Reliability Analysis Using Deep Learning and Semi-supervised Techniques

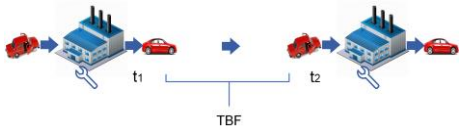
Chong Chen and Ying Liu
Institute of Mechanical and Manufacturing Engineering,
School of Engineering, Cardiff University

Abstract

Over the last decades, reliability prediction has gained much more attention. The maintenance of automobiles can be carried out before the critical failures occur if the time between failure(TBF) can be predicted. The purpose of this research is to study the TBF through a data-driven approach. Deep learning, as a tool capable of capturing the highly complex and non-linearly patterns, has not been properly introduced into reliability analysis. Furthermore, as the large sum amount of the real-world labelled data is hard to obtain, the semi-supervised learning approach has also been explored so that eventually an integrated approach for reliability analysis is targeted in this study.

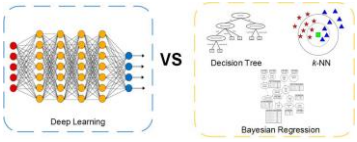
Research Motivation

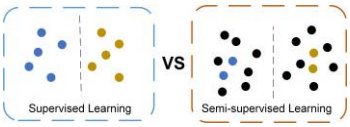
- Reliability has to be maintained to avoid failures. If the failure time of a product can be forecasted, maintenance can be scheduled and performed in advance to avoid the critical failure and further loss.
- Reliability analysis is an essential part of operation and inventory management with any maintenance department in a commercial setup.
- Reliability analysis has been widely performed using both statistical and machine learning techniques. Deep learning, as a group of emerging powerful machine learning algorithms, has been gaining much momentum in different fields reflecting its superior capability in dealing with the big data. It is also applicable in reliability analysis.
- In the meantime, to harness the merits of deep learning, it requires a large sum of real-world data, because the models trained using synthetic data or simulated data cannot always demonstrate a decent generalised ability during validation.



Research Questions

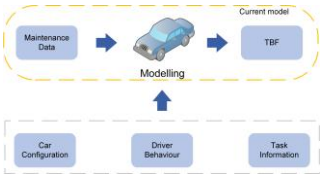
1. Is the performance of deep learning generally better than the traditional machine learning algorithms in reliability analysis?





2. When the labelled training data is limited, how would a semi-supervise learning method be engaged so that the performance of a deep learning approach won't be necessarily jeopardized?

3. With a joint deep learning and semi-sup approach finalized, how would big data analytics, concerning human behaviour, vehicle design, operational task, etc., impact reliability modelling and predictive maintenance?



Work to Date

- A large quantity of real-world automobile engine maintenance data (more than 3 million entries) has been collected.
- Data pre-processing and feature selection has been carried out with insights gained concerning some of the most effective ones in modelling.
- The cleaned data have been explored in building an initial deep learning model based on deep neural network (DNN).
- A number of baseline models have also been developed using classic machine learning algorithms, e.g., *k*-nearest neighbours (*k*-NN), decision tree (DT), and Bayesian regression (BR).
- So far, experimental studies have shown the merits of the developed DNN model in terms of both model correlation coefficient and Root-Mean-Square-Error, while further exploration and validation are needed to turn it into a useful tool in reliability analysis.

The maintenance data used in the modelling

Feature	Description
nRepair	The times of engine experienced maintenance
Vage	The age of the vehicle
AcmM	The cumulative miles when the failure occurs
Maint	The failure type
Model	The model of the vehicle
Model_Year	The year of the first production
VAGC	The type of vehicle

The results of the modelling

Metrics	DNN	*k*-NN	DT	BR
Model correlation coefficient	**0.864**	0.654	0.738	0.772
Root-Mean-Square-Error (Days)	**316**	506	462	424
maximum error (Days)	**2020**	2847	2655	2352

Conclusions and Future Work

1. TBF prediction plays an important role in reliability analysis. With the accurate prediction of TBF, the occurrence of critical failure can be possibly avoided.
2. Reliability analysis has been widely implemented using statistical approaches. However, machine learning techniques are also useful in this field, especially when a sizable dataset is.
3. Deep learning, as a useful bioavailable data analytic tool, shows merits in this initial case compared to the classic machine learning algorithms.
4. More data concerning vehicle design, human driving behaviour, operational task, and so on can inject a positive impact on TBF modelling. The exact effect awaits further exploration.
5. While deep learning is data-hungry and a large quantity of labelled real-world data is hard to be obtained, semi-supervised is deemed a viable possibility yet to be explored.
6. An advanced feature selection algorithm, which can identify the essential features effectively and efficiently, will be designed.