Signatures to help interpretability of anomalies

Emmanuel Gangler

LPC
Particules pLasmas Univers applicationS
Laboratoire de Physique de Clermont

cnrs

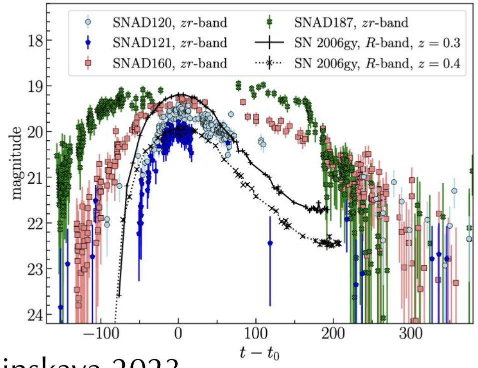
UNIVERSITÉ Clermont Auvergne

What is an anomaly ?

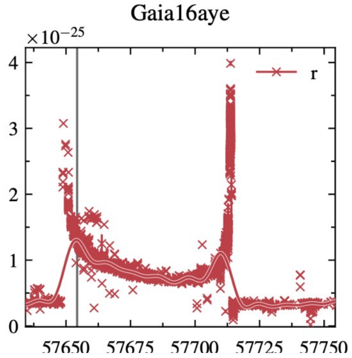
An anomaly is a pattern that does not conform to expected normal behavior, and is suspected to be generated by a different mechanism.

Some challenges :

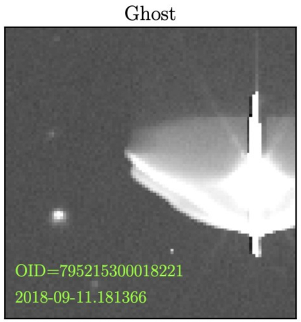
- Normal behavior may be not known
  - Usually in large data sets → Unsupervised learning, fast algorithms.
- Interpreting the underlying behavior/properties of anomalous data
  - Field expert knowledge needed



SLSN
Pruzhinskaya 2023



Binary microlens event
Pruzhinskaya 2019



Bogus finder
Malanchev 2021

What is an anomaly ?

An anomaly is a pattern that does not conform to expected normal behavior, and is suspected to be generated by a different mechanism.

Some challenges :

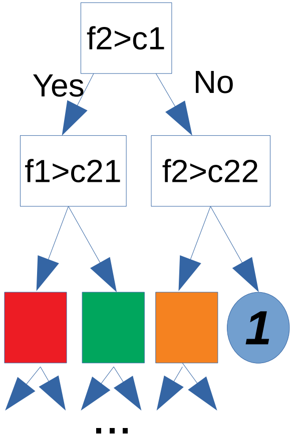
- Normal behavior may be not known
  - Usually in large data sets → Unsupervised learning, fast algorithms.
- Interpreting the underlying behavior/properties of anomalous data
  - Field expert knowledge needed

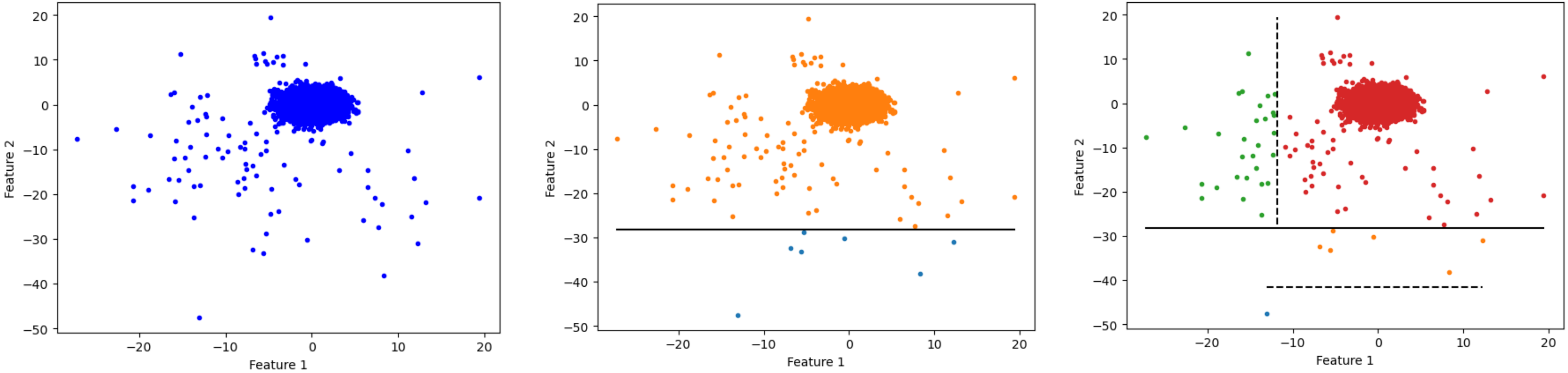
How to help the astronomer ?

- Find anomalies
- **Understand** Machine decision
- Find **more interesting** anomalies
- Discard uninteresting anomalies

Isolation Forest Density Estimation

- **Random Tree:**
  - Select a feature randomly
  - Select a random threshold within the range spanned by the feature for the (sub) sample
  - Repeat for each subsample
  - Stop when only 1 point in the sub-partition
- **Random Forest:**
  - Average depth on T trees (proxy for local density)

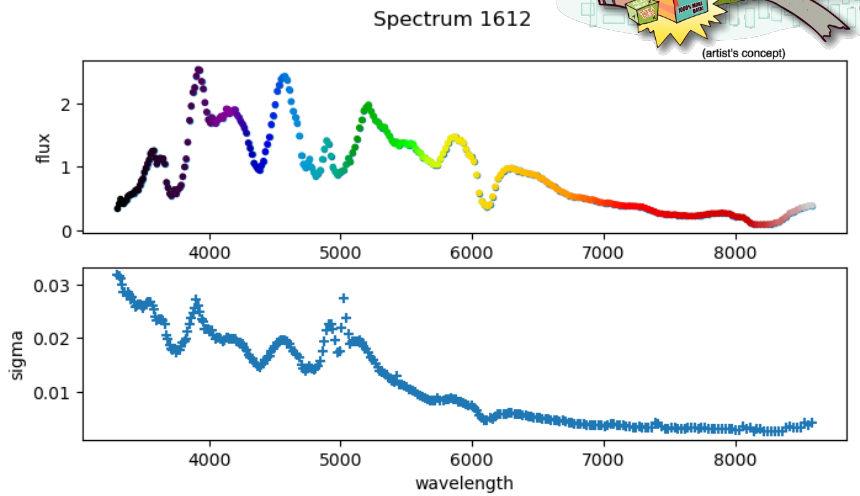




Depth 0 Depth 1 Depth 2

SNFactory dataset

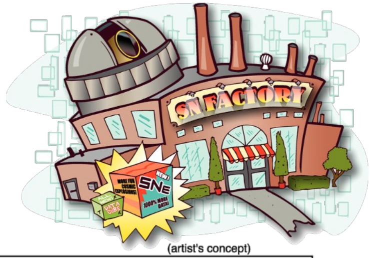
(artist's concept)



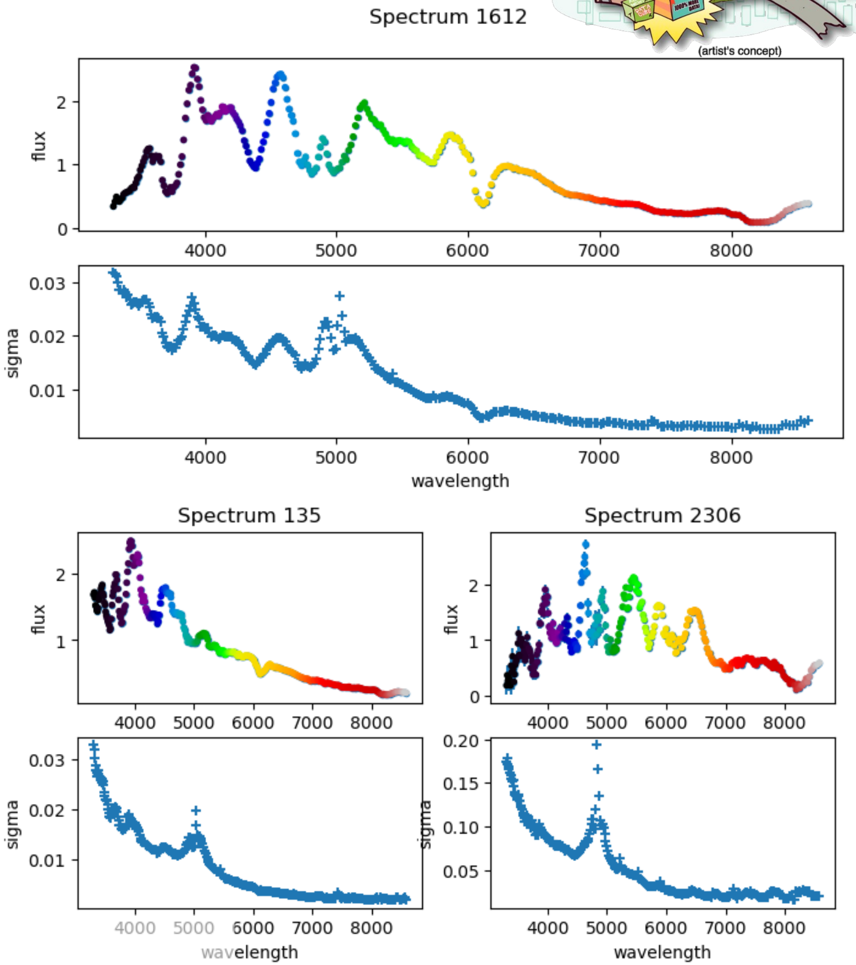
- Public astronomical dataset
  arXiv:2005.03462
  - **2323 spectra** of Type Ia supernovae
    - 288 spectral bins
    - With noise estimate

      → **576 features**

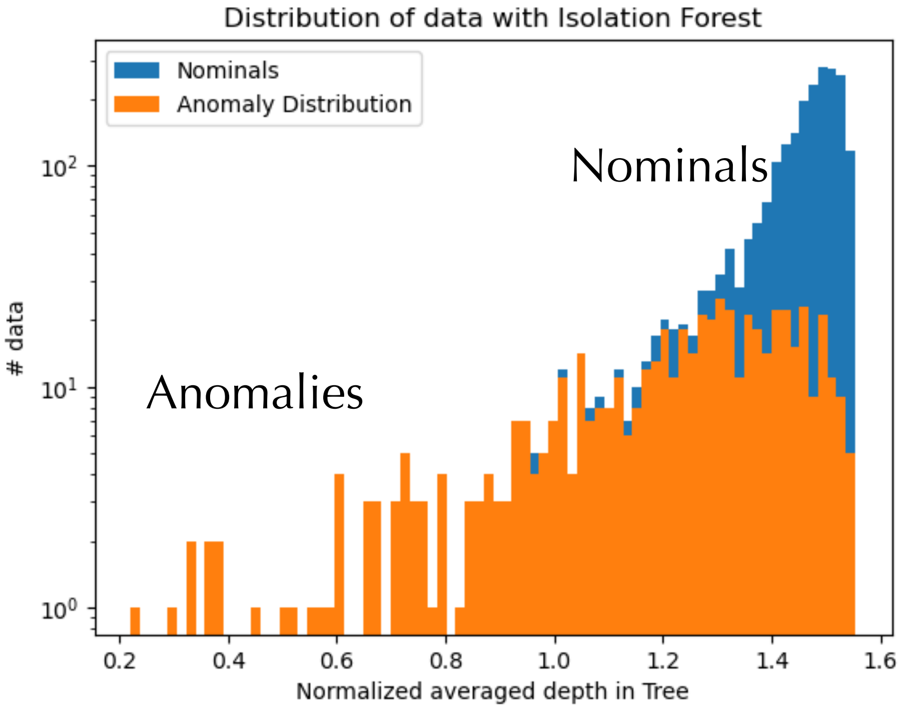
SNFactory dataset



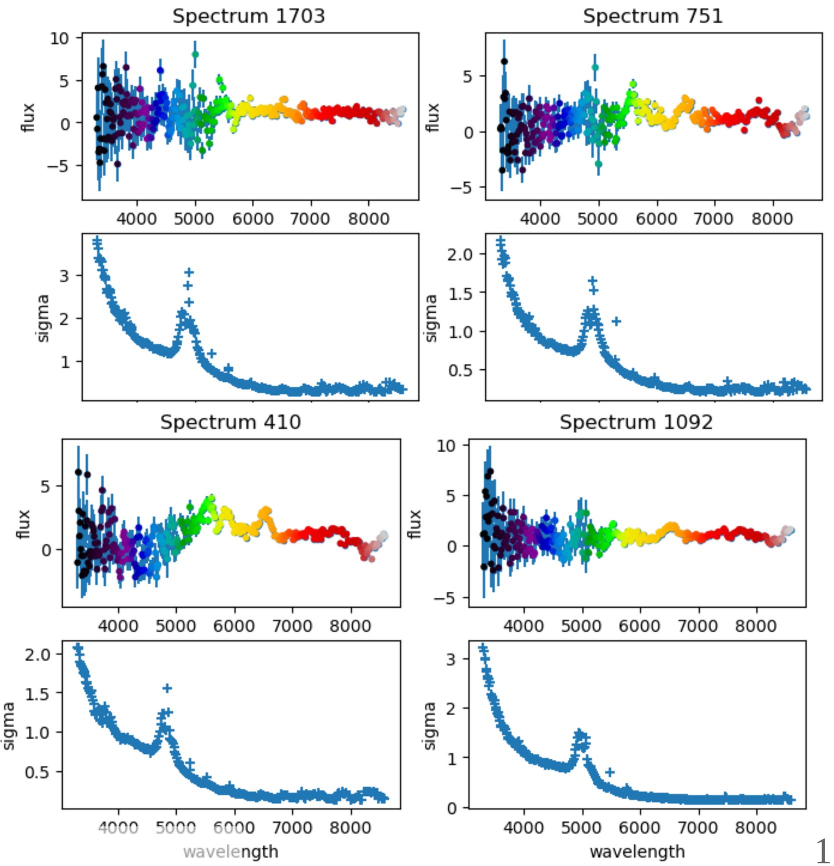
- Public astronomical dataset
- Interest of this dataset for anomalies
  - High **internal variability**
  - **Expert knowledge** for anomalies
  - Many **different class** of anomalies
  - **Noisy data** making the task difficult
  - Local **data artifacts**



Isolation Forest for SNFactory dataset



Top 4 outliers

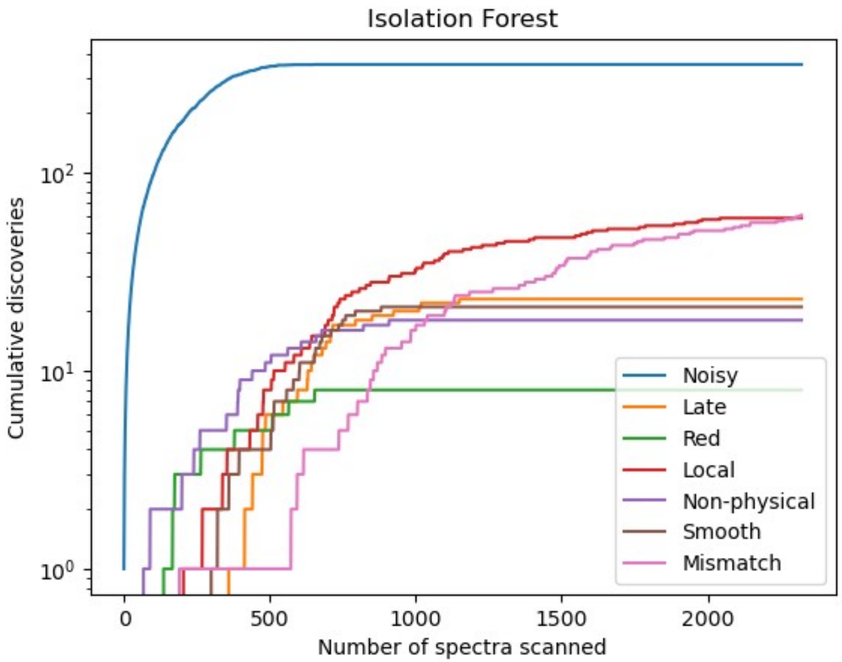


Outliers are anomalies of the **"noisy"** type

Discovery efficiency

- IF very efficient for **dominant anomaly**
  - → Noisy data dominate (AUC=0.985)

- Less efficient for **other classes**
  - **Rank** of last anomaly type dicovered: 360 (expected 326)
  - For non-noise anomalies : AUC=0.61

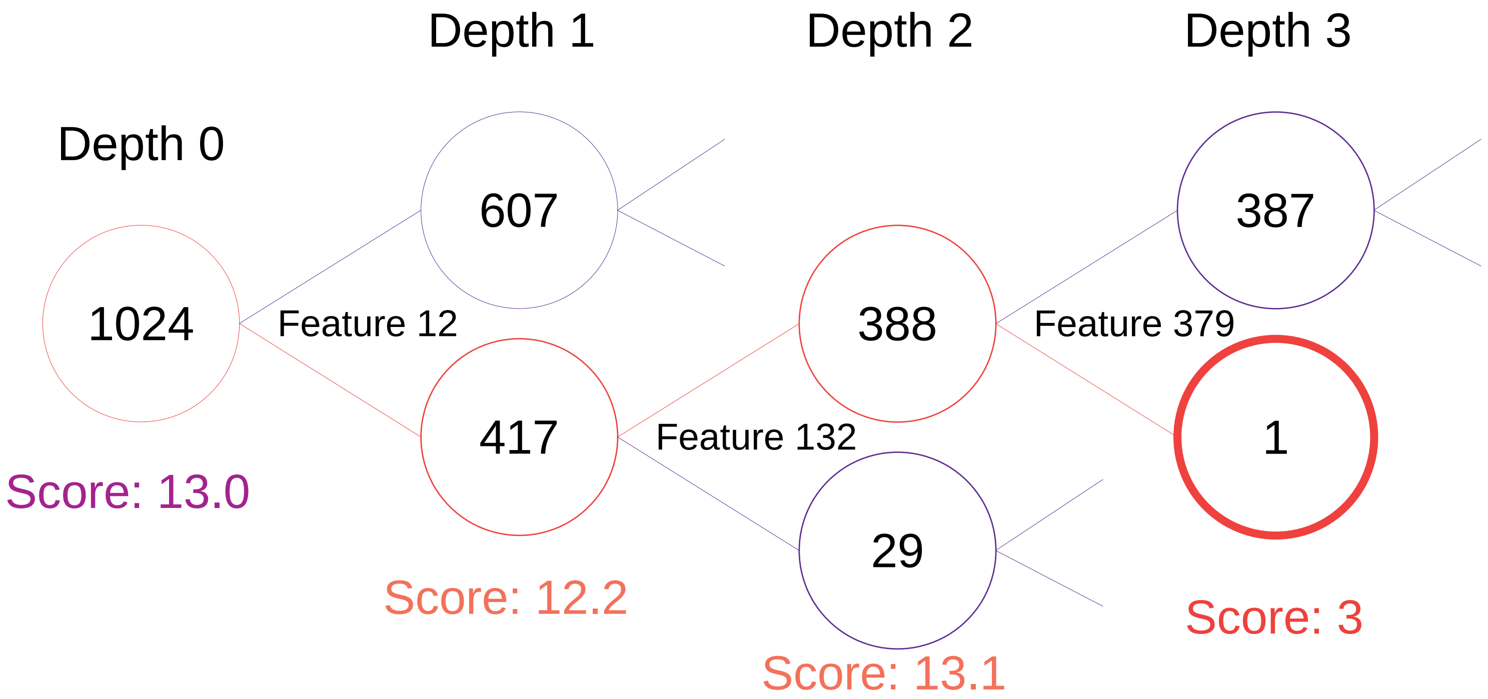


Some common Questions:

- **Why** are some data taged as anomalies ?
- Are there **different classes** of anomalies ?
- Can I improve discovery of **new anomalies** ?
- Can I find **more anomalies** of a given kind ?

Anomaly signature

- Anomaly score for 1 tree



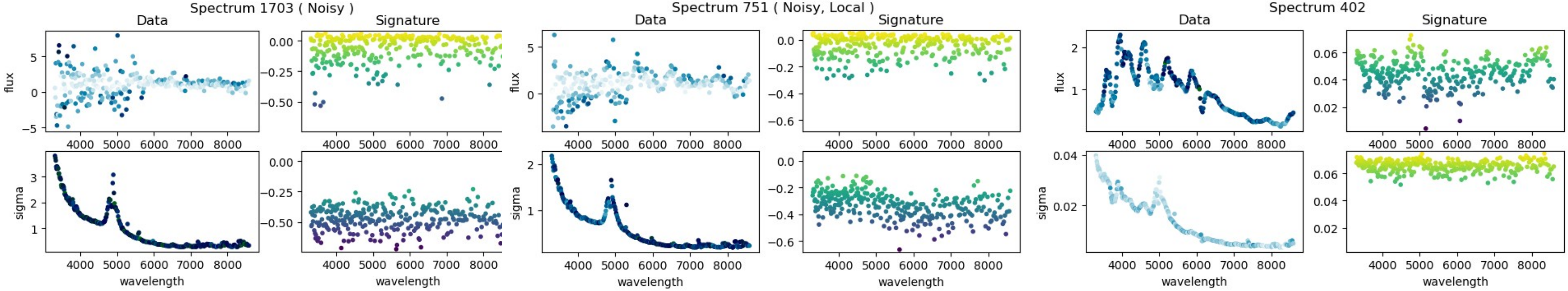
For this outlier :

Feature 12 = - 0.8 Feature 132 = + 0.9 **Feature 379 = - 10.1**

Signature & interpretability

Top 1 Anomaly | Top 2 Anomaly | Top 1 Nominal

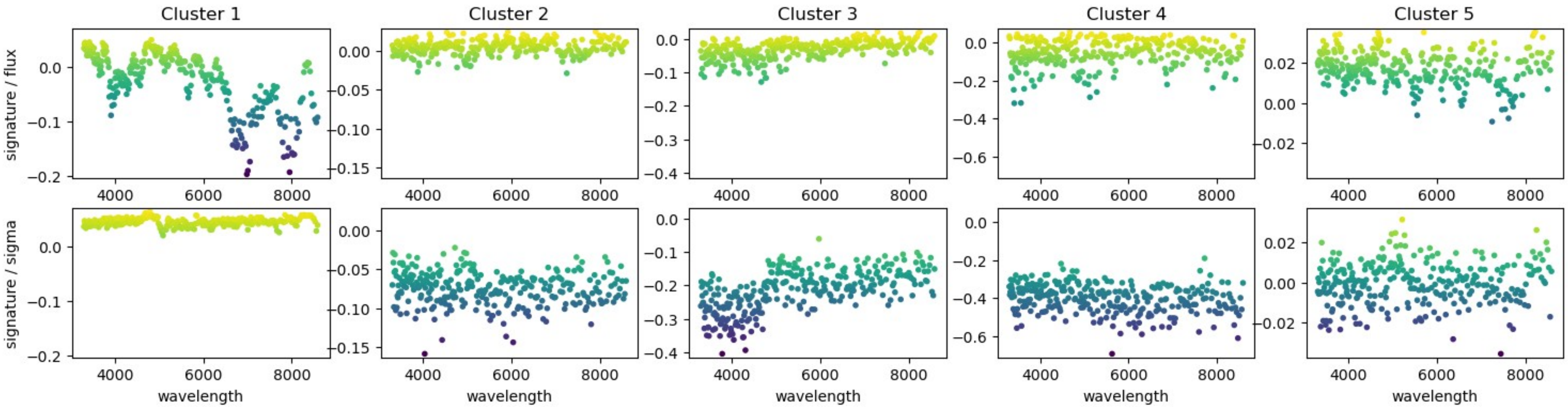
Data Signature | Data Signature | Data Signature



- Signature highlights where the data is anomalous
  - Negative score = anomalous
  - Interpretation : decision based on sigma
  - Positive score = nominal

Signatures & Clustering

- K-means on signatures for **top 10 % anomalous data** (232 spectra)
  - Very unbalanced : 90 % of those are tagged “Noisy”
  - Contains 7 % of nominals



- All anomalies belong to Cluster 1 !
- Only 39 elements : easier to analyse
- Still 2 classes of anomalies not found… + Choice of K is empirical

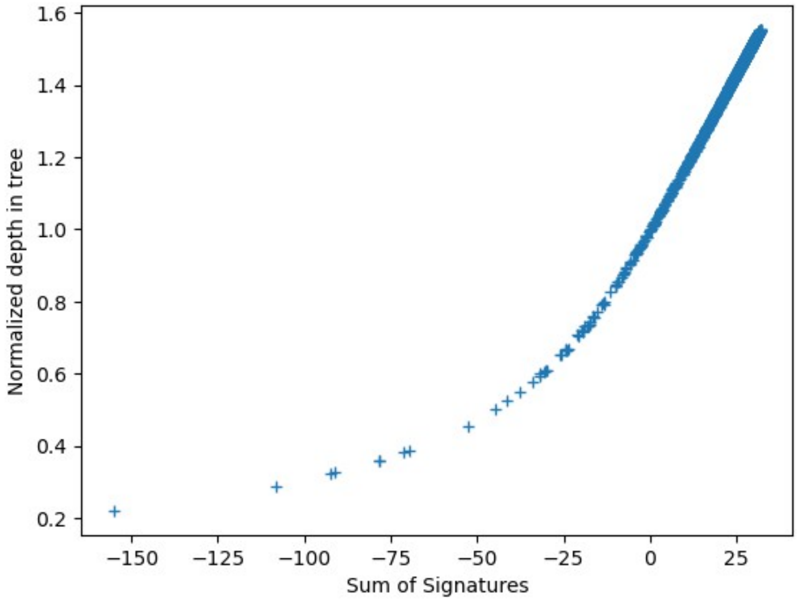
Recursive approach to novelty discovery with signatures

- Signature allow to derive a weighted anomaly score

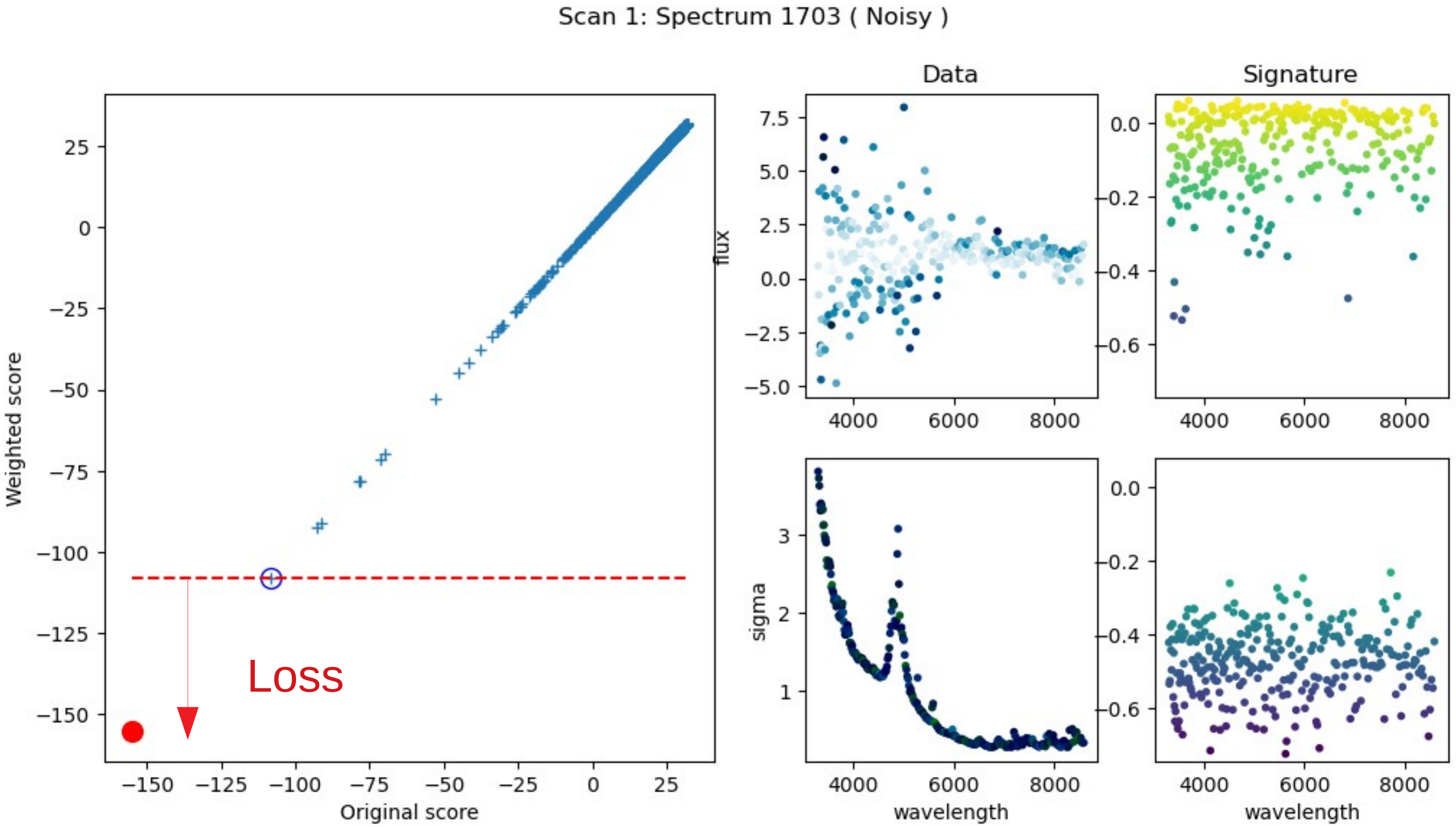
$$S^j = \sum_i \alpha_i S^j_{f_i}$$

  - Initialization of weights to 1

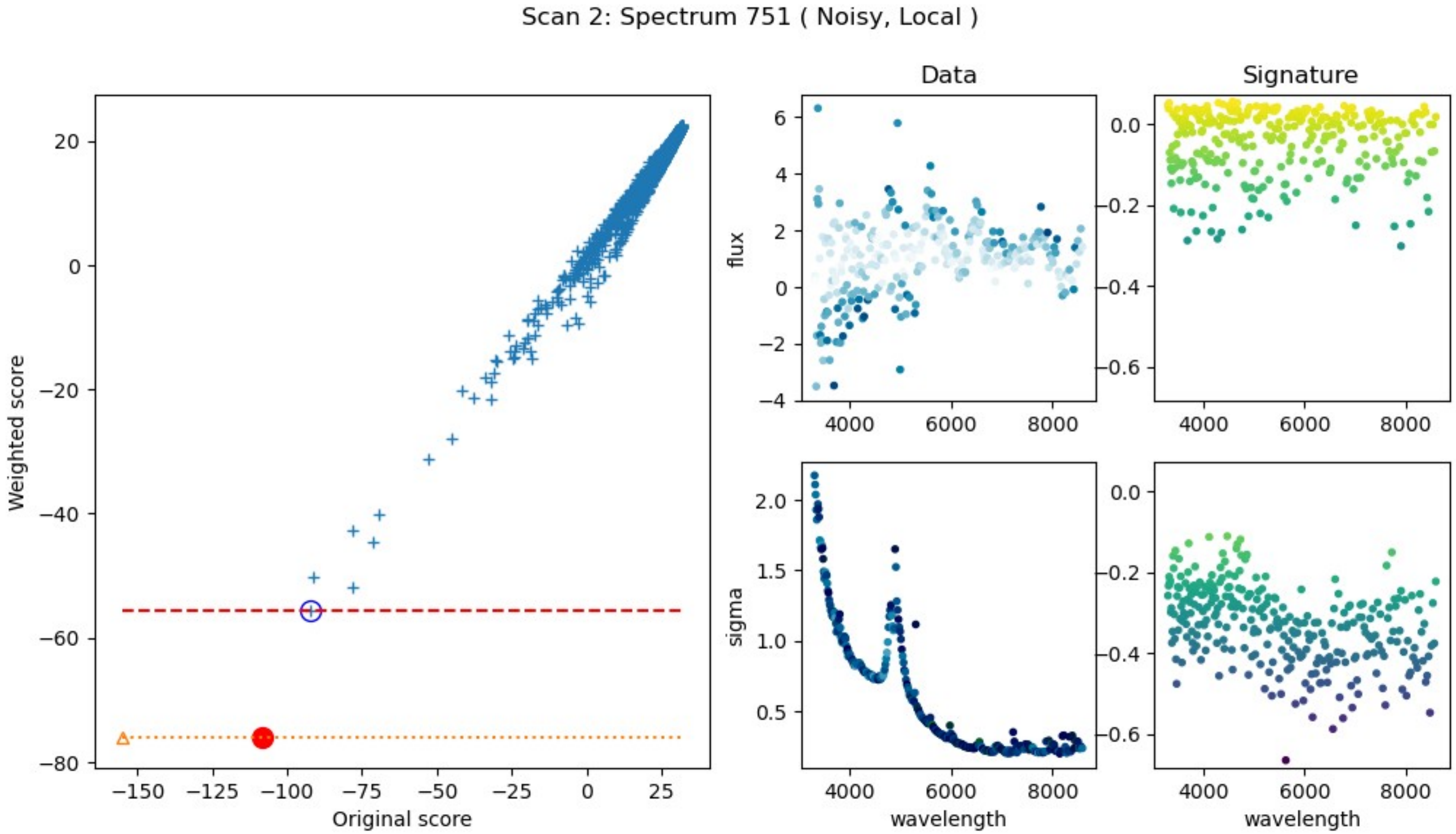
- For each data examined by the expert :
  - Tag as either wanted or unwanted
  - Update the weights / Hinge loss function
  - Propose to the expert the next mostly anomalous



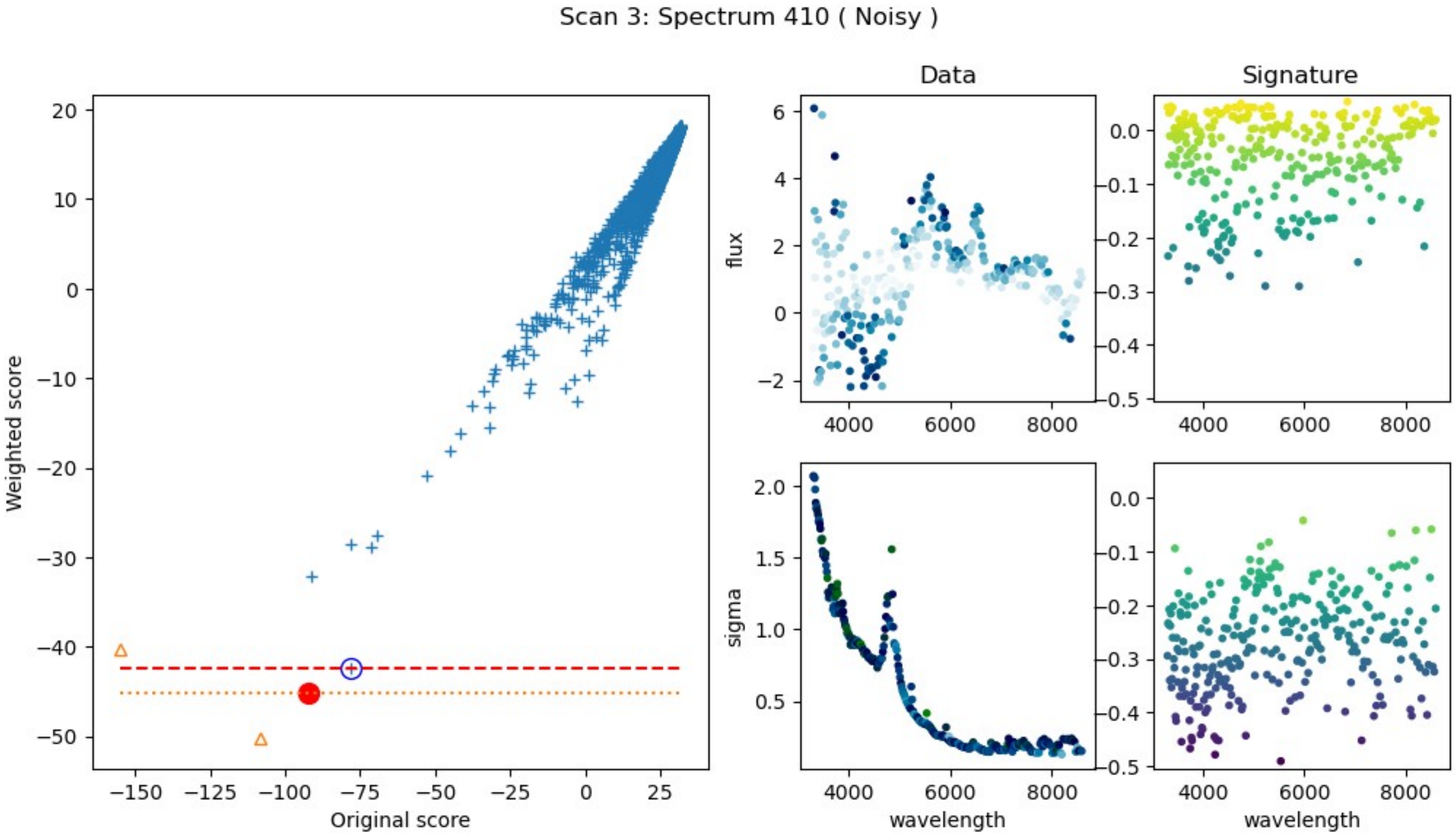
Recursive approach to novelty discovery with signatures



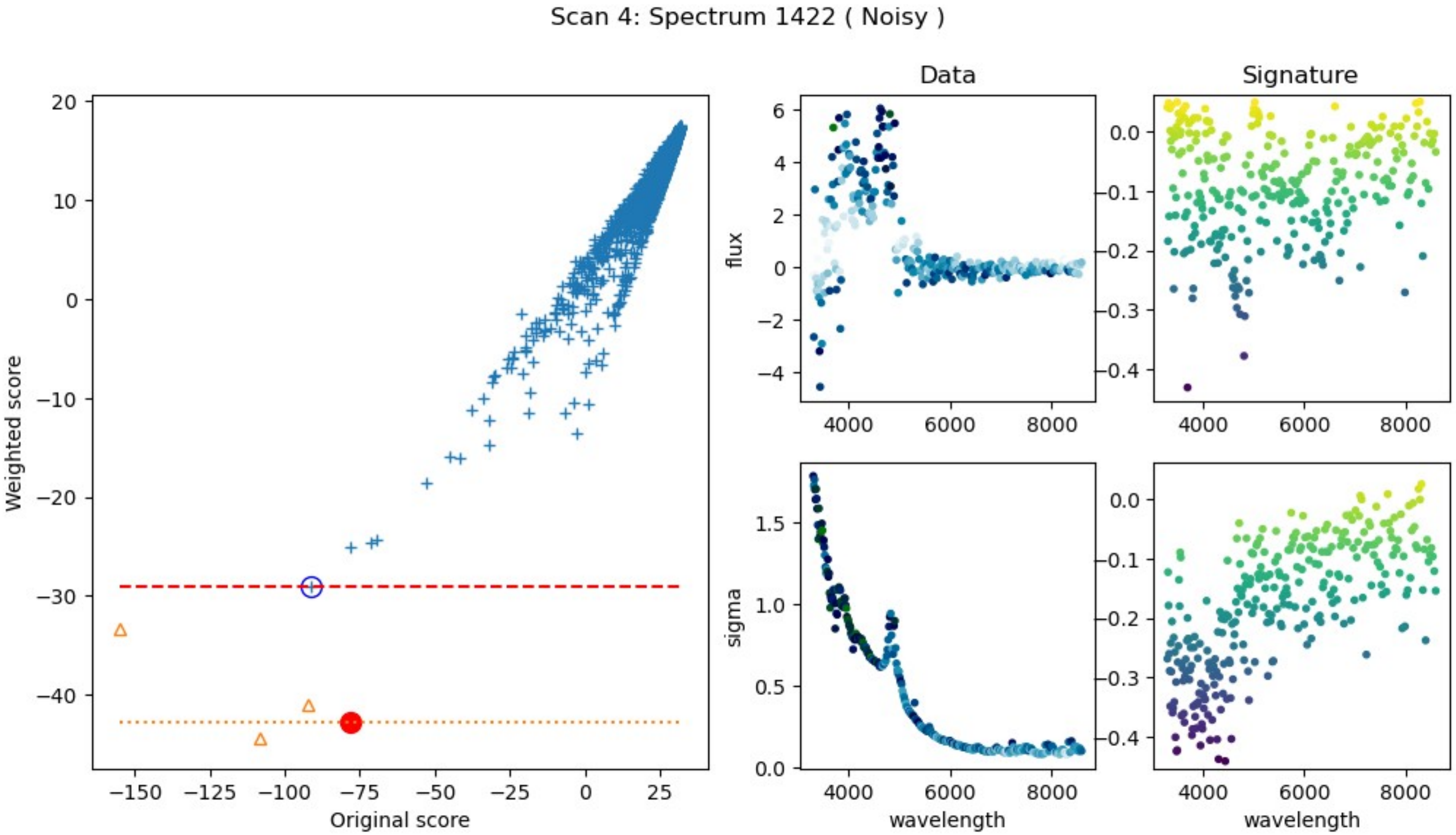
Recursive approach to novelty discovery with signatures



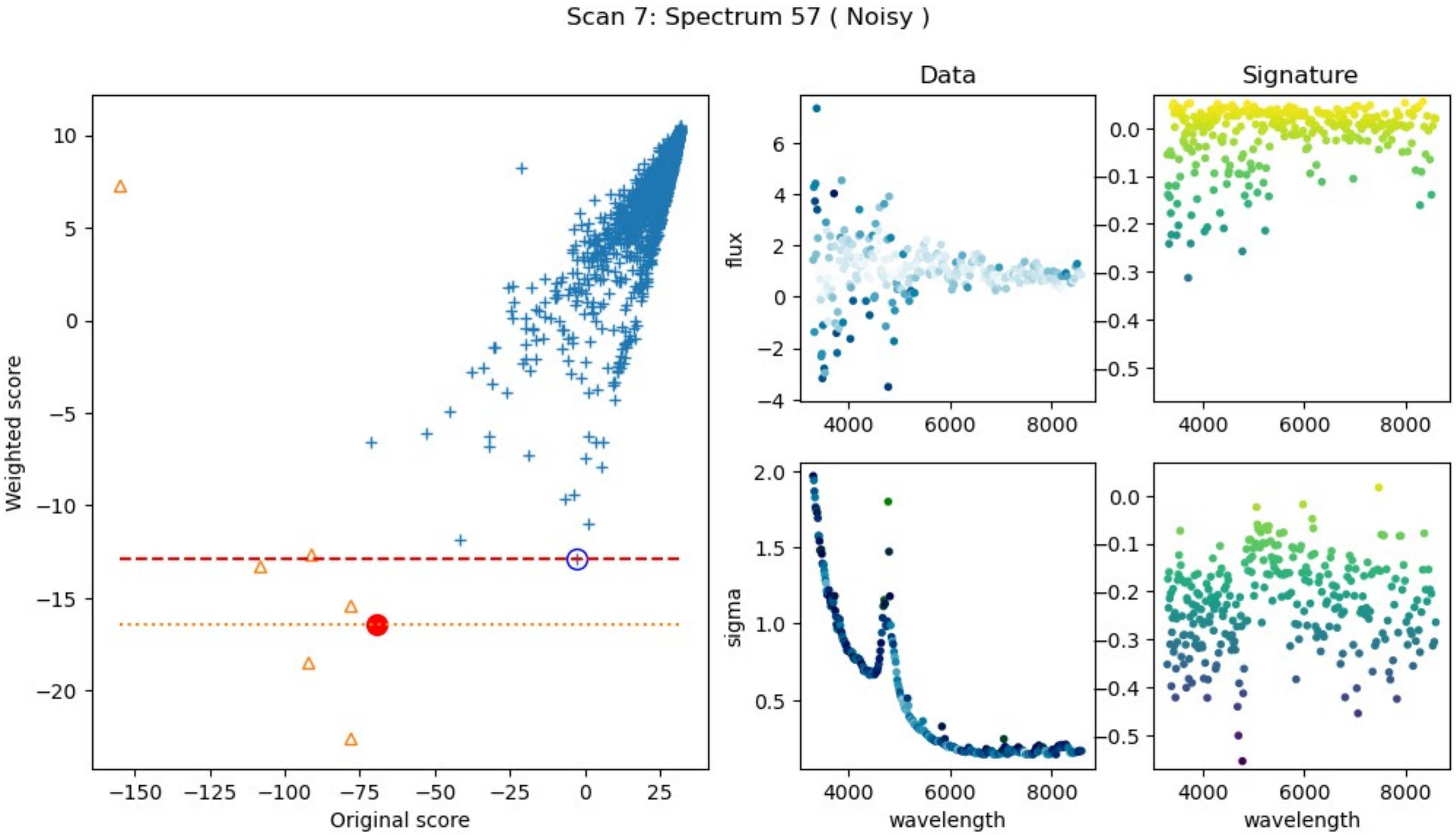
Recursive approach to novelty discovery with signatures



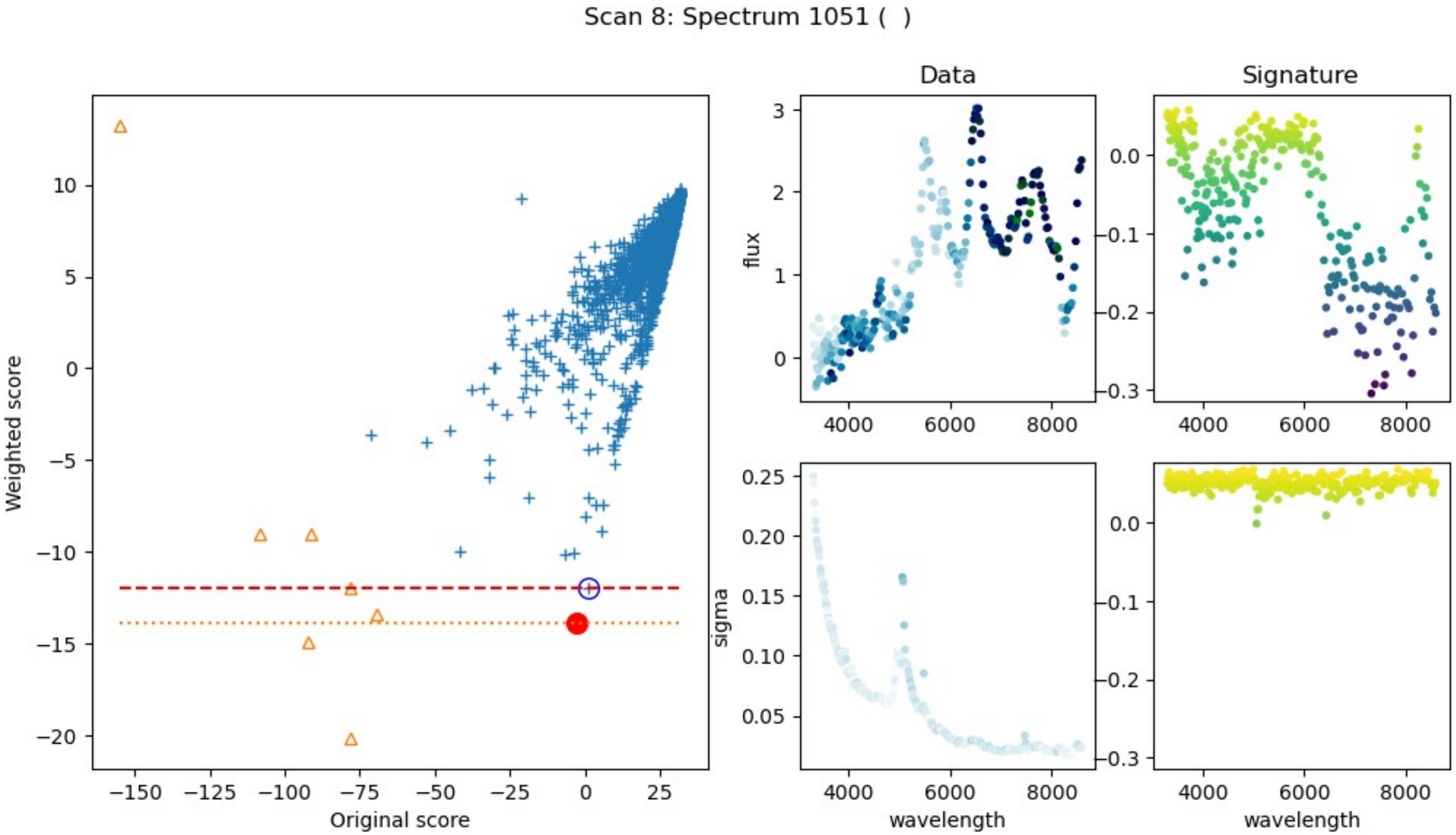
Recursive approach to novelty discovery with signatures



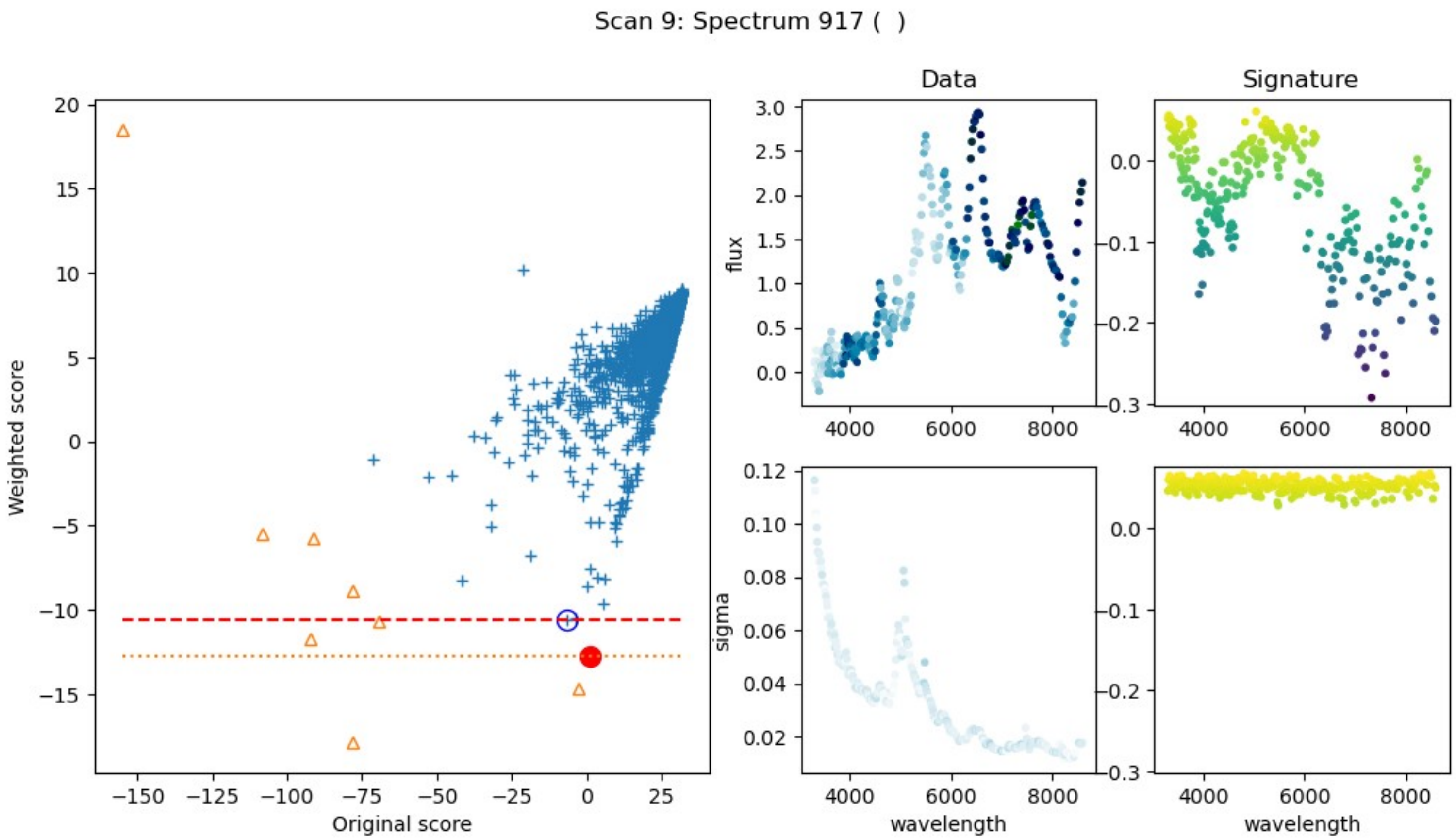
Recursive approach to novelty discovery with signatures



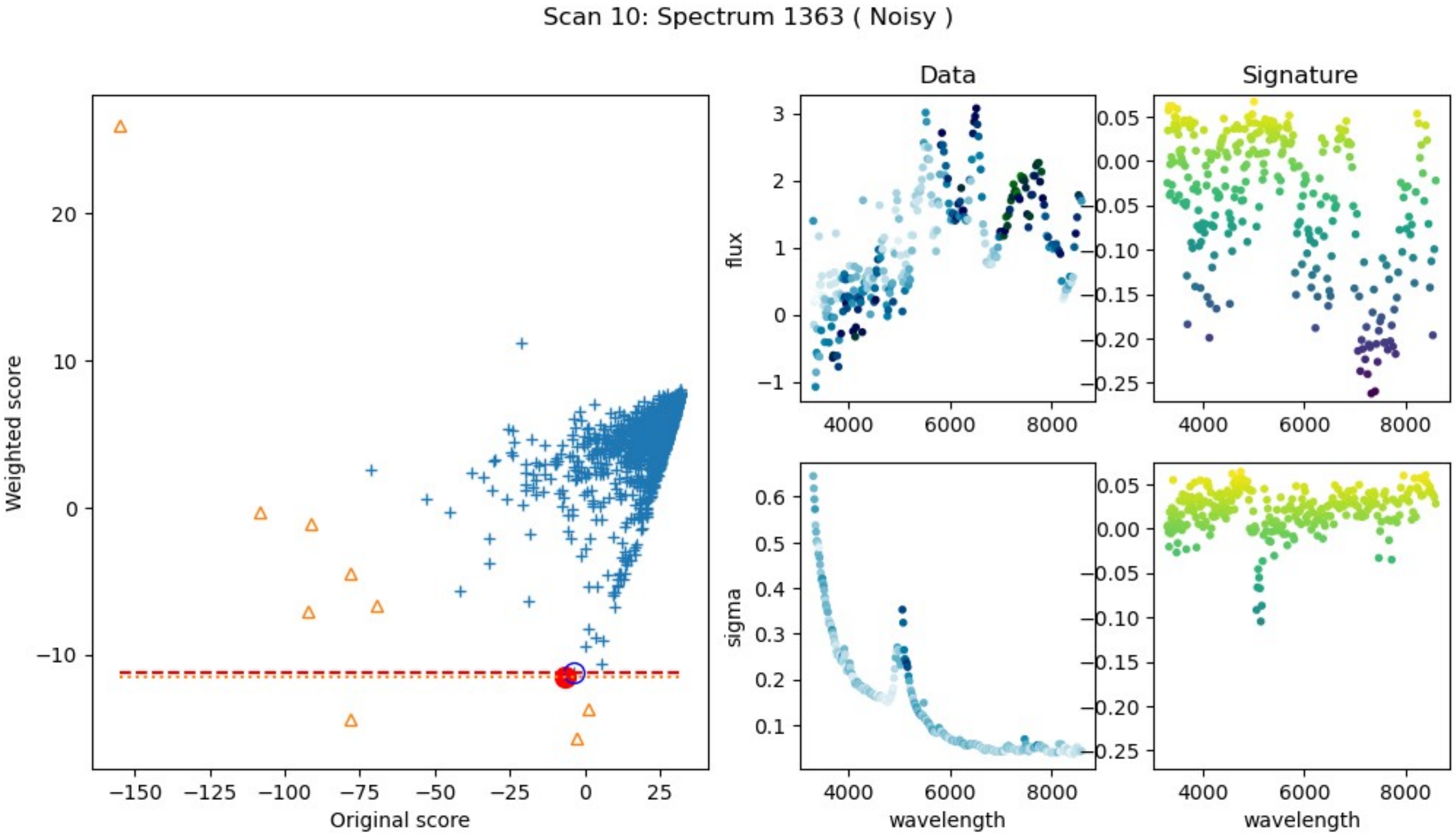
Recursive approach to novelty discovery with signatures



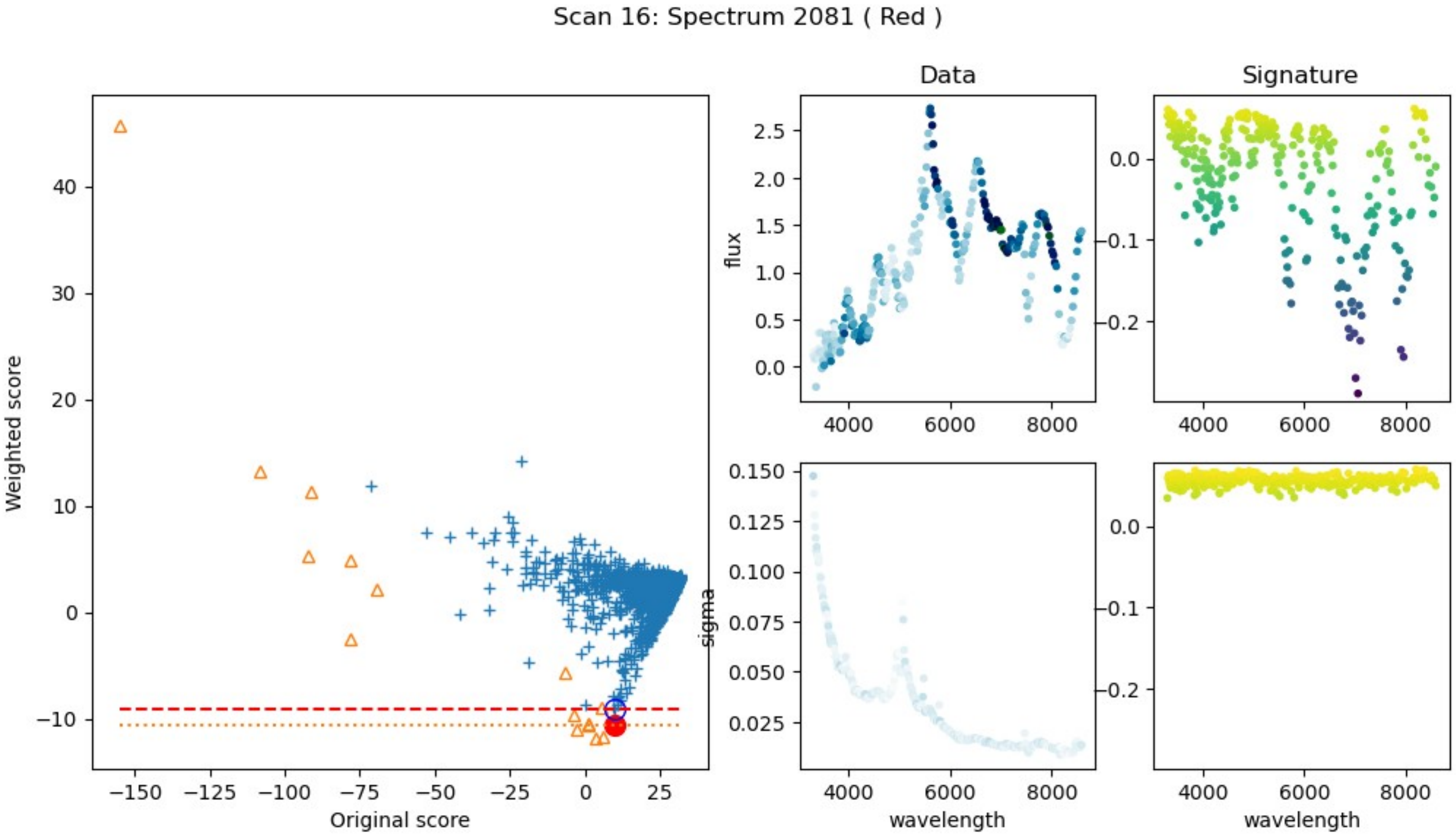
Recursive approach to novelty discovery with signatures



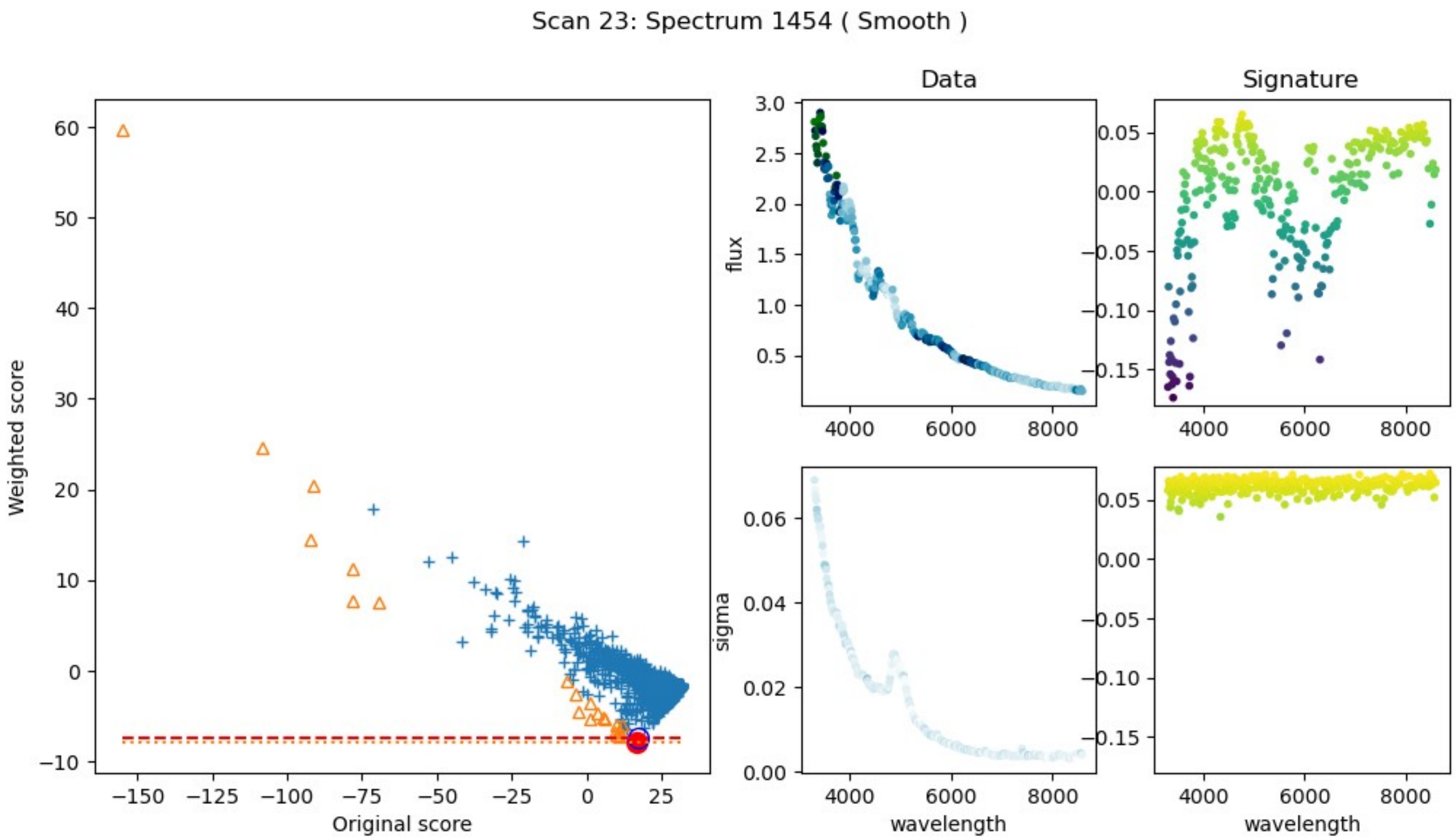
Recursive approach to novelty discovery with signatures



Recursive approach to novelty discovery with signatures



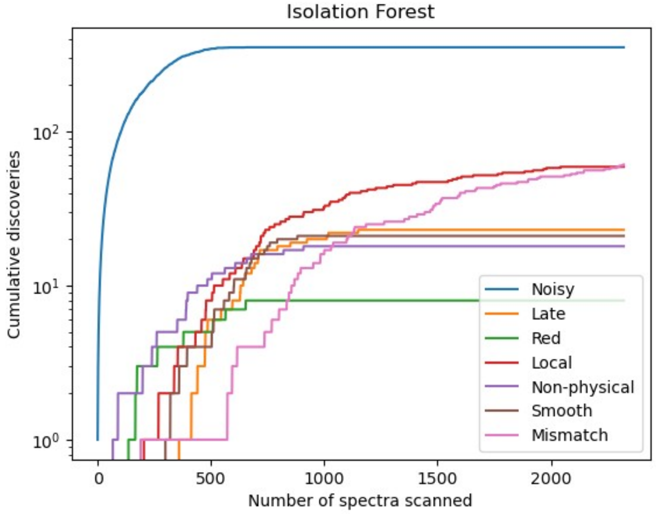
Recursive approach to novelty discovery with signatures



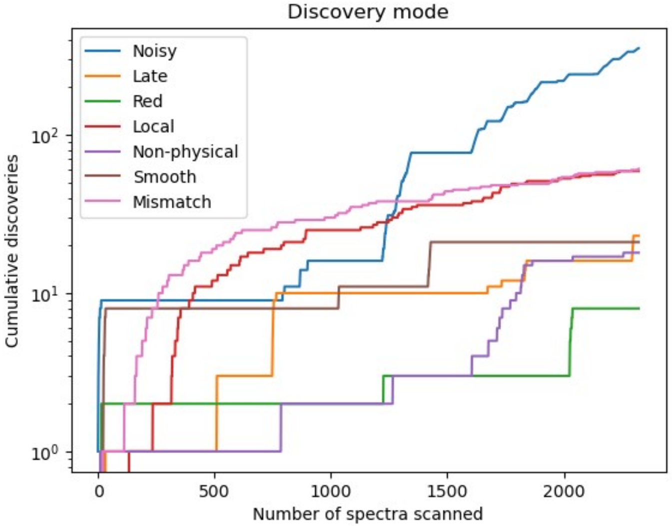
Different tasks :

Factor x3 speed-up in discoveries

- Isolation Forest
  - Default
  - *AUC(Noisy)*=0.98
  - AUC(Others)=0.60
  - Rank of last class=360



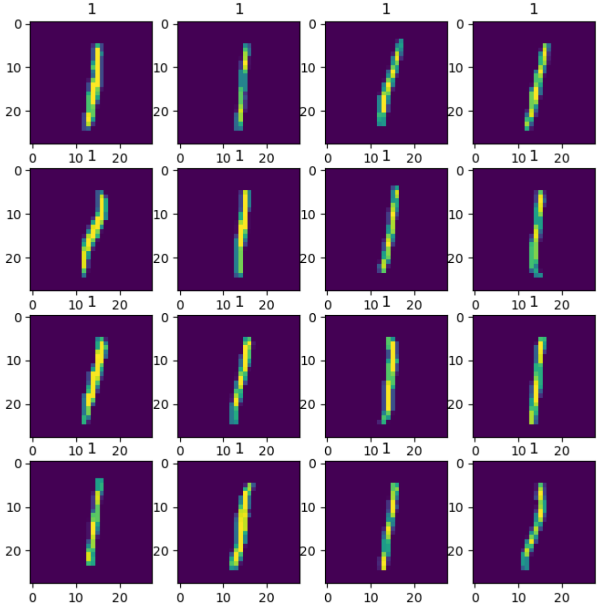
- Discovery mode
  - Optimize on novelty
  - AUC(Noisy)=0.25
  - AUC(Others)=0.51
  - ***Rank*** of last class=133



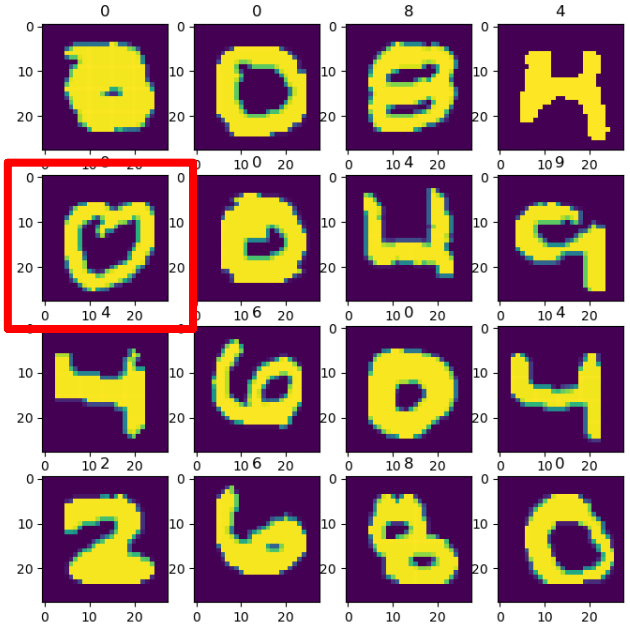
Using Signatures for Similar Anomalies

Anomaly search within MNIST:

- Handwritten digits
- Curated dataset

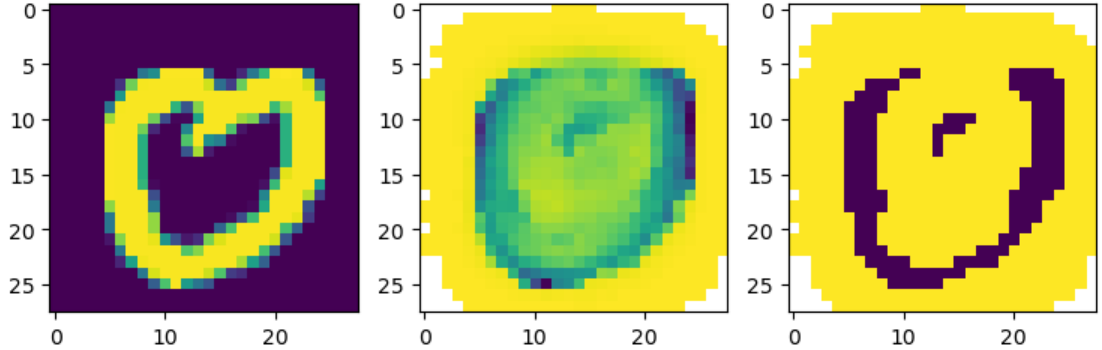


16 most Nominals



16 most Outliers

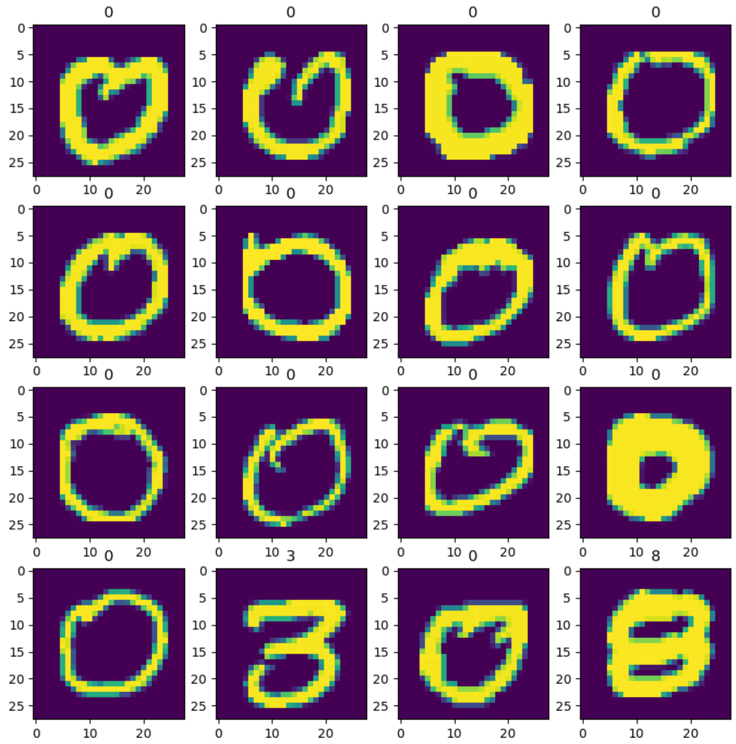
Using signature to select more of the same



Anomaly → Signature → Feature Mask

↓

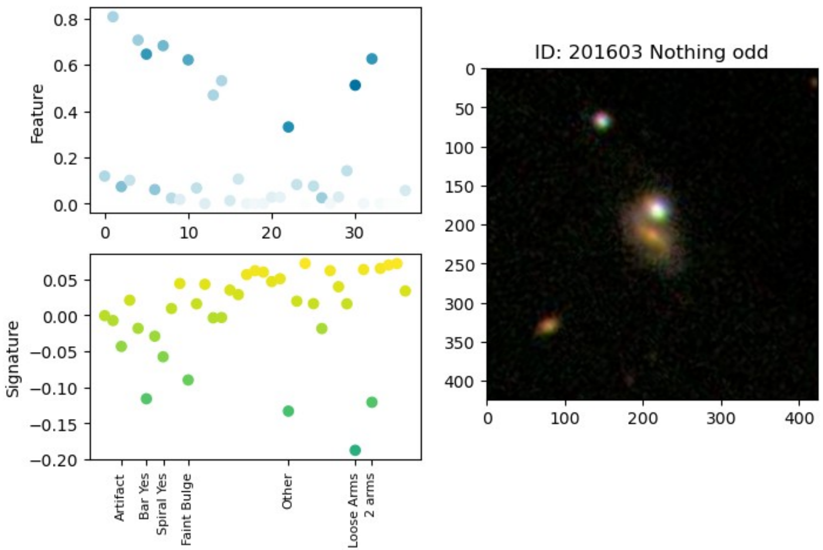
Retraining with selected features

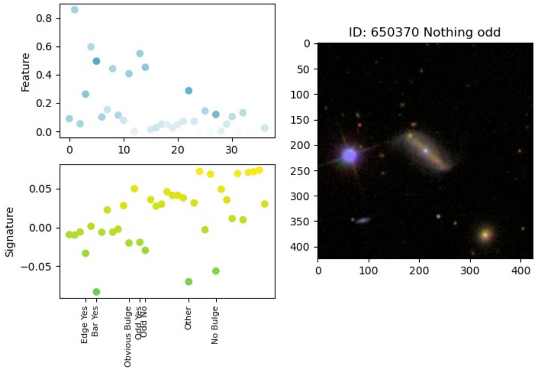


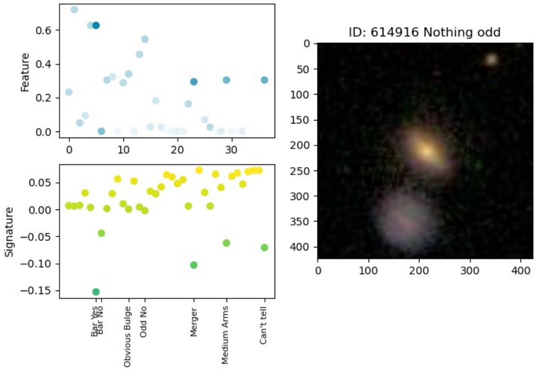
→ Similar Anomalies

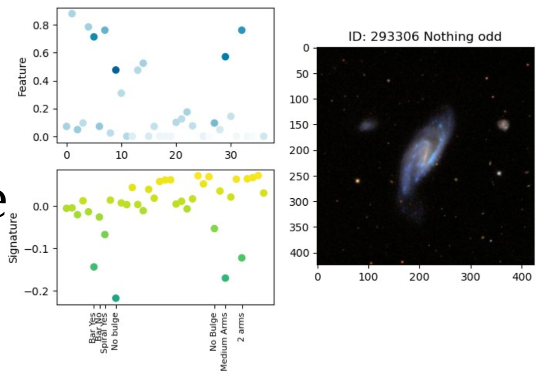
Galaxy Zoo 2 anomalies

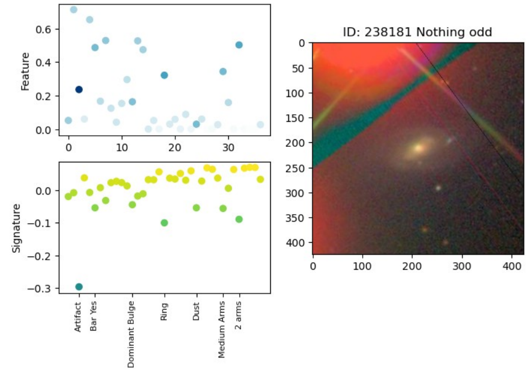
- Anomalies = Rare patterns in human decisions





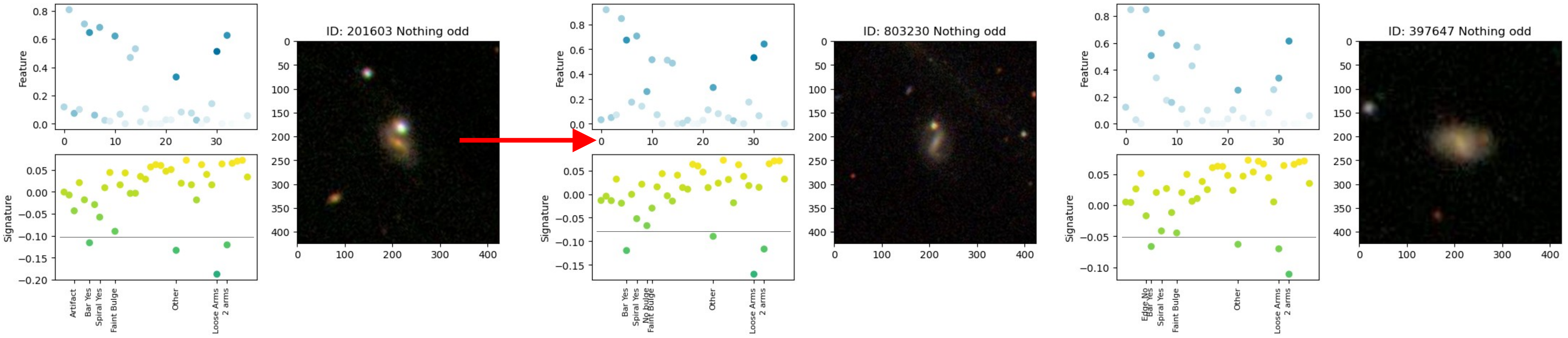


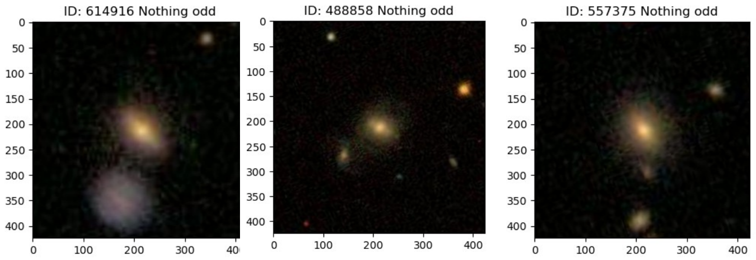




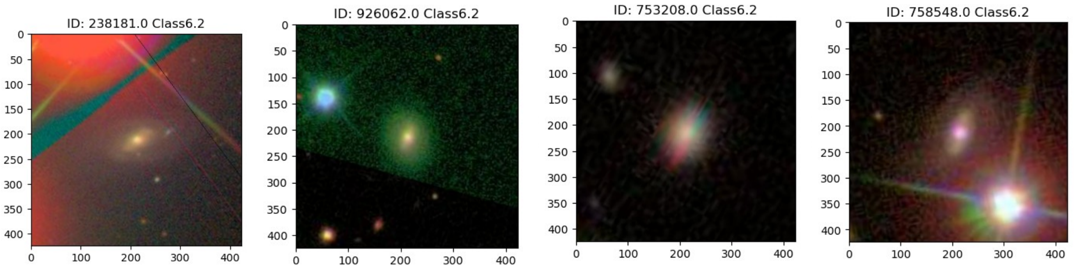
- No obvious pattern in signature
  - → **diversity**

Anomalies by similar signatures





...



...

Rare on "Merger + Bar"

Rare on "Artifact"

Conclusions

- **Anomaly Signature** is a metric for feature importance
- **Domain agnostic / method aware** (Isolation Forest)
  - Works with any tabluar data (features, vectorized images...)
- ***Many use cases:***
  - Interpretability of the decisions
  - Visualisation of outliers
  - Feature selection
  - Categorization of outliers
  - Active learning of anomalies

This is only the beginning !

Stay tuned on

SNAD

https://snad.space

https://github.com/snad-space/coniferest