Flight Data Monitoring (FDM)
Unknown Hazards detection during Approach Phase using Clustering Techniques and AutoEncoders

Antonio Fernández Llamas
Innaxis
af@innaxis.org
Introduction

- Airlines safety departments analyse Flight Data Monitoring (FDM) to inspect safety occurrences
- FDM data is known for presenting a very high variable dimensionality
- Human experts use rule-based systems based on thresholds exceedance
- Rare events or anomalies are the hardest to manually detect
- Anomaly detection using Machine Learning is definitely a challenging concept in aviation field.
Objectives

• Develop a forensic tool that semi-automatically tags abnormal flights

• Speed up the flight inspection processes, looking at operations that might hide safety implications.

• Focus on outliers detected during approach procedures

• Take advantage of the majority of the samples present in most of aviation datasets -> normal flights.

• Detect unknown hazards which might reveal new possible safety concerns
Problem assessment

- Diagnostic analysis for **flights landing on LEBL 25R**
- Descriptive and predictive analysis.
- Main outcome: predict outlier approaches using an AutoEncoder.

<table>
<thead>
<tr>
<th>Descriptive analysis</th>
<th>Predictive analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Unsupervised learning</td>
<td>- Semi-supervised learning</td>
</tr>
<tr>
<td>- Clustering techniques</td>
<td>- AutoEncoders</td>
</tr>
<tr>
<td>- Outlier detection</td>
<td>- Binary classification based on reconstruction error</td>
</tr>
</tbody>
</table>
Data sources

- **FDM, METAR** (in LEBL) and **Flight Plans**.

- The FDM dataset is composed by 35,000 landing attempts in LEBL 25R.

- FDM is known for presenting a very high variable dimensionality.

- We had more than 150 variables with a resolution of up to 8 samples per second.

- **DataBeacon** infrastructure used to process the data, using a 5-node (m5.4xlarge) cluster.
Feature engineering (I)

- Multi-step data preparation pipeline (cleaning, de-identification, merging and sampling)
- We sampled the series from 12 NM to threshold to touchdown (TD) point, with a step of 0.5 NM.
Feature engineering (II) - Pipeline

for each approach

- static info
- weather
- timepoints
- features

read
preprocess

aggregate
write

G/A?
YES
NO
Feature engineering(III)- Features

- We engineered a total of 825 features per landing attempt

<table>
<thead>
<tr>
<th>Group</th>
<th>Features</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft energy</td>
<td>Air speed. Ground speed. Energy level. Aircraft mass.</td>
<td>FDM</td>
</tr>
<tr>
<td>Aircraft configuration</td>
<td>Flaps orientation</td>
<td>FDM</td>
</tr>
<tr>
<td>Crew coordination</td>
<td>Autopilot status. Pilot flying</td>
<td>FDM</td>
</tr>
<tr>
<td>Pilot awareness</td>
<td>Current time, Distance travelled, Total Time Flown</td>
<td>FDM</td>
</tr>
</tbody>
</table>
Descriptive analysis(I) - tSNE

- Dimensionality reduction is needed -> tSNE
- It inspects the input statistical properties and manages to represent the same data using less dimensions.
Descriptive analysis(II) - HDBSCAN

- HDBSCAN is a hierarchical clustering algorithm.

- 9 clusters of data identified.

- The algorithm flagged several points as noise, that were not assigned to a cluster.
Descriptive analysis(III) - GLOSH outliers

- Anomalies are detected by measuring a local deviation of an observation with respect to its neighbours.

- Outliers can be located within a region or cluster, not only globally.
Descriptive analysis(IV) - Outliers histogram

95% quantile threshold
Outlier A

Ground speed

Vertical descent over time

CAS

Wind direction
Outlier B

Vertical descent over time

AP status
0 is manual flight
1 is captain’s AP
2 is first officer’s
3 is dual

Flaps (rad)
Predictive analysis (I) - AutoEncoder

- **Train** only using normal approaches to learn the behaviour of a non-anomalous procedure.
- The AutoEncoder will be tuned to recreate “normal” inputs minimizing the loss.
- If we test using outliers, the **reconstruction error** will be high.
- Set a **threshold** to solve a *binary classification* problem.
Predictive analysis (II) - Severity regions

Threshold 1 (THR1) = near zero
Threshold 2 (THR2) = 70% quantile
Threshold 3 (THR3) = 90% quantile
Threshold 4 (THR4) = 95% quantile

Severity regions:
Normal (OS=0) → 12,011 flights
Very Low (THR1<OS<THR2) → 12,489 flights
Low (THR2<OS<THR3) → 7,000 flights
Medium (THR3<OS<THR4) → 1,750 flights
High (OS>THR4) → 1,750 flights

*Note: OS = outlier score
### Predictive analysis (III) - Train, valid, test

<table>
<thead>
<tr>
<th>INPUT</th>
<th>lstm1:LSTM</th>
<th>(None, 1, 825)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lstm2:LSTM</td>
<td>LSTM: 256</td>
</tr>
<tr>
<td></td>
<td>lstm3:LSTM</td>
<td>LSTM: 128</td>
</tr>
<tr>
<td>ENCODER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE</td>
<td>lstm4:LSTM</td>
<td>LSTM: 64</td>
</tr>
<tr>
<td>DECODER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OUTPUT</td>
<td>lstm5:LSTM</td>
<td>LSTM: 128</td>
</tr>
<tr>
<td></td>
<td>lstm6:LSTM</td>
<td>LSTM: 256</td>
</tr>
<tr>
<td></td>
<td>lstm7:LSTM</td>
<td>(None, 1, 825)</td>
</tr>
</tbody>
</table>

- LSTM layers - useful to extract patterns from sequential data
- LSTM layer input shape is: \((n\_samples, n\_timesteps, n\_features)\)
- We trained the AutoEncoder with samples having an \(OS = 0\)
- We used a 10% of these flights as validation dataset
- The rest of the samples with an \(OS > 0\) will be our testing dataset
Predictive analysis (IV) - MSE Loss

Train loss

Threshold = 0.03

Test loss
The model struggles to classify samples with low outlier scores because of their similarity with “normal” samples. In the other hand, the model perfectly recognizes high severity outliers.
Results (II) - Confusion matrix

Confusion matrix

Predicted label
accuracy=0.8307; misclassification=0.1693
Conclusions and future work

• The results obtained are promising, providing an extremely useful tool for FDM and safety analysts.

• Overall, the combination of clustering and AutoEncoders enabled to solve a very complex problem both from the data science and aviation safety perspective.

• Benefits are not only to speed up the detection of risk events, but also to support the recognition of unknown hazards.

• In future researches, this could enable to develop an application for an airline safety department where anomalous approaches are tagged at the end of the day, to be later examined by a safety expert.
Flight Data Monitoring (FDM) Unknown Hazards detection during Approach Phase using Clustering Techniques and AutoEncoders

THANKS!

Antonio Fernández Llamas
af@innaxis.org