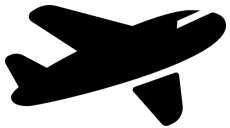
Anomaly Detection - Exercise

Algorithms in Machine Learning, ISAE-SUPAERO

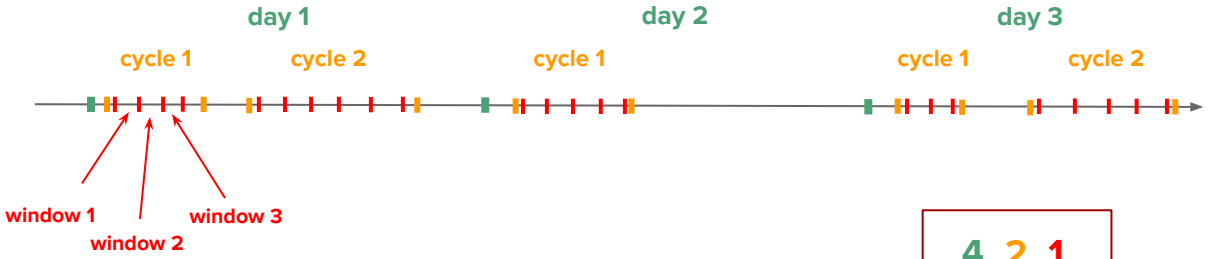
Fabrice Jimenez
Data Scientist & Artificial Intelligence Engineer
Airbus Commercial Aircraft

The dataset

Aircraft systems are recording values of parameters such as speed, temperature, pressure, electrical current values...



11 parameters recorded:
p1, p2, p3, ..., p11



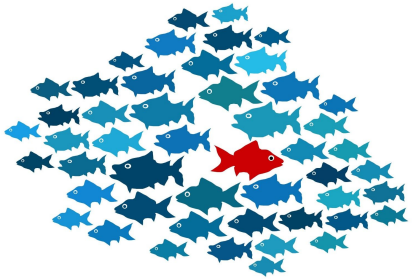
Window = section of measures, up to 100 points

1 parameter "day_cycle_window" :

4_2_1

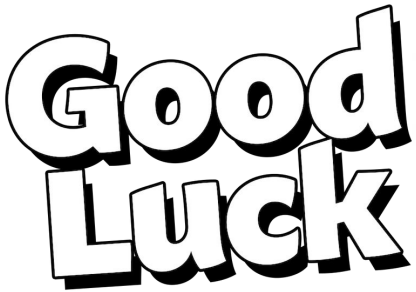
day 4, cycle 2, window 1

The question

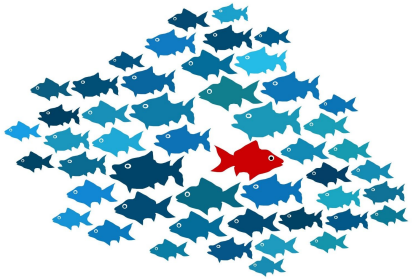


An aircraft system expert comes to see you (data scientist) with this dataset, and asks you to:
“Build an algorithm to detect windows that are abnormal.”

- 1/ With this information and no more, formulate the problem, and tell him what is feasible and what is not.
- 2/ Develop an approach to answer his question in the best way possible.
- 3/ Present your findings to the expert, in a way he can understand and help you validate your results…



Other interesting datasets



- **1/ KDD Cup 1999: network intrusion detection - famous, OK but balanced dataset**
  http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
- **2/ Large collection of datasets on different tasks and data types:**
  http://odds.cs.stonybrook.edu/
- **3/ Numenta Anomaly Benchmark - generated dataset, but used for detailed benchmark:**
  https://github.com/numenta/NAB