BENAZIR BHUTTO SHAHEED UNIVERSITY

SENSOR NETWORK FAULT DETECTION SYSTEM
(PROJECT REPORT)

Project Submission Details

Submitted by: Ali Hassan, Abdul Wahab, Anas Hussain

Roll Number: B2331012, B2331006, B2331018

Supervisor: PROFESSOR ANWAR ALI

University Name: BBSU

Year: 2025

Project Objective:

To develop a system that identifies faulty vehicle sensor readings in real-time. Sensors include accelerometer axes (Accel_X, Accel_Y, Accel_Z), GPS speed (GPS_Speed), and engine RPM (Engine_RPM). The system should detect anomalies, flag likely faulty sensors, and retain extreme outlier values during preprocessing.

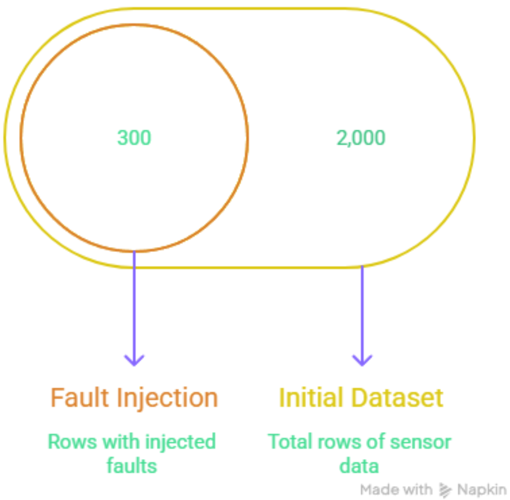
1. Data Generation

- **Dataset Size:** 2,000 rows
- **Sensors:** RPM, GPS_Speed, Accel_X, Accel_Y, Accel_Z
- **Timestamps:** Randomized within Jan 1, 2024 – Dec 31, 2024, sorted chronologically
- **Fault Injection:**
  - 15% of all sensor cells were injected with extreme random faults (spikes or drops)
  - Fault magnitude: ±15–30× standard deviation of the respective sensor

Example Row (Faulty Included):

Timestamp	RPM	GPS_Speed	Accel_X	Accel_Y	Accel_Z
2024-08-19 06:26:27	1835	50	0.0035	-0.0455	9.813

- Data saved as: sensor_fault_dataset_v2.csv



2. Data Preprocessing

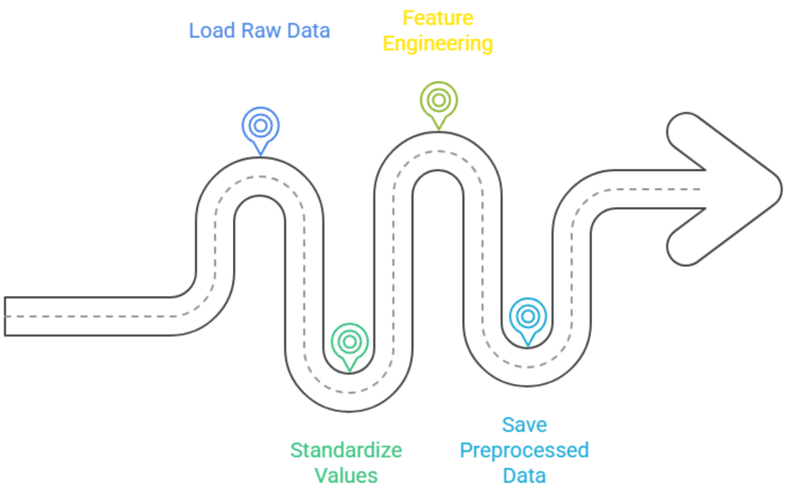
Goals:

- Retain all original sensor readings, including faulty values
- Standardize features for anomaly detection
- Avoid any imputation or filtering that might remove outliers

Steps Taken:

1. Load Raw Data
2. Standardize Sensor Values (using StandardScaler)
3. Feature Engineering:
    - Added Delta_GPS_Speed = GPS_Speed.diff()
    - Added Speed_RPM_Ratio = GPS_Speed / Engine_RPM
4. Save Preprocessed Data:
    - File: vehicle_sensor_data_preprocessed.csv

Key Note: Extreme values (faults) were preserved to ensure anomalies remain detectable.



3. Model Comparison and Selection

Models Evaluated:

1. PCA (Principal Component Analysis – Linear Reconstruction)

- Computes reconstruction error to detect deviations from normal sensor correlations

2. Isolation Forest (Tree-based Anomaly Detection)
   - Learns normal multivariate correlations
   - Flags anomalies as deviations from learned patterns
3. Gaussian Mixture Model (Probabilistic Density Estimation)
   - Uses probability density to detect low-likelihood observations as anomalies

Ground Truth Anomalies: - Rows containing extreme faults: 63 anomalies identified

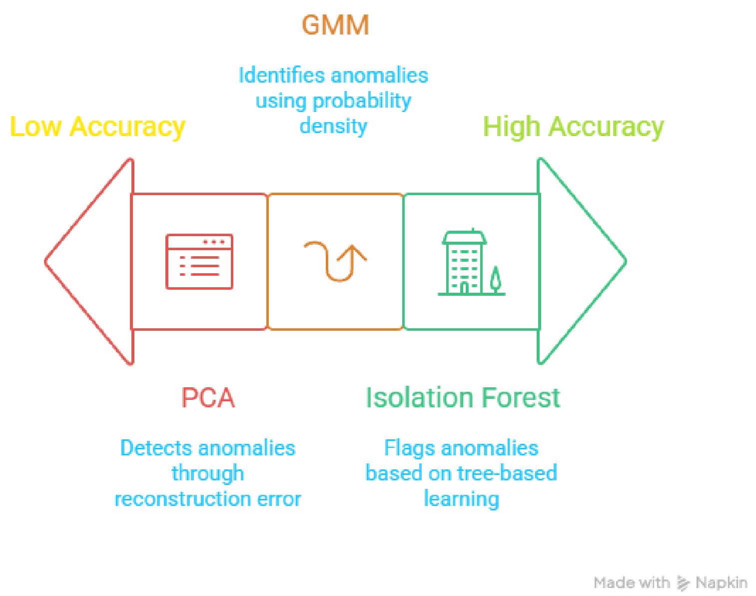
Training Process: - Scaled sensor data using StandardScaler - Trained all three models on the preprocessed dataset - Evaluated using AUC-ROC score (higher = better detection)

Evaluation Results:

Model	AUC-ROC Score
Isolation Forest	0.9964
GMM (Density Estimate)	0.9201
PCA (Reconstruction Error)	0.4721

Conclusion:

- Best Model: Isolation Forest (AUC = 0.9964)
- Isolation Forest successfully captures normal correlations and flags outliers without removing extreme values
- Proceeded with Isolation Forest for fault attribution



4. Fault Attribution Logic

Objective: Identify which sensor(s) caused the anomaly.

Method:

1. Compute the net acceleration magnitude:

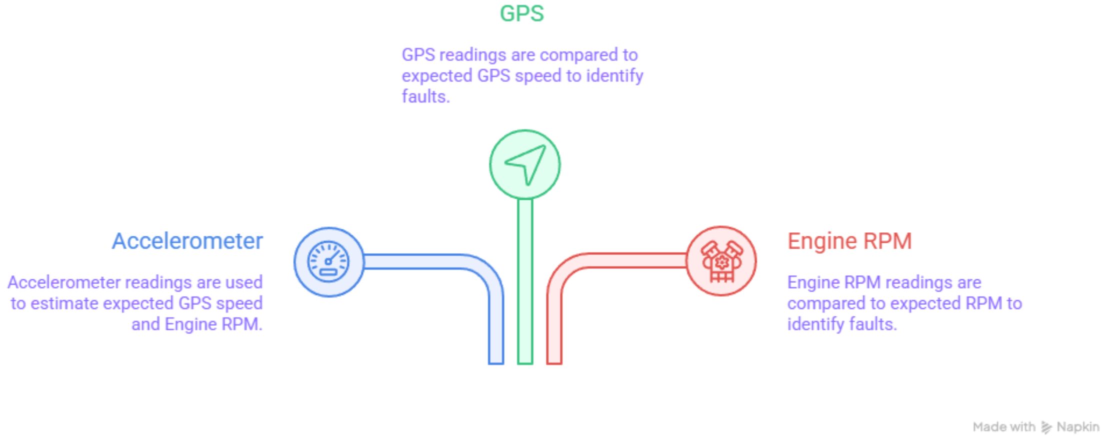
accel_mag = sqrt(Accel_X^2 + Accel_Y^2 + (Accel_Z - 9.8)^2)

2. Estimate expected GPS speed and Engine RPM from accelerometer:

Expected_GPS = k1 * accel_mag Expected_RPM = k2 * accel_mag

- Constants k1 and k2 are calculated from training data

3. Flag a sensor as faulty if its actual reading deviates more than 50% from the expected value
4. Combine physics-based attribution with Isolation Forest anomaly score to identify likely faulty sensors



5. Model Deployment

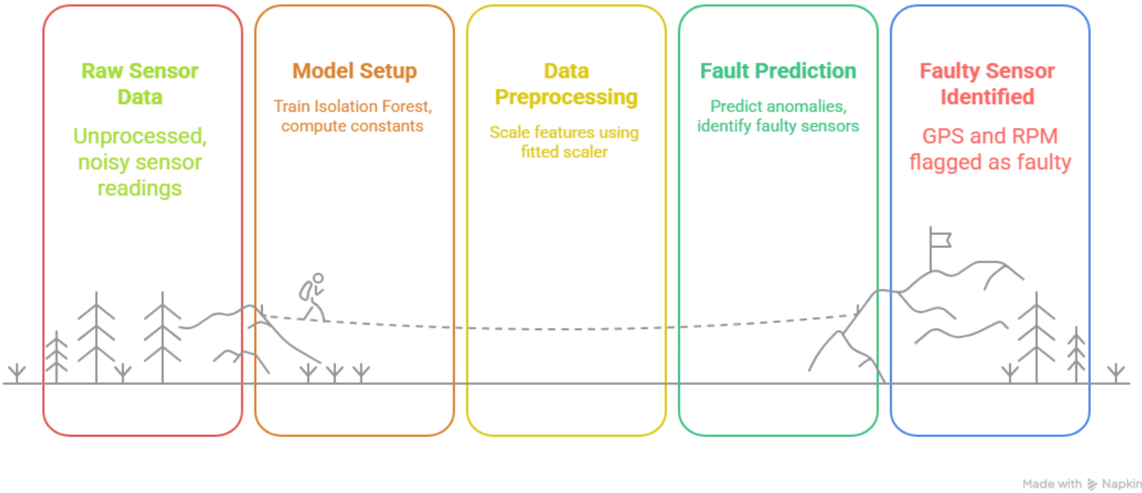
Workflow:

1. Setup Model: Train Isolation Forest, compute constants (k1, k2)
2. Preprocess Incoming Data: Scale features using previously fitted scaler
3. Predict Fault:
   - Input: 5 sensor readings (Accel_X, Accel_Y, Accel_Z, GPS_Speed, Engine_RPM)
   - Output:
     - Anomaly detected?
     - Faulty sensor(s)
     - Expected vs actual values

Manual Test Example:

Sensor	Actual	Expected	Faulty
GPS_Speed	40.0	15.2	Yes
RPM	2600	1000	Yes
Accel_X/Y/Z	0.5/0.1/9.8	-	No

Result: GPS and RPM flagged as faulty due to significant deviation.



6. Key Achievements

- Generated a realistic dataset with 2000 rows and injected extreme faults
- Preserved faults during preprocessing, enabling accurate anomaly detection
- Compared 3 models and selected Isolation Forest for its high detection accuracy (AUC=0.9964)
- Developed physics-based correlation formulas to attribute faults to specific sensors
- Implemented real-time fault detection interface with clear reporting

7. Future Work

- Include temporal correlations (time series analysis) for better anomaly detection
- Integrate sensor fusion techniques for more robust fault attribution
- Deploy in real vehicle telemetry systems for live sensor monitoring

Summary:

This project demonstrates a robust end-to-end workflow for detecting faulty vehicle sensors. By combining tree-based anomaly detection with physics-informed correlations, it reliably flags anomalous readings and attributes faults to specific sensors while preserving extreme values for analysis.