Model Deployment, Surrogate Models and Anomaly Detection

Jaime Coello de Portugal

PAUL SCHERRER INSTITUT

PSI

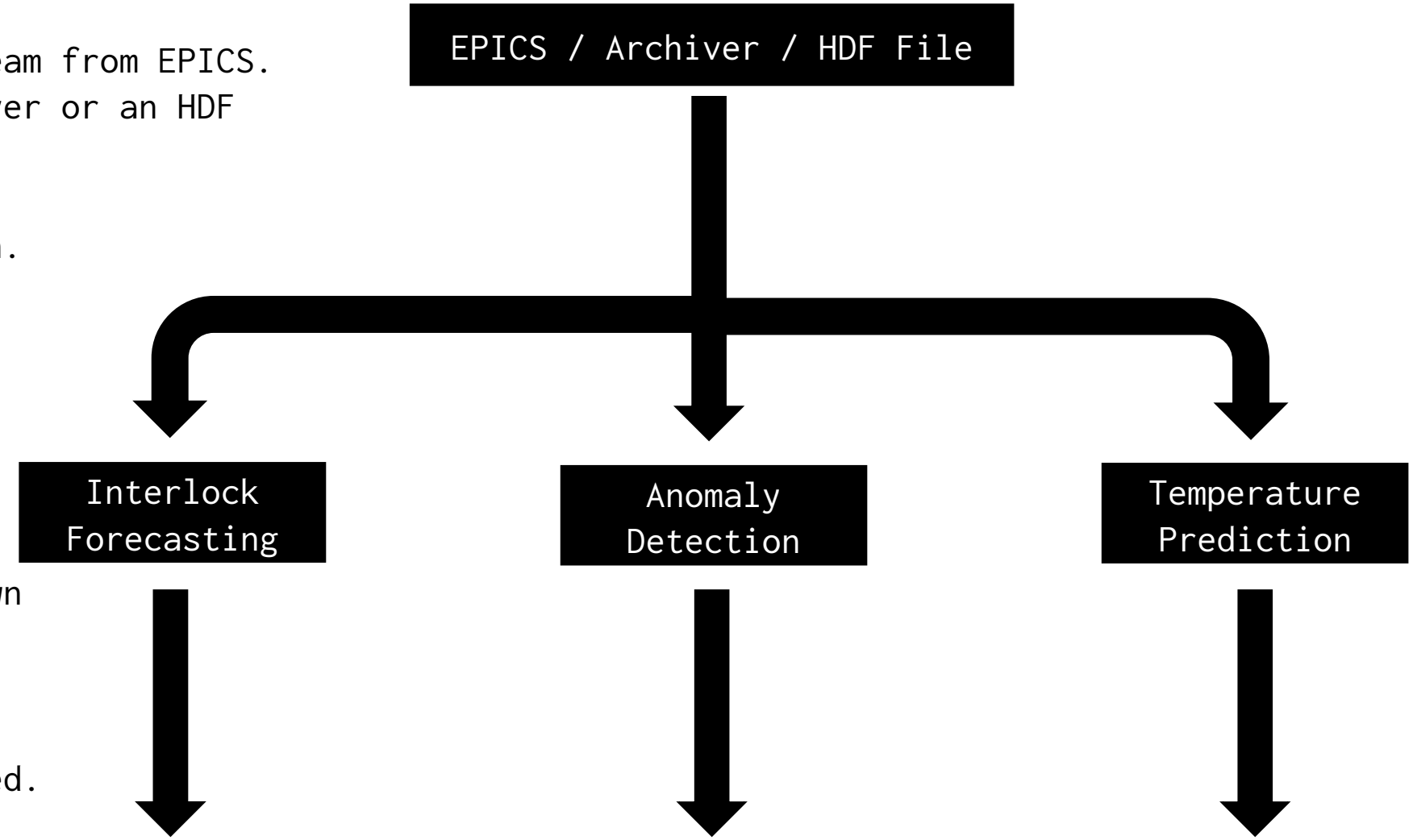
Many thanks to Jochem Snuverink and Melissa Zacharias

Data Collection:

- Alignment of the data stream from EPICS. Or loading from the Archiver or an HDF File.
- Missing data interpolation.
- Windowing as requested.

Model output computation:

- Each model produces its own results.
- Additional diagnostics information can be provided.



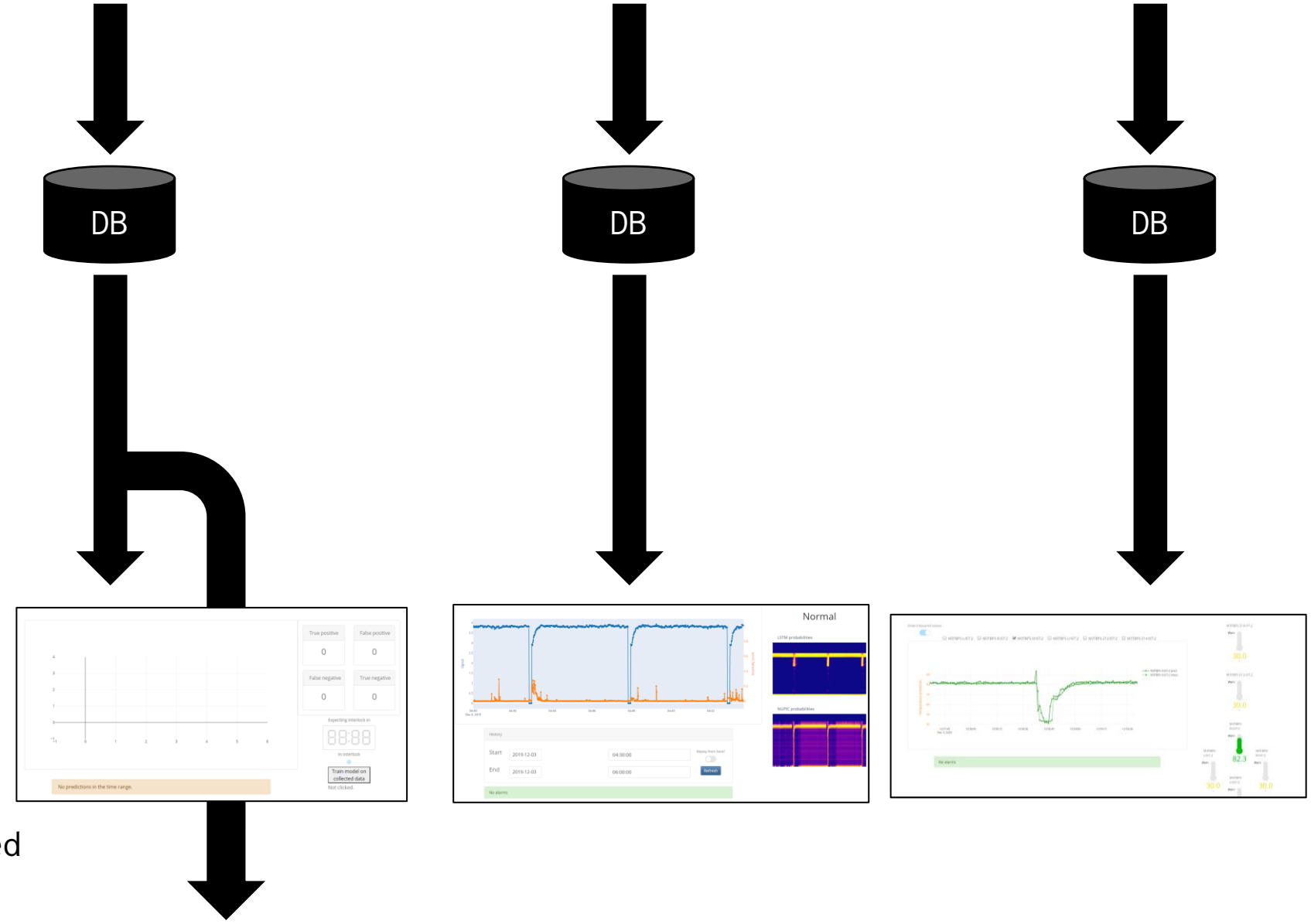
Models deployment

Data persistence:

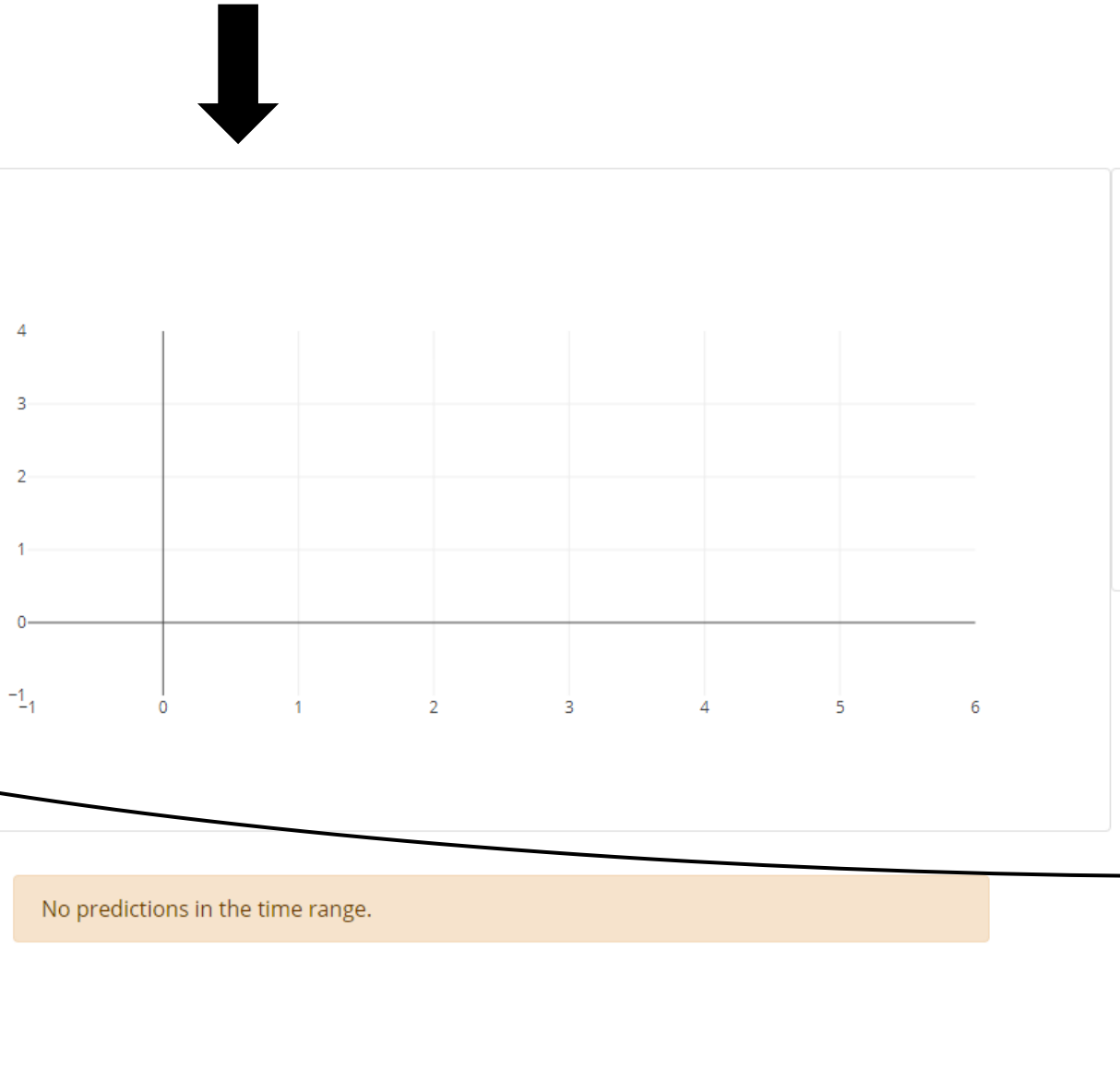
- Input, prediction and diagnostic information are stored in a database.
- Each entry is keyed with the acquisition timestamp.
- All data is available for live display or later analysis.

Display:

- Live or historic display for each of the models.
- Each display can be customized for the specific project.



- As the input data is recorded in the database, it can be used to train the model.
- The interactive training GUI is accessible from the display.

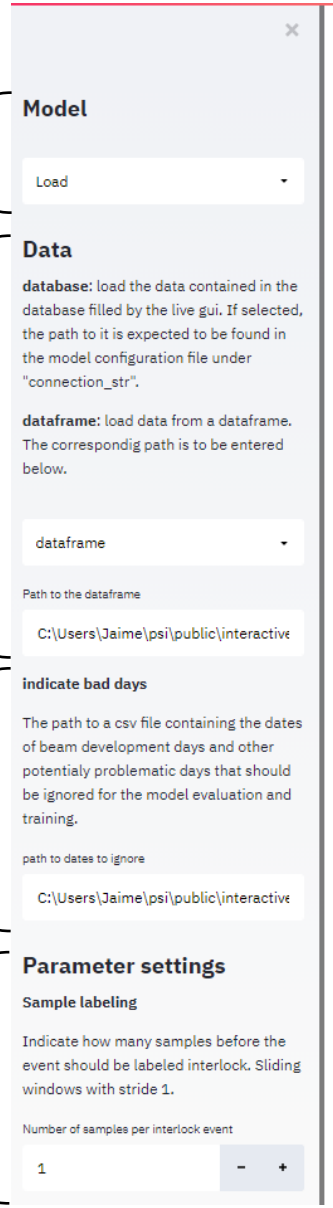


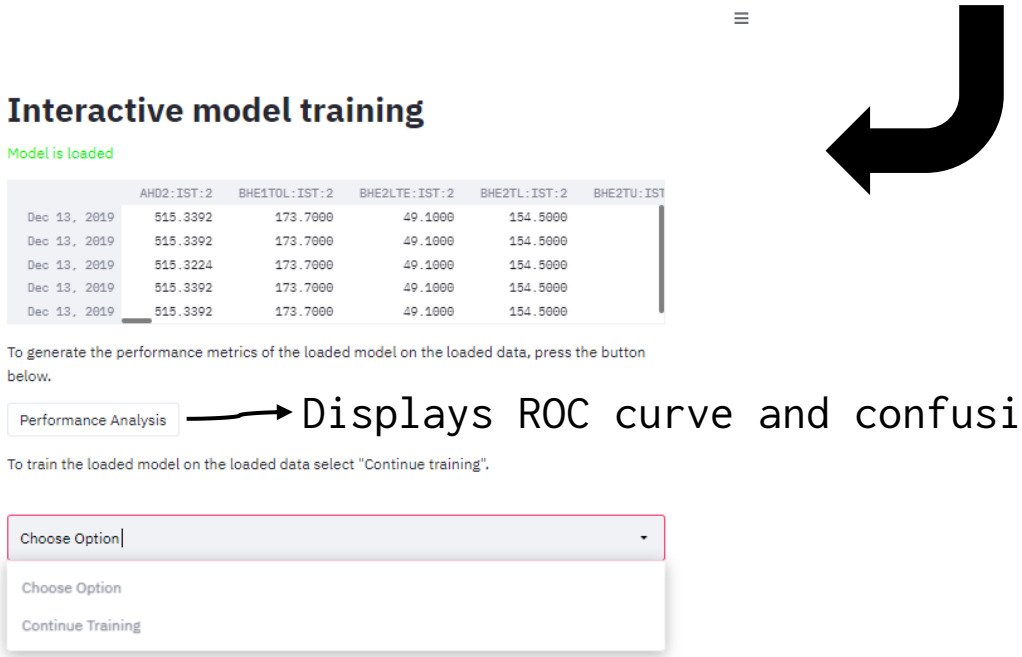
Takes the same configuration file as the live display.

Data can be loaded from the live display DB or from a HDF file.

A CSV file with days to ignore can be given (like beam development days)

Configure sample labelling





Displays ROC curve and confusion matrix

Generates a new, retrained model and stores it to be loaded in the live GUI

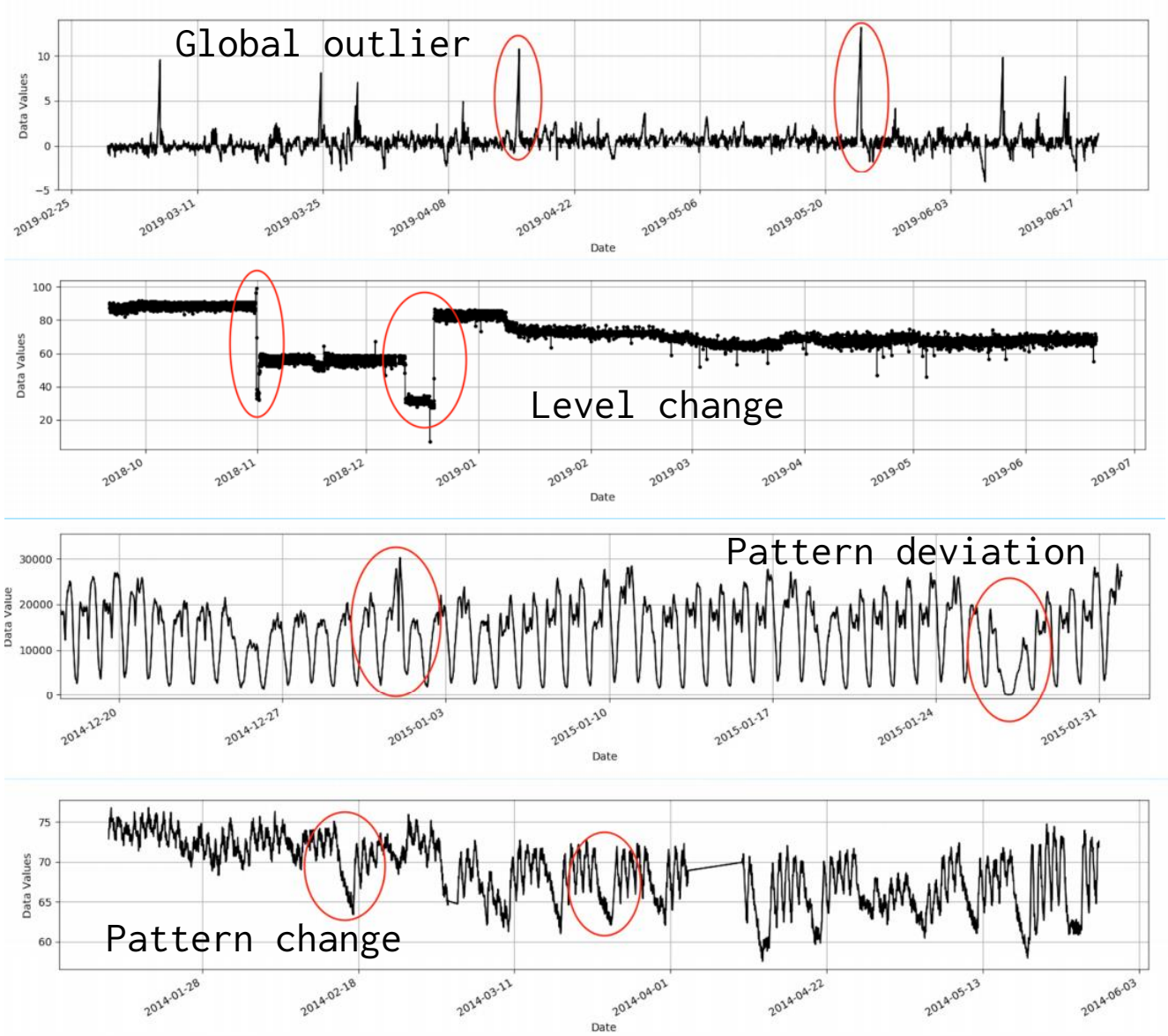
Models deployment conclusions

- All the machine learning models we developed are deployed using an unified system.
- The system can be relatively easily customized to deploy any live ML model on a machine controlled by EPICS.
- All input information, model results and diagnostic information are recorded in a database.
- The information from the database can be shown live or as historical data.
- In the case of the interlock forecasting model, the database information can be used to interactively retrain the model.
- The interactive training GUI allows to check the performance and redeploy retrained models on live recorded data.

"An observation which deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism."

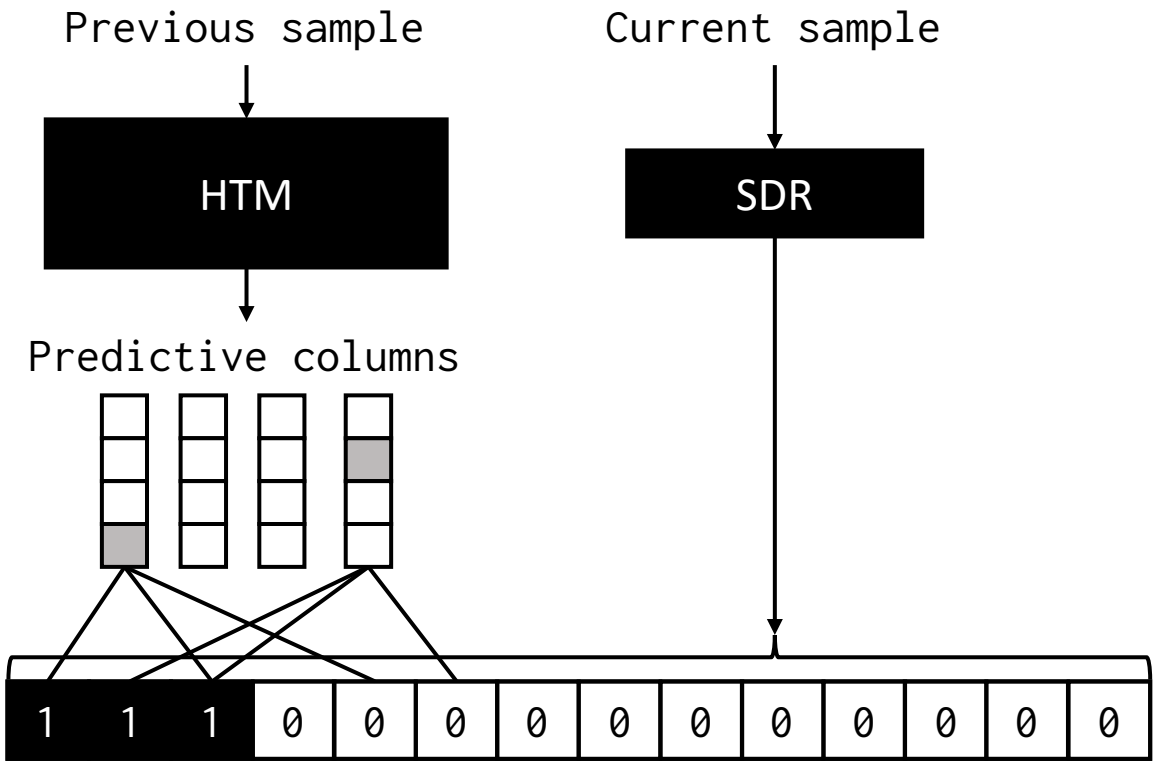
— Hawkins, 1980

- Try to find observation or sequences that deviate from the "normal behaviour".
- Experts would recognize these anomalous patterns easily but cannot be monitoring the huge amount of data some systems produce.
- E.g: Credit card fraud detection, intrusion detection in cybersecurity, or **fault diagnosis in industry**.
- Specific e.g: At HIPA the MHB7R:ILOG:2 temperature detector broke down without anyone noticing.



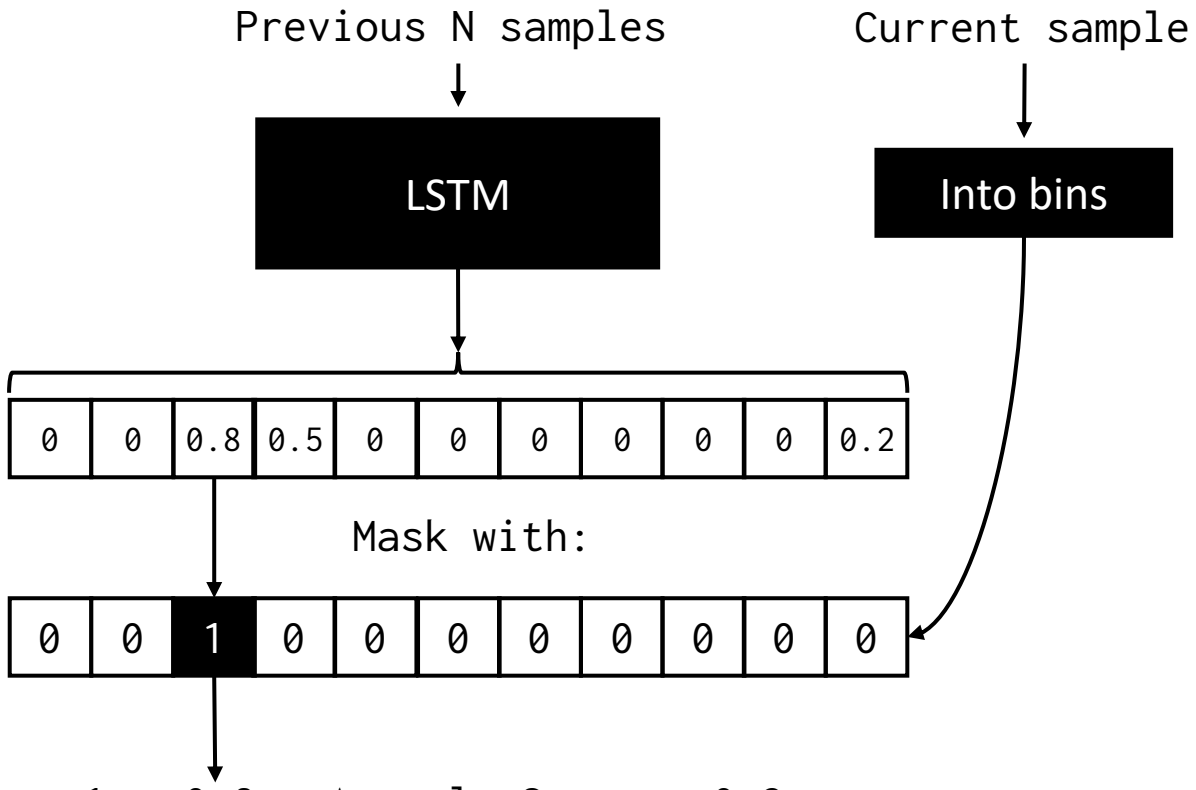
Plots from: Ted D'Ottavio et al, ***Experience Using NuPIC to Detect Anomalies in Controls Data****, ICALEPCS 2019*

HTM Anomaly Score:



$$\text{Anomaly Score} = \frac{\text{Correctly predicted inputs}}{\text{Number of connections}}$$

LSTM Anomaly Score:

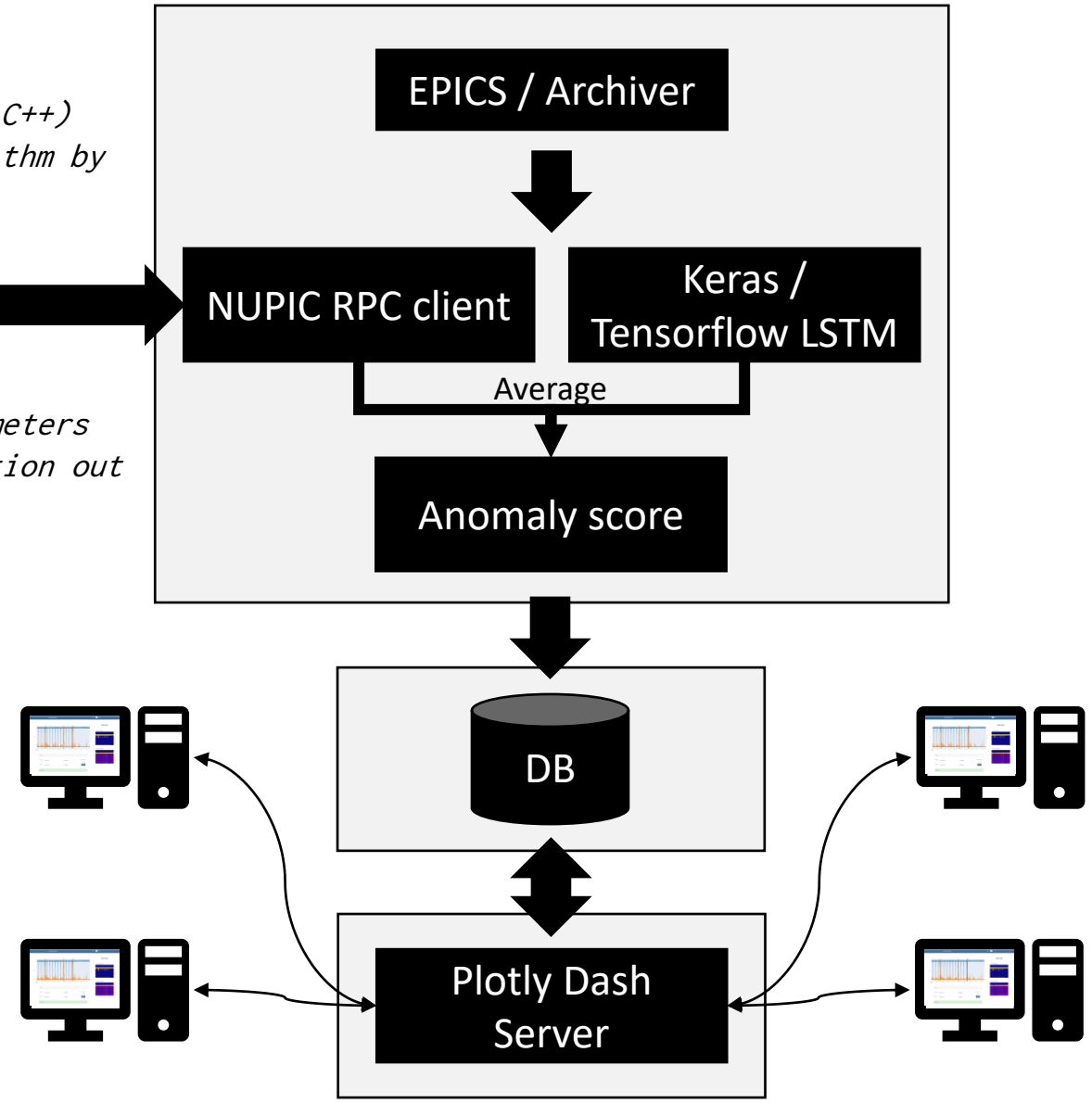


Training penalty = Anomaly Score + $\gamma \sum |prediction|$

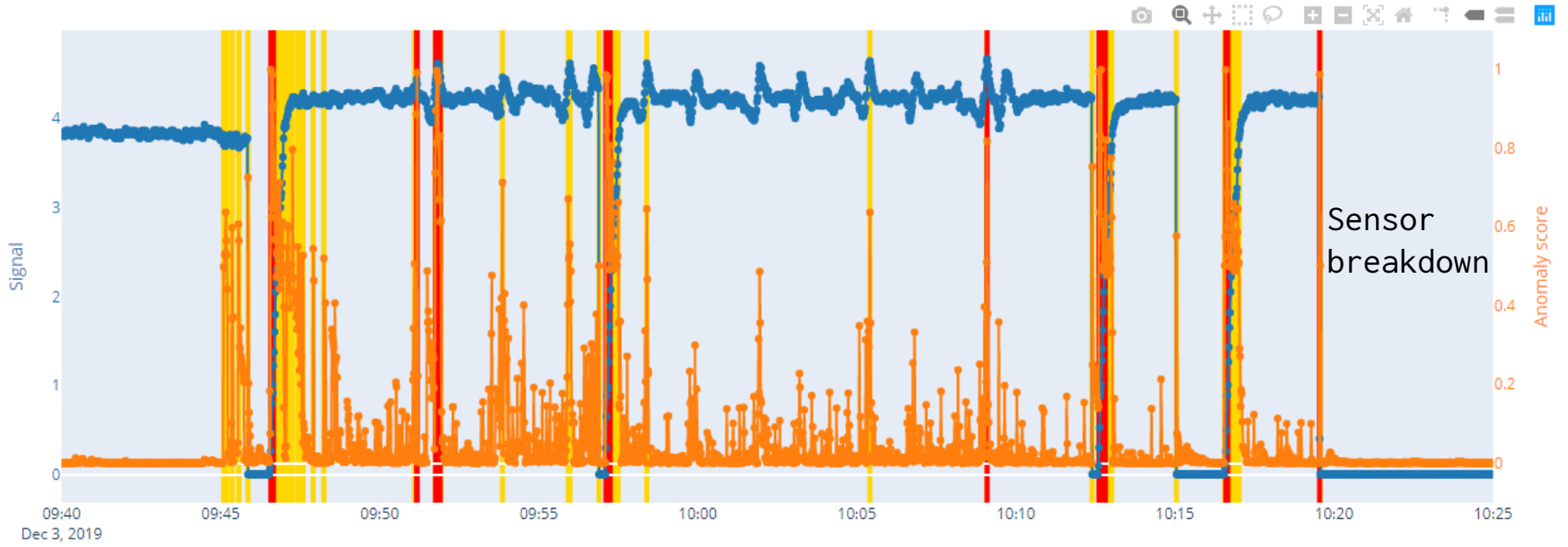
Anomaly detection: HTM and LSTM combination

NUPIC is a Python 2 (wrapping C++) implementation of the HTM algorithm by Numenta (numenta.com).

They provide a set of hyperparameters that works good for anomaly detection out of the box



Anomaly detection in HIPA



History

Start 2019-12-03 09:40:00

End 2019-12-03 10:25:00

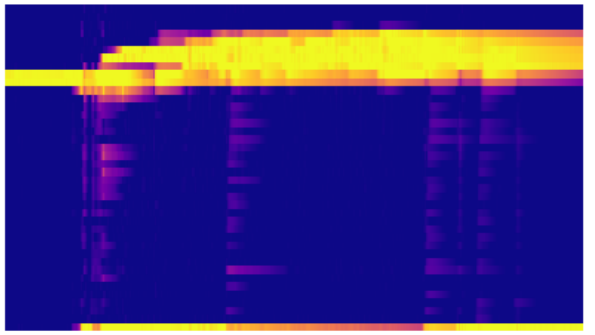
Replay from here?

Refresh

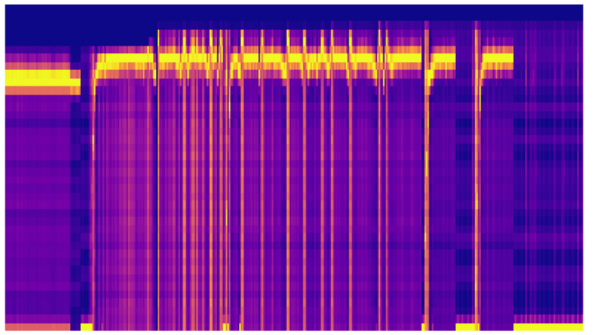
No alarms

Normal

LSTM probabilities



NUPIC probabilities

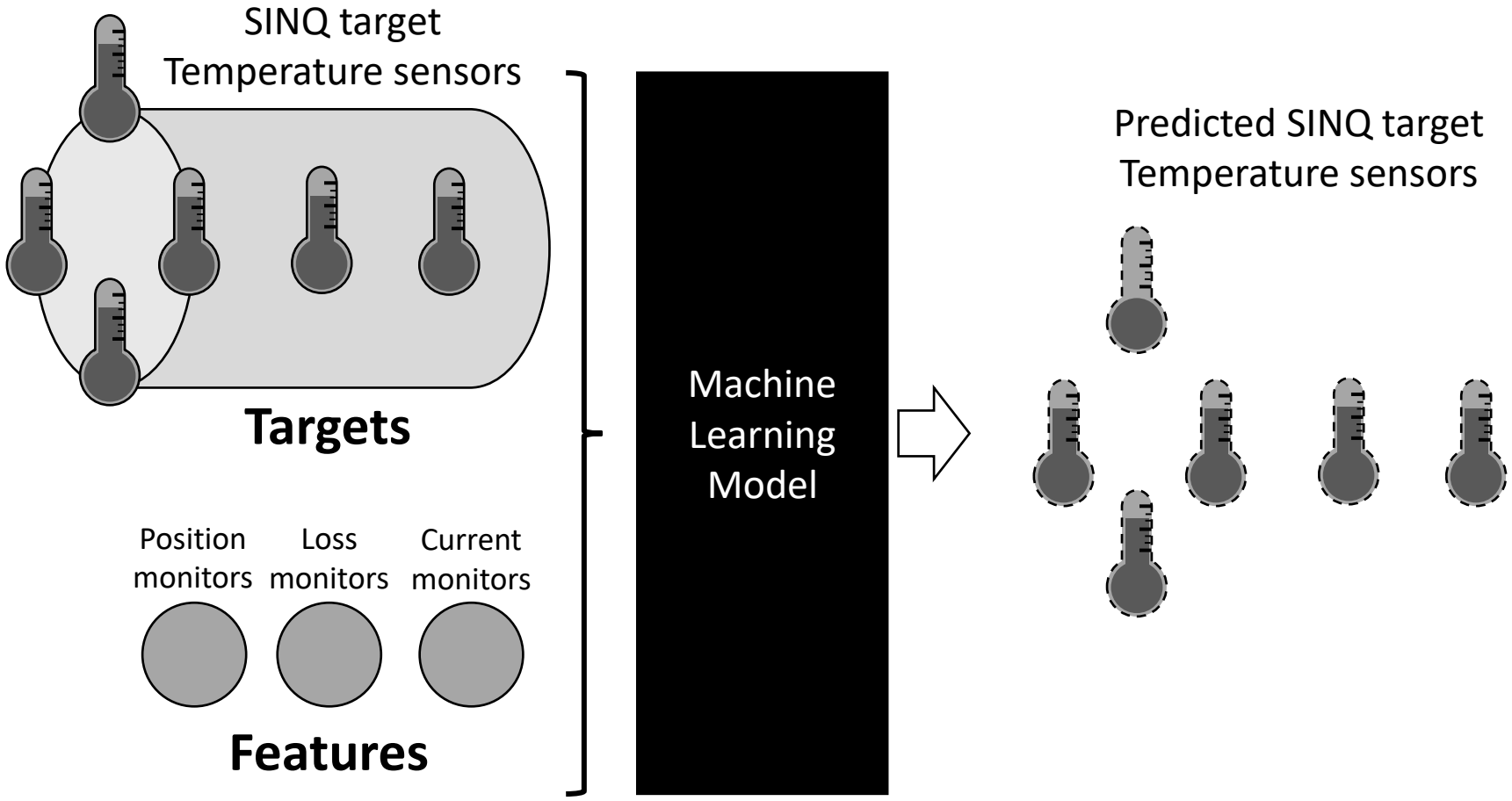


Live anomaly detection in PROSCAN

- Anomalies detected live in the vacuum level in the cyclotron:

Anomaly detection conclusions

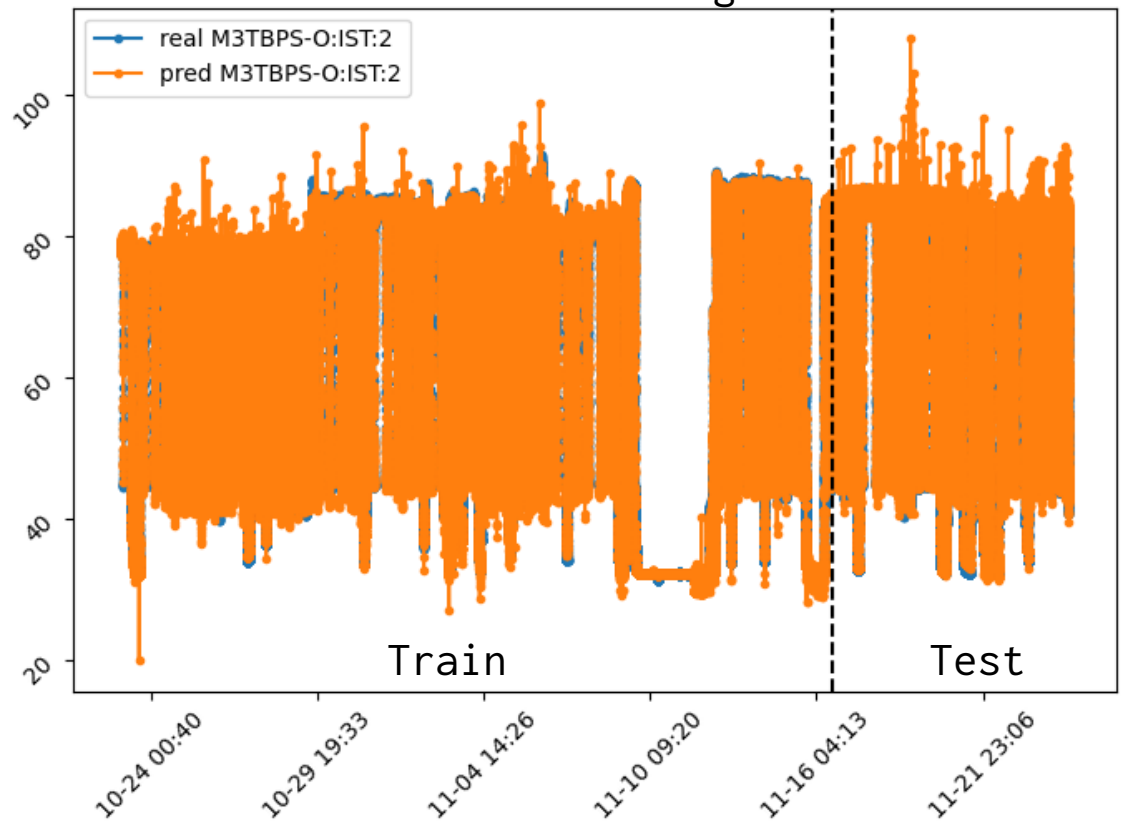
- The combination of the HTM network and the LSTM seems to provide a good real time anomaly detection system.
- The system works well on archived and live data, adapts to both stable and periodic behaviour.
- It would have alerted the experts of the failure of the MHB7R:ILOG:2 sensor tenths of minutes in advance.
- Tested on real time on PROSCAN.
- Preparing a publication of the method and the use at HIPA.


SINQ target
Temperature sensors
Targets
Position monitors
Loss monitors
Current monitors
Features
Machine Learning Model
Predicted SINQ target
Temperature sensors

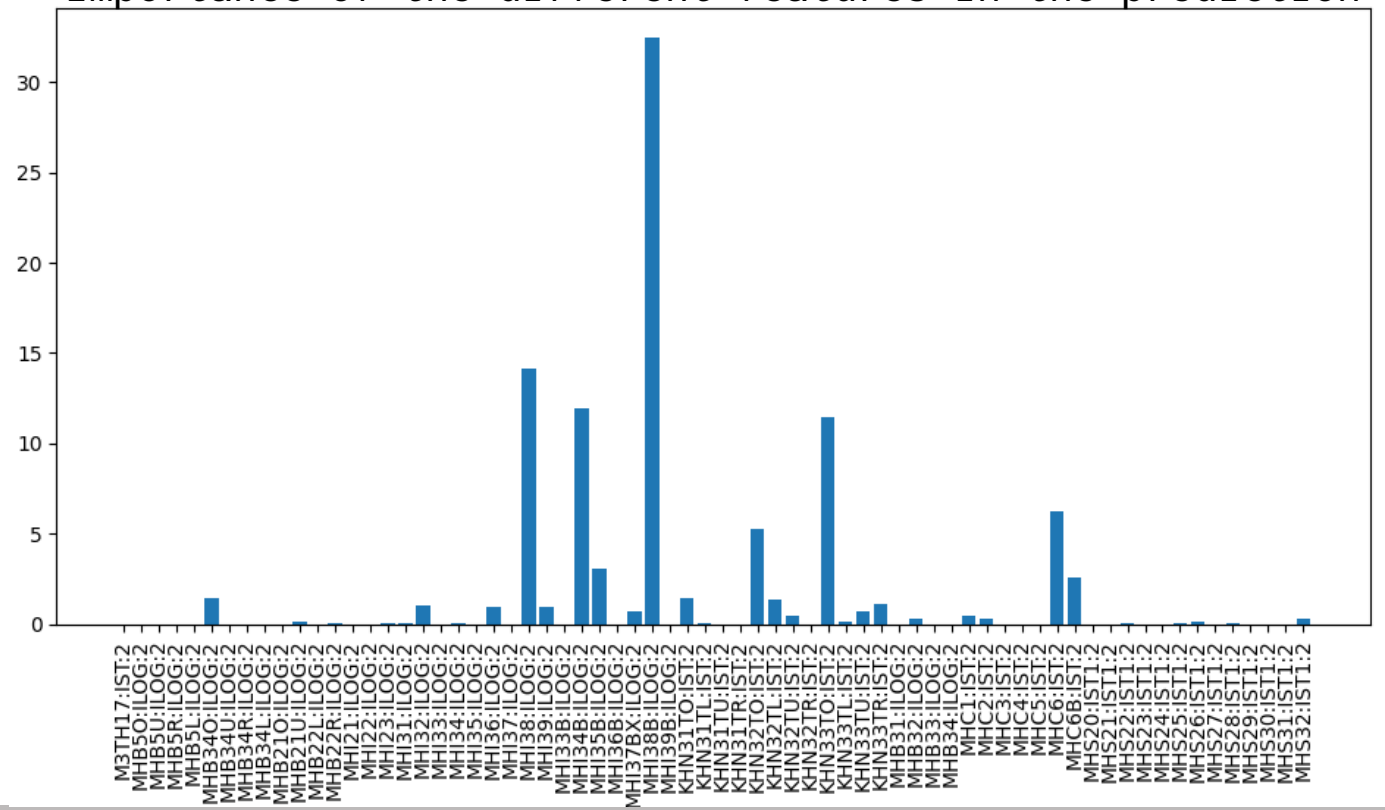
SINQ temperature sensors: Virtual diagnostics

- Model for M3TBPS-O:IST:2 (one of the four external blades).
- Uses 68 monitors all over the machine, many in the SINQ line.
- The model is a Gradient Boosting model (tree ensemble) made with CatBoost.

Prediction vs. actual reading of M3TBPS-O:IST:2

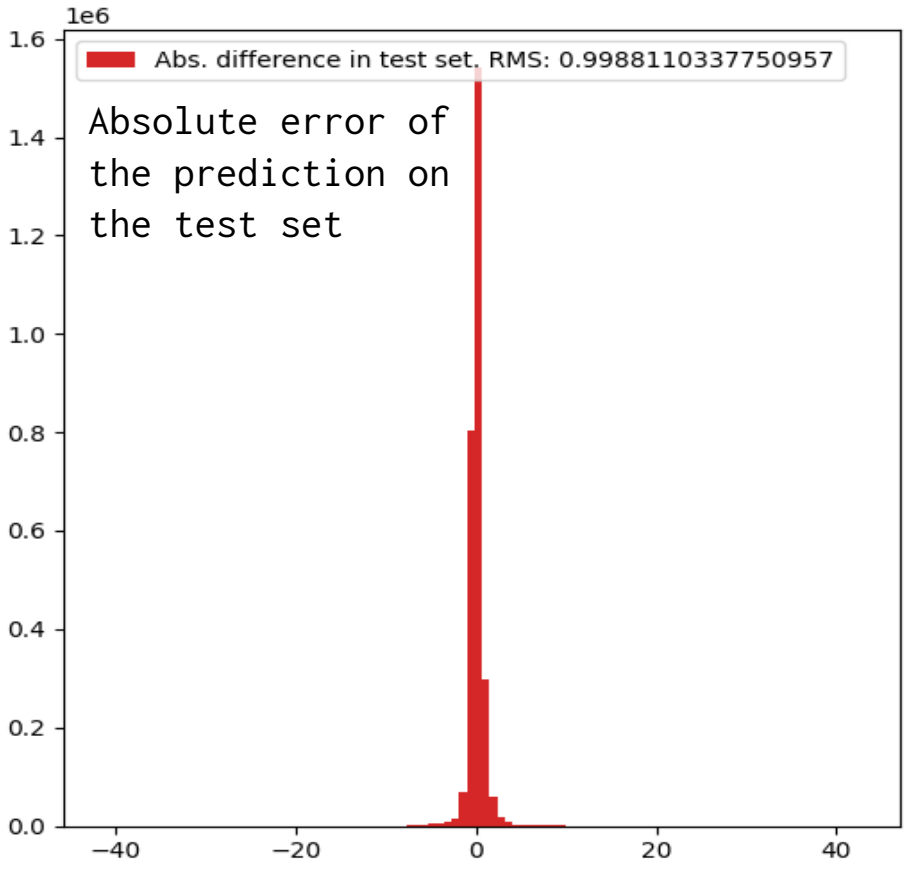


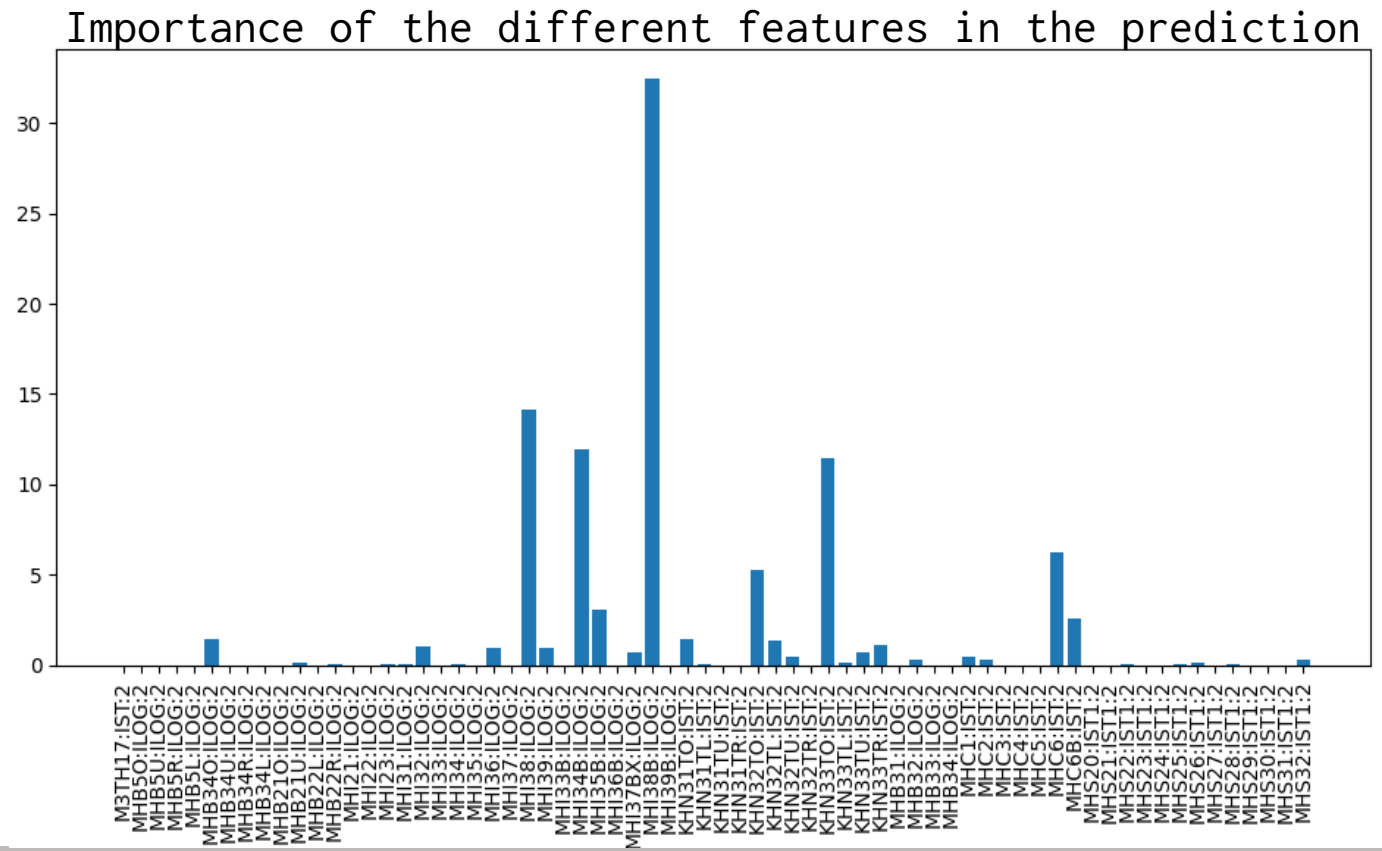
Importance of the different features in the prediction



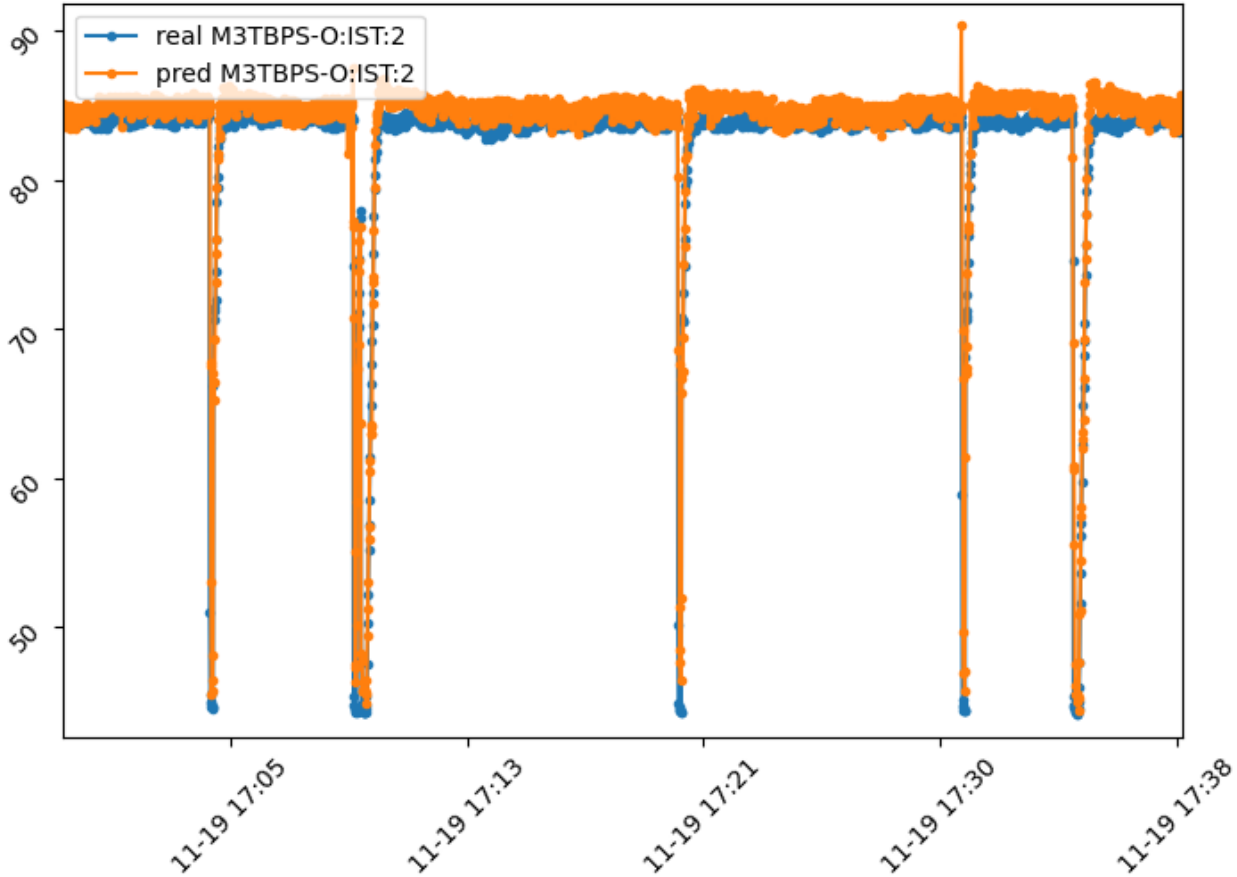
SINQ temperature sensors: Virtual diagnostics

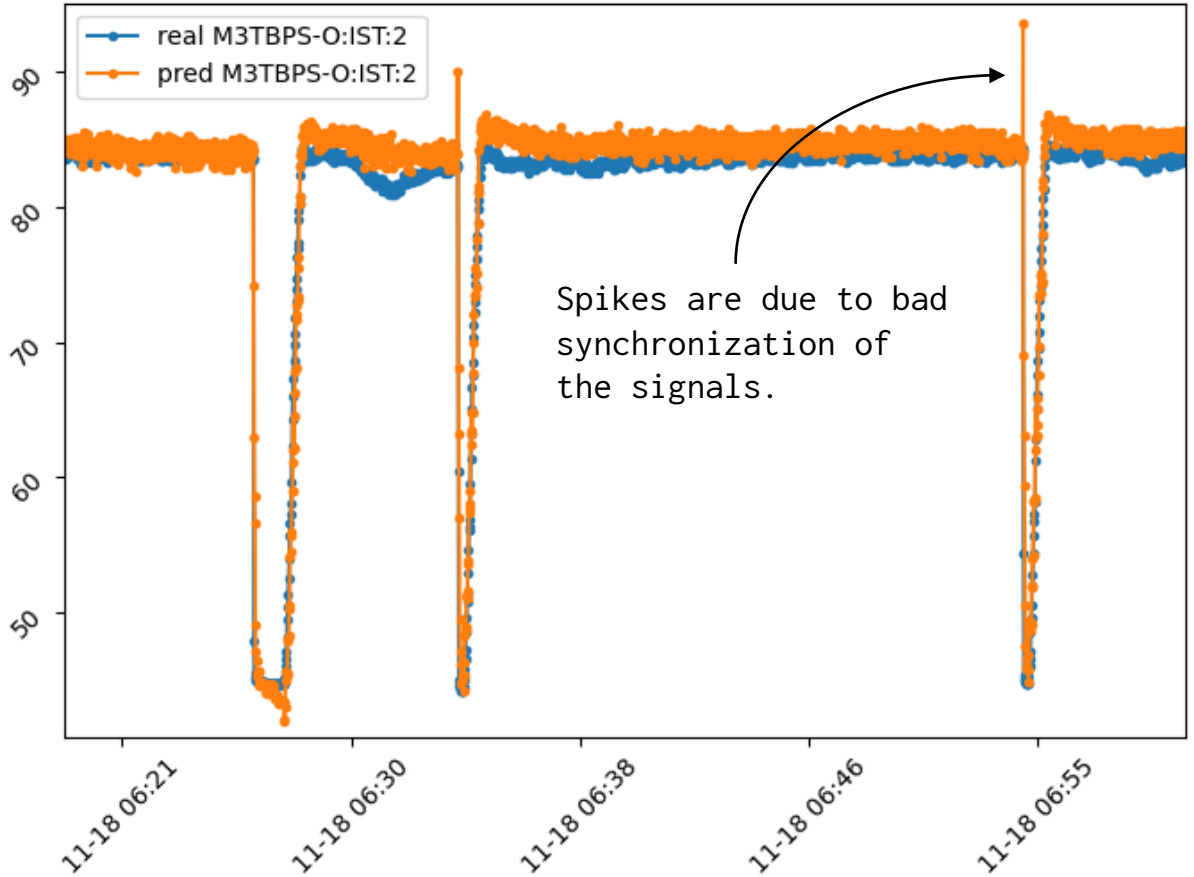
- Model for M3TBPS-O:IST:2 (one of the four external blades).
- Uses 68 monitors all over the machine, many in the SINQ line.
- The model is a Gradient Boosting model (tree ensemble) made with CatBoost.



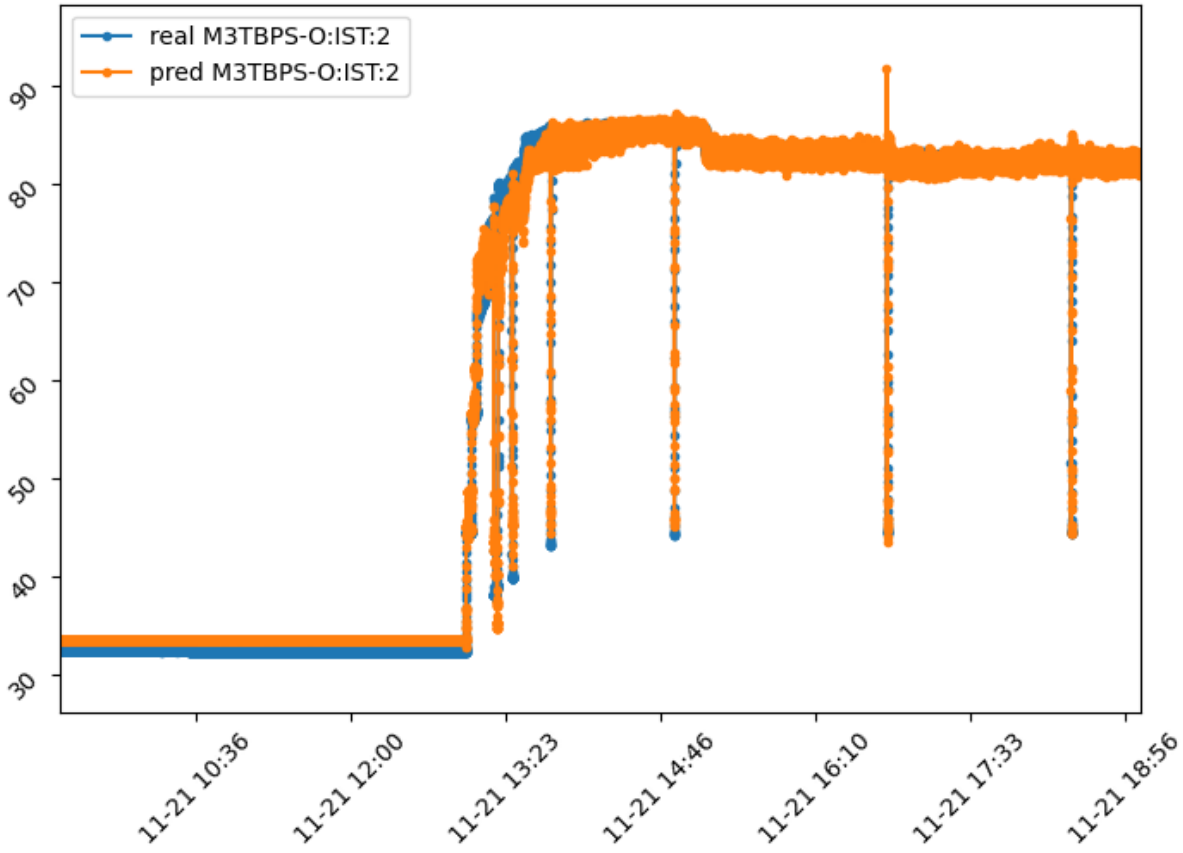


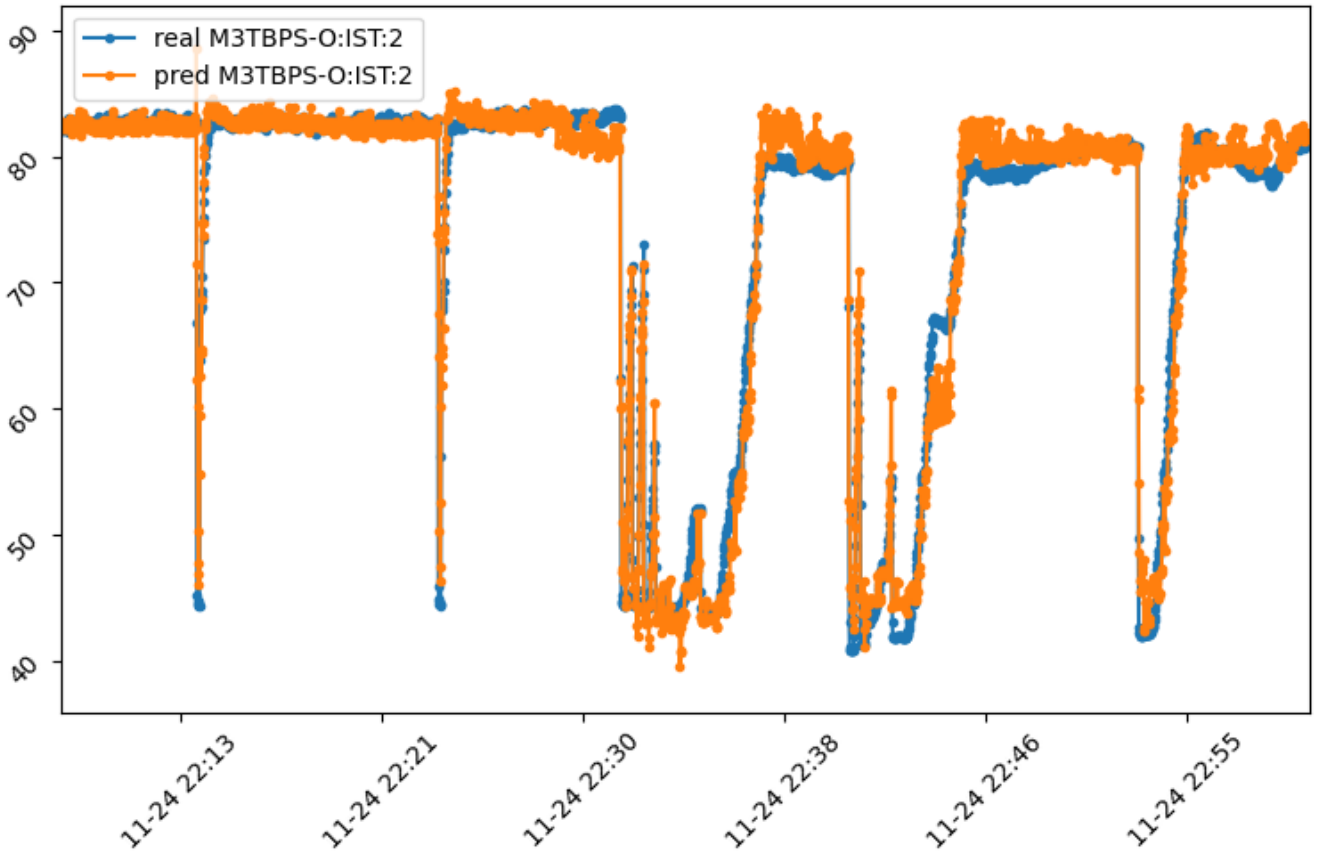
Zooming in some examples in the test set:





Zooming in some examples in the test set:





SINQ temperature sensors: Virtual diagnostics

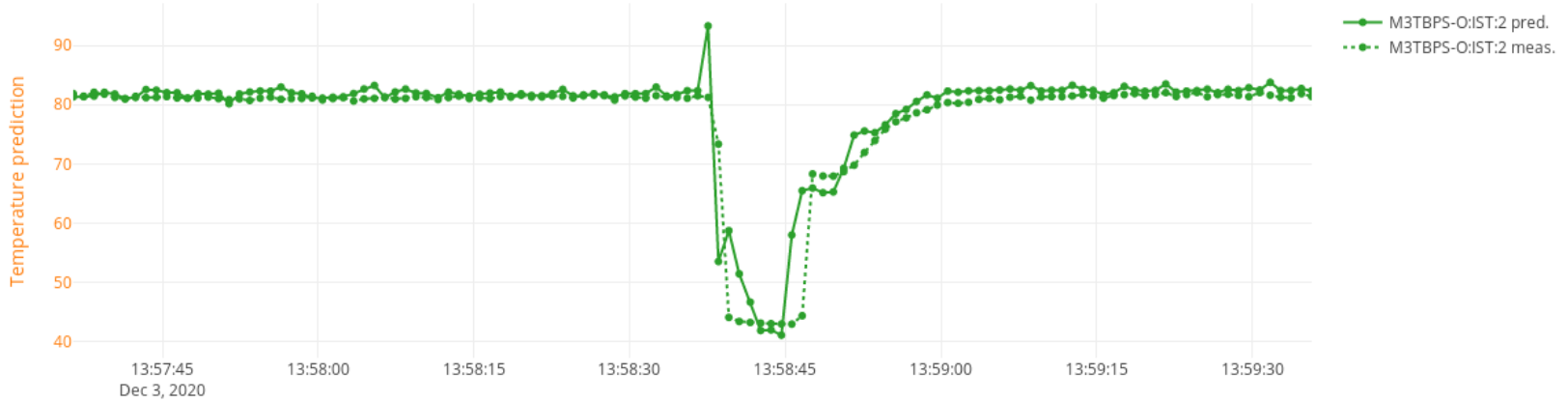
SINQ temperature predictions

History Live

Show measured values

M3TBPS-L:IST:2 M3TBPS-R:IST:2 M3TBPS-O:IST:2 M3TBPS-U:IST:2 M3TBPS-Z12:IST:2 M3TBPS-Z14:IST:2

Model results running live at HIPA many days after the training set:



No alarms

M3TBPS-Z14:IST:2
Warn
30.0
c

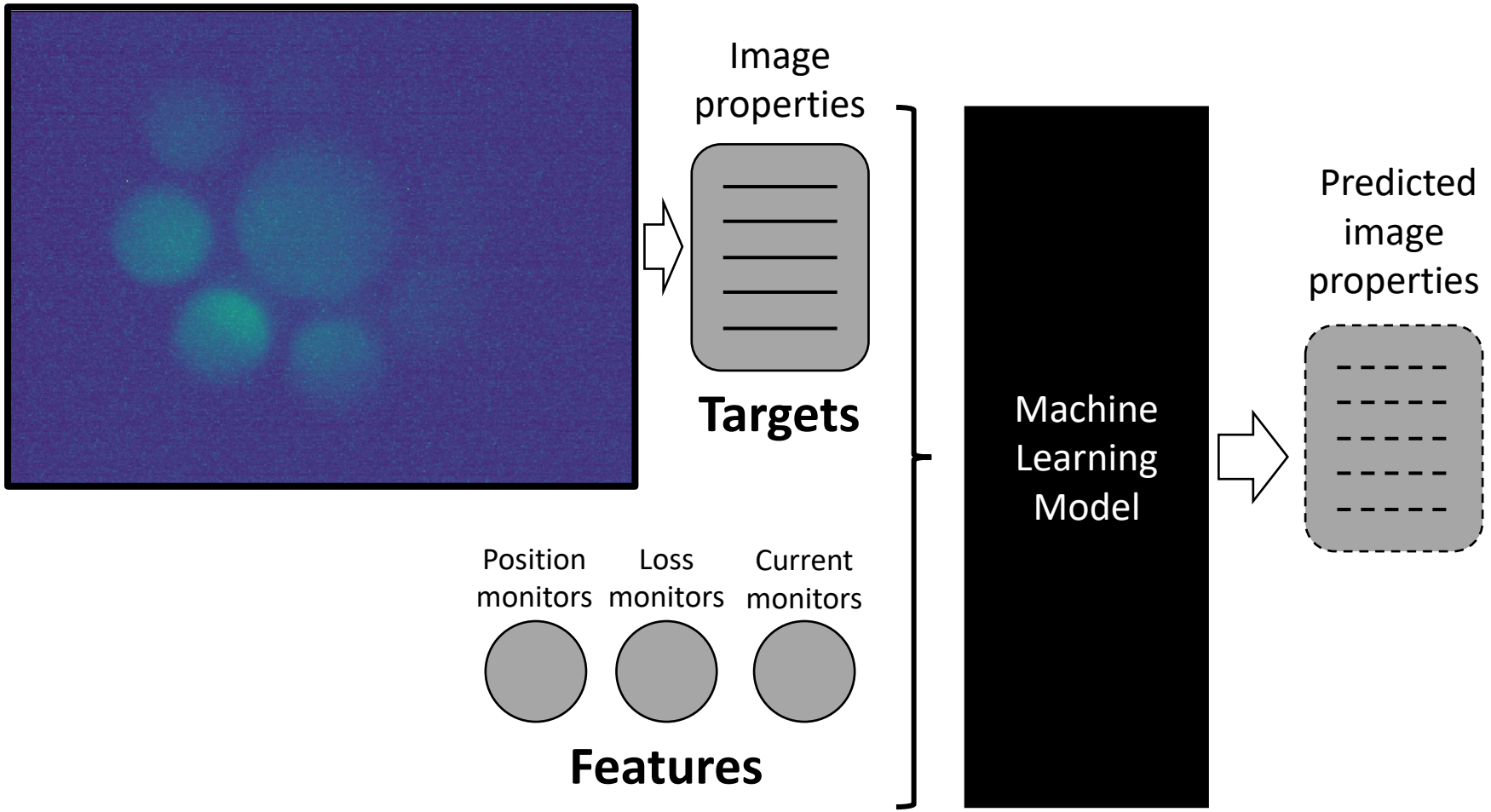
M3TBPS-Z12:IST:2
Warn
30.0
c

M3TBPS-O:IST:2
Warn
82.3
c

M3TBPS-L:IST:2
Warn
30.0
c

M3TBPS-R:IST:2
Warn
30.0
c

M3TBPS-U:IST:2
Warn



- Good models are available, but the camera image processing is being improved.

Virtual diagnostics conclusions

- A model for M3TBPS-O:IST:2 has been developed and predicts the signal of the sensor within 1 degree.
- The model uses just passive beam diagnostic sensors.
- The predictions stays correct on data observed weeks after the training set, no long term drifts.
- In 2021, the temperature sensor array is being changed to new locations in the target. The model could provide temperature predictions in a location where no sensors are available.
- A report of these models is planned to document the development.