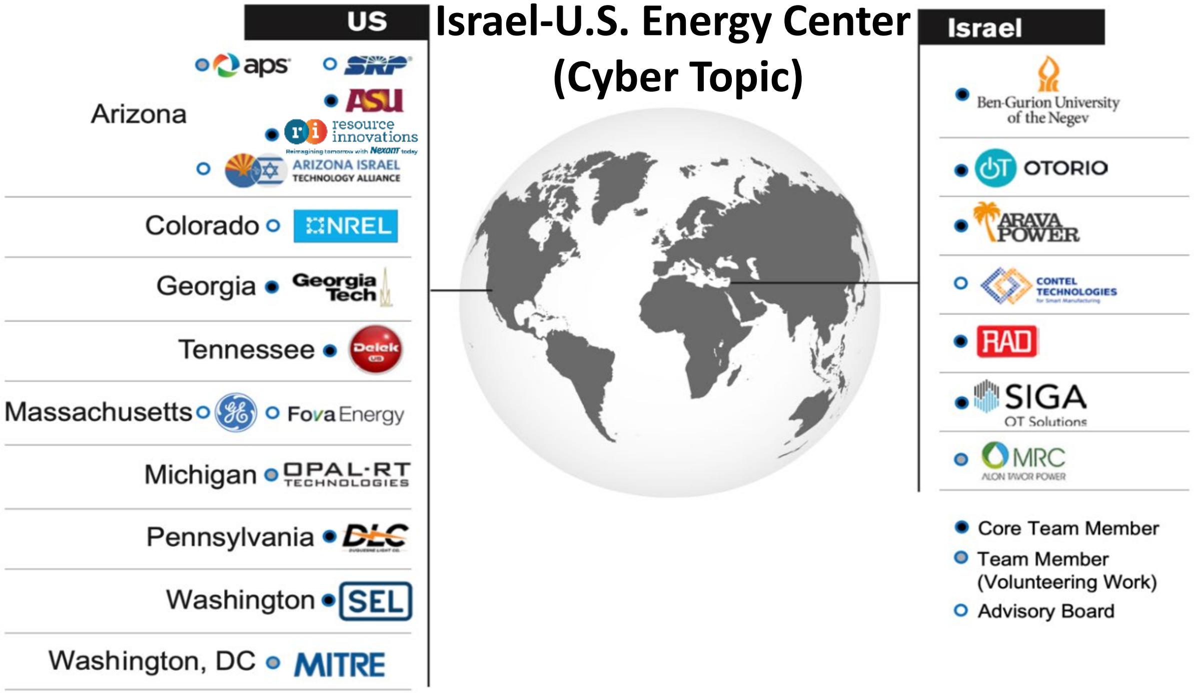
Task 12 - Explainable cyber A.I. analytics



Israel-U.S. Energy Center (Cyber Topic)

US

Arizona ◉ aps ○ SRP ● ASU ● resource innovations ○ ARIZONA ISRAEL TECHNOLOGY ALLIANCE

Colorado ○ NREL

Georgia ● Georgia Tech

Tennessee ● Delek US

Massachusetts ○ GE ○ Fova Energy

Michigan ◉ OPAL-RT TECHNOLOGIES

Pennsylvania ● DLC

Washington ● SEL

Washington, DC ◉ MITRE

Israel

● Ben-Gurion University of the Negev

● OTORIO

● ARAVA POWER

○ CONTEL TECHNOLOGIES

● RAD

● SIGA OT Solutions

◉ MRC ALON TAVOR POWER

● Core Team Member

◉ Team Member (Volunteering Work)

○ Advisory Board

Prepared for

Eitan Yudilevich, Eynan Lichterman,
and **Tal Fischelovitch**

BIRD

Aug 24, 2022

Goal

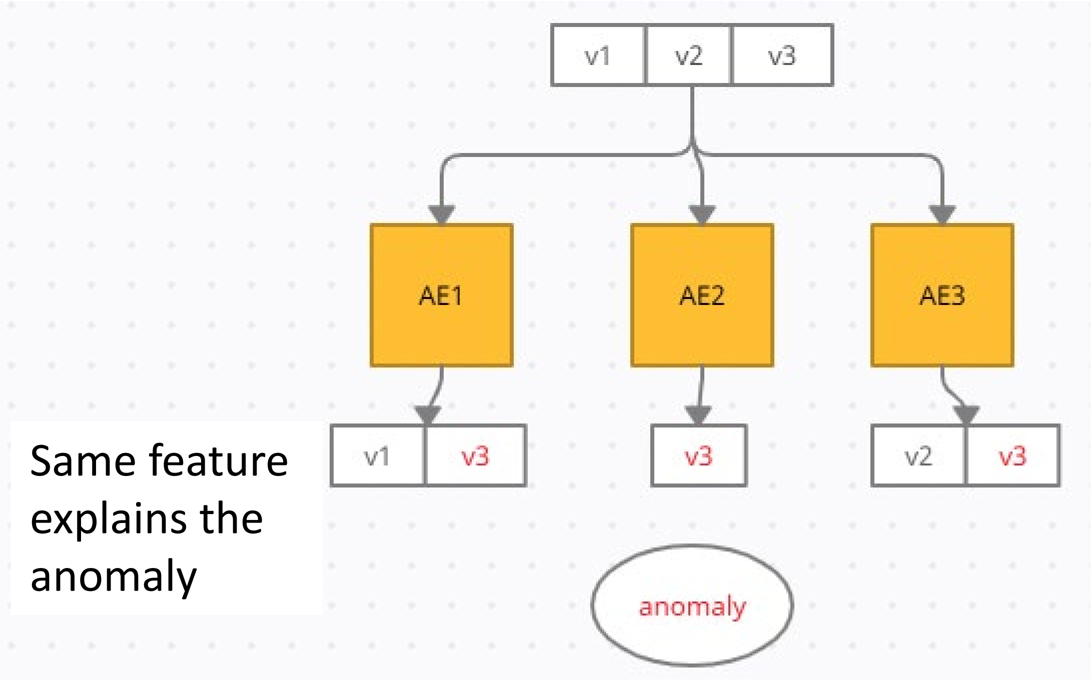
Goal:

- Minimize false positives anomalies
- Developing an evaluation method for the explanations

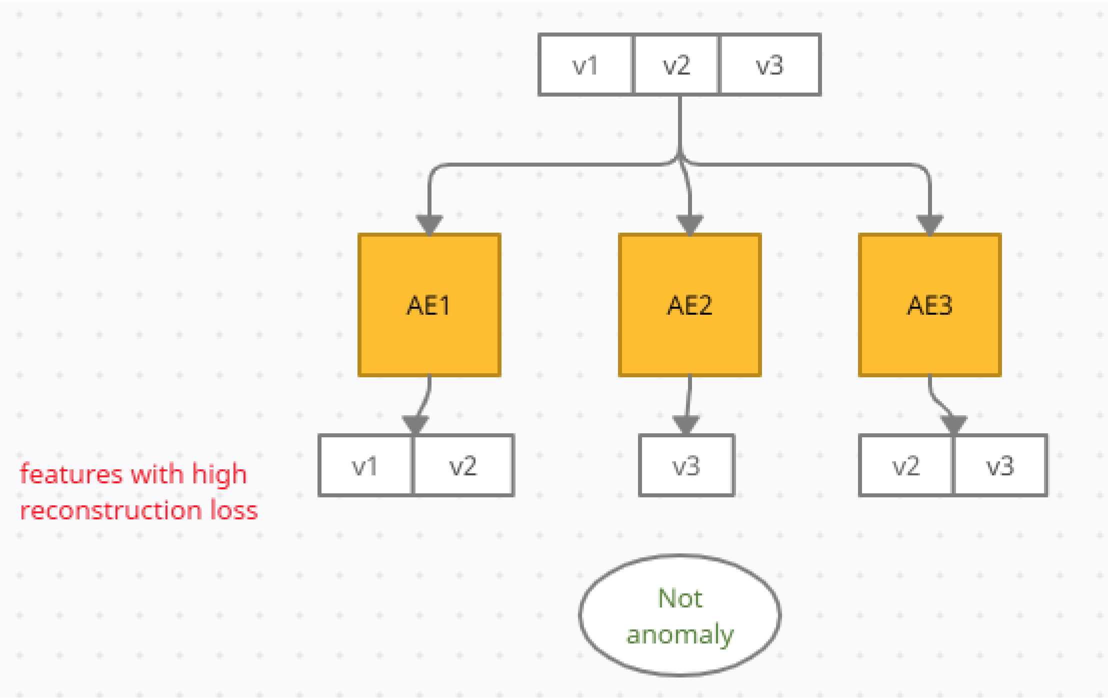
Method: Ensemble of anomaly detectors models

- Train multiple models independently
- Decide using ensemble of anomaly detectors’ **explanations** which records are anomalous

Unsupervised\Supervised anomaly detectors

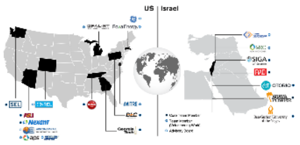


Same feature explains the anomaly



features with high reconstruction loss

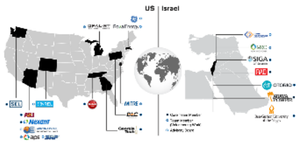
Experiments



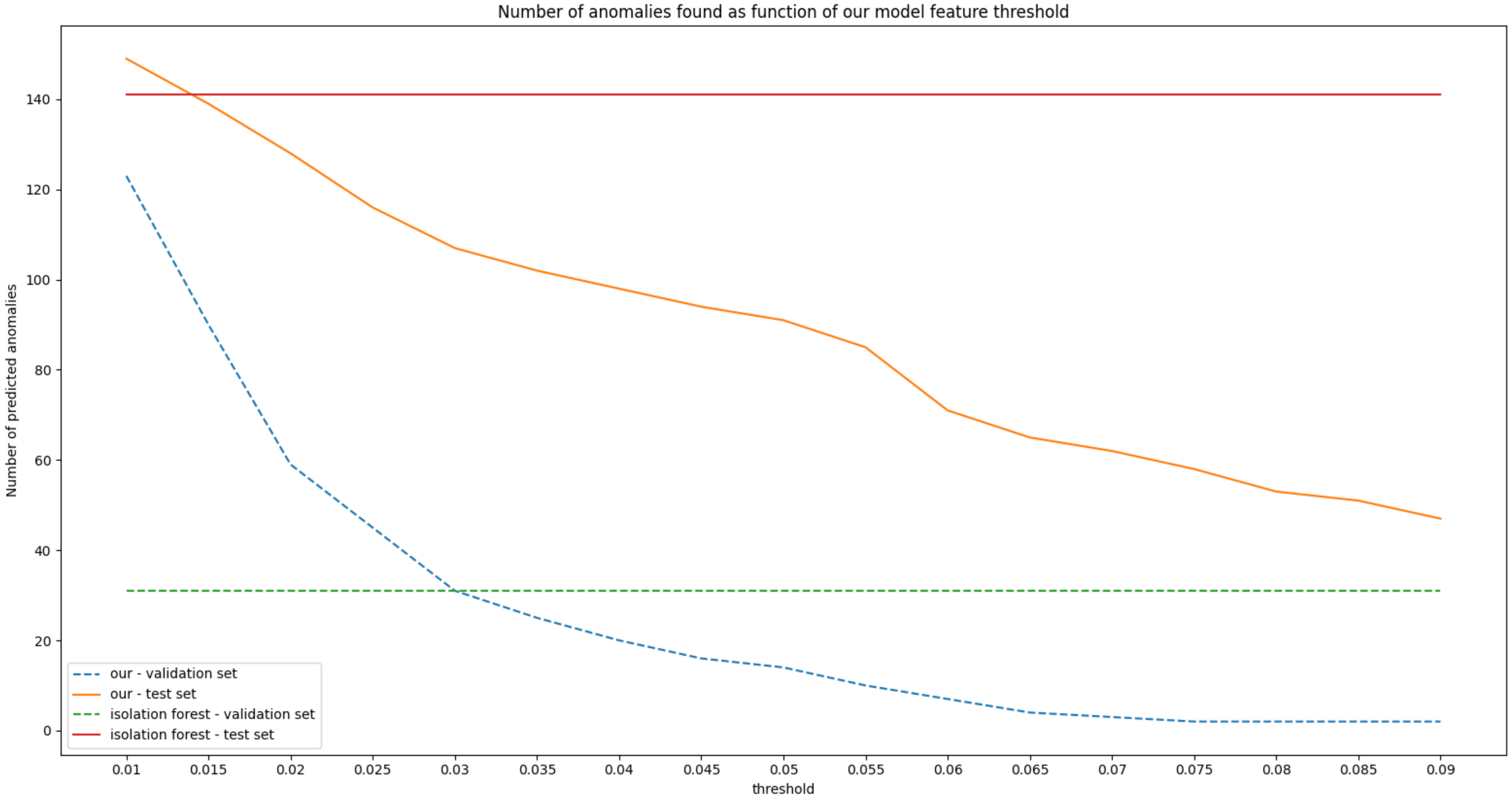
- Compared the results to isolation forest

(as in the paper: Kitsune: An Ensemble of Autoencoders for Online Network Intrusion Detection)

- The isolation forest algorithm is used to detect anomalies. Instead of modeling the normal points, it detects anomalies based on isolation (the distance a data point is from the rest of the data)

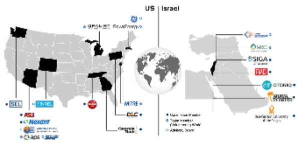


Cardio dataset (validation set size: 311, test set size: 176)

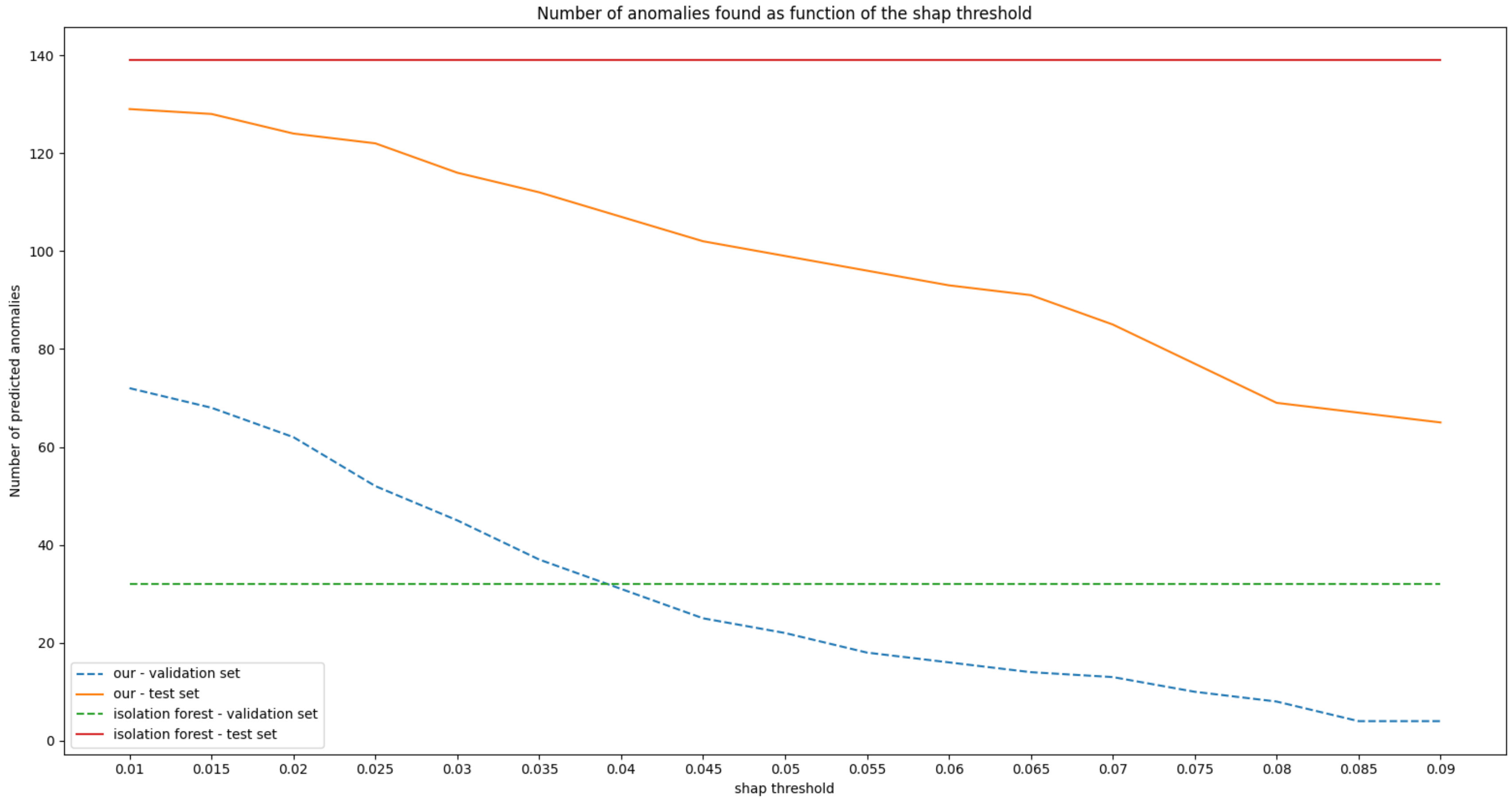


Number of anomalies found as function of our model feature threshold

There is a trade-off between the number of anomalies we find in the test set and the number of anomalies that are misclassified in the validation set. In the case of high thresholds, we have very few false positives.

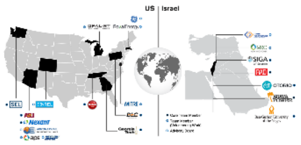


Cardio dataset (validation set size: 311, test set size: 176)



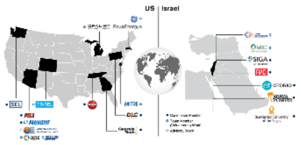
Again, there is a trade-off between the number of anomalies we find in the test set and the number of anomalies that are misclassified in the validation set. When we increase the shap value threshold to 0.04 or above then we have less false positive rate than the isolation forest.

Tool development



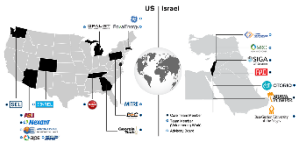
- The developed algorithm can be used by domain experts
- Input: A list of transactions
- Output: A list of anomalies, ordered by severity

The use of the data provided by the industrial partners



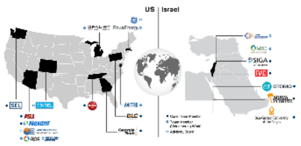
- We need the data to train a model that detects anomalies
- We prefer tagged data for validation (the method is unsupervised)

Data



- **System**: IT and OT from the same network
- **Output**: Anomalous entities by date&time
- **Network size**: Medium
- **Data:** both raw data and processed tabular data (sessions)
- **Max dataset size**: 1Tb
- **Time resolution**: Maximal
  - If its not possible to provide within the size limitation maximal resolution for all entities, please provide maximal resolution for part of the entities, and a sample for the rest.
- **Time frame:** At least one month (same period for IT and OT data)

Benefits for industrial partners



- Save anomalies’ analysts time by investigating less events
- Increase trust in the AI systems that reveals the anomalies

- **Next Step:** Meet stakeholders that may be interested