Predictive Maintenance in Oil and Gas Dataset by using Naive Bayes and Gaussian Elimination Method

Alex Goh Chee Hong[1], Tole Sutikno[2], Shahreen Kasim[1,*]

[1]Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, 86400 Batu Pahat, Malaysia

[2]Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

[*]Corresponding author's email: shahreen@uthm.edu.my

ABSTRACT: Maintenance is an important process that improve the overall performance of a system in a long-term. In this paper we are focusing on predictive maintenance. By using sensor and run's data, estimation can be done based on the health status of the machine. This can improve the performance and life span of the machine, at the same time it also able to reduce the total maintenance cost due to predictive maintenance is preventive maintenance. In order to perform predictive maintenance, we are using Naives Bayes classifier to classify pass and fail data point. For linearization of data, we are using Gaussian Elimination Method to obtain a normal distribution data. The data and then will be processed with K-Nearest Neighbour to produce a full data sheet for visualization purpose.

Keywords: *Predictive Maintenance, Hybrid Algorithm, Naives Bayes*

1. INTRODUCTION

Maintenance is a process that preserve a condition, situation or state. Maintenance usually involves fixing, repairing, improving of an event to ensure the event can be carried out smoothly. Maintenance is required as an event will be facing issues, errors or fault if were left over a period. Hence, maintenance is essential to ensure the continuation of a process.

There are various types of maintenances available, such as run to failure, schedule maintenance, preventive maintenance, and predictive maintenance. Table 1 shows the summary of each maintenance.

Table 1

Type of Maintenances

Type of Maintenances	Cost	Automation	Efficiency
Run to Failure	Low	No	Low
Scheduled Maintenance	High	No	Low
Preventive Maintenance	Average	Semi	Medium
Predictive maintenance	low	Yes	High

In this paper we are focusing on Predictive maintenance where maintenance is done by using machine learning algorithm.

2. METHODLOGY

Predictive maintenance is developed based on the sensor's reading and previous run data. The dataset used are oil and Gas data. Once classify the data using Numpy and Pandas library, we will be carrying out classification and normalization.

2.1 Naives Bayes Classification

Naives Bayes classification is used to differentiate between two different data points. In this research it is used to differentiate the passing and failing run based on the data set.

2.2 Gaussian Elimination Method

Gaussian Elimination Method is used after we done classifying using Naives Bayes Classifier. Gaussian Elimination method is used to produce a normal distribution reading to have an actual percentage of a passing and failing run. It is important to do so as it is allowed us to differentiate when a run is failing, we are aware of it and maintenance can be carried out.

3. RESULT AND DISCUSSION

Naïve Bayes classifier is suitable to be used in this predictive maintenance model as it already helped to improve the overall performance of prediction by classifying success and fail data into the correct category and then only do Gaussian Distribution which is used to determine the probability of the failure to occur.

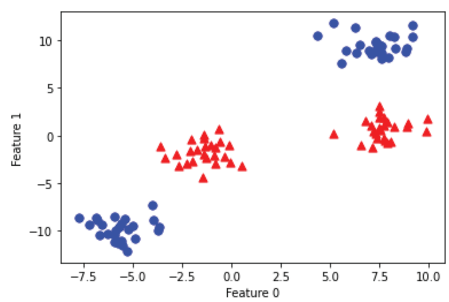

Figure 1: Naives Bayes Classifier Classify Result

Based on Figure 1, we can see that data are being classified into correct cluster. As we can see that red colour is a failing data while blue colour is a success passing data. After we done classification, we will do a linear curve to determine the failing and passing value on each run.

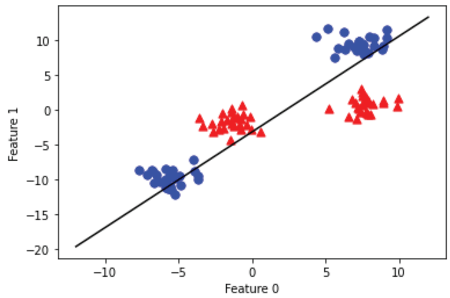

Figure 2 : Naives Bayes Classifier with Linear Regression

Based on Figure 2 we can see that the passing run is lie on the linear regression line while for the failing it lies further away from the regression line. The regression line is indicating a full cycle run. This means that, when the reading of data is too extreme it will mostly fail which this data is very useful for predicting when there is a failure.

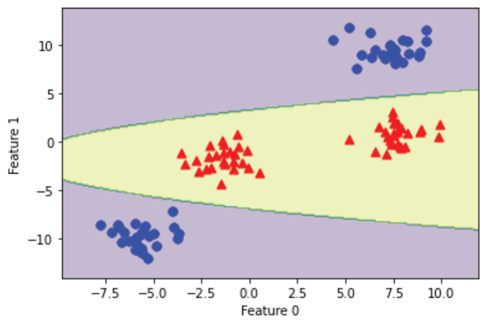

Figure 3 : Normal Distribution graph with Gaussian Distribution

Based on Figure 3, we can see that the failure data point had been captured inside the bell curve which eventually will lead to a success of predicting failure run. This had proved that the hybrid algorithm is working for predicting failure. Moreover, passing case also categorized outside the bell curve which is also very important evidence to prove that it will not trigger any maintenance event outside the curve as there is no success run within the bell curve. Hence whenever there are run that lies within the bell curve, means that a maintenance will be required.

4. CONCLUSION

This paper had proven that Predictive Maintenance can be done with sensor's data and previous run data sheet. For future works, this research is going to include K-Nearest Neighbour algorithm to improve the reading and better prediction accuracy.

ACKNOWLEDGEMENT

This study was supported by the Ministry of Higher Education under Fundamental Research Grant Scheme (FRGS/1/2019/ICT02/UTHM/02/3).

REFERENCES

[1] Rivera, D. L., Scholz, M. R., Fritscher, M., Krauss, M., & Schilling, K. (2018). Towards a Predictive Maintenance System of a Hydraulic Pump∗. *IFAC-PapersOnLine*.Siew-Kien Mah and Sim-Hui Tee,"Computational Analysis of PTEN Gene Mutation," International Journal on Advanced Science, Engineering and Information Technology, vol. 2, no. 5, pp. 383-386, 2012. [Online]. Available: http://dx.doi.org/10.18517/ijaseit.2.5.229.

[2] Vianna, W. O. L., & Yoneyama, T. (2018). Predictive Maintenance Optimization for Aircraft Redundant Systems Subjected to Multiple Wear Profiles. *IEEE Systems Journal*. https://doi.org/10.1109/JSYST.2017.2667232Liu, W., Li, C., Xu, Y., Yang, H., Yao, Q., Han, J., . . . Li, X. (2013). Topologically inferring risk-active pathways toward precise cancer classification by directed random walk. Bioinformatics, 29(17), 2169-2177. doi:10.1093/bioinformatics/btt37

[3] Butler, J., & Smalley, C. (2017). An introduction to predictive maintenance. *Pharmaceutical Engineering*. Misman, M. F., Mohamad, M. S., Deris, S., & Hashim, S. Z. (2014). A group-specific tuning parameter for hybrid of SVM and SCAD in identification of informative genes and pathways. International Journal of Data Mining and Bioinformatics IJDMB, 10(2), 146. doi:10.1504/ijdmb.2014.064013

[4] Selcuk, S. (2017). Predictive maintenance, its implementation and latest trends. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*. https://doi.org/10.1177/0954405415601640Landi, M. T., Caporaso, N. E., Jen, J., Dracheva, T., Rotunno, M., Figueroa, J. D., . . . Shih, J. H. (2008). Gene Expression Signature of Cigarette Smoking and Its Role in Lung Adenocarcinoma Development and Survival. PLoS ONE, 3(2). doi:10.1371/journal.pone.0001651

[5] Cline, B., Niculescu, R. S., Huffman, D., & Deckel, B. (2017). Predictive maintenance applications for machine learning. In *Proceedings - Annual Reliability and Maintainability Symposium*. https://doi.org/10.1109/RAM.2017.7889679X. Wang, E. Dalkic, M. Wu, C. Chan. Gene Module Level Analysis: Identification to Networks and Dynamics. (2008). Current Opinion in Biotechnology. 19(5): 482-491.

[6] Dey, A. (2016). Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technologies*.S. A. Gagliano, A. D. Paterson, M. E. Weale, and J. Knight. Assessing models for genetic prediction of complex traits: a comparison of visualization and quantitative methods. (2015). BMC genomics, 16(1), 405.

[7] Kelleher, J. D., Namee, B. Mac, & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics*. *Igarss 2014*. Yang, S., & Naiman, D. Q. (2014). Multiclass cancer classification based on gene expression comparison. Statistical Applications in Genetics and Molecular Biology, 0(0). doi:10.1515/sagmb-2013-0053.

[8] Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers and Industrial Engineering*.

[9] Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*.

[10] Paolanti, M., Romeo, L., Felicetti, A., Mancini, A., Frontoni, E., & Loncarski, J. (2018). Machine Learning approach for Predictive Maintenance in Industry 4.0. In *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications, MESA 2018*.