ISSN 2007-9737

OBD-II Driving Behavior Assessment: Analysis for Driving Styles Detection

Fabian Navarrete-Rocha*, Jorge I. Galván-Tejada, José M. Celaya-Padilla

Universidad Autónoma de Zacatecas, Unidad Académica de Ingeniería Eléctrica, Mexico

{fabiannavarrete, gatejo, jose.celaya}@uaz.edu.mx

Abstract. This paper presents a novel approach for driving style classification using exclusively On-Board Diagnostics (OBD-II) vehicle data. Key parameters including throttle position, revolutions per minute (RPM), and mass air flow (MAF), among others were collected via an ELM-327 adapter from a 2018 Honda HR-V along predefined urban routes driven by five participants. Applied to a dataset of 7,349 OBD-II observations collected from five drivers, K-means clustering with elbow method validation and a Silhouette Score of 0.63 identified three behaviorally meaningful driving profiles: normal, moderate, and dynamic, with k=3 selected based on its interpretability and alignment with the experimental conditions. The results demonstrate effective behavioral segmentation using only internal vehicle parameters, eliminating the need for external sensors. This method proves particularly valuable for cost-sensitive applications, offering reliable driver behavior characterization through widely accessible telematics. Future work will leverage the generated clusters as ground-truth labels for supervised learning algorithms — including Support Vector Machines (SVM), Random Forest, and K-Nearest Neighbors (KNN) — to develop real-time classification systems.

Keywords. Driving styles, road safety, machine learning.

1 Introduction

Driver behavior classification relies on pattern recognition methods applied to data obtained either from driving simulations or controlled field experiments, making data acquisition strategy and experimental design critical components of any effective identification model [10].

In recent years, driver behavior analysis has emerged as a key tool for enhancing road safety, optimizing driver assistance systems, and personalizing the driving experience. Current research has explored this domain through advanced technologies such as smartphone-embedded inertial sensors and emotion-detection cameras. However, many of these approaches face significant barriers to widespread adoption, including high implementation costs, considerable technological complexity, and difficulties in replication within resource-constrained environments [5]. In this context, second-generation On-Board Diagnostics (OBD-II) represents a more accessible, cost-effective, and broadly available alternative.

Although OBD-II has traditionally served mechanical diagnostics and fault-code retrieval purposes, its potential as a standalone data source for driving style characterization remains largely underexplored — particularly without reliance on supplementary technologies such as external cameras or additional sensors [7]. The advantages of OBD-II are notable: low acquisition cost, straightforward implementation, and strong scalability make it well-suited for large-scale deployment. Its passive data-collection capability is especially relevant, as it enables detailed driver behavior analysis without requiring real-time intervention or any modification of vehicle dynamics, resulting in a less intrusive and more user-accepted approach [1]. This study investigates the potential of using exclusively OBD-II port data to detect and characterize driving

behavior patterns. Variables such as RPM, vehicle speed, MAF, and throttle position provide direct insight into individual driving styles. While these metrics have been employed in prior research for similar purposes, they were typically combined with additional sensing technologies, diverting attention from the intrinsic analytical value of OBD-II data alone. Exploiting this data source as a standalone input represents an emerging and promising research direction. The overarching goal of this research is to objectively identify risk profiles and classify drivers using OBD-II data in conjunction with machine learning techniques, establishing a foundational framework for potential integration with Advanced Driver-Assistance Systems (ADAS). Such integration could enable personalized system interventions tailored to specific drivers and environments, thereby enhancing overall system effectiveness [3, 2].

The methodological simplicity of this approach offers meaningful advantages. Rather than competing with complex systems such as fully featured ADAS platforms, this work seeks to develop a robust, standalone framework for driver classification relying exclusively on OBD-II data — an underexplored field with substantial expansion potential. While future applications could include feeding behavioral data into broader systems or informing road safety policies, the autonomous classification capability introduced here already constitutes a relevant contribution to the study of human behavior in driving contexts.

Beyond its immediate applications in driver monitoring and fleet management, the framework proposed in this study holds significant relevance for the field of robotics and autonomous systems. Accurate driving style classification constitutes a foundational component for the development of intelligent navigation systems, adaptive driver assistance, and robot driver architectures. In the context of autonomous vehicles, the ability to recognize and adapt to human driving patterns enables more natural and socially compliant motion planning, facilitating smoother human-robot interaction in shared traffic environments. Furthermore, the behavioral profiles generated by this approach can serve as reference models for training autonomous agents to replicate or anticipate human driving decisions, contributing to the broader goal of developing robust, human-aware robotic mobility systems.

2 Related Work

The use of machine learning and deep learning techniques has substantially advanced the development of driver distraction detection and driving style monitoring methodologies [8]. Karrouchi et al. demonstrated that a multimodal driver behavior analysis system — combining CAN bus vehicle data (engine speed, acceleration, and fuel consumption) with facial recognition — achieves 99.1% accuracy in detecting aggressive driving styles and fatigue states. Their implementation of a universal monitoring unit and integrated analysis of seven driving parameters alongside facial signals established quantitative thresholds for driving style classification, validating the system's real-world applicability. These high-precision results highlight significant potential for enhancing road safety policies and informing targeted driver training programs [6].

Xia and Chen conducted a comprehensive survey of trajectory planning algorithms for autonomous vehicles, examining two distinct methods for estimating collision risks from sensory data, alongside techniques for generating human-preferred trajectories. Through a systematic review of recent literature, the authors classify and compare local planning approaches while identifying their distinctive features. The work incorporates behavior-driven strategies to improve trajectory naturalness and presents a multi-perspective comparative analysis revealing methodological interconnections, concluding with open challenges and future directions to improve system safety and efficacy in real-world traffic environments [9]. Zhang, Zhang, and Zhao proposed an intelligent driving style identification method using Next-Generation Simulation (NGSIM) data, addressing limitations of prior approaches that tend to analyze specific scenarios in a fragmented manner.

Their study classified car-following and lane-changing behaviors through driving state extraction and characteristic parameter

Table 1. Comparative analysis of related driving style classification approaches

Reference	Data source	Method	Driving styles	External sensors	Real-world data	Cost
Karrouchi et al. [6]	CAN Bus + facial recognition	Multimodal analysis	Aggressive, fatigue	Yes	Yes	High
Zhang et al. [11]	NGSIM trajectory data	PCA + K-means, SVM, RF, KNN	Conservative, moderate, aggressive	Yes	No	Medium
Huang et al. [4]	NGSIM naturalistic data	Inverse reinforcement learning	Personalized	Yes	No	High
Kumar & Jain [7]	OBD-II	Machine learning	Multiple	No	Yes	Low
Proposed work	**OBD-II only**	**K-means, Elbow Method**	**Normal, moderate, dynamic**	**No**	**Yes**	**Low**

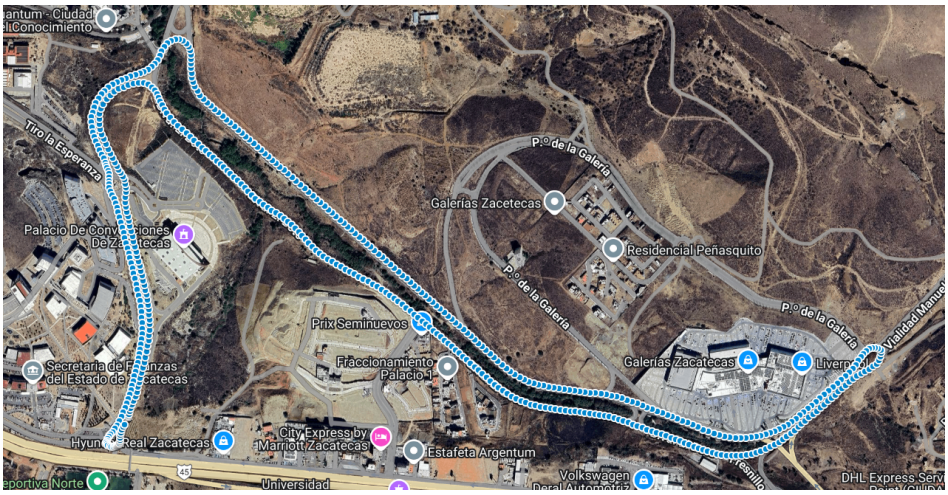

Fig. 1. Satellite view of the defined 6 km route in Zacatecas

optimization, employing correlation analysis and Principal Component Analysis (PCA) for dimensionality reduction. The K-means algorithm was subsequently applied to cluster driving styles into three categories: conservative, moderate, and aggressive. Classification performance was evaluated using SVM, Random Forest, and KNN, with SVM achieving superior accuracy (>97%). These results validate the efficacy of a naturalistic data-driven approach, with practical implications for driver training, cost optimization, and personalized autonomous driving systems [11].

Huang, Wu, and Lv proposed a driving behavior modeling framework employing Inverse Reinforcement Learning (IRL) to derive personalized reward functions from naturalistic NGSIM driving data. The methodology addresses the continuous-state problem through latent intention discretization, enabling effective modeling of complex driving behaviors. The system combines a polynomial trajectory sampler with an interactive vehicle model, demonstrating robust capability in capturing individual driver preferences with only marginal precision degradation during testing. Comparative analyses confirmed the framework's superiority over conventional approaches, particularly in vehicle interaction scenarios requiring reliable trajectory planning — with significant implications for the development

of advanced driver assistance and autonomous vehicle control systems [4].

Table 1 summarizes a comparative analysis of the most relevant related works alongside the proposed approach. As evidenced, while prior studies achieve high classification accuracy, they predominantly rely on external sensors, simulation-based datasets, or multimodal data sources that increase implementation complexity and cost. In contrast, the proposed framework leverages exclusively OBD-II internal vehicle parameters collected under real-world conditions, offering a low-cost, scalable, and non-intrusive alternative that addresses a gap in the existing literature.

Table 2. Data collected through OBD-II

Variable	Description	Units
THROTTLE POS	Throttle position	%
SPEED	Vehicle speed	km/h
RELATIVE THROTTLE POS	Relative throttle position	%
RPM	Revolutions per minute	rpm
ENGINE LOAD	Engine load	%
CATALYST TEMP B1S1	Catalyst Temperature	°C
INTAKE PRESSURE	Intake air pressure	kPa
MAF	Mass air flow	g/s
ABSOLUTE LOAD	Absolute engine load	%
Class	Driving Style	1–3

3 Materials and Methods

3.1 Data Collection

The experiment was designed to maintain a controlled environment in which driving conditions were not affected by heavy traffic, allowing meaningful differences across driving styles to be observed and captured. Data collection was conducted using a single common vehicle and accessible OBD-II-compatible technology.

All data were gathered along a predefined 6 km urban route (See Fig. 1), selected to minimize the influence of external factors such as traffic congestion. Each participating driver completed three distinct experimental conditions:

1. Maintaining a speed between 40 and 50 km/h, representing a normal driving style.
2. Maintaining a speed between 50 and 60 km/h, representing a moderate driving style.
3. Maintaining a speed between 60 and 70 km/h, representing a dynamic driving style.

In all conditions, each driver completed the route three times to ensure behavioral repeatability and generate sufficient data for reliable classification. Five drivers with diverse profiles were selected to participate under identical environmental conditions.

Participants were instructed to respect speed limits and the natural flow of traffic without altering their habitual driving style, thereby enabling the capture of a naturalistic behavioral range under comparable circumstances. During each trip, an ELM-327 OBD-II adapter transmitted data in real time to a laptop running a custom Python script, which collected and stored the parameters listed in Table 2.

It is worth noting that the complete dataset comprises recordings from 21 drivers with diverse profiles regardless of age, experience, or gender. However, a deliberate decision was made to conduct the present analysis using a representative subset of five participants. This choice was motivated by two complementary objectives: first, to assess the reproducibility of the experimental protocol and validate that meaningful behavioral segmentation can be achieved even under minimal sampling conditions; and second, to demonstrate the feasibility of the proposed framework as a proof-of-concept before scaling to the full dataset. The diversity of the selected participants, spanning different driving profiles, further supports the representativeness of this subset. The complete 21-driver dataset will be incorporated in future work to evaluate the robustness and generalizability of the classification model across a broader population.

The variables obtained via the OBD-II port during driving tests enable precise driver behavior

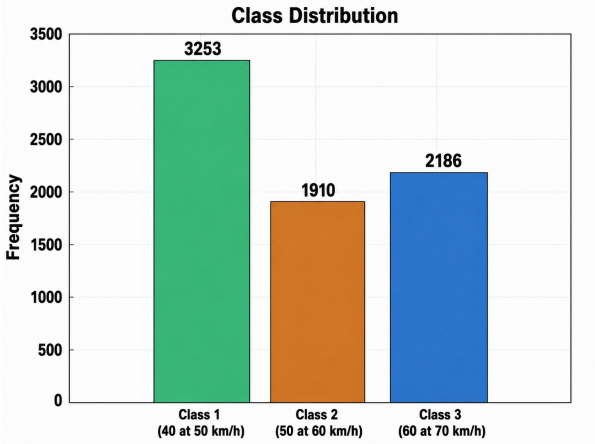

Fig. 2. Class distribution

profiling by recording real-time vehicle interactions across diverse driving conditions. Below is a technical description of each parameter:

— **Throttle Position - %:** Quantifies accelerator pedal engagement, reflecting power demand. Abrupt variations or sustained high values may indicate aggressive driving.

— **Vehicle Speed - km/h:** A core behavioral indicator; persistent speed limit violations or sudden fluctuations suggest risk-prone behavior.

— **Relative Throttle Position - %:** Measures accelerator position within its operational range, revealing power delivery patterns, progressive vs. aggressive.

— **Revolutions Per Minute - rpm:** Engine speed patterns identify gear-shift habits. Sustained high RPM correlates with sportive/aggressive driving.

— **Engine Load - %:** Represents motor exertion intensity. Chronic high loads typify forceful driving styles.

— **Catalyst Temperature - °C:** While primarily engine-related, elevated temperatures may result from rapid acceleration or sustained high-speed operation.

— **Intake Air Pressure - kPa:** Manifold air pressure fluctuations reflect power demand dynamics, with frequent peaks suggesting strenuous driving.

— **Mass Air Flow - g/s:** Air intake volume, combined with RPM and load, estimates fuel efficiency and driving smoothness.

— **Absolute Engine Load - %:** Normalized motor effort relative to maximum capacity, indicating driving intensity.

— **Class Driving Style - 1 to 3:** Behavioral classification label, e.g., 1: Normal (40–50 km/h), 2: Moderate (50–60 km/h), 3: Dynamic Style (60–70 km/h).

During the experimental data acquisition phase involving five participating drivers, three distinct driving style classes were operationally defined based on predetermined threshold parameters. The class distribution across collected instances is presented in Fig. 2. It is important to note that the Class variable included in Table 2 serves solely as a reference label assigned during data collection to identify the experimental condition under which each observation was recorded.

This variable was explicitly excluded from the feature set prior to applying the K-means clustering algorithm, ensuring that the unsupervised classification was driven entirely by the intrinsic structure of the OBD-II parameters. Following data collection and database generation, a systematic methodological pipeline was implemented. The process began with comprehensive data cleaning to remove extraneous characters, incomplete records, and outliers that could compromise analytical integrity. Upon completion of data purification, the K-means clustering algorithm was applied to identify common behavioral patterns and categorize driving styles into distinct groups. The complete methodological workflow is illustrated in Fig. 3.

3.2 Data Preprocessing

During the preprocessing stage, data cleaning was performed on the raw information collected through the OBD-II system. This phase involved:

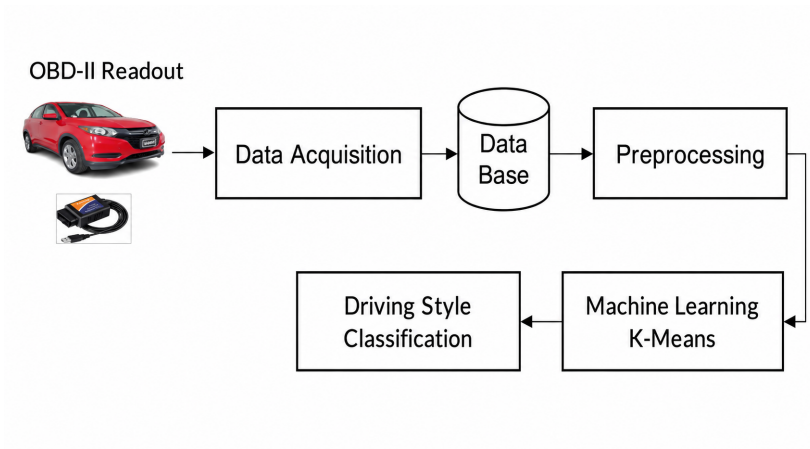

Fig. 3. Methodology implemented for driving classification

1. Removal of non-numeric characters embedded within data fields, particularly measurement unit indicators (e.g., “km/h”, “%”, “rpm”) that were stored alongside numerical values.
2. Text sanitization through string processing techniques designed to isolate and preserve only numeric values.
3. Retention of all collected features, as all variables described in Table 2 were deemed relevant for subsequent analysis.

These measures addressed a critical data quality issue: the presence of unit descriptors within numeric fields prevented proper quantitative conversion and downstream processing. The cleansing procedure ensured dataset integrity and full compatibility with the analytical algorithms applied in subsequent stages.

Following data cleaning, a normalization procedure was implemented to ensure that all variables contributed equally to the clustering model. This step is particularly critical when using distance-based algorithms such as K-means, where features with larger numerical ranges can disproportionately influence results if left unscaled. The selected technique was Z-score standardization, which transforms each observation according to its deviation from the feature mean, as expressed in Equation (1):

$$Z_i = \frac{x_i - \mu}{\sigma}. \tag{1}$$

This transformation produces variables with zero mean and unit standard deviation ($\mu = 0, \sigma = 1$), effectively eliminating scale bias introduced by differing measurement units or physical magnitudes. Although alternative normalization methods exist—such as min-max scaling, which constrains values to the [0, 1] interval—Z-score standardization demonstrates superior robustness in the presence of outliers, a particularly valuable property in driving data analysis where anomalous recordings may arise from exceptional driving events or sensor malfunctions.

3.3 Machine Learning

For driving style classification, unsupervised machine learning techniques were applied, specifically:

— The K-means clustering algorithm, to identify and group behavioral patterns within the dataset.

— The Elbow Method, to objectively determine the optimal number of clusters.

Since no pre-existing labels were available for the collected data, an unsupervised approach was selected to enable unbiased segmentation of driving behavior based solely on the intrinsic structure of the OBD-II variables. The elbow method was used to identify the inflection point at which improvements in within-cluster cohesion begin to diminish, providing an objective criterion for selecting the most representative number of clusters.

3.4 K-means

K-means is a widely adopted unsupervised learning algorithm that partitions a dataset into k distinct groups, assigning each observation to the cluster whose centroid is nearest in feature space. The algorithm follows four steps:

1. Initialization: k centroids are randomly selected from the observation space as provisional cluster centers.
2. Assignment: Each data point x_i is allocated to the nearest centroid c_j according to the Euclidean distance metric:

$$d(x_i, c_j) = \sqrt{\sum_{n=1}^{m} (x_{in} - c_{jn})^2}, \tag{2}$$

   where m denotes the dimensionality of the feature space.
3. Update: Each centroid is recalculated as the mean of all data points assigned to its cluster:

$$c_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i. \tag{3}$$

4. Iteration: Steps 2 and 3 are repeated until centroid positions stabilize, i.e., until $\Delta c_j < \epsilon$.

K-means offers two principal advantages for driving behavior analysis: computational efficiency and scalability to high-dimensional datasets, including OBD-II telemetry streams. In the present study, this algorithm successfully identified latent driving patterns across participants, enabling representative cluster formation without requiring pre-labeled training data.

3.5 Elbow Method

The elbow method is a standard technique for determining the optimal number of clusters k in partitioning algorithms such as K-means. It evaluates the Within-Cluster Sum of Squares (WCSS) as a function of increasing cluster count, computed as follows:

$$WCSS = \sum_{j=1}^{k} \sum_{x_i \in C_j} \|x_i - c_j\|^2, \tag{4}$$

where x_i is a data point belonging to cluster C_j, c_j is the corresponding centroid, and $\|x_i - c_j\|^2$ is the squared Euclidean distance between the point and its centroid.

As k increases, WCSS decreases monotonically, since a greater number of clusters reduces the average distance between points and their centroids. However, beyond a certain value of k, the marginal reduction in WCSS becomes negligible. This inflection point—where the WCSS curve transitions from steep descent to near-plateau—forms a characteristic “elbow” shape and is interpreted as the optimal cluster count. This criterion balances model accuracy with parsimony, avoiding both under-segmentation (too few clusters) and over-segmentation (too many clusters). In this study, the elbow method played a central role in objectively determining the number of driving style groups present in the collected data.

It is acknowledged that K-means assumes statistical independence among observations, thereby disregarding the sequential and temporal nature inherent to driving data. Temporal dynamics—such as the sustained duration of abrupt accelerations or prolonged high-RPM episodes—are not captured by this static clustering approach.

This represents a recognized limitation of the present framework, and future work will explore time-aware modeling strategies, such as Hidden Markov Models or recurrent neural network architectures, to incorporate temporal dependencies into driving behavior characterization.

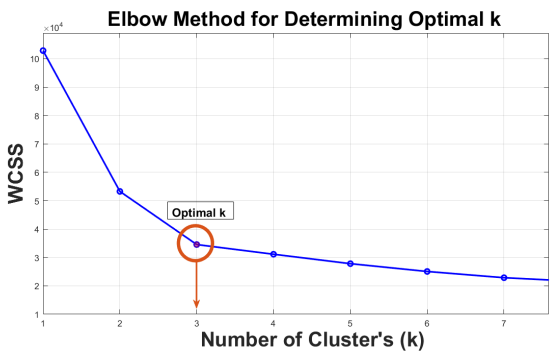

Fig. 4. Elbow method for determining optimal k

4 Results

To analyze driving styles, K-means clustering was applied to the normalized OBD-II dataset. The algorithm grouped recorded driving patterns into homogeneous categories by identifying structural similarities across observations.

Table 3. Cluster metrics

Cluster	THROTTLE POS (%)	RPM	MAF (g/s)	Classification
1	20.03	1377	7.97	Normal
2	21.26	1486	9.05	Moderate
3	23.48	1593	10.73	Dynamic

The variables considered—including vehicle speed, RPM, throttle position, engine load, and mass air flow—collectively enabled the classification of drivers based on their behind-the-wheel behavior, yielding three interpretable and distinct driving profiles: normal, moderate, and dynamic.

The optimal number of clusters was determined using the elbow method, which evaluates the variation in WCSS across candidate values of k. As shown in Fig. 4, the WCSS curve exhibits a clear inflection point at $k = 3$, beyond which further increases in cluster count yield only marginal reductions in within-cluster dispersion. This result indicates that three clusters effectively capture the natural structure of the data while maintaining model simplicity. Consequently, K-means was configured with $k = 3$ for all subsequent analyses.

The observed pattern confirms that additional clusters beyond $k = 3$ would produce diminishing returns in terms of within-cluster cohesion, offering no meaningful improvement in behavioral discrimination. The elbow method's identification of $k = 3$ thus provides a data-driven and interpretable foundation for the driving style classification framework.

Prior to presenting the clustering results, a correlation matrix was constructed to examine linear relationships among the collected OBD-II variables Fig. 5. This analytical step facilitates the interpretation of how specific vehicle parameters jointly reflect driving behavior patterns and provides context for understanding the characteristics of each emergent cluster.

The analysis revealed strong inter-variable dependencies within two primary groups. First, THROTTLE POS and RELATIVE THROTTLE POS exhibited near-perfect correlation ($r = 0.9966$), with both also showing strong correlation with MAF ($r > 0.95$), indicating that these features collectively capture engine power demand. Second, ENGINE LOAD demonstrated high correlation with both INTAKE PRESSURE ($r = 0.9752$) and ABSOLUTE LOAD ($r = 0.9293$), consistent with expected engine behavior under varying load conditions. In contrast, SPEED showed minimal correlation with the remaining features, suggesting it provides independent and complementary information to the model.

These findings informed the subsequent feature selection strategy: rather than using all variables redundantly, representative features were selected from each correlated group to reduce dimensionality while preserving explanatory power. The variables retained for cluster interpretation were THROTTLE POS, throttle engagement; RPM, engine exertion; ENGINE LOAD and ABSOLUTE LOAD, overall engine workload; SPEED, independent behavioral indicator with direct road safety relevance; and MAF, air intake volume, serving as a proxy for fuel consumption and driving smoothness.

Table 3 presents the mean values of three representative variables—THROTTLE POS, RPM, and MAF—for each identified cluster, along with the corresponding behavioral classification.

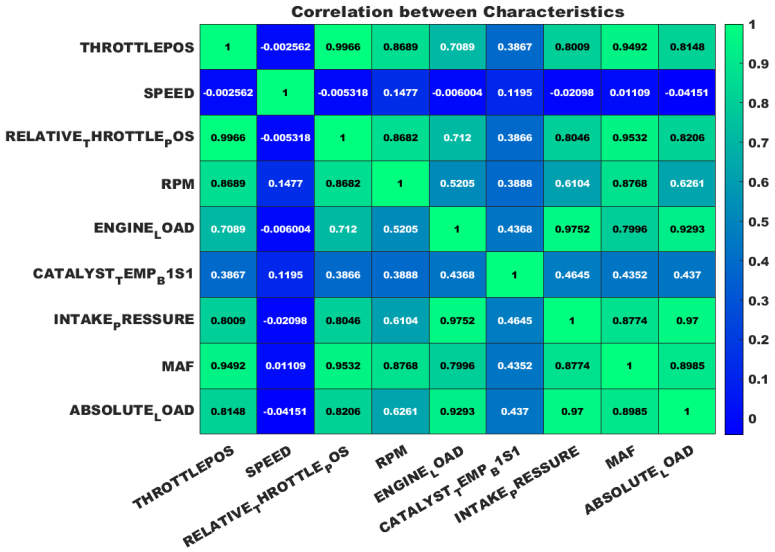

Fig. 5. Correlation analysis of vehicle parameters

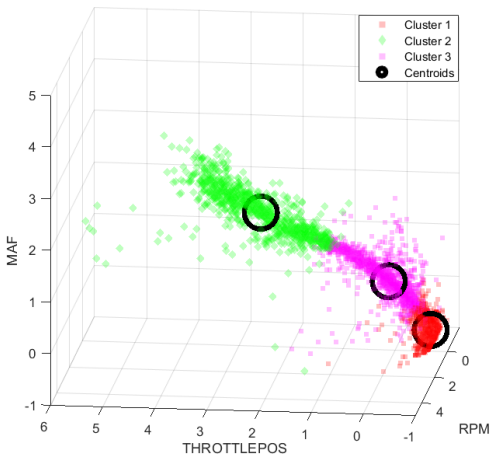

Fig. 6. K-means clusters with centroids

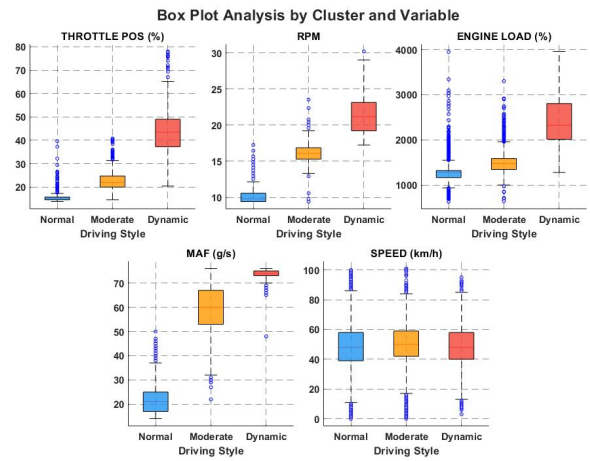

Fig. 7. Box plot analysis

The three clusters correspond to clearly distinct driving profiles. Cluster 1, Normal, exhibited the lowest values across all key parameters, reflecting conservative throttle use, low engine exertion, and minimal air flow—characteristic of fuel-efficient, low-intensity driving.

Cluster 2, Moderate, displayed intermediate values, with balanced increases in throttle position and RPM, indicating an adaptive driving style that transitions between economy and performance.

Cluster 3, Dynamic, showed the highest parameter values across all metrics, with elevated throttle engagement, peak RPM, and maximum MAF, reflecting aggressive acceleration patterns and sustained high-intensity engine operation.

These results confirm the model's effectiveness in classifying the full behavioral spectrum—from

economical to performance-oriented driving—using exclusively OBD-II data. The cluster separation is further reinforced by the 3D visualization presented in Fig. 6, which projects the data along the three most discriminative axes, THROTTLE POS, RPM, and MAF, and visually validates the coherence and separation of the identified groups.

To further characterize the distribution of key OBD-II parameters across the identified clusters, Fig. 7 presents box plots for the five most representative variables: throttle position, RPM, engine load, MAF, and vehicle speed. The visualizations confirm the progressive behavioral separation across clusters, with the Dynamic style consistently exhibiting higher medians and greater variability in engine-related parameters, while the Normal style shows compact distributions centered at lower operational values. The Moderate cluster occupies an intermediate position across all variables, reinforcing the behavioral gradient captured by the K-means model.

5 Conclusions and Future Work

This study successfully demonstrated that driving styles can be effectively identified and clustered using exclusively OBD-II port variables, specifically throttle position, RPM, and mass air flow, without the need for external sensors or supplementary technologies. The K-means clustering algorithm, validated through the elbow method and supported by a Silhouette Score of 0.63 at $k = 3$, consistently identified three behaviorally meaningful and well-separated driving profiles: normal, moderate, and dynamic. Correlation analysis further supported optimal feature selection, reducing redundancy while preserving the model's explanatory power. Taken together, these findings confirm the viability of unsupervised learning as a standalone framework for driver behavior characterization in cost-sensitive and resource-constrained applications.

From a robotics perspective, the behavioral profiles generated by this framework represent a foundational contribution toward the development of intelligent autonomous systems. The ability to classify human driving styles in real time constitutes a critical enabler for adaptive motion planning in autonomous vehicles, human-aware robot driver architectures, and intelligent navigation systems capable of anticipating and replicating naturalistic driving decisions. Integration of this classification framework with ADAS platforms could further enable context-aware, personalized system interventions tailored to individual driver profiles.

From a practical standpoint, this approach offers a scalable and non-intrusive solution for driver monitoring that leverages widely available OBD-II hardware, making it particularly well-suited for large-scale deployment in fleet management, insurance telematics, and driver training programs.

Several limitations of the present study should be acknowledged. First, the current results are based on a representative subset of five drivers from a complete dataset of 21 participants; while this subset was sufficient to validate the reproducibility of the experimental protocol, broader generalization requires evaluation across the full dataset. Second, the experimental design involved predefined speed ranges, which may influence the natural separation of clusters; future work should assess the framework's performance under fully unconstrained naturalistic driving conditions. Third, the static nature of K-means does not capture temporal dynamics inherent to driving behavior, a limitation that will be addressed through time-aware modeling approaches in subsequent studies.

Future work will focus on integrating the unsupervised clusters generated in this study as ground-truth labels for supervised classification algorithms, including SVM, Random Forest, and KNN, to evaluate predictive power and generalizability in real-time classification scenarios. Additionally, expanding the dataset to include a greater number of drivers, vehicle types, and route conditions will be essential to assess the robustness and external validity of the classification model.

References

1. **Chang, H., Park, D. (2020).** Potentialities of vehicle trajectory big data for monitoring

potentially fatigued drivers and explaining vehicle crashes on motorway sections. Sustainability, Vol. 12, No. 15. DOI: 10.3390/su12155877.

2. **Garrosa, M., Olmeda, E., Fuentes del Toro, S., Díaz, V. (2021).** Holistic vehicle instrumentation for assessing driver driving styles. Sensors, Vol. 21, No. 4. DOI: 10.3390/s21041427.

3. **He, Y., Yang, S., Zhou, X., Lu, X. Y. (2022).** An individual driving behavior portrait approach for professional driver of hdvs with naturalistic driving data. Computational Intelligence and Neuroscience, Vol. 2022, pp. 3970571. DOI: 10.1155/2022/3970571.

4. **Huang, Z., Wu, J., Lv, C. (2021).** Driving behavior modeling using naturalistic human driving data with inverse reinforcement learning. IEEE Transactions on Intelligent Transportation Systems. DOI: 10.1109/TITS.2021.3088935.

5. **Jia, Y., Zhang, Z., Li, X., Chen, X., Qu, D. (2025).** Driving style tendency quantification method based on short-term lane change feature extraction. Sustainability, Vol. 17, No. 8. DOI: 10.3390/su17083563.

6. **Karrouchi, M., Nasri, I., Rhiat, M., Atmane, I., Hirech, K., Messaoudi, A., Melhaoui, M., Kassmi, K. (2023).** Driving behavior assessment: a practical study and technique for detecting a driver's condition and driving style. Transportation Engineering, Vol. 14, pp. 100217. DOI: 10.1016/j.treng.2023.100217.

7. **Kumar, R., Jain, A. (2023).** Driving behavior analysis and classification by vehicle obd data using machine learning. The Journal of Supercomputing, Vol. 79, No. 16, pp. 18800–18819. DOI: 10.1007/s11227-023-05364-3.

8. **Sangwan, A. K., Sangwan, D. (2021).** A computer vision-based approach for driver distraction recognition using deep learning and genetic algorithm based ensemble. CoRR, Vol. abs/2107.13355. DOI: 10.48550/arXiv.2107.13355.

9. **Xia, T., Chen, H. (2024).** A survey of autonomous vehicle behaviors: trajectory planning algorithms, sensed collision risks, and user expectations. Sensors, Vol. 24, No. 15. DOI: 10.3390/s24154808.

10. **Yang, G., Ridgeway, C., Miller, A., Sarkar, A. (2024).** Comprehensive assessment of artificial intelligence tools for driver monitoring and analyzing safety-critical events in vehicles. Sensors, Vol. 24, No. 8. DOI: 10.3390/s24082478.

11. **Zhang, S., Zhang, Z., Zhao, C. (2024).** A method of intelligent driving-style recognition using natural driving data. Applied Sciences, Vol. 14, No. 22. DOI: 10.3390/app142210601.

Article received on 10/03/2026; accepted on 18/07/2026.
**Corresponding author is Fabian Navarrete-Rocha.*