Driving Style Classification Using Deep Temporal Clustering with Enhanced Explainability

Ritu Patel, Sarthak Siddhpura, Moin Vinchhi, Vrunda Patel, Vishv Boda
Ahmedabad University, Ahmedabad, Gujarat 380009, India

I. Research Contributions by Paper

- Proposed a deep temporal clustering method for driving style classification, integrating temporal features to capture variations within a single trip across three maneuver types: accelerating, braking, and maintaining.
- Developed a three-phase methodology: (1) segmenting trips into maneuver categories using vehicle speed and DTW-based K-means, (2) learning latent representations with a temporal autoencoder (MLP + BiLSTM), and (3) clustering via KL divergence minimization to classify styles into aggressive, moderate, and conservative.
- Enhanced model explainability using SHAP values to assess feature importance (e.g., throttle, RPM) in clustering outcomes, improving interpretability over traditional "black box" deep learning approaches.
- Collected a car-following dataset from 25 drivers in a driving simulator, sampling at 5 Hz, with features including speed, acceleration, gap, RPM, throttle, brake, and steering, normalized to [0,1].
- Demonstrated superior performance over K-Shape clustering (Silhouette Coefficient 0.642 vs. 0.608), with clusters better aligned to representative driving behaviors (e.g., aggressive style showing higher velocity and throttle).
- Highlighted intra-trip driving style diversity, showing drivers may exhibit different styles (e.g., moderate accelerating, conservative braking) across maneuvers, validated by consistent patterns across multiple trips.

II. Key Insights From Paper

The paper by Feng et al. offers valuable insights for our research on identifying abnormal driving behavior using trajectory datasets and probabilistic-statistical methods, particularly through its innovative deep temporal clustering approach. A central contribution is the segmentation of driving trips into maneuver-specific categories—accelerating, braking, and maintaining—using vehicle speed and Dynamic Time Warping (DTW) with K-means clustering. This aligns closely with our assumption that drivers behave similarly on the same road patch, providing a framework to establish "normal" behavior statistically within spatial-temporal segments. For our binary classifier, this suggests a strategy where trajectories from a road patch can be segmented into short temporal windows (e.g., 1-second sequences as in the paper) and clustered to define a probabilistic baseline of normalcy. DTW's ability to measure sequence similarity is particularly relevant, enabling us to compare a given trajectory's spatial-temporal features (e.g., position, velocity) against cluster centroids. Deviations exceeding a statistical threshold—such as distance from the centroid or variance in speed—could then be flagged as abnormal, leveraging the paper's finding that aggressive behaviors (e.g., high velocity, abrupt changes) distinguish clusters. This temporal granularity enhances our ability to detect micro-level anomalies within a broader context, supporting a prob-stat approach rooted in unsupervised learning.

Another key insight is the use of a temporal autoencoder with bidirectional LSTM (BiLSTM) layers to learn latent representations of driving data, reducing dimensionality while preserving temporal correlations. For our project, this technique could transform high-dimensional trajectory data (e.g., GPS coordinates, speed, direction) into a compact form suitable for clustering and classification. The paper's clustering process, which minimizes KL divergence between latent representations and an auxiliary target distribution, introduces a probabilistic refinement step that we can adapt. By clustering trajectories into normal behavior groups (e.g., moderate driving on a road patch), we could assign soft probabilities (akin to the paper's q_{ij}) to each trajectory's membership, then use a threshold to classify outliers as abnormal. The paper's superior performance over K-Shape (Silhouette Coefficient 0.642 vs. 0.608) underscores the efficacy of incorporating temporal features, suggesting that our classifier could benefit from similar deep learning enhancements to capture subtle deviations—like sudden lane changes or erratic braking—that traditional methods might miss. Applied to naturalistic trajectory datasets, this could yield a robust statistical model of normal driving, with abnormalities detected as low-probability events relative to cluster norms.

The integration of SHAP values to explain feature contributions (e.g., throttle and RPM as dominant in accelerating maneuvers) provides a third critical insight, enhancing interpretability—a priority for validating our classifier's decisions. In our context, applying SHAP to a supervised model fitted to clustering labels could reveal which spatial-temporal features (e.g., speed variance, heading changes) most influence normal vs. abnormal classifications on a road patch. This aligns with our prob-stat focus by offering a way to quantify feature importance probabilistically, ensuring our binary classifier is not a black box but a transparent tool for safety applications. The paper's finding that driving styles vary within a trip—e.g., a driver being moderate in accelerating but conservative in braking—challenges our assumption of uniform behavior, suggesting a need to model normalcy at a maneuver level rather than assuming consistency across a patch. By combining the paper's segmentation, deep clustering, and explainability techniques, we can develop a classifier that leverages trajectory data to define normal behavior statistically, detect anomalies with precision, and provide interpretable results, advancing our goal of enhancing road safety through data-driven insights.

III. Unaddressed Issues and Assumptions

- **Unaddressed Issues:**
  - Lacks real-world validation, relying solely on simulated data, which may not capture the complexity of naturalistic trajectories needed for robust anomaly detection.
  - Does not address scalability of the deep clustering approach for large-scale trajectory datasets, a critical concern for applying the method to extensive road networks.
  - Limited discussion on contextual factors (e.g., road type, traffic conditions) that could affect maneuver classification and anomaly detection on specific road patches.
  - No evaluation of real-time processing feasibility, essential for deploying a binary classifier in dynamic driving scenarios.
- **Assumptions Made:**
  - Assumes three maneuver categories (accelerating, braking, maintaining) sufficiently represent driving behavior, potentially overlooking other patterns (e.g., turning, merging) relevant to trajectory analysis.
  - Assumes simulated data at 5 Hz is representative of real-world driving, which may not account for sensor noise or lower sampling rates in naturalistic datasets.
  - Assumes driving styles are stable within maneuver segments, whereas external factors (e.g., sudden obstacles) could introduce variability not captured in clustering.
  - Assumes feature normalization to [0,1] preserves all relevant information, potentially masking outliers critical for abnormal behavior detection.

IV. Motivation Behind Choosing This Paper

This paper was selected for its novel deep temporal clustering approach, which directly supports our research on abnormal driving detection using trajectory datasets and probabilistic-statistical methods. Its segmentation of trips into maneuvers and use of DTW and K-means provide a practical framework to define normal behavior statistically on road patches, aligning with our binary classification goal. The temporal autoencoder and KL divergence-based clustering offer advanced techniques to process spatial-temporal features, enhancing our ability to detect anomalies in high-dimensional data. Additionally, the incorporation of SHAP values for explainability addresses a key need in our prob-stat approach, ensuring interpretable results. The paper's focus on intra-trip variability and safety-relevant features complements our aim to model driving behavior probabilistically, making it a critical resource for developing an effective and transparent classifier.

References