892233391152349
```

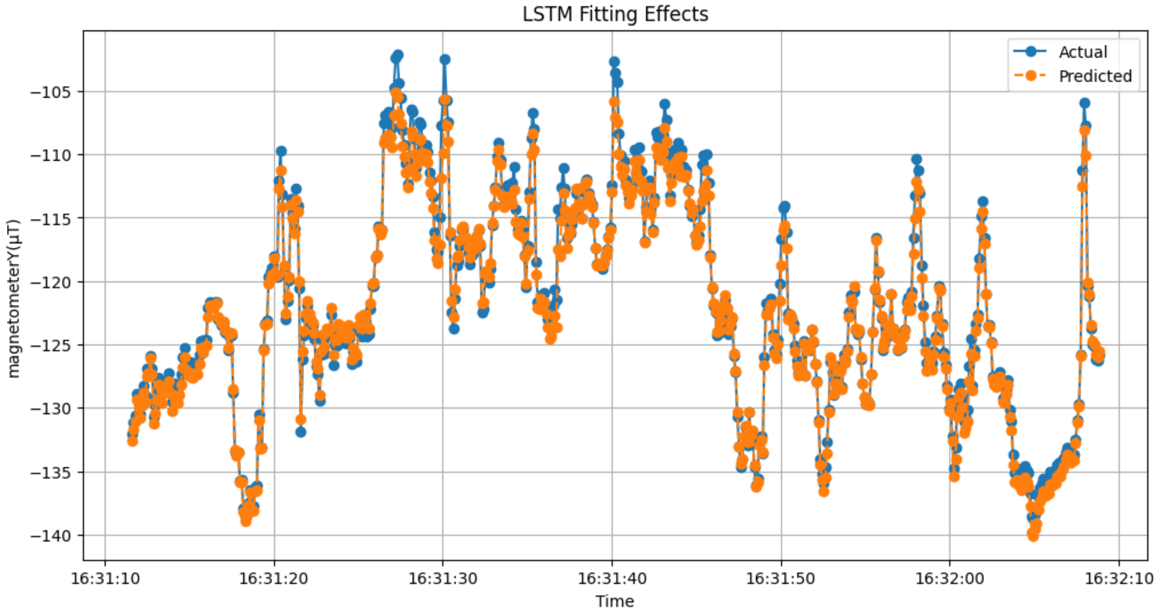

fig. 11: Actual vs Predicted values of MagnetometerY

```
mse 1.2762781568333765
rms 1.1297248146488492
```

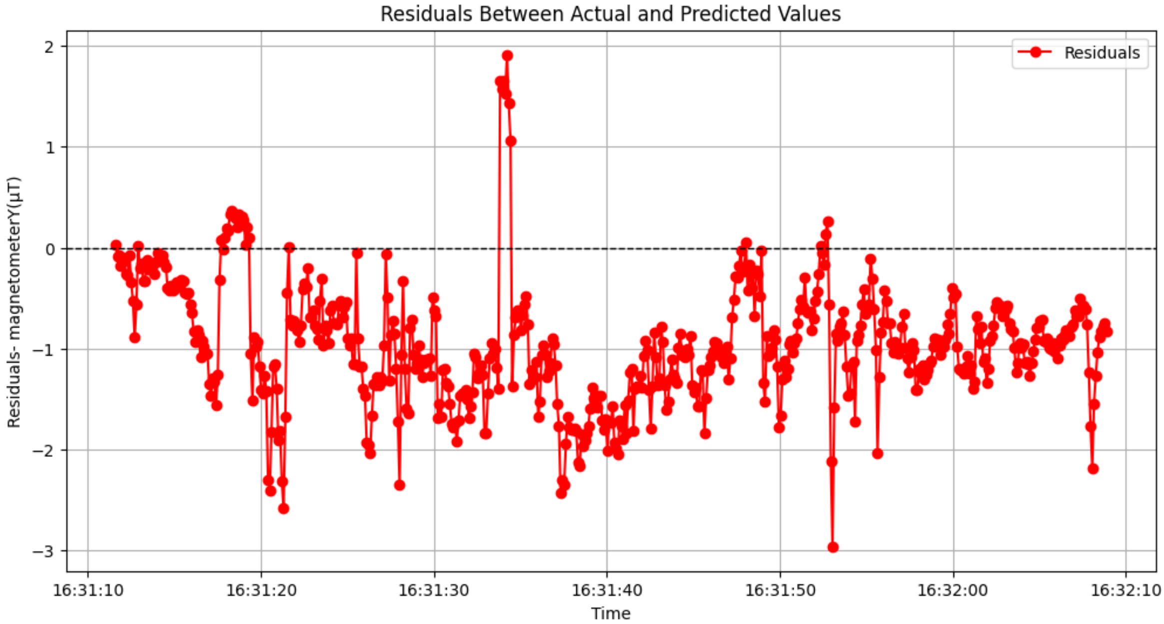

fig. 12: Residuals-Magnetometer(Y)

```
MAE: 0.884333126207921
RMSE: 1.2239935798790433
R2: 0.9661961919152107
```

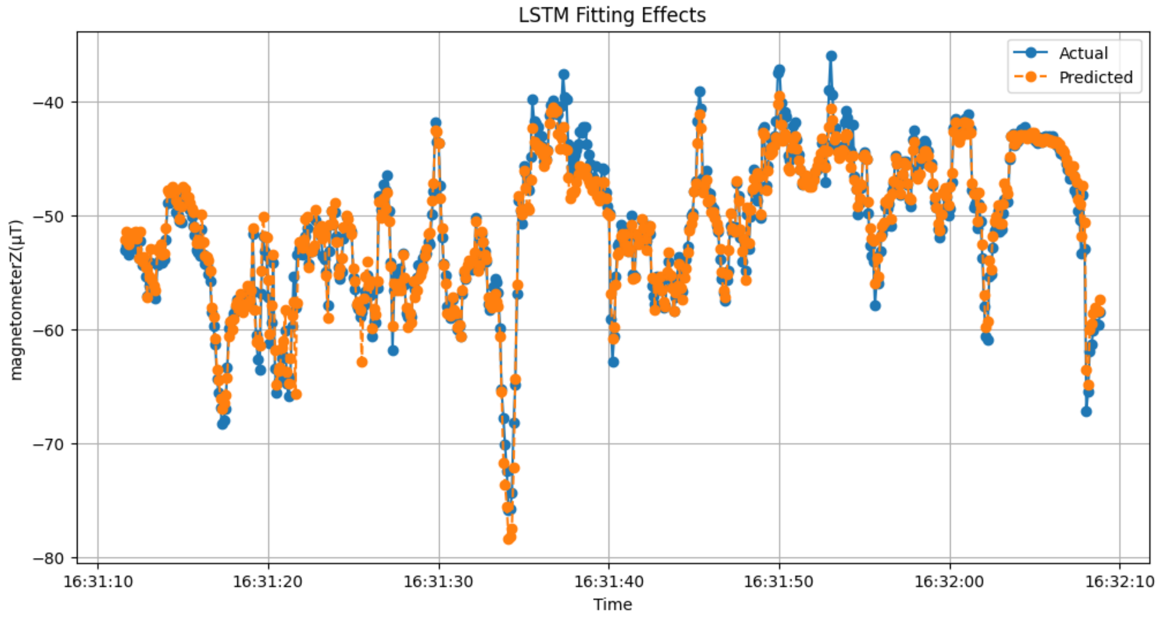

fig. 13: Actual vs Predicted values of MagnetometerZ

```
mse 1.9555852233073072
rms 1.398422405179246
```

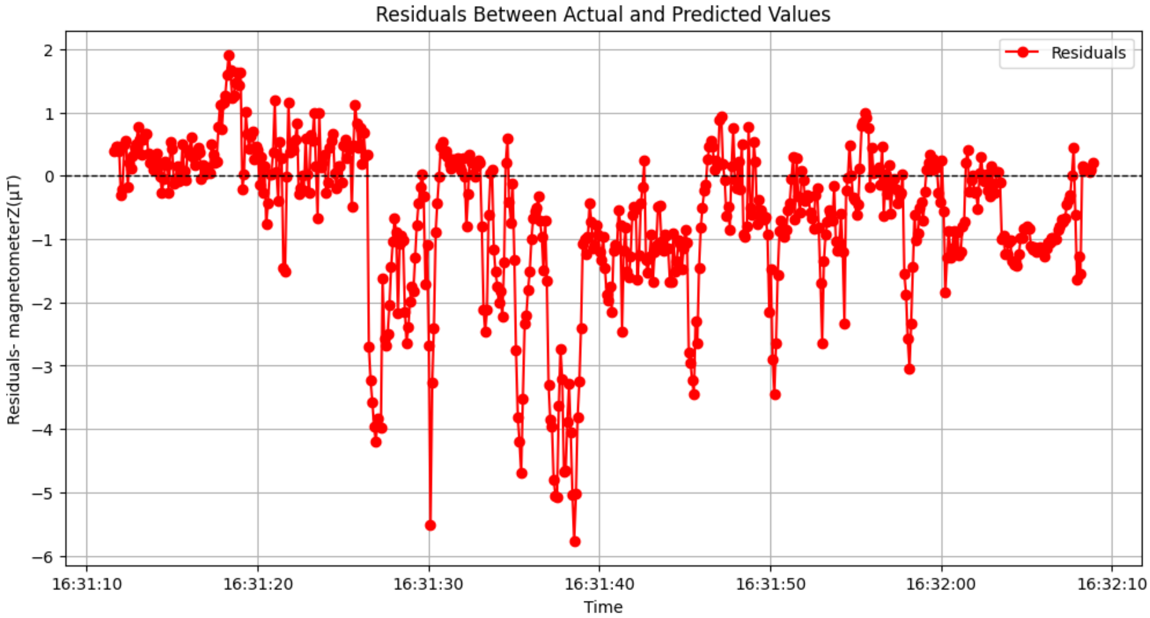

fig. 14: Residuals-Magnetometer(Z)

```
MAE: 0.3376563051303978
RMSE: 0.4951716337693991
R2: 0.6013571699872229
```

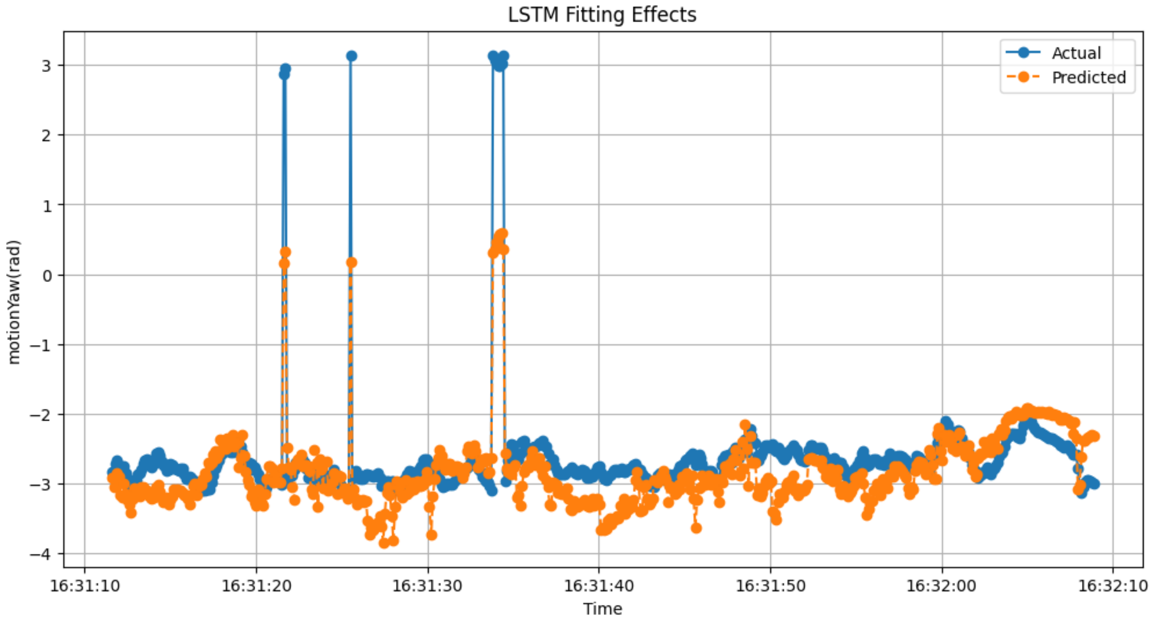

fig. 15: Actual vs Predicted values of motionYaw(rad)

```
mse 0.2451949468898559
rms 0.4951716337693991
```

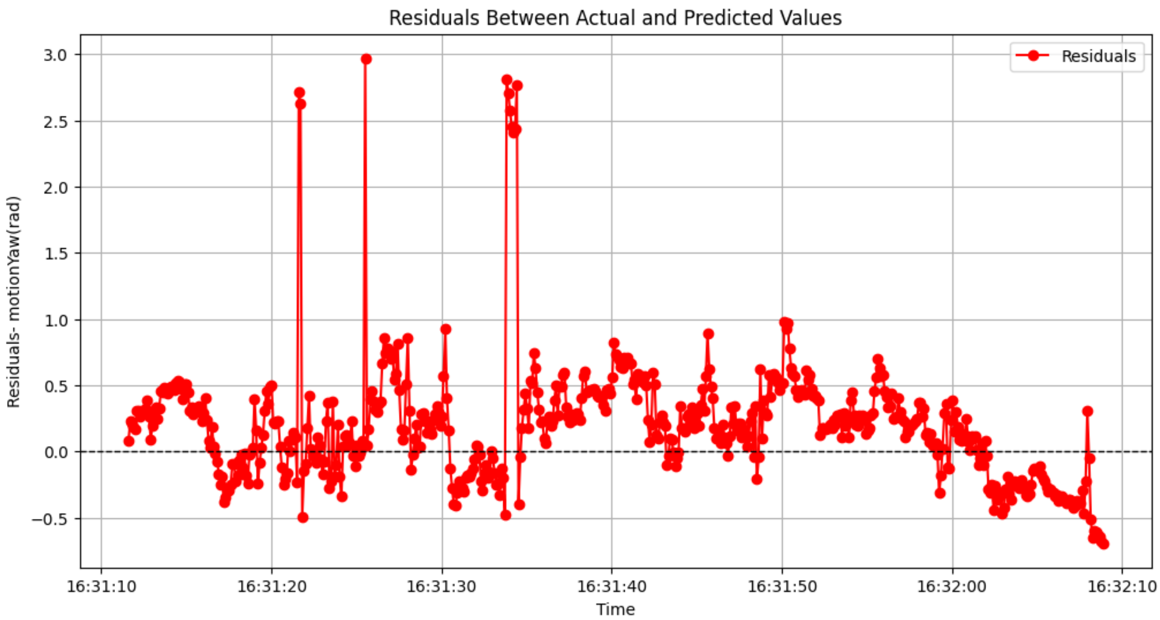

fig. 16:Residuals- motionYaw

```
MAE: 0.1562989585408793
RMSE: 0.20295714420596198
R2: -0.5675365644291273
```

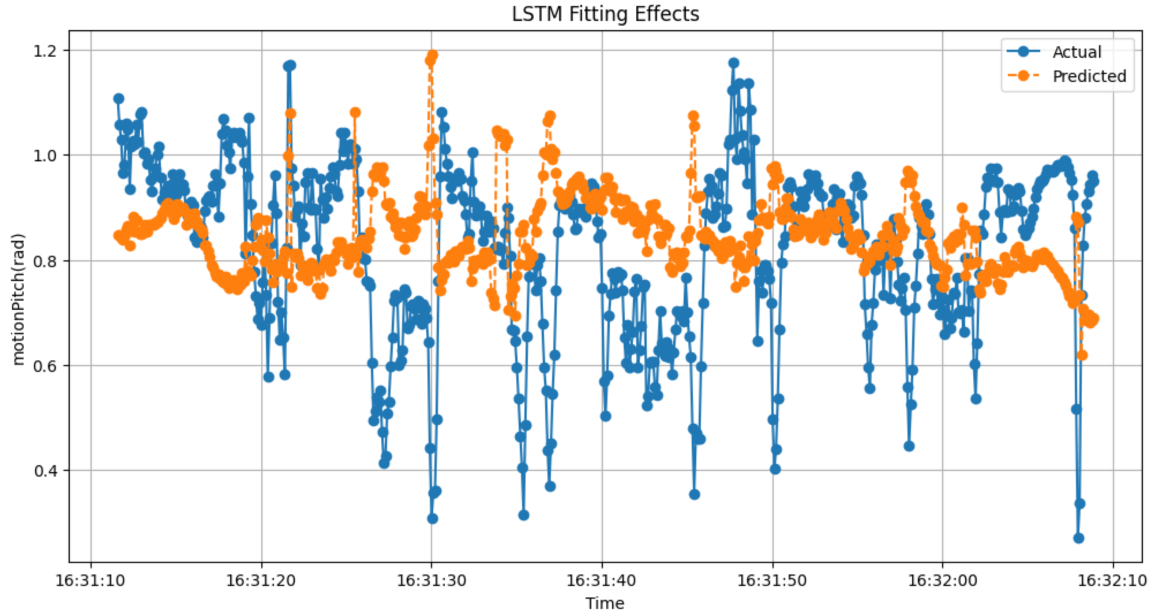

fig. 17: Actual vs Predicted values of motionPitch(rad)

```
mse 0.0007958682985853257
rms 0.028211137846342278
```

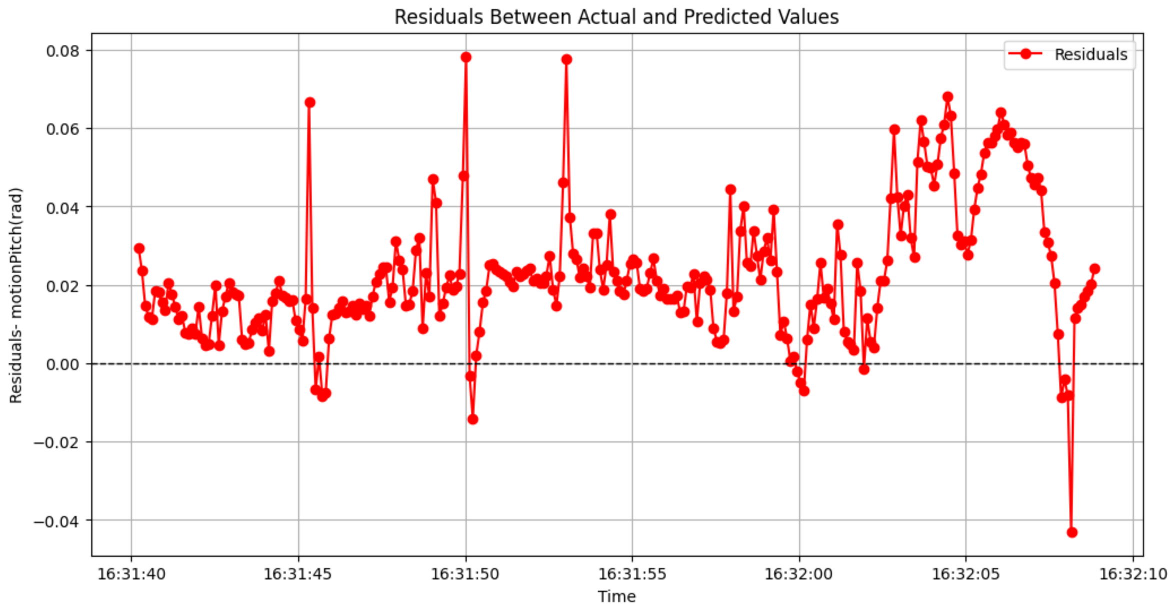

fig. 18: Residuals - motionPitch(rad)

```
MAE: 0.03085785649869797
RMSE: 0.04117053522059662
R2: 0.9996804316677876
```

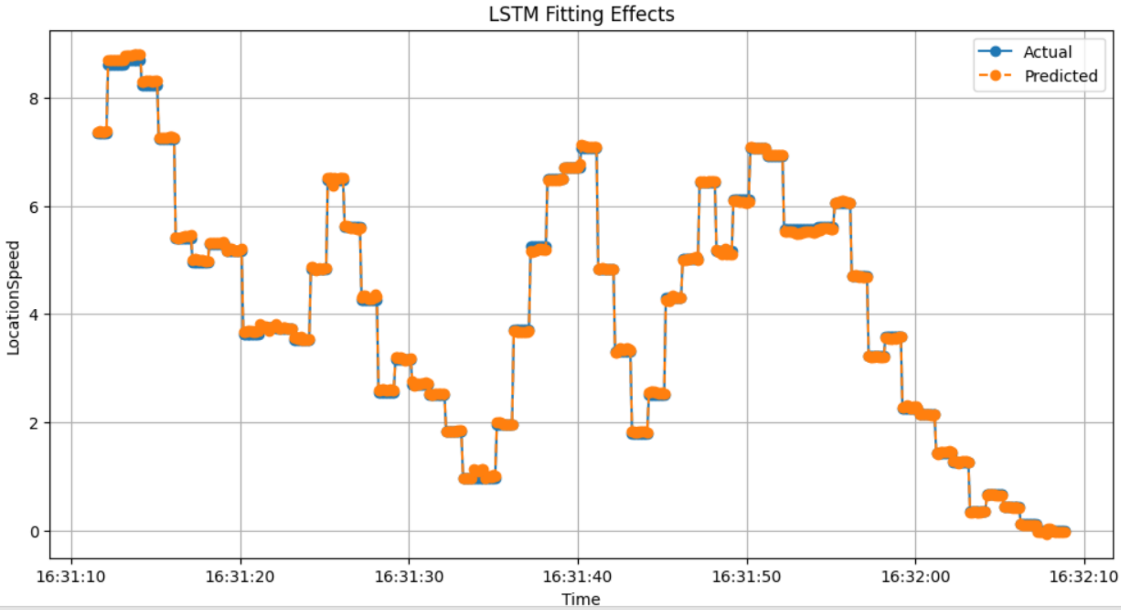

fig. 19: Actual vs Predicted values of LocationSpeed

```
0.0016950129703503865
0.04117053522059662
```

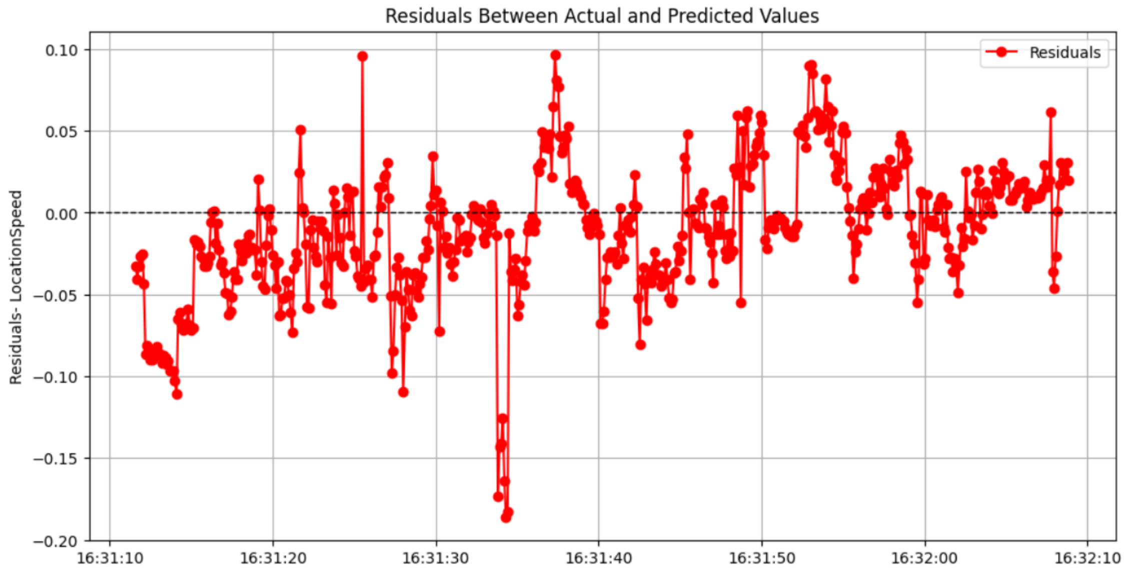

fig.20: Residuals of LocationSpeed

For labeled data:

Classification report for comparison of models

1. LSTM

```
In [88]: # Print the metrics
         print(f"Accuracy: {accuracy * 100:.2f}%")
         print(f"Precision: {precision:.2f}")
         print(f"Recall: {recall:.2f}")
         print(f"F1 Score: {f1:.2f}")

         Accuracy: 73.87%
         Precision: 0.75
         Recall: 0.74
         F1 Score: 0.74
```

2. CNN

```
# Calculate overall precision, recall, and F1 score
precision = precision_score(y_test, y_pred_classes, average='weighted')
recall = recall_score(y_test, y_pred_classes, average='weighted')
f1 = f1_score(y_test, y_pred_classes, average='weighted')

print(f'Overall Precision: {precision:.2f}')
print(f'Overall Recall: {recall:.2f}')
print(f'Overall F1 Score: {f1:.2f}')
```

```
27/27 [==============================] - 0s 1ms/step
Overall Accuracy: 0.78
Overall Precision: 0.76
Overall Recall: 0.78
Overall F1 Score: 0.77
/usr/local/lib/python3.10/dist-packages/sklearn/metrics/_classification.
  _warn_prf(average, modifier, msg_start, len(result))
```

3. Random Forest

```
In [16]: print(classification_report(y_test, y_pred))
              precision    recall  f1-score   support

           0       0.90      0.91      0.91       163
           1       0.84      0.94      0.89        87
           2       0.00      0.00      0.00         3
           3       0.96      0.71      0.81        34

    accuracy                           0.89       287
   macro avg       0.67      0.64      0.65       287
weighted avg       0.88      0.89      0.88       287
```

4. XGBoost

```
[9]  # Evaluate the classifier
     accuracy = accuracy_score(y_test, y_pred)
     confusion = confusion_matrix(y_test, y_pred)
     report = classification_report(y_test, y_pred)
```

```
[10] print("Accuracy:", accuracy)
     print("Confusion Matrix:")
     print(confusion)
     print("Classification Report:")
     print(report)
```

```
Accuracy: 0.9128919860627178
Confusion Matrix:
[[149  10   0   4]
 [  5  82   0   0]
 [  3   0   0   0]
 [  3   0   0  31]]
Classification Report:
              precision    recall  f1-score   support

           0       0.93      0.91      0.92       163
           1       0.89      0.94      0.92        87
           2       0.00      0.00      0.00         3
           3       0.89      0.91      0.90        34

    accuracy                           0.91       287
   macro avg       0.68      0.69      0.68       287
weighted avg       0.90      0.91      0.91       287
```

Algorithm	Accuracy	Precision	Recall	F1-Score
LSTM	0.73	0.75	0.74	0.74
CNN	0.78	0.76	0.78	0.77
Random Forest	0.89	0.88	0.89	0.88
XG-Boost	0.91	0.90	0.91	0.92

Table 1: Comparison of algorithms

	LSTM			**RESIDUAL**	
Parameters	**MAE**	**RMSE**	**R2_SCORE**	**MSE**	**RMS**
Accelerometer	0.0335	0.0457	0.9293	0.0024	0.0498
Location speed	0.0308	0.0411	0.9996	0.0016	0.0411
magnetometer	0.7692	0.9970	0.9775	0.9941	0.9970
gyro rotation	0.0138	0.035	0.9981	0.001241	0.03523
motion pitch	0.1562	0.2029	-0.5675	0.000795	0.02821

Table 2: Evaluation Metrics for Individual Parameters

3.8 Conclusion

The project "Behavioral Analysis and Risk Assessment of Two-Wheeler Drivers" presents a comprehensive analysis of driving patterns and risk assessment for two-wheeler riders. Machine learning models were applied to predict driving behavior, with a particular focus on the LSTM, CNN, Random Forest, and XG Boost algorithms for labeled data and LSTM Residual for non - labeled data. These models were evaluated based on various performance metrics, including accuracy, precision, recall, F1-score, R2 score. The Random Forest and XGBoost models outperform others, achieving an accuracy of 89% and 91.2%, respectively. The project involved an extensive literature survey, which highlighted the need for proactive solutions tailored specifically for two-wheeler safety. Existing systems were found to be primarily designed for cars, leaving a gap in holistic support for two-wheeler riders.

This project represents a significant step toward improving road safety and addressing the pressing issue of two-wheeler rider accidents. The methodologies and findings presented here lay the foundation for future research and initiatives aimed at creating a safer environment for all road users in India.

References

1. Bouaouni, Yacine & Yahia, Rayane & Boubezoul, Abderrahmane. (2021). Driving-Pattern Identification and Event Detection Based on an Unsupervised Learning Framework: Case of a Motorcycle-Riding Simulator. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3130400.
   https://www.researchgate.net/publication/356483474

2. Deenadayalan, Jaganathan & Vinayagam, Prabhu & Chinnasamy, A & Suresh, A. (2022). Detection of Two wheeler Driver Safety Using Machine Learning. 2020.
   https://www.researchgate.net/publication/365265615

3. Jaikishan Damani, Perumal Vedagiri,
   Safety of motorised two wheelers in mixed traffic conditions: Literature review of risk factors,
   Journal of Traffic and Transportation Engineering (English Edition), Volume 8, Issue 1,2021
   https://www.sciencedirect.com/science/article/pii/S2095756421000027

4. Z. Chen, J. Yu, Y. Zhu, Y. Chen and M. Li, "D3: Abnormal driving behaviors detection and identification using smartphone sensors," *2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, Seattle, WA, USA, 2015, pp. 524-532, doi: 10.1109/SAHCN.2015.7338354.
   https://ieeexplore.ieee.org/document/7338354

5. Eleni I. Vlahogianni, George Yannis, John C. Golias,
   Detecting Powered-Two-Wheeler incidents from high resolution naturalistic data,
   Transportation Research Part F: Traffic Psychology and Behaviour,
   Volume 22,2014
   https://www.sciencedirect.com/science/article/abs/pii/S1369847813001101

6. Ghandour, R.; Potams, A.J.; Boulkaibet, I.; Neji, B.; Al Barakeh, Z. Driver Behavior Classification System Analysis Using Machine Learning Methods. *Appl. Sci.* 2021, *11*, 10562.
https://doi.org/10.3390/app112210562

7. Yongfeng Ma, Zhuopeng Xie, Shuyan Chen, Fengxiang Qiao, Zeyang Li,
Real-time detection of abnormal driving behavior based on long short-term memory network and regression residuals, Transportation Research Part C: Emerging Technologies,
Volume 146, 2023, 103983, ISSN 0968-090X,
https://doi.org/10.1016/j.trc.2022.103983.

8. Y. Cai, R. Zhao, H. Wang, L. Chen, Y. Lian and Y. Zhong, "CNN-LSTM Driving Style Classification Model Based on Driver Operation Time Series Data," in *IEEE Access*, vol. 11, pp. 16203-16212, 2023, doi: 10.1109/ACCESS.2023.3245146.
https://ieeexplore.ieee.org/abstract/document

9. S. Bouhsissin, N. Sael and F. Benabbou, "Driver Behavior Classification: A Systematic Literature Review," in *IEEE Access*, vol. 11, pp. 14128-14153, 2023, doi: 10.1109/ACCESS.2023.3243865.
https://ieeexplore.ieee.org/document

10. G. Castignani, T. Derrmann, R. Frank and T. Engel, "Driver Behavior Profiling Using Smartphones: A Low-Cost Platform for Driver Monitoring," in *IEEE Intelligent Transportation Systems Magazine*, vol. 7, no. 1, pp. 91-102, Spring 2015, doi: 10.1109/MITS.2014.2328673.
https://ieeexplore.ieee.org/abstract/document