Machine Learning approach to identify Machine Failure

Juhi Soni
Information Technology
TH OWL University of Applied Sciences
Lemgo, Germany
juhi.soni@stud.th-owl.de

I. Motivation

Every year industries are spending millions on the maintenance of equipment. Ineffective maintenance results in machine downtime, which eventually leads to an increase in cost. The rapid advancements of AI and data availability are affecting the industries in various areas such as quality and production improvement, scheduling, and maintenance management. Machine learning/ANN techniques enabled by the growing abilities of the hardware, cloud-based solutions, and newly introduced state-of-the-art algorithms have given fruitful solutions in these areas [9]. Over the years, different algorithms, architecture, and methodologies have been proposed for maintenance management such as watchdog agent (a design enclosed with various machine learning algorithms), predictive maintenance framework, etc. [11].

With the help of the Internet of things (IoT) devices, we can not only collect the process data but also collect the data related to the physical health of the machine such as pressure, temperature, torque, tool wear, etc. Predictive maintenance with the scope of preventive maintenance is a principal technique to deal with maintenance issues. We can avoid unnecessary equipment replacement, machine downtime and find the root cause of the fault using a predictive maintenance strategy. By doing this, we can save costs and improve efficiency [9]. In predictive analytics, we analyze historical data to predict future outcomes. In maintenance, predictive analytics is used to predict the type of failure and the time to complete failure [11].

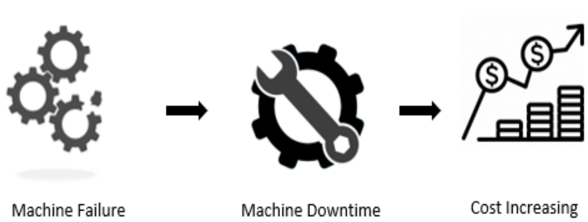

Fig. 1. Machine Maintenance Problem [18, 20, 22]

A. Problem Description

To develop predictive maintenance model using machine learning algorithms to identify machine failure. In order to build a machine failure identification model using machine learning algorithms, a predictive maintenance dataset from the UCI Machine Learning repository is used. The data set consists of 10 000 data points stored as rows with 14 features in columns. This dataset is a synthetic dataset that reflects real predictive maintenance encountered in industry.

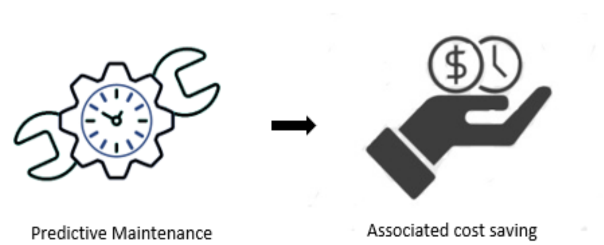

Fig. 2. Predictive Maintenance [19, 21]

B. Paper Organization

The rest of this paper is structured as follows. Section II describes State of the Art (solutions exist for the problem). Section III describes algorithms used to train predictive maintenance model (Support vector machine, Random Forest and Neural network). Results of these three algorithms are added in section IV. Finally, Section V describes the summary of problem and applied approach.

II. State of the Art

This section describes the scientific research areas in the field of Predictive maintenance. Predictive maintenance refers to the intelligent monitoring of equipment to avoid equipment breakdown and associated cost. It evolves from visual inspection to the automated process using advanced technologies like sensors, neural network, pattern recognition, signal processing, machine learning, etc [12].

A. Machine Learning

Machine learning is a part of Artificial Intelligence. It focuses on how computer systems learn and improve automatically through experience. Various machine learning algorithms are developed to cover a broad range of data and problems [6]. Machine learning has gradually improved the state of art in many different tasks such as speech recognition, computer vision, optimization, predictive analytics, etc. Machine learning is classified into three categories [11].

- **Supervised Learning:** Supervised learning algorithms build a mathematical model from a dataset that contains both the inputs and the desired outputs [6]. The goal of an algorithm is to learn a function that can be used to predict the output associated with new inputs.
- **Unsupervised Learning:** Unsupervised learning algorithms take a dataset that contains only inputs and find structure in the data such as grouping or clustering of data points. The algorithms learn from the test data that has not been labeled, classified or categorized [14].
- **Reinforcement Learning:** Reinforcement learning algorithms use a reward mechanism. The training data in reinforcement learning are assumed to provide only an indication as to whether an action is correct or not; if the action is correct, there will be a reward, and if the action is incorrect, it looks for finding the correct action [6].

Other kind of learning algorithms include Semi-supervised learning, Self learning, Feature learning, Sparse dictionary learning, Anomaly detection and Association rules [14].

B. Machine Learning Approach For PDM

Machine learning method decision forest has been implemented on a real cutting machine woodworking machinery that is a machining center on the wood industry [12]. They used 30% of collected data as a training set to train a machine learning model using the random forest[3] method to estimate the spindle health status. Results show 95% accuracy on a real world data set of 530731 data readings with 15 different features.

A multiple classifier PdM, which uses SVM [1] and k-NN classifiers for integral type faults detection has been demonstrated for a semiconductor manufacturing Ion-Implanter-related maintenance task [9]. SVM and k-NN classifier approach has proven effective in minimizing cost compared to the classical PvM techniques. In this case, SVM offers superior performance to k-NN classifiers with better accuracy, precision, and recall.

A machine learning approach to predict equipment failures in advance and improve rail network velocity by reducing derailments or reducing intermediate maintenance calls due to false alarms using a vast amount of historical data, in combination with failure data, maintenance action data, inspection schedule data, train type data, and weather data [4]. SVM classifier is trained with RBF kernel for bearing alarm prediction, and TPR and FPR are used for SVM model evaluation. Equipment failures cost significantly to railway operations. This solution is capable to save several thousand to hundreds of thousands of dollars depending on alarm location and traffic [4].

Identifying fault in equipment is one of the most integral components of predictive maintenance [11]. Exhaust fan vibration data, PCA, hierarchical clustering, K-means, and C-means algorithms are used to train machine learning model for abnormality detection. The result shows that this model has successfully identified the normal, warning, and healthy state of the machine. It indicates that clustering algorithms are indeed a better tool in fault detection in a machine.

A decision tree classifier is used to build a predictive maintenance model using a synthetic dataset to predict machine failure and its type. The decision tree provides high-quality results, but it fails to detect a considerable number of cases [15].

For Predictive maintenance of nuclear infrastructure, machine learning algorithms SVM and logistic regression are used to train a maintenance model using data collected from various sensors from nuclear power plant equipment. This maintenance framework provides high accuracy with no overhead and can identify a machine failure [16].

Binary logistic regression model to classify machinery conditions using oil samples from a diesel engine, this model classifies data into four classes and compared it with SVM and ANN [17]. Unsupervised ML techniques local outlier factor (LOF) and clustering-based outlier detector are used to detect the location of faulty cells within the electric vehicle batteries [10]. This approach was the most robust and they used data from a vehicle fleet to detect frequencies of faults. They used incremental learning approach to classify known and unknown fault types using ensemble's base classifiers algorithms for Automotive Systems [8]. It allows to adapt new fault types without re-training the entire model.

C. Neural Networks Approach For PDM

Neural-network-based decision support system developed for predictive maintenance of rotational equipment. This system aimed to minimize operational costs. In this approach, an artificial neural network model is trained to estimate the life percentile and failure times of roller bearings [2].

In the oil and gas industry equipment failure can be predicted using predictive maintenance to prevent major loss using Long Short-Term Memory (LSTM) networks [13]. This model is trained on data obtained from real-world plant and several simulations of the model are performed to get the optimum parameters of the model for higher accuracy of the prediction [13].

III. Solutions

Support Vector Machine, Random Forest, and Neural network algorithms are used to develop a Predictive Maintenance model to identify machine failure.

A. Predictive Maintenance Dataset

Predictive Maintenance Dataset contains 14 features. The below list describes all features, and among five are the independent variables.

1) UID: unique identifier ranging from 1 to 10000.
2) Product ID: consisting of a letter L, M, or H as product quality variants and a variant-specific serial number.
3) Air Temperature: between 295 to 305k.
4) Process Temperature: between 304 to 314k.
5) Rotational Speed: between 1160 to 2886rpm.
6) Torque: between 3.8 to 76.6Nm.
7) Tool wear: 5/3/2 minutes of tool wear.
8) TWF: Tool wear failure.
9) HDF: Heat dissipation failure, if the difference between air- and process temperature is below 8.6 K and the tool rotational speed is below 1380 rpm [15].
10) PWF: Power failure, if this power is below 3500 W or above 9000 W, the process fails, the product of torque and rotational speed equals the power required for the process [15].
11) OSF: Overstrain failure, if the product of tool wear and torque exceeds 11,000 minNm [15].
12) RNF: Random failures.
13) Machine failure: whether the machine has failed in this particular datapoint for any of the above failure modes are true [15].

Total six features (five independent variables and one dependent variable) out of 14 features from the dataset are used to train model. In dataset a Machine Failure label indicates, whether the machine has failed or not. Below figures show the histogram plot of all six variables. From the plots, we can see that there are no abnormalities in the data.

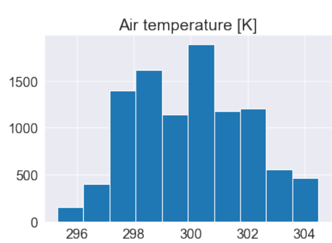

Fig. 3. Air temperature

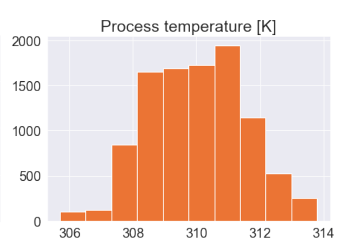

Fig. 4. Process temperature

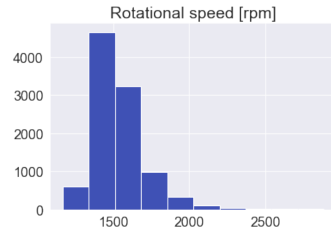

Fig. 5. Rotational speed

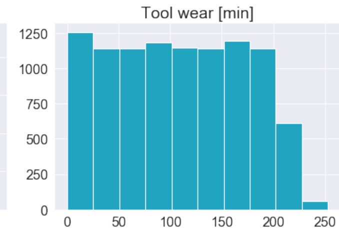

Fig. 6. Tool wear

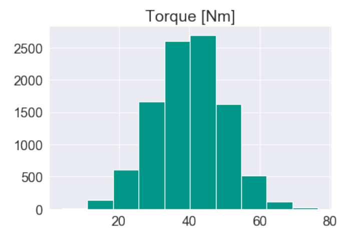

Fig. 7. Torque

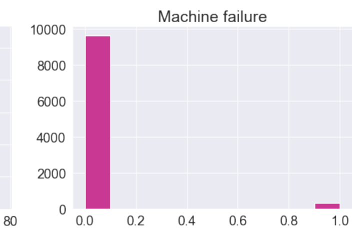

Fig. 8. Machine Failure

B. Support Vector Machine

A support vector machine is a supervised learning algorithm that finds the decision boundary to differentiate two different classes of objects. It is a linear or non-linear classifier that separates the two classes of data points. For non-linear data points, it uses the kernel function to find a decision boundary between data points.

Formalization: Algorithm take an array of data as input.

$$(N, K) \rightarrow \sum_{i=1}^{n} \alpha_i \; y_i \; K(x_i, x) + b \rightarrow f(x)$$

N is training data, K = kernel function

Algorithm 1: Support Vector Machine

Input: Data N
Output: Identify Different Classes of Object

Function SVC (X, y, k) **:**
  V $\leftarrow \phi$;
  for $x \in X$ **do**
    x_k = k(x);
    v = v - x_k ;
    $V \leftarrow V \cup v$;
  end
  return V
End Function;

C. Random Forest

A random forest is an ensemble learning algorithm that builds small decision trees with few features. Each tree in the random forest splits data points into two different classes and the average of these decision trees becomes the prediction model.

Formalization: Algorithm take an array of data as input.

(N, n, e) $\rightarrow$ $\frac{1}{n}$ $\sum_{i=1}^{n}$ f_n(x, e) $\rightarrow$ f(x)

N is training data, e is entropy, n is no. of decision tree

Algorithm 2: Random Forest

Input: Data N

Output: Identify Different Classes of Object

Function RF (X, y, n, e) **:**
  DF $\leftarrow \phi$;
  for $x \in X$ **do**
    $x_i \leftarrow BTS(X)$;
    $f \subset F$;
    $d_i \leftarrow f_n(x, f, e)$;
    $DF \leftarrow DF \cup d_i$;
  end
  return DF
End Function;

D. Neural Networks

A neural network is a network of neurons whose architecture is based on the brain. In a neural network, the weight reflects the connection between neurons. ANN is invented to solve artificial intelligence problems like image analysis, pattern recognition, predictive modeling, adaptive control where a neural network can be trained via dataset [5]. It learns from the data and performs actions based on experience and learning.

A neural network contains an input layer, single or multiple hidden layers, and output layer. Although there are networks available that consist of a single layer, or even one element [5]. To train a neural network, input data is passed on to an input layer where initial weights are chosen randomly. The input layer passes it to hidden layers and the hidden layer forwards it to the output layer. Each of these layers of neurons is making a decision by weighing up the results from the previous layers, and the output layer gives results using a bias value [7].

Formalization: Algorithm take an array of data as input.

(N, hl, a) $\rightarrow$ f (b + $\sum_{i=1}^{n}$ $x_i w_i$) $\rightarrow$ f(x)

N is training data, a is activation function, hl is no. of hidden layers

Algorithm 3: Neural Network

Input: Data N

Output: Identify Different Classes of Object

Function ANN (X, y, hl, f) **:**
  for $i \in hl$ **do**
    for $x \in i$ **do**
      y_i = f(b + $x_i w_i$) ;
      δ = y(x) - y_i ;
      update w_i and b ;
  return $Optimalclassifier$
End Function;

E. Implementation

To train the Predictive Maintainance model, first split the dataset into training and testing data. Air temperature and process temperature are defined in kelvin, the rotational speed is defined in rpm (Revolutions per minute), torque is defined in Nm (Newton-meters), and tool wear is defined in a minute. So, to get the data points on the same scale, applied feature scaling on all five features to avoid dominance between features. After feature scaling, support vector classifier, random forest, and neural network models are trained with training data. The confusion matrix and accuracy value is used to evaluate these models.

Algorithm 4: Predictive Maintenance model

Input: Data N

Output: 1/0 Identify Machine Failure

Function CA (Y_actual, Y_pred) **:**
  $acc = \frac{TP+TN}{TP+TN+FP+FN}$;
  return acc
End Function;

$D \leftarrow N$;
$X \leftarrow D[IV]$;
$y \leftarrow D[DV]$;
$X\_T, X\_t, y\_T, y\_t \leftarrow splitData(X, y, 0.25)$;
for $x \in X\_T$ **do**
  $x \leftarrow \frac{x-\mu}{\sigma}$;
end
$SC \leftarrow SVC(X\_T, y\_T, k)$;
$RC \leftarrow RF(X\_T, y\_T, n, e)$;
$AN \leftarrow ANN(X\_T, y\_T, hl, f)$;
$Y_{svc} \leftarrow SC(X\_T)$;
$Y_{rf} \leftarrow RF(X\_T)$;
$Y_{ann} \leftarrow AN(X\_T)$;
$AC_{sv} \leftarrow CA(y\_t, Y_{svc})$;
$AC_{rf} \leftarrow CA(y\_t, Y_{rf})$;
$AC_{ann} \leftarrow CA(y\_t, Y_{ann})$;

IV. Experiments and Results

A. Experiments

This dataset was first trained with a linear support vector machine. This linear support vector model has given lower accuracy (0.9648) and a higher error rate (0.0352). So, to improve the model performance dataset trained with RBF kernel, which has given lower error rate and higher accuracy compared to a linear model. The neural network model was first trained with 15 hidden layers and ReLU (Rectified Linear Unit) activation function, and it gives 0.978 accuracy and a 0.22 error rate. Later it trained with 150 hidden layers in order to increase the accuracy of a model, but it gives 0.98 accuracy and 0.2 error rate. So, an increased number of hidden layers haven't much improved the performance of a model.

B. Results

In this section, the result of the three models is described. Fig. 9 depicts a confusion matrix of the support vector model. This model has correctly predicted 2434 (true positive + true negative) data points out of 2500 data points.

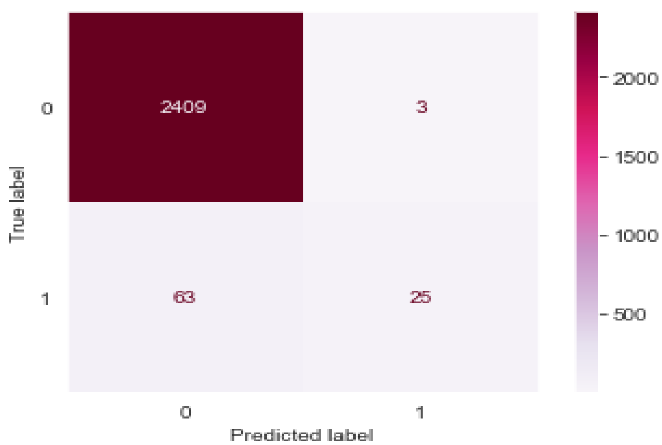

Fig. 9. SVM confusion matrix

Fig. 10 displays the confusion matrix of random forest. 2402 and 54 matrix values indicate the true positive and true negative value and 10 and 34 indicate the false positive and false negative value.

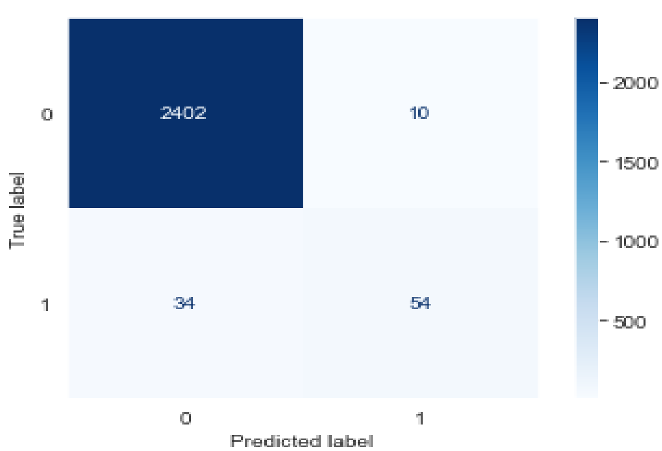

Fig. 10. Random Forest confusion matrix

Fig. 11 shows the plot of the confusion matrix of the neural network. This model has correctly predicted 2445 (true positive + true negative) data points and 55 (false positive + false negative) data points are incorrectly predicted from 2500 data points.

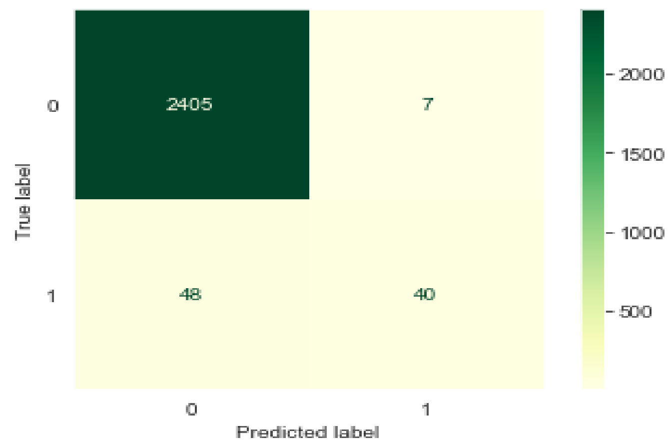

Fig. 11. Neural Network confusion matrix

The accuracy and error rate of all three models (Support Vector Machine, Random Forest, and Neural Networks) are shown in the below table. Random Forest has performed well among all three algorithms for this dataset. It has the highest accuracy and lowest error rate compared to the other two algorithms. On the other hand SVM has a lowest accuracy and highest error rate.

Table-1

Evaluation method	**SVM**	**RF**	**ANN**
Accuracy	0.9736	0.9824	0.978
Error Rate	0.0264	0.0176	0.022

V. Summary and Outlook

In this paper, support vector machine, random forest, neural network model are trained with predictive maintenance dataset to identify machine failure. From the dataset five features air temperature, process temperature, rotational speed, torque, and tool wear are used as independent variables, and machine failure is used as a dependent variable. Feature scaling method is used to normalize the range of independent variables. Data is split into training and testing set in order to evaluate the machine learning and neural network algorithm.

Confusion matrix, accuracy, and error rate are used to evaluate all three models. From the results, we can see that random forest has performed well among all with a testing data set. It has high accuracy and a low error rate compared to the other two models. All three models are able to predict the failure of the machine, but random forest provides higher accuracy with no performance overhead.

References

[1] Vikramaditya Jakkula. “Tutorial on support vector machine (svm)”. In: *School of EECS, Washington State University* (2006).

[2] Sze-jung Wu et al. “A Neural Network Integrated Decision Support System for Condition-Based Optimal Predictive Maintenance Policy”. In: (2007).

[3] Thais Mayumi Oshiro, Pedro Santoro Perez, and José Augusto Baranauskas. “How many trees in a random forest?” In: *International workshop on machine learning and data mining in pattern recognition*. 2012.

[4] Hongfei Li et al. “Improving rail network velocity: A machine learning approach to predictive maintenance”. In: *Transportation Research Part C: Emerging Technologies* (2014).

[5] Sonali B Maind, Priyanka Wankar, et al. “Research paper on basic of artificial neural network”. In: *International Journal on Recent and Innovation Trends in Computing and Communication* (2014).

[6] M. I. Jordan and T. M. Mitchell. “Machine learning: Trends, perspectives, and prospects”. In: *Science* (2015).

[7] Michael A Nielsen. *Neural networks and deep learning*. Determination press San Francisco, CA, 2015.

[8] Chaitanya Sankavaram et al. “Incremental Classifiers for Data-Driven Fault Diagnosis Applied to Automotive Systems”. In: (2015).

[9] Gian Antonio Susto et al. “Machine Learning for Predictive Maintenance: A Multiple Classifier Approach”. In: *IEEE Transactions on Industrial Informatics* (2015).

[10] Yang Zhao et al. “Fault and defect diagnosis of battery for electric vehicles based on big data analysis methods”. In: (2017).

[11] Nagdev Amruthnath and Tarun Gupta. “A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance”. In: *2018 5th International Conference on Industrial Engineering and Applications (ICIEA)*. 2018.

[12] Marina Paolanti et al. “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.

[13] Tayaba Abbasi, King Hann Lim, and Ke San Yam. “Predictive maintenance of oil and gas equipment using recurrent neural network”. In: 2019.

[14] Wikipedia contributors. “Machine learning, The Free Encyclopedia”. In: (2020).

[15] Stephan Matzka. “Explainable Artificial Intelligence for Predictive Maintenance Applications”. In: IEEE. 2020.

[16] “Predictive maintenance architecture development for nuclear infrastructure using machine learning”. In: *Nuclear Engineering and Technology* (2020).

[17] Andreas Theissler et al. “Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry”. In: (2021).

[18] Clipartmax. URL: https://www.clipartmax.com/download/m2H7i8K9N4Z5m2Z5_helicopter-maintenance-repair-icon-png/.

[19] Comarch. URL: https://www.comarch.com/artificial-intelligence-management/predictive-maintenance/.

[20] Iconfinder. URL: https://www.iconfinder.com/icons/3756117/charge_cost_earning_expenditure_expense_increase_icon.

[21] Imgbin. URL: https://imgbin.com/png/3BsEZh2w/cost-reduction-saving-computer-icons-money-png.

[22] Istockphoto. URL: https://www.istockphoto.com/de/grafiken/system-failure-icon.