Aircraft data analysis and predictive maintenance: The experience of the “Arrowhead-fPVN” project

ITADATA2025: The 4th Italian Conference on Big Data and Data Science

Gianluca Boleto [1] Gianluca Sommariva [1] Luca Oneto [1] **Stefano Rovetta** [1]
Alberto Calvo [2] Luca Mortorano [2] Claudio Porretti [2] Davide Anguita [1]

[1]DIBRIS – University of Genova, Italy

[2]Leonardo S.p.A., Turin, Italy

2025-09-11

The “Arrowhead fPVM” project

- Horizon Europe project “Arrowhead Flexible Production Value Networks,” grant agreement No 101111977 of the Chips Joint Undertaking.
- The Arrowhead microservice-based platform enables interoperability and automation in complex systems (IoT, System of Systems)
- The project explores applications of this technology to **automated translation of data models** in key industrial areas
- The authors collaborate to its application in the field of aircraft predictive maintenance.
- Project website: `https://fpvn.arrowhead.eu/fpvn-arrowhead/`

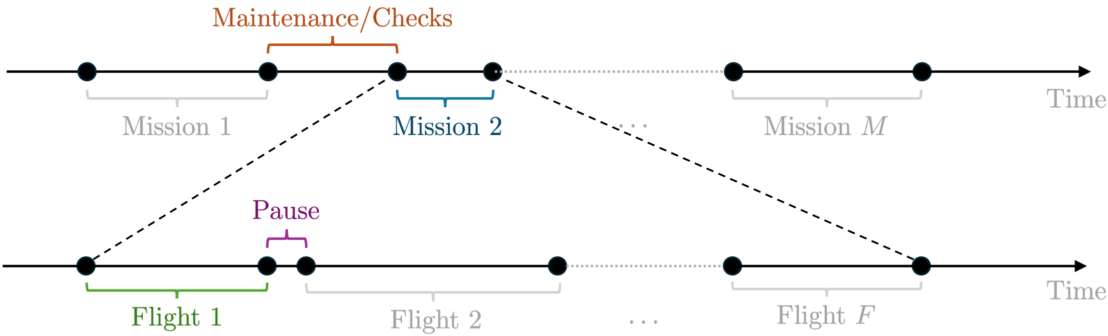
Motivation of the work

- In aeronautics, predictive maintenance plays a crucial role
- Aircraft continuously generate data:
  - various metrics (e.g. flight parameters, system status)
  - diagnostic messages (e.g. fault codes)

  Sometimes maintenance interventions are needed.
- In this project we compare
  - traditional feature engineering based on domain expertise

    vs
  - learned representations

The goal



- Problem 1: **Should a given component be removed/replaced?** (“removal prediction”, a classification task)
- Problem 2: **How long will a given component last before needing replacement?** (“remaining useful life”, a regression task)

The Role of Data

- Several approaches to prediction using explicit simulation…
- However, complex systems are difficult to model in closed form!
  $\Rightarrow$ **Machine learning** is used to provide predictions
- The project benefits of on-board data collection facilities
- However, raw data are seldom ready for providing information, unless they are trivially related, for instance, to specific failures.
- We explore how data representation techniques impact machine-learning based predictive performance.
- Experiments on real-world data from Leonardo S.p.A.

The data

- Flight data from
  - 26 parameters
  - from 22 aircraft
  - monitored over an eight-year period
  - several thousands of flights
  - more than 35,000 measurements of aircraft parameters
- In total, tens of millions of measurements were gathered, covering both in-flight and on-ground operations.

Input/Output	Type	Description	When Collected
Inputs (Predictors)	Scalars	Start date and start time	Flight
		In-flight flag	
		Last flight date of the same aircraft	
		Days from last removal of the same component	
	D-Series	Parameters measurement: 26 different lists, one for each parameter.	Flight
		Ground configuration records: 124 different lists, one for each configuration parameter.	Pause
		Flight configuration records: 253 different lists, one for each configuration parameter.	Flight
	S-Series	Faults triggered: more than 3000 lists, one for each different type of fault that could be activated during flights.	Flight
Outputs (to be Predicted)	Binary	Removal	Maintenance
	Days	Remaining Useful Life	Maintenance

Methodology

Comparing:

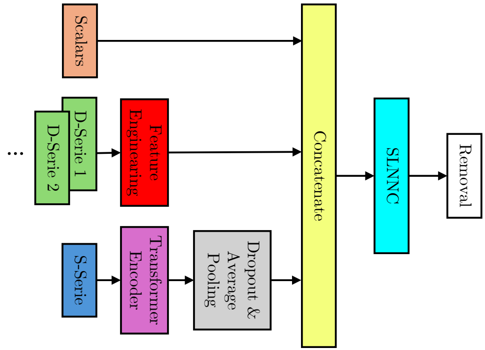
- SHALLOW MODELS

— Require feature engineering

\+ Less demanding in terms of training data

Random Forests, linear / Gaussian Support Vector Machines, XGBoost, Gaussian Processes, K-Nearest Neighbors

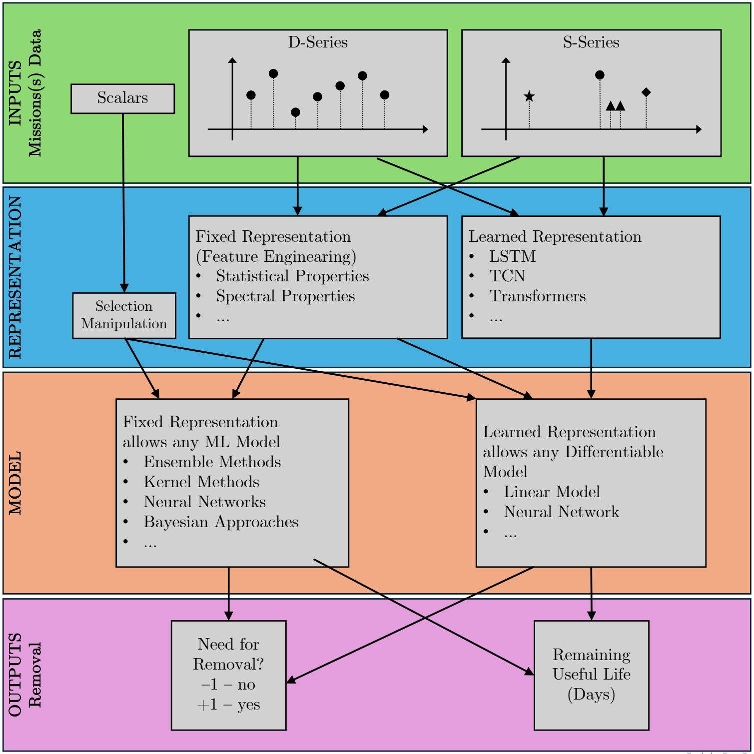
- DEEP MODELS



— Require lots of data and (exponentially) as much training time

\+ Learn data representations without the need of an expert

Comparing workflows



Some findings

- XGBoost is the best-performing shallow model
- Among the deep models, some achieve a better performance (i.e., need to select a suitable model)
- A combined feature engineering+learning strategy improves over both shallow and deep baselines
- Recurrent network models do not necessarily improve over the baseline, while the Transformer does so significantly
- Predicting removals turns out to be a very challenging task, with a high number of false negatives
- Remaining useful life prediction accuracies are very high

The end – Thanks