

An approach to improve asset maintenance and management priorities using machine learning techniques

Anu Hanish Nithin, Michael Hobbs, Srinivas Sriramula & Yaji Sripada

To cite this article: Anu Hanish Nithin, Michael Hobbs, Srinivas Sriramula & Yaji Sripada (2022) An approach to improve asset maintenance and management priorities using machine learning techniques, Safety and Reliability, 41:3, 151-169, DOI: [10.1080/09617353.2022.2142011](https://doi.org/10.1080/09617353.2022.2142011)

To link to this article: <https://doi.org/10.1080/09617353.2022.2142011>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 23 Nov 2022.



Submit your article to this journal [↗](#)



Article views: 667



View related articles [↗](#)



View Crossmark data [↗](#)



ARTICLE

OPEN ACCESS



An approach to improve asset maintenance and management priorities using machine learning techniques

Anu Hanish Nithin^{a,b}, Michael Hobbs^b, Srinivas Sriramula^a and Yaji Sripada^c

^aSchool of Engineering, University of Aberdeen, Aberdeen, UK; ^bOptimal (Optima Asset Maintenance Solutions Ltd.), Aberdeen, UK; ^cSchool of Natural and Computing Sciences, University of Aberdeen, Aberdeen, UK

ABSTRACT

Numerous services are available today to develop an optimised asset management solution to enhance asset operations by improving the system availability, decreasing down-time and operation and maintenance costs. Three cases of engineering problems are explored in this paper, with data-driven machine learning solutions proposed for these problems. The first case refers to the labour-intensive nature of criticality analysis which are used in asset management to prioritise assets. A machine learning solution is proposed by the development of a *trained criticality analysis model*, with a classification error of 12.35%, which could help in a better prediction of the end result by automating the process i.e. training the model. The second case looks at an application of machine learning on *asset health prediction* by analysing failure patterns and parameters for a machine. The model was evaluated with an error loss of 0.0024. The third case looks at an integration of the priorities related to asset maintenance and management through the development of a text classification machine learning *service selector* (landscape) tool and explores improvising the end-user selection of the services based on their challenges and perceived pain-points related to asset management. The model was evaluated with an accuracy of 84%.

ARTICLE HISTORY Received 31 March 2022; Revised 20 October 2022; Accepted 20 October 2022

KEYWORDS Asset management; machine learning; criticality analysis; predictive maintenance; reliability engineering; deep learning

CONTACT Srinivas Sriramula s.sriramula@abdn.ac.uk School of Engineering, University of Aberdeen, Aberdeen, UK

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

According to ISO 55000 (International Standard, 2014), an asset is defined as 'an item, thing or entity that has potential or actual value to an organization' and asset management is defined as a 'coordinated activity of an organization to realize value from assets'. The asset management and maintenance activities receive significant importance due to the high potential consequences involved in terms of cost, risk and safety of personnel, environment, and reputation. Significant time and effort are invested in improving asset operations by optimising the reliability and availability, reducing the operation and maintenance costs and by adopting digitalisation to improve decision-making. The Institute of Asset Management (IAM)'s conceptual model on asset management states there are six subject groups, namely strategy and planning, asset management decision-making, life cycle delivery, asset information, organisation and people and risk and review (The Institute to Asset Management, 2015). The assets are classified into physical, financial, human, information and intangible assets (Hastings, 2021). Examples of physical assets include equipment, systems, components and plants in the process, mining, and chemical industries and these remain the focus of this paper within the IAM's subject groups of asset management. Asset maintenance management, as a research topic is gaining popularity with techniques such as Reliability-Centered Maintenance; Reliability, Availability and Maintainability (RAM) studies; and optimisation techniques used to optimise the maintenance costs and schedules. These techniques are also relevant to estimate remaining useful life of critical components (Prakash & Kaushik, 2020; Rahdar et al., 2020) by considering stochastic processes and degradation models using sensor data and updating them with posterior information in a Bayesian analysis framework.

Machine learning is widely used in asset maintenance and management owing to the Fourth Industrial Revolution and availability of big data that are interconnected. It is a very efficient tool in processing huge amounts of data and identifying patterns and classifying them to understand the asset better and to facilitate proper management of assets. The dataset is used as an input to the computer as a prior experience and the computer is trained to learn the patterns of the dataset to give a 'trained model'. This will aid in predicting the likely pattern of a new dataset when presented to the trained model. This can uncover the hidden patterns of the new dataset which would improve the decision making and optimise the asset operations and availability. Various applications of the machine learning in asset management include the development of digital twin and predictive analytics for fault diagnostics and anomaly detection to monitor the health of the equipment (Xu & Saleh, 2021). Methods such as random forest decision trees, artificial neural networks, support vector machines are proposed in the literature to model fault detection and reliability prediction systems considering both sensor and

synthetic data for industrial equipment such as pumps, bearing, gearbox, air compressors, gas turbines etc (Carvalho et al., 2019; Riverol & Pilipovik, 2021).

Three cases of engineering problems are discussed in this paper, with data-driven machine learning solutions proposed to these problems. The first case refers to the labour-intensive nature of criticality analysis which are used in asset management to prioritise assets (components or equipment). A machine learning solution is proposed by the development of a trained criticality analysis model which could be used for automation purposes of this analysis. The second case looks at the application of machine learning on predictive analytics by analysing failure patterns and parameters for a machine. The third case describes a common challenge faced by the end-users in selecting the right services from a wide range of services for their asset operations. A development of a text classification machine learning technique is proposed to solve this problem. The following sections provide the background and literature review to the fields of interest in asset management and machine learning. The three case presentations, problems and methodologies for the proposed solutions are presented followed by results and discussion.

2. Background and literature review

2.1. Asset maintenance and management activities

The asset maintenance and management activities are described through a landscape model which encompasses typical services provided by a physical asset maintenance firm. A combination of these services would result in an optimised asset management solution that would satisfy the needs and challenges of the asset operators to improve the asset system availability and reduce the downtime of the assets, while following a systematic planning of operation and maintenance activities. The core services are defined as follows:

- Digital Twin—The concept of digital twin was first proposed by NASA in the aerospace industry in 1969, with the term ‘digital twin’ coined by Michael Grieves in 2002. The technology has gained popularity over the recent years in the energy sector for asset management and is being mentioned as a promising technique in Industry 4.0. A digital twin may be defined as a digital replica of a physical asset, complete with different operating contexts, and may be simulated over numerous scenarios to arrive at a decision on asset management. The data from the physical asset are linked to the digital twin platform to continuously update the digital twin models to complement the physical entities. The digital twin models may be a combination of statistically

derived models and artificial intelligence techniques. The main advantages of using a digital twin include improving the decision processes for operation and maintenance of the assets, predicting the asset behaviour and equipment failures over the life cycle of assets and simulating over any phase of the life cycle of the asset to determine the total cost of ownership and to optimise the reliability, availability and maintainability of the assets (Jiang et al., 2021; Macchi et al., 2018; Stavropoulos & Mourtzis, 2022).

- Online Condition Monitoring—The remote and online monitoring of the assets is achieved through the installation of smart sensors. The signals produced are processed, and the potential failures of the equipment can be predicted well in advance to prevent unscheduled downtime and repairs of the equipment. However, due to the costs of installing condition monitoring equipment and monitoring, it should only be considered if the information it provided would be beneficial to overall uptime. The requirement for the online condition monitoring tasks on critical assets may be determined by applying criticality analysis or reliability-centered maintenance principles which are described in the following subsections. The parameters that could be utilised to continuously monitor the health of the equipment are for example—pressure, temperature, vibration, flow rates. The sensors used for collection of these parameters vary from piezoelectric pressure sensors, thermocouples, thermistors for temperature to displacement probes, accelerometers for vibration and so forth. The online condition monitoring data can be interconnected to other services such as digital twin technology and predictive analytics (Basson, 2018; Stavropoulos & Mourtzis, 2022).
- Predictive Analytics—Artificial intelligence techniques are used to train and learn from the sensor data coming from the assets to understand the equipment health. The model is trained, tested and validated from huge amounts of historical data which reflects the normal and faulty states of the equipment. The historical data is collected through the sensor measurements of operating parameters such as temperature, vibration, flowrates, pressure etc. The features from these data are extracted and pattern recognition machine learning tools are used to achieve a desired trained model. The trained model then predicts the state of the equipment when new set of data is presented to it. The equipment failures and faults are detected at the earliest to prevent unwanted downtime and repair work. The model requires no level of traditional programming and detects the hidden pattern within the dataset (Liu et al., 2018; Pandya et al., 2018). The results from the predictive analytics module may be used for the determination of

maintenance intervals for reliability-centered maintenance projects or for optimisation of spares.

- **Reliability Engineering and Methods**—Reliability is defined as the probability that an asset, equipment or a component will not fail over a specified time interval. The equipment failures over time may be best represented by different probability distribution functions and various statistical inference techniques such as maximum likelihood estimation. Statistical methods are used to identify the best fit probability distribution function from the historical failure data. This probabilistic analysis forms the basis of a significant step in Reliability, Availability and Maintainability (RAM) studies. Operational availability may be defined as the fraction of total time that the equipment is functioning, and the maintainability is defined as the probability that a component or a piece of equipment is being maintained over a specified time interval. The key performance metrics of RAM studies are mean time to failure and mean time to repair which are derived from the probability distribution functions of failure history and repair time of the equipment. A Reliability Block Diagram (RBD) may be first modelled and simulated over the life cycle to quantitatively determine the system availability (Calixto, 2016). Reliability-Centered Maintenance (RCM) technique helps to determine the maintenance strategies of the equipment by considering their functions, functional failures within their operating context, followed by the selection of tasks (according to an algorithm) to prevent or mitigate the consequences. A criticality analysis of each failure mode may also be performed, as a part of the RCM studies. The reader is directed to authors' previous papers on RCM for detailed reference (Nithin, Obisesan, et al., 2021; Nithin, Sriramula, et al., 2021).
- **Inventory Management**—It is often a challenge faced by the asset operators to optimise the spares holding to avoid unnecessary downtime or excessive storage costs. Not having the right spares at the right time leads to production unavailability whilst waiting on spares. Conversely, unwanted stock holdings increase overhead costs such as storage costs, administrative and preservation costs. To optimise spares holdings and inventory management, a trade-off between the unavailability costs, the overhead costs and type of spares need to be considered. To achieve this cost optimisation, the annual demand of spares must be determined from the reliability techniques such as probabilistic analysis of historical failure data, RCM, RAM, Poisson process etc. The economic order quantity can be evaluated from the annual demand of the spare, and the ordering and holding cost of spares. The safety stock and

reordering point are two parameters which have to be considered to evaluate the spare movement and hence, to optimise the spares holding (Ferrari et al., 2006; Gulati & Smith, 2009; Smith, 2011).

- Computerised Maintenance Management System (CMMS)—The industry uses these for work management and to store information about the asset and business operations to improve the enterprise asset performance. Asset registers are built in the CMMS for a hierarchical representation of assets within the system, with information on failure history, repair costs, production downtime, maintenance intervals, condition monitoring measurements, criticality rankings and resources such as labour, materials, and specialist tools. The maintenance tasks and plans derived through reliability methods such as RCM are uploaded into the CMMS, and these should be reviewed regularly for continual improvement of the asset maintenance activities. The workorders for any maintenance tasks are stored and raised within CMMS and provides a platform for workorder backlog management, routeing, and scheduling of resources according to the maintenance tasks (Gulati & Smith, 2009; Nithin, Obisesan, et al., 2021).

2.2. Machine learning techniques

Machine learning techniques are broadly classified into supervised learning, unsupervised learning, and reinforcement learning. Supervised learning encompasses datasets where inputs (explanatory variables) and outputs (labels or responses) are available. In other words, the computer is trained to learn from the input-output data pairs. The machine learning techniques of regression and classification are used for supervised learning, an example of this would be email spam classification, where a large dataset of emails that are spam and not spam are used to train the computer. Unsupervised learning techniques are applicable to unlabelled datasets, and the response prediction is based on clustering or grouping variables of the same behaviour based on the features. An example of unsupervised learning would be clustering or market segmentation to determine the most potential market from a wide population based on their inherent interests. Reinforcement learning refers to automatic learning from mistakes and uses feedback loops to ultimately reduce the mistakes. An example of reinforcement learning is a self-driving car (Gupta & Sehgal, 2021). The statistical tool of linear regression forms the basis of the simplest machine learning technique. A linear approximation of the explanatory and dependent variables is used in linear regression method and used for predictions and relations between the variables (Gupta & Sehgal, 2021).

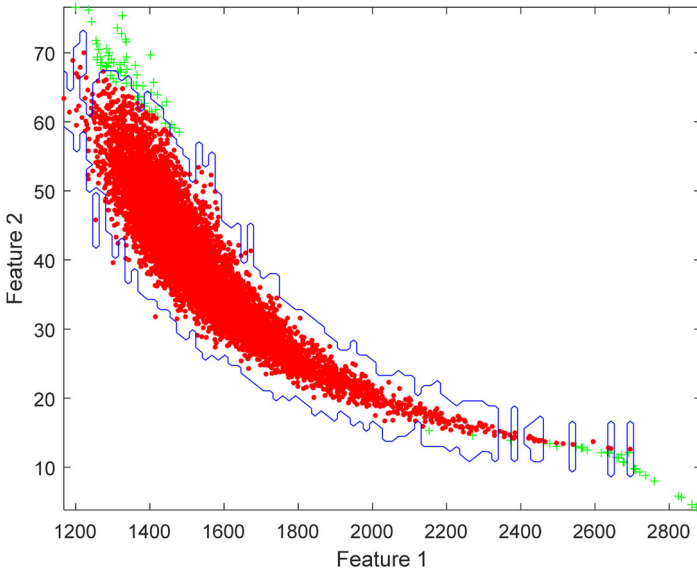


Figure 1. Non-linear boundary of SVM classification technique.

The Support Vector Machine (SVM) classification technique is described here, as a precursor for the following section on applications. SVM essentially finds a hyper-plane between the data to determine a decision boundary for classification. The support vectors are the points closest to the decision hyper-plane. Therefore, the probability of classifying a datapoint correctly increases with the distance of the support vector from the decision hyper-plane. The goal of the SVM is to minimise the cost function that comprises of a training error cost and a penalty term. The training error term refers to the minimal distance of each training example from the hyperplane and the penalty term refers to a factor that accounts for error. Mathematically, a SVM cost function for a labelled dataset of N pairs of x_i input and y_i output is represented as (Forsyth, 2019):

$$S(a, b, \lambda) = \left[\frac{1}{N} \sum_{i=1}^N \max(0, 1 - y_i(a^T x_i + b)) \right] + \lambda \left(\frac{a^T a}{2} \right)$$

where a and b represent the parameters of hyperplane or decision boundary separating the y_i class labels. λ is the regularisation parameter and along with the penalty term $\frac{a^T a}{2}$ ensures that the loss due to learning is minimised for new data which are not accounted for in the training dataset. For non-linear datasets, SVM uses Kernel functions to classify the data by transforming them to higher-dimensional space (Gupta & Sehgal, 2021). **Figure 1** shows a pictorial representation of SVM non-linear Kernel boundary that classifies two groups of data, where data have been compiled from an online database (Matzka, 2020a).

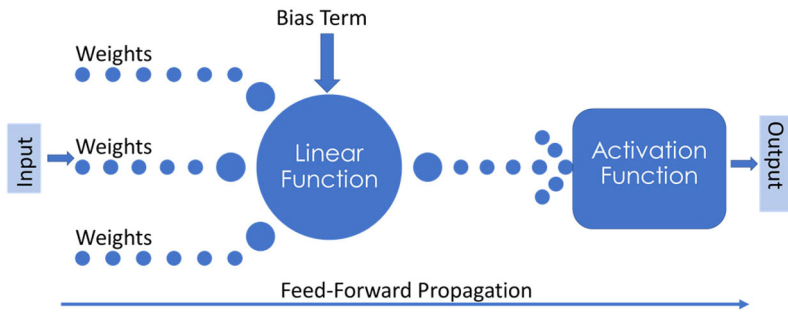


Figure 2. Schematic representation of a neuron (adapted from (López-Monroy & García-Salinas, 2022)).

Another area of machine learning of interest to this paper is the technique of deep learning that is widely used for natural language processing. Deep Learning is based on artificial neural networks, which are inspired by the functionality of biological neural nets as found in the brain. The neuron is a linear function of input values, a set of weights and a bias term. An activation function is then applied on the linear function of the neuron to give the output predicted response and is termed as the feed-forward propagation, as presented in Figure 2.

Several types of activation functions such as sigmoid, Tanh and ReLU (Rectified Linear Unit) activation functions are used in the literature. The loss functions for each of the actual and predicted responses are evaluated during the learning phase as a measure of performance, which basically represents a percentage of error in prediction. The main objective of the deep learning technique is to minimise the cost function of the summation of the loss functions over the entire training dataset and the weights in the initial layers are iteratively updated to meet this objective and is termed backpropagation algorithm (López-Monroy & García-Salinas, 2022). For most natural language processing, the sequence of the data points (words) should be preserved. For this purpose, Recurrent Neural Networks (RNN) are used to repeatedly process and update the hidden states by looking at the input and the previous hidden states, thereby preserving the sequence of the data. A popular model of RNN is the long short-term memory network (LSTM). As the name suggests, LSTM has gates to update, forget the previous memory or hidden states and store these in memory cells, therefore has the capability to find the long and short-term relations between the sequences of the data points and to decide the information that it intends to store or forget (López-Monroy & García-Salinas, 2022; Tsantekidis et al., 2022). A schematic and simplified representation of a standard LSTM architecture is presented in Figure 3.

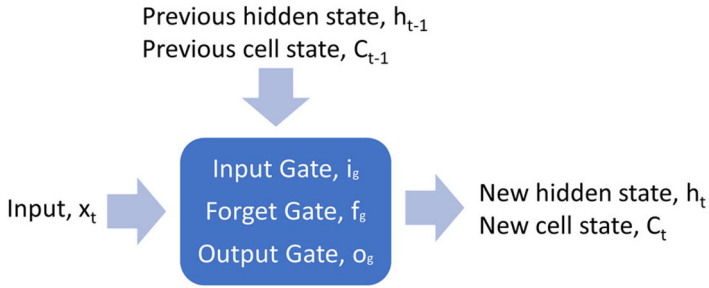


Figure 3. Simplified representation of LSTM architecture (adapted from (López-Monroy & García-Salinas, 2022)).

Mathematically, a LSTM network could be represented as the sigmoid and tanh activation functions of the elements of the network (López-Monroy & García-Salinas, 2022):

$$\begin{aligned}
 i_g &= \sigma(x_t U^i + h_{t-1} W^i) \\
 f_g &= \sigma(x_t U^f + h_{t-1} W^f) \\
 o_g &= \sigma(x_t U^o + h_{t-1} W^o) \\
 \tilde{C}_t &= \tanh(x_t U^c + h_{t-1} W^c) \\
 C_t &= \sigma(f_g \bullet C_{t-1} + i_g \bullet \tilde{C}_t) \\
 h_t &= \tanh(C_t) \bullet o_g
 \end{aligned}$$

where i_g is the input gate, f_g is the forget gate, o_g is the output gate, σ represents the sigmoid function, U^i, U^f, U^o, U^c are the weight matrices that connects the input to hidden layer; W^i, W^f, W^o, W^c are weights that connects the hidden-to-hidden layer; \tilde{C}_t is the candidate hidden state which is an activation function of input and the previous hidden state; C_{t-1}, h_{t-1} and C_t, h_t are cell states and hidden states of the previous and current states, respectively.

3. Applications of machine learning on asset management services

3.1. Development of a novel trained machine learning model of criticality matrix

Criticality analysis is a popular method in reliability engineering to prioritise and rank the critical assets within a system (Basson, 2018). A criticality matrix is developed with likelihood of failure occurrence and consequence of failure on each axis, as shown in Figure 4. The likelihood or the probability of failure occurrence is classified into five categories, ranging from remote occurrence to frequent occurrence. The consequence of failure

		Likelihood of Failure Occurrence				
		Remote	Unlikely	Possible	Probable	Frequent
Consequence of Failure	Catastrophic	Yellow	Yellow	Red	Red	Red
	Critical	Green	Yellow	Yellow	Red	Red
	Major	Green	Green	Yellow	Yellow	Red
	Moderate	Green	Green	Green	Yellow	Yellow
	Minor	Green	Green	Green	Green	Yellow

Figure 4. An example of criticality matrix with prioritisation bands.

ranges from minor failures to catastrophic failures, in terms of the impact of failure.

However, the traditional method of criticality analysis is found to be very time-consuming. A novel SVM application is proposed to develop a trained model of the criticality matrix. For the applicability of SVM, the likelihood of failure occurrence is derived from the mean time to failure of the various assets. The consequence of failure only in terms of repair time is considered in this paper. The minimum and maximum values of the mean time to failure and repair time are derived from the industrial reliability data handbook OREDA (SINTEF & OREDA, 2009). Following this, the data points within the minimum and maximum points were simulated and ranked under four categories of criticality such as minor, major, moderate and catastrophic. This simulated data appended with real data, of 10,000 data points, were used to train the SVM model. A sample of the training data is depicted in Figure 5.

As seen in Figure 6, the trained criticality matrix model with a classification error of 12.36%, was developed through SVM multi-label classification which could be then used for classification and prioritisation of new assets, while facilitating automation and minimising manual intervention. The classification error is obtained through the K -fold cross-validation technique (Bangert, 2021). Here, the data is iterated ' K ' times to split between training and validation subsets, with the subsets varying for each iteration. The final classification error of 12.36% is the average number of errors in classifications within ' K ' iterations. A confusion matrix is illustrated in Figure 7 to show the True Positive Rates (TPR) which denote the percentage of observations correctly classified for each class. False Negative Rates (FNR) are also depicted in the figure to represent the observations classified incorrectly.

The classification error could be further minimised by considering the quadratic relationships between the variables or different machine learning approaches. The comparisons between various approaches could be a subject of future work. With this trained model as shown in Figure 6, the manual intervention for the classification of numerous assets could be

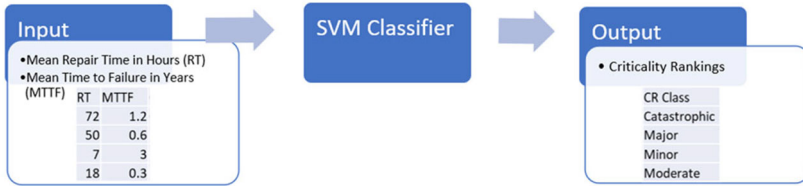


Figure 5. Sample of training data for SVM model.

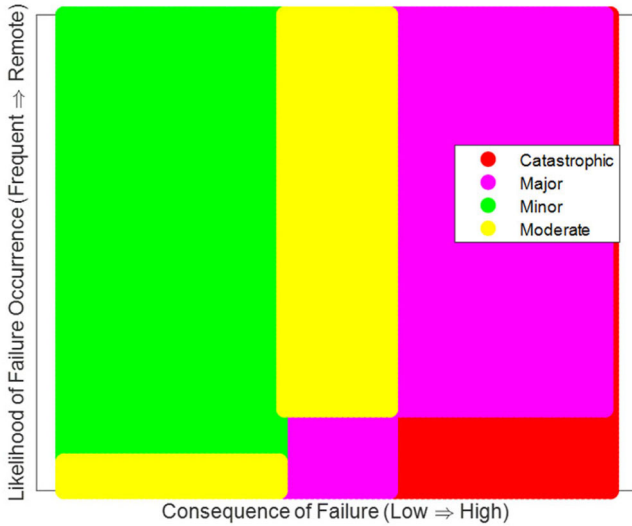


Figure 6. SVM classification decision boundary regions for criticality matrix model.

True Class	Predicted Class				Summary	
	Catastrophic	Major	Minor	Moderate	TPR	FNR
Catastrophic	71.3%	28.7%			71.3%	28.7%
Major	2.7%	90.1%		7.1%	90.1%	9.9%
Minor			95.1%	4.9%	95.1%	4.9%
Moderate		11.2%	13.9%	74.9%	74.9%	25.1%

Figure 7. Confusion matrix illustrating the True Positive Rates (TPR) for SVM model.

Table 1. Sample of training data for SVM asset health prediction.

Air temperature [K]	Process temperature [K]	Rotational speed [rpm]	Torque [Nm]	Tool wear [min]	Failure type
300.8	309.4	1342	62.4	113	'Heat Dissipation Failure'
298.1	308.6	1551	42.8	0	'No Failure'
298.2	308.7	1408	46.3	3	'No Failure'
298.1	308.5	1498	49.4	5	'No Failure'
298.4	308.2	1282	60.7	216	'Overstrain Failure'
298.3	308.1	1412	52.3	218	'Overstrain Failure'
298.9	309.1	2861	4.6	143	'Power Failure'
298.9	309	1410	65.7	191	'Power Failure'
297	308.3	1399	46.4	132	'Random Failures'
301	310.6	1493	37.8	206	'Tool Wear Failure'

minimised, which could instead be focussed on the negligible percentage of data classified incorrectly. It should also be noted that there are other parameters such as production loss, downtime, and redundancy used for criticality analysis and those could be appended to this model as future work.

3.2. Asset health prediction using machine learning techniques

The section explores predictive maintenance from sensor data obtained from an online database (Matzka, 2020a). The sensor data consists of 10,000 data points with information regarding air temperature, process temperature, rotational speed, torque and tool wear, and five failure modes of tool wear failure, heat dissipation failure, power failure, overstrain failure and random failures (Matzka, 2020b). Although several models for predictive maintenance are prevalent in literature, this example is presented for general understanding of the process. The sensor data is analysed through SVM to predict the health of the machine. A sample of training data is given in Table 1.

The boundaries of SVM classification using Gaussian Kernel, with an error loss of 0.0024 and K -fold loss classification error of 0.28%, derived as explained in Section 3.1, are given in Figure 8. The dataset was also split into 80% training dataset and 20% test dataset, which gave a higher test accuracy of 99.3% than the training accuracy which suggests the trained model does not overfit the data. The results align with the literature (Matzka, 2020b), however, a different technique, explainable decision trees was used to arrive at the same conclusion. Figure 8(a) shows a correlation between the torque and rotational speed that may lead to power failure. Figure 8(b) shows the correlation of tool wear and torque that may cause

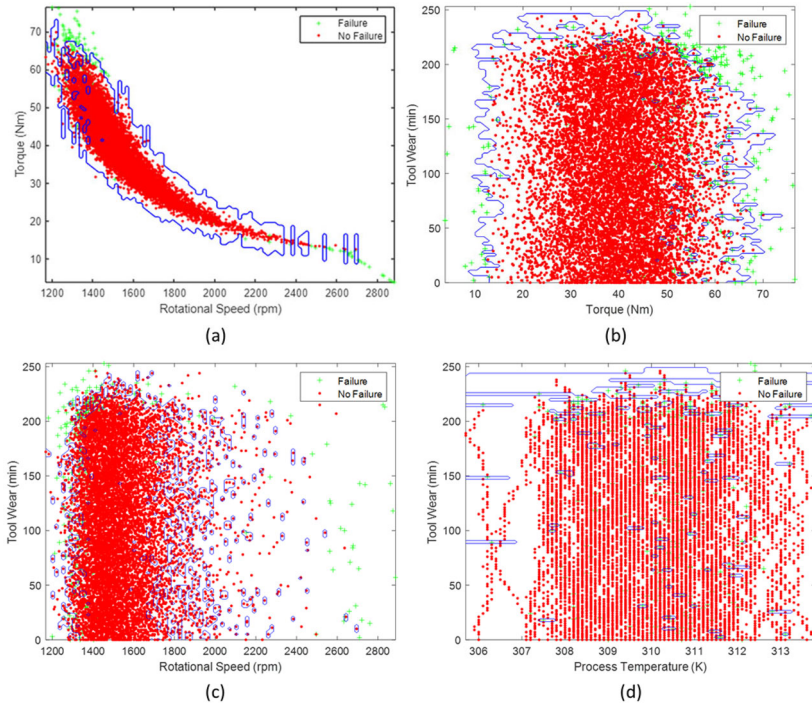


Figure 8. SVM boundaries for asset health prediction.

overstrain failure. The random failures due to tool wear are also evident in Figure 8(c,d).

3.3. Development of a text classification machine learning model for service selection

The natural language processing technique LSTM is applied to the development of a service selector tool which could help the user to choose from the variety of core services such as digital twin, online condition monitoring, predictive analytics, reliability engineering and methods, inventory management, data storage systems which are described in Section 2.1. The tool has been developed in MATLAB version R2021b (The MathWorks & Inc, R2021b). Figure 9 shows the main steps of the service selector (landscape) tool which presents the generic steps of a machine learning task with the novelty of the addition of a layer of word embedding that depicts the relationship between words in a text document. The first step comprised of the data collection and compilation of various core services through expert opinion on the main challenges of asset management and the class labels, the data enhanced using word tuning techniques. The data, with 2278 rows of textual data, is partitioned into training and validation data. Figure 10

Table 2. Sample of training data for LSTM text classification.

Input text	Class labels (output)
Rank or prioritise the critical systems or equipment based on failures.	Criticality Analysis
Predict health of equipment using sensor data through machine learning techniques.	Predictive Analytics
Develop maintenance strategies by studying function and function failure for systems and components.	Reliability-Centered Maintenance

input as shown in [Figure 11](#). For training the LSTM network, a number of hidden layers, along with word embedding layer is used. The word embedding layer maps the words that are similar to each other and places them together in the vector space. The output layer of the LSTM model will be the multi-class classification layer. The model was trained and evaluated with an accuracy of 84%. The trained model is used for prediction of new data for the service selector (landscape) tool as shown in [Figure 12](#).

4. Results and discussion

Machine learning applications in the asset management sector have gained popularity due to their capability to recognise hidden patterns in any data set and have been used to the development of core services mentioned in the paper. A brief introduction to the machine learning fundamentals and two techniques of SVM and deep learning method of LSTM is given. Three illustrative examples for engineering problems are presented which could be used as a machine learning alternative to traditional methods of asset management and maintenance activities. Firstly, the application of SVM multi-label classification to the development of a criticality matrix trained model is provided which would let the end user to rank or prioritise a new set of data by presenting to the trained model. Secondly, an example for asset health prediction using SVM binary classification is provided to study the health of the equipment.

While the first two classification cases are trained on tabular numerical data, the third problem of the service selector tool is trained on textual data. The service selector (landscape) model that brings together the main core services related to asset management would help the asset operators with improved decision-making. Services such as digital twin technology, predictive analytics, online condition monitoring, reliability engineering techniques and methods, inventory optimisation, data cleansing and asset register build and storage through CMMS system are discussed briefly. LSTM is used as a text classification predictor of the service selector (landscape) model that lets the end user to choose from a variety of services based on their challenges and pain points related to asset management.

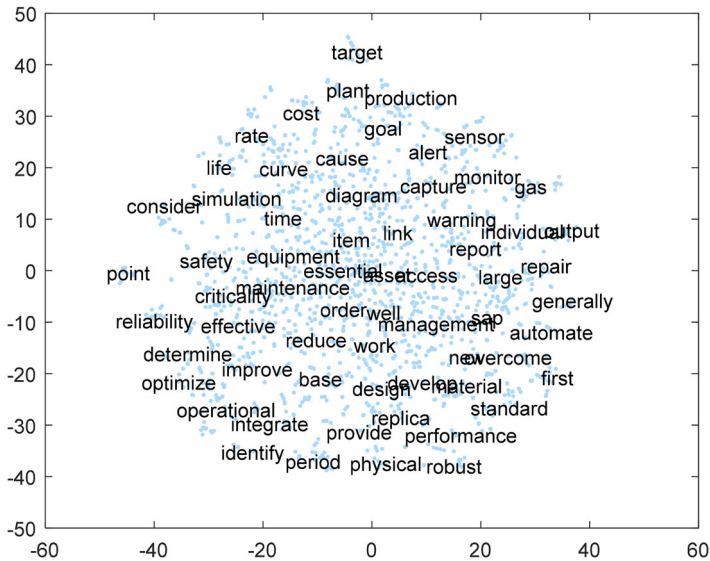


Figure 11. Word encoding into vector space.

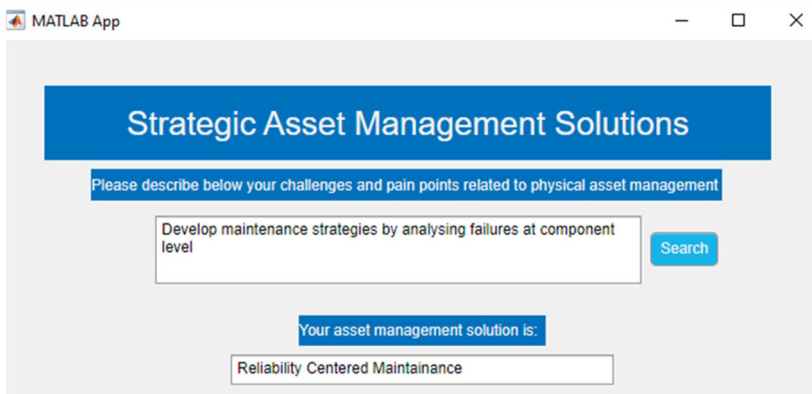


Figure 12. Service selector (landscape) tool for option selection, developed in MATLAB.

The authors initially used the SVM technique for the natural language processing case study. However, this technique gave a lower accuracy and overfitted the data. It was evident that another layer of word embedding was required to understand the relationship between the words and deep learning was required to process the textual data and subsequently, LSTM was chosen. Data collection and word tuning used for the broader topic of asset management terminologies and definitions proved to be the most challenging aspect for the development of the model. The accuracy of the model could be further enhanced by continued development of the textual

input data and could also be verified using other deep learning techniques which could be a subject of future work.

5. Conclusions

This paper looked at three engineering problems related to asset management and maintenance priorities. Data-driven machine learning solutions were proposed that would automate the work processes. A multi-label SVM classification machine learning model for criticality matrix, with a classification error of 12.36%, is proposed as an alternative to the traditional method of criticality analysis. For criticality model, there are other parameters such as production loss, downtime, and redundancy used for criticality analysis and those could be appended to the trained model as future work. The binary SVM classification model for asset health prediction to predict the equipment failure has been modelled with an error loss of 0.0024. The models could be extended by comparing with other machine learning techniques other than SVM. The text classification LSTM machine learning model for the service selector tool was trained with an accuracy of 84%. The complimentary services along with main services could be integrated within the model as a future recommendation.

Disclosure statement

The first and second authors of this paper are currently working with Optimal, and the research outcomes may benefit Optimal. We have disclosed those interests fully to Taylor & Francis, and we have in place a plan for managing any potential conflicts arising from that involvement.

Funding

This research project was funded and supported by Innovate UK and Optimal (Optima Asset Maintenance Solutions Ltd.) as part of the Knowledge Transfer Partnerships Scheme.

ORCID

Srinivas Sriramula  <http://orcid.org/0000-0001-6158-8686>

References

- Bangert, P. (2021). Chapter 3 - machine learning. In P. Bangert (Ed.), *Machine learning and data science in the oil and gas industry* (pp. 37–67). Gulf Professional Publishing. <https://www.sciencedirect.com/science/article/pii/B9780128207147000030>
- Basson, M. (2018). *RCM3 risk-based reliability centered maintenance*. Industrial Press.

- Calixto, E. (2016). Chapter 4 - reliability, availability, and maintainability (RAM analysis). In E. Calixto (Ed.), *Gas and oil reliability engineering* (2nd ed., pp. 269–470). Gulf Professional Publishing. <https://www.sciencedirect.com/science/article/pii/B978012805427700004X>
- Carvalho, T. P., Soares, F., Vita, R., Francisco, R., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- Ferrari, E., Pareschi, A., Regattieri, A., & Persona, A. (2006). Statistical management and modeling for demand of spare parts. In H. Pham (Ed.), *Springer handbook of engineering statistics* (pp. 905–929). Springer London.
- Forsyth, D. (2019). SVMs and random forests. In D. Forsyth (Ed.), *Applied machine learning* (pp. 21–48). Springer International Publishing.
- Gulati, R., & Smith, R. (2009). *Maintenance and reliability best practices*. Industrial Press.
- Gupta, P., & Sehgal, N. K. (2021). Machine learning concepts. In P. Gupta & N. K. Sehgal (Eds.), *Introduction to machine learning in the Cloud with Python: Concepts and practices* (pp. 3–22). Springer International Publishing.
- Hastings, N. (2021). *Physical asset management: With an introduction to the ISO 55000 series of standards*. Springer.
- International Standard. (2014). *Asset management - overview, principles and terminology. Report No. ISO 55000* (1st ed.). International Standard.
- Jiang, F., Ma, L., Broyd, T., & Chen, K. (2021). Digital twin and its implementations in the civil engineering sector. *Automation in Construction*, 130, 103838. <https://doi.org/10.1016/j.autcon.2021.103838>
- Liu, R., Yang, B., Zio, E., & Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 108, 33–47. <https://doi.org/10.1016/j.ymssp.2018.02.016>
- López-Monroy, A. P., & García-Salinas, J. S. (2022). Chapter 9 - neural networks and deep learning. In A. A. Torres-García & C. A. Reyes-García, L. Villaseñor-Pineda, & O. Mendoza-Montoya, (Eds.) *Biosignal processing and classification using computational learning and intelligence* (pp. 177–196). Academic Press. <https://www.sciencedirect.com/science/article/pii/B978012820125100021X>
- Macchi, M., Roda, I., Negri, E., & Fumagalli, L. (2018). Exploring the role of digital twin for asset lifecycle management. *IFAC-PapersOnLine*, 51(11), 790–795. <https://doi.org/10.1016/j.ifacol.2018.08.415>
- Matzka, S. (2020a). Explainable artificial intelligence for predictive maintenance applications. In *2020 Third International Conference on Artificial Intelligence for Industries (AI4I)* (pp. 69–74). IEEE.
- Matzka, S. (2020b). UCI machine learning repository, AI4I 2020 predictive maintenance dataset. University of California, School of Information and Computer Science. <https://archive.ics.uci.edu/ml/datasets/AI4I+2020+Predictive+Maintenance+Dataset>
- Nithin, A. H., Obisesan, A., Sriramula, S., & Ebinum, T. (2021). *Reliability and cost estimation for maintenance planning of oil and gas assets* [Paper presentation]. ASME 2021 40th International Conference on Ocean, Offshore and Arctic Engineering.
- Nithin, A. H., Sriramula, S., & Ebinum, T. (2021). Reliability-centered maintenance and cost optimization for offshore oil and gas components. *Journal of Physics: Conference Series*, 1730(1), 012007. American Society of Mechanical Engineers (ASME). <https://doi.org/10.1088/1742-6596/1730/1/012007>

- Pandya, D., Srivastava, A., Doherty, A., Sundareshwar, S., Needham, C., Chaudry, A., & Krishnalyer, S. (2018). *Increasing production efficiency via compressor failure predictive analytics using machine learning* [Paper presentation]. Offshore Technology Conference (p. D011S002R006). Offshore Technology Conference. <https://doi.org/10.4043/28990-MS>
- Prakash, G., & Kaushik, A. (2020). A change-point-based Wiener process degradation model for remaining useful life estimation. *Safety and Reliability*, 39(3–4), 253–279. <https://doi.org/10.1080/09617353.2020.1801165>
- Rahdar, M. H., Nasiri, F., & Lee, B. (2020). Availability-based predictive maintenance scheduling for vibrating-grate biomass boilers. *Safety and Reliability*, 39(2), 165–187. <https://doi.org/10.1080/09617353.2020.1759261>
- Riverol, C., & Pilipovik, V. (2021). A reliability prediction methodology based on improved radial basis function neural network (RBFNN): a condensate light crude oil stabilization facility as case study. *Safety and Reliability*, 40(3), 157–164. <https://doi.org/10.1080/09617353.2021.1989227>
- SINTEF & OREDA. (2009). *OREDA: Offshore reliability data handbook: Vol. 1 – topside equipment*. OREDA.
- Smith, D. J. (2011). *Chapter 16 - quantified reliability centered maintenance* (pp. 255–262). Elsevier. <http://www.sciencedirect.com/science/article/pii/B9780080969022000167>
- Stavropoulos, P., & Mourtzis, D. (2022). Chapter 10 - digital twins in industry 4.0. In D. Mourtzis (Ed.), *Design and operation of production networks for mass personalization in the era of cloud technology* (pp. 277–316). Elsevier. <https://www.sciencedirect.com/science/article/pii/B9780128236574000105>
- The Institute to Asset Management. (2015). *Asset management - an anatomy. Report No. Version 3*. The Institute to Asset Management.
- The MathWorks, Inc. (R2021b). Natick, Massachusetts, United States. Matlab.
- Tsantekidis, A., Passalis, N., & Tefas, A. (2022). Chapter 5 - recurrent neural networks. In A. Iosifidis & A. Tefas (Eds.), *Deep learning for robot perception and cognition* (pp. 101–115). Academic Press. <https://www.sciencedirect.com/science/article/pii/B9780323857871000105>
- Xu, Z., & Saleh, J. H. (2021). Machine learning for reliability engineering and safety applications: Review of current status and future opportunities. *Reliability Engineering & System Safety*, 211, 107530. <https://doi.org/10.1016/j.res.2021.107530>