Machine Learning for (Autonomous) Vehicle Data Analysis

Short description:

This workshop introduces participants to practical machine learning techniques for analyzing autonomous vehicle data. Using real-world sensor datasets (camera, LiDAR, and time-series telemetry), participants will learn how to apply supervised, unsupervised, and deep learning methods to tasks such as object detection, trajectory prediction, and anomaly detection. The workshop combines conceptual explanations with hands-on exercises, highlighting challenges specific to autonomous driving, including data complexity, robustness, and safety considerations. It is aimed at participants interested in applied machine learning for intelligent transportation systems.

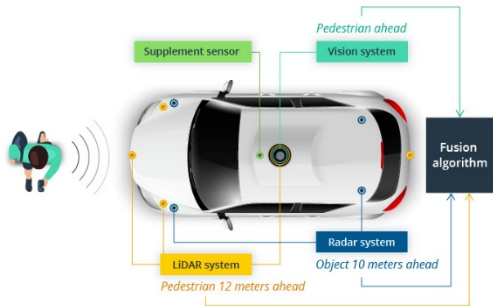
Goal:

Provide participants with a structured overview of concrete ML methods used in autonomous vehicle (AV) data analysis, focusing on why, when, and how each method is applied.

Theoretical part break-down:

PART 1 – Autonomous Vehicle Data & Problem Framing

- AV data ecosystem: camera, LiDAR, radar, GPS, CAN bus, IMU
- Characteristics of AV data:
  - High dimensionality
  - Multi-modality
  - Temporal dependence
- Typical ML problem formulations:
  - Classification (object type, driving events)
  - Regression (speed, distance, time-to-collision)
  - Sequence prediction (trajectories)
- Data quality, labeling challenges, and uncertainty



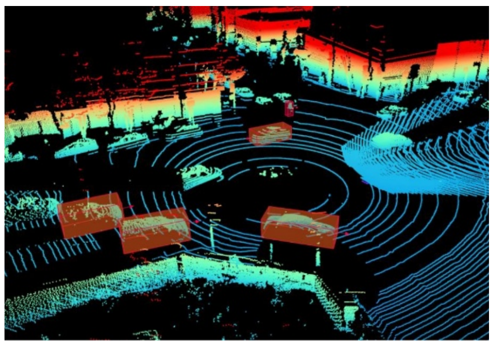
PART 2 – Supervised Learning for Perception & Prediction

Methods:

- Logistic Regression & Regularized Linear Models
- Random Forests
- Gradient Boosting (XGBoost / LightGBM)

Applications:

- Object classification from sensor features
- Driving maneuver recognition
- Short-horizon trajectory prediction



PART 3 – Deep Learning for Visual & Spatial Data

Methods:

- Convolutional Neural Networks (CNNs)
- Transfer learning for AV perception
- Autoencoders for feature extraction

Applications:

- Camera-based object detection
- LiDAR point cloud representations
  (2D projections vs embeddings)
- Dimensionality reduction for downstream tasks

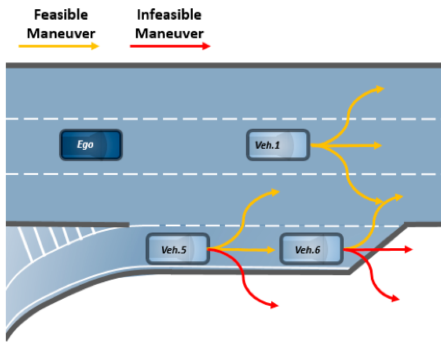
PART 4 – Temporal Models & Behavior Modeling

Methods:

- Hidden Markov Models (HMMs)
- Recurrent Neural Networks (LSTM, GRU)
- Temporal Convolutional Networks
- Transformers for time-series data

Applications:

- Driver behavior modeling
- Trajectory forecasting
- Event detection in driving sequences



PART 5 – Unsupervised Learning, Anomalies & Explainability

Methods:

- K-means, DBSCAN
- Isolation Forest
- Autoencoder-based anomaly detection
- SHAP and feature attribution methods

Applications:

- Detection of unusual driving situations
- Sensor faults and rare events
- Model transparency in safety-critical contexts

Hands-on Practical Workshop:

The hands-on practical workshop is organized as a multiple day intensive program in which participants apply the machine learning methods introduced in the theoretical part to a real-world autonomous driving dataset. Starting from data exploration and preprocessing, participants build supervised baseline models for perception and driving-event classification, then progress to deep learning and temporal models for behavior analysis and trajectory prediction. In the final stage, they address safety-oriented aspects such as anomaly detection, explainability, and robustness analysis. Working in teams on a unified case study, participants develop an end-to-end machine learning pipeline for autonomous vehicle data, gaining practical experience with both model performance and interpretability trade-offs in safety-critical systems.