(IFIP) Networking 2025 Conference
(NETWORKING 2025)
26-28 May, Limassol Cyprus

Interpretable Outlier and Anomaly Detection for Mobile Networks from Small Tabular Data

Juan Marcos Ramírez*, Pablo Rojo[a], Vincenzo Mancuso[b]*, and Antonio Fernández-Anta[c]*

*IMDEA Networks Institute, [a]Nokia MN, [b]University of Palermo, [c]IMDEA Software Institute

[Developing the Science of Networks]

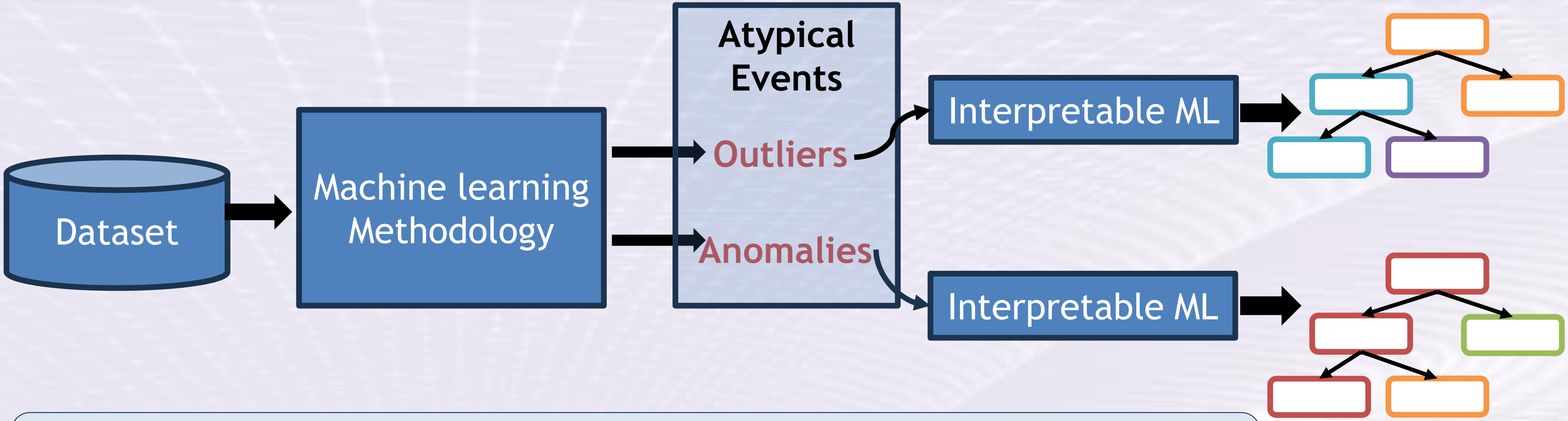
Motivation

- 5G networks provide high-speed connectivity and fast data transfer rates
- The growing size and complexity pose a challenge for monitoring network functioning.
- Operators consistently perform measurement campaigns to assess network performance.
- Drive-test is a strategy that captures datasets with many variables and limited samples.
- The drive-test data structure makes deep learning approaches impractical.

How We Address Atypical Event Detection in Mobile Networks

Methodology based on classical machine learning to **detect atypical events** (outliers & anomalies) from drive test data.

Interpretable machine learning module to describe which variables and samples are most strongly related to atypical events.

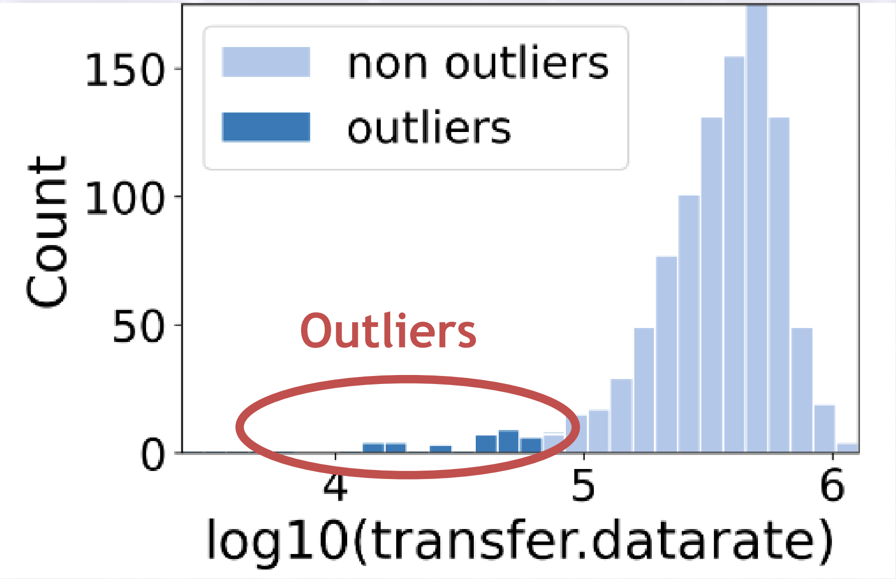


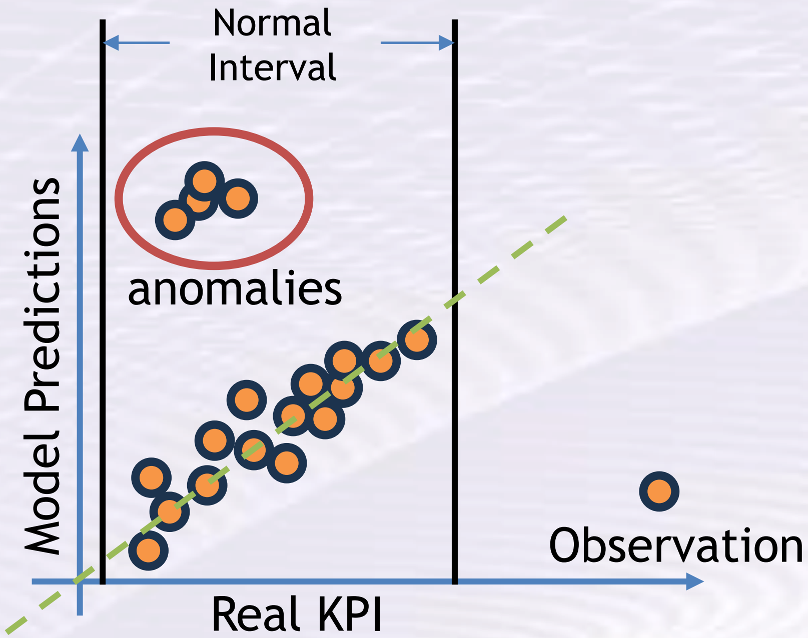
ROAD Index: a novel metric that quantifies the relationship between the pattern of atypical events and underlying binary patterns in features

Definitions

Outliers: KPI values significantly deviate from normal behavior (low throughput, and long session duration).

Anomalies: KPI values within expected range but severely deviated from model predictions.

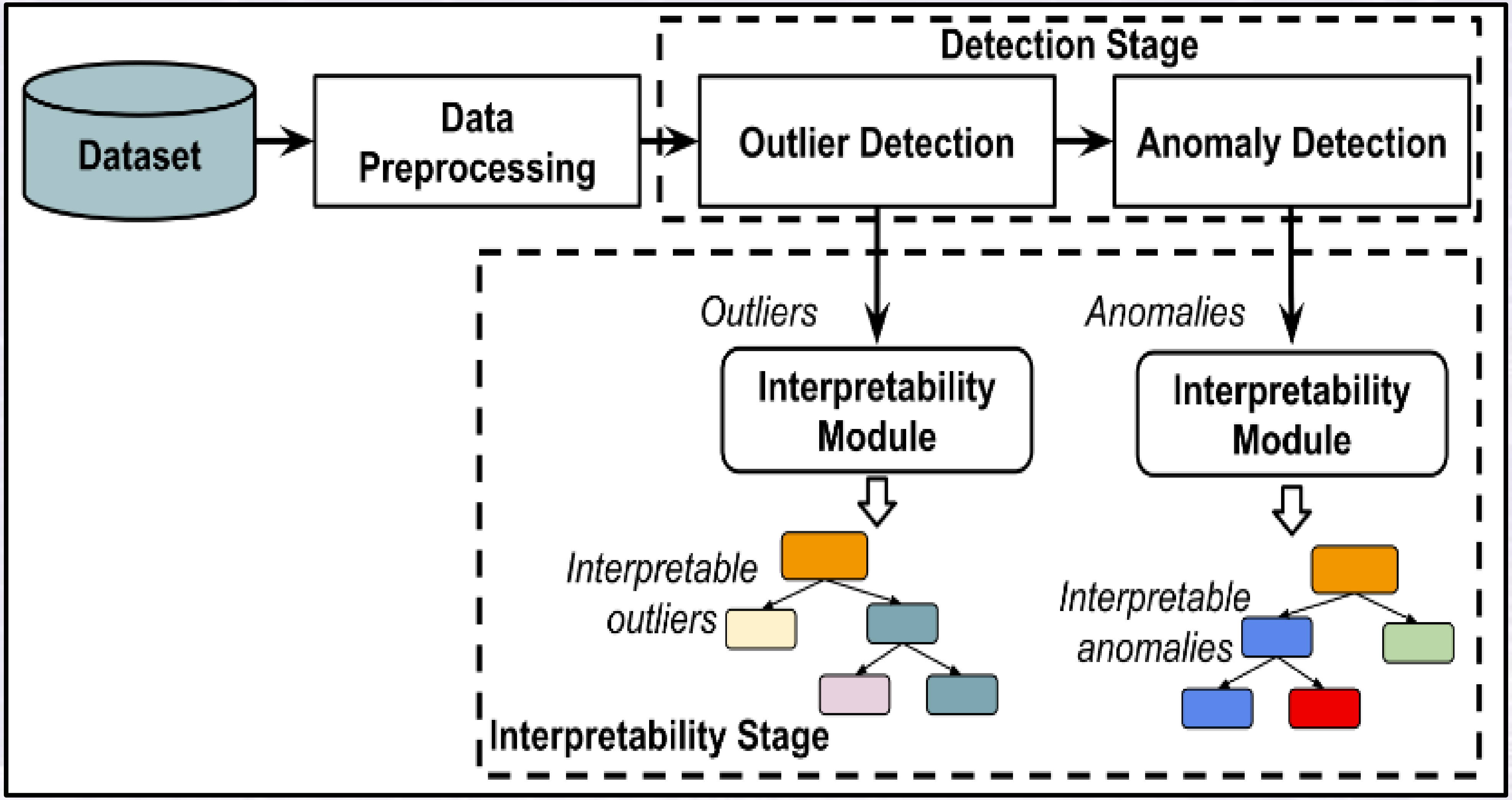




These atypical samples may indicate severe performance degradation or faulty network components.

These observations are typically attributed to suboptimal operational conditions (unexpected delays or resource allocation)

Interpretable Outlier and Anomaly Detection (ROAD) Methodology Overview

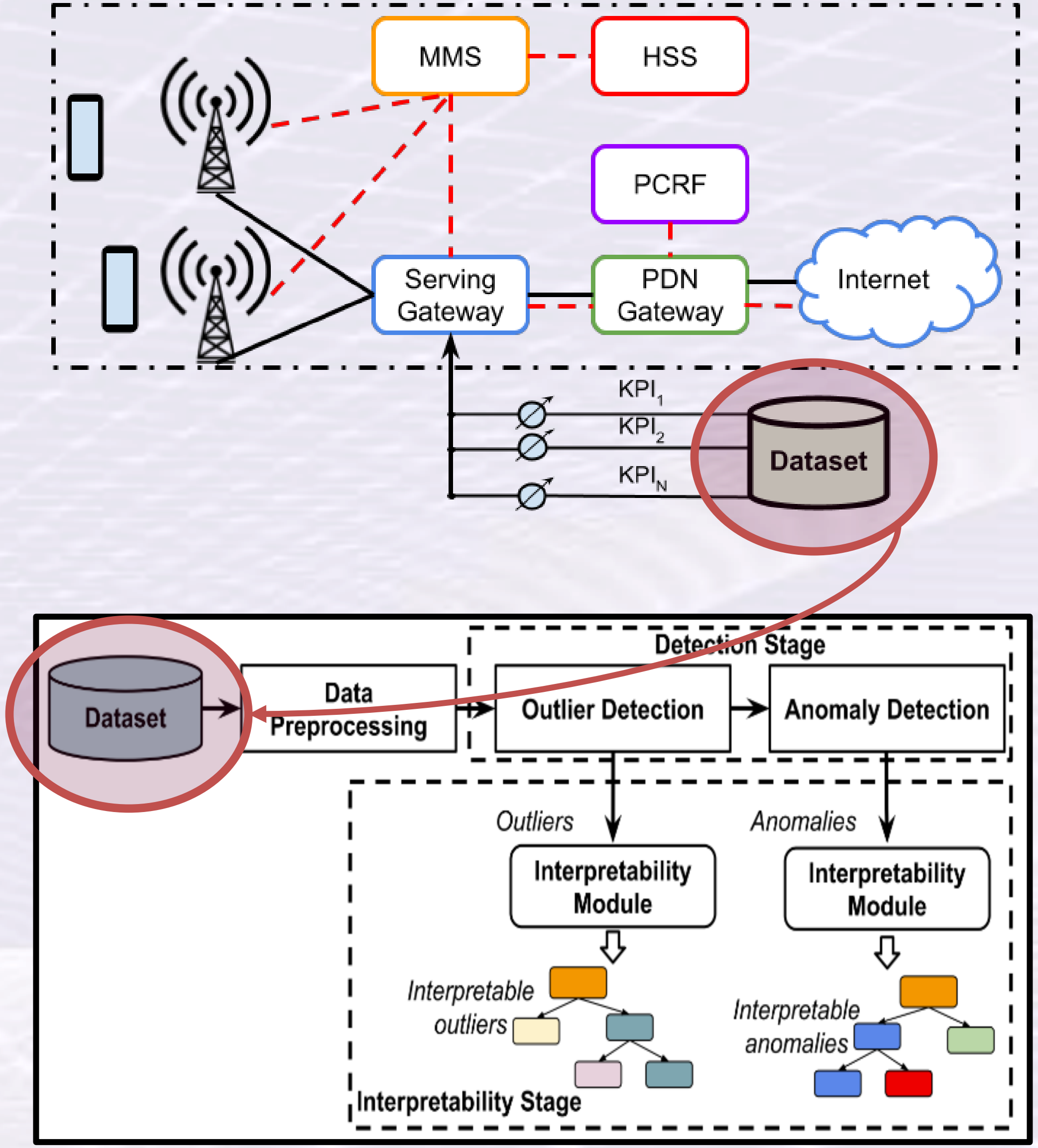


Flowchart of the ROAD methodology

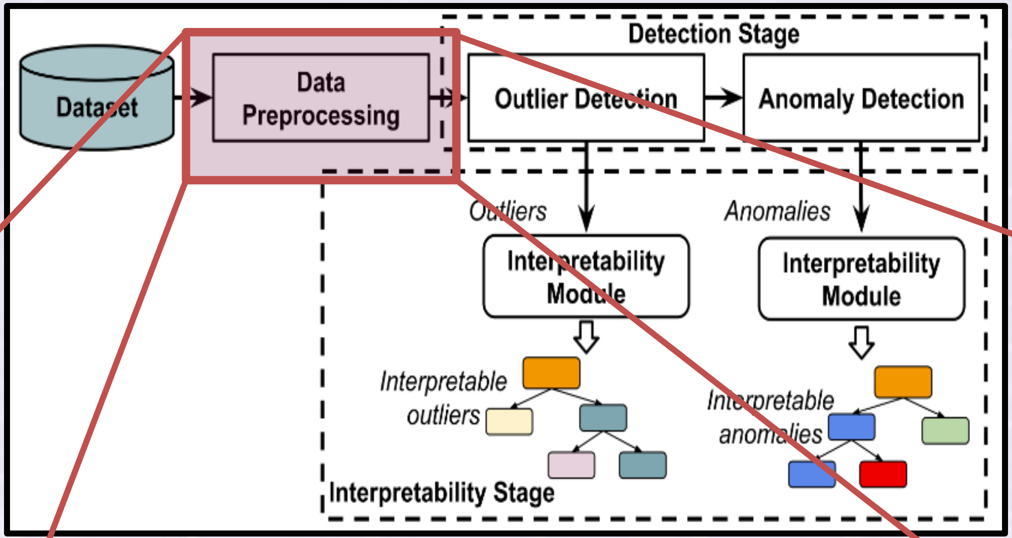
Drive-tests in Mobile Networks

- Field-based performance measurement methodology.
- Collected by mobile test equipment (e.g., smartphones, modems in vehicles).
- Captures KPIs: throughput, latency, signal strength, etc.
- Provides ground truth for real network experience
- Limited spatial and temporal coverage
- Small sample sizes limit deep learning application

Mobile Networks Components and Data Collection



Data Preprocessing



1. Data Cleaning

- **Handle missing values**
  - Drop features with > 1% of null cells
  - Fill remaining null cells with the column **median**
- **Remove non-informative features**
  - Drop features with constant values (zero variance)

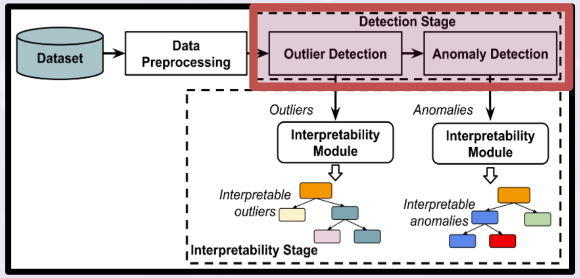
2. Multicollinearity reduction

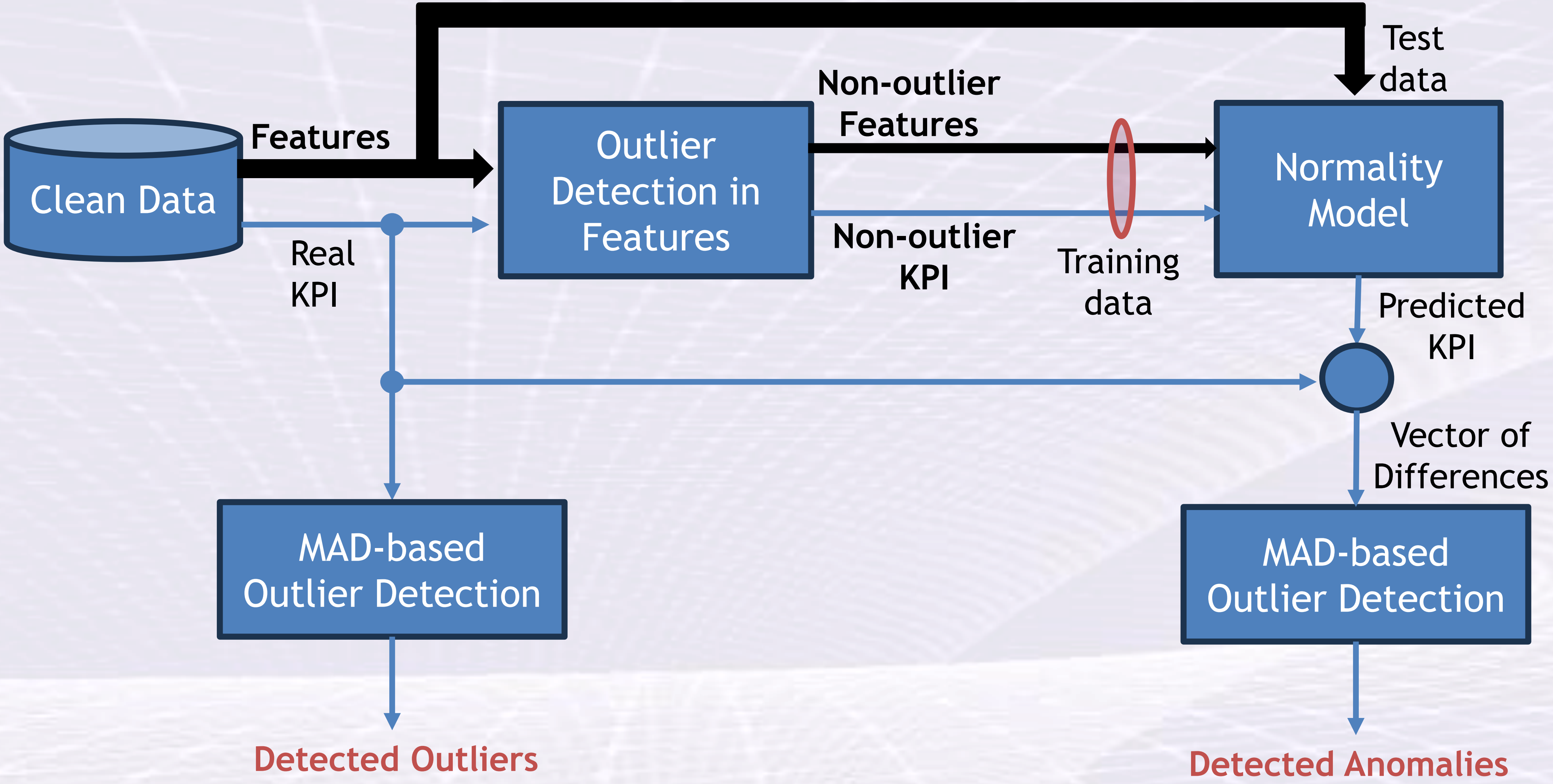
- **Variance Inflation Factor (VIF)**
  - Iteratively remove features with **VIF >1000**
  - Stops when all VIF values are below the threshold

3. Filter KPI-correlated features

- Build a **vector of correlation factors** between the KPI and features
- **Cluster the vector of correlation factors** (model-based clustering)
- **Removes the cluster with the highest correlation to the KPI**

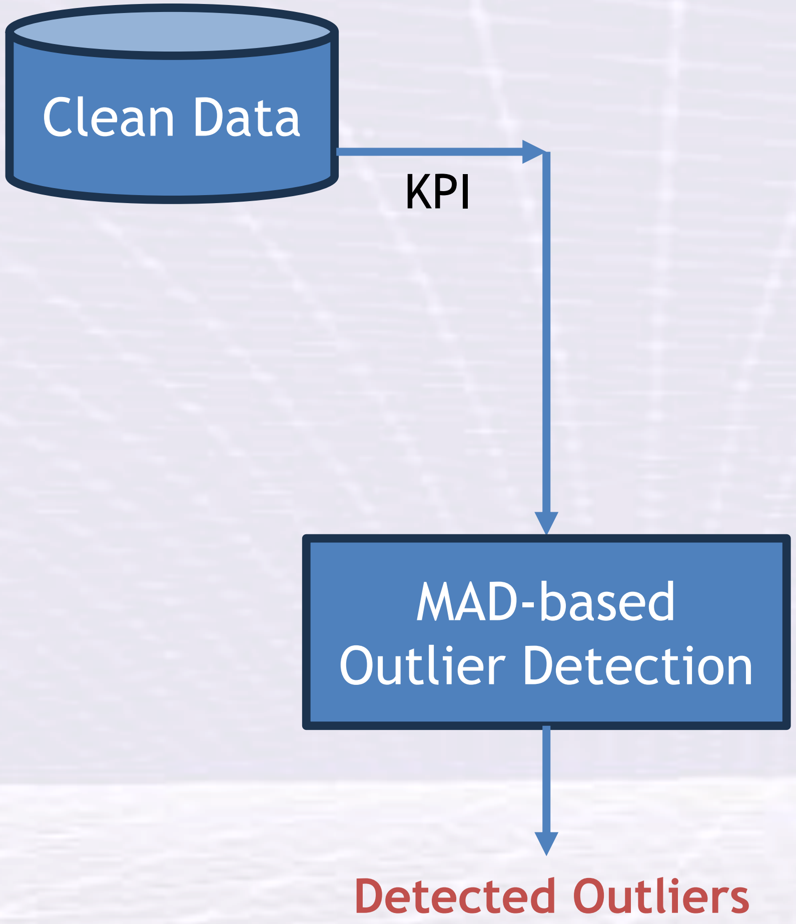
Outlier and Anomaly Detection





Outlier Detection

- We detect outliers by analyzing KPI deviations directly.



Step 1: Apply logarithmic transformation to the KPI values.

- Helps separate extreme low-performance values.

Step 2:Compute the Median Absolute Deviation (MAD) using:

- l_y represents the logarithmic version of the KPI vector

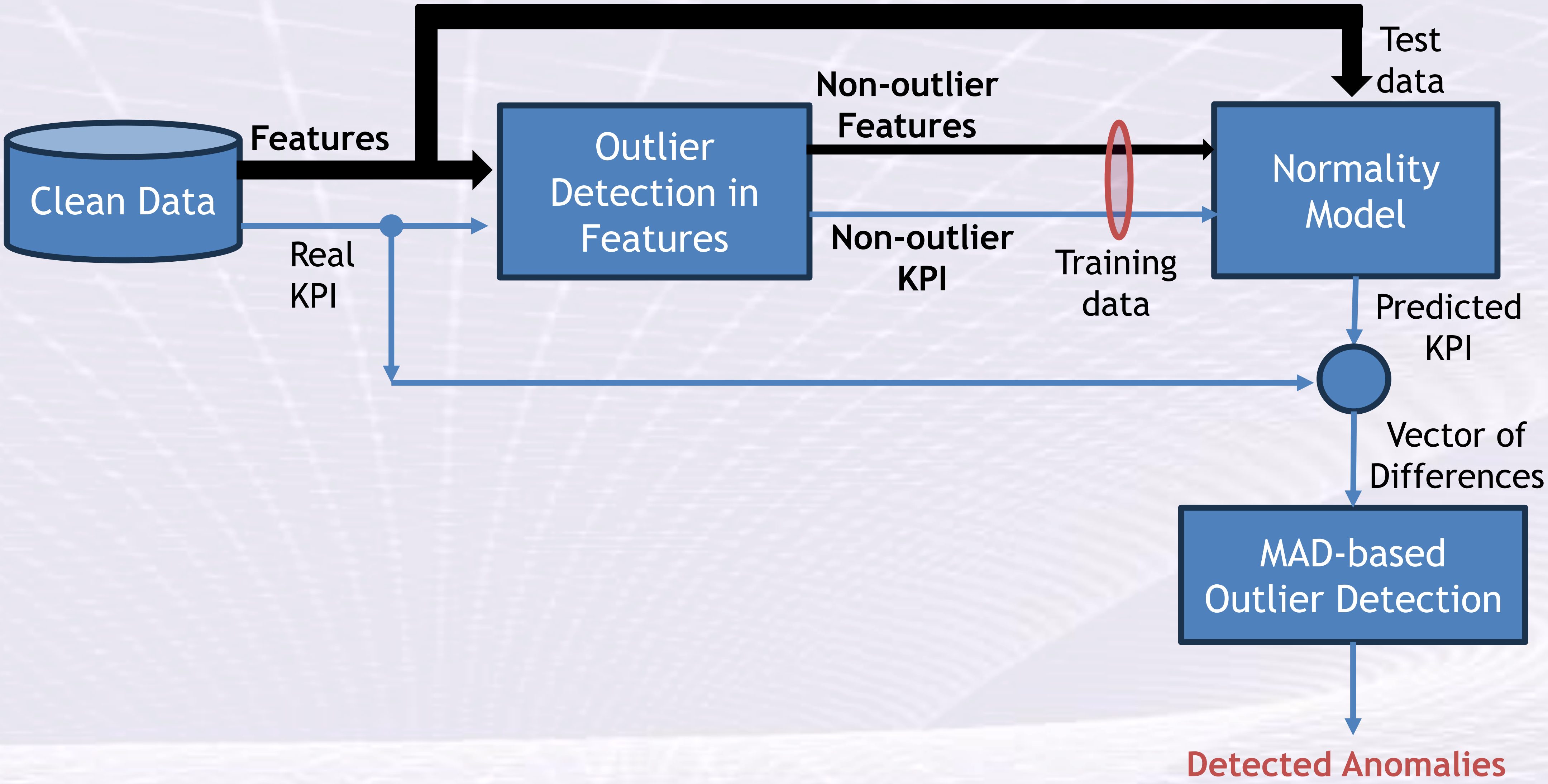
$$\mathrm{MAD} = \frac{1}{Q_3}\mathrm{median}\left(|\mathbf{l_y} - \mathrm{median}(\mathbf{l_y})|\right)$$

Step 3:Use the MAD criterion to detect outliers [1].

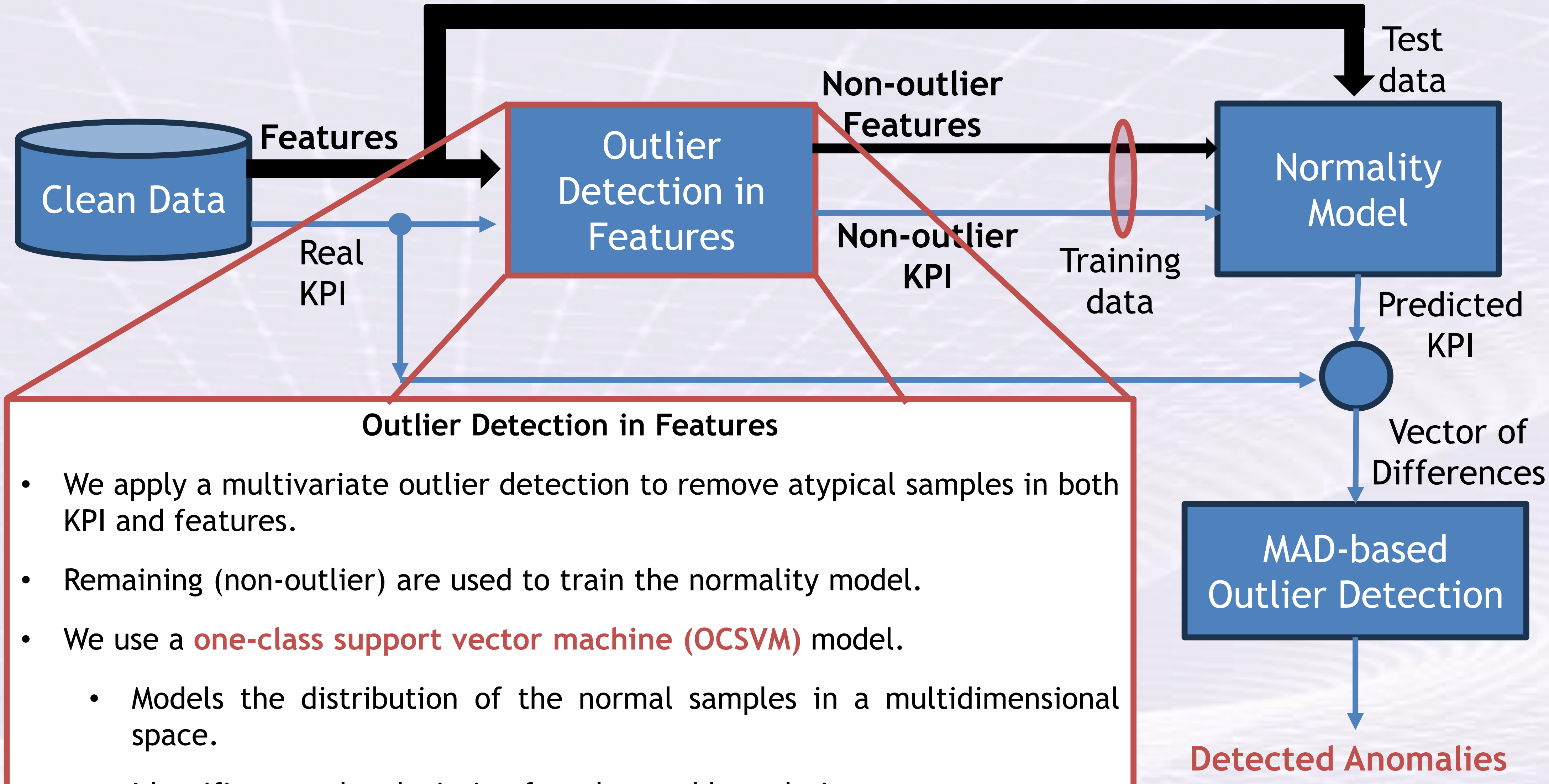
- A sample (l_{yi}) is detected as outlier if:

$$\mathrm{median}(\mathbf{l_y}) - 3 \cdot \mathrm{MAD} < l_{y_i} < \mathrm{median}(\mathbf{l_y}) + 3 \cdot \mathrm{MAD}$$

Anomaly Detection



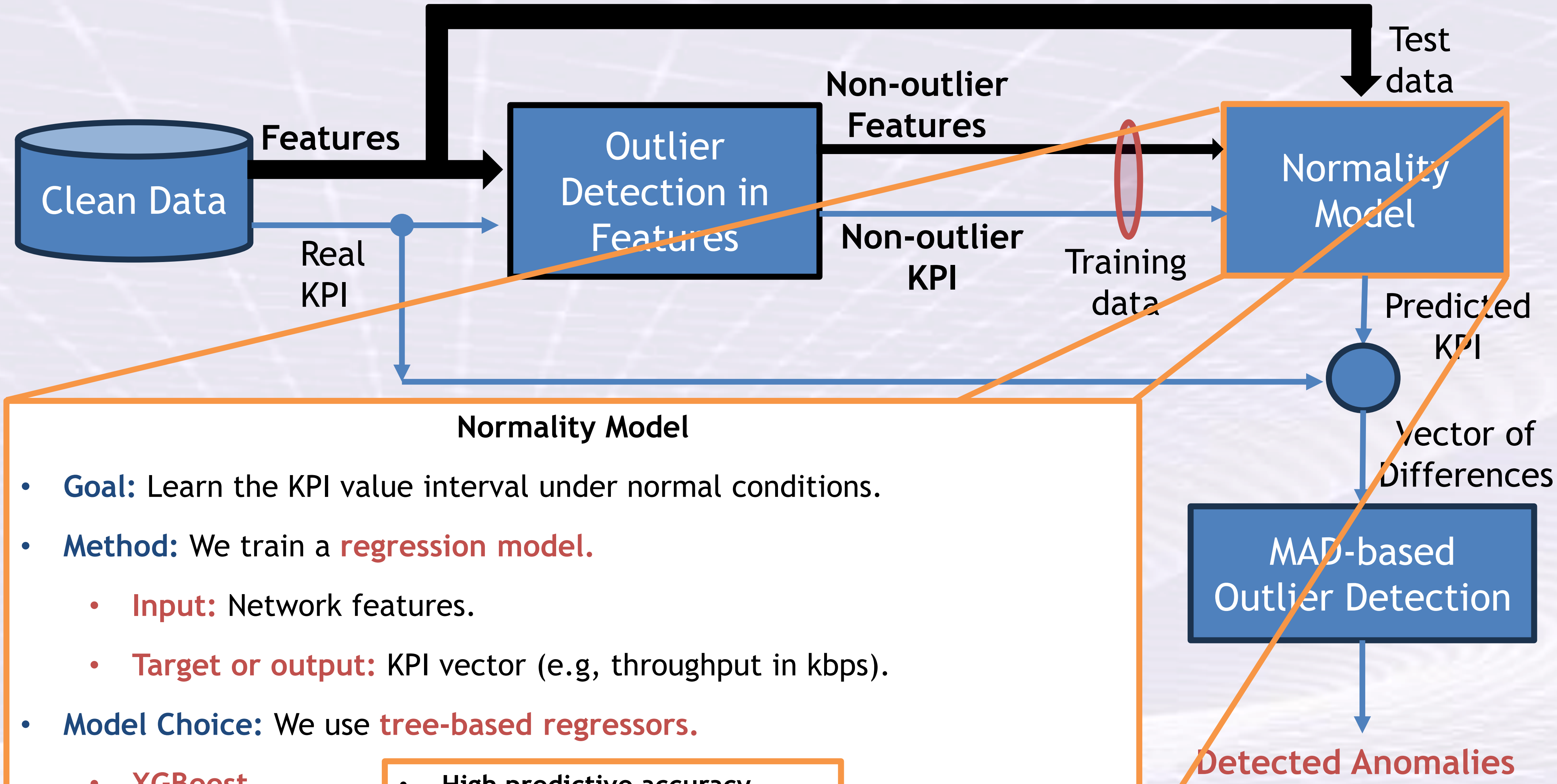
Anomaly Detection



Outlier Detection in Features

- We apply a multivariate outlier detection to remove atypical samples in both KPI and features.
- Remaining (non-outlier) are used to train the normality model.
- We use a one-class support vector machine (OCSVM) model.
  - Models the distribution of the normal samples in a multidimensional space.
  - Identifies samples deviating from learned boundaries.

Anomaly Detection

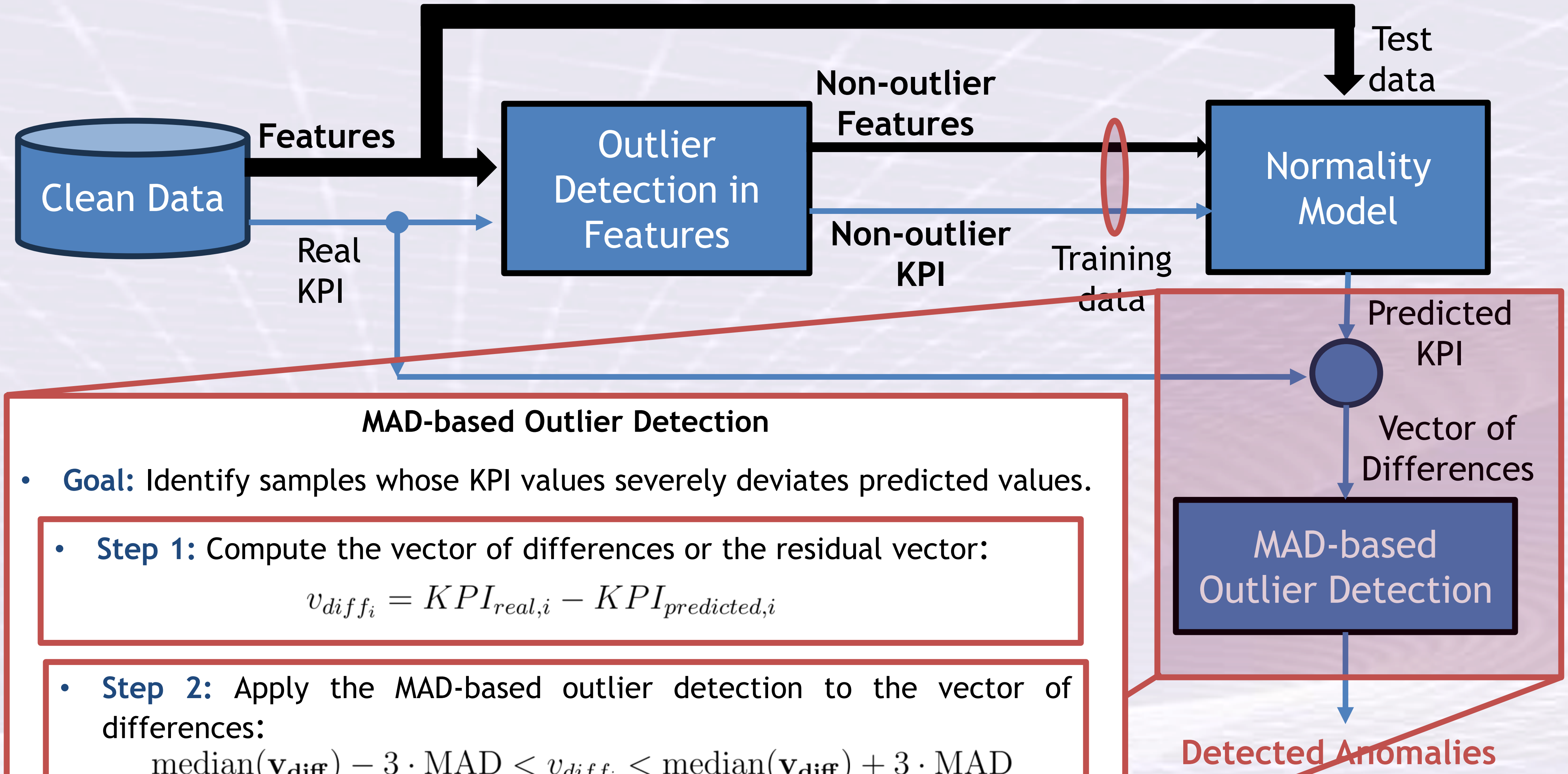


Normality Model

- **Goal:** Learn the KPI value interval under normal conditions.
- **Method:** We train a **regression model.**
  - **Input:** Network features.
  - **Target or output:** KPI vector (e.g, throughput in kbps).
- **Model Choice:** We use **tree-based regressors.**
  - **XGBoost.**
  - **LightGBM**

- **High predictive accuracy**
- **Support interpretability tools**

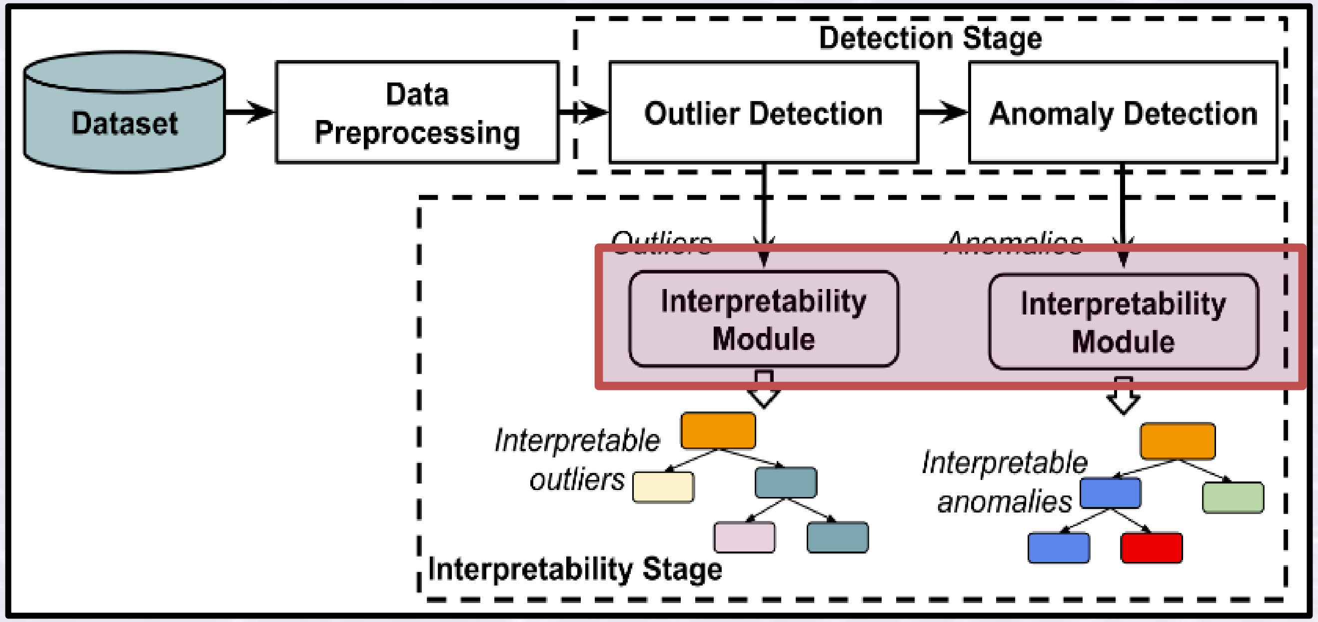
Anomaly Detection



MAD-based Outlier Detection

- **Goal:** Identify samples whose KPI values severely deviates predicted values.
  - **Step 1:** Compute the vector of differences or the residual vector:
    $$v_{diff_i} = KPI_{real,i} - KPI_{predicted,i}$$
  - **Step 2:** Apply the MAD-based outlier detection to the vector of differences:
    $$\mathrm{median}(\mathbf{v_{diff}}) - 3 \cdot \mathrm{MAD} < v_{diff_i} < \mathrm{median}(\mathbf{v_{diff}}) + 3 \cdot \mathrm{MAD}$$

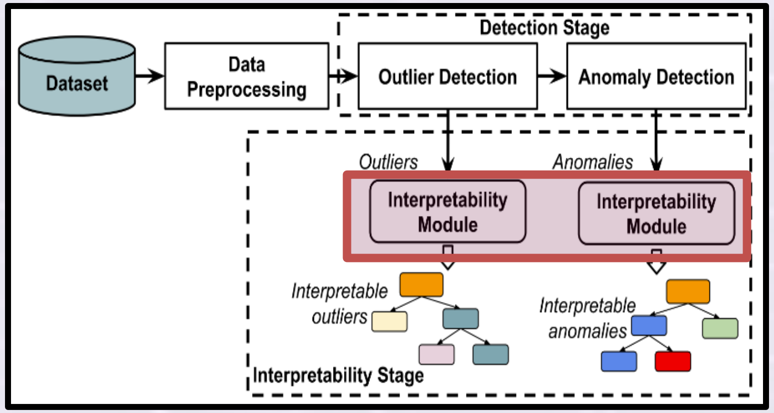
Interpretability Module: Understanding Atypical Behavior



- **Objective:** identify features and samples most associated with atypical events (outliers or anomalies).
- **Operation:** Runs **separately** for outliers and anomalies.
- **Composed of three key sub-phases:**

ROAD-based Feature Selection → Unsupervised Sample Labeling → Decision Tree Classification

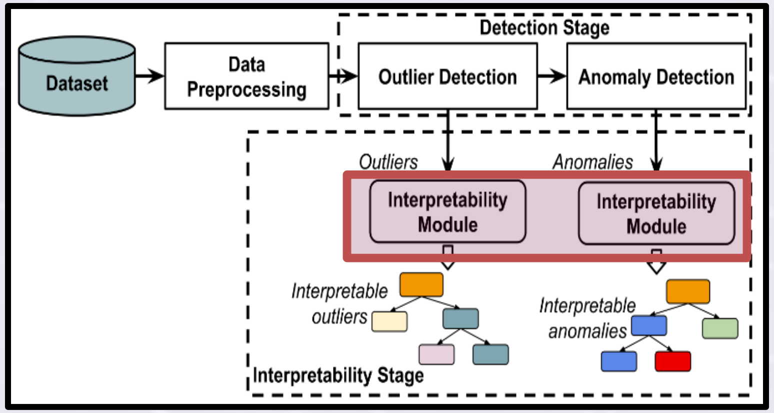
Interpretability Module



ROAD-based Feature Selection

- **1D Clustering**
  Apply a series of 1D model-based clustering to each feature.
- **Binary Pattern Extraction**
  Convert each clustering into binary patterns.
- **ROAD Index Computation**
  For each feature, the ROAD index is the maximum Jaccard between the atypical events (outliers or anomalies) and binary patterns obtained from 1D clustering

Interpretability Module

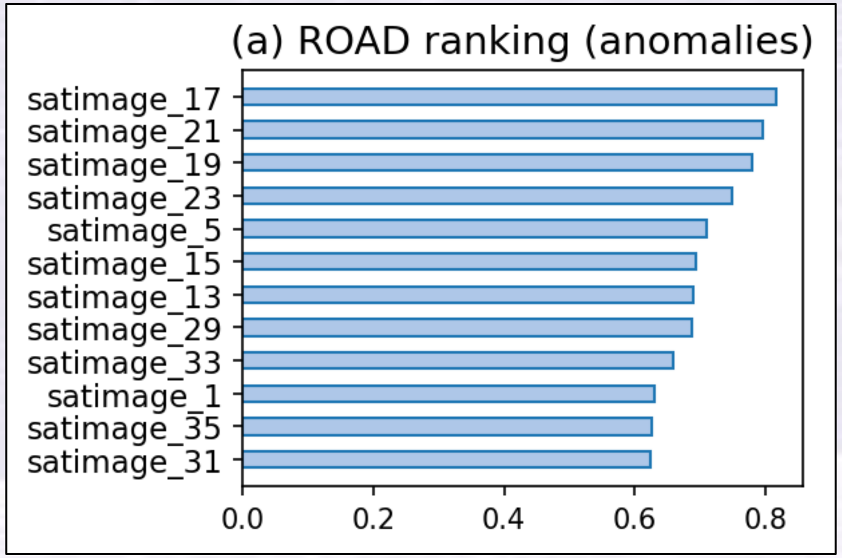


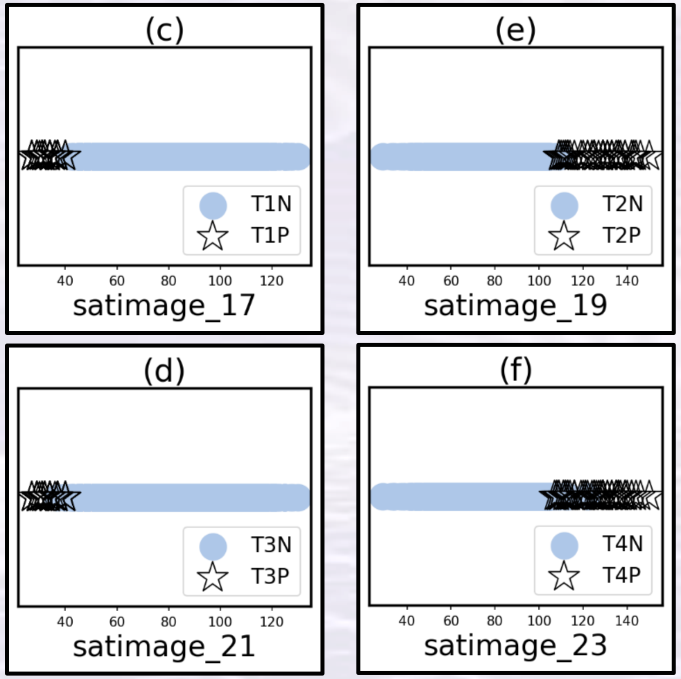
ROAD-based Feature Selection

- **1D Clustering**
  Apply a series of 1D model-based clustering to each feature.
- **Binary Pattern Extraction**
  Convert each clustering into binary patterns.
- **ROAD Index Computation**
  For each feature, the ROAD index is the maximum Jaccard between the atypical events (outliers or features) and binary patterns obtained from 1D clustering

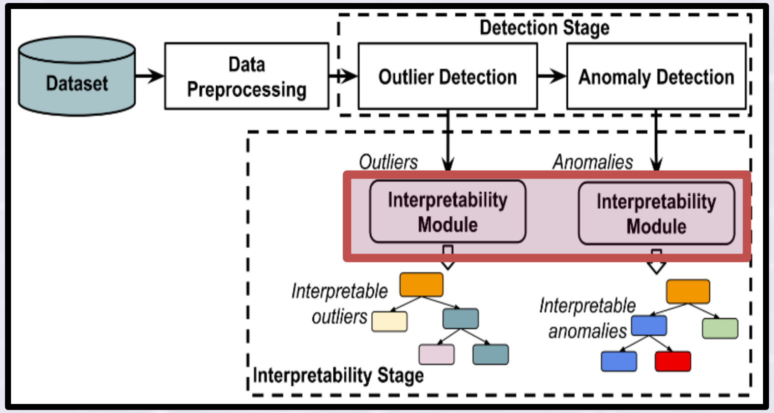
ROAD Index

- **Range:** 0 (no relation) → 1 (strong alignment with atypical samples)
- **Interpretation:** A high ROAD score means the feature is strongly associated with outliers/anomalies.





Interpretability Module

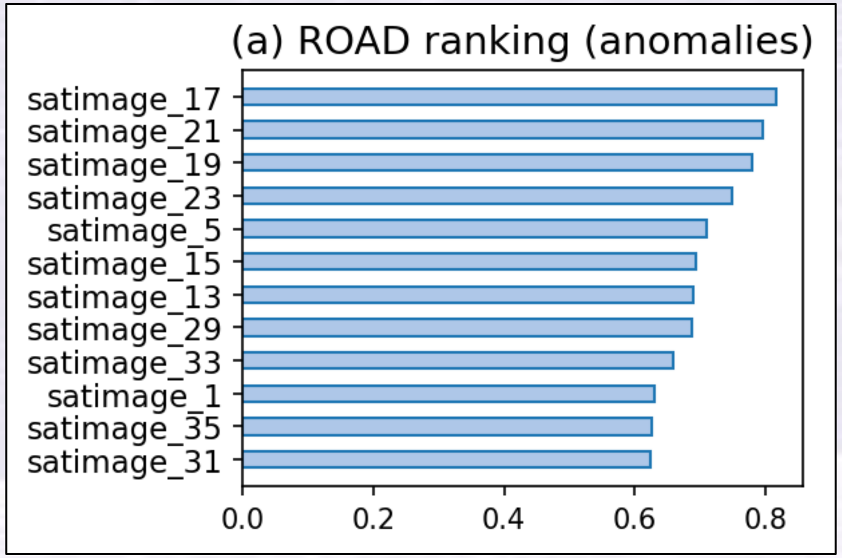


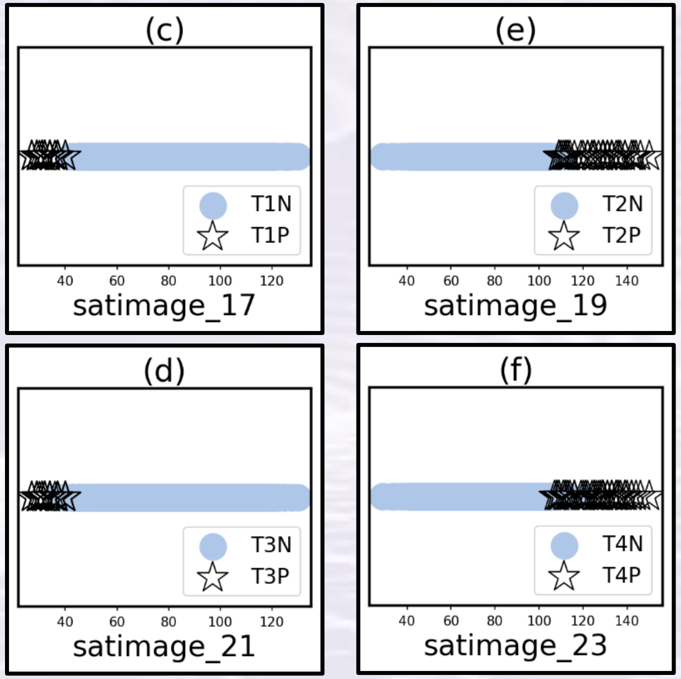
ROAD-based Feature Selection

- **1D Clustering**
  Apply a series of 1D model-based clustering to each feature.
- **Binary Pattern Extraction**
  Convert each clustering into binary patterns.
- **ROAD Index Computation**
  For each feature, the ROAD index is the maximum Jaccard between the atypical events (outliers or features) and binary patterns obtained from 1D clustering
- **Feature Selection**
  Select features with the largest ROAD indices.

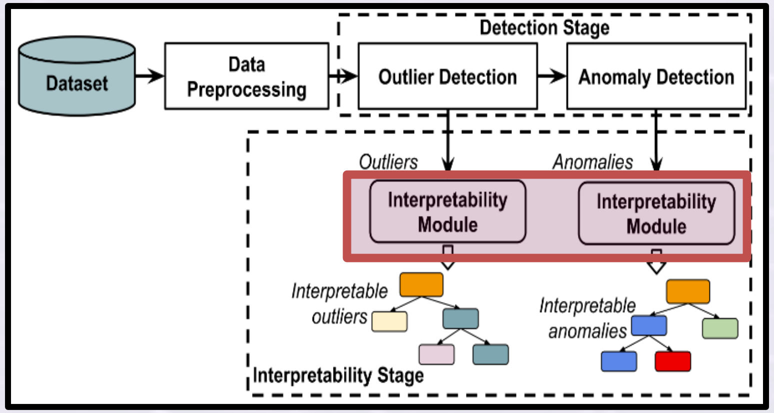
ROAD Index

- **Range:** 0 (no relation) → 1 (strong alignment with atypical samples)
- **Interpretation:** A high ROAD score means the feature is strongly associated with outliers/anomalies.





Interpretability Module



ROAD-based Feature Selection

- **1D Clustering**
Apply a series of 1D model-based clustering to each feature.

- **Binary Pattern Extraction**
Convert each clustering into binary patterns.

- **ROAD Index Computation**
For each feature, the ROAD index is the maximum Jaccard between the atypical events (outliers or features) and binary patterns obtained from 1D clustering

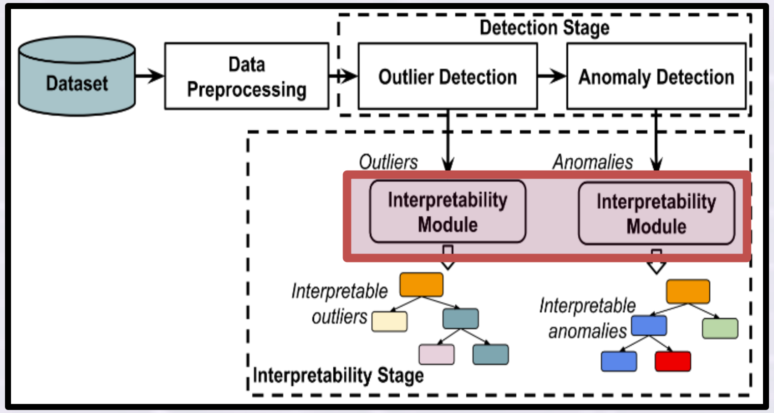
- **Feature Selection**
Select features with the largest ROAD indices.

Unsupervised Pattern Labeling

Each Feature has a binary pattern associated with the ROAD index.

- **Problematic Samples**
For each feature:
If the binary pattern is 1
  - Problematic Sample=True (TP$)
Otherwise
  - Problematic Sample=False (TN$)

- **Labeling Construction**
The module concatenates the flags for each sample:
If all flags are False:
  - Label = ‘Compliant’



Interpretability Module

ROAD-based Feature Selection

- **1D Clustering**

Apply a series of 1D model-based clustering to each feature.

- **Binary Pattern Extraction**

Convert each clustering into binary patterns.

- **ROAD Index Computation**

For each feature, the ROAD index is the maximum Jaccard between the atypical events (outliers or features) and binary patterns obtained from 1D clustering

- **Feature Selection**

Select features with the largest ROAD indices.

Unsupervised Pattern Labeling

Each Feature has a binary pattern associated with the ROAD index.

- **Problematic Samples**

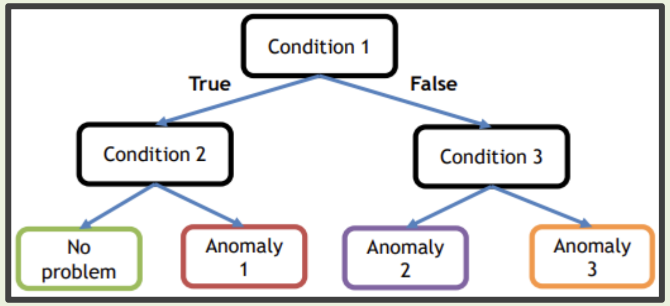
For each feature:
If the binary pattern is 1
- Problematic Sample=True (TP$)

Otherwise
- Problematic Sample=False (TN$)

- **Labeling Construction**

The module concatenates the flags for each sample:
If all flags are False:
- Label = ‘Compliant’

Decision Tree Classification



Decision trees are **interpretable** models and easy to understand

- **Input:** Network Features
- **Target:** Built Labels

Decision trees can be analyzed to understand which parameters and thresholds are associated with atypical events (**Interpretability!!!**).

Datasets

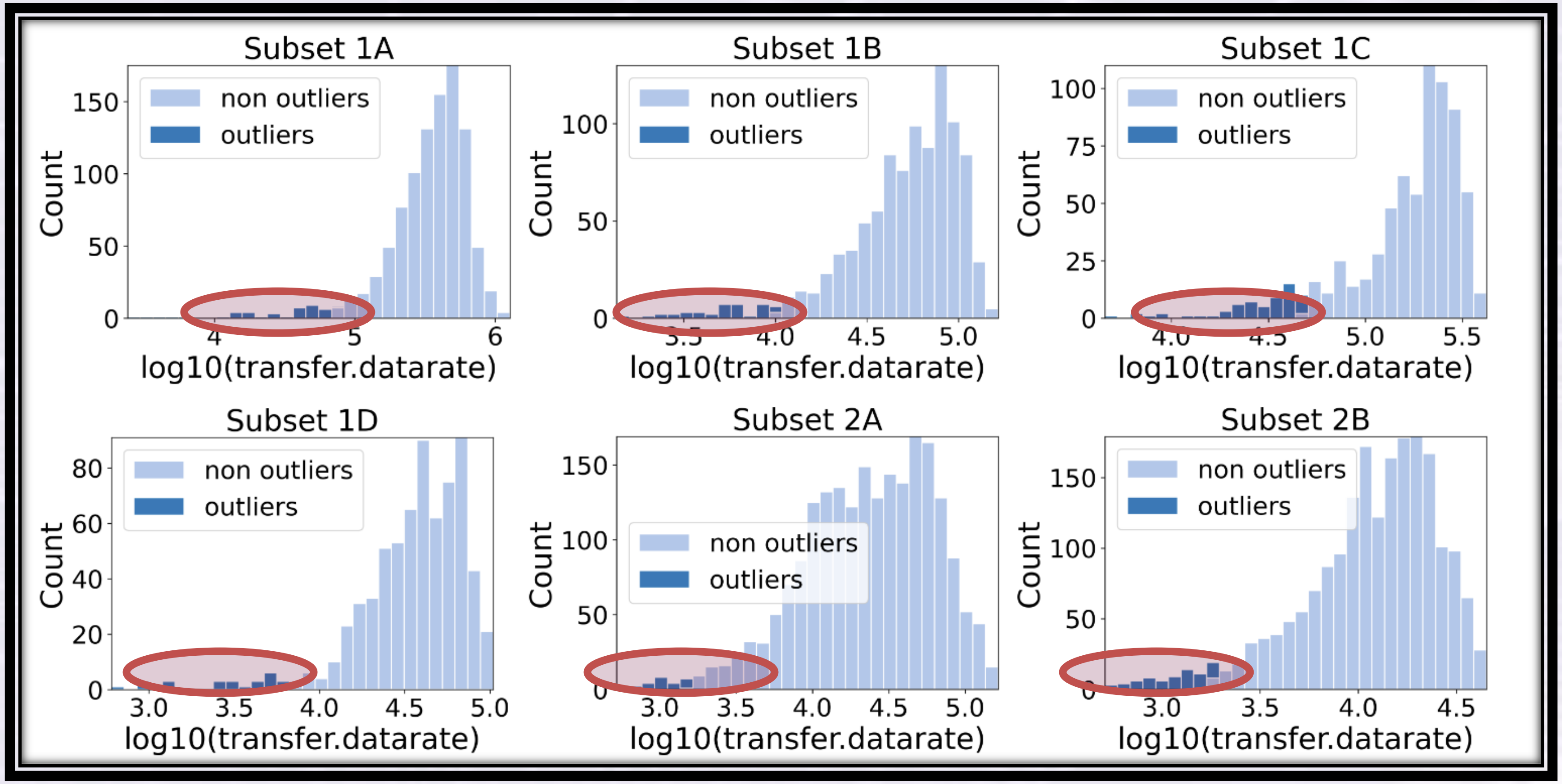
- We use two datasets (Dataset 1 and Dataset 2) collected by Nokia during drive-test campaigns across various mobile networks in Europe
  - Dataset 1 was divided into four subsets (Subset 1A, 1B, 1C, and 1D), with each subset corresponding to a specific type of test.
  - Dataset 2 was segmented into two subsets (Subset 2A and 2B).

TABLE I
SUMMARY OF THE CHARACTERISTICS OF EACH DATASET USED TO EVALUATE THE PERFORMANCE OF THE PROPOSED METHODOLOGY.

Dataset	Subset	Test Type	Rows	Cols
Dataset 1	Subset 1A	Capacity DL	1001	189
	Subset 1B	Capacity UL	971	161
	Subset 1C	HTTP Transfer DL	709	188
	Subset 1D	HTTP Transfer UL	688	166
Dataset 2	Subset 2A	HTTP File DL	2095	119
	Subset 2B	HTTP File UL	1985	102

Outlier Detection

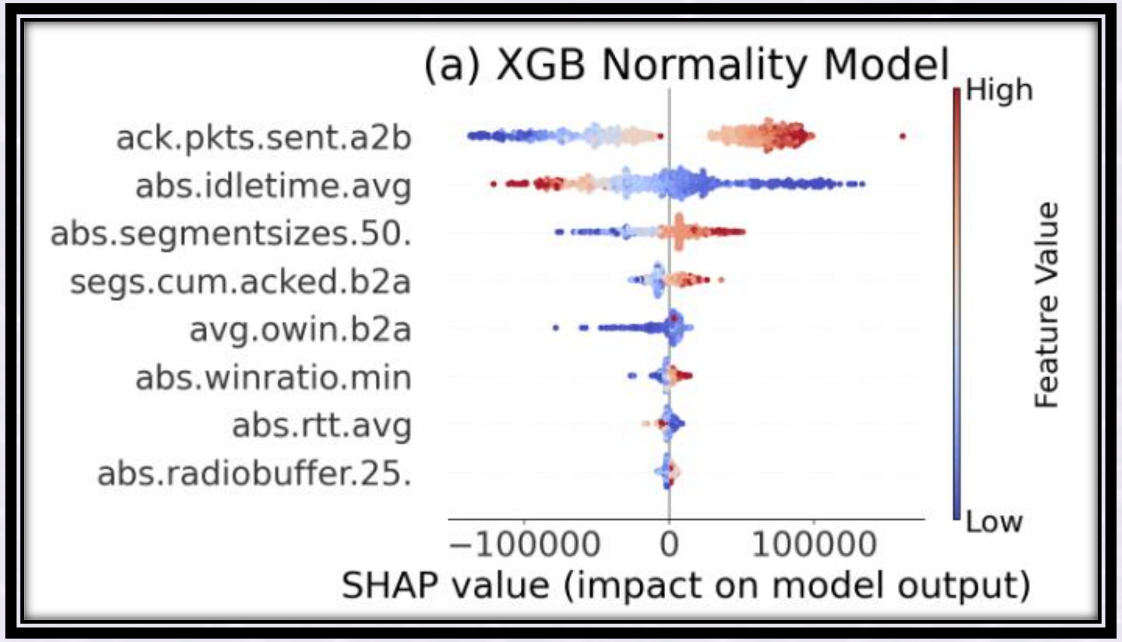
Histogram of the logarithmic version of the throughput for the six datasets

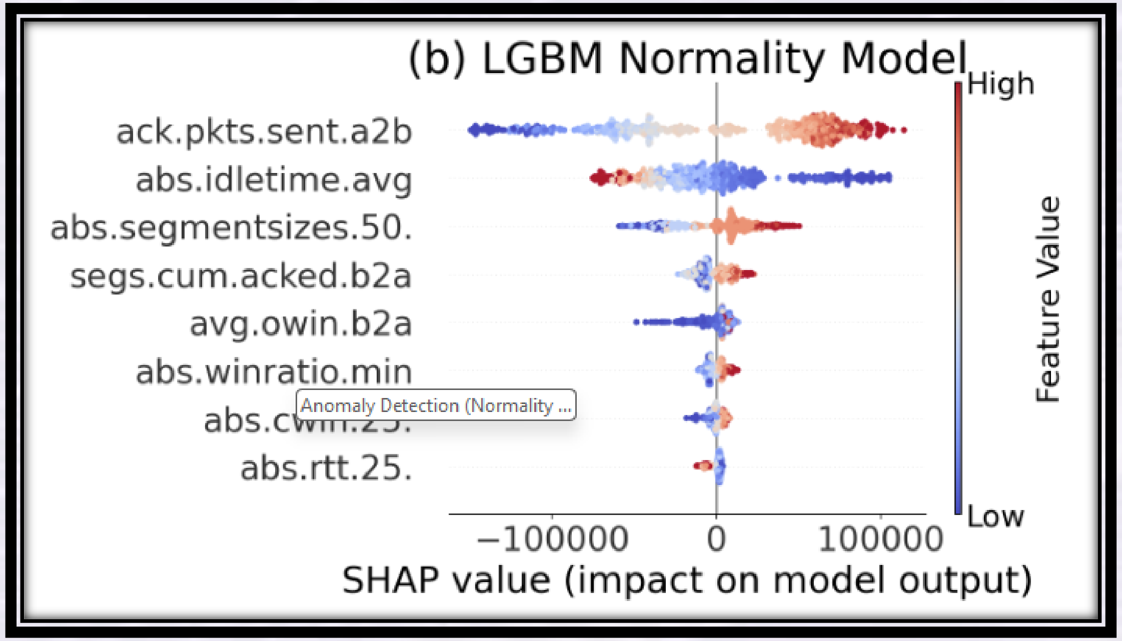


- Our approach **identifies instances with significant low throughput** (below acceptable levels)
- These events can be related to long webpages loading times or interruptions in video streaming

Anomaly Detection (Normality Model Interpretability)

SHAP values using (a) XGBoost and (b) LGBM





- We can observe features influence the normality model
  - High ack.pkts.sent.a2b and low abs.idletime.avg → High throughput
  - Low ack.pkts.sent.a2b and high abs.idletime.avg (delays) → Low throughput

Outlier and Anomaly Detection

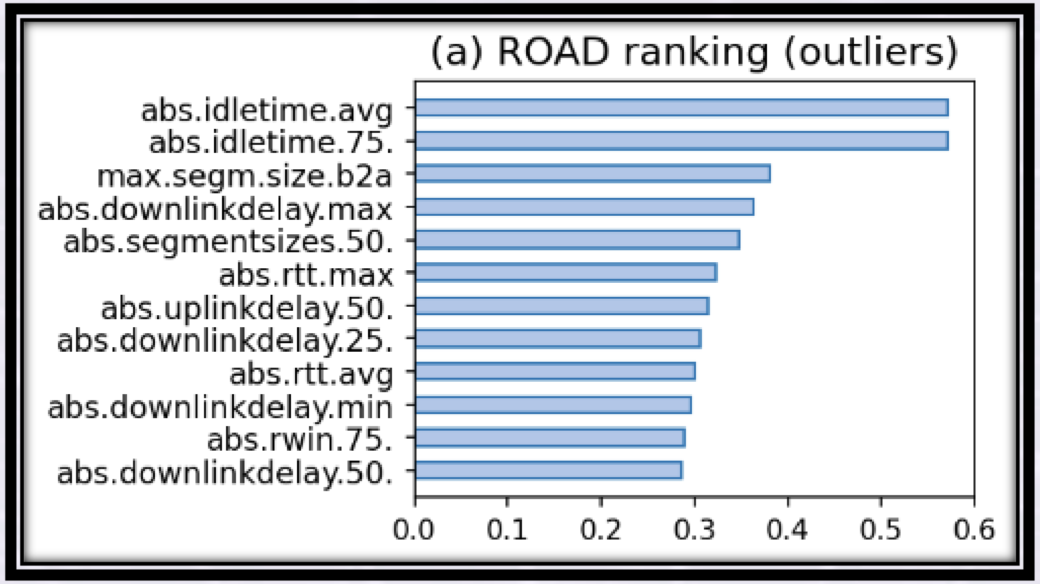
- We compare the detection performance of the proposed methodology to the explainable machine learning anomaly detection (XMLAD) approach (detects outliers only) [2].
- Both approaches exhibits competitive performance in detecting outliers.
- Our approach outperforms in detecting normal (non-outlier or true negative) samples.

TABLE III
ACCURACY METRICS OBTAINED BY DETECTION METHODS OF ANOMALOUS SCENARIOS USING THE REAL DATASETS

Dataset	Metric	Method		
		XMLAD	Our approach	
		Outliers	Outliers	Anomalies
Subset 1A	Sensitivity	0.92	1.00	1.00
	Recall	0.36	0.99	1.00
	F1-score	0.10	0.94	0.95
Subset 2A	Sensitivity	0.69	1.00	0.57
	Recall	0.40	1.00	0.98
	F1-score	0.17	1.00	0.64
Subset 1B	Sensitivity	0.94	0.73	0.48
	Recall	0.61	1.00	1.00
	F1-score	0.54	0.81	0.65

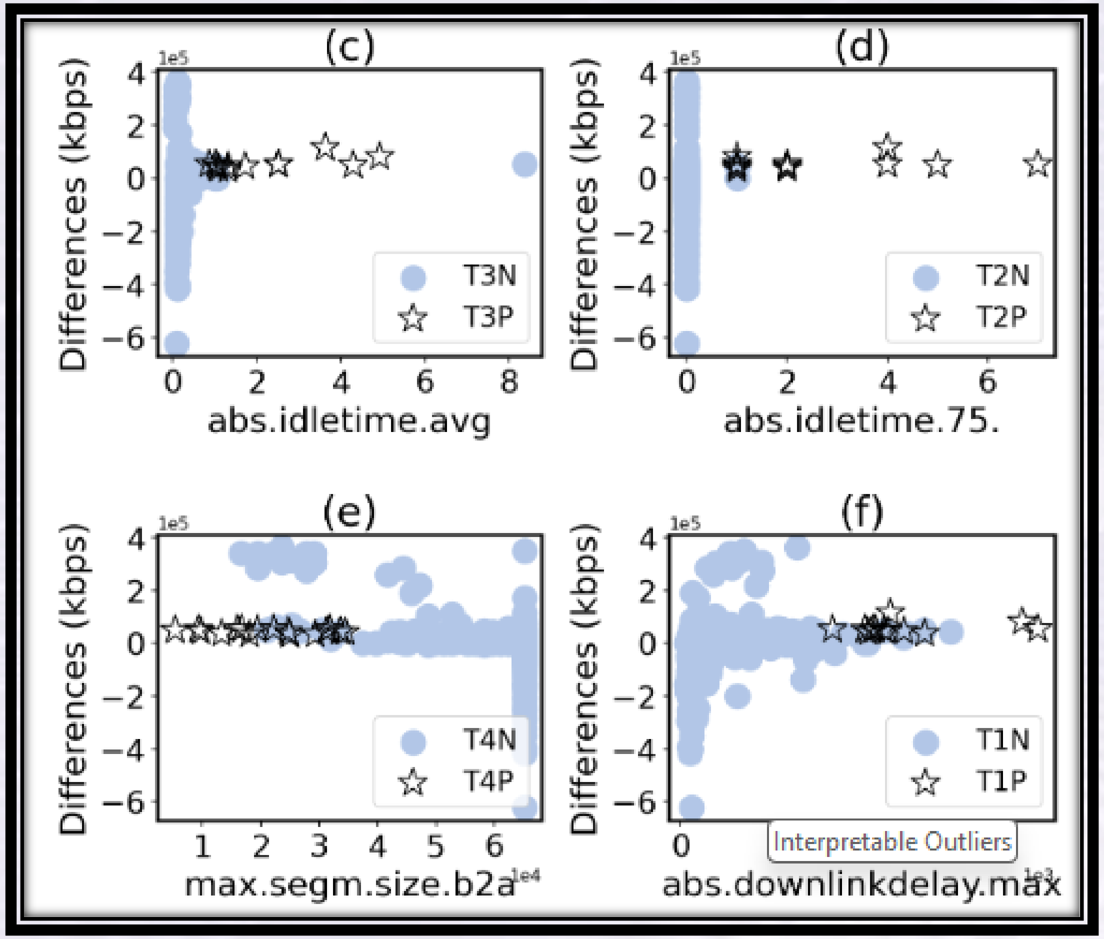
institute imdea networks

Interpretable Outliers



ROAD Ranking for Outliers

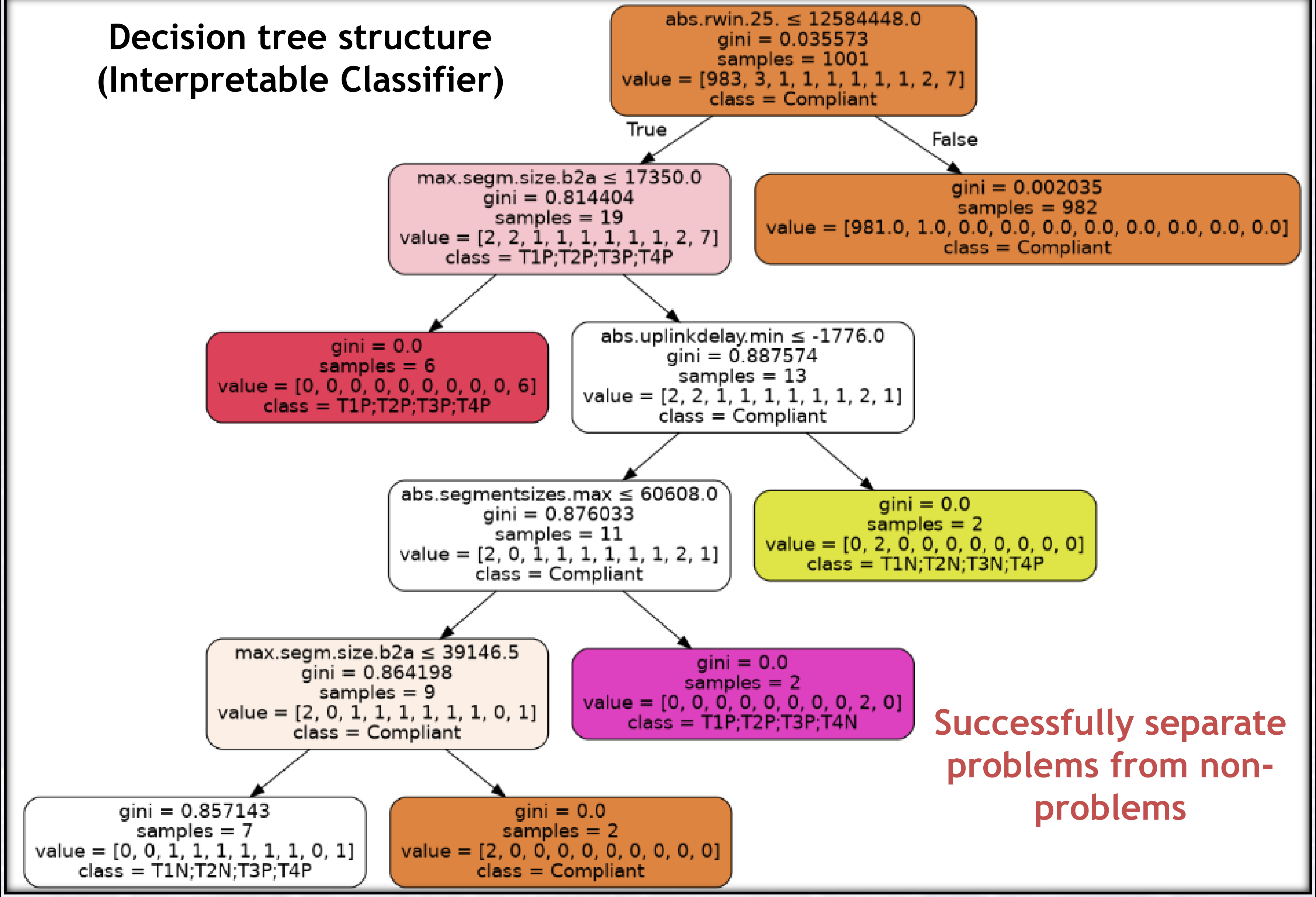
- Outliers are closely related to
  - Large delays (idle time and downlink delays)
  - Low segment sizes



Differences versus feature values

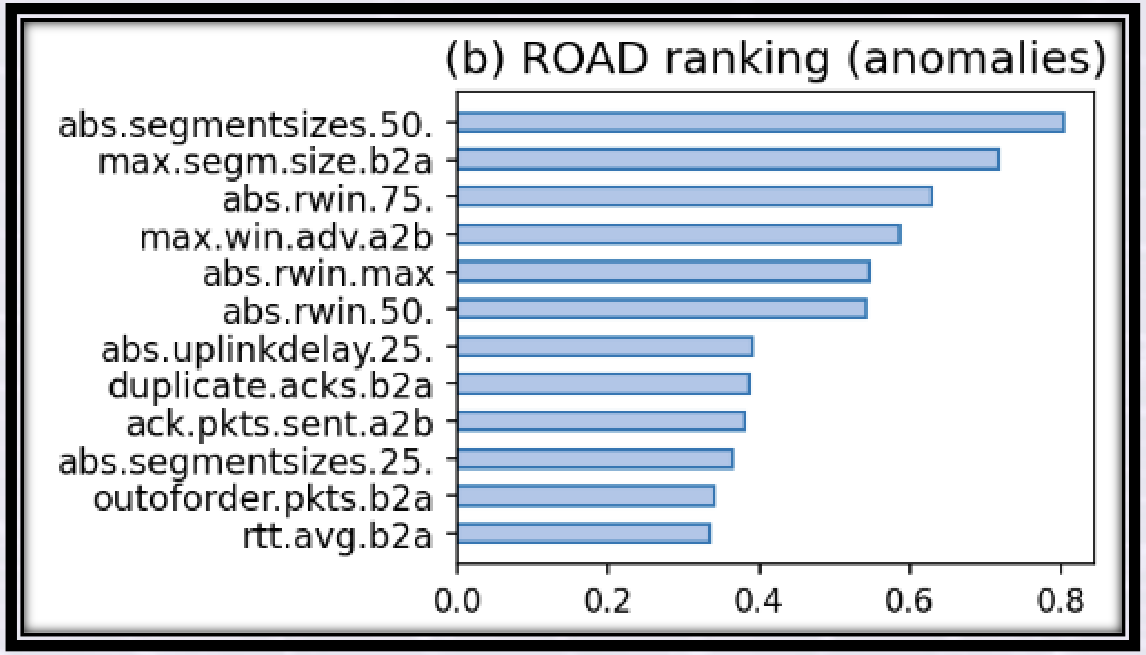
Interpretable Outliers

Decision tree structure (Interpretable Classifier)



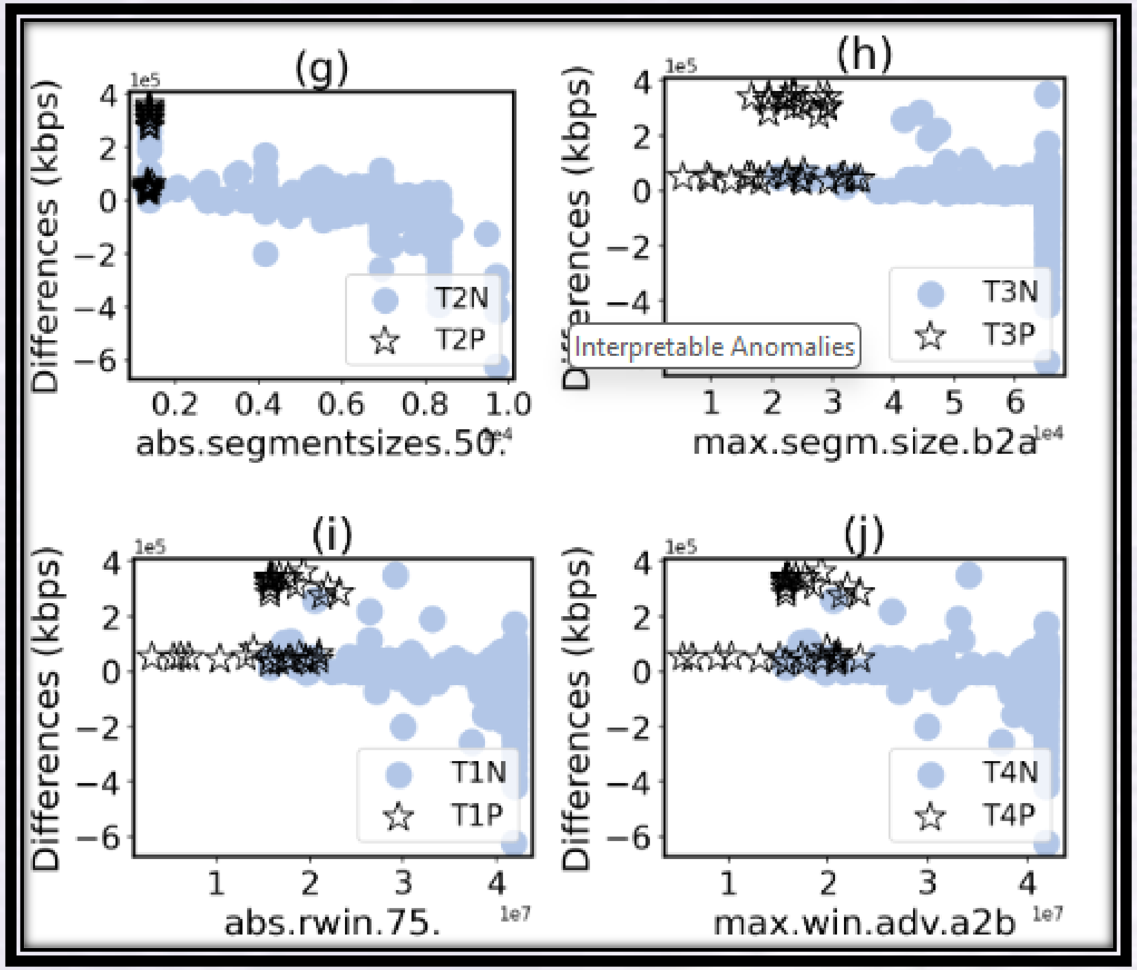
Successfully separate problems from non-problems

Interpretable Anomalies



ROAD Ranking for Anomalies

- Anomalies are associated with **low segment sizes**.
- Small segment sizes suggest the network operates with **excessive retransmissions or flow control restrictions**.

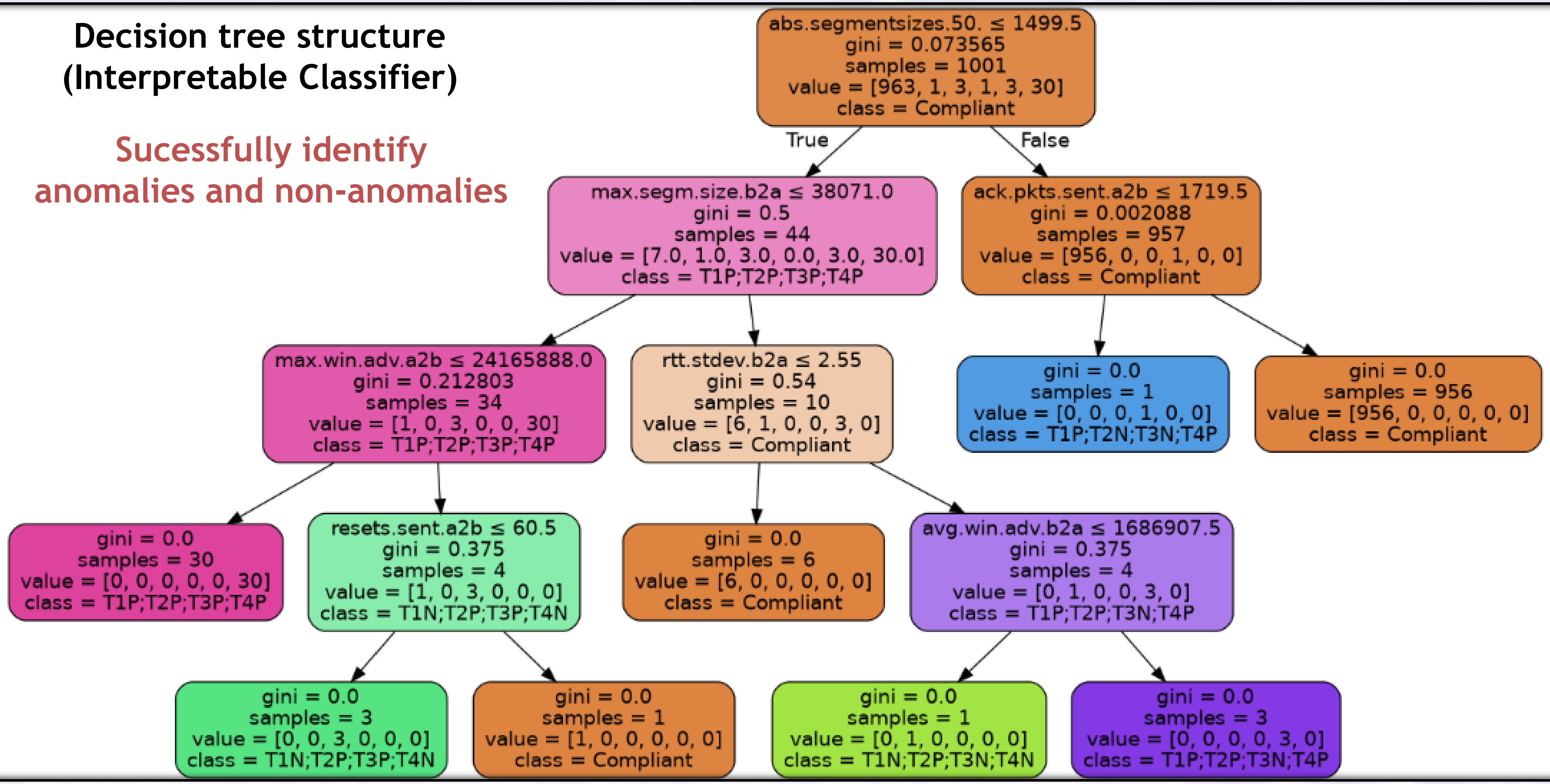


Differences versus feature values

Interpretable Anomalies

Decision tree structure (Interpretable Classifier)

Sucessfully identify anomalies and non-anomalies



Conclusions and Future Work

- We delineate and distinguish between outliers and anomalies, while presenting a comprehensible methodology for their detection and characterization.
- In addition to detection, this study introduced a novel interpretability module that identifies patterns closely associated with atypical samples.
- We also introduced the ROAD index to quantify the relationship between atypical samples (outliers or anomalies) and binary patterns extracted from input features.
- ROAD index was used to select features closely related to atypical behavior.
- For future work, we aim to perform controlled measurements to reliably label the ground truth samples and evaluate the methodology using deep learning architectures.

Acknowledgments

- This paper has been funded by project PID2022-140560OB-I00 (DRONAC) funded by MICIU/AEI/10.13039/501100011033 and ERDF, EU.

Questions?

References

- [1] Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology, 49(4), 764-766.
- [2] Ramírez, J. M., Díez, F., Rojo, P., Mancuso, V., & Fernández-Anta, A. (2023). Explainable machine learning for performance anomaly detection and classification in mobile networks. Computer Communications, 200, 113-131.

imdea networks institute