Machine Learning for Predictive Maintenance

Dwi Kusumaningrum, Nani Kurniati, Budi Santosa
Department of Industrial and Systems Engineering
Sepuluh Nopember Institute of Technology
Surabaya, 60111, Indonesia
dwi.19024@mhs.its.ac.id, nanikur@ie.its.ac.id, budi_s@ie.its.ac.id

Abstract

Downtime due to sudden machine failure will cause much loss to the company. To overcome this, companies need to develop suitable maintenance strategy. Nowadays, machines in smart manufacturing provide large volume data to monitor machine conditions. Big data analytics becomes needed for processing large data, especially for predicting machine failure in Industry 4.0. Predictive maintenance works better than corrective or preventive maintenance. It can continuously monitor diagnostic and prognostic processes to predict future failures and the equipment's remaining useful life (RUL). Using real multi-sensor data and machine failure reports in industrial equipment, machine learning models can study data patterns and build failure prediction models based on real-time condition monitoring. The purpose of this study is to construct a diagnostic and prognostic model with tuning optimal machine learning parameters in support vector machine and random forest (RF) for classification of equipment conditions and RUL so that company can find out the prediction of future failure times. Also, the comparison of machine learning parameters and methods is carried out to determine which model has the highest accuracy. Based on each model's accuracy, RF performed better than SVM in diagnostic and prognostic models.

Keywords
Machine Learning, Predictive Maintenance, Diagnostic, Prognostic, Remaining Useful Life

1. Introduction
Smart manufacturing is physical and digital integration in manufacturing from upstream to downstream by utilizing advanced information to increase flexibility and adaptability to change with the Internet of Things (IoT) (Moyne and Iskandar, 2017). The data generated by the supporting machines for smart manufacturing has a large volume. Large-scale data processing (big data analytics) is necessary for data processing to support company decision making. Data can come from variations in production machines and equipment as well as factory locations. One of the data used in supporting industry 4.0 is sensor data to monitor the machine's reliability. Sensor data can be used to make maintenance planning decisions to avoid unnecessary equipment replacement, reduce unplanned machine downtime, energy efficiency, and save cost that usually occurs in the company.
Based on a case study on one manufacturing company in Indonesia that already uses smart manufacturing industrial equipment, machine availability only reaches 70 – 80 %. This number can lead to several low productivities in production. With provided large volume of sensor-data, the company will have the opportunity to optimize maintenance strategy and reduce unplanned machine downtime to increase availability time and production productivity.
There are three types of machine maintenance strategies: corrective maintenance, preventive maintenance, and predictive maintenance (Carvalho et al, 2019). Corrective and preventive maintenance exclude diagnostic and prognostic activities in determining maintenance. In predictive maintenance, condition-based maintenance (CBM) is carried out. CBM discusses the maintenance plan's optimization by relying on sensor data generated by the engine in real-time and preventing system failures by knowing when and how the machine will be maintained. Developing a maintenance strategy using CBM method will need a good quality of data representing the existing machines' situation and state.
CBM data processing with machine learning can continuously analyze patter data from the latest data by itself. Machine learning in predictive maintenance can be used to diagnose and prognostic the machine's deterioration against existing parameters. However, not many studies have carried out both processes using real industrial data. Several

studies were found that discussed a diagnostic or prognostic process and only used simulation data. Praveenkumar et al. (2014) used support vector machines (SVM) to diagnose failure identification in automotive transmissions with a good accuracy level. The research by Prytz et al. (2015) used random forest (RF) as a classification algorithm and comparison of the selection of different data parameters to predict the damage of several components of commercial vehicles to model accuracy. Both SVM and RF algorithms have good accuracy in overcoming several problems such as inconsistent datasets and unbalanced datasets. Also, Mathew et al. (2018) map sensor data for the diagnosis process, which is to classify based on machine conditions using an artificial neural network that results in good accuracy.
Besides, several studies have method of diagnosis and prognosis in the application of predictive maintenance. Soylemezoglu et al. (2011) and Susanto and Kurniati (2020) used the Mahalanobis Taguchi System to apply diagnostics and prognostics in machine maintenance. In this research, the actual machine condition and the threshold for normal and abnormal sensors can be found. Tran and Yang (2012) performed diagnostics and prognostic machine conditions using machine learning using PCA and SVM algorithms to process larger data volumes faster than without machine learning. Previous research shows that predictive maintenance with machine learning has promising results on the implementation in industries.
In this research, the author proposes a predictive maintenance approach using machine learning using real industrial equipment `data based on supervised learning. This multi-sensor data will be processed using several machine learning methods, namely SVM, and random forest. The aims to obtain the diagnostic model for motor EPFAN machine conditions using machine learning to classify normal and abnormal conditions. Also, to get a prediction model for machine failure using machine learning for machine RUL classification. With these models, the company will have increased availability of existing machines in the company. Some of these methods are used to perform diagnosis and prognosis in the classification of machine conditions and machine RUL, which have a high degree of accuracy. Furthermore, the accuracy of machine learning methods is compared to determine which method has the best results.

2. Literature Review

2.1 Predictive Maintenance

Predictive maintenance is a method of periodic monitoring of the actual condition of the machine, operating efficiency, and other indicators, and the system will provide the data needed to ensure the maximum interval between repairs and minimize the amount and cost of unplanned downtime on machines (Mobley R K, 2002). Predictive maintenance can make productivity, product quality, the overall effectiveness of production better. The use of several tools such as vibration monitoring, thermography, and tribology can be a comprehensive method of applying predictive maintenance to determine the actual condition of the system based on actual data on maintenance activities. Predictive maintenance is a machine activity that is based on machine conditions using the actual current state of the machine rather than relying on average life statistics such as mean time to failure for scheduling maintenance activities.

2.1 Support Vector Machine

Support vector machine is a technique for classifying, which is included in supervised learning. Given the input data $x_i(i = 1,2,\dots,M)$ M is the number of samples. According to Santosa and Umam (2018), there are two main approaches in multi-class SVM, namely:

1) One Versus All
This method constructs k an SVM model where k is the number of classes. SVM class *i* is trained with all existing data samples in class *i* with a positive label and all other examples are labeled negative. An example is given *l* training data $x_1, y_1), \dots, (x_l, y_l)$ where $x_i \in R^n, i = 1, \dots, l$ and $y_i \in \{1, \dots, k\}$ is the class of x_i so the *i*th SVM satisfies the following problem:

$$minimize\colon \frac{1}{2}\left|\left|w^i\right|\right|^2 + C\sum_{i=1}^{M} \xi_j^i (w^i)^T \quad (1)$$

$$subject\ to$$

$$(w^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i \text{ jika y = i}, \quad (2)$$

$$\xi_j^i \geq 0, j = 1, \dots, l$$

Where the x_i training data is mapped to a higher dimensional space with the function ϕ and C is the penalty parameter, where the penalty can reduce the number of errors from training.

2) One Versus One

This method constructs a k (k-1) / 2 classification in which each foreigner is trained on data from two classes. For training data from classes i and j, it can be solved with the following problems:

$$minimize: \frac{1}{2}\left\|w^{ij}\right\|^2 + C\sum_t \xi_j^{ij}(w^{ij})^T \tag{3}$$

$$subject\ to$$

$$(w^{ij})^T\phi(x_t) + b^{ij} \geq 1 - \xi_t^{ij} \text{ jika } y_t = j, \tag{4}$$

$$\xi_t^{ij} \geq 0, j = 1, \dots, l$$

.

2.2 Random Forest

The random forest algorithm is an ensemble learning method that builds the aggregate of decision trees from the training dataset. Each decision tree produces a response in the form of a predictor value. In each decision tree, each internal node shows a test on the attribute, each branch describes the results of the test, and each leaf node represents a class label for classification or the response of the regression (Wu et al. 2017) a decision tree that has a continuous response is called a regression tree. To reduce variance and avoid overfitting. Some essential concepts related to RF are bootstrap aggregation or bagging, slipping, and stopping criterion.

Bootstrap or bagging aggregation produces *B* training dataset D_i known as bootstrap sample. With replacement sampling or bootstrapping, multiple observations are repeated for each D_i. The bagging method can reduce variance and overfitting.

Feature bagging is random sampling of m variables for each note and choose the best division among these variables rather than choosing the split between predictors. If the partition is divided into *M* region $R_1, R_2, \dots, R_m$then the response can be modeled as a constant c_m in each region.

$$f(x) = \sum_{m=1}^{M} c_m I(x \in R_m) \tag{5}$$

The criterion for separating each node is to minimize the sum of square. So that $\widehat{c_m}$ is the mean of y_i in the R_m region.

$$\widehat{c_m} = ave(y_i|x_i \in R_m \tag{6}$$

The separator variable j and the dividing point s, then a pair of half-planes is

$$R_1(j,s) = \{X|X_j \leq s\}\ dan\ R_2(j,s) = \{X|X_j \leq s\} \tag{7}$$

Which must fulfill the following objective functions:

$$min_{\ j,s}\left[min_{c1}\sum_{x_i \in R_{1(j,s)}}(y_i - c1)^2 + min_{c2}\sum_{x_i \in R_{2(j,s)}}(y_i - c2)^2\right] \tag{8}$$

For every j and s, the minimization is solved by:

$$\widehat{c_1} = ave(y_i|x_i \in R_1(j,s)) \text{ and } \widehat{c_2} = ave(y_i|x_i \in R_2(j,s))$$

Stopping criterion is a separation process that continues until the number of records is below the threshold and five is used as the threshold. For each tree that is constructed, a prediction at the new point x can be made by averaging the predictions of all individual regression trees B on x:

$$\widehat{f_{rf}^{B}}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \tag{9}$$

3. Research Method

The proposed model's critical phases are data acquisition, data pre-processing, feature selection, build classification model, and validation. Each step of this proposed model is essential and adds valuable insight into its performance. Figure 1 gives a scheme that is showing activities to do to implement PdM.

Data acquisition for this research is an essential phase because the performance will depend on the correctness of a dataset. Accurate and good quality of data will give a good result that can represent the proposed model's effectiveness. Research data are collected from numerous sources such as periodical sensor data, downtime machine, and maintenance activities report. After data acquisition, data pre-processing needed. Data pre-processing includes data labeling, data cleaning, and data transformation. This phase is essential because classifiers cannot process raw data. After all, some incomplete and missing data will cause data noise. Also, sensor data will need to be transformed to the same range because each sensor has different range of values, so computation will be more effective and prevent data from bias. So, the model output will give accurate accuracy. Transformed data will be split into test datasets and training datasets. Training dataset will be trained with various algorithm which is Support Vector Machine (SVM), and Random Forest (RF) based on their proven ability in classification problem and data label to build industrial equipment diagnostic and prognostic model until it has good accuracy. If the model has not reached good accuracy yet, hyperparameter tuning will be tuning the model's parameter. After the model achieves good accuracy, the proposed model is tested with test dataset to validate the output accuracy and optimized parameters with 10-fold time-series cross-validation. After the model is validated, algorithms' accuracy will be compared to determine which algorithm has the highest accuracy. The highest accuracy algorithm will be chosen to be the main algorithm in the proposed predictive maintenance model.

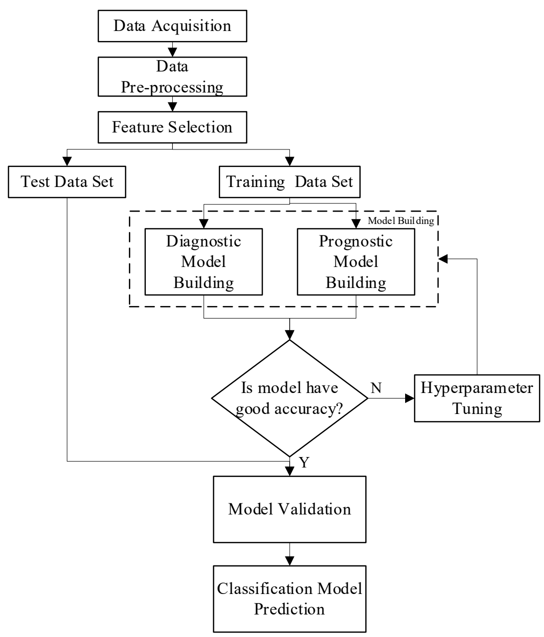

Figure 1. Proposed method

4. Data Collection

Data collection starts with collecting sensor data that contains electric current, temperature, and vibration that are generated periodically. Also, supporting data such as downtime and maintenance activities activity reports from 2017 - 2019. This data collection will support data labeling to construct input data for the proposed model. In data labeling for diagnostic, data will be divided into normal and several classes abnormal that contains various machine failure causes, which can be seen through supporting data such as downtime and maintenance reports. In data labeling for prognostic, data will be divided into two classes of RUL range. Based on expert judgment at a manufacturing company,

data label will be divided into RUL for more than one month and less than one month which can be seen through supporting data such as downtime reports.
After labeling data, data will be pre-processed to clean data noise and transform sensor data to a standardized scale range. After sensor data standardized, sensor data with label data will become input for building the diagnostic and prognostic model.
In building diagnostic and prognostic models, this research used SVM, and RF calculated with Python software for analyzing equipment sensor data. This phase aims to obtain diagnostic and prognostic model based on machine condition with machine learning algorithms that have highest accuracy and obtain prognostic model based on RUL range. To get the highest accuracy, each algorithm's parameter will be tuning with grid search to determine each algorithm's best parameter. Table 1 until 2 shows the parameter used in the hyperparameter tuning process. The parameter scenarios are obtained from previous such as Erfanifard et al. (2014), in tuning SVM parameters, Sonobe et al. (2014) in tuning RF parameters.

Table 1. Parameters used for SVM model

Parameter	1	2	3	4
Kernel	Linear	Polynomial	RBF	Sigmoid
C	10	100	200	300
Gamma	0.1	0.2	0.3	0.4

Table 2. Parameters used for RF model

Parameter	1	2	3
Splitting Criteria	Entropy	Gini	-
N Decision Trees	50	100	150

5. Results and Discussion

5.1 Diagnostic Modelling

The accuracy of SVM, RF best parameters on 20% test and 80% training dataset for diagnostic modelling after validated with 10-fold cross-validation is shown in Table 3.

Table 3. Grid-search Result Diagnostic Model

Algorithm	Default Parameter				After Tuning			
	Parameter	Accuracy	Weighted-Precision	Weighted -Recall	Parameter	Accuracy	Weighted -Precision	Weighted -Recall
SVM	kernel = RBF	0,927	0,93	0,93	kernel = RBF	0,940	0,94	0,94
	C = 1				C = 300			
	Gamma = scale				Gamma = 0,4			
RF	criterion = gini	0,974	0,98	0,98	criterion = gini	0,974	0,98	0,98
	n_estimators = 100				n_estimators = 50			

According to Table 3, the best accuracy, precision, and recall in diagnostic modeling are obtained with RF algorithm compared with SVM in this case study. RF has successfully reached the accuracy of 97% from total data labels. Moreover, precision and recall that obtained with RF consistently reach the highest value among other algorithms. This result is validated with 10-fold accuracy that shows in Figures 2. Figure 2 shows that each algorithm does not indicate overfitting because the model tested with different test data and does not tell a significant difference among each fold.

The grid-search results show that the tuning parameter does not significantly differ in model accuracy, especially on the RF algorithm. Figure 3, shows that accuracy from small to large amount of decision trees constantly around 97%. This result indicates that with a smaller number of decision trees, the model can reach the same accuracy with larger number of trees.

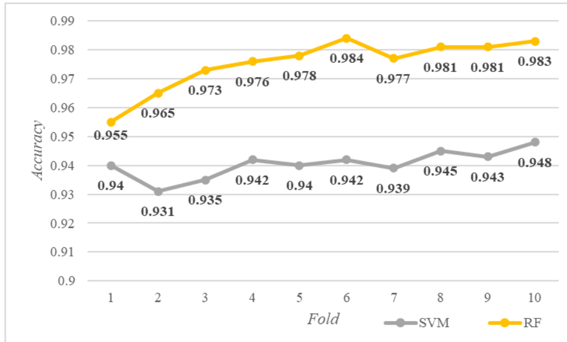

Figure 2. 10-fold time-series cross-validation diagnostic modelling

Moreover, different using kernel and C on SVM algorithm will have different accuracy. This accuracy movement shows in figure 4. With using RBF kernel and different C value, SVM algorithm accuracy have better results from another kernel. With smaller C, hyperplane margin will increase and cause larger errors in the model (Luts et al., 2010). But if using too much C will cause overfitting into the model (Santosa & Umam, 2018), it will need n-fold cross-validation to validate the results.

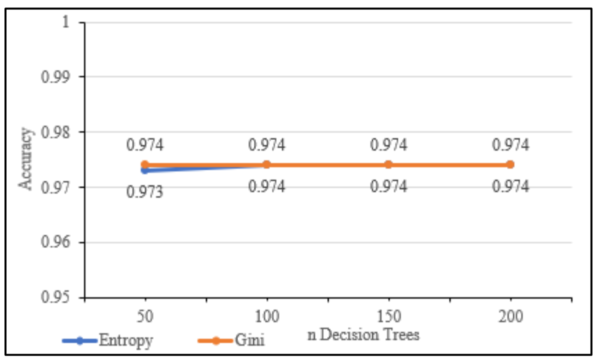

Figure 3. Accuracy RF based on number of decision trees in diagnostic modelling

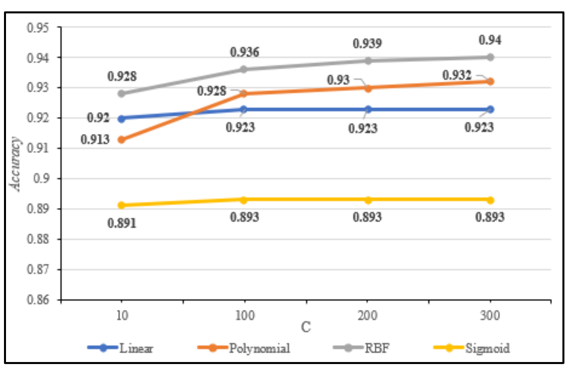

Figure 4. Accuracy SVM based on kernel and C value diagnostic modelling

5.2 Prognostic Modelling

The accuracy of SVM, RF best parameters on 20% test and 80% training dataset for prognostic modeling after validated with 10-fold cross-validation is shown in Table 4.

Table 4. Grid-search Result Prognostic Model

Algorithm	Default Parameter				After Tuning			
	Parameter	Accuracy	Weighted-Precision	Weighted-Recall	Parameter	Accuracy	Weighted-Precision	Weighted-Recall
SVM	kernel = RBF C = 1 Gamma = scale	0,857	0,860	0,860	kernel = RBF C = 300 Gamma = 0,4	0,867	0,88	0,87
RF	criterion = gini n_estimators = 100	0,896	0,896	0,896	criterion = gini n_estimators = 150	0,897	0,91	0,91

According to Table 4, the best accuracy, precision, and recall in prognostic modeling are obtained with RF algorithm compared with SVM in this case study. RF has successfully reached the accuracy of 89,6% from total data labels. Moreover, precision and recall that obtained with RF consistently achieve the highest value compared to SVM. This result is validated with 10-fold accuracy that shows in Figures 5. Figure 5 shows that each algorithm does not indicate overfitting because the model tested with different test data and does not tell a significant difference among each fold.

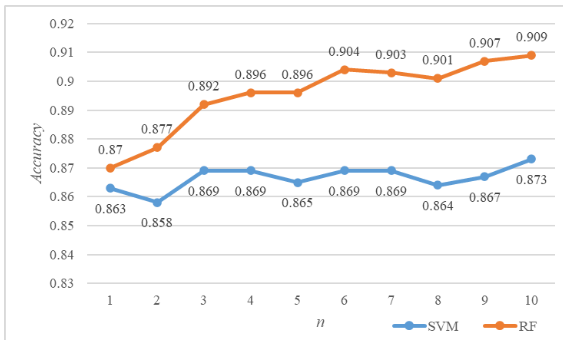

Figure 5. 10-fold time series cross validation prognostic modelling

The grid-search results on Table 4 show that tuning parameter does not have a significant difference towards both model accuracies. Figures 6 shows that accuracy from small to large amount of decision trees constantly around 89%. This result shows that there is no significant difference tuning hyperparameter tuning in prognostic modeling. Moreover, using kernel RBF and Poly and C value on the prognostic model with SVM algorithm does not differ from accuracy. Constant accuracy movement for RBF and poly shows in figure 7. It shows that kernel sigmoid has the smallest accuracy compared to another algorithm.

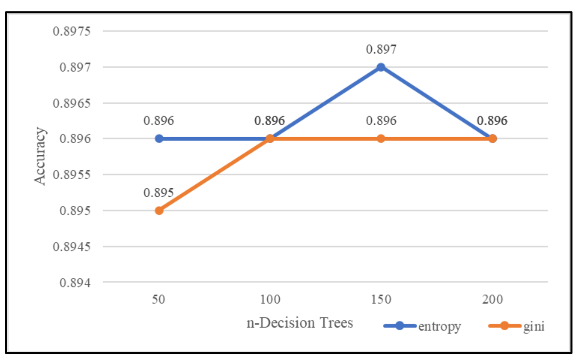

Figure 6. Accuracy RF based on number of decision trees in prognostic modelling

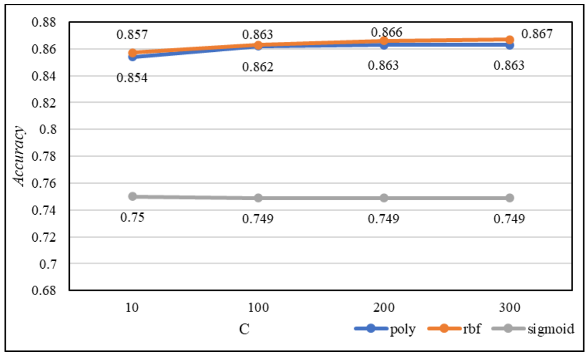

Figure 7. Accuracy SVM based on kernel and C value prognostic modelling

6. Conclusion

This research proposed a fault diagnostic algorithm based on two algorithms for multi-class classification of normal and abnormal conditions and fault prognostic for RUL classes in predictive maintenance of Motor EPFAN with sensor data. Different machine learning techniques, namely SVM and RF have been used to proposed model scenarios. According to the discussion and result analysis, the proposed diagnostic and prognostic model could predict classes of each model. Each algorithm has good accuracy for the diagnostic and prognostic model. But the highest accuracy, precision, and recall were obtained with random forest. Therefore, RF is a suitable algorithm for Motor EPFAN fault diagnosis and prognostic. Using this proposed model, similar industries can optimize sensor data utilization to make better decision-making for maintenance strategies and predict remaining RUL. Other machine learning techniques and parameter will be explored more to improve the performance of research results in the future.

References

Bansal, Dheeraj, et al. "A Real-Time Predictive Maintenance System for Machine Systems." International Journal of Machine Tools and Manufacture, vol. 44, no. 7–8, pp. 759–66, doi:10.1016/j.ijmachtools.2004.02.004, 2004.

Carvalho, Thyago P., et al. "A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance." Computers and Industrial Engineering, vol. 137, no. August, Elsevier, 2019, p. 106024, doi:10.1016/j.cie.2019.106024, 2019.

Erfanifard, Yousef, et al. "Tree Crown Delineation on UltraCam-D Aerial Imagery with SVM Classification Technique Optimised by Taguchi Method in Zagros Woodlands." International Journal of Image and Data Fusion, vol. 5, no. 4, pp. 300–14, doi:10.1080/19479832.2014.936527, 2014.

Mathew, Vimala, et al. "Prediction of Remaining Useful Lifetime (RUL) of Turbofan Engine Using Machine Learning." IEEE International Conference on Circuits and Systems, ICCS 2017, vol. 2018-Janua, no. Iccs, pp. 306–11, doi:10.1109/ICCS1.2017.8326010, 2018.

Mobley R K. "Introduction to Predictive Maintenance." Elsevier Science, https://books.google.com.my/books?hl=en&lr=&id=SjqXzxpAzSQC&oi=fnd&pg=PP1&dq=predictive+maintenance+in+food+&ots=iFvSQDKogg&sig=OiSs4k_l1HgMYOVmjrU9BjnQW68&redir_esc=y#v=onepage&q=predictive maintenance in food&f=false, 2002.

Moyne, James, and Jimmy Iskandar. "Big Data Analytics for Smart Manufacturing: Case Studies in Semiconductor Manufacturing." Processes, vol. 5, no. 3, doi:10.3390/pr5030039, 2017.

Praveenkumar, T., et al. "Fault Diagnosis of Automobile Gearbox Based on Machine Learning Techniques." Procedia Engineering, vol. 97, Elsevier B.V., pp. 2092–98, doi:10.1016/j.proeng.2014.12.452, 2014.

Prytz, Rune, et al. "Predicting the Need for Vehicle Compressor Repairs Using Maintenance Records and Logged Vehicle Data." Engineering Applications of Artificial Intelligence, vol. 41, pp. 139–50, doi:10.1016/j.engappai.2015.02.009, 2015.

Santosa, Budi, and Ardian Umam. Data Mining Dan Big Data Analytics. 2nd ed., Penebar Media Pustaka, 2018.

Sonobe, Rei, et al. "Parameter Tuning in the Support Vector Machine and Random Forest and Their Performances in Cross- and Same-Year Crop Classification Using TerraSAR-X." International Journal of Remote Sensing, vol. 35, no. 23, pp. 7898–909, doi:10.1080/01431161.2014.978038, 2014.

Soylemezoglu, Ahmet, et al. "Mahalanobis-Taguchi System as a Multi-Sensor Based Decision Making Prognostics Tool for Centrifugal Pump Failures." IEEE Transactions on Reliability, vol. 60, no. 4, pp. 864–78, doi:10.1109/TR.2011.2170255, 2011.

Susanto, B. S. P., and N. Kurniati. "Multi Sensor-Based Failure Diagnosis Using the Mahalanobis Taguchi System." IOP Conference Series: Materials Science and Engineering, vol. 847, no. 1, 2020, doi:10.1088/1757-899X/847/1/012036, 2020.

Tran, Van Tung, and Bo Suk Yang. "An Intelligent Condition-Based Maintenance Platform for Rotating Machinery." Expert Systems with Applications, vol. 39, no. 3, Elsevier Ltd, pp. 2977–88, doi:10.1016/j.eswa.2011.08.159, 2012.

Wu, Dazhong, et al. "A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests." Journal of Manufacturing Science and Engineering, Transactions of the ASME, vol. 139, no. 7, pp. 1–9, doi:10.1115/1.4036350, 2017.

Biographies

Dwi Kusumaningrum is a final year master's student at the Department of Industrial and Systems Engineering, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia. She has received a bachelor's degree from Department of Industrial and Systems Engineering Sepuluh Nopember Institute of Technology, Surabaya, Indonesia. She has research interest in manufacturing, data mining, and reliability.

Nani Kurniati is an active associate professor at the Department of Industrial and Systems Engineering, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia. She has received Ph.D. degree from Department of Industrial Management from National Taiwan University of Science and Technology (NTUST), Taiwan. She is interested in and has experience researching the reliability and maintenance field and the quality engineering and management field.

Budi Santosa is an active professor at the Department of Industrial and Systems Engineering, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia. He has received Ph.D. degree from School of Industrial Engineering, University of Oklahoma. He is interested in and has experience in data mining and metaheuristic optimization research field.