Machine Learning of Engine Health Monitoring Data: Development of a machine learning model for damage prediction of real flight missions

Technische Universität Berlin

Faklutät V für Verkehrs- und Maschinensysteme
im Fach Festigkeit und Lebensdauer

Bachelorarbeit

zur Erlangung des akademischen Grades
Bachelor of Science (B.Sc.)
im Studiengang Physikalische Ingenieurwissenschaft

Verfasser	Samuel Oliver Bryson
Matrikelnr.	352667
Ausgabedatum	29.01.2020
Eingereicht am	XX.XX.2020
Betreuung	Prof. Dr.-Ing. Robert Liebich
	Dr. Stephan Pannier
	M.Eng. Tolga Yağcı
	Tobias Werder, M.Sc.

Eidesstattliche Erklärung

Hiermit erkläre ich, Samuel Oliver Bryson (geboren am 21.12.1991 in Nottingham, Großbritannien), dass ich die vorliegende Abschlussarbeit eigenständig und ohne unerlaubte fremde Hilfe angefertigt habe sowie keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Berlin, den XX.XX.XXX   _____

                                                          Samuel Oliver Bryson

List of Abbreviations

AI Artificial intelligence

DL Deep learning

EHM Engine Health Monitoring

FE Finite Element

HP high-pressure

HPT high-pressure turbine

IP intermediate-pressure

LP low-pressure

ML machine learning

TET turbine entry temperature

1 Introduction

In recent years, many industries have experienced enormous changes due to an influx of enormous quantities of data, a greater availability of data storage and high computational power, and an ever-increasing armoury of machine learning algorithms and methods. The aviation industry is no exception to this: the Airbus A350 XWB, for example, comes equipped with approximately 6 000 sensors that produce 300 GB of data every day; the next generation Airbus A380 will come with 10 000 sensors on a single wing alone (Rajaraman, 2016).

Modern airliners are reliant on a constant stream of Engine Health Monitoring (EHM) data for monitoring the status of the aircraft during the flight. A post-flight review of the data and a subsequent adjustment of flight manoeuvres can offer significant improvements in fuel efficiency and minimise component fatigue. As an optional aftermarket service, Rolls-Royce offers Efficiency Management and Efficiency Optimisation, which use EHM data with "cutting-edge technology such as machine learning and applied analytics" to optimise customers' use of their engines (Rolls-Royce plc, 2020).

The ability to process this data in a reliable and efficient way is paramount for carrying out these analyses. Using highly optimised, specially designed programming libraries, data-orientated machine learning approaches are more accessible and more powerful than ever, and can produce excellent results in detecting patterns in data and evaluating these accordingly.

1.1 The Engine

The Civil Aerospace department of Rolls-Royce designs and manufactures primarily high-bypass turbofan jet engines, which offer an ideal arrangement for civil aircraft flying below the speed of sound (Rolls-Royce plc, 2015).

Engines consist of four main components (fan, compressor, combustion chamber and turbine) which correspond approximately to the four stages of a Brayton cycle (intake, compression, combustion, expansion). The engine draws in air, compresses it, burns fuel in the compressed air and forces this air out through the rear nozzle, while extracting some energy at the turbine stage to continue powering the fan and compressor (Rolls-Royce plc, 2015).

To increase thrust and efficiency, many modern civil aircraft engines are equipped
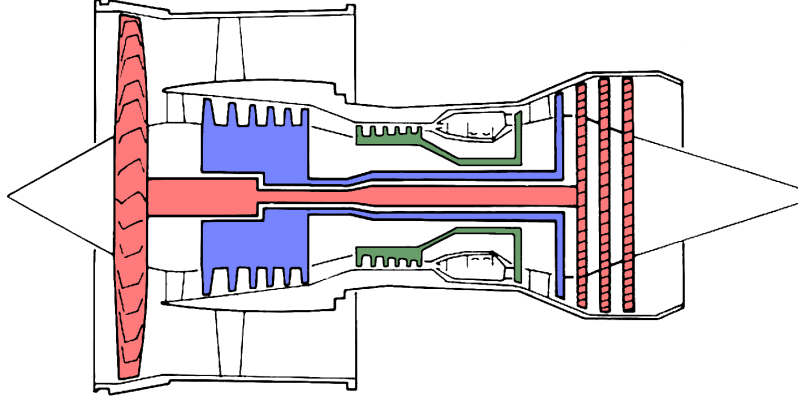
Figure 1: A schematic diagram of a triple-spool high-bypass engine with LP, IP and HP components shown in red, blue and green, respectively (based on (Rolls-Royce plc, 2015)).

with two concentric shafts, split into low-pressure (LP) and high-pressure (HP), which connect respective sets of compressors and turbines (Spittle, 2003). Some, including the Rolls-Royce Trent family of engines, also have a third shaft, coupling the intermediate-pressure (IP) compressor and turbine. Figure 1 shows a rough configuration of such an engine.

While this arrangement has greatly improved efficiency, it comes with the compromise of significantly increased temperatures within the engine, particularly as the air passes from the combustion chamber to the high-pressure turbine (HPT). The turbine entry temperature (TET) has risen to such an extent that HPT blades now operate in temperatures far above their melting point, requiring great improvements in the materials, structures and systems used, such as extracting cooler air from the compressor stage and releasing this through air holes in the blades to create an insulating barrier between the blade and hot air (Spittle, 2003).

1.2   Damage

Components used in such extreme conditions experience degradation. They therefore cannot be used indefinitely and must be removed from service after a certain amount of time to avoid potentially catastrophic failure. The amount of time for which the component is permitted for service, referred to as its Approved Life, is determined in safety analyses (EASA, 2015). Approved Life is measured in Engine Flight Cycles, to be referred to as *cycles* in the following. (These aspects will be discussed in more

detail in Section **??**.)

The degradation of the engine through its use is referred to as damage, whose unit is also cycles. The extent of damage is dependent on many parameters: Since operators use their aircraft for different purposes, flight parameters such as duration and altitude vary greatly and cause different levels of damage.

In a Finite Element (FE) context, components are modelled digitally and separated into a finite number of individual elements, the corner points of which are called nodes. Areas of particular interest within the component (usually where stresses are expected to be highest) are referred to as features.

Within Rolls-Royce, damage from flight missions is currently calculated using one of two internal tools for flight profile analysis: SA66 and Perseus (see Sections **??** and **??**, respectively). The former is ideal for processing many flights in a short amount of time, but is restricted to a low number of features due to the time involved in manually setting up the surrogate FE model for each feature. The latter can be described as a brute-force method that determines damage for all surface nodes, but is restricted by the amount of time required to process a single flight.

1.3   Artificial Intelligence

Artificial intelligence (AI), the field of research that occupies itself with giving machines the ability to think and learn, has been the focus of much research in recent years due to the increasing capabilities of hardware to tackle the challenges it involves, as well as its profitability in industries looking to implement automation of routine labour or optimise data analysis tasks (Goodfellow, Bengio, & Courville, 2016).

One subfield of AI is machine learning (ML), in which machines extract information and patterns from data without thorough or explicit instructions, usually making use of highly contrived data from which noise (non-critical background information) has been removed. Classifying objects based on a finite set of input values (e.g. birds based on their weight, wingspan, the colour of their back and whether they have webbed feet (Harrington, 2012)) is an ideal task for machine learning.

ML tasks are generally split into two categories: supervised and unsupervised learning (Kelleher, Namee, & D'Arcy, 2015). The former involves training the model to associate input data with known output data and using it on previously unseen data;

in the latter, a model is given data and instructed to find patterns or groups based on input data alone (Goodfellow et al., 2016).

Deep learning (DL) is an application of ML that employs more complex models capable of making sense of noisier, largely unprocessed data, such as sound signals and images (Goodfellow et al., 2016). Classifying birds could in this case involve extracting necessary information from an image or diagram of the bird, requiring far less manual measurement or input.

DL has seen a huge rise in popularity in recent years, with applications including speech recognition (Deng & Li, 2013), medical diagnoses (Lee, Park, Joo, & Moon, 2018), stock market prediction (Krollner, Vanstone, & Finnie, 2010) and many others (Kelleher et al., 2015). One particular challenge among the DL research community is time series data (Yang & Wu, 2006), a sequential collection of values recorded over time. Time series data remains a great challenge due to its noisy, multidimensional nature (Kelleher et al., 2015) and the dfficulties involved in developing algorithms that can also interpret the temporal information held in the signal (Bagnall, Lines, Bostrom, Large, & Keogh, 2017).

1.4 Motivation

It is in the interest of Rolls-Royce to improve the speed and accuracy of EHM data analysis to gain an overview of the performance and remaining Approved Life of in-service engines.

The goal of the present thesis is therefore to identify a supervised machine learning approach that offers a robust, verifiable, comprehensive, fast and sufficiently accurate means of processing EHM data to determine the extent of damage incurred by surface nodes of a component during real flight missions.

2 Data

The following will include an overview and visualisation of the data acquired for the research, and how it was processed.

2.1 EHM Data

The majority of data used for the research is EHM data. This data was recorded by sensors in 231 BR725 engines during a total of 14 045 flights, and returned on a voluntary basis to Rolls-Royce by customers for analysis. Each aircraft using the BR725 has two engines; to minimise the amount of data used, the values used in this thesis were taken from the left engine only.

2.2 Overview

The flights range in length from 1 013 to 57 062 seconds (approximately 0.28 to 15.85 hours), with a mean length of 10 182,82 seconds and a standard deviation of 9 561,03 seconds (see Figure 2).

Each flight is represented by a CSV file with 216 columns, comprising one timestamp and 215 values per second of recording time.

The four flight parameters extracted were altitude (ALT), rotational speed of the high-pressure shaft (NH), and pressure and temperature of air exiting the compressor (P30 and T30, respectively). These are shown for one typical flight in Figure 3.
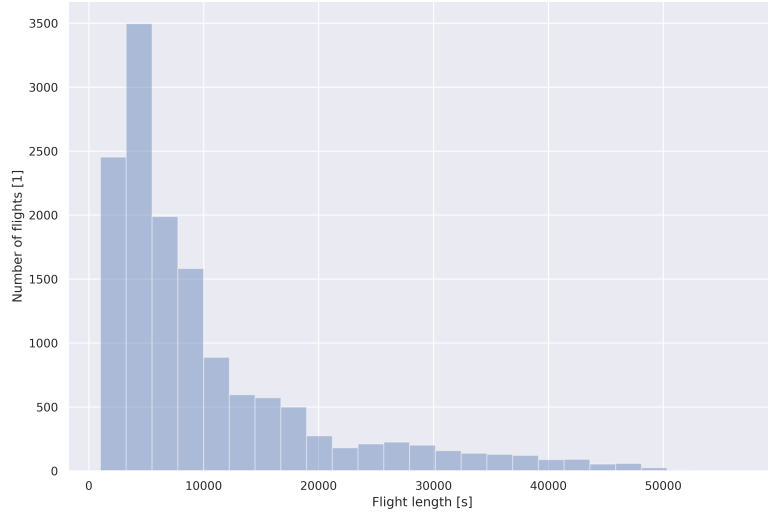
2.3 Flight Phases

Each flight can be split into several phases: preflight, taxi out, take-off, climb, cruise, descent, reverse thrust and taxi in.
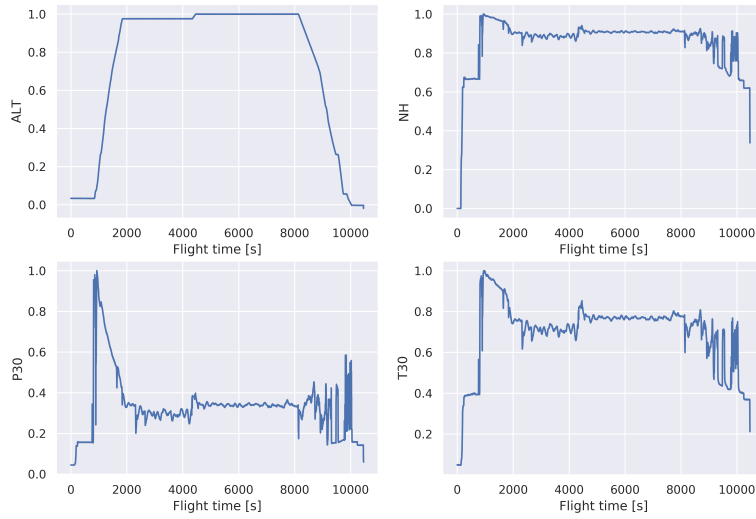
2.4 Output

The flights were processed using SA66 (Section **??**) to determine the number of cycles consumed

Figure 2: A histogram of the lengths of 14045 flights

Figure 3: ALT, NH, P30 and T30 of a randomly selected flight. All parameter values are normalised.

References

Bagnall, A., Lines, J., Bostrom, A., Large, J., & Keogh, E. (2017, May). The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, *31*(3), 606–660. Retrieved 2020-03-26, from `http://link.springer.com/10.1007/s10618-016-0483-9` doi: 10.1007/s10618-016-0483-9

Deng, L., & Li, X. (2013, May). Machine Learning Paradigms for Speech Recognition: An Overview. *IEEE Transactions on Audio, Speech, and Language Processing*, *21*(5), 1060–1089. (Conference Name: IEEE Transactions on Audio, Speech, and Language Processing) doi: 10.1109/TASL.2013.2244083

EASA. (2015). *Certification Specifications for Engines (CS-E).* European Union Aviation Safety Agency. Retrieved 2020-03-18, from `https://www.easa.europa.eu/document-library/certification-specifications/cs-e-initial-issue` (Library Catalog: www.easa.europa.eu)

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press. (Google-Books-ID: omivDQAAQBAJ)
        \url{http://www.deeplearningbook.org}

Harrington, P. (2012). *Machine learning in action.* Shelter Island, N.Y: Manning Publications Co. (OCLC: ocn746834657)

Kelleher, J. D., Namee, B. M., & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies.* MIT Press. (Google-Books-ID: uZxOCgAAQBAJ)

Krollner, B., Vanstone, B., & Finnie, G. (2010, January). Financial Time Series Forecasting with Machine Learning Techniques: A Survey..

Lee, W., Park, S., Joo, W., & Moon, I.-C. (2018, November). Diagnosis Prediction via Medical Context Attention Networks Using Deep Generative Modeling. In *2018 IEEE International Conference on Data Mining (ICDM)* (pp. 1104–1109). (ISSN: 1550-4786) doi: 10.1109/ICDM.2018.00143

Rajaraman, V. (2016, August). Big data analytics. *Resonance*, *21*(8), 695–716. Retrieved 2020-03-18, from `https://doi.org/10.1007/s12045-016-0376-7` doi: 10.1007/s12045-016-0376-7

Rolls-Royce plc. (2015). *The Jet Engine* (5th edition ed.). Chichester, West Sussex: Wiley-Blackwell.

Rolls-Royce plc. (2020). *Airlines CareServices.* Retrieved 2020-03-18, from `https://www.rolls-royce.com/products-and-services/civil-aerospace/aftermarket-services/airlines.aspx` (Library Catalog: www.rolls-royce.com)

Spittle, P. (2003, November). Gas turbine technology. *Physics Education*, *38*(6), 504–511. Retrieved 2020-03-18, from `http://stacks.iop.org/0031-9120/38/i=6/a=002?key=crossref.5d4ad51e02e416f47faed1970f94d375` doi: 10.1088/0031-9120/38/6/002

Yang, Q., & Wu, X. (2006, December). 10 CHALLENGING PROBLEMS IN DATA MINING RESEARCH. *International Journal of Information Technology & Decision Making*, *05*(04), 597–604. Retrieved 2020-03-26, from `https://www.worldscientific.com/doi/abs/10.1142/S0219622006002258` doi: 10.1142/S0219622006002258