Predictive Maintenance and Monitoring of Industrial Machine using Machine Learning

3 authors:

Kausha I. Masani
Marwadi Education Foundation
1 PUBLICATION   28 CITATIONS

SEE PROFILE

Parita Oza
Nirma University
66 PUBLICATIONS   666 CITATIONS

SEE PROFILE

Smita Agrawal
Nirma University
72 PUBLICATIONS   721 CITATIONS

SEE PROFILE

PREDICTIVE MAINTENANCE AND MONITORING OF INDUSTRIAL MACHINE USING MACHINE LEARNING

KAUSHA I. MASANI, PARITA OZA AND SMITA AGRAWAL*

Abstract. Machine learning is one of the break-through technologies of the modern digital world. It's applications are found in various research domain such as medicine, image processing, production and manufacturing, aviation and autonomics and many more. To efficiently run a machine, it's maintenance and its monitoring automation system play key role.The major problem we are targetting is to overcome the lack of an automation system which can give accuracy rate of the production machine at a given instance of time. Also the important energy meter parameters required to make power report in automation system for addressing the production issues, at given interval of time, were also not recorded. Thus in this paper, we describe how machine learning techniques is used for prediction of accuracy of running production machine. To address this issues, we have used supervised machine learning technique of Binary decision tree using CART method and for power report, while the data is fetched using RS232 to RS485 convertor via Modbus communication protocol. Using CART we have predicted the machine accuracy at a given time with specific energy meter readings as its input features. This paper discusses the problem definition identified, data analysis of energy meter data and it's fetching and at the end ML techniques applied to predict accuracy of running production machine. In the end we prepare various power reports of the different machines from the fetched parameters as well as produce a graphical warning of deteriorate performance of the machine at a given instance of the time.

Key words: Machine Learning, Decision tree, CART, Binary Classification, Supervised Machine Learning technique, Energy Meter, ModBus Communication Protocol, Power.

AMS subject classifications. 97R40

1. Introduction. The world is moving towards fast automation, an era of the digital and advanced technology-driven world of production. The production of any product depends upon the performance of the machine used for its manufacturing process. The maintenance and monitoring of this machine hence is a vital aspect to increase the sale of the product as well as to maintain the quality of the product.

We addressed the area of machine maintenance and monitoring via a machine learning model, for a single machine of a cement production plant. In this paper, we discuss the problem area, use of a machine learning model to address the problem, selected features for the model and their criticality in detail and output of the model. The supporting idea for this research work was found in [1-16] where fault detections were made in the automatic meter reading systems, to improve scalability and reliability.

2. Problem. During measure the machine accuracy for the given set of data, specific problems may arise. There is no such system which can fetch the data from energy meters of the machine and prepare data for the ML model. Also, there is no such ML model which can predict the accuracy of the machine for the given set of the data.

Here we proposed an ML model which can predict the accuracy in which machine is running for the given set of energy meter data and predict the accuracy of the machine for the future process before it comes to breakdown point. This paper also discussed the method used to fetch the energy meter data from a remote location to the local system using RS232 to RS485 converter.

3. Introduction to ML and It's Techniques. Machine learning is fast taking over the manufacturing industry to provide prognostics for the industrial machine to increase the efficiency of the machine and to prevent machine failure. The use of machine learning and its techniques depends on the type of application and data available. Machine learning has mainly four types, i.e. supervised learning, Semi-supervised learning, unsupervised learning and reinforcement learning.

Supervised learning is the method which we are adapting to predict the performance of the pre-decided machine. Supervised learning provides a variety of techniques for classification regression analysis of the data. These techniques are Bayesian Networks, Support vector machines, Naive Bayes, Decision trees, k-NN, Neural

---
*Department of Computer Science and Engineering, Nirma University, Ahmedabad, India ({kaushamasani@gmail.com, parita.prajapati@nirmauni.ac.in, smita.agrawal@nirmauni.ac.in})

networks and so forth [2]. All these serve the purpose according to the data and nature off the applications [14,15].

Some brief explanation about these is mentioned below [2]:

- Baysian Network [2]: It has interpretation ability to define problem into the structural relationship among predictors. Due to this training time takes less computational period. The additional advantage is that there are no free parameters to be set.
  However, disadvantages for Supervised ML method is that it is unable to perform when dataset grows as well as it is unable to handle high dimensional data.
- Logistic Regression [2]: This method has been used when the dependent variable or target variable is divided. In this method, the probabilistic interpretation is excellent, and the model can be updated to take new data easily. The main disadvantage of this method is that it needs a large sample size to achieve stable results. This model has its application in the area which deals with crash types, injury severity, voters types, and so forth.
- K Nearest Neighbor [2]: k-NN is non-parametric classification algorithm. It assigns to a non-labelled point, which is the nearest class of previously labelled point. It is widely used for multi-label applications and multi-modal classes. Though it has lower efficiency, it is also addressed as a simple lazy learning method. The performance of this model all depends on selection value of 'k'. The major drawback of this method is that its sensitive to irrelevant features, affected adversely by noise, its performance also varies with varied data.
- SVM [2]: Support Vector Machines named SVM, have complex algorithms which provide high accuracy. With appropriate kernel, they can work well even if the data is not linearly separable. It avoids overfitting, has the flexibility for selection of kernels for nonlinearity. It has good generalisation ability, but it has some disadvantages also. It is very complex, along with slower training speed. Also, its performance mostly depends on the selection of parameters this method majorly used in text-classification applications.
- Decision Tree [2]: Decision tree is easy to interpret and explain. Additionally, they can easily handle the interaction between the features. Though non-parametric, outliers do not affect the model. Some famous algorithm is ID3, CART, C4.5, and C5.0 according to the various splitting data such as Info Gain, Gini Index, Gain Ratio and Gini Co-efficient. A decision tree can handle a variety of the data, missing values, redundant attributes; have good generalisation abilities; are robust to noise; provides high-performance data with very small computational data. These use divide and conquer approach, which performs accurately with highly relevant features.
  Due to these reasons, the relevancy of dataset and the availability of highly relevant features, a Decision tree is finalized as the model for the present system in this paper.

4. Proposed ML model. The proposed model adapted for this application is the Decision tree with binary tree classification. Cross-validation and entropy are the methods used for splitting the sampled data and calculating the highest variation parameter, respectively.

Dependent Variable. The dependent variable as explained in [3] determines the goal of the study; the user chooses it. In our example, power is selected as the dependent variable. Below the root node, we find the next level of the tree. Here, the tree selects variable current-phase as a predictor for the dependent variable and separates all households according to the predictor's values. The separation of data is called a split.

Selection of Particular Split. A split in a decision tree relates to the predictor with the maximum separating power. In other words, the best split does the best job in the creation of nodes where a single class dominates. In our example, power best splits the data based on its entropy, as power has highest entropy amongst all parameters. There are several methods of calculating the predictor's power to separate data. One of the best-known methods is based on the Gini coefficient of inequality [3].

Gini Coefficient. The Gini coefficient is [13],essentially, a measure of how well the predictor separates the classes contained in the parent node. Computation of the Gini coefficient is illustrated in Fig. 4.1[3]. Explaining with an example to explain Gini- coefficient more in simple form. Suppose if we compute the graph between richest to poorest members of the society, in terms of wealth, we get diagonal corresponds to an equal
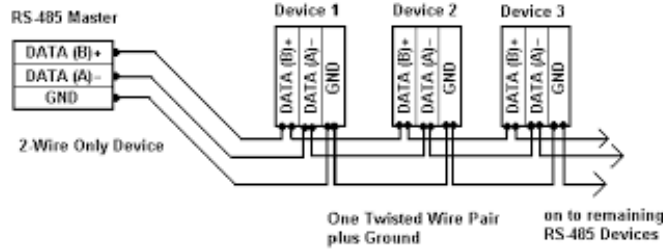
Fig. 4.1. *Example to Explain Gini-Coefficient*

Fig. 5.1. *Master-Slave Serially Connection Used In ModBus Protocol*

distribution of wealth, and the curve above it represents a real economy, where there is always some inequality in the income distribution [3].

The Gini coefficient is calculated as the area between the curve and the diagonal divided by the area below the diagonal. In Fig. 4.1 the computation of the Gini coefficient of inequality Classification and Regression Trees (CART) use the Gini's measure of inequality for selecting splits. With the help of the Gini index, it is decided which attribute is to be split and what is the splitting point of the splitting of the attribute. The Gini index is minimized at each split so that the tree becomes less diverse as we progress. The class histograms are built for each successive pairs of values of attributes. At any particular node, after obtaining all the histograms for all attributes, the Gini index for each histogram is computed. The histogram that gives the least Gini index gives us the splitting point P for the node under consideration.

The most widely used measure of impurity (at least by CART) is according [4]

$$\text{Gini index: } i(p) = \sum_{i \neq j} p_i p_j = 1 - \sum_{j} p_j^2$$

We define the impurity of a tree to be the sum over all terminal nodes of the impurity of a node multiplied by the proportion $p_i p_j$ of cases that reach that node of the tree [4]. The stopping criterion is when the Gini index at a node becomes zero, as this implies that all data records at that node have been classified completely.

5. Fetching The Data For The ML Model. The real-time data or input features used for the ML model are first being fetched from energy meters, which are located at the various remote location of the cement plant. These fetched data is then dumped into the database created for the same. The fetched data is then used for the preparation of the input file for the ML model.

The energy meter of the cement mill machine is Modbus enabled. Thus it shall pass the values over LAN using Modbus communication protocol. This protocol is a serial communication protocol, widely used in industrial machine communications. This idea for fetching the data through a convertor was sourced from inspiration from [5]. The connection diagram of a master device with a serially connected slave device is shown in Fig. 5.1.

The system build using the convertor is shown in the Fig. 5.2.

As shown in Fig. 5.2 the data from energy meter is fetched via RS232-Rs485 convertor into the developed SCADA-like system in a remote computer and then it is dumped into the database. 1440 data records per energy meter are fetched in 1 day. So the pool of data for data processing and analysis becomes large when total records are multiplied by total no of energy meters.

6. Data Pre-processing and Analysis. The features for the proposed system are selected from amongst 32 parameters, based on their criticality concerning the system. These parameters are average voltageCurrent phase-1, current phase-2, current phase-3, kWh. Below mentioned is the standard power equation for the
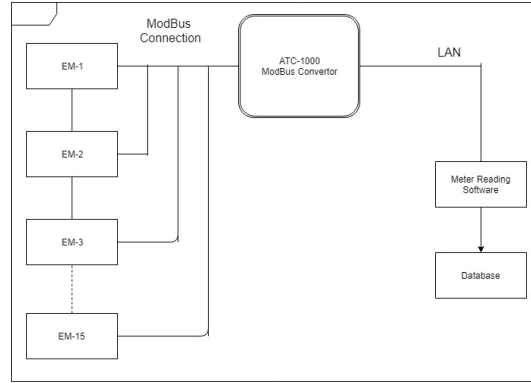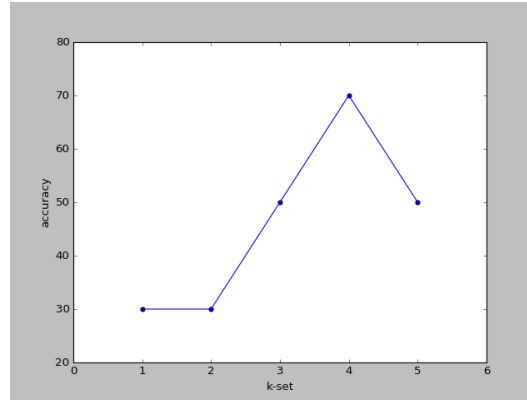
Fig. 5.2. *Block Diagram of Data Fetching Mechanism*

Fig. 7.1. *Reultant Graph Representation*

purpose to understand why power is critical and shall be having more entropy, which can be explained like:

$$(6.1) \qquad\qquad \text{Power Equation: } P = \sqrt{3} \cdot V_L \cdot I_L \cdot Cos\phi$$

where:

V_L = Voltage Line
I_L = Current Line
$Cos\phi$ = Power factor (0.8)

As observed from the above equation, power depends upon its supplied current, voltage and power factor. The power factor we are taking as a constant value. The I_L in the equation varies depending upon the load of the material machine is currently running for. Thus current phases 1, 2 and 3 are also considered equally critical along with power.

The data received is converted into a decimal form from the exponential form for analysis purpose. Once the CSV file is ready, we are ready to deploy it as an input file to the proposed model.

7. Test Results. When the program is deployed with input file in the form of CSV, initially the whole records are made in the set of 5 sets. These sets are then processed with the ML model giving the graphical representation results, as shown in Fig. 7.1.

As we observed that, the whole dataset is sampled into 5 sets where each of them giving efficiency rate individually. We can deduce that the dataset of set one has accuracy the lowest along with dataset two. However, as the workload of the machine is increased, the power and efficiency of the machine also rise as depicted in the

above graph. This shows the required result of this whole system, i.e. the maximum efficiency of the running machine before it reaches its breakdown point.

This system helps the engineers to avoid untimely breakdown of the system during peak production hours. The implementation of this system not only helps with knowing the efficiency at a given point of time for the data but also helps in making the production report of the designated department in any manufacturing unit. This system not only helps in improving the working of an industrial machine but also helps in monitoring its performance during it is running time.

8. Conclusion. Machine learning is currently taking over not only the robotics section of the technology but its also take place faster in the industrial and production sector of the manufacturing industry. The adaptation and implementation of ML techniques are now catching up the trend and is fast moving towards automation with quality production.

The main aim here is to implement one such method to predict the accuracy of the production machine, to improve the production quality as well as quantity and remaining useful lifetime of the machine. The paper also tries to provide one such novel method in the subject area of automation in the industrial machine with the help of machine learning, and it is available techniques. This outcome of the system not only shall help to prepare various power reports of the different machines but also shall be able to generate a graphical warning of deteriorating performance of the machine at a given instance of the time.

REFERENCES

[1] Y. Kou, G. Cui, J. Fan, X. Chen, W. Li, Machine learning based models for fault detection in automatic meter reading systems, 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), pp. 684-689, 10.1109/SPAC.2017.8304362

[2] A. Singh, N. Thakur, A. Sharma, A Review of Supervised Machine Learning Algorithms, 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1310–1315, 2016

[3] S. Sivagama Sundhari, A knowledge discovery using decision tree by Gini coefficient, 2011 International Conference on Business, Engineering and Industrial Applications, pp. 232-235, 2011

[4] Z. Jian, W. Zhaowei, Q. Changsong, Study on the key quality parameter decision for multi-process by CART method, 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), vol. 13, pp. V13-612-V13-616, 2010

[5] A. Kanaway, A. Sane, Machine learning for predictive maintenance of industrial machines using IoT sensor data, 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 87-90, 2017

[6] X. Zhao, M. Li, J. Xu, G. Song, Multi-Class Semi-Supervised Learning in Machine Condition Monitoring, 2009 International Conference on Information Engineering and Computer Science, pp. 1-4, 2009

[7] N. Amruthnath, T. Gupta, A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance, 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), pp.355-361, 2018

[8] L. Kovács, G. Z. Terstyánszky, Diagnosising Faults by Supervised and Unsupervised Learning, 1999 European Control Conference (ECC), pp. 4238–4242, 1999

[9] V. Mathew, T. Toby, V. Singh, B.M. Rao, M. G. Kumar, Prediction of Remaining Useful Lifetime (RUL) of turbofan engine using machine learning, 2017 IEEE International Conference on Circuits and Systems (ICCS), pp.306-311,2017

[10] S. Dhankhad, E. Mohammed, B. Far,Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study, 2018 IEEE International Conference on Information Reuse and Integration (IRI), pp. 122-125, 2018

[11] M. Balamurugan, S. Kannan, Performance analysis of cart and C5.0 using sampling techniques, 2016 IEEE International Conference on Advances in Computer Applications (ICACA), pp. 72-75, 2016

[12] L. Yifei, W. He, P. Shuai, S. Weiqiong, D. Ning, W. Fang, Application of supervised machine learning algorithms in diagnosis of abnormal voltage, pp. 1-5, China International Conference on Electricity Distribution (CICED), 10.1109/CICED.2016.7576347,2016

[13] Y. Kou, G. Cui, J. Fan, X. Chen, W. Li, Towards Statistical Modeling and Machine Learning Based Energy Usage Forecasting in Smart Grid, SIGAPP Appl. Comput. Rev., vol 15, pp. 6–16, 2015

[14] S.S. Agrawal, A. Patel, CSG cluster: A collaborative similarity based graph clustering for community detection in complex networks, International Journal of Engineering and Advanced Technology, 8 (5), pp. 1682-1687, 2019

[15] S. Agrawal, A. Patel, Clustering Algorithm for Community Detection in Complex Network: a Comprehensive Review, Recent Patents on Computer Science, 2019, 12: 1.

[16] N. Patel, P. Oza P., S. Agrawal, Homomorphic Cryptography and Its Applications in Various Domains. In: Bhattacharyya S., Hassanien A., Gupta D., Khanna A., Pan I. (eds) International Conference on Innovative Computing and Communications. Lecture Notes in Networks and Systems, vol 55, Springer, 2019

K.I. Masani, P. Oza, S. Agrawal