Increasing the Detection Performance of Automatic Safety Monitoring Tools with XGBOOST

Christian Verdonk Gallego (CRIDA)
Chen Xia (CRIDA)
Miguel García Martínez (CRIDA)
Diego Piva (CRIDA)
Javier Perez Heras (ENAIRE)

8th December 2021 – Safety Session

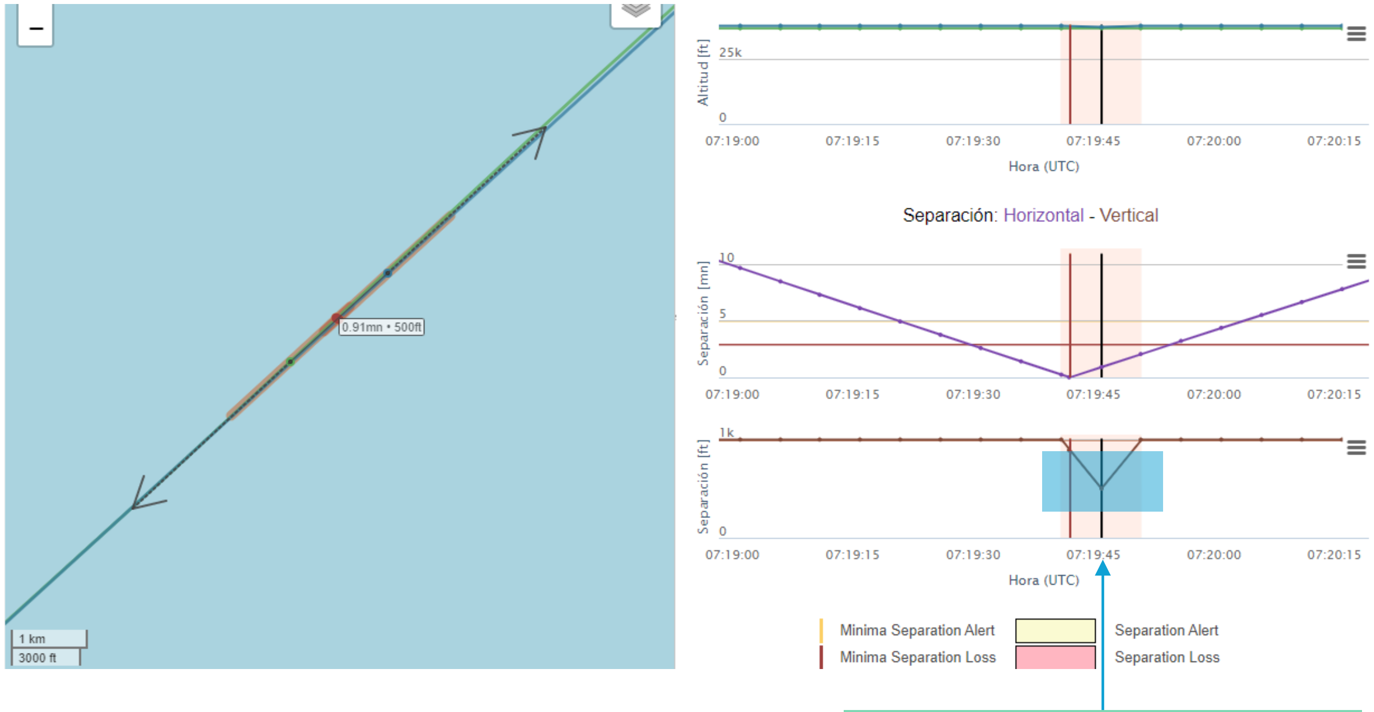
1. Introduction
2. Challenges
3. Methodology
4. Results
5. Conclusions

Introduction

- All ANSPs need to have an *Automatic Safety Monitoring Tool (ASMT)*
- The ASMT tool main function is to identify all *Separation Management Infringements (**SMIs**)*
- A SMI is defined in terms of the *horizontal* and *vertical* distances between aircraft. For en-route airspace, these are defined as **5NM** and **1,000ft** respectively.
- **ENAIRE's** ASMT is developed by CRIDA, within the framework of **PERSEO**.
- The Safety Monitoring Team of ENAIRE reported that the usability of the tool was not within the performance required, due to a high number of **False Positives.**

Introduction - II

- PERSEO's *SMI Process* computes **all** horizontal and vertical distances of aircraft pairs, and stores those that have a horizontal distance below 20NM *(an interaction)*.
- PERSEO processes all *interactions* and those that infringe the separation minima, depending on the Operating Environment, are stored as SMIs.
- Due to the precision of Mode C (100ft), there might be fluctuations in the altitude of the aircraft. This problem can also be aggravated by de-correlations, or lack of coverage in the border areas of radar coverage.
  - To overcome these, different heuristics were programmed for "filtering" those SMI that are **non-genuine.**



Jump on the Relative Vertical Distance
Non-Genuine SMI

Introduction (III)
What was our performance?

ENAIRE provided a "ground-truth", i.e. genuine SMI reported by the ATC and / or pilots

$$TPR = \frac{TP}{TP + FN}$$

$$TNR = \frac{TN}{TN + FP}$$

$$FPR = \frac{FP}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

$$F1 = 2\frac{Precision * TPR}{Precision + TPR}$$

Metric	Score
Sensitivity	96.5%
Specificity	69.5%
FP Rate	30.5%
Precision	73.1%
F1	83.2%

Some of the "genuine" SMI were hidden due to they also suffered from anomalies in the vertical profile.

Almost 1 in 3 were falsely classified as "Genuine SMI"

Although we have a "Good" sensitivity, the FPR degraded the precision!

Problem Statement and Challenges

Goal: The goal of the PERSEO ASMT is to provide to the safety practitioner a list of genuine SMIs that occur within a given period and airspace, minimising the number of non-genuine ones (i.e. minimisation of the FPR)

Scope: SMI occurred within Spanish En-Route airspace (above FL245).

Objective:

- Definition of a **classifier** which to differentiate between **genuine** and **non-genuine SMIs.**
- The model should be embedded in the ETL processes that feed CRIDA's Data warehouse.
- The model should retrofit the data stored, from **2013.** This is a **limitation**, as we cannot use new data and transformations need to be kept to a mínimum.

Process

01

Data Understanding

Taxonomy of cases

Classification in Genuine or not?

02

Data Preparation

Selection of Sample for Training and Testing

Selection of Features

03

Modelling

Algorithm selection

Hyperparameters Finding

04

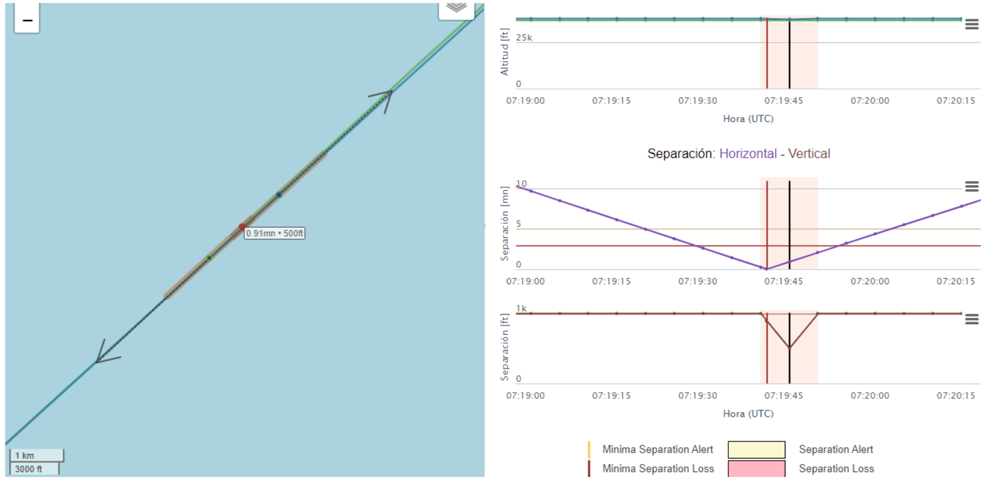
Results Evaluation

Validation of results

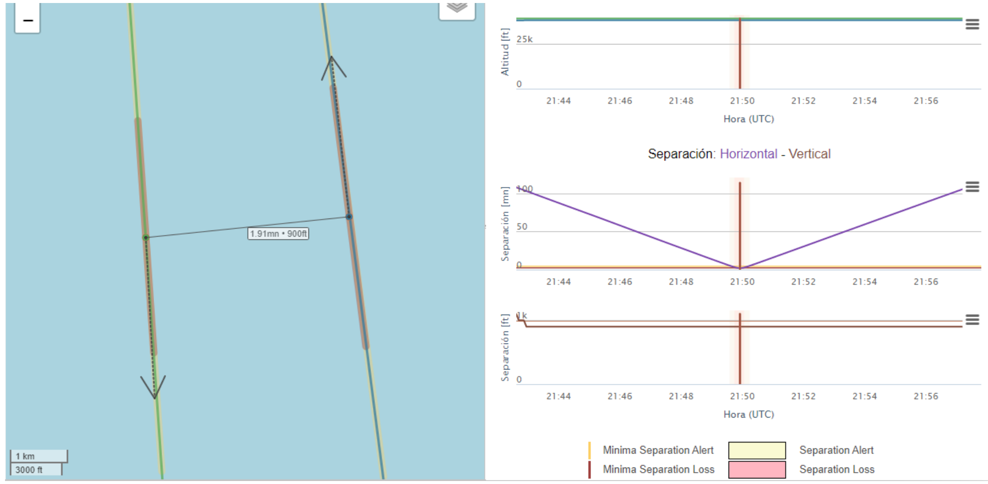
Data Understanding

Category	Definition	**Genuine?**
A	Tracks affected by fluctuations or other kind of anomalies	0
B1	Flight established at FL XX100 o XX900 ft during a short period of time: between cruising flights (including holding pattern)	0
B2	Flight established at FL XX100 o XX900 ft during a short period of time: at least one of them is climbing or descending (including holding pattern)	1
C	Flight established at XX100 o XX900 ft	0
D	At least one of the two flights is climbing or descending	1
E	Small infringement and short duration	1
M	Between military flights	Out of the scope
V	Between VFR (visual flight rule) flights.	Out of the scope
Level Bust	Level Bust	1

Case A



Case C



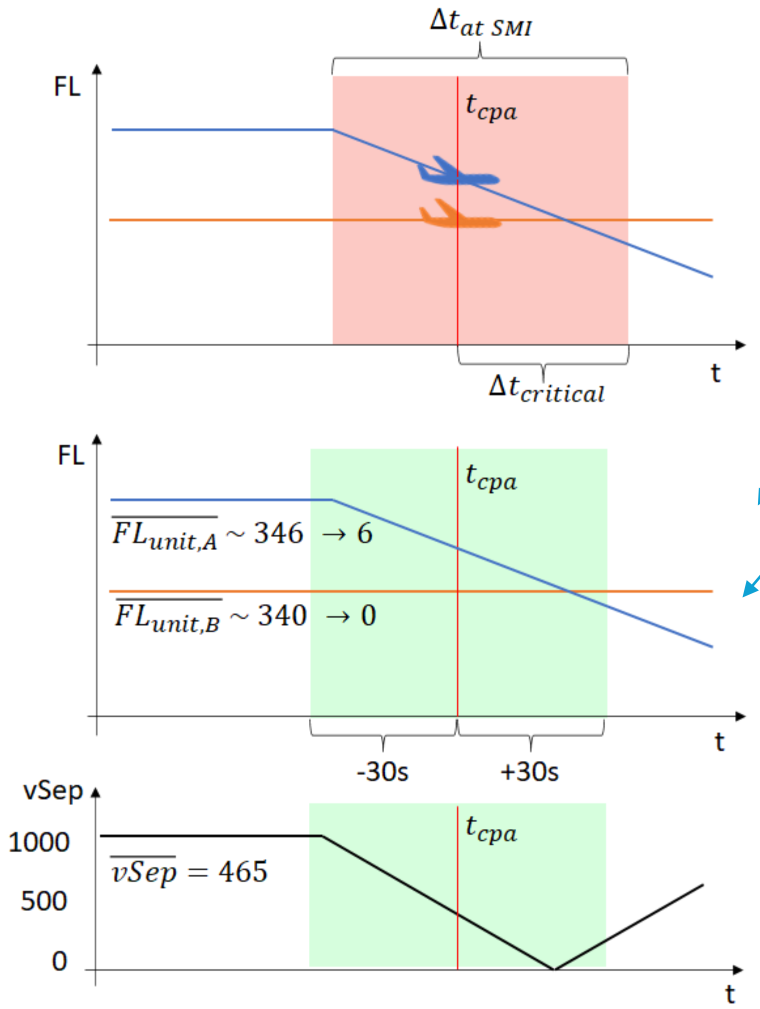
Data Preparation
Selection of Sample for Training and Testing

Sample:

- Balanced sample from 2018 and 2019
- **Genuine cases** were identified from Ground Truth, and complemented with randomly selected cases from CRIDA's DWH, after a labelling by the team
- Test sample is from **2020**
- **Features** are basic transformations from the outputs of the **PERSEO** core ASMT Process.

Parameter's name (In Spanish)	Parameter	Type	
FLTTYP1	Flight type of flight 1	1: M or S 0: rest	Initial PERSEO ASMT Core Process
FLTTYP2	Flight type of flight 2	1: M or S 0: rest	
%H (x_H)	Horizontal separation/ Horizontal Separation minima	Percentage	
%V (x_V)	Vertical separation / Vertical Separation Minima	Percentage	
Alt. (Alt_)	Altitude	Numeric values (discrete FL)	
Anomalías en traza (Anomal_asEnTraza)	Original track anomaly identified by PERSEO.	1: no anomaly 2: fluctuation (100) 3: fluctuation (500) 4: garbling	
Convergente	Convergency state when the SMI starts.	1: convergent 0: divergent	
Duración (s) (Duracion_s_)	Duration of the SMI	Numerical values	
Duración crítica (s) (Duracion_critica_s_)	Duration of from instant of min. separation until the end of the SMI	Numerical values	

Selection of Features



Mean_alt_A	**Mean of the modulus operation of the FL and 10, of Flight 1 during t_{CPA} ± 30s**	**Numerical values**	Transformed from Surveillance after the core PERSEO ASTM Process
Mean_alt_B	Mean of the modulus operation of the FL and 10 for Flight 2 during t_{CPA} ± 30s	Numerical values	
std_vSep	Standard deviation of the vertical separation during t_{CPA} ± 30s	Numerical values	
mean_vSep	Mean vertical separation during t_{CPA} ± 30s	Numerical values	
std_hSep	Standard deviation of the horizontal separation during t_{CPA} ± 30s	Numerical values	
mean_hSep	Mean horizontal separation during t_{CPA} ± 30s	Numerical values	

Modelling

Model Selection

- **XGBoost** (eXtreme Gradient Boosting)
- Selected because:
  - Reported high performance in classification tasks.
  - It could be easily fitted within our **JAVA-based ETL pipelines**
- Hyperparameter selections:
  - Using Scikit-learn package (Sk-Learn)
  - We used the GridSearchCV method (i.e. cross

Hyperparameters

Objective Function: Binary Logistic

Max_Depth: [3,5,7,**9]**

Min_child_weight [**1**,3,5]

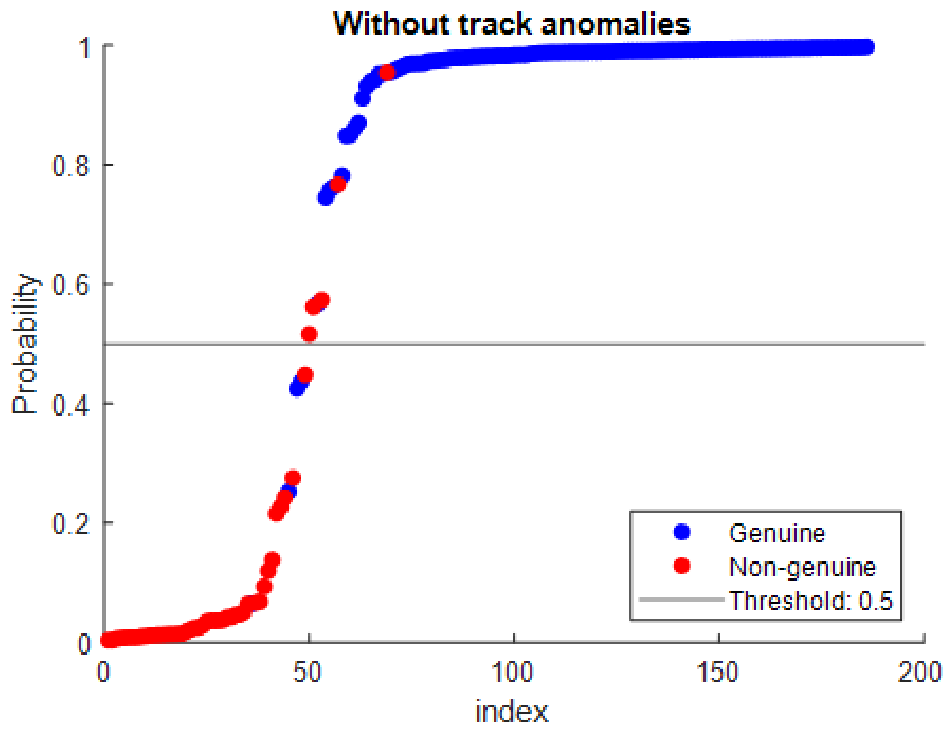
eta: [**0.05**, 0.1,0.15, 0.2]

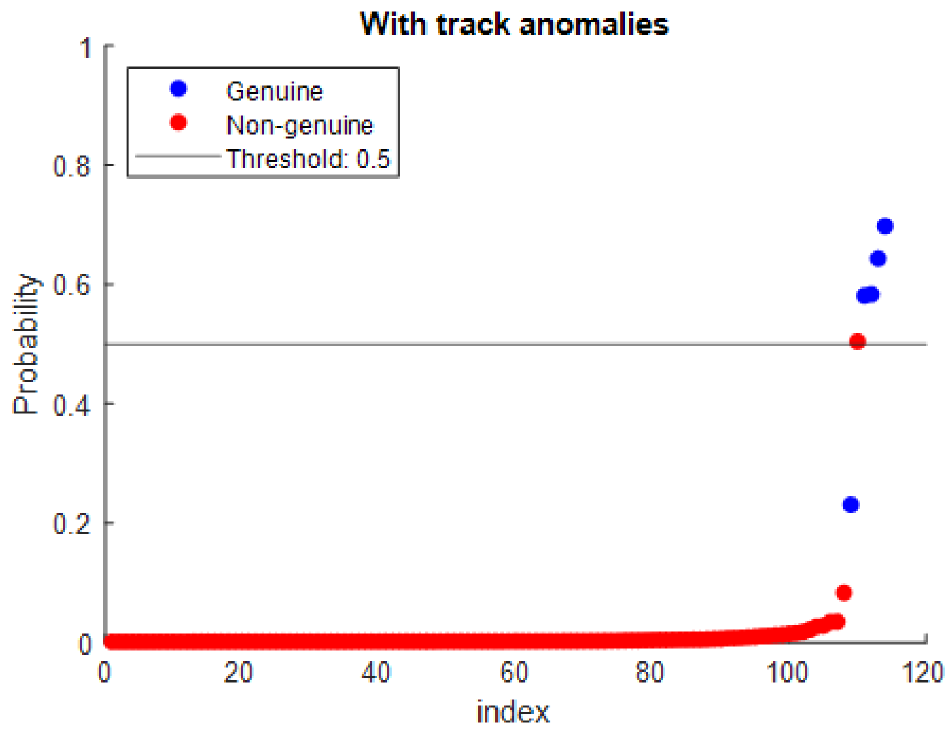
Gamma: [0.1,0.2, **0.3**, 0.4, 0.5]

Subsample [0.6, 0.7, 0.8, **0.9**,1]

Colsample_bytree = 0.5

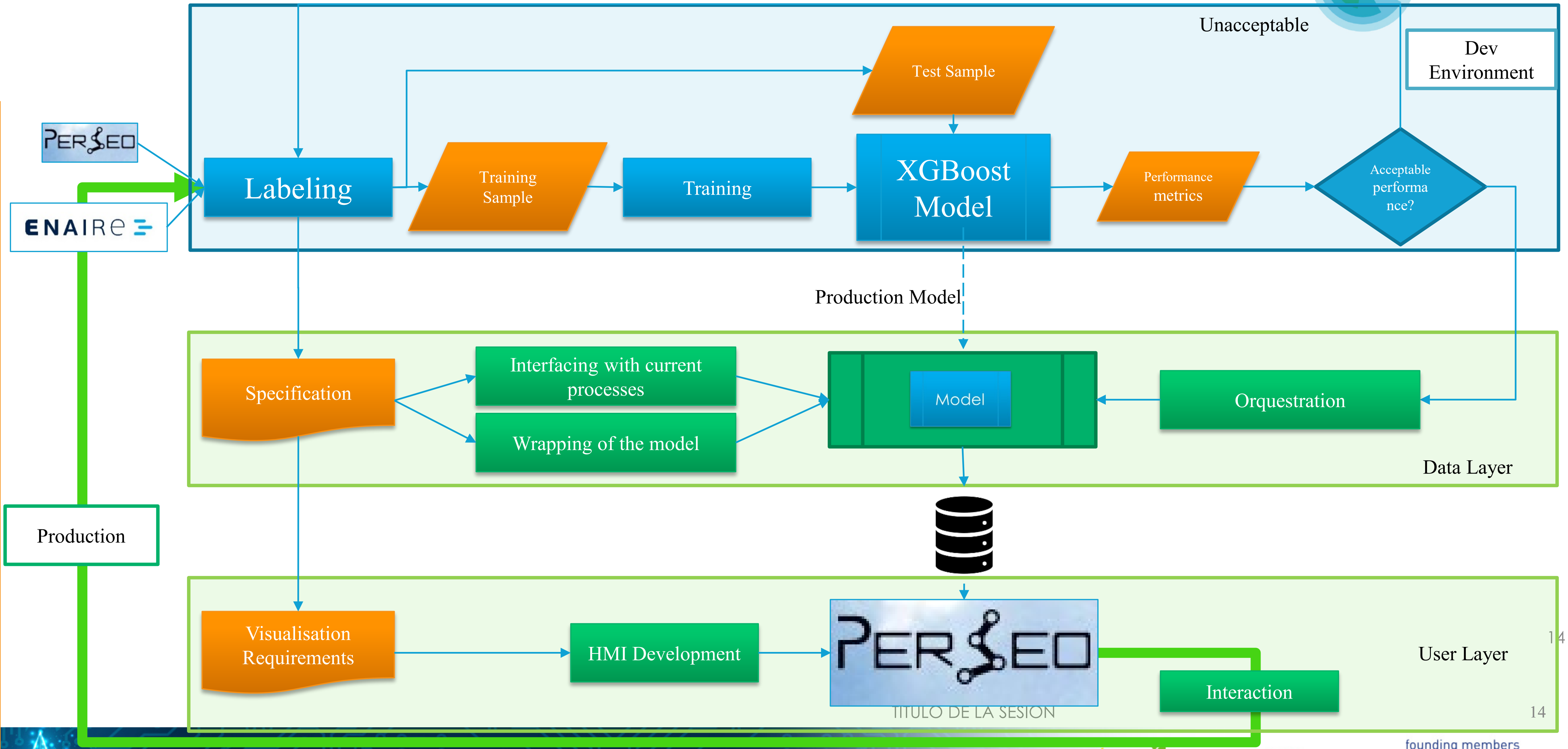
Results Evaluation





	Sensitivity	*Specificity*	*FP Rate*	*Precision*	*F1*
Filters Baseline	96.5%	69.5%	30.5%	73.1%	83.2%
XGBOOST	96.5%	96.2%	3.8%	95.8	96.1%

ETL Architecture



Conclusions

- This paper has presented the workflow followed for deploying a Machine Learning model (XGBoost) in production, also retrofitting for all the data available (since 2013) to increase the detection performance of SMIs.
- The paper presented a taxonomy of vertical anomalies in the tracks, features that were selected for the model as well as the hyperparameters selected.
- The performance of the model was verified against the baseline performance, gaining full acceptability from the Safety Monitoring Team of ENAIRE.
- The FPR decreased 27 points, meanwhile the precision increased by a 22%.
- During this year, this model has also been validated in TMA (over 10,000ft), and now is being extended to Final Approach and Wake-Vortex separations
- The full deployment of these technologies can pave the way for an effective and trustable deployment of Safety Intelligence.