

Machine Learning-Based Abnormality Detection Approach for Vacuum Pump Assembly Line

Paras Garg¹, Amitkumar Patil¹, Gunjan Soni^{1*}, Arvind Keprate², Seemant Arora³

¹Malviya National Institute of Technology, Jaipur

²Oslo Metropolitan University, Norway

³Bosch Ltd., Italy

2019pie5423@mnit.ac.in

2019RME9109@mnit.ac.in

*gsoni.mech@mnit.ac.in

arvindke@oslomet.no

arora@it.bosch.com

Abstract

The fundamental basis of Industry 4.0 is to make the manufacturing sector more productive and autonomous. In the manufacturing sector, practitioners always long for product quality improvement, reducing reworking costs, enhancing first pass yield in production or assembly line, in this regard anomaly detection, is becoming popular and is widely used. With the integration of anomaly detection models and Artificial Intelligence-based condition monitoring systems, industries have attained promising results in achieving these goals. However, it is a highly complex task to extract meaningful information from the large amount of data generated by the manufacturing systems. Hence, in this paper, effective machine-learning-based anomaly detection and prediction model has been proposed. A two-phase model is presented in this study and is validated for abnormality detection in the assembly line of vacuum pumps. In the first phase, a random forest algorithm is used to predict the pump vacuum. Based on the actual and predicted values, the error is computed. Then, EWMA (Exponentially Weighted Moving Average) chart is employed to detect the anomalies. In the second phase of the proposed model, based on the EWMA chart and calculated error, anomaly prediction is done. For better prediction results, statistical features are extracted from the error values and used as input for the second phase. To validate the proposed approach, other machine learning models SVR, Decision Tree, Logistic Regression, KNN and SVC have been compared. A statistical method EWMA chart is also integrated with random forest.

Keywords: Abnormality detection, Abnormality Prediction, Machine learning, process control, Assembly line, mechanical vacuum pump.

I. Introduction

New technologies have been adopted by the industrial industry as quality measurement tools, resulting in a data-rich environment that lays the groundwork for the use of machine learning (ML) techniques for data collected from various sources to reduce production costs and improve product quality [1]. Assembly machines are considered one of the vital components in the manufacturing

sector and are frequently employed in production lines; each machine is responsible for guaranteeing the end product's quality [2].

Anomaly detection technique is used in many areas like healthcare, computer security, credit card image processing, and much other application [2], [4].

The following are types of anomalies:

- a. Point anomalies: If one piece of data differs significantly from the rest, it is considered anomalous. Detecting credit card fraud based on "amount spent" is a business use case [5].
- b. Contextual anomalies: When data behave contextually different from the whole dataset is called a contextual anomaly, for example, if water consumption in summer is more than the rest of the year. Contextual anomaly is also known as aperiodic anomaly. Collective anomalies: A group of data samples working together to find abnormalities is known as collective anomalies.
- c. Business use case: Someone tries to copy data from a remote workstation to a local host without permission, which would be identified as a possible cyber assault [5]. Reducing the reworking cost and enhancing the first pass yield, the false alarm can cause waste of capital, time, and productivity, so a machine-learning-based anomaly detection algorithm would work best to improve accuracy and prevent failure. The objectives addressed in this study are depicted in Table 1.

Table 1: Objectives

Model	Purpose	Input features	Output
Vacuum performance prediction	Vacuum prediction and construction of EWMA chart based on the error term	X1-x14	Y: minimum vacuum created by the pump
Abnormality prediction	Prediction of anomaly by taking EWEWMA chart as input		Y: 0 means abnormal 1 means normal

II. Literature Review

Anomaly detection has been a center of many studies and research articles throughout the years, as per Xu et al. in [2], it is a major field. As a result, various methods and techniques for identifying anomalies and improving the effectiveness of current strategies for detecting anomalies have been provided in the existing literature. This section will examine the most cutting-edge approaches in the area. Numerous approaches were presented in [2], with the primary objective of determining the relative efficacy of techniques for extremely large unlabeled benchmark datasets. The difficulty comes when we have to find or draw the boundary between normal and abnormal data. X.liu et al. developed an approach that employed an ensemble learning algorithm to find abnormalities in the exceedingly difficult dataset. Anomalies are discovered using the Angle-based methods k-Nearest Neighbors Outlier Detection which is combined to form the input to the local outlier factor (LOF). The authors used Rank power, precision, and area beneath roc curve characteristics to quantify their model's performance in order to discover various models would deliver the best results. Fast ABOD (-angle-based outlier detection) and KNN had the best area under the receiver operating characteristic curve (AUROC) performance in the dataset, recognizing up to 75% of real positive anomalies. The authors of [7] also employed a large-scale, high-

dimensional data collection method. They also ran across the same problems as those detailed in [8]. Researched to have a better understanding of the many types of classification algorithms that might be used for deep intrusion detection. Finally, they proposed combining a one-class SVM (OC-SVM) with a deep neural network (OC-NN) model to create a hybrid model [9]. Local outlier factor and isolation forest model also perform well on -large-scale data, according to [10] and [11]. Additionally, they utilized Precision, Recall, and F1-score as performance metrics in addition to F1-score, except for Galante, who included an extra approach, One-Class Support Vector Machines (OCSVM) as well as AUROC as a statistic. Additionally, this group of test participants ran into the same problem while working with high-dimensional space, which was rectified using PCA. We obtained a 64 percent F1 score by extracting new characteristics that better properly characterize the data.

The authors of [6] compared 10 distinct benchmark datasets using several unsupervised methods. They computed seven sets of outlier scores using -one-class SVM, KNN, CBLOF, and HBOS. Additionally, they had to consider how they determined the boundary between anomalous and non-anomalous data values. The authors conducted tests to evaluate these approaches' performance on a variety of datasets of varied sizes and in high-dimensional space. Additionally, they employed AUROC as a performance metric, along with Precision and Rank Power. To do this, it is observed that a dataset with more no of variables AUROC is better compared to low no. of variables in the case of KNN and LOF. They discovered abnormalities in the network traffic dataset when they utilized cluster-based anomaly detection algorithms such as NKICAD, K-means, CBLOF, and LDCAF [7]. The authors used an unclassified dataset but used statistical analytic approaches to produce labels for their algorithms. To begin, the authors validated their methodology by researching a diverse group of benchmark datasets. They then applied this approach to their dataset of network traffic. On average, these cluster-based methods can identify 87 percent of network traffic abnormalities. In [8], OCSVM, LOF, and Random Forests were used to detect abnormalities. Thus, the author in [8] stated that their dataset is free of anomalies that would render it non-specific in high-dimensional data, resulting in an AUROC of 70%. That is, they employed semi-supervised anomaly detection to train their models but did so contrary to the findings in [9], [6]. Additionally, the authors in [10] employed similar approaches, but with the addition of the following metrics: F-score, Recall, and Precision. Simultaneously, the method was reported to have a high rate of detection and a low incidence of false positives. Unsupervised approaches include cluster analysis and classification inside a single class, as well as autoencoders. Additionally, unsupervised learning techniques such as the convolutional autoencoder (which has been used to identify faulty components in the manufacturing industry) and the generative adversarial network which has been adopted in the manufacturing sector for identifying defective components. The major approaches for detecting time series anomalies are recurrent autoencoder (RAE) [11] and one-class SVM [12]. When abnormal conditions arise, these models are taught to recognise the process's true state, and any deviation from that state is considered abnormal. This product has applications in finance [13], security [14], and healthcare [15]. When it comes to anomaly prediction, the problem is far more challenging, and little effort has been made to address it. There is a suggestion to anticipate energy price anomalies using the pattern sequence forecasting (PSF) technique. A wavelet-based multi-feature classification model and extreme learning machine technique were developed for predicting the possibility of a stock market anomaly happening. Later in the petroleum industry, a nonlinear adaptive ELM (also known as an adaptive sparse dynamic window filter) and a wavelet transform-based method was employed. Compared to other models, the experimental findings indicated that the forecast of the anomaly was improved [16]. The majority of the evaluated studies assessed just the outcomes of previous research on run-to-run identification of aberrant patterns in SMT using standard Shewhart control charts, this shows that little scientific research has been conducted on

prognosis techniques. The proposed research employs advanced machine learning techniques to develop a framework for the prediction of aberrant assembly line trends. This approach contributes to increased overall assembly line productivity while also lowering the cost of fault repair.

III. FRAMEWORK AND MODELLING

I. Framework

For this problem, a two-phase model is proposed as shown in Figure 1. A mechanical vacuum pump assembly line dataset is used in this study. For confidentiality purposes, the name of the industry is not revealed here. In this work, a two-phase model is proposed, first, the vacuum created by the pump is predicted and the error term between actual and predicted vacuum is calculated. Then an EWMA chart of prediction error is constructed. This EWMA chart is used as input for the second phase of the model which is the abnormality prediction phase.

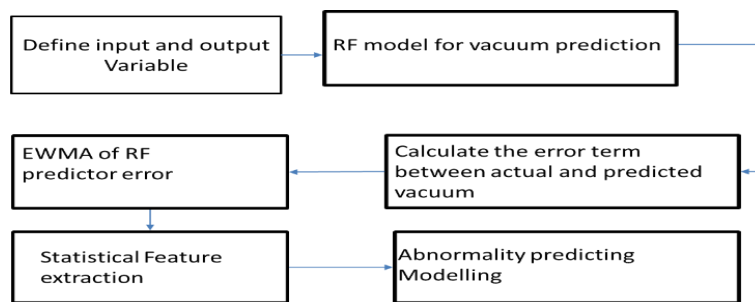


Figure 1: Framework of proposed methodology

The adopted dataset has a total of six workstations. For the development of the vacuum prediction model, all the variables from workstation 1 to workstation 5 has been taken as input for our model, and the minimum vacuum collected in workstation 6 is taken as the output for our model. The random forest-based algorithm is used as a model for this study and the further error term is calculated between the predicted and actual vacuum, thereafter based on the error term, an EWMA chart is constructed. This EWMA chart is hereby used as the input for abnormality prediction modeling. To produce the best results, the frame size is set to nine observations in this work. To do so, the predicted error observations are divided into chunks of nine observations and the statistical feature are extracted. Normal and Abnormal data shows different features. Many statistical features are extracted from each set of observations and these extracted features are used as the input for abnormality prediction modeling as proposed in the framework. The list of statistical features is given in Table 2.

Table 2: Statistical features

Features	Calculation formula
----------	---------------------

Mean	$X_{mean} = \frac{1}{n} \sum_{i=1}^n x_i$
Variance	$\sigma^2 = \sum \frac{(x - \mu)^2}{N}$
Skewness	$X_{skew} = \frac{1}{n(X_{std})^3} \sum_{i=1}^n (x_i - X_{mean})^3$
Kurtosis	$X_{kurt} = \frac{1}{n(X_{std})^4} \sum_{i=1}^n (x_i - X_{mean})^4$
RMS	$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{x}_i - x_i)^2}{n}}$
Maximum	$X_{max} = \max(x_i); (i = 1, 2, 3 \dots, n)$
Minimum	$X_{min} = \min(x_i); (i = 1, 2, 3 \dots, n)$
Coefficient of Variation	$CV = \frac{X_{std}}{X_{mean}} \times 100\%$

II Random forest model:

Ensemble learning methods are used for supervised learning using Random Forest Regression. When using this ensemble learning, several machine learning models are combined to yield a more accurate prediction than a single model. Even though the trees don't interact with each other, we can realize that they line up in parallel. The Random Forests construct multiple decision trees and the mean of all of the classes is used as the prediction of all of the trees. The steps of the Random Forest algorithm are discussed briefly as follows:

1. Pick randomly from the training set the k data points you would like to use.
2. Use this dataset to build a decision tree.
3. You must decide on the number of trees you want to build, and then you must repeat the first two steps.
4. Make four of your N-tree trees predict the value of y for the new data point and average their predictions.

Table 3: Pseudo Random Forest Algorithm

Algorithm 1: Random Forest for Regression and classification

1. For a = 1 to A, the following is true:
 - (a) From the training data, create a bootstrap sample Z of size N.
 - (b) Fit the bootstrapped data to a random-forest tree T a by iteratively repeating the following procedures for each terminating node of the tree until the minimal node size n min is attained.
 - i. At random, choose m variables from the predictor variables.
 - ii. Choose the optimal variable/split-point from them.
- iii. Dividing the node into two child nodes

2. Produce the tree ensemble $\{T_a\}_1^a$

To create a forecast about a new point x , use the following syntax:

Regression: $\hat{f}_{rf}^A(x) = \frac{1}{A} \sum_{a=1}^A T_a(x)$

Classification: Let $\hat{C}_a(x)$ be the class prediction of the a_{th} random-forest tree.

Then $\hat{C}_{rf}^A(x) = \text{majorityvote}\{\hat{C}_a(x)\}_1^A$

II. RF-Based EWMA Chart Construction Phase

EWMA control based on RF regression in the first phase Chart was developed. The independent's simultaneous impacts minimum vacuum variables generated by the pump are considered. The process is shown in the figure. 4. To begin, first define the input and output variable with the help of the design of the experiment (DOE). Then a nonlinear regression model is precisely suited. X through Y are derived from the set of remarks (X_i, Y_i) , and the mathematical model defines the real link between X and Y.

$$Y = X\beta + \delta \tag{1}$$

Where Y is the minimum vacuum generated by the pump and X is a vector representation of a set of k control variables is a vector with k + 1 parameters that reflects the actual value. Δ represents the error term in a relationship between variables that are distributed regularly and independently regularly and independently with a nearly uniform distribution. The standard deviation is σ and the mean is 0. Nevertheless, observing the entire population and determining the truth is tough. As a result, the term is calculated based on a set of assumptions. By decreasing the total, we may obtain a set of data that is representative. Equation 2 represents the computation of β .

$$\min_{\beta} \sum_{i=1}^n \delta_i^2 \tag{2}$$

$\delta = Y - X\beta$ by definition, as a result of equation (1).

The equation is differentiated with respect to and the resulting is set to zero to minimize the expression in (2).

We can discover the value by solving that, and we can use equation (3) to forecast the value of Y for minimal vacuum by solving that.

$$\hat{Y} = f(X, \hat{\beta}) \tag{3}$$

Here Y and X are the predictors, respectively. After computing Y and utilizing the residual of a subsequent matched observation occurring at time t, the formulas error = $\hat{Y}(t) - Y(t)$ are used to compute X (t) and Y (t). Errors are charted using the EWMA control chart of error terms. If the sample statistic plotted deviates from the bounds, the process has gotten out of hand. Even if all samples are controlled in a variety of real-world circumstances, considerable variance is still a possibility. The patterns on the chart, in general, do not follow a random pattern.

This is quite advantageous, as it provides critical diagnostic information. When you observe abnormality patterns, it indicates that the process is not proceeding as planned and that it must be restored to regular operating circumstances.

IV. Control chart showing the projected error term using EWMA

Regression control charts are designed to track dependent variables. This is typically done while the operation is being performed. Numerous more process variables might have a significant influence on the output concurrently. The EWMA control chart was developed to increase control precision. Shewhart control charts constrained them by failing to account for minor and moderate changes. Since they employ information from the most recent observation, the estimate's accuracy is less likely to change over time. When you notice a slight shift, the EWMA chart records the change. Most current observations, as well as the most recent data evidence from the past, has greater weight while explaining the process. FirstEWMA regression control chart must be produced. The first step of process function \hat{f} , is estimated and then used to forecast \hat{Y} , whenever you wish Next, design an EWMA control the error term vector, seen here as a chart, utilizes exponential smoothing. The equation should look like this. When the stage is done, the function f has moved, indicating that the process it previously represented no longer exists, and the error term has been shifted to minimize errors.

$$Z_i = \lambda R_i + (1 - \lambda)Z_{i-1} \quad (4)$$

Where R_i is the i th observation residual time $I = 1, 2, \dots, n$) and λ is the exponential smoothing constant that we determine; a bigger value of implies that the most recent observation has a stronger influence.

The EWMA chart is created under the premise that the error term follows a normal distribution, although EWMA is insensitive to this assumption.

V. Abnormality Prediction Phase using EWMA chart of error as input

In the second step, a KNN classifier is used to anticipate abnormalities. The input is the produced by EWMA chart. The duration of observation is set at nine in this study to achieve the best results. Normal and abnormal data have distinct characteristics. Numerous statistical characteristics are collected from each set of observations, and these recovered features are utilized as input for the framework's suggested abnormality prediction modeling.

KNN Classifier Algorithm: KNN assumes that items identical to the training examples exist nearby. What I want to say is that similar things are nearby. A part of what we need as children are used to finding the distance between points on a graph with KNN. There are several methods of measuring distance, and it may be useful to use a different one depending on the problem we are trying to solve. Although the straight-line distance (also known as the Euclidean distance) is prominent and readily understood, it is not always the best choice.

Table 4: KNN Classifier Algorithm

Algorithm 2: KNN Pseudocode

1. Fill the database with data.
2. When you have chosen the number of neighbors, initialize K to that number. For each piece of data, there is an example in the data.
3. To find out how far apart two data points are, you must use a formula that contains the distance between them.
4. In an ordered collection, add the distance and the index of the example.

5. Use the shorter distances as the start and progressively use the longest distances as the finish to get an ordered list of distance and index values (in ascending order).
6. Choose the first -K-sorted items, see which of the selected K items have their labels.
7. Return the mean of the K-labeled variables if regression is required.
8. If you wish to return the mode of the K labels, classify the items.

The process of selecting a proper value for K: When we run the KNN algorithm several times with different values of K, we pick the K that provides the best results with the fewest mistakes while also allowing the system to perform properly with new data that it has never seen before.

IV. Experimentation and Results

I. Data Description: The data obtained from six workstations consists of two types of variables. The first one is assembly parameters like speed, torque, angle at which different components of the mechanical pump are mounted and the second type of variable is the functional variable which is a test variable used to check the performance of mechanical vacuum pump like the pneumatic test, vacuum test, leakage test pressure. Figure 2 is the typical flow chart that is followed in this work.

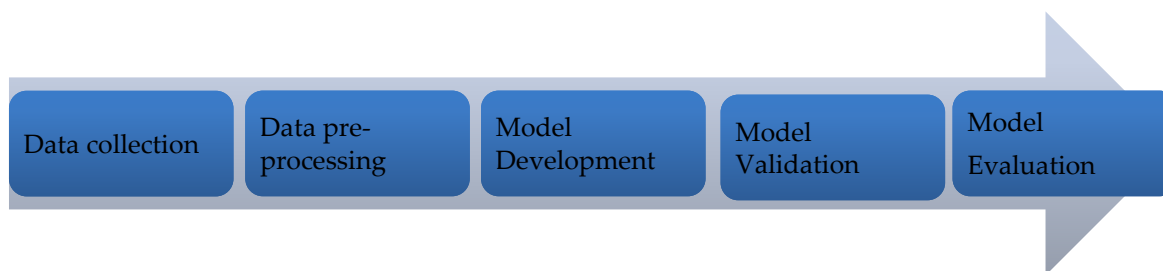


Figure 2: Flow Chart

II. Data Pre-Processing: Dataset preparation is required for machine learning algorithms to utilize the dataset. Duplicating datasets, redistributing data, and organizing the data, high values, distinctive characteristics, measurements. Feature scaling and duplicate removal processes are also performed in the data preprocessing step. The dataset used in this study consists of 40 columns and 25000 rows. Features of some of the columns are described in the Table 5.

III. Vacuum performance prediction model: A machine learning model is executed on the given data and compared the performance of other models to determine which one was the best. The algorithms that were chosen are Random forest, SVR, and Decision Tree. To test the consistency, multiple runs are performed for each algorithm. According to the authors of [18], anomaly detection models are prone to failure. Due to its uncontrolled nature, it suffers from volatility. Additionally, Unsupervised learning, as discussed in Section II examination of comparable works. The accuracy of several models is seen in Table 6. Random forest with hyper parameter optimization is the underlying model with 70.48 percent accuracy. SVR is the -second-best model for vacuum prediction in this dataset.

Table 5: Description of Dataset

Feature	Type	Description
Last Station	Numeric	This tells about the last station of pump
Entry Time	Time	Time of starting assembly
Exit Time	Time	Time of completing assembly
Cycle Time	Numeric	Time is taken to complete assembly and testing of

one pump		
Final outcome Ok	Binary	It tells the pump working properly
Final outcome Not Ok	Binary	It tells the pump is not working properly
Piece Reworked	Numeric	How many pieces reworked
Workstation	Numeric	At which workstation a pump is assembled
Couple	Numeric	Torque at which screw is inserted
Angle	Numeric	The angle at Which Screw is inserted
Pressure	Numeric	To check leakage
Flow rate	Numeric	Flow rate at which air is leaking from the pump

Table 6: Prediction accuracy of different models

Sr.no	Model	Accuracy	Standard Deviation
1	Random Forest	70.48%	3.40%
2	Decision Tree	41.23%	6.56%
3	SVR	68.86%	3.75%
4	RF-Regressor (n=14)	68.78%	3.76%
5	PCA+Regressor(n=5)	68.66%	3.48%
6	PCA+Regressor(n=7)	66.24%	3.27%

IV.RF-based EWMA chart: This complicated link between the minimum vacuum and the specified independent variables is captured by the RF predictor. Fivefold -cross-validation is performed to test the model's effectiveness. Several RF prediction models can have a higher probability of correct identification (also known as index measures or indicators) performance-related measures, like mean-absolute error (MAE). In mathematical terms, they are equal to (5), (6), and (7), respectively. R^2 Indicates how well the variation in the dependent variable (minimum vacuum) is explained by the factors specified. To say that the model fits exactly with the data is to say that the model has a R^2 of 1. MAE is used to calculate the prediction error (i.e. the absolute difference between the real and predicted values)The RMSE limits a larger disagreement, whereas the MAE averages the expected deviations. Y values have the same RMSE and MAE units. Both ranges are 0 to infinity. On the whole, small RMSE and MAE values indicate that the model is fitting the data well. The red point in the chart below is showing the detection of and an anomaly.

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - Y_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \tag{5}$$

$$RSME = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2} \tag{6}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \tag{7}$$

Where Y_i and \hat{Y}_i signify the observed and expected minimum vacuum a, and n signifies the total number of pumps Figure 3 showing the EWMA chart of errors.

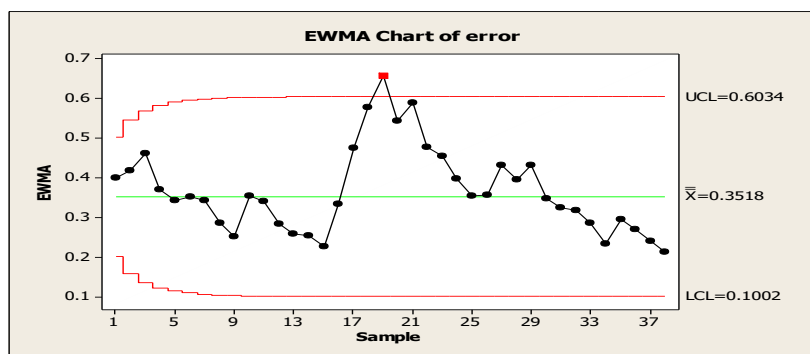


Figure 3: EWMA Chart

V. Anomaly detection Model: Using the information from first model, we constructed a machine learning model and evaluated the performance of many models to choose the best one. The selected algorithms are Logistic Regression, Random forest, SVC, and KNN. The accuracy of these models is shown in Table 7.

Table 7: Accuracy of different model

Model	Accuracy
Logistic Regression	0.97
KNN	0.98
SVC	0.96
Random forest	0.98

Model Validation: Single main validation strategy: percentage split is used because the datasets had a significant number of samples. This approach evaluates models using a fraction of the available data. Divide the dataset in half to create a training set of 80% of the data and a test set of 20% of the data

Model Evaluation: Given that the ultimate goal is assurance of quality and to provide only correct results, the ideal model should be capable of distinguishing aberrant data points from normal data points. The model must be capable of discriminating between right and incorrect positive occurrences. To evaluate our model we used two performance indicators confusion matrix and F1-score. For a binary classification model, one can use several performance indicators from the confusion matrix. A Confusion matrix is a $N \times N$ matrix that is used to evaluate the performance of a classification model, where N is the number of target classes. In the matrix, the machine learning model's prediction values are compared to the actual target values. By assessing the model's overall performance and pointing out any type of mistake, we get a holistic picture of how our classification model is doing.

From Table 8, by comparing the performance of different models it can be stated that the proposed model works effectively for anomaly detection and prediction of vacuum pump assembly line. Random forest showing best result across all performance metrics.

Table 8: comparison of different prediction accuracy

Model	Accuracy (%)	Recall (%)	Specificity (%)	Precision (%)	F1-score (%)	Training Time
Logistic Regression	93.40	88.0	87.4	97	92.28	3.1
KNN	93.40	87.9	86.4	98	92.67	4.3
SVC	92.04	85.4	84.3	96	90.39	3.5
Random Forest	94.85	91.2	90.2	98	94.47	4.2

V. Conclusion and Future Scope

In manufacturing sectors, it is quite challenging to attain product quality improvement, reducing reworking costs, enhancing first pass yield in production or assembly line. Anomaly detection approaches have shown promising results in overcoming the challenges to achieve the goals. In this paper, RF-EWMA based anomaly detection and prediction model has been proposed. A two-phase model is presented in this study for abnormality detection in the assembly line of vacuum pumps. Results show that the number of rejected pumps can be reduced by implementing this framework. The random forest model is showing the best accuracy for vacuum performance prediction. KNN and random forest are showing the best accuracy in classification models. It is to be noted that the initial cost for implementation is higher, however, in long run, it will truly create value.

This work can be further developed in numerous different ways in the future. Feature selection strategies can improve the model's efficiency. Another way, developing a classification model that can differentiate aberrant patterns, such as trend shift and cyclic pattern, from each other, which helps with diagnosing system faults and enhancing fault diagnostic decision-making. Additional modeling may be performed to enhance the proposed framework's flexibility by suggesting machine modifications via process parameter control to restore the process to normal condition when the model detects an abnormal state. Finally, a mix of Adaboost-based machine learning algorithms may be used to enhance the model. As a result, multivariate SPC approaches open up new avenues of investigation.

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Conflicting Interests:

The authors declare that there is no conflict of interest.

References

- [1] Pittino, F., Puggl, M., Moldaschl, T. and Hirschl, C. (2020). Automatic anomaly detection on in-production manufacturing machines using statistical learning methods, *Sensors*, 20:23-44.
- [2] Liu, X. and Nielsen, P. S. (2016). Regression-based online anomaly detection for smart grid data. arXiv:1606.05781.

- [3] Schreyer, M., Sattarov, T., Borth, D., Dengel, A. and Reimer B. (2018). Detection of anomalies in -large-scale accounting data using deep autoencoder networks. arXiv:1709.05254.
- [4] Chahla, C., Snoussi, H., Merghem, L. and Esseghir, M. (2019). A novel approach for anomaly detection in power consumption data. in Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods, 483-490.
- [5] Abdelrahman, O. and Keikhosrokiani, P. (2020). Assembly Line Anomaly Detection and Root Cause Analysis Using Machine Learning. IEEE Access.
- [6] Goldstein, M. and Uchida, S. (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. PLoS ONE, 11:1-31.
- [7] Alelaumi, S., Wang, H., Lu, H. and Yoon, S. W. (2020). A Predictive Abnormality Detection Model Using Ensemble Learning in Stencil Printing Process. IEEE Transactions on Components, Packaging and Manufacturing Technology.
- [8] Liu, F. T., Ting, K. M. and Zhou, Z. H. (2012). Isolation-based anomaly detection. ACM Trans. Knowl. Discov. Data, 6:1-44.
- [9] John, H. and Naaz, S. (2019). Credit card fraud detection using local outlier factor and isolation forest. Int. J. Comput. Sci. Eng., 7:1060-1064.
- [10] Karami, A. and Guerrero-Zapata, M. (2015). A fuzzy anomaly detection system based on hybrid PSO-k means algorithm in content-centric networks. Neurocomputing, 149:1253-1269.
- [11] Malhotra, P., Vig, L., Shroff, G. and Agarwal, P. (2015). Long short term memory networks for anomaly detection in time series. in Proceedings, 89:9-94.
- [12] Khreich, W., Khosravifar, B., Hamou-Lhadj, A. and Talhi, C. (2017). An anomaly detection system based on variable N-gram features and one-class SVM. Inf. Softw. Technol., 91:186-197.
- [13] Jurgovsky, J. et al., (2018). Sequence classification for credit-card fraud detection. Expert Syst. Appl., 100:234-245.
- [14] Singh, R. Kumar, H. and Singla, R. K. (2015). An intrusion detection system using network traffic profiling and online sequential extreme learning machine. Expert Syst. Appl., 42:8609-8624.
- [15] Chauhan, S. and Vig, L. (2015). Anomaly detection in ECG time signals via deep long short-term memory networks. in Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA) 1-7.
- [16] Hosseinioun, N. (2016). Forecasting outlier occurrence in stock market time series based on wavelet transform and adaptive ELM algorithm. J. Math. Finance, 6:127-133.
- [17] Alpaydin, E. Introduction to Machine Learning, vol. 1107, 2nd ed., Cambridge MA, USA: MIT Press, 2014.
- [18] Breiman, L. (2001). Random forests. Mach. Learn., 45:5-32.
- [19] Chalapathy, R. and Chawla, S. [2019]. Deep learning for anomaly detection: A survey, arXiv: 1901.0340.
- [20] Xu, X., Liu, H., Li, L. and Yao, M. (2018). A comparison of outlier detection techniques for high-dimensional data. Int. J. Comput. Intell. Syst., 11:652-662.
- [21] Galante, L. (2019). A comparative evaluation of anomaly detection techniques on multivariate time series data. Int. J. Comput. Sci. Eng., 18:17-29.
- [22] Ahmed, M. and Mahmood, A. N. (2015). Novel approach for network traffic pattern analysis using clustering-based collective anomaly detection. Ann. Data Sci., 2:111-130.
- [23] Schapire, R. E. and Freund, Y. (2012). Boosting: Foundations and Algorithms. Cambridge, MA, USA: MIT Press.