Urban Traffic Pattern Clustering for Smart Cities

Author: Asad Khan

Abstract:

This thesis investigates the use of unsupervised machine learning techniques—specifically K-Means, DBSCAN, and Hierarchical Clustering—to analyze urban traffic patterns in smart city environments. Effective traffic management remains a major challenge due to increasing urbanization, irregular congestion, and unpredictable mobility behaviors. This study presents a clustering-based analytical framework for grouping traffic zones based on temporal flow characteristics and identifying abnormal traffic events such as unexpected congestion spikes or road disruptions. The methodology is evaluated using large-scale GPS, sensor, and IoT-generated datasets. Experimental results demonstrate that clustering reveals hidden mobility structures, improves peak-hour zone detection, and enables more accurate anomaly identification compared to rule-based systems. The findings provide valuable insights for adaptive traffic control, urban planning, and real-time congestion monitoring.

Keywords: Smart Cities, Traffic Analysis, Clustering, Anomaly Detection, Unsupervised Learning, Urban Mobility

Introduction:

Rapid population growth in metropolitan areas has intensified traffic congestion, resulting in long commute times, energy inefficiency, and environmental burden. Traditional traffic modeling techniques often fail to capture the variability and complexity of modern mobility patterns. Machine learning–based clustering provides a data-driven approach to discovering underlying traffic structures without predefined labels, enabling more intelligent and sustainable traffic management systems.

Proposed Methodology:

This research applies unsupervised learning algorithms to spatiotemporal traffic datasets obtained from GPS trajectories, road sensors, and city IoT networks. K-Means is used to segment regions based on average speed, volume, and time-of-day flow characteristics. DBSCAN identifies density-based congestion clusters and naturally handles outliers. Hierarchical Clustering provides multi-level pattern interpretation across districts. Preprocessing includes noise reduction, normalization, and temporal window segmentation. Anomaly detection is performed by evaluating cluster deviation scores and identifying traffic points that significantly differ from cluster norms, indicating events such as accidents, blockages, or sudden flow surges.

Results Summary:

Experimental evaluation reveals that the clustering framework effectively identifies peak-hour congestion zones and uncovers mobility trends not visible through traditional statistical analysis. DBSCAN demonstrates superior performance in detecting abnormal traffic behaviors due to its ability to isolate low-density anomalies. Visualized cluster distributions assist urban authorities in optimizing signal timing, planning alternate routes, and predicting high-load areas during peak periods.

Conclusion:

This study establishes clustering as an effective mechanism for analyzing and managing urban traffic patterns in smart cities. By uncovering hidden mobility structures and enabling real-time anomaly detection, machine learning-driven traffic clustering supports the development of intelligent, adaptive, and efficient transportation systems. Future work may integrate deep learning, graph neural networks, and edge-based IoT processing to build fully automated traffic prediction and control pipelines.