Article

Not peer-reviewed version

Driver Behavior Profiling Through Jerk Dynamics and Statistical IMU Descriptors

Danut Dragos Damian , Felicia Michis , Luminita Moraru *

Posted Date: 29 April 2026

doi: 10.20944/preprints202604.2048.v1

Keywords: driving profiling; IMU data; jerk; jerk-related features; driving score; classification

 doi:10.20944/preprints202604.2048.v1

Article

Driver Behavior Profiling Through Jerk Dynamics and Statistical IMU Descriptors

Danut Dragos Damian [1], Felicia Michis [2] and Luminita Moraru [1,3,4,*]

[1] The Modelling & Simulation Laboratory, Dunarea de Jos University of Galati, 800123 Galati, Romania
[2] Emil Racovita High School of Galati, 800340 Galati, Romania
[3] Karadeniz Technical University, Faculty of Science, Department of Physics, Trabzon 61080, Türkiye
[4] Department of Physics, School of Science and Technology, Sefako Makgatho Health Sciences University, Medunsa 0204, South Africa
* Correspondence: luminita.moraru@ugal.ro

Abstract

This study proposes a transparent, data-driven framework for behavior recognition based exclusively on IMU measurements, hypothesizing that vehicular jerk-based features can help in differentiating driving behavior. Unlike studies relying on direct jerk values, our approach derives novel findings from jerk-based features. For rolling windows of 300 samples, a comprehensive set of statistical and dynamic descriptors is extracted, including amplitude, variance, standard deviation, coefficient of variation, standard error, skewness, and kurtosis, as well as jerk-based features such as jerk_std, jerk_variance, jerk_amplitude, and jerk_spikes. Statistical analysis is used to identify features with strong discriminative power. Effect sizes, measured by Cohen's d, quantify the difference between normal and aggressive driving styles. The selected features are used to compute the Driving Score (DS) and provide a driver's profile. Experimental results reveal a correlation between lower DS scores (<50) and windows characterized by high jerk variability, large amplitude fluctuations, and frequent spikes. Conversely, higher DS scores (>70) indicate smooth and stable motion patterns. The robustness of the proposed framework is evaluated using several machine learning classifiers as baselines, with the most important jerk-based features as inputs. For the aggressive driver class, the DBS model reports a Recall of 0.952 and an F1 of 0.925. For the normal driver class, the DBS model reports a Recall of 0.839 and an F1 of 0.879. The model has a total accuracy of 0.907. Also, Logistic Regression and ensemble models like XGB and RF perform well. The proposed framework offers an explainable, computationally efficient alternative to conventional machine-learning classifiers for identifying aggressive drivers.

Keywords: driving profiling; IMU data; jerk; jerk-related features; driving score; classification

---

1. Introduction

Driving behavior significantly influences (positively or negatively) both road safety and the environment. A key component of modern road safety is responsible driving, reflected through a driver's ability to anticipate, adapt, and maintain vigilance during travel. Despite the continuous development of intelligent vehicle technologies such as adaptive cruise control, automated braking, lane-keeping systems, and real-time driver monitoring, human behavior remains the dominant factor influencing crash occurrence. Speed is a variable parameter directly linked to the driver, especially during vehicular jerk and traffic conflicts. According to global evaluations, aggressive maneuvers, instability in motion, and abrupt changes in acceleration remain strongly associated with unsafe driving patterns, which can be effectively detected using motion-derived indicators such as jerk, amplitude variability, skewness, and kurtosis [1–6].

This work offers a framework beyond machine learning and deep learning solutions for classifying driver behavior. These approaches often rely on complex representations or heavily

 doi:10.20944/preprints202604.2048.v1

preprocessed data, which can reduce interpretability. In contrast, statistical IMU descriptors such as amplitude variance, standard deviation, skewness, kurtosis, and jerk-based dynamics provide an interpretable and theoretically grounded way to describe driving behavior.

Given that driver profiling is a multifaceted approach, investigating and developing new quantitative statistical measures that capture aspects of driving style could be beneficial. Motivated by these findings, this work introduces a transparent and data-driven methodology for profiling driver behavior using the Driving Behavior Scoring algorithm. The approach relies on IMU signals and computes a comprehensive set of statistical and dynamic descriptors without any filtering stage. In a first step, the algorithm segments the data into sliding windows and calculates jerk along the three axes. Next, it extracts key features including amplitude, variance, standard deviation, coefficient of variation, standard error, skewness, kurtosis, and jerk-related features (jerk_std, jerk_variance, jerk_amplitude, jerk_spikes). These features are normalized using Min–Max scaling and combined into a behavioral score ranging from 0 to 100 that reflects the degree of calmness or aggressiveness for each driving window. A threshold of 50 has been selected to classify behavior as normal or aggressive. It was set based on the jerk distributions from the empirical data (detailed below). Finally, a classification is performed as benchmarking to validate the robustness of the proposed approach. This study addresses a gap in the current literature. While many existing studies rely heavily on jerk parameters, several practical barriers still hinder the timely, consistent, and scalable assessment. First, relatively little research has explored statistical and dynamic descriptors based on jerk-based feature dynamics. This approach relies on the understanding that driver data in different datasets exhibit varying driving styles. When large positive or negative jerk values are present, a uniform approach becomes difficult. Consequently, using dynamic descriptors based on jerk-related features offers a novel and effective alternative. Second, due to the high variability of raw data, we assess the robustness of the proposed approach using various machine learning models as a baseline, in an intra- study on both datasets. Unlike studies relying on direct jerk values, our approach derives novel findings from jerk-based features. This study makes the following key contributions:

(1) Present a methodological framework for quantifying driving behavior using interpretable IMU-based features grounded in well-established statistical principles. The significance of feature selection from IMU data is explored in relation to its impact on driver behavior.

(2) Evaluate the variations and interpretability of statistical features in driver profiling based on the Driving Score (DS) measure. Using statistical analysis, the proposed driving scoring approach improves the transparency and interpretability of drivers' behavior. This goes beyond feature-based analysis or neural network classification, which usually lack explainability.

(3) Demonstrate robustness of jerk-based feature dynamics by employing machine learning baselines across multiple metrics.

(4) Conduct a thorough within- and intra-study evaluation across both datasets.

This approach addresses the increased need for robust and transparent monitoring strategies that complement intelligent transportation systems and support the design of predictive, personalized safety solutions.

2. Related Works

As transportation systems become more automated, advanced analytics are essential for detecting subtle deviations in driving styles. Smartphone-based sensors, inertial measurement units, and on-board telematics have made real-time vehicle dynamics monitoring affordable. This continuous stream of measurable behavioral signatures is critical [7]. Moreover, despite existing methods for classifying driving behavior based on relative power acceleration [8] and vehicular jerk thresholds [9], no research has examined statistical differences and influence areas using real-world data. The greater variability in accelerometer signals, along with increased asymmetry (skewness), higher kurtosis, and notable jerk fluctuations, is directly linked to risky behaviors, including harsh braking, aggressive acceleration, lane instability, and quick steering corrections [10–12]. To assess a

 doi:10.20944/preprints202604.2048.v1

driver's behavior and operational performance, it is important to focus on features that accurately reflect their impact. One essential characteristic is the jerk effect, which results from sudden acceleration and deceleration. Jerk values measure how quickly accelerations change, effectively indicating how smoothly a driver operates. Hayati et al. [10] provide a comprehensive review of jerk's relevance in university science and engineering. The discussion focuses on jerk's usefulness in traditional land-based vehicles and in anti-jerk controller design for autonomous vehicles. Extensive research has dealt with driving behavior parameters. They were motivated by the need for transparent analytical models capable of interpreting IMU signals and deriving meaningful behavioral categories [13–16]. Redhu & Siwach [13] investigated the effect of traffic jerk using linear stability analysis. Jerk, a factor contributing to traffic congestion, was analyzed using a lattice hydrodynamics model. This model examined vehicle braking and acceleration patterns and their dependence on the jerk parameter. They found that the jerk parameter significantly contributes to traffic jams and reducing it requires a greater emphasis on driver anticipation and normal behavior. To enhance driving safety, comfort, and operability in complex urban environments, Sun et al. [15] developed a framework for verifying online driving styles through a personalized intention-aware automated driving strategy. They evaluated driving style using various longitudinal stimuli and applied different weight coefficients to classify it as steady, general, or radical. Komavec et al. [17] evaluated driver performance and the likelihood of unsafe driving using data from a simulated driving session. They proposed a risk assessment score to estimate a driver's propensity for risky behavior. Two machine learning models were then employed to classify drivers as either risky or non-risky. Machine learning algorithms are especially effective for risk assessment when the driving score stays interpretable and offers valuable feedback on driver profiling, but there is a lack of general explainability. Feng et al. [18] explored the usefulness of vehicle longitudinal jerk in identifying aggressive drivers. Their findings showed aggressive drivers had significantly higher values for both positive and negative jerk-based metrics. They concluded that a large negative jerk is effective in identifying aggressive drivers. Medarevic et al. [19] proposed a rule-based Driver Scoring System model to analyze driving simulator data and identify driver profiles. Their approach involves clustering distinct driver profiles. Three clusters are formed, and similar driving behavior reveals the driver profile (less or more aggressive). Tawadros et al. [20] proposed a method based on torque computation for jerk estimation. Automation usually leads to increased jerk and slower shift times compared to a skilled driver. They proposed a wireless torque measurement device to estimate the jerk characteristic. They conducted extensive simulations and experimental measurements and concluded that a low-cost torque sensor and Bluetooth communication could quantify the experimentally derived jerk. El Mourabit et al. [21] proposed a framework for Instantaneous Time-to-Collision computation based on estimated relative distance, velocity, acceleration, and jerk to assess collision risk. They used a constant-jerk model to describe the changing velocities and accelerations as significant factors influencing drivers' behavior.

Despite these advances, assessing drivers' dangerous tendencies remains limited to questionnaires, expensive measurement tools, or intensive computational algorithms built based on advanced machine learning/deep learning models. Existing approaches rarely address statistical and dynamic descriptors, especially in scenarios involving jerk-based feature dynamics.

3. Materials and Methods

This paper aims to conduct a detailed analysis of IMU data to recognize different driving behaviors and classify them into two categories: normal (labeled 0) and aggressive (labeled 1).

3.1. Data Quality

This study also aims to evaluate the impact of data quality on publicly available datasets used to analyze driver behavior scoring systems. We therefore selected two datasets. Data quality significantly influences key driver behavioral patterns. Two key factors significantly impacted data quality. Firstly, the reliability of the captured data can be compromised by inherent noise,

 doi:10.20944/preprints202604.2048.v1

necessitating additional validation and scrutiny. Secondly, the technical characteristics of the IMUs, i.e., specifically the three-axis linear accelerometers, play a crucial role in data quality. This study uses two multi-center datasets to improve generalization, diversity, and heterogeneity. These datasets include IMU information and data labelling. We used data from both the Mendeley [22] and Kaggle [23] datasets.

The Mendeley dataset contains information like longitude, latitude, speed, distance, time, heading, accelerometer (Acc_X, Acc_Y, Acc_Z), and gyroscope signals recorded using IMU sensors. It also includes a label column indicating driving behavior (0 for normal and 1 for aggressive). The Mendeley dataset yielded 140 rolling/sliding windows for accelerometer signals.

The Kaggle dataset contained smartphone sensor data from Android devices, specifically accelerometer and gyroscope readings. This dataset was collected under realistic traffic conditions. The experiment involved driving the same stretch of road at three different speeds: slow, normal, and aggressive. However, the slow data was removed from the current study. The Kaggle dataset case used 41 rolling/sliding windows for accelerometer signals.

For both datasets, the instance numbers were 1927 for aggressive driving and 2197 for normal driving. The balance between classes is kept. Each window is a time series containing 300 samples. The obtained time-series windows contain data from both normal and aggressive drivers. For each window, 42 features were computed. The initial feature datasets contain 80,934 data points for aggressive driving and 92,274 for normal driving. After the Driving Score selection procedure (detailed below), a reduced subset of 22 statistically significant features remained. This results in 3080 samples for the Mendeley dataset and 902 for the Kaggle dataset. A further reduction was performed when only jerk-based features (jerk_std_x, jerk_variance_x, jerk_amplitude_x, jerk_spikes_x, jerk_std_y, jerk_max_z, jerk_amplitude_z) were considered for classification, corresponding to 980 samples for the Mendeley dataset and 140 for the Kaggle dataset. The Mendeley dataset was used for training, while Kaggle was used for testing.

3.2. Jerk Features

Vehicular jerk features, linked to driving volatility assessment, help identify deviations from normal driving and quantify speed variations during a journey. An acceleration profile illustrates how a driver accelerates and decelerates, while a jerk profile shows the rates of acceleration and deceleration. This is crucial for assessing abrupt changes in driving behavior. Jerk J[ms^{-3}] is related to acceleration a[ms^{-2}] and velocity v[ms^{-1}] by the equation [24]:

$$J = \frac{da}{dt} = \frac{d^2 v}{dt^2} \tag{1}$$

Figure 1 illustrates the proposed model framework. The feature engineering step is included to improve the performance of driving profiling. Here, apart from well-known features such as amplitude, variance, standard deviation, coefficient of variation, standard error, skewness, and kurtosis, some new derived jerk-based indicators, such as jerk_std, jerk_variance, jerk_amplitude, and jerk_spikes, are generated for each window. The IMU channels (Acc_X, Acc_Y, Acc_Z) are processed to generate a behavioral score for each window, enabling the classification of driving patterns into classes 0 and 1. Six machine learning models serve as baselines, providing a comprehensive performance benchmark.

The proposed framework for feature selection and DS computation is shown in Figure 2.

The Driving Score (DS) quantifies driving quality by combining acceleration, speed, and jerk data. Initially, data is normalized to a 0–100 range and a vector of normalized characteristics is constructed as:

$$f_{norm} = (f_{norm}^1, f_{norm}^2, \ldots, f_{norm}^N) \tag{2}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2026 doi:10.20944/preprints202604.2048.v1

Each feature gets the same weight $w_i = \frac{1}{N}, \forall i \in \{1,2,\ldots,N\}$. The RawScore is defined as,

$$RawScore = \sum_{i=1}^{N} w_i f_{norm}^{i} = \frac{1}{N}\sum_{i=1}^{N} f_{norm}^{i} \tag{3}$$

The RawScore is converted into a Driving Score defined on the range [0, 100] by the relation:

$$DrivingScore = 100(1 - RawScore) \tag{4}$$

The Driving Score (DS) captures the full spectrum of dynamic driving conditions, and the proposed score range is as follows.

Table 1. The proposed score range.

Driver's profiling	Minimum DS	Maximum DS
Calm/normal driver	75	100
Aggressive driver	49.9	0

Time-series data analysis reveals a wider range of positive jerk dynamics during acceleration than during braking/deceleration. Consequently, to report a stable driving attitude, a threshold of 50 aligns with central tendency and effectively serves as a functional boundary distinguishing between normal and aggressive drivers.

Cohen's d measure characterizes the effect size, i.e., quantifies the magnitude of statistical significance difference in standard deviation units, by relating the mean difference to variability. A large Cohen's d value indicates the mean difference/effect size is large compared to the data's variability. Essentially, the difference between normal and aggressive drivers is not only real but also significant.

$$Cohen's\ d = \frac{window\ A\ mean - window\ B\ mean}{average\ standard\ deviation\ across\ both\ windows} \tag{5}$$

The interpretation of the Cohen's d measure is as follows: small (0.2), medium (0.5), and large (0.8) [25].

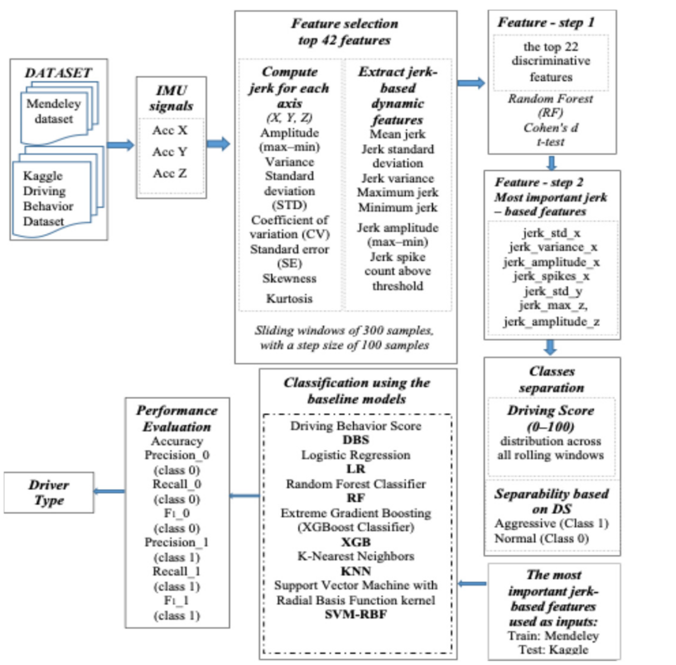

Figure 1. Methodology overview.

```
Input: Accelerometer data
(signals on the X, Y, Z axes)

1. Load IMU data
2. Segment data using sliding windows:
   For start = 0 to (Length - WindowSize) step
Step:
      window = data[start : start +
WindowSize]
3. For each window, compute jerk:
   jerk_x = derivative (Acc_X)
   jerk_y = derivative (Acc_Y)
   jerk_z = derivative (Acc_Z)
4. Extract statistical features for each axis:
   amplitude = max(signal) - min(signal)
   variance = Var(signal)
   std = Std(signal)
   coef_var = std / mean(signal)
   std_error = std / sqrt(N)
   skewness = Skew(signal)
   kurtosis = Kurtosis(signal)
5. Extract jerk-based dynamic features:
   jerk_mean
   jerk_std
   jerk_variance
   jerk_max
   jerk_min
   jerk_amplitude = jerk_max - jerk_min
   jerk_spikes = count (|jerk| > threshold)
6. Create feature vector for current window:
FeatureVector = [all features from steps 4–5]
7. Repeat for all windows
8. Select reduced feature subset:
   SelectedFeatures =
  [jerk_std_x, jerk_variance_x,
  jerk_amplitude_x, jerk_spikes_x,
  jerk_std_y, jerk_max_z, jerk_amplitude_z]
9. Split datasets:
   Training set (Mendeley)
   Testing set (Kaggle)

Output:
DrivingScore for each window
DriverClass (0 = Calm, 1 = Aggressive)
Evaluation metrics: Accuracy, Precision_0,
Recall_0, F1_0, Precision_1, Recall_1, F1_1

10. Compute DrivingScore:
   Normalize features using Min–Max
scaling (fit on training data, apply to test data)
   Compute Driving Score according to
Equations (2)–(4)
distribution alignment:
Δ = mean (DS_train) − T
DS_test = DS_test − Δ
11. Classify driving behavior:
   If DrivingScore ≥ T:
      DriverClass = 0 (Normal)
   Else:
      DriverClass = 1 (Aggressive)
12. Train baseline machine learning models:
Logistic Regression (LR), Random Forest
(RF), XGBoost (XGB), K-Nearest Neighbors
(KNN), Support Vector Machine (SVM)
13. Generate predictions on test data:
   y_pred = predicted labels
   y_true = ground truth labels
14. Compute evaluation metrics:
 Accuracy = (TP + TN) / (TP + TN + FP + FN)
   Precision_1 = TP / (TP + FP)
   Recall_1 = TP / (TP + FN)
   F1_1 = 2 * (Precision_1 * Recall_1) /
  (Precision_1 + Recall_1)
   Precision_0 = TN / (TN + FN)
   Recall_0 = TN / (TN + FP)
   F1_0 = 2 * (Precision_0 * Recall_0) /
   (Precision_0 + Recall_0)
End.
```

Figure 2. DS pseudocode and classification step using the baseline models.

4. Experimental Results

The experimental results from applying the proposed Driving Behavior Scoring framework to the IMU dataset are used to explore statistical differences between calm and aggressive driving windows. This includes examining the dynamic behavior captured by jerk-based indicators and assessing the Driving Score's effectiveness in distinguishing between the two classes. In a first step, both datasets provide 42 features. For the Mendeley dataset, the feature importance is shown in Figure 3.

The 42 IMU signal features reveal that jerk dynamics clearly separate classes 0 and 1. The top three most predictive features for normal driving style are: jerk_variance_z (4684.14), jerk_variance_x (1535.74), and jerk_variance_y (1090.73). The top three most predictive features for aggressive driving style are: jerk_variance_x (6465.34), jerk_variance_y (4707.68), and jerk_amplitude_x (587.82). They indicate increased instability and irregularity in the motion patterns.

Jerk_std_x is 77.88 (class 1), up from 34.52 (class 0), jerk_std_z increasing from 62.74 in class 0 to 138.6 in class 1, and jerk_variance_z increasing from 4684.14 in class 0 to 20922.75 in class 1. Significant changes in acceleration occur during aggressive driving. Shock intensity and frequency indicators follow similar patterns. Jerk_spikes_z rises from 271.28 (class 0) to 287.6 (class 1), whereas jerk_spikes_x and y follow similar trends (from 238.66 (class 0) to 282.6 (class 1) and from 218.57(class 0) to 270.38 (class 1). These increases suggest class 1 segments have more abrupt, irregular, high-intensity dynamics than class 0. This is encouraged by signal extremes jerk_max_z from 235.28 (class 0) to 542.49 (class 1), on Ox from 89.52 (class 0) to 203.78 (class 1), and on Oy, from 124.57 (class 0) to 288.36 (class 1). Jerk_amplitude_z rises from 503.06 (class 0) to 1109.63 (class 1). Furthermore, in the class 1 case, jerk_mean_x and jerk_mean_z are consistently higher, indicating dynamic activity during aggressive driving.

doi:10.20944/preprints202604.2048.v1

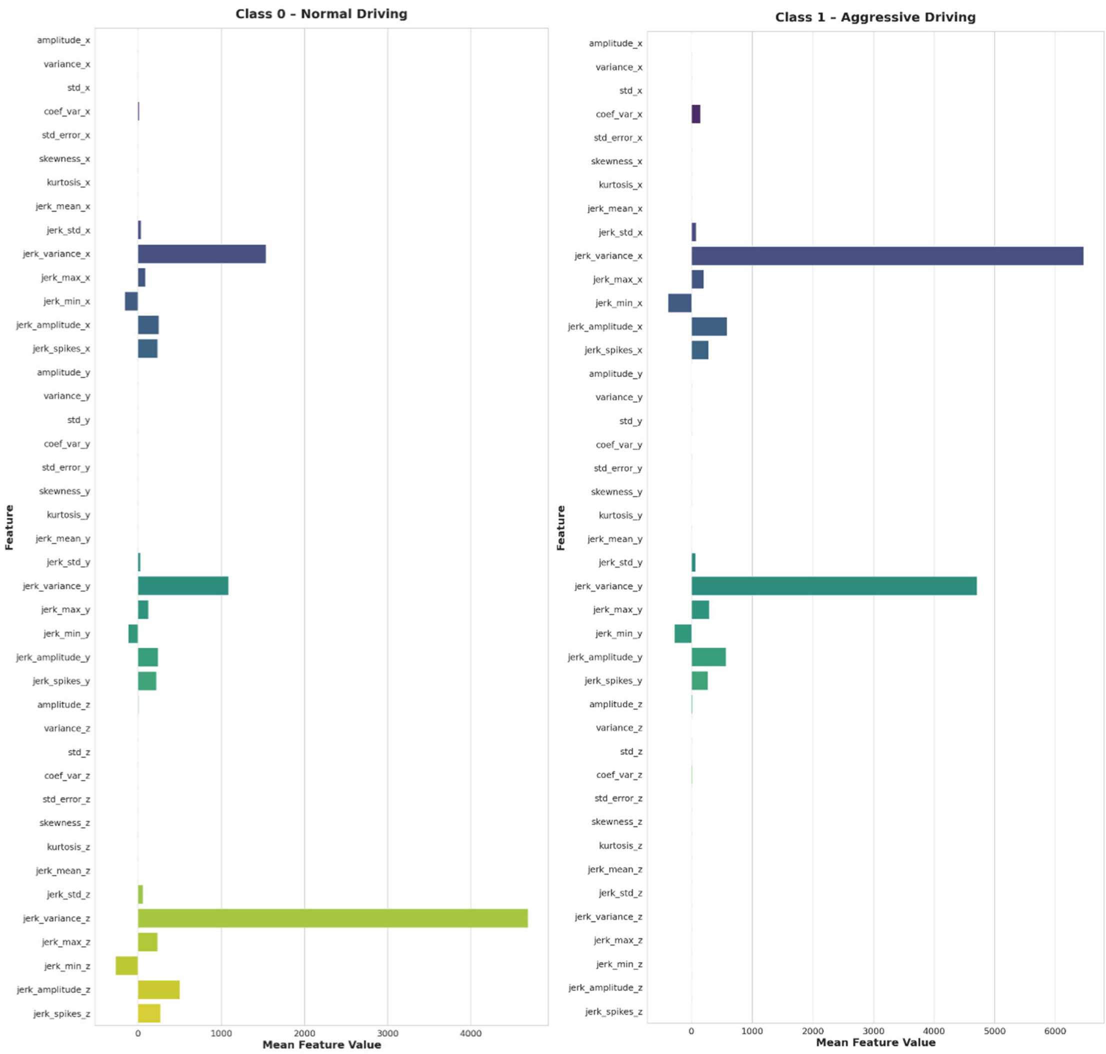

Figure 3. Feature importance of the DBS model for the Mendeley dataset. The analysis of the 42 IMU signal features shows that jerk dynamics effectively distinguish between classes 0 and 1, while the classical statistics feature fails. For example, jerk_variance_x and jerk_variance_y make major contributions to both classes, while jerk_variance_z significantly contributes to class 0.

For the Kaggle dataset, the extracted statistical descriptors are presented in Figure 4. The top three most predictive features for normal driving style are: jerk_variance_z (4957.69), jerk_variance_x (2897.64), and jerk_variance_y (3129.72). The top three most predictive features for aggressive driving style are: jerk_variance_z (6118.22), jerk_variance_y (4724.43), and jerk_ variance_x (4701.71). The jerk features related to the standard deviation metric show that the normal class has jerk_std_z (69.58), jerk_std_y (55.55), and jerk_std_x (53.65). For class 1, jerk_std_z (77.64), jerk_std_y (68.52), and jerk_std_x (68.26). These differences are quite evident, demonstrating a clear distinction between the two driving styles. For jerk amplitude: jerk_amplitude_x goes up from 349.97 (class 0) to 488.59 (class 1), jerk_amplitude_y varies from 369.10 (class 0) to 452.26 (class 1), and jerk_amplitude_z goes up from 497.16 (class 0) to 582.04 (class 1). The jerk_spikes feature variations are smaller. Jerk_spikes_x ranges from 274.30 (class 0) to 281.11 (class 1) while jerk_spikes_y varies from 277.20 (class 0) to 278.58 (class 1). Finally, jerk_spikes_z ranges from 279.30 (class 0) to 280.37 (class 1). The jerk_max values vary as follows: jerk_max_x ranges from 171.73 (class 0) to 229.54 (class 1), jerk_max_y from 189.48 (class 0) to 229.68 (class 1), and jerk_max_z from 244.48 (class 0) to 298.96 (class 1). Similarly, jerk_min_x ranges from -178.25 (class 0) to -259.05 (class 1) while jerk_min_y varies from -179.62 to -

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2026 doi:10.20944/preprints202604.2048.v1

222.58 and jerk_min_z from -252.68 to -283.08. This suggests braking events are becoming more forceful.

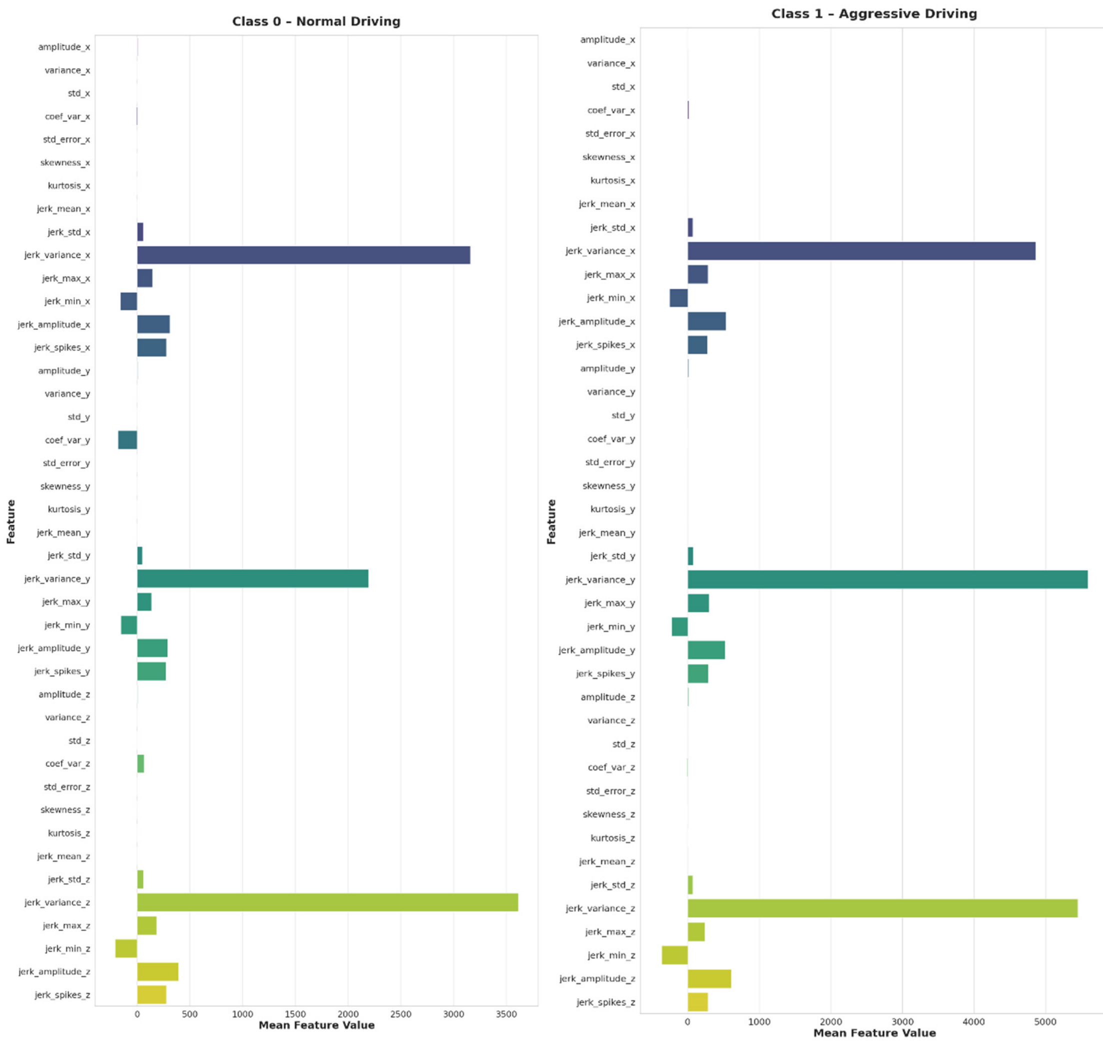

Figure 4. Feature importance of the BDS model for the Kaggle Driving Behavior Dataset. The analysis of the 42 IMU signal features shows that jerk dynamics effectively distinguish between classes 0 and 1, while the classical statistics feature fails. For example, most balanced contributions of jerk_variance_x, jerk_variance_y, and jerk_variance_z are observed for both classes.

Following the comparative evaluation described above and illustrated in Figures 3 and 4, we employed the Random Forest (RF) algorithm to select the top-performing features from all 42 analyzed. Figure 5 shows the ranking of IMU characteristics for classification, based on the RF feature priority ranking. Thus, RF selects only 22 meaningful features.

To confirm these findings, for selected features, a statistical analysis using t-tests, separately for both datasets (Tables 2 and 3), is performed. The statistics in Table 2 show that the two classes differ significantly, particularly in jerk qualities. Classes 0 and 1 show substantial differences, with p-values $< 10^{-15}$ and Cohen's d > 2.0. This shows jerk-based signals are a useful indicator of drivers' behavior. Using p-values ($p < 0.05$), impact sizes (|Cohen's d| > 0.8), and category relevance (importance > 0.01), we confirm the soundness of the selected 22 relevant features. For all features, the distributions corresponding to Class 1 (aggressive driving) are consistently shifted toward higher values and exhibit much larger variability compared to Class 0. The t-test results confirm the relevance of these indicators in identifying the vehicle's dynamic instability, reinforcing the class separation. Eleven of the 22 selected features are jerk-based.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2026 doi:10.20944/preprints202604.2048.v1

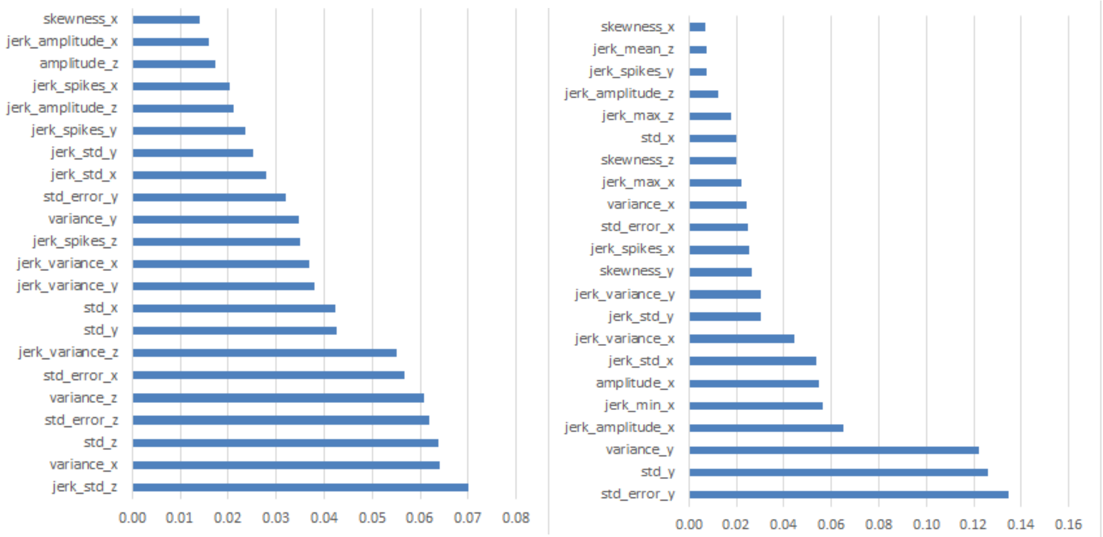

Figure 5. The most important IMU characteristics for classification, based on the Random Forest feature priority ranking: (a) for the Mendeley dataset; (b) for the Kaggle Driving Behavior Dataset.

Table 2. The average values of the most relevant features based on t-test and Cohen's d results for the Mendeley dataset.

Feature	Mean Class 0	Mean Class 1	Mean Diff	p-value	Cohen's d
std_y	0.4205	0.9758	0.5553	6.89E-25	2.2608
std_error_y	0.02428	0.05634	0.03206	6.78E-25	2.2507
jerk_std_x	34.52	77.89	43.36	5.55E-25	2.2336
jerk_std_y	28.85	66.36	37.52	7.56E-25	2.2205
std_x	0.5201	1.1962	0.6761	1.41E-24	2.2074
std_error_x	0.03	0.0691	0.039	1.41E-24	2.2074
std_error_z	0.05276	0.11831	0.06555	3.40E-26	2.1863
std_z	0.9138	2.0491	1.1354	3.40E-26	2.1863
jerk_std_z	62.74	138.6	75.87	1.22E-25	2.1479
jerk_spikes_z	271.29	287.61	16.32	3.79E-18	2.0522
jerk_spikes_y	218.57	270.38	51.81	1.82E-16	1.9915
amplitude_z	7.109	16.96	9.851	3.70E-22	1.9321
amplitude_x	4.253	9.702	5.449	2.35E-19	1.9094
jerk_spikes_x	238.66	282.6	43.93	2.32E-15	1.8837
jerk_amplitude_z	503.06	1109.63	606.57	3.95E-20	1.8785
jerk_amplitude_x	250.55	587.82	337.27	2.29E-20	1.8708
jerk_min_x	-161.03	-384.04	-223.01	5.23E-20	1.8382
jerk_variance_x	1535.74	6465.34	4929.6	5.73E-21	1.8287
variance_y	0.23475	1.01416	0.77941	5.52E-21	1.8273
jerk_max_z	235.28	542.5	307.21	2.91E-20	1.8243
variance_x	0.35649	1.52971	1.17322	4.18E-20	1.7712
amplitude_y	3.5339	8.2574	4.7235	4.82E-19	1.7684

Table 3. The average values of the most relevant features based on t-test and Cohen's d results for the Kaggle Driving Behavior Dataset.

Feature	Mean Class 0	Mean Class 1	Mean Diff	p-value	Cohen's d
std_error_y	0.048772	0.065949	0.017177	1.52E-11	3.098565
std_y	0.844749	1.142266	0.297517	1.32E-11	3.092563

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2026 doi:10.20944/preprints202604.2048.v1

variance_y	0.724149	1.311692	0.587543	1.53E-11	3.088703
jerk_std_x	53.64572	68.26217	14.61645	4.73E-09	2.573756
jerk_variance_x	2897.641	4701.71	1804.069	2.41E-08	2.487013
amplitude_x	5.842132	8.461601	2.619469	1.11E-08	2.426849
std_x	0.888451	1.232249	0.343797	5.93E-08	2.259893
std_error_x	0.051295	0.071144	0.019849	5.93E-08	2.259893
variance_x	0.804261	1.547805	0.743544	3.32E-07	2.169465
jerk_std_y	55.55076	68.52391	12.97315	1.23E-07	2.089843
jerk_variance_y	3129.723	4724.427	1594.704	1.57E-07	2.06342
jerk_amplitude_x	349.9736	488.5905	138.6169	5.49E-07	2.003472
jerk_max_x	171.7273	229.5377	57.81031	4.50E-05	1.47686
jerk_spikes_x	274.3	281.1053	6.805263	0.000106	1.413204
amplitude_y	5.912154	7.70982	1.797666	0.000729	1.244735
jerk_amplitude_z	497.1614	582.0357	84.8743	0.000751	1.173531
jerk_max_z	244.4822	298.9603	54.47812	0.000904	1.172567
kurtosis_x	0.747122	1.61447	0.867348	0.002304	1.110563
skewness_x	0.021984	0.357469	0.335484	0.003437	1.018505
jerk_amplitude_y	369.099	452.2575	83.15848	0.019511	0.809856
std_error_z	0.057506	0.063991	0.006485	0.015902	0.805656
std_z	0.99604	1.108361	0.112321	0.014501	0.802423

The repeated statistically significant differences ($p < 0.05$) in the Kaggle dataset (Table 3) demonstrate a statistically significant difference between the two classes. Also, the practical significance is conveyed through effect sizes, as large Cohen's d values exist. In the Kaggle dataset case, ten of the 22 selected features are jerk-based.

The distribution of Driving Score values is shown in Figures 6 and 7.

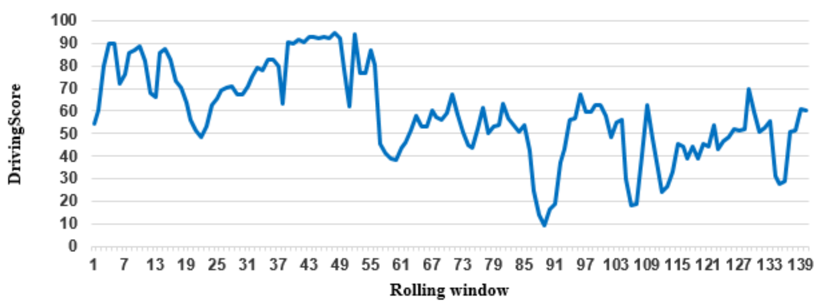

Figure 6. The distribution of the Driving Score measure for the Mendeley dataset.

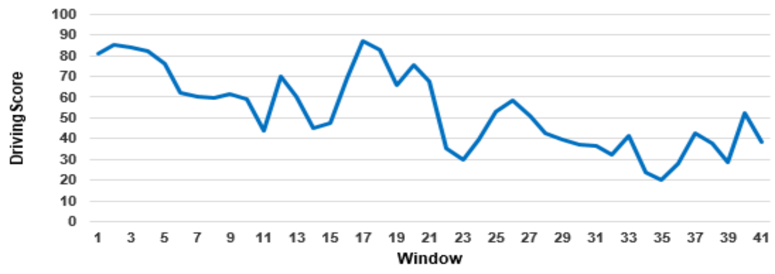

Figure 7. The distribution of the Driving Score measure for the Kaggle Driving Behavior Dataset.

Figure 6 shows the DS distribution across all rolling and sliding windows, revealing a wide score range from 9.27 to 93.88. This encompasses both highly unstable driving segments and remarkably stable intervals. Windows with excessive jerk movement correlate directly with low DS values, while those with stable motion and minimal jerk variability correlate with high DS values. The DS dropped sharply to values between 9.27 (window 89) and 18.83 (window 106), marking the lowest scores in the entire sequence. This indicates the most hazardous driving style. Low-scoring windows correspond to abrupt changes in vehicle motion, elevated jerk intensity, high variance, and frequent spike events, indicating an aggressive behavior. The jerk dynamics vary strongly in these windows. The jerk_amplitude_x values, for instance, are over 500, which means that the acceleration changes a lot from one peak to the next. There are also more spike events and jerk variability between windows 57 and 62 (DS ~ 38–45) and 105 and 108 (DS ~ 18–39). This suggests the driver was accelerating rapidly, decelerating quickly, or making sudden steering adjustments. Windows from 42 to 52 show consistently higher scores. This reflects smooth acceleration profiles, reduced jerk variability, and minimal peak activity, aligning with the characteristics of normal calm driving. Windows with DS values consistently above 80, like windows 39–49 (scores up to 94.48), have lower jerk-based metrics. For instance, there is a lower jerk_amplitude (250.55 on the x-axis and 503.06 on the z-axis), which means that the acceleration limits are lower, and a low jerk_std (34.52 on the x-axis and 62.74 on the z-axis), which means that the acceleration profiles are smooth and gradual. This continuous scoring range allows the model to capture not only the two final behavioral classes but also the subtle transitions between the two states, which are visible in the (64–85) and (122-128) intervals.

Figure 7 provides the DS distribution across all rolling and sliding windows for the Kaggle dataset. This reveals a score range from 20.20 to 86.87, suggesting a generally moderate driver profile. The DS reaches its lowest values between 20.21 (window 35) and 23.65 (window 34), indicating the most unstable windows are linked to aggressive driving. The DS values for windows 36 (28.17) and 39 (28.84) are almost identical. Low scores ranging from 30 to 40 are also observed for windows 23 (30.06), 32 (32.43), 22 (35.54), 31 (36.46), 30 (37.26), 38 (37.54), 41 (38.45), 29 (39.35), and 24 (39.63). This suggests recurring fluctuations in driving style. Windows with moderate scores between 40 and 70 are placed in intervals 6–10 and 25–28, suggesting transitions between stable and unstable driving. This is evident from variations in acceleration and jerk without extreme levels. These segments indicate an uneven driving style. Transitions between the normal and aggressive driving styles are observed for the following windows: 33, 37, 28, 11, 14, 15, 27, 40, 25, 26, 10, 8, 13, 7, 9, 6, 19, 21, 16, 12, 20, 5. DS values range from 41.48 to 75.91. These indicate an inconsistent driving style. The higher DS values are found in windows 1 (80.62), 4 (82.29), 18 (82.71), 3 (84.16), 2 (84.98), and 17 (86.87), indicating a calm and controlled driving style.

To further evaluate and validate the robustness of jerk-based feature dynamics, we carried out a series of quantitative analyses on newly generated datasets containing jerk-based features. These new datasets contain only jerk-based features (jerk_std_x, jerk_variance_x, jerk_amplitude_x, jerk_spikes_x, jerk_std_y, jerk_max_z, jerk_amplitude_z*)*, as are presented in Tables 3 and 4. The jerk-based features extracted from the Mendeley dataset are used as the training set, while those from the Kaggle dataset are used for testing. Table 5 presents the baseline machine learning classifiers used to confirm the robustness of our approach. The performance metrics of the proposed approach and additional baseline models, for both classes 0 and 1, are reported in Table 6. The corresponding confusion matrices for each model are illustrated in Figure 8.

Table 5. Classifiers used as a baseline against the proposed model DBS.

Abbreviation	Full Name
DBS	Driving Behavior Score
LR	Logistic Regression
RF	Random Forest Classifier
XGB	Extreme Gradient Boosting (XGB Classifier)
KNN	K-Nearest Neighbors
SVM-RBF	Support Vector Machine with Radial Basis Function kernel

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 29 April 2026 doi:10.20944/preprints202604.2048.v1

Table 6. Performance comparison between models across the newly generated jerk-based features dataset. Evaluation metrics include accuracy, specificity, sensitivity, and F1-score.

Model	Accuracy	Precision (class 0)	Recall (class 0)	F1 (class 0)	Precision (class 1)	Recall (class 1)	F1 (class 1)
DBS	0.907	0.922	0.839	0.879	0.899	0.952	0.925
LR	0.900	0.875	0.875	0.875	0.917	0.917	0.917
RF	0.879	0.855	0.839	0.847	0.894	0.905	0.899
XGB	0.864	0.836	0.821	0.829	0.882	0.893	0.888
KNN	0.857	0.800	0.857	0.828	0.900	0.857	0.878
SVM_RBF	0.857	0.790	0.875	0.831	0.910	0.845	0.877

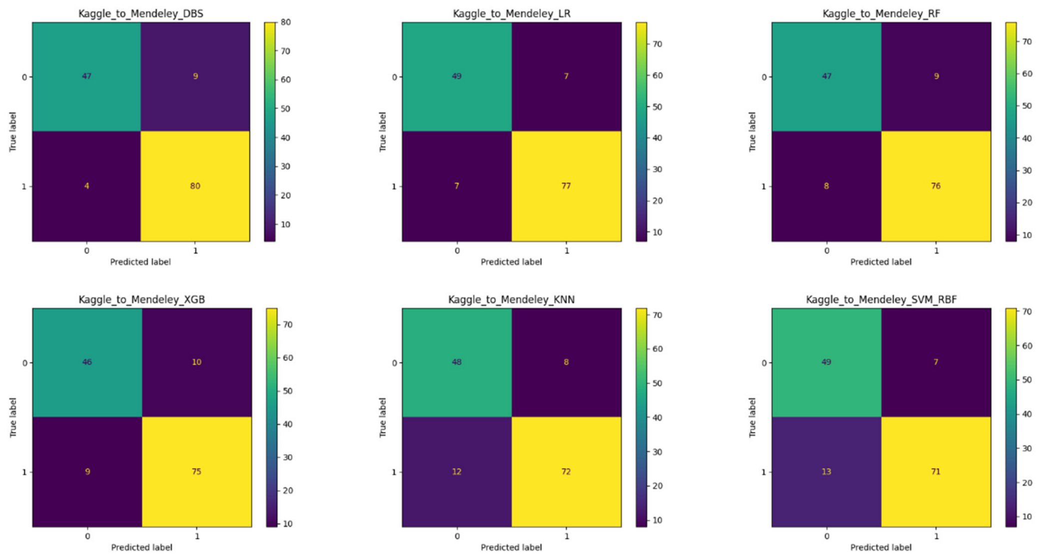

Figure 8. Confusion matrices for each model.

The Driving Behavior Score (DBS) model identifies 80 true positives and 4 false negatives, with a Recall (class 1) of around 0.952 and an F1 (class 1) of about 0.925. The model also had a total accuracy of 0.907. There was a trade-off between sensitivity and specificity since DBS produced 9 false positives for the normal class (class 0). Logistic Regression has an accuracy of 0.90, and shows the best balance between the two classes in terms of precision and recall. Ensemble models like XGB and RF also perform well, with an overall accuracy of around 0.85. KNN and SVM_RBF classifiers show a slightly lower classification performance with an accuracy of 0.857.

4. Discussion

In this paper, we presented a new approach for driver behavior profiling using dynamic information. This study interpreted and data-driven profiled driving behavior using statistical descriptors derived from IMU signals. To validate the proposed method, we have analyzed two datasets. We used a well-established Mendeley dataset cleaned of noise and other influences. The Kaggle dataset is somewhat limited as it uses one driver on two routes, and the car was slow. Data in Figures 3 and 4 displayed the feature importance of the DBS model. The extracted statistical descriptors show clear differences between the two behavior classes with consistent differences for jerk-derived features across all IMU axes. The Random Forest algorithm selects just 22 meaningful features from the top 42 performers (Figure 5). The statistical significance between the selected characteristics of the two classes was assessed using t-tests, separately for both datasets (Tables 2 and 3). The statistics in Tables 2 and 3 confirm that the two classes differ significantly, particularly in jerk qualities. Moreover, the practical significance of these results is reinforced through effect sizes. The

 doi:10.20944/preprints202604.2048.v1

large effect sizes (Cohen's d > 1.7) confirm the significant differences between driving style classes as established by jerk-based signals. This suggests the difference between the two group means is greater than the standard deviation and often exceeds the variability within each group. Consequently, the distributions are clearly distinct with minimal overlap. These differences suggest real shifts in driving behavior. The proposed framework considers both statistical and practical significance, minimizing noise or measurement errors.

Furthermore, the data in Figures 6 and 7 illustrate the distribution of DS values, supporting the interpretability of the proposed method. This step assesses the model's robustness and generalization capabilities. It is worth noting that the composite DS does not make binary classifications; instead, it detects subtle changes in driving behavior. The strong negative correlation between DS and jerk-based indicators demonstrates that jerk is the most significant factor differentiating various driving behaviors. Higher jerk amplitudes and greater jerk variability lead to lower composite scores, while stable jerk profiles result in higher scores. This behavior supports the notion that jerk measures driving smoothness and validates the proposed composite scoring system. In the proposed scoring system, drivers with DSs above the average threshold of 50 do not need further risk management. However, those with lower DSs may benefit from auxiliary driving style corrections.

As data in Table 6 and Figure 8 show, baseline models' performance suggests the proposed approach is suitable, and the data quality (derived from jerk-based features) is adequate for machine learning classification. While the Logistic Regression classifier serves as a strong baseline model, achieving similar overall accuracy to our method, the proposed DBS approach offers more balanced performance across classes. This is achieved while maintaining a high F1 score for both classes.

Due to the lack of existing research reporting experimental protocols and evaluations of DS related to statistical and dynamic descriptors, particularly in jerk-based feature dynamics, a direct comparison with prior studies is not feasible. However, based on baseline model comparison, we demonstrated that our proposed approach is a viable solution. We conducted a thorough within- and intra-study evaluation across both datasets.

To some extent, we can say that the reported findings align with the majority of existing studies demonstrating the potential of jerk to identify aggressive drivers by using data collected via sensors embedded in mobile phones [26]. Both acceleration and jerk characteristics contribute to perceived motion intensity. Consistent with the literature, the results indicated that perceived motion intensity depended both on acceleration and jerk [27]. The key difference and novelty are that our results were not derived directly from jerk values but rather from jerk-based features.

Beyond these task-specific observations, several limitations of the study should also be acknowledged. Firstly, the Mendeley dataset documentation simply states that the sensor data was recorded from an Android phone mounted on a dashboard while driving and includes a label column for driving behavior. It does not specify how many drivers (different people) contributed their data. The dataset description does not list a count of subjects or drivers. Secondly, the Kaggle dataset omits key predictive features from the DS value computation. For instance, it only retains jerk_spikes_x, ignoring other spike variations that signify the abrupt, high-intensity dynamics characteristic of erratic or risky driving. DS combines acceleration, speed, and jerk data to determine a driver profile, and missing features can impact this. Thirdly, when the aggressive driving was assessed, only speeding was chosen to label this behavior. Future studies will examine various aspects of aggressive driving, including weaving through traffic or illegal lane crossing. To further enhance the accuracy of driving profile identification, specific turning behaviors, such as left and right turns, may be included. Also, we will examine adaptive thresholding, individualized scoring profiles, and multimodal data fusion to increase system accuracy and application.

Overall, the results demonstrate consistent improvements in driving behavior profiling across both datasets. The proposed approach effectively transfers jerk- discriminative representations into a robust composite scoring system. Furthermore, the proposed approach is less affected by variations in IMU modality acquisition and dataset-specific differences.

 doi:10.20944/preprints202604.2048.v1

5. Conclusions

The proposed method revealed 22 jerk-based and variability features with the best discriminative power, with effect sizes up to Cohen's d > 1.7 and p-values < 10^{-15}, after assessing over 140 rolling/sliding windows for the Mendeley dataset and 41 windows for the Kaggle dataset, each containing 300 samples. The reported results demonstrate the overall performance and reliability of the proposed scoring-based classification. It consistently aligns with the statistical characteristics of IMU signals and is therefore suitable for real-time driver monitoring.

In addition, systematic comparisons with conventional machine learning models across multiple performance metrics further support the favorable performance of the proposed approach. The comparison was conducted within- and intra-study evaluation across both datasets. This demonstrates the robustness of jerk-based feature dynamics and their independence from the quality of the raw data.

The proposed method is transparent, computationally lightweight, and robust. It works on interpretable physical indicators, making it suitable for real-time telematics, driver-monitoring, and mobile sensing applications. The potential applications of the proposed approach include telematics platforms, driver-assistance systems, and mobile monitoring applications where transparency, interpretability, and low computational cost are essential.

Author Contributions: FAM and DDD: Writing—review & editing, Writing—original draft, Software, Methodology, Investigation, Formal analysis. LM: Conceptualization, Writing—review & editing, Funding acquisition, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Informed Consent Statement: Not applicable.

Data Availability Statement: We utilized a publicly available dataset to train our models. The datasets can be accessed using the following links: https://data.mendeley.com/datasets/5stn873wft/1, https://www.kaggle.com/datasets/outofskills/driving-behavior (accessed on 1st February 2026).

References

1. Cheng, R.; Liu, F.; Ge, H. A new continuum model based on full velocity difference model considering traffic jerk effect. *Nonlinear Dynamics* **2017**, *Volume 89 (1*), pp. 639–649. https://doi.org/10.1007/s11071-017-3477-2
2. Zhai, C.; Wu, W. Analysis of drivers' characteristics on continuum model with traffic jerk effect. *Physics Letters A* **2018,** *Volume 382 (47),* pp. 3381–3392. https://doi.org/10.1016/j.physleta.2018.09.029
3. Feng, F.; Bao, S.; Sayer, J.; Flannagan, C.; Manser, M.; Wunderlich, R. Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data. *Accident Analysis & Prevention* **2017**, *Volume 104*, pp. 125–136.
4. https://doi.org/10.1016/j.aap.2017.04.012
5. Haselberger, J.; Schick, B.; Müller, S. Self-perception versus objective driving behavior: Subject study of lateral vehicle guidance. *Transportation Research Part F: Psychology and Behaviour* **2025**, *Volume 109*, pp. 272–298.
6. https://doi.org/10.1016/j.trf.2024.12.012
7. Qi, J.; Zhu, R.; Liu, C.; Mauricio, A.; Gryllias, K. Anomaly detection and multi-step estimation based remaining useful life prediction for rolling element bearings. *Mechanical Systems and Signal Processing* **2024**, *Volume 206*, p. 110910.
8. https://doi.org/10.1016/j.ymssp.2023.110910

9. Fernandes, P.; Ferreira, E.; Macedo, E.; Coelho, M. C. Unraveling roundabout dynamics: Analysis of driving behavior, vehicle performance, and exhaust emissions. *Transportation Research Part D: Transport and Environment* **2024,** *Volume 133*, p. 104308.
10. https://doi.org/10.1016/j.trd.2024.104308
11. Mantouka, E.; Barmpounakis, E.; Vlahogianni, E.; Golias, J. Smartphone sensing for understanding driving behavior: Current practice and challenges. *International Journal of Transportation Science and Technology*, **2021**, *Volume 10, Issue 3*, pp. 266-282.
12. https://doi.org/10.1016/j.ijtst.2020.07.001
13. Shahariar, G.H.; Sajjad, M.; Suara, K.A.; Jahirul, M.I.; Chu-Van, T.; Ristovski, Z.; Brown, R.J.; Bodisco, T.A.. On-road CO2 and NOx emissions of a diesel vehicle in urban traffic. *Transportation Research Part D: Transport and Environment* **2022**, *Volume 107*, p.103326. https://doi.org/10.1016/j.trd.2022.103326.
14. Ferreira, E.; Fernandes, P.; Bahmankhah, B.; Coelho, M.C. Micro-analysis of a single vehicle driving volatility and impacts on emissions for intercity corridors. *International Journal of Sustainable Transportation* **2022**, *Volume 16 (8)*, pp. 681–705.
15. https://doi.org/10.1080/15568318.2021.1919797.
16. Hayati, H.; Eager, D.; Pendrill, A.-M.; Alberg, H. Jerk within the Context of Science and Engineering—A Systematic Review. *Vibration* **2020**, *Volume* 3, pp. 371-409. https://doi.org/10.3390/vibration3040025
17. Shuai, Z.; Dong, A.; Liu, H.; Cui, Y. Reliability and Validity of an Inertial Measurement System to Quantify Lower Extremity Joint Angle in Functional Movements. *Sensors* **2022**, *Volume 22*, p. 863. https://doi.org/10.3390/s22030863
18. Chen, S.; Xiao, Y.; Wang, Y.; Xie, Y.; Zhu, T.; Duan, R.; Chen, J. Skew-normal distributions for modeling asymmetric moving tendencies in pedestrian trajectories. *Neurocomputing* **2026**, *Volume 661*, p. 131934. https://doi.org/10.1016/j.neucom.2025.131934
19. Redhu, P.; Siwach, V. An extended lattice model accounting for traffic jerk. *Physica A: Statistical Mechanics and its Applications* **2018**, *Volume 492,* pp. 1473–1480. https://doi.org/10.1016/j.physa.2017.11.074
20. Fernandes, P.; Tomás, R.; Acuto, F.; Pascale, A.; Bahmankhah, B.; Guarnaccia, C.; Granà, A.; Coelho, M.C. Impacts of roundabouts in suburban areas on congestion-specific vehicle speed profiles, pollutant and noise emissions: An empirical analysis. *Sustainable Cities and Society* **2020**, *Volume 62*, p. 102386. https://doi.org/10.1016/j.scs.2020.102386.
21. Sun, B., Deng, W., Wu, J., Li, Y., Wang, J. An intention-aware and online driving style estimation based personalized autonomous driving strategy. *International Journal of Automotive Technology* **2020**, *Volume 21(6)*, pp. 1431-1446.
22. https://doi.org/10.1007/s12239-020-0135-3
23. Karrouchi, M.; Nasri, I.; Rhiat, M.; Atmane, I.; Hirech, K.; Messaoudi, A.; Melhaoui, M.; Kassmi, K. Driving behavior assessment: A practical study and technique for detecting a driver's condition and driving style, *Transportation Engineering* **2023**, *Volume 14*, p. 100217. https://doi.org/10.1016/j.treng.2023.100217
24. Komavec, M.; Kaluža, B.; Stojmenova, K.; Sodnik, J. Risk assessment score based on simulated driving session, In: Proceedings of the Driving Simulation Conference 2019 Europe VR, Driving Simulation Association, Strasbourg, France, 2019 September 4th-6th, pp. 67-74
25. Feng, F.; Bao, S.; Sayer, J.R.; Flannagan, C.; Manser, M.; Wunderlich, R. Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data, *Accident Analysis & Prevention* **2017**, *Volume 104*, pp. 125-136. doi: 10.1016/j.aap.2017.04.012
26. Medarević, J.; Tomažič, S.; Sodnik, J. Simulation-based driver scoring and profiling system. *Heliyon* **2024**, *Volume 10*, Issue 22, p. e40310. https://doi.org/10.1016/j.heliyon.2024.e40310.
27. Tawadros, P.; Awadallah, M.; Walker, P.; Zhang, N. Using a low-cost bluetooth torque sensor for vehicle jerk and transient torque measurement. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* **2020**, *Volume* 234, pp. 423–437. https://doi.org/10.1177/0954407019861613
28. El mourabit, A.; Bouazizi,O.; Oussouaddi, M.; El Abidine Z.; Ismaili, A.; Attaoui, Y.; Chentouf, M. Instantaneous time to collision estimation using a constant jerk model and a monocular camera. *Engineering Science and Technology, an International Journal* **2025**, *Volume 64*, p. 102011. https://doi.org/10.1016/j.jestch.2025.102011

 doi:10.20944/preprints202604.2048.v1

29. Nazirkar, S. Phone sensor data while driving a car and normal or aggressive driving behaviour classification [Data set]. Mendeley Data, V1. (2021). https://doi.org/10.17632/5stn873wft.1 (accessed on 1st February 2026).
30. Kaggle. Driving Behavior Dataset. Available online: https://www.kaggle.com/datasets/outofskills/driving-behavior (accessed on 1st February 2026).
31. Zhang, L., Peng, K., Zhao, X., Khattak, A.J. New fuel consumption model considering vehicular speed, acceleration, and jerk. *Journal of Intelligent Transportation Systems* **2023,** *Volume 27 (2),* pp. 174–186. https://doi.org/10.1080/15472450.2021.2000406
32. Magnusson, K. A Causal Inference Perspective on Therapist Effects. PsyArXiv **2023** https://doi.org/10.31234/osf.io/f7mvz
33. Mantouka, E. G.; Barmpounakis, E.N.; Vlahogianni, E. I. Identifying driving safety profiles from smartphone data using unsupervised learning. *Safety Science* **2019**, *Volume 119*, pp. 84-90, https://doi.org/10.1016/j.ssci.2019.01.025
34. de Winkel, K.N.; Soyka, F.; Bülthoff, H.H. The role of acceleration and jerk in perception of above-threshold surge motion. *Experimental Brain Research* **2020,** *Volume 238 (3)*, pp.699–711. https://doi.org/10.1007/s00221-020-05745-7