Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 272 (2025) 594–600

Procedia Computer Science

www.elsevier.com/locate/procedia

The 2nd International Workshop on Intelligent Mobile Systems based on Internet of Things
October 28-30, 2025, Istanbul, Türkiye

Driver Behaviour Analysis Using Telematics Sensor Data and Deep Learning Models

Didar Yedilkhan[a], Nurbakhyt Agybetov[a], Beibut Amirgaliyev[a,*]

[a]*Astana IT University, Astana, 010000, Kazakhstan*

Abstract

This research evaluates the effectiveness of deep learning approaches for detecting aggressive driving behaviours through time-series data collected from inertial sensors. The publicly available Driving Behaviour Dataset allowed us to develop and test three neural network models including Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Gated Recurrent Unit (GRU) using tri-axial accelerometer and gyroscope readings. The data preparation involved z-score normalization followed by sliding window segmentation and one-hot encoding of behaviour labels. The evaluation of model performance relied on standard classification metrics which included accuracy alongside precision, recall, F1-score and confusion matrices. The LSTM architecture achieved the highest classification accuracy at 96% which demonstrates its ability to detect driving behaviour patterns across time. The research demonstrates how deep learning models especially sequence-based LSTMs, show promise for real-time driver behaviour recognition, which could lead to applications in intelligent transportation systems and advanced driver-assistance technologies.

Keywords: Driving Behaviour; Deep Learning; LSTM; CNN; GRU; Time-Series Classification; Inertial Sensors; Accelerometer; Gyroscope; Intelligent Transportation Systems.

* Corresponding author. Tel.: +7-701-033-9111.
E-mail address: Beibut.amirgaliyev@astanait.edu.kz

10.1016/j.procs.2025.10.253

1. Introduction

The worldwide speed of urbanization puts additional strain on urban infrastructure networks, which include transportation systems. The well-developed public transport system in London helps minimize mobility inequality between different population groups. The widespread vehicle ownership in car-dependent cities, including Dubai and Astana, as well as many American urban areas, results in traffic congestion, air pollution, and increased road accident risks. The World Health Organization reports that human error causes more than 1.19 million annual traffic deaths worldwide [2]. The United Nations Sustainable Development Goals include road safety improvement as a global priority through SDG 3.6 and SDG 11.2 and Vision Zero initiatives. Technology provides scalable solutions beyond infrastructure and legal reforms. The research investigates mobile-based assessment methods that use smartphone sensor data to evaluate driving conduct. These systems enhance driver knowledge while decreasing the likelihood of accidents. Machine learning algorithms demonstrate superior performance in detecting unsafe behaviour from time-series data compared to traditional rule-based systems because they offer better adaptability and accuracy.

2. Literature Review

Over the past decade, research on driver behaviour has shifted from rule-based approaches to data-driven machine learning models, enabled by advancements in sensors and computational power [12]. Early studies applied simple thresholding methods (e.g., acceleration > 3 m/s^2 or deceleration < -4 m/s^2) [16], but these lacked flexibility across different environments and vehicle types. To address these limitations, supervised learning models such as Random Forest, k-Nearest Neighbours (k-NN), and Support Vector Machines (SVM) were introduced. Lee and Jang [9] demonstrated that even basic classifiers could effectively predict unsafe behaviours using naturalistic driving data. Johnson and Trivedi [6] employed SVMs with smartphone accelerometer data to detect events like hard braking and lane changes. Shoaib et al. [15] used ensemble learning on mobile motion data, achieving classification accuracies above 80%. Tree-based models like XGBoost have gained popularity for their accuracy and interpretability. Zhao et al. [18] demonstrated that axis-specific acceleration and angular velocity features were crucial in identifying aggressive patterns with over 90% accuracy. Deep learning methods, particularly RNNs and LSTMs, further improved recognition of temporal dependencies in sensor data. Kim et al. [8] achieved F1-scores above 90% using accelerometer and gyroscope signals, although such models are limited by high computational demands [11]. Hybrid architectures and reinforcement learning (RL) have also been explored [1][17], but practical deployment faces challenges in safety, generalizability, and interpretability. A major limitation in the literature is poor transferability—models trained in one region often underperform elsewhere due to variability in traffic, weather, and road types [18]. In conclusion, while deep learning models provide higher accuracy, classical ML approaches offer better efficiency and are more suitable for real-time, embedded applications. This research compares both categories to find a balance between predictive performance and computational feasibility for real-world deployment.

3. Materials and Methods

3.1. Dataset and Exploratory Analysis

This study utilizes the publicly available Driving Behaviour Dataset [1], designed for classifying aggressive driving using inertial sensor data. It comprises six time-series features recorded by the MPU6050 sensor module—three-axis accelerometer (AccX, AccY, AccZ) and gyroscope (GyroX, GyroY, GyroZ). Data were collected under real driving conditions using three vehicles (Ford Fiesta 1.4, Ford Fiesta 1.25, and Hyundai i20), each operated by three different drivers. This diversity enhances the generalizability of model training. Each sample is labelled with one of four aggressive driving manoeuvres: sudden acceleration, hard braking, sharp right turn, and sharp left turn. An initial exploratory data analysis included histograms (Fig.1) and a Pearson correlation heatmap (Fig. 2). Distributions showed skewness in GyroZ and AccZ, indicating frequent sharp manoeuvres. Correlation analysis revealed strong relationships between GyroZ and AccY (r = 0.82), and a moderate inverse relationship between GyroX and GyroZ (r = –0.46), providing insight into dynamic vehicle behaviour and informing model architecture.

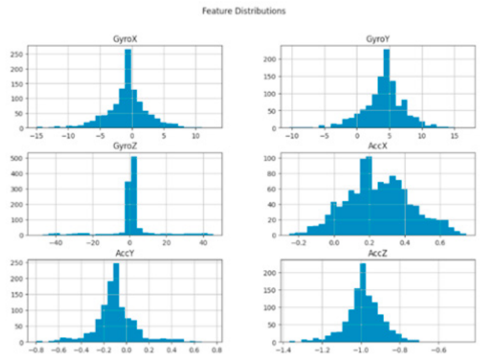

Fig. 1. Feature-wise histogram

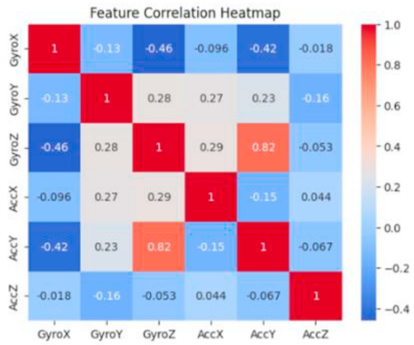

Fig.2. Pearson correlation heatmap

3.2. Data Preprocessing

Data preprocessing began with the removal of incomplete or noisy sensor samples to ensure the integrity of the dataset. All features were standardized using z-score normalization (Equation 1), resulting in a mean of zero and a variance of one per feature. A sliding window approach was applied to preserve temporal dynamics, segmenting the signal into overlapping sequences of 14-time steps (~7 seconds). Each window was labelled based on the final time step, reflecting the dominant behaviour at the end of the sequence. Class labels were first encoded as integers and then transformed into one-hot vectors to align with the categorical cross-entropy loss function. The dataset was finally split into training and testing subsets using a 70/30 ratio while preserving class balance to ensure fair evaluation.

$$x' = \frac{(x-\mu)}{\Omega} \tag{1}$$

3.3. Types of Models

The use of deep learning models for sequential classification tasks is widespread because they excel at modeling intricate temporal relationships in time-series data [19, 20]. The research chose LSTM, GRU and 1D-CNN architectures because they have shown success in comparable applications. LSTM stands as a basic yet powerful model for time-series modeling despite its straightforward nature.

A. Long Short-Term Memory (LSTM) (Fig.3)
The LSTM model contained a single LSTM layer with 64 hidden units followed by a dense layer with ReLU activation and a final softmax output layer for classification (Fig. 3). A dropout layer with a rate of 0.3 was applied after the LSTM layer to prevent overfitting. The input, forget, and output gates in LSTM networks enable these models to learn long-term dependencies in sequential data which makes them suitable for time-series modeling tasks.

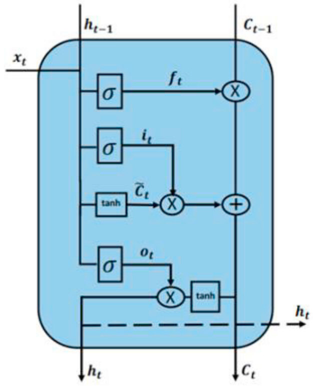

Fig.3. LSTM architecture (adapted from Saleh et al., 2017[13])

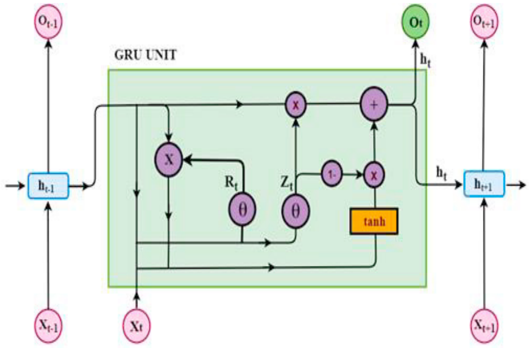

Fig. 4. GRU architecture (adapted from Doniec et al., 2020 [4])

B. 1D Convolutional Neural Network (CNN)

The CNN architecture was designed to capture local temporal patterns within the time-series data. It included a one-dimensional convolutional layer with 64 filters and a kernel size of 3, followed by max pooling, dropout (rate = 0.3), flattening, and a dense output layer (Fig. 5). CNNs are advantageous in terms of training speed and parallelization and are particularly effective at extracting short-range dependencies in sequential inputs.

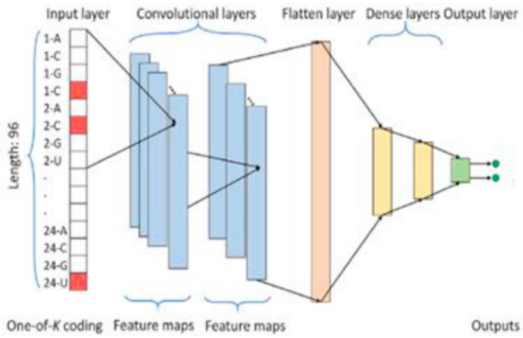

Fig. 5. 1D Convolutional Neural Network (CNN) architecture (adapted from Zhao et al., 2018 [21])

C. Gate Recurrent Unit (GRU)

The GRU model adopted the same architecture as the LSTM model by substituting the LSTM layer with a GRU layer that had 64 units (Fig. 4). All other architectural components remained unchanged. GRUs simplify the internal mechanisms of LSTMs by combining the forget and input gates into a single update gate, reducing the number of parameters and computational requirements while maintaining strong performance in modelling temporal sequences.

3.4 Protocol for Training and Evaluation

The Adam optimizer (learning rate = 0.001) and categorical cross-entropy were used as the loss function for all models. They were trained for 20 epochs with a batch size of 32. No early stopping or learning rate scheduling was used to make sure that evaluations were always the same. Using stratified sampling to keep the class distribution, the dataset was split into training (70%) and testing (30%) sets. We kept track of how the training and validation accuracy, and loss changed at each epoch to monitor the models' learning progress. Figure 6 shows the results for the LSTM model.

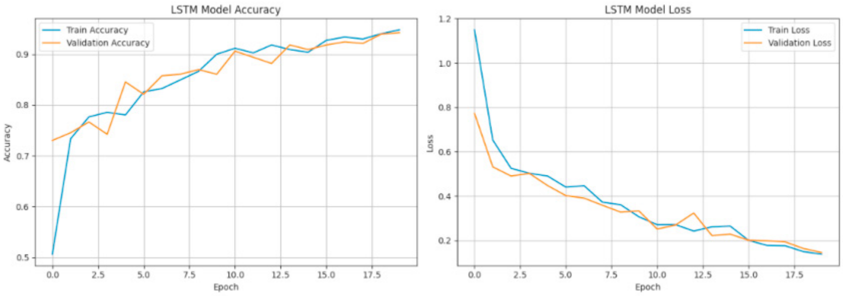

Fig. 6. Training and Validation Accuracy and Loss Curves for the LSTM Model

The learning curves show that the model is stable and can generalize well. The validation accuracy is very close to the training accuracy, indicating that the model has learned temporal dependencies without overfitting. The fact that both training and validation loss dropped a lot in the first few epochs shows how well LSTM can pick up on important temporal features early on in training.

4. Results

The evaluation of traditional and deep learning models used standard performance metrics which included accuracy, precision, recall, F1-score and confusion matrices. Random Forest produced the highest accuracy rate at 62% while XGBoost achieved 61%. XGBoost demonstrated excellent precision in identifying sudden braking events (Class 3). The SVM model demonstrated poor performance since it reached only 40% accuracy and struggled to distinguish between acceleration and braking events.

Table 1. Performance comparison of all models

Model	Accuracy	Precision (Macro)	Recall (Macro)	F1-Score (Macro)
LSTM	96%	0.95	0.96	0.96
GRU	92%	0.91	0.94	0.92
CNN	88%	0.88	0.89	0.88
XGBoost	61%	0.63	0.62	0.62
Random Forest	62%	0.64	0.63	0.64
SVM	40%	0.48	0.42	0.38

4.1. Deep Learning Models

The deep learning models achieved better results than the traditional methods. The LSTM model achieved an accuracy of 96%, effectively capturing temporal dependencies and maintaining a balanced precision and recall across all four classes. GRU followed with 92% accuracy, showing strength in detecting left-turn manoeuvres (Class 2). The CNN model achieved 88% accuracy, performing well in identifying acceleration and braking, though it was less effective in distinguishing between turning behaviours. A detailed comparison of all models is provided in Table 1, which highlights the superior performance of LSTM across all evaluation metrics.

4.2. Confusion Matrix Insights

The LSTM model demonstrated the most balanced classification results because it misclassified behaviour types very rarely. GRU and CNN showed occasional mistakes in turning action identification because their sensor signatures overlapped. All models received training through TensorFlow 2.14 in Google Colab's GPU environment to guarantee both reproducibility and computational speed. Moreover, XGBoost determined AccX and GyroZ as its most important features through permutation importance analysis, which confirmed their importance in measuring longitudinal acceleration and angular motion. The feature importance bar chart visually demonstrated how these critical dynamics enabled the detection of aggressive manoeuvres, including sudden acceleration, braking, and sharp turns.

5. Conclusion

This research proved that deep learning approaches particularly recurrent neural networks work well for aggressive driving behavior classification through time-series sensor data analysis. The LSTM model demonstrated the best performance among all evaluated models by achieving 96% accuracy and showing outstanding capabilities to understand temporal relationships in sequential motion data. GRU achieved similar performance results with reduced computational requirements which makes it appropriate for embedded system applications with limited resources. The CNN model showed consistent accuracy in detecting acceleration and braking behaviors even though its accuracy was slightly lower than other models. Deep learning models outperformed traditional models including Random Forest and XGBoost in this study yet XGBoost feature importance analysis revealed AccX and GyroZ as essential factors for detecting longitudinal and angular driving dynamics. The research shows promising results but it has certain boundaries. The study used a controlled dataset which included a restricted number of drivers and vehicles which may limit the models' ability to work in diverse real-world driving situations. The study used only four behavior labels which might not fully represent the actual complexity of driving behaviors. The future research should collect more diverse driving data and add GPS and OBD-II sensor information to improve behavioral context. The models need to be tested in real vehicles for real-time deployment to prove their suitability for use in advanced driver-assistance systems (ADAS) and insurance telematics solutions.

Acknowledgements

This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No.BR24992852 “Intelligent models and methods of Smart City digital ecosystem for sustainable development and the citizens’ quality of life improvement”).

References

[1] Al Abri, K. A., Jabeur, N., Gharrad, H., & Yasar, A. U. H. (2023). An intelligent divide-and-conquer approach for driving style management. *Personal and Ubiquitous Computing, 27*(5), 1729–1746. https://doi.org/10.1007/s00779-023-01740-1
[2] Behera, S., Bhardwaj, B., Rose, A., Hamdaan, M., & Ganesan, M. (2022). DriveSense: Adaptive system for driving behaviour analysis and ranking. In D. J. Hemanth (Ed.), *Machine learning techniques for smart city applications: Trends and solutions* (pp. 45–58). Springer International Publishing. https://doi.org/10.1007/978-3-031-08859-9_5
[3] Bibri, S. E., & Krogstie, J. (2017). On the social shaping dimensions of smart sustainable cities: A study in science, technology, and society. *Sustainable Cities and Society, 29*, 219–246. https://doi.org/10.1016/j.scs.2016.11.004
[4] Doniec, R. J., Sieciński, S., Duraj, K. M., Piaseczna, N. J., Mocny-Pachońska, K., & Tkacz, E. J. (2020). Recognition of drivers’ activity based on 1D convolutional neural network. *Electronics, 9*(12), 2002. https://doi.org/10.3390/electronics9122002
[5] Guo, F. (2019). Statistical methods for naturalistic driving studies. *Annual Review of Statistics and Its Application, 6*(1), 309–328. https://doi.org/10.1146/annurev-statistics-030718-105153
[6] Johnson, D. A., & Trivedi, M. M. (2011, October). Driving style recognition using a smartphone as a sensor platform. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 1609–1615). IEEE. https://doi.org/10.1109/ITSC.2011.6083078
[7] Toledo, T. (2007). Driving behaviour: Models and challenges. *Transport Reviews, 27*(1), 65–84. https://doi.org/10.1080/01441640600823940
[8] Darwish, K., & Ali, M. (2023). Driving behaviours recognition using deep neural networks. *Embedded Self-Organising Systems, 10*(5), 9–12. https://doi.org/10.14464/ess.v10i5.592
[9] Lee, J., & Jang, K. (2024). Characterizing driver behaviour using naturalistic driving data. *Accident Analysis & Prevention, 208*, 107779. 10.1016/j.aap.2024.107779

[10] Lv, Z., Zhang, S., & Xiu, W. (2020). Solving the security problem of intelligent transportation system with deep learning. *IEEE Transactions on Intelligent Transportation Systems, 22*(7), 4281–4290. https://doi.org/10.1109/TITS.2020.2980864
[11] Mantouka, E., Barmpounakis, E., Vlahogianni, E., & Golias, J. (2021). Smartphone sensing for understanding driving behaviour: Current practice and challenges. *International Journal of Transportation Science and Technology, 10*(3), 266–282. https://doi.org/10.1016/j.ijtst.2020.07.001
[12] Ossen, S., & Hoogendoorn, S. P. (2011). Heterogeneity in car-following behaviour: Theory and empirics. *Transportation Research Part C: Emerging Technologies, 19*(2), 182–195. https://doi.org/10.1016/j.trc.2010.05.006
[13] Saleh, K., Hossny, M., & Nahavandi, S. (2017). Driving behaviour classification based on sensor data fusion using LSTM recurrent neural networks [Conference paper]. In *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Yokohama, Japan, October 16–19, 2017 (pp. 1–6). IEEE. https://doi.org/10.1109/ITSC.2017.8317835
[14] Ślusarczyk, B., & Kozłowska, M. A. (2023). Levels of road safety in the European Union and its spatial diversity. *Humanities and Social Sciences, 30*(2), 121–133. https://doi.org/10.7862/rz.2023.hss.20
[15] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. (2014). Fusion of smartphone motion sensors for physical activity recognition. *Sensors, 14*(6), 10146–10176. https://doi.org/10.3390/s140610146
[16] Wang, J., Ravi, V., Flint, J., & Alwan, A. (2022). Unsupervised instance discriminative learning for depression detection from speech signals. In *Interspeech 2022* (Vol. 2022, pp. 2018–2022). International Speech Communication Association. https://doi.org/10.21437/interspeech.2022-10814
[17] World Health Organization. (2023). *Global status report on road safety 2023: Summary*. World Health Organization.
[18] Zhao, X., Yang, H., Yao, Y., Guo, M., & Chai, S. (2022). *Research on traffic safety evaluation method based on aggressive driving behaviours and traffic flow characteristics*. SSRN. https://doi.org/10.2139/ssrn.4167534
[19] Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2019). Deep learning for time series classification: a review. *Data mining and knowledge discovery*, *33*(4), 917-963.
[20] Karim, F., Majumdar, S., Darabi, H., & Chen, S. (2017). LSTM fully convolutional networks for time series classification. *IEEE access*, *6*, 1662-1669.