Alpha-i Detection Service for Aerospace

The ADS for Aerospace is a platform which give the end user the power to analyse flight sensors data and predict anomalous behaviors by controlling the actions of a machine learning algorithm. The clean and simple web user interface permits the execution of the following actions:

1. Build a dataset by uploading flights data
2. Train or Retrain a ML model on a specific part of the dataset
3. Running Detection on a uploaded data
4. View and Explore the flight data as well the root cause of an anomaly

Dataset

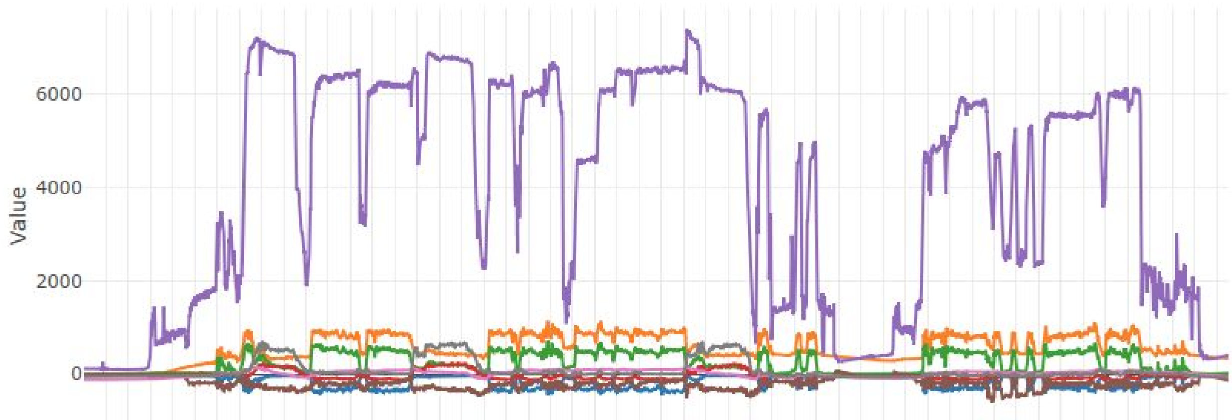
Define the data type

Before loading data in the platform, the administrator must define Class of Data by configure some parameters as the number of sensors, the sample rate, and other parameter as required by the underlying model.
This operation can be done only through an api call.

Build your dataset

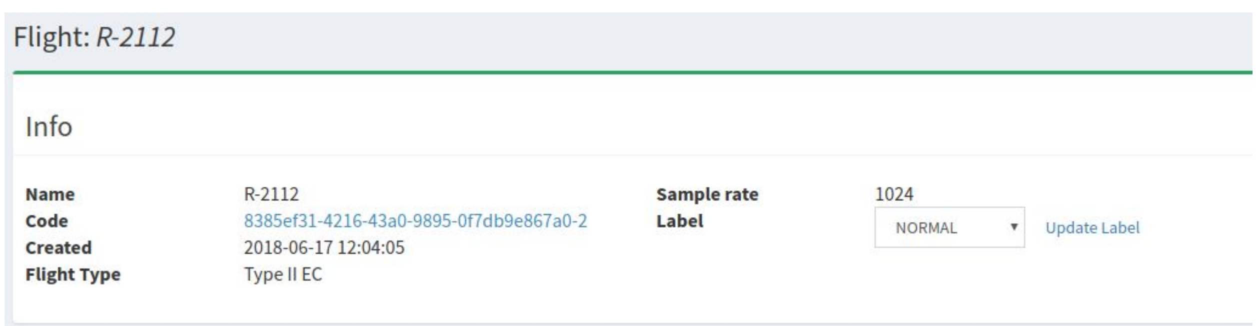
Once the class of data are defined the user is then allowed to upload files containing the flight data.
The format accepted is an HDF file containing one single flight in a pandas dataframe format.
Each column of the dataframe contains the data for a specific sensor.
The row contains data for a single timestep.
Once uploaded, an overview of its content is plotted on screen

.



Label your data

Once a file is uploaded its content can be labeled as ***normal*** or ***abnormal***. This is the case when the content of the flight is known and it is going to be used as a part of a training set.



The flights of which the user doesn't know its status are left *unlabeled* and subject of a *detection*

Train your model

Before being capable to run a Anomaly detection the ml algorithm needs to be trained on normal flight data.
To do that the user must define the following.

Flight type
Select from the existing types of flight data.
Parent training
Select a parent training if you want the new model to originate from a pre-existing model configuration.
Domain

Select between time and frequency (fft) domain.
Downsample factor
Factor by which the resolution of the input data is reduced. E.g. A downsample factor of 100 means that input data sampled at 1kHz will be processed at a resolution of 10Hz by means of averaging.
Number of iterations
Number of iterations used to train the model. We recommend: 50,000 for a training from scratch; 10,000 for a re-training

Difference between training and retraining

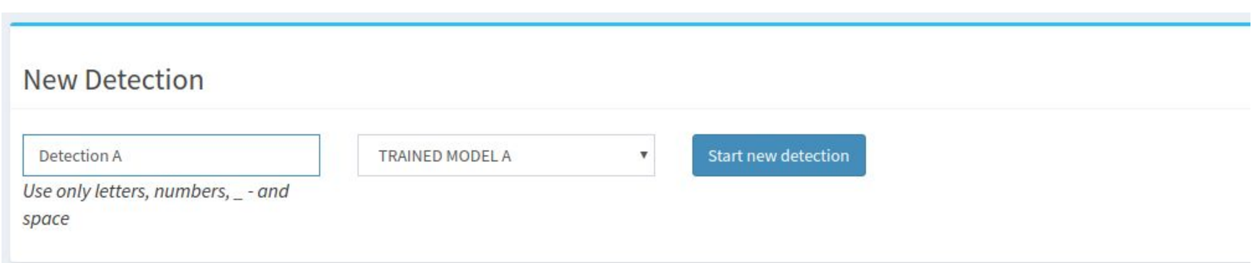
When training, the algorithm start with no-knowledge of data.

When re-training, the algorithm loads a previous neural network configuration and creates a new one by analysing the new added data.
This makes the model an organic entity which improve its detection capability as more data is added to the platform.

Training task are long running processes that run asynchronously.

Detect anomalies

Once the platform is loaded with a training model for a specific data type, a detection task can be created.
This can be done on the Flight page using the form *New Detection* where the user must choose with which trained model he wants to perform the detection and a Detection Name



A detection task might take some time to complete but by running it in the background the platform leave the user free to uploading other data or navigate and explorer previous detections.
After a detection task completes, a root cause analysis task is fired.

Detection Result

On detection completion the result are plotted in the detection with a filled red line page alongside the normalized content of the flight.

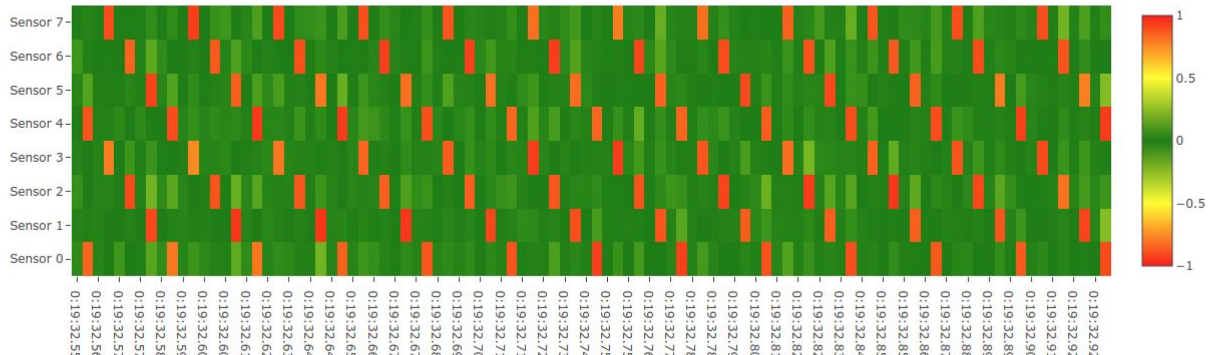
The higher the value of the red line, the higher the probability of anomaly.

Root cause Analysis

The root cause analysis is a machine learning task where the most anomalous chunk of the detection result (and its previous and next) are deeply analysed to uncover which component of the signal is causing the anomaly.
The result plot is composed by an heatmap which color temperature shows how much the real chunk is different from an optimal model-generated one according to the data seen during the training.
The second part of the diagnostic result is the plot which shows the real content of the chunk analysed.

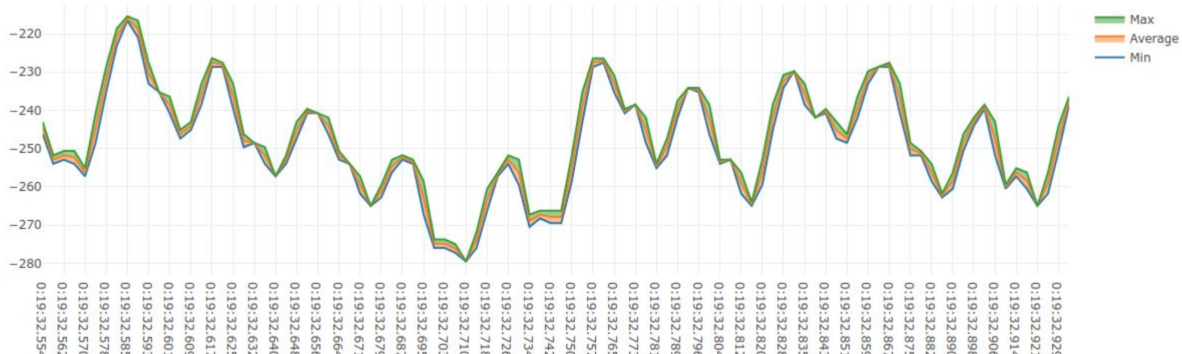
The diagnostic depends on which training set has been used for the detection. A *time domain* detection will show a time domain heatmap.
A *frequency domain* detection will show a spectrum analysis heatmap.

Time Domain RCA

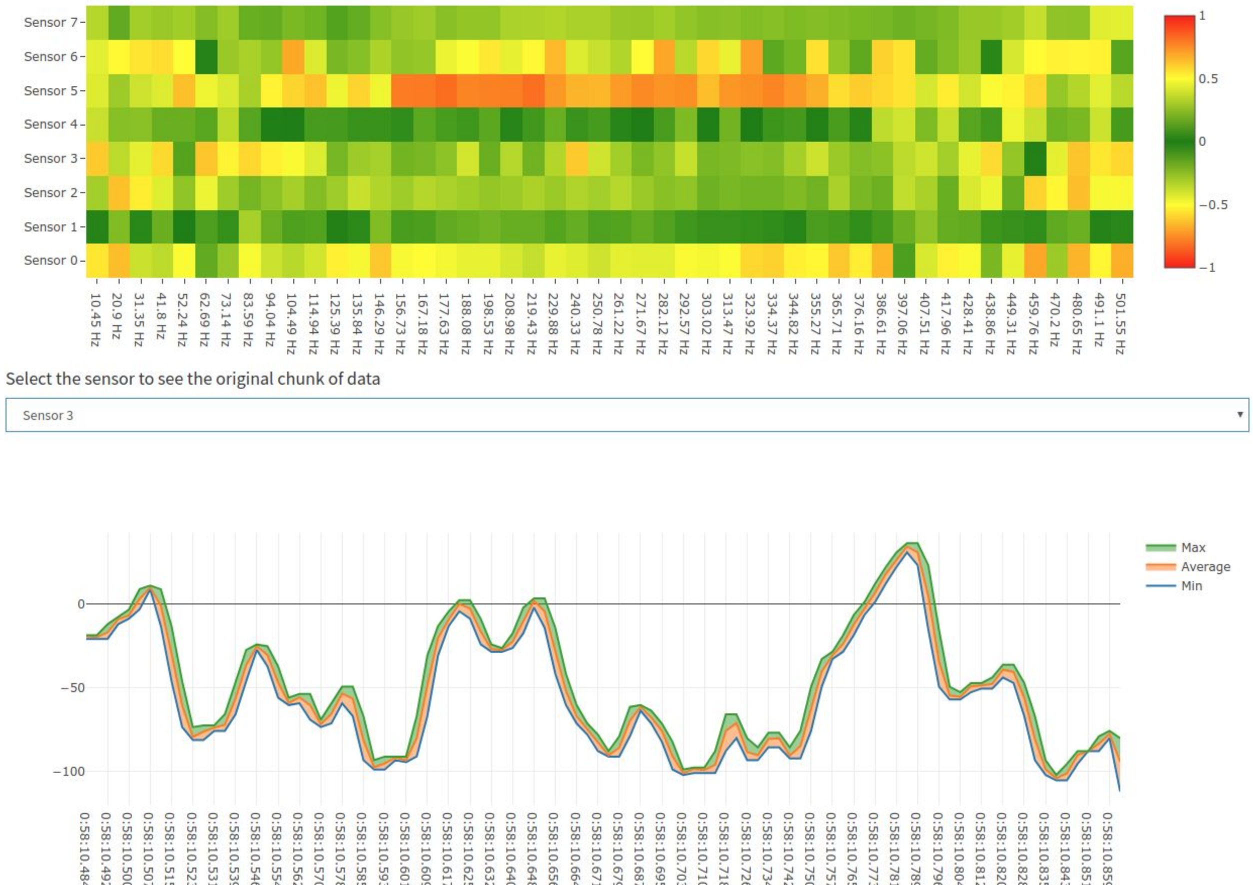


Select the sensor to see the original chunk of data

Sensor 1 ▾



Frequency Domain RCA



Screencast

At the following link you find a screencast demonstrating the use of the platform: https://vimeo.com/302251572