Structuring Data for Intelligent Predictive Maintenance in Asset Management ⋆

Oluseun Omotola Aremu * **Adrià Salvador Palau** **
Ajith Kumar Parlikad ** **David Hyland-Wood** ***
Peter Ross McAree *

* *School of Mechanical and Mining Engineering, The University of Queensland, QLD, 4072, AU (e-mail: o.aremu, p.mcaree@ uq.edu.au)*
** *Institute for Manufacturing, University of Cambridge, Cambridge, CB2 1PZ UK (e-mail: as2636, aknp2@ cam.ac.uk)*
*** *School of Information Technology and Electrical Engineering, The University of Queensland, QLD, 4072, AU (e-mail: d.hylandwood@uq.edu.au)*

Abstract: Predictive maintenance (PdM) within asset management improves savings in operational cost, productivity, and safety management capabilities. While PdM can be administered using various methods, growing interest in Artificial Intelligence (AI) has lead to current state of the art PdM relying on machine learning (ML) technology. Like other tools used in PdM for asset management, standards for applying ML technology for PdM are required. This work introduces a standard of practice in regards to usage of asset data to develop ML analytic tools for PdM. It provides a standard method for ensuring asset data is in a form conducive to ML algorithms, and ensuring retention of asset information necessary for optimum PdM during the data transform. In the ML domain, it has been proven through research initiatives that the data structure used to train and test ML algorithms has a great impact on their performance and accuracy. Using poorly trained models for estimation due to improper data usage, can leave some AI-based PdM tools vulnerable to high rates of inaccurate estimations. Thus, leading to value loss during an asset's life cycle.

Keywords: Asset management, predictive maintenance, artificial intelligence, information sharing, big data

1. INTRODUCTION

Predictive maintenance (PdM) is the monitoring of an asset or system's condition over it's life cycle to provide a prognosis to when maintenance is required (Grall et al., 2002). Positive contributions of PdM show it is a key component for asset management at the operational and enterprise level in many industries (Faiz and Edirisinghe, 2009). As an asset's life cycle data is composed of many measurements that translate to large volumes of data, data-driven algorithms are often used for PdM analytics (Alaswad and Xiang, 2017). Along with the maturity of data-driving technology, current PdM tools are increasingly dependent on machine learning (ML) based artificial intelligence (AI) technology (Walker et al., 2013).

Although ML algorithms have proven to be state of the art in improving prognostics and condition monitoring capabilities, they are not perfect. Their success can be marred by the data form or structure used to train and test algorithms (Peng et al., 2010). These limitations can result in increased computation complexity, time, and reduced accuracy (Martínez-Álvarez et al., 2016). Thus, rendering prognosis PdM tools ineffective in generating real-time estimations of an asset's future state and accurate results (Saxena et al., 2010). Currently, there is no architecture or framework that provides a standard of practice for how data should be structured per method of PdM analysis dependent on ML. Thus, data-driven PdM algorithms could be limited in the ability to provide accurate and current information depending on choice in ML-based analysis.

High dimensional (multivariate) data structures with continuous time series data samples, along with variation in data sparsity, highlight the limitations of ML algorithms (Martínez-Álvarez et al., 2016). These variations in data are common and make it difficult for ML algorithms to contrive information from observed data. As raw asset data commonly contain characteristics of such variations, they are not conducive to ML algorithms in a generic form.

⋆ Acknowledgement: Manufacturing for providing the computer used to perform the simulations. This research was supported by the Australian Government Research Training Program Scholarship, American Australian Association Education Fund Fellowship, National GEM Consortium, Sustain-Owner (Sustainable Design and Management of Industrial Assets through Total Value and Cost of Ownership), a project sponsored by the EU Framework Programme Horizon 2020, MSCA-RISE-2014: Marie Skłodowska-Curie Actions Research and Innovation Staff Exchange (Rise) (grant agreement number 645733 Sustain-owner H2020-MSCA-RISE-2014).

The inability of ML algorithms to manage high-dimensional data coupled with variation in sample density is known as the *curse of dimensionality* (James et al., 2013). The *curse of dimensionality* defines a fixed condition of ML algorithms in which the higher the dimension space, the more dense data samples required (Stuart and Peter, 2016). Essentially, for a ML algorithm, a 100 by 200 point cloud in which each point could be a variable, can lead to a dimensionality of 20,000 which will lead to an observed sparsity of data. This generated sparsity inhibits ML algorithms from deriving strong statistical inferences based on data point relations, which is the core principle of most ML algorithms (Theodoridis and Koutroumbas, 2009).

The *curse of dimensionality* further emphasizes the importance of properly using data in ML-based applications. Operations using ML-based analytics in PdM of assets with multivariate sensing, can be exposed to poor algorithm performance due to unmanaged high-dimensionality. This can lead to logistic inefficiencies in operations. Thus, resulting in a potential economic disaster from improper management of high value assets.

This work proposes a framework that provides a method for achieving a lower dimension representation of an asset's life cycle while minimizing loss of critical information. It introduces a method to transform asset life cycle data into a form better suited for ML based PdM tools, which enables current and accurate AI-based PdM of assets.

In Section 2 the current state of ML within PdM is described, and further details on requirements of data structured for ML is provided. Section 3 presents main challenges within data manipulation of asset life cycle data, and provides a guiding principal for solutions necessary to structure data for PdM dedicated ML analytic tools. Then, Section 4 provides a method for correlating information of domain knowledge to fault modes present in an asset's life cycle data. Metrics used to evaluate the reduced form of a data set are explained in Section 5. Subsequently, Section 6 provides methods used to assess the performance of ML algorithms using the reduced data form. Next, Section 7 describes a proposed data manipulation framework to achieve a lower dimension form. In closing, Section 8 provides a discussion detailing the proposed method of data dimension reduction, followed by Section 9 which concludes the presented work.

2. ARTIFICIAL INTELLIGENCE IN PREDICTIVE MAINTENANCE

In this work, AI is limited to ML through methods of pattern recognition (Stuart and Peter, 2016). With respect to classes of ML algorithms employed, this work is bounded to the categories of clustering, classification, regression, and probabilistic decision models. These are methods commonly used in varying combinations to develop and apply ML analytics used for PdM within asset management (Tobon-Mejia et al., 2011).

The common requirement of PdM in industries, is to provide information capable of influencing or guiding decision making during condition monitoring of an asset or process, and determining its probability of failure (Su et al., 2006). Within manufacturing, PdM can be used to predict events such as component outages in an assembly process, while in utilities, it can be used to predict outages in power equipment such as wind turbines (Kothamasu et al., 2006). In other industries such as mining, it can be used to achieve optimal maintenance scheduling of high value assets (Makinde and Ramatsetse, 2016).

In attempt to meet this requirement across diverse industries, current PdM technology is now more dependent on data-driven ML based algorithms (Susto et al., 2015). These algorithms require access to high volumes of reliable historical data. Thus, for PdM analytical tools to work effectively, there must be an adequate amount of context rich asset historical data; an adequate amount is defined by requirements of industry and application. As the volume of sensors and measurements increase to capture critical asset information, the need for a standardized framework to properly curate and structure asset data for ML-based PdM is required.

3. CHALLENGES IN STRUCTURING ASSET DATA FOR MACHINE LEARNING

When reducing data sets to a lower dimension representation, there is a balance between retained information and dimension size that must be maintained. Raw high-dimensional data sets contain large volumes of data, but it is difficult to deduce useful information (Guyon and Elisseeff, 2003). With an obtained reduced dimension form, one is able to manage the data in a tangible manner, but has now reduced the likelihood of containing the highest amount of useful information. This balance is further tested when information relating to domain knowledge is required. Also, the higher the dimensionality of a data set, the lower the results and performance of ML algorithms (Theodoridis and Koutroumbas, 2009). Reducing computational complexity while improving result accuracy is key to improving the effectiveness of any ML algorithm.

With respect to PdM, the challenge in developing an effective data restructuring process to achieve a reduced dimension form of an asset's life cycle requires a balance between: (1) maintaining the highest correlation to the raw data set by retaining the highest possible measure of domain knowledge related information, (2) ensure the new data form has reduced the complexity of computation through computation time and reduce space, and (3) minimize unnecessary introduction of new parameters when deriving new variables from the original data.

4. DOMAIN KNOWLEDGE IN ASSET DATA

Commonly, ML tools for PdM consist of various algorithms used to identify fault modes and capture their effects within an asset's life cycle. In Susto et al. (2015) the authors provide a multiple classifier approach by combining *support vector machines* (Theodoridis and Koutroumbas, 2009) and *k-nearest neighbors* (Theodoridis and Koutroumbas, 2009). The method was tested in the failure prediction of a semiconductor system based on observed fault types. Alternatively, in Peng et al. (2010), the authors describe the implementation of an artificial neural network to predict system failure features (fault modes) and remaining useful life (RUL) estimations. In Zhang et al. (2006) the authors describe the use of *hidden Markov models* (Rabiner, 1989) in combination with *principal component analysis* (Theodoridis and Koutroumbas, 2009) to conduct advanced diagnostics by creating an adaptive fault prediction model.

Literature shows that faults and their implications are the essential knowledge for PdM in the domain of asset management. Therefore this work assumes faults and their implications correlate to a type of information contained in an asset's life cycle data and are translatable to a type of domain knowledge representation (An et al., 2013). *Fault identification* and *fault classification* are not within the scope of this work, but quantifying and retaining their effects as domain knowledge during data manipulation is key to ensure the lower dimension data form remains conducive to PdM analytics.

This work proposes *Rényi entropy* in (1) (Rényi et al., 1961), as a method to measure the presence of domain knowledge within asset data. Rényi et al. (1961) explains that entropy can be used as a method of measuring the "unexpectedness" of data in a variable or feature, by quantifying its contained uncertainty. In *Rényi entropy*, $H_\alpha(X)$, X is a feature with x samples, and $\alpha\epsilon[0,\infty]$ denotes the entropy order of the probability mass function, $p(x)$. α can be viewed as the distribution type used to weight $H_\alpha(X)$ in relation to probabilities of $p(X = x)$. For example, in the limit for $\alpha \to 0$, $H_0(X)$ follows a uniform distribution and all probabilities of $p(X = x)$ are weighted equally.

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \int p^\alpha(x)dx. \tag{1}$$

According to Rényi et al. (1961), the greater the entropy of a feature, the more information it contains. Assuming that an asset's nominal condition has a known entropy measure, once the effects of faults are present in an observation, a change in entropy is expected. Thus, the measure of entropy in a data set with known faults can stand as a measure of the influence of faults on the data set.

5. IDEAL PROPERTIES OF STRUCTURED DATA

There exist many properties that characterize an ideally-structured data set for efficient ML algorithm functionality. In alignment with asset management constraints, this work will focus on the following:

Retained essential features: reduces chance of over-fitting during algorithm training, which can lead to an inaccurate model (Keogh et al., 2001).
Indexing: data structures with discretized observations improve an algorithm's ability to analyze each data sample, thus allowing for faster identification of data points with minimal computation (Stuart and Peter, 2016).
Memory constraints: data sets requiring less memory capacity, ensure more of the information in the data set can be interpreted or captured in one instance (Padillo et al., 2016).

6. ALGORITHM PERFORMANCE ON STRUCTURED ASSET DATA

Quantitative evaluation methods are necessary to determine if the performance of ML algorithms is improved by the lower dimension data form. If a lower dimension data form does not accurately maintain the data's discriminatory information contained in the original data, ML algorithms will be ineffective (Theodoridis and Koutroumbas, 2009). The following are metrics considered as ideal performance indicators to evaluate algorithm performance when using a lower dimension data form (Lin et al., 2007):

Classification: assign an observed variable or set of data points to a predefined class or label (Theodoridis and Koutroumbas, 2009).
Anomaly detection: determine if newly observed data points do not match a predefined pattern (Bonissone and Iyer, 2007).
Clustering: determine the natural grouping of observed data points in a data set (Theodoridis and Koutroumbas, 2009).

Classification, *anomaly detection*, and *clustering* are adequate measures to determine an algorithm's effectiveness and provide individual metrics that evaluate an algorithm's performance per learned data set (Lin et al., 2007). By evaluating the performance of an algorithm per metrics of each method, one can quantitatively evaluate the effectiveness of the achieved data structure. Also, algorithm performance within each category provide insight to the degree at which the new data form has maintained the original discriminative information and assumed original domain knowledge specific information.

7. A REDUCED DATA FORM OF ASSET DATA

7.1 General Process

As shown in Fig. 1, *preprocessing* is the first step in reducing a data set to a lower dimension form. It requires removal of outliers, noise reduction, and completion of missing data (Theodoridis and Koutroumbas, 2009). Next, *feature selection* is used to identify a feature subset with the highest discriminatory information and variance (Guyon and Elisseeff, 2003).

Methods such as *hypotheses testing* (Theodoridis and Koutroumbas, 2009) are used to evaluate features individually, while some techniques like *correlation criteria* (Guyon and Elisseeff, 2003) evaluate feature subsets. Other techniques select features based on classification performance and select feature subsets that improve a classifier's accuracy (Guyon and Elisseeff, 2003). Finally, *dimension reduction* through linear methods such as *principal component analysis* is applied to achieve a lower dimension space representation of the original data set (Roweis and Saul, 2000).

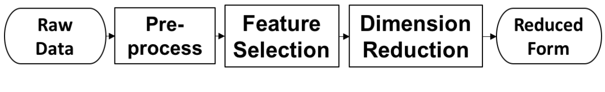

Fig. 1. General data manipulation framework to achieve a lower dimensional data form

This general process is acceptable for some data types such as images, text, or signal data, but its conditions and criteria for effective manipulation do not translate well for asset management data due to:

- Interpretation of feature characteristics vary between general data and asset data, for example:
  - In asset data, low variance can be an indicator of a nominal state, while a high variance feature could be an unreliable sensor (Wang, 2010).
- Dimension reduction and feature selection methods do not provide a method to measure retention of information pertaining to domain knowledge.
- No consideration for diversity of features. In asset data, different features can measure different aspects of an asset and can require separate data manipulation processes.

7.2 Proposed Data Transform Framework

This work presents an asset management dedicated data transform process that achieves a lower dimension data form conducive to AI-based PdM. This framework, shown in Fig. 2, reduces asset data into a lower dimension form while retaining information necessary for PdM. Achieving a data form ideal for ML dependent PdM analytics, requires processes and methods dedicated to both ML and asset management. These processes and methods are explained in the following subsections.

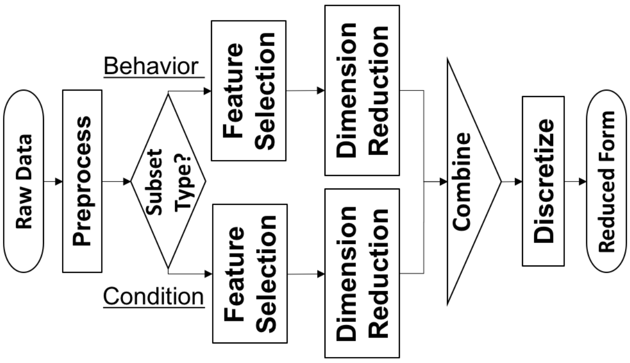

Fig. 2. Proposed data manipulation framework to achieve a lower dimension data form specific to PdM

7.3 Measured Condition and Measured Behavior

Measurements that describe an asset's life cycle can be divided into asset's input settings and output responses (Wang, 2010). Although this is not always the case, it is common among most assets and this work adopts this concept. An asset's measurements within the framework depicted in Fig. 2, will be assessed per subsets of *measured conditions* and *measured behaviors* throughout its life cycle. *Measured conditions* describe an asset's usage patterns (work load, flying altitude, operating speed) and its environmental conditions (ambient temperature, pressure, asset model). Thus, output properties of an asset are *measured behaviors* of the system's output due to usage (power consumption, current draw, outlet temperature).

Presumably, *measured behavior* data will contain domain context information such as fault modes and their effects on the asset's performance. The *measured conditions* of an asset are linked to its *measured behavior*, similarly to how results of an experiment are connected to the experimental set-up. In normal conditions, the relations linking the condition (input) with the behavior (output) will be constant. Therefore, domain specific information such as fault modes will also be captured within these relations. Both subsets contain individually pertinent information, and should not be combined during data manipulation. A dedicated process to address the proper combined feature generation from the two subsets is required, as shown in Fig. 2. Premature combined data manipulation could: (1) generate features that do not truly capture domain knowledge with respect to the asset's condition, and (2) result in a loss of information critical to understanding the implications of fault modes.

7.4 Feature Selection per Condition and Behavior Data

Feature selection is an important process. It reduces the effects of "bad data" by removing features that do not adequately measure an asset's life cycle. Ideal methods will select features based on statistical relevance, and their ability to describe domain specific information in *measured conditions* and *measured behavior* subsets. The measure of a feature's *entropy*, as described in (1), is used as the primary selection criteria. Then, correlation methods can be used to ensure the feature subset has minimal correlation and high content of discriminatory information.

It is in this process that both domain knowledge and statistical relevance are combined. As the results are further propagated through the remainder of the data transform framework, its success bares a strong influence on the final result of the data transform process. Thus, one must be confident in the output of this process.

Along with meeting the requirements set by correlation and entropy, the feature subset must also be tested to determine its fit for use with ML algorithms as described in section 6. Generally, selected feature will lead to larger distinctions between classes, but have high measures of associative characteristics within each class for a classifier (Theodoridis and Koutroumbas, 2009). For anomaly detection, the selected features will produce higher confidence and consistency in detection of outliers (Theodoridis and Koutroumbas, 2009). Using cluster evaluation, generated clusters will have low intra-cluster distances and high inter-cluster distances (Theodoridis and Koutroumbas, 2009).

7.5 Dimension Reduction in Condition and Behavior Data

In this process, dimension reduction is based on the data's original manifold (lower-dimensional constraint in which the data is embedded) to ensure a confident representation of the original data's feature relations without assuming linear or non-linear dependencies (Lafferty and Lebanon, 2005). Manifold learning algorithms are more conducive to dynamic systems, and aid in reducing loss of information by allowing multi-dimensional representation of the data (Han et al., 2013). *Diffusion mapping, kernel principal component analysis,* and *isometric mapping* amongst other similar methods are common methods of manifold based *dimension reduction* (Han et al., 2013).

Similar to general testing results mentioned in Section 7.4, effective dimension reduction will also maintain or improve the standards for clustering, anomaly detection, and classification.

7.6 Combined Representation of Subsets

Combining *measured condition* and *measured behavior* features to a singular lower dimension representation is a nontrivial process. Fig. 3 provides a representation of behavior and condition data relationship and how they will be combined. Also, (2) describes the mapping of *measured condition* and *measured behavior* to a singular representation:

$$\varphi(C, B, t) : D \rightarrow LC. \tag{2}$$

$\varphi(C, B, t)$ maps the raw data, $D \epsilon R^{m+n}$, of arbitrary sample size, to a singular series, $LC : [x(c_1, b_1, t_0), ..., x(c_c, b_n, t_f)]$. LC describes the asset's life cycle with respect to condition observations, $C_{1:c}$, behavior observations, $B_{1:b}$, and time of observation, $t \epsilon [t_0 : t_f]$. m is the number of features describing the condition subset, and c is the number of identified condition modes. While n, is the number of features in the behavior subset, and b is the number of determined fault modes, ranging between *nominal* and *failure* as shown in Fig. 3.

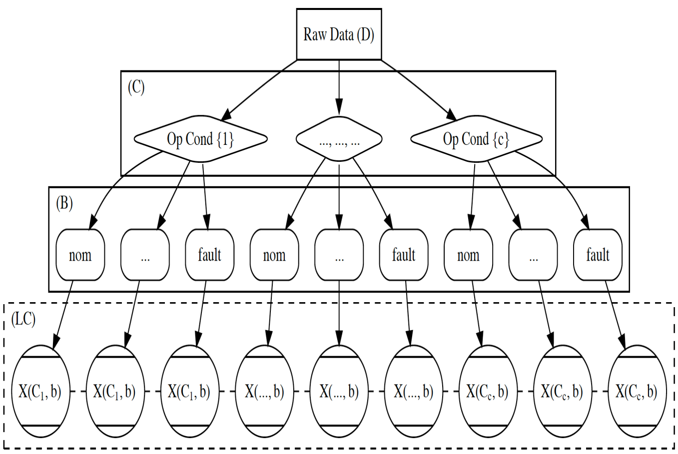

Fig. 3. Subset combination process to achieve a singular representation of an asset's life cycle (LC) based on mapping of identified *condition* (C) and *behavior* (B) modes

The output, LC, describes an asset's condition per information gathered from each subset in a singular data representation. This stage of the data manipulation process is the highlight of the proposed framework. It produces a uniform combined representation of an asset's life cycle, while containing information about it's operating condition and responses in a lower dimension form.

7.7 Discretization to Sequential Form

Discretization reduces dimension within individual feature samples. Improper application can result in reduction or complete loss of discriminatory information, and domain knowledge information maintained or generated in previous stages. Common methods of discretization include *discrete Fourier transform*, *discrete time wavelet transform*, and *Haar wavelet transform* (Theodoridis and Koutroumbas, 2009). They are applicable in various forms, but often contain flaws that are prohibitive to the goal of the concept presented in this work. For example, these methods are weak in maintaining correlation of data point characteristics, such as distance measurement between the discretized form and original data form. Also, they perform poorly as a second stage dimensionality reduction method as they do not often reduce the dimension within individual feature samples.

A more applicable method is SAX (Symbolic Aggregate approXimation) (Lin et al., 2007). SAX uses a framework that maintains the original feature's lower-bound distance measure per sample in the final discretized form. SAX uses Piecewise Aggregate Approximation (Lin et al., 2007) as its primary method to perform lower-bound distance estimations. Through this method of discretization, it is expected that information necessary to conduct effective ML-based PdM analysis can be retained. In previous applications, SAX has shown promising results (Georgoulas et al., 2015; Sun et al., 2014), and it aims to address the challenge of discretization posed in this work.

8. DISCUSSION

The main contribution of the proposed data transform framework is to minimize loss of information during data manipulation while ensuring the preservation of domain knowledge pertinent to PdM analysis. Although the framework can account for domain knowledge presence in the data through measures of entropy, an open question remains:

> *Is information about fault presence fully preserved in a dimension reduction of the original data?*

This work uses entropy to measure loss of information during data manipulation, with the assumption that the retained information also contains information describing fault implications. Further research, that is better aligned with *fault identification*, is required to determine if methods like entropy, that measure information within a variable, can be used to differentiate between fault types. Then one can further assume the same methods can be used to measure retention of fault implications during dimension reduction of asset data.

9. CONCLUSION

This work highlights the possibility of improving the performance and results of ML algorithms for AI-based PdM analysis by ensuring the asset data is properly structured for ML algorithms. This is accomplished through the introduction of a novel framework dedicated to standardizing data transform for AI-based PdM dependent on ML algorithms. Also, the framework provides a reduced form of an asset's life cycle data while retaining information specific to asset management. The reduced representation is expected to be smaller in data size and therefore reduces constraints of data size during data transfer. Thus, other research areas that focus on data transfer between assets or systems can benefit from this work. Overall, the proposed framework is a tool that can improve various aspects of asset management.

REFERENCES

Alaswad, S. and Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157, 54–63.

An, D., Choi, J.H., and Kim, N.H. (2013). Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab. *Reliability Engineering and System Safety*, 115, 161–169.

Bonissone, P.P. and Iyer, N. (2007). Soft Computing Applications to Prognostics and Health Management (PHM): Leveraging Field Data and Domain Knowledge. *Computational and Ambient Intelligence. Proceedings of 9th International Work Conference on Artificial Neural Networks, IWANN 2007*, 928–939.

Faiz, R.B. and Edirisinghe, E.A. (2009). Decision making for predictive maintenance in asset information management. *Interdisciplinary Journal of Information, Knowledge, and Management*, 4, 23–36.

Georgoulas, G., Karvelis, P., Loutas, T., and Stylios, C.D. (2015). Rolling element bearings diagnostics using the Symbolic Aggregate approXimation. *Mechanical Systems and Signal Processing*, 60, 229–242.

Grall, A., Dieulle, L., Bérenguer, C., and Roussignol, M. (2002). Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE transactions on reliability*, 51(2), 141–150.

Guyon, I. and Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research (JMLR)*, 3(3), 1157–1182.

Han, Y., Xu, Z., Ma, Z., and Huang, Z. (2013). Image classification with manifold learning for out-of-sample data. *Signal Processing*, 93(8), 2169–2177.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning*, volume 112. Springer.

Keogh, E., Chakrabarti, K., Pazzani, M., and Mehrotra, S. (2001). Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases. *Knowl. Inf. Syst.*, 3(3), 263–286.

Kothamasu, R., Huang, S.H., Verduin, W.H., Kothamasu, R., Huang, S.H., and Verduin, W.H. (2006). System health monitoring and prognostics- a review of current paradigms and practices. *International Journal of Advanced Manufacturing*, 28.

Lafferty, J. and Lebanon, G. (2005). Diffusion Kernels on Statistical Manifolds. *Journal of Machine Learning Research*, 6(1), 129–163.

Lin, J., Keogh, E., Wei, L., and Lonardi, S. (2007). Experiencing SAX: a Novel Symbolic Representation of Time Series. *Data Min Knowl Disc*, 15, 107–144.

Makinde, O. and Ramatsetse, B. (2016). Robotic System Methodology for Automated Maintenance of a Reconfigurable Vibrating Screen. (July).

Martínez-Álvarez, F., Troncoso, A., Quintián, H., and Corchado, E. (2016). A Data Structure to Speed-Up Machine Learning Algorithms on Massive Datasets Francisco. In *International Conference on Hybrid Artificial Intelligence Systems*, volume 9648, 365–376.

Padillo, F., Luna, J.M., Cano, A., and Ventura, S. (2016). A Data Structure to Speed-Up Machine Learning Algorithms on Massive Datasets. In *International Conference on Hybrid Artificial Intelligence Systems*, volume 9648, 365–376.

Peng, Y., Dong, M., and Zuo, M.J. (2010). Current status of machine prognostics in condition-based maintenance: A review. *International Journal of Advanced Manufacturing Technology*, 50(1-4), 297–313.

Porte, J.D.L. and Herbst, B. (2008). An introduction to diffusion maps. In *Proceedings of the 19th Symposium of the Pattern Recognition Association of South Africa (PRASA 2008)*, 15–25. Cape Town, South Africa.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.

Rényi, A. et al. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California.

Roweis, S.T. and Saul, L.K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500), 2323–2326.

Saxena, A., Roychoudhury, I., and Celaya, J.R. (2010). Requirements Specifications for Prognostics : An Overview. In *Proceedings of AIAA Infotech@Aerospace 2010*, April, 3398–3398.

Stuart, R. and Peter, N. (2016). Artificial intelligence-a modern approach 3rd ed.

Su, Y.C., Cheng, F.T., Hung, M.H., and Huang, H.C. (2006). Intelligent prognostics system design and implementation. *IEEE Transactions on Semiconductor Manufacturing*, 19(2), 195–206.

Sun, Y., Li, J., Liu, J., Sun, B., and Chow, C. (2014). An improvement of symbolic aggregate approximation distance measure for time series. *Neurocomputing*, 138, 189–198.

Susto, G.A., Schirru, A., Pampuri, S., McLoone, S., and Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, 11(3), 812–820.

Theodoridis, S. and Koutroumbas, K. (2009). *Pattern Recognition, Fourth Edition*. Academic Press.

Tobon-Mejia, D., Medjaher, K., Zerhouni, N., and Tripot, G. (2011). Hidden Markov models for failure diagnostic and prognostic. *Prognostics and System Health Management Conference*, (PHM"11), 1–8.

Walker, R., Perinpanayagam, S., and Jennions, I.K. (2013). Rotordynamic faults: recent advances in diagnosis and prognosis. *International Journal of Rotating Machinery*, 2013.

Wang, T. (2010). *Trajectory Similarity Based Prediction for Remaining Useful Life Estimation*. Ph.D. thesis, University of Cincinnati.

Zhang, Q., Benveniste, A., and Zhang, Q. (2006). Early warning of slight changes in systems and plants with application to condition based maintenance.