Aircraft Maintenance Prediction Tree Algorithms

A

Dima Alberg
Shamoon College of Engineering, Israel

Yossi Hadad
Shamoon College of Engineering, Israel

INTRODUCTION

Maintenance of aircraft components have always been an important consideration in aviation. Accurate prediction of possible failures will increase the reliability of aircraft components and systems and decrease maintenance of big future failures. The scheduling of maintenance operations help determine the overall maintenance and overhaul costs of aircraft components. Maintenance costs constitute a significant portion of the total operating expenditure of aircraft systems (Kadir, Onur, & Harun, 2020).

According to Fan (2015) there are three main types of maintenance for equipment: corrective maintenance, preventive maintenance, and predictive maintenance. Corrective maintenance helps manage repair actions and unscheduled fault events, such as equipment and machine failures. When aircraft equipment fails while it is in use, it is repaired or replaced. Preventive maintenance can reduce the need for unplanned repair operations. It is implemented by periodic maintenance to avoid equipment failures or machinery breakdowns. Tasks for this type of maintenance are planned to prevent unexpected downtime and breakdown events that would lead to repair operations. Predictive maintenance, as the name suggests, uses some parameters which are measured while the equipment is in operation to guess when failures might happen. It intends to interfere with the system before faults occur and help reduce the number of unexpected failures by providing the maintenance personnel with more reliable scheduling options for preventive maintenance. Assessing system reliability is important to choose the right maintenance strategy.

The operation and maintenance of modern predictive sensor-equipped systems such as aircraft generates vast amounts of numerical and symbolic data streams. According to Tawaikuli et al. (2020) m Multi Sensor Data Fusion oulti sensor data streams collection and preparation is an expensive, resource consuming and complex phase often performed centrally on raw data for a specific application. These data streams are generated by thousands of sensors installed in various components of the aircraft and then sent in real-time to relational databases storages in ground stations. Before being transmitted to the ground, a number of on-board computer systems monitor and analyze the data stream in order to make sure that various systems of the aircraft are operating properly. However, once the data stream is stored in central databases, further data analysis is rarely performed. This paper presents an algorithm that makes use of this data stream in order to develop interval Machine Learning ML regression tree models to predict the need for replacement of various aircraft components before they become non-operational. The end goal is to implement this ML model in a flight monitoring system that will receive the real-time multi-sensor data input from aircraft fleet, analyze it, and output alerts in the form of appropriate replacement rules when there is a need for component maintenance.

DOI: 10.4018/978-1-7998-9220-5.ch120

The monitoring system will use the automatically generated multi-sensor data stream from the aircraft, and induce an interval regression tree model described in this paper, to detect component problems and recommend their replacement. Such a system could help improve the airline's operation by: reducing the number of delays, reducing maintenance costs, helping achieve better maintenance planning, and increasing the level of safety. The approach proposed in this paper applies techniques from the fields of machine learning on big amounts of complex historical data in order to develop the predictive models required by the monitoring system. The approach described addresses four fundamental difficulties with existing data mining approaches: automatic selection of relevant data, automatic labeling of instances, an evaluation method that accounts for dependencies between the instances, and a scoring function measuring the extent to which the results fit the domain requirements. By addressing these four issues, we believe that the proposed approach will help extend the range of potential applications for ML techniques. Examples of other applications that can benefit from the approach developed in this paper are: prediction of problems in complex systems (e.g.: trucks, ships, trains, and cars), prediction of problems with complex industrial equipment for which a lot of data is continuously acquired, and prediction of critical events in medical applications (e.g. Emergency Room care). The fact that the proposed approach relies on a minimal amount of domain specific information will also facilitate the adaptation to other applications.

The paper is organized as follows: The related work is described in the next section. In Section 3 the INGPRET algorithm is introduced and its computational complexity and performance metrics are analyzed. In Section 4 some experimental results are reported for two real-world temporal data sets. Section 5 discusses the main features of the proposed algorithm and identifies some future research directions.

BACKGROUND

Actually, most of the regression tree algorithms apply binary recursive partitioning, since each node is always split into two child nodes, and are recursive, because the process is repeated at every node. It is also possible to split the data into three or more subsets or child nodes. Regression trees provide quite simple and easily interpreted regression models with reasonable accuracy. However, according to Breiman et al. (1984), these methods are known for their split instability. Finally, the interested reader may find a more detailed survey of regression tree methods in (Alberg, Last, & Kandel, 2012).

XGBoost stands for eXtreme Gradient Boosting and is a scalable implementation of gradient boosting regression trees (Chen & Guestrin, 2016). Since its release in 2016, XGBoost has been a very popular machine learning method, and it has a highly impressive winning record when it comes to machine learning competitions. XGBoost has already been used in several aviation risk assessment applications (Zhang & Mahadevan, 2019), (Shen & Wei, 2020) and represents a boosting ensemble of regression and decision trees. It is worth noting that XGBoost performance is not affected by multicollinearity (highly correlated explanatory variables), which is often highly present in multi-sensor data.

Interval prediction is an important part of the forecasting process and is intended to enhance the accuracy of point estimation. An interval forecast usually consists of upper and lower limits between which the future value is expected to lie with a prescribed probability. The limits are sometimes called forecast limits (Wei, 2006) or prediction bounds (Brockwell & Davis, 1991), while the interval is sometimes called a confidence interval (Granger & Newbold, 1986) or a forecast region (Hyndman, 1995). We prefer the more widely-used term "prediction interval," as used by (Chatfield, 2001) [REMOVED TA FIELD] and (Harvey, 1989), both because it is more descriptive and because the term "confidence interval" is usually applied to interval estimates for fixed but unknown parameters. In contrast, a predic-