# Anomalies Detection in Smart Manufacturing Using Machine Learning and Deep Learning Algorithms

**Mohamed Gamal**
Production Engineering Department, Faculty of Engineering,
Nahda University, Beni-Suef 62521, Egypt
mohammed.gammal@nub.edu.eg

**Ahmed Donkol**
Electrical Engineering Department, Faculty of Engineering,
Nahda University, Beni-Suef 62521, Egypt
ahmed.donkol@nub.edu.eg

**Ahmed Shaban**
Mechanical Engineering Department, Faculty of Engineering,
Fayoum University, Fayoum 63514, Egypt
ahmed.shaban@fayoum.edu.eg

**Francesco Costantino, Giulio Di Gravio and Riccardo Patriarca**
Mechanical and Aerospace Engineering Department,
Sapienza University of Rome, Via Eudossiana, 18, 00184 Rome, Italy
francesco.costantino@uniroma1.it; giulio.digravio@uniroma1.it; riccardo.patriarca@uniroma1.it

## Abstract

Nowadays, the rapidly changing of manufacturing environment has pushed companies to achieve more customer satisfaction by enhancing product quality, reducing production cost, and realizing sustainability. Anomaly detection has a strong influence on the quality of products and it is usually conducted through visual quality inspection. The visual quality inspection of a product can be performed either manually or automatically. The manual inspection suffers from being a monotonous task, leading to overlooked errors and subjective assessments. Accordingly, the manufacturing industry has high ambitions to rely upon automated quality inspection systems to cope with the requirements of smart manufacturing and the emergence of industry 4.0. Efficient utilization of big data can enable the development of intelligent quality inspection systems. Machine learning as one of the prevailing data analytics methods is widely used to support and improve the performance of the automated quality inspection systems. This research compares the performance of Recurrent Neural Networks (RNN) like Multilayer Perceptron (MLP) with the traditional machine learning algorithms (TMLA) for anomalies detection in manufacturing such as Decision Trees, Random Forest (RF), k-Nearest-Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes, and Logistic Regression (LR). A data set for faults is adapted from the literature to fairly compare the performance of these algorithms considering different accuracy measures such as accuracy, precision, sensitivity, and F1-score.

## Keywords
Smart manufacturing, industry 4.0, anomalies detection, machine learning, deep learning.

## 1. Introduction

One of the most vital day-to-day practices in industry is anomaly detection. There are two major types of anomaly detection, manually and automatically inspection. Basically, manual inspection is a repetitive task that easily leads to missed mistakes and self-assessments. This is particularly challenging in cases where defects occur infrequently. Furthermore, since manual marking of defects is extremely time-consuming, anomaly detection in manufacturing systems has gained popularity among researchers and industry staff. This has motivated the researchers to develop efficient automated anomaly detection systems to overcome the issues associated with the manual inspection (Haselmann, Gruber and Tabatabai, 2019). To enhance product quality, reduce cost, and get more sustainability, anomalies must be detected and causes of those anomalies must be eliminated or minimized. Visual product quality inspection can be related to the field of anomaly detection, which is defined as the detection of patterns that deviate from expected behavior. Furthermore, anomaly detection aids in the tracking of standard daily exercise profiles for any device or system. In such situations, defects visual inspection for fabricated components, either automatically or manually, is a standard procedure. The accurate identification of unusual events provides the decision maker with the opportunity to follow up on the framework in order to effectively avoid, correct, or react to the circumstances associated with them. In either case, the complexity of industry processes and the exponential growth of data makes detecting manufacturing system anomalies difficult. As a result, the industry has high hopes for smart manufacturing, which automates every form of surface inspection.

Smart manufacturing is a modern manufacturing worldview in which manufacturing tools and machines are network connected, sensors monitored, and controlled by advanced computational intelligence. Smart manufacturing aims to improve quality of products, productivity and sustainability. Statistics demonstrated that 82% of the organizations utilizing smart manufacturing techniques have encountered expanded proficiency, and 45% of the organizations experienced expanded consumer loyalty (Wang *et al.*, 2018). As smart manufacturing frameworks complex nature increased, exponential growth of data has been found in manufacturing. The effective use of big data would give knowledge to achieve improvement in anomalies detection. Machine learning, as one of the predominant data analytics strategies, has been generally used to devise complex models and algorithms that loan themselves to get knowledge from data in this particular field (Liu *et al.*, 2019).

This paper presents a comparative study of traditional machine learning algorithms (TMLA) for anomalies detection in manufacturing with Neural Deep Learning. In particular, this research compares the performance of Recurrent Neural Networks (RNN) like Multilayer Perceptron (MLP) with the traditional machine learning algorithms (TMLA) for anomalies detection in manufacturing such as Decision Trees, Random Forest (RF), k-Nearest-Neighbor (KNN), Support Vector Machine (SVM), Naïve Bayes, and Logistic Regression (LR). The comparisons have been carried out using the Semeion data set as a training set to compare the performance of the various machine learning and Deep learning techniques in terms of accuracy and precision (Tian, Fu and Wu, 2015). The results obtained are very promising and can be beneficial for anomalies detection and quality improvement in smart manufacturing. Thus, supervised machine learning techniques will be very supportive in early diagnosis and prognosis of manufacturing anomalies detection. This research encourages production managers to utilize the ML to automate and improve fault detection instead of monotonous manual inspection. Also, ML techniques are powerful tools in manufacturing processes that can be utilized to gain beneficial insights into the behavior of manufacturing systems, improving the quality of decisions.

The paper is structured as follows: the related work is summarized in Section 2. The research methodology including the different ML algorithms to compare is presented in Section 3. The implementation and results analysis are provided in Section 4. The conclusions and future work are given in Section 5.

## 2. Related Work

Many hypothetical and observational works have demonstrated that machine learning methods, including the utilization of data mining, recognition of patterns, and artificial neural networks, are promising for various manufacturing applications (Shaban and Shalaby, 2010; Köksal, Batmaz and Testik, 2011; Shaban and Shalaby, 2012; Wen *et al.*, 2012; Kateris *et al.*, 2014; Buczak and Guven, 2016; Quatrini *et al.*, 2020b). Nonetheless, the application of ML in industry is dependent on internet of things technology, and industrial big data stay rare. Thusly, more endeavors are expected to encourage the relevant system, techniques and applications (Min *et al.*, 2019).

Past research has shown that ML and big data are as a rule progressively used in a variety of industry areas. Hybrid artificial intelligence and ML approaches multi-strategy were presented for managing multifaceted nature, and changes and vulnerabilities in industry (Monostori, 2003). Examples are the application of ML into condition based monitoring (Quatrini *et al.*, 2020a), quality inspection (Kang, Catal and Tekinerdogan, 2020), safety management (Akel *et al.*, 2021), and several industrial applications (Bertolini *et al.*, 2021). Machine learning techniques and data mining were proposed in metal industry application (Pham *et al.*, 2004). ML algorithms were adopted and validated for plastic molding industry quality control through a real case application (Tellaeche and Arana, 2013). The huge number of crude data gathered from physical manufacturing destinations or created in different information systems causes overwhelming information, over-burden issues and majority of traditional techniques for data mining are not yet ready to process large data for management smart production (Cheng *et al.*, 2018), the role of enormous data in supporting smart manufacturing by providing an overview of the historical perspective on the data lifecycle in manufacturing (Min *et al.*, 2019).

ML algorithms are applied to fathom various Data Mining errands. For classification errands, the most adequate algorithms are KNN, Naive Bayes, and Support Vector Machine (Kotenko, Saenko and Branitskiy, 2019). Deep learning gives successful methods that can learn features consequently at various abstraction levels, permitting learning complex contribution to-yield works legitimately from data, without relying upon feature extractors, which can be of incredible advantage for modern industrial applications. For the most part, ANN, SVM and DL strategies will in general perform better when managing multi-dimensions and persistent features; while Naive Bayes and KNN will in general perform better when managing discrete features. On the other hand, Naive Bayes and KNN are on the whole reasonable with clear physical significance, while SVM, ANN and deep learning techniques have poor interpretability (Liu *et al.*, 2018).

Contingent upon the application, the performances of different algorithms are not quite the same. Firstly, K-NN is easy to implement and can be used for both classification and regression tasks but needs large computation, lots of storage space and the selection of K influence right arrow too match (Liu *et al.*, 2018). Secondly, Naive Bayes requires little storage, good physical ability and robust to missing values but it needs prior probability, prior assumptions and may has combinatorial explosion and computation problem (Liu *et al.*, 2018). Thirdly, SVM has high classification accuracy and can deal with high dimensional features but it has low efficiency for big data and with no physical significance. On the other hand, deep learning needn't bother with the feature extractor because it learns features and perceiving deficiency consequently, take in progressively complex1structure from data. Because of the deep architecture, deep learning doesn't need to extract feature because features are learnt and faults are recognized automatically, and for the deep architecture, it learns from data complex structures. But it has some limitations as it needs large samples, long training time, and without physical meaning (Liu *et al.*, 2018). Throughout the current research status in the field of imagery, deep learning leads almost all frontier development directions, but there are still some other limitations. For example, deep learning can be compared with TMLA under the support of computing resources and data volume; this comparison will show excellent effects (Chen *et al.*, 2019).

Feature based methods have been used for many years in the field of industrial inspection and recognition. Iglesias et (2018) extracted 71 features from 3D and color 2D data to indicate the quality of slate slab, e.g., surface uniformity, material defects and warping. Zeng et al. (2016) proposed a man-made clear visual feature based on strobe light to highlight the edges between the seam and metal on weld; then the accurate seam edges can be computed by threshold and edge extraction. Fortunately, as more and more factories focus on the accumulation of defective product data, this issue can be mitigated. Moreover, many viable solutions have been found in literature, including data generation and augmentation, transfer learning, unsupervised learning and semi-supervised learning. There may be problems if the local images are directly fed to a neural network, mainly unnecessary computation brought by background content and noise interference. In addition, training and applying a satisfactory neural network are typically computational expensive (Wang *et al.*, 2018).

Most of the learning-based inspection techniques are trained with the entire training data which may not be available before the task. Moreover, after the learning model has been built, the newly observed defective samples are hard to utilize. Online machine learning allows model updating and optimization at each identification step, which makes it a promising method for handling manufacturing data stream (Wang *et al.*, 2018). The literature indicates that there is a need to compare the widely available machine-learning methods for anomaly detection in different manufacturing environments and applications.

# 3. Methodology

## 3.1 Machine Learning Algorithms

Machine learning (ML) can be defined as a subset of Artificial Intelligence (AI) that provides the ability to model systems based on a data set used for the purpose of training in contrast to the typical approach of coding all possible outcomes beforehand. The main machine learning objective is to permit a system to gain information from an earlier time or present and utilize the information to settle on expectations or choices to address future events. Multiple approaches and techniques are present to make systems that can learn. Some of them are decision trees, SVM, KNN, Naïve Bayes, logistic regression, and MLP artificial neural network. A brief note on those techniques is presented in this section.

### 3.1.1 Decision Tree

A decision tree is a recursive split of the instance space that is used to classify data (Safavian and Landgrebe, 1991). The decision tree is made up of nodes that create a rooted tree, which is a directed tree with no incoming edges and a node named "root.", each of the other nodes has one incoming edge. An internal or test node is a node with outgoing edges. All additional nodes are referred to as leaves (also known as terminal or decision nodes), each internal node in a decision tree divides the instance space into two or more sub-spaces based on a discrete function of the input attribute values, each leaf is given a class to symbolize it. Instances are identified by travelling them from the tree's root to a leaf and classifying them based on the results of the tests along the way. As such, a decision tree is a tree in which each branch node represents a choice between a number of alternatives, and each leaf node represents a classification or decision as shown in Figure 1.

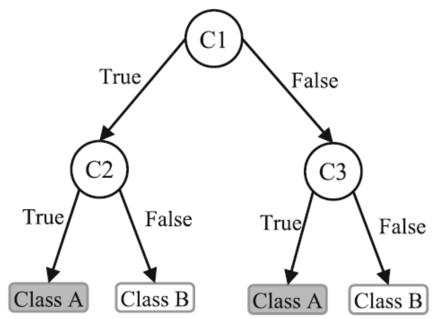


Figure 1. How Decision Tree Works

### 3.1.2 Random Forest (RF)

A random forest (RF) is an ensemble classifier made up of numerous DTs, much like a forest is made up of many trees as depicted in Figure 2 (Uddin *et al.*, 2019). Over fitting of the training data is common deep with DTs, resulting in a large variance in classification results for a minor change in the input data. They are extremely sensitive to their training data, making them prone to making mistakes with the test dataset. Different sections of the training dataset are used to train the different DTs of RF. To categorize a new sample, the sample's input vector must be passed down with each DT of the forest. The classification conclusion is subsequently determined by each DT considering a separate section of the input vector. The forest then decides whether the classification with the most "votes" (for discrete classification outcome) or the average of all trees in the forest should be used (for numeric classification outcome). Because the RF algorithm takes into account the results of several different DTs, it can reduce the variance caused by only considering one DT for the same dataset. Figure 2 shows an illustration of the RF algorithm (Uddin *et al.*, 2019).

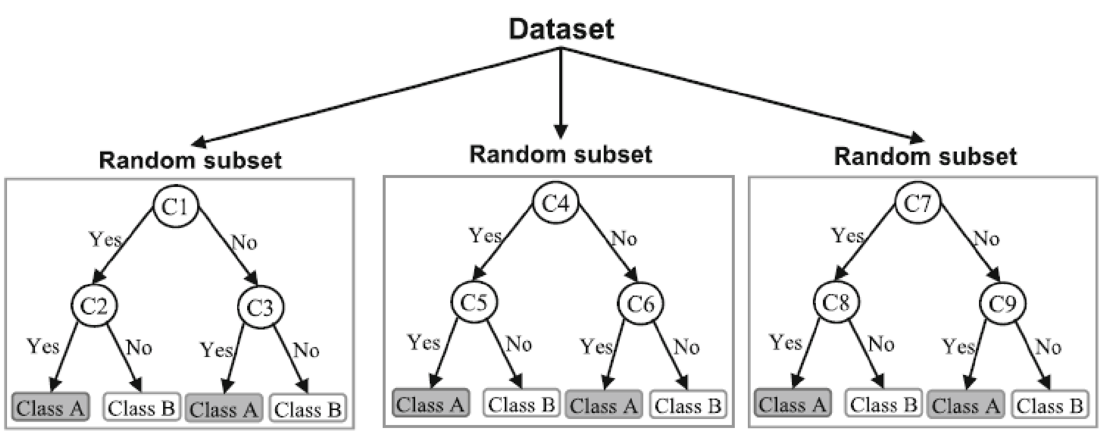


Figure 2. How RF Works

### 3.1.3 K-Nearest Neighbor (KNN)
One of the simplest and earliest classification techniques is the K-nearest neighbour (KNN) algorithm (Sharma, Aggarwal and Choudhury, 2018). The KNN algorithm does not need to take probability values into account. The number of nearest neighbours in the KNN algorithm represented by the ‘K’ is considered to take a ‘vote’. Changing the value of ‘K' can result in various classification results for the same sample object. Figure 3 depicts the KNN's classification process for a new object. When K= 3, the new object (star) is classed as ‘blue,' but when K= 6, it is labelled as ‘red.'

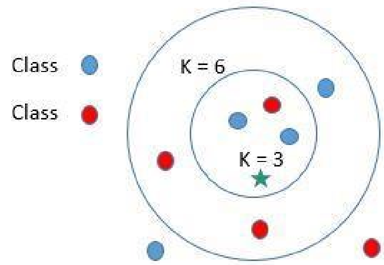


Figure 3. Illustration KNN algorithm

### 3.1.4 Naïve Bayes
Naïve Bayes Classifiers are probabilistic in nature and are defined by applying the Bayes theorem to them (Liu *et al.*, 2018). It is naive because it believes that all features are independent of one another, which is rarely the case in real-world situation. Naïve Bayes has proven to be effective for a wide range of machine learning tasks.

### 3.1.5 Support Vector Machine (SVM)
SVM algorithm can classify both linear and non-linear data, each data item is initially mapped onto an n-dimensional feature space, where n is the number of features (Liu *et al.*, 2018). The hyper plane that splits the data items into two groups is then identified. With the marginal distance for both classes maximized and classification errors minimized, the marginal distance between the decision hyper plane and the class's nearest instance is the distance between the decision hyper plane and the class's nearest instance. Each data point is initially plotted as a point in an n-dimensional space (where n is the number of features), with the value of each feature being the value of a given coordinate. To complete the classification, the hyper plane that separates the two classes by the greatest margin must be located. A basic SVM classifier is seen in Figure 4.

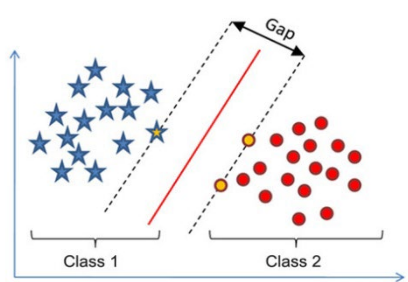


Figure 4. How SVM Works

### 3.1.6 Logistic Regression (LR)
When the value of the target variable is categorical in nature, LR is utilized as a classification procedure (Penumuru, Muthuswamy and Karumbu, 2020). Also, LR is utilized when the data in question has binary output, such as when it belongs to one of two classes or is either a 0 or a 1. This approach is utilized to develop a trained model for predicting with a ridge estimator. It is easy to implement and straightforward, does not require any assumptions regarding the distribution of independent variable (s). LR-based models can be updated easily, and it has a nice probabilistic interpretation of model parameters, but If the target variables are discrete, logistic regression should not be employed.

### 3.1.7 MULTI LAYER PERCEPTRON (MLP)

MLP stands for multilayered nonlinear neural network, it has numerous layers, including an input layer, one or more hidden layers, and an output layer as shown in Figure 5 (Liu *et al.*, 2018). Each neuron in the hidden layer has a nonlinear activation function, such as sigmoid or tan sigmoid; each neuron is defined by its activation function and is connected to all neurons in the next layer. Each connection is defined by its weight factor or synaptic weight. Figure 5 illustrates a typical configuration of a multilayer perceptron-based neural network. L is the structure's number of layers, the input layer is on the top, and the output layer is on the bottom, from the second to the (L-1)$^{th}$ layer, there are concealed layers. As can be seen, the network is completely interconnected, allowing each layer's neuron to communicate with each layer's neuron, this allows for forward transmission of data from the input layer to the output layer via the hidden layers.

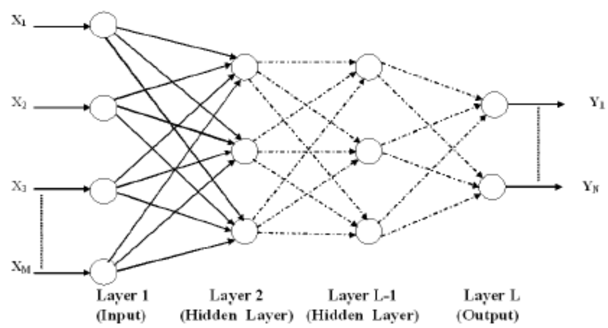


Figure 5. A Typical MLP- NN Configuration

For some tasks, it's hard to specify what features should be extracted to provide the artificial intelligence algorithms; in deep learning strategies can defeat the previously mentioned insufficiencies. Deep Learning (DL) is an on-going pattern of machine learning that has surfaced as a strategy for detection of patterns from complex process, by raw signals as input data. Deep learning gives a viable method to learn consequently at various degrees of abstraction, permitting learning complex contribution to-yield directly works from data, without relying upon feature extractors, which is an incredible advantage for anomalies recognition in various industries. The low-level cost of enormous data sets stockpiling and high computational performance has driven the rise of deep learning.

## 3.2 Data Set Description

This research compares different machine learning algorithms based on a dataset taken from the literature. The data set relates to Steel plate faults, by Semeion, Research of Sciences of Communication. In this dataset, a superficial fault of a stainless-steel leaf represented by 7 faults as shown in Table 1. Each fault is characterized by 27 attributes and features representing the geometric shape and contour of each fault (see Table 2).

Table 1. Faults types and its sample sizes (Tian, Fu and Wu, 2015)

| Fault Number | Faults type | Sample size |
|---|---|---|
| 1 | Pastry | 158 |
| 2 | Z scratch | 190 |
| 3 | K scratch | 391 |
| 4 | Stains | 72 |
| 5 | Dirtiness | 55 |
| 6 | Bumps | 402 |
| 7 | Other Faults | 673 |

Table 2. Steel plates independent attributes (Tian, Fu and Wu, 2015)

| Number. | Attribute. | Number. | Attribute. |
|---|---|---|---|
| 1 | X minimum | 15 | Edges Index |
| 2 | X maximum | 16 | Empty Index |
| 3 | Y minimum | 17 | Square Index |
| 4 | Y maximum | 18 | Outside X Index |
| 5 | Pixels Areas | 19 | Edges X Index |
| 6 | X perimeter | 20 | Edges Y Index |
| 7 | Y perimeter | 21 | Outside Global index |
| 8 | Sum of Luminosity | 22 | Log of Areas |
| 9 | Minimum of Luminosity | 23 | Log X Index |
| 10 | Maximum of Luminosity | 24 | Log Y Index |
| 11 | Length of Conveyer | 25 | Orientation Index |
| 12 | TypeofSteel_A300 | 26 | Luminosity Index |
| 13 | TypeofSteel_A400 | 27 | Sigmoid of Areas |
| 14 | Steel Plate Thickness | | |

### 3.3 Performance Metrics

The performance measures used for comparing and evaluating the performance of machine learning techniques include: accuracy, sensitivity, precision and F1 Score, that can be defined as follows:

- Accuracy measure of the correct prediction in correspondence to the wrong ones.
- Recall (Sensitivity) measure the system's effectiveness in predicting positives and determining costs.
- Precision measure the degree of correctness in determining the positive outcomes that may be defined as precision. It is the ratio between true positives and the overall set of positives.
- F1 Score is the weighted average of Precision and Sensitivity. This measure hence, considers both types of false values. F1 score is considered perfect when at 1 and is a total failure when at 0.

These metrics can be represented mathematically as follows (Sharma, Aggarwal and Choudhury, 2018):

- Accuracy (%) = (T P + T N) / (T P + T N + F P + F N).
- Sensitivity (%) = T P / (T P + F N).
- Precision (%) = TP / (TP + FP).
- F1 Score (%) = (2*(Precision*Sensitivity)) / (Precision +Sensitivity).

where TN, FN, TP, and FP represent true negative, false negative, true Positive, and false positive, respectively.

## 4. Implementation and Result Analysis

A comparative study using Random Forest, Decision trees, KNN, SVM, Naïve Bayes, Logistic Regression, with MLP is conducted. We have used numpy, pandas and Scikit-learn which are open source machine learning libraries in Python. An open source web application named as Jupyter Notebook is used to run the program. For training and testing categories, roughly 50% of all examples are picked for training and 50% for testing (see Table 3).

Table 3. Composition of training and testing datasets

| Fault number | Training data set (number of examples, %) | | Testing data set (number of examples, %) | |
|---|---|---|---|---|
| 1 | 80 | 50.60% | 78 | 49.40% |
| 2 | 100 | 52.60% | 90 | 47.40% |
| 3 | 200 | 51.20% | 191 | 48.80% |
| 4 | 40 | 55.60% | 32 | 44.40% |
| 5 | 30 | 54.50% | 25 | 45.50% |
| 6 | 200 | 49.80% | 202 | 50.20% |
| 7 | 350 | 52.00% | 323 | 48.00% |

The results of all performance measures (accuracy, precision, sensitivity and F1-score) for single categories classification under the different machine learning algorithms are reported in Tables 4, 5, 6 and 7, respectively. In Table 4, it can be observed that RF provides a promising accuracy performance in detecting the seven faults in comparison to all other ML algorithms where it achieves an accuracy that varies with the faults from 81% to 99% with an average accuracy of 93.29% for all faults. The DT, LR , SVM, and KNN provide an average accuracy that ranges from 82.86% to 91.14% while the MLP and Naïve Bayes provide the lowest average accuracies. The results of the other performance measures indicate the consistency of the RF algorithm as it provides a superior performance with respect to all the performance measures for all faults (see Tables 4-7). However, it can be observed that all algorithms fail to achieve a good detection performance for faults 6 and 7 as compared to other faults. Therefore, these tested ML algorithms are not the most suited for detecting these two faults. Generally, it can be observed that faults can be categorized into three categories; first one concludes those faults that can be easily classified like fault 4; the second category like fault 7 that cannot be easily classified or determined and the third category includes those faults which ML algorithms classification ability differ from classifying it. So, it is recommended to find the suitable ML algorithms to classify those faults in the last category considering the results previously mentioned vital metrics.

Table 4. Single categories Accuracy (%)

| **Class** | **Decision trees** | **KNN** | **RF** | **SVM** | **Naïve Bayes** | **LR** | **MLP** |
|---|---|---|---|---|---|---|---|
| Fault1 | 91% | 84% | 94% | 92% | 31% | 92% | 92% |
| Fault2 | 95% | 83% | 98% | 91% | 74% | 91% | 49% |
| Fault3 | 97% | 92% | 98% | 79% | 93% | 94% | 91% |
| Fault4 | 99% | 94% | 99% | 97% | 87% | 99% | 94% |
| Fault5 | 98% | 96% | 98% | 98% | 33% | 98% | 98% |
| Fault6 | 81% | 71% | 85% | 79% | 44% | 78% | 54% |
| Fault7 | 77% | 60% | 81% | 66% | 51% | 66% | 39% |
| Average | 91.14% | 82.86% | 93.29% | 86.00% | 59.00% | 88.29% | 73.86% |

Table 5. Single categories precision (%)

| **Class** | **Decision trees** | **KNN** | **RF** | **SVM** | **Naïve Bayes** | **LR** | **MLP** |
|---|---|---|---|---|---|---|---|
| Fault1 | 91% | 85% | 93% | 84% | 89% | 88% | 93% |
| Fault2 | 96% | 83% | 97% | 82% | 92% | 82% | 90% |
| Fault3 | 97% | 92% | 98% | 63% | 93% | 94% | 92% |
| Fault4 | 99% | 93% | 99% | 93% | 97% | 99% | 98% |
| Fault5 | 98% | 96% | 98% | 95% | 97% | 98% | 95% |
| Fault6 | 81% | 71% | 84% | 62% | 84% | 68% | 78% |
| Fault7 | 77% | 60% | 81% | 43% | 73% | 78% | 73% |
| Average | 91.29% | 82.86% | 92.86% | 74.57% | 89.29% | 86.71% | 88.43% |

Table 6. Single categories Sensitivity (%)

| **Class** | **Decision trees** | **KNN** | **RF** | **SVM** | **Naïve Bayes** | **LR** | **MLP** |
|---|---|---|---|---|---|---|---|
| Fault1 | 91% | 84% | 94% | 92% | 31% | 92% | 92% |
| Fault2 | 95% | 83% | 98% | 91% | 74% | 91% | 49% |
| Fault3 | 97% | 92% | 98% | 79% | 93% | 94% | 91% |
| Fault4 | 99% | 94% | 99% | 97% | 87% | 99% | 94% |
| Fault5 | 98% | 96% | 98% | 98% | 33% | 98% | 98% |
| Fault6 | 81% | 71% | 85% | 79% | 44% | 78% | 54% |
| Fault7 | 77% | 60% | 81% | 66% | 51% | 66% | 39% |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Average | 91.14% | 82.86% | 93.29% | 86.00% | 59.00% | 88.29% | 73.86% |

Table 7. F1-score for Single Categories (%)

| Class | Decision trees | KNN | RF | SVM | Naïve Bayes | LR | MLP |
|---|---|---|---|---|---|---|---|
| Fault1 | 91% | 84% | 92% | 88% | 38% | 88% | 88% |
| Fault2 | 96% | 83% | 97% | 86% | 79% | 86% | 58% |
| Fault3 | 97% | 92% | 98% | 70% | 93% | 94% | 91% |
| Fault4 | 99% | 94% | 99% | 95% | 91% | 99% | 95% |
| Fault5 | 98% | 96% | 98% | 96% | 47% | 98% | 96% |
| Fault6 | 81% | 71% | 83% | 69% | 44% | 69% | 58% |
| Fault7 | 77% | 60% | 80% | 52% | 48% | 52% | 28% |
| Average | 91.29% | 82.86% | 92.43% | 79.43% | 62.86% | 83.71% | 73.43% |

The comparison results of the different machine learning algorithms (based on the average scores of the performance measures over all faults) are summarized in Table 8 and depicted graphically in Figure 6. The results clearly confirm that RF is the most effective algorithm in faults detection with respect to all the performance measures. The second-best algorithm in detecting the faults is the decision tree. On the other hand, naïve Bayes is the lowest effective one, and by comparing TMLA with MLP as one of the artificial neural networks, despite MLP make good classifier algorithm, TMLA are superior to MLP and doing well in such classification cases.

Table 8. Average scores of the performance measures for the different machine learning algorithms

| | Decision trees | KNN | RF | SVM | Naïve Bayes | LR | MLP |
|---|---|---|---|---|---|---|---|
| Accuracy (%) | 91.14% | 82.86% | 93.29% | 86.00% | 59.00% | 88.29% | 73.86% |
| Precision (%) | 91.29% | 82.86% | 92.86% | 74.57% | 89.29% | 86.71% | 88.43% |
| Sensitivity (%) | 91.14% | 82.86% | 93.29% | 86.00% | 59.00% | 88.29% | 73.86% |
| F1-Score (%) | 91.29% | 82.86% | 92.43% | 79.43% | 62.86% | 83.71% | 73.43% |

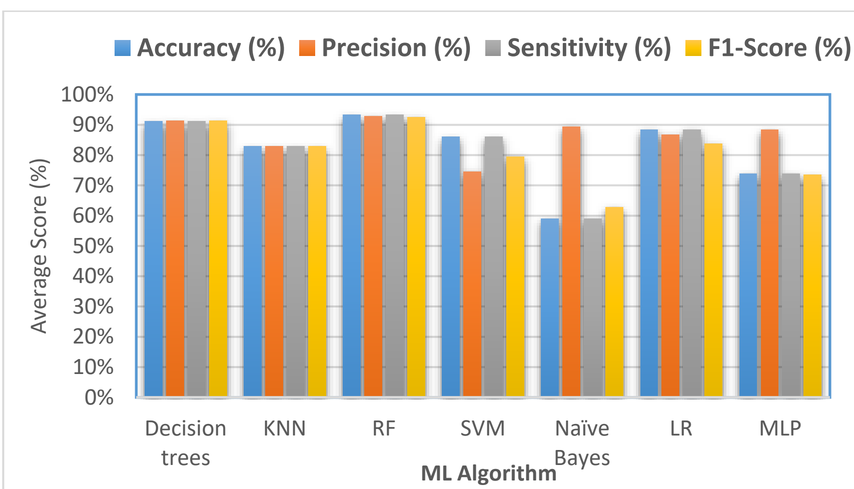


Figure 6. Comparison of different ML algorithms based on the average scores of the performance measures.

## 5. Conclusion

This paper presented a comparative study of different machine learning algorithms, for anomalies detection in manufacturing using the dataset of Steel plate faults, by Semeion, Research of Sciences of Communication. A framework has been developed for the multi classification of steel plates faults from fault (1-7) by investigation the use of mostly popular traditional machine learning; Decision trees, KNN, SVM, Naïve Bayes, RF, Logistic Regression and artificial neural network namely multilayer perceptron (MLP). The accuracy, precision, sensitivity, and F1-score are considered the key performance metrics for the comparative study. It has been observed that some ML Algorithms had an accuracy above 93% in detecting faults. The RF has the best recall, accuracy, precision and F1 score performance average measures over the competitors. By comparing TMLA with MLP as one of the ANN, TMLA are superior to MLP. From the results, the faults can be classified into three categories; (i) easily classified faults like fault 4, (ii) faults cannot be easily classified like fault 7, (iii) faults detection performance depends on the type of ML algorithms. Therefore, it is recommended that, to find the suitable ML algorithm to classify those faults. The performance of the methods achieved of test data is reached 93% that guide production managers can utilize the ML to automate and improve fault detection instead of monotonous manual inspection. In addition, the significant and rapid improvements in data integration technologies make ML a powerful tool in manufacturing processes and can be utilized to gain beneficial insights into the behavior of manufacturing systems, improving the quality of decisions.
Future research should consider testing the adopted machine learning algorithm in this paper with a variety of applications in manufacturing. In addition, other machine learning algorithms or hybrid algorithms should be developed and tested for improving the detecting accuracy. In particular, there is a need to find some appropriate machine learning algorithms to handle the cases such as faults 6 and 7.

## References

Akel, N. *et al.* (2021) 'Business Intelligence for the Analysis of Industrial Accidents Based on Mhidas Database', *Chemical Engineering Transactions*, 86, pp. 229–234. doi: 10.3303/CET2186039.

Bertolini, M. *et al.* (2021) 'Machine Learning for industrial applications: A comprehensive literature review', *Expert Systems with Applications*. Pergamon, 175, p. 114820. doi: 10.1016/J.ESWA.2021.114820.

Buczak, A. L. and Guven, E. (2016) 'A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection', *IEEE Communications Surveys and Tutorials*. Institute of Electrical and Electronics Engineers Inc., 18(2), pp. 1153–1176. doi: 10.1109/COMST.2015.2494502.

Chen, X. *et al.* (2019) 'Research on deep learning in the field of mechanical equipment fault diagnosis image quality', *Journal of Visual Communication and Image Representation*. Academic Press Inc., 62, pp. 402–409. doi: 10.1016/j.jvcir.2019.06.007.

Cheng, Y. *et al.* (2018) 'Data and knowledge mining with big data towards smart production', *Journal of Industrial Information Integration*. Elsevier B.V., 9, pp. 1–13. doi: 10.1016/j.jii.2017.08.001.

Haselmann, M., Gruber, D. P. and Tabatabai, P. (2019) 'Anomaly Detection Using Deep Learning Based Image Completion', in *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*. Institute of Electrical and Electronics Engineers Inc., pp. 1237–1242. doi: 10.1109/ICMLA.2018.00201.

Iglesias, C., Martínez, J. and Taboada, J. (2018) 'Automated vision system for quality inspection of slate slabs', *Computers in Industry*. Elsevier B.V., 99, pp. 119–129. doi: 10.1016/j.compind.2018.03.030.

Kang, Z., Catal, C. and Tekinerdogan, B. (2020) 'Machine learning applications in production lines: A systematic literature review', *Computers & Industrial Engineering*. Pergamon, 149, p. 106773. doi: 10.1016/J.CIE.2020.106773.

Kateris, D. *et al.* (2014) 'A machine learning approach for the condition monitoring of rotating machinery †', *Journal of Mechanical Science and Technology*, 28(1), pp. 61–71. doi: 10.1007/s12206-013-1102-y.

Köksal, G., Batmaz, I. and Testik, M. C. (2011) 'A review of data mining applications for quality improvement in manufacturing industry', *Expert Systems with Applications*. Pergamon, pp. 13448–13467. doi: 10.1016/j.eswa.2011.04.063.

Kotenko, I., Saenko, I. and Branitskiy, A. (2019) 'Improving the Performance of Manufacturing Technologies for Advanced Material Processing Using a Big Data and Machine Learning Framework', in *Materials Today: Proceedings*. Elsevier Ltd, pp. 380–385. doi: 10.1016/j.matpr.2018.12.162.

Liu, J. *et al.* (2019) 'Anomaly Detection in Manufacturing Systems Using Structured Neural Networks', in *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*. Institute of Electrical and Electronics Engineers Inc., pp. 175–180. doi: 10.1109/WCICA.2018.8630692.

Liu, R. *et al.* (2018) 'Artificial intelligence for fault diagnosis of rotating machinery: A review', *Mechanical Systems and Signal Processing*. Academic Press, pp. 33–47. doi: 10.1016/j.ymssp.2018.02.016.

Min, Q. *et al.* (2019) 'Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry', *International Journal of Information Management*. Elsevier Ltd, 49, pp. 502–519. doi: 10.1016/j.ijinfomgt.2019.05.020.

Monostori, L. (2003) 'AI and machine learning techniques for managing complexity, changes and uncertainties in manufacturing', in *Engineering Applications of Artificial Intelligence*. Pergamon, pp. 277–291. doi: 10.1016/S0952-1976(03)00078-2.

Penumuru, D. P., Muthuswamy, S. and Karumbu, P. (2020) 'Identification and classification of materials using machine vision and machine learning in the context of industry 4.0', *Journal of Intelligent Manufacturing*. Springer, 31(5), pp. 1229–1241. doi: 10.1007/s10845-019-01508-6.

Pham, D. T. *et al.* (2004) 'An application of datamining and machine learning techniques in the metal industry', in *Proceedings of the 4th CIRP International Seminar on Intelligent Computation in Manufacturing Engineering (ICME-04)*. Sorrento (Naples), Italy.

Quatrini, E. *et al.* (2020a) 'Condition-based maintenance-An extensive literature review', *Machines*, 8(2). doi: 10.3390/MACHINES8020031.

Quatrini, E. *et al.* (2020b) 'Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities', *Journal of Manufacturing Systems*, 56, pp. 117–132. doi: 10.1016/j.jmsy.2020.05.013.

Safavian, S. R. and Landgrebe, D. (1991) 'A Survey of Decision Tree Classifier Methodology', *IEEE Transactions on Systems, Man and Cybernetics*, 21(3), pp. 660–674. doi: 10.1109/21.97458.

Shaban, A. and Shalaby, M. A. (2010) 'A double neural network approach for the automated detection of quality control chart patterns', *International Journal of Rapid Manufacturing*. Inderscience Publishers, 1(3), p. 278. doi: 10.1504/ijrapidm.2010.034250.

Shaban, A. and Shalaby, M. A. (2012) 'A double neural network approach for the identification and parameter estimation of control chart patterns', *International Journal of Quality Engineering and Technology*, 3(2). doi: 10.1504/IJQET.2012.049681.

Sharma, S., Aggarwal, A. and Choudhury, T. (2018) 'Breast Cancer Detection Using Machine Learning Algorithms', in *Proceedings of the International Conference on Computational Techniques, Electronics and Mechanical Systems, CTEMS 2018*. Institute of Electrical and Electronics Engineers Inc., pp. 114–118. doi: 10.1109/CTEMS.2018.8769187.

Tellaeche, A. and Arana, R. (2013) 'Machine learning algorithms for quality control in plastic molding industry', in *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*. doi: 10.1109/ETFA.2013.6648103.

Tian, Y., Fu, M. and Wu, F. (2015) 'Steel plates fault diagnosis on the basis of support vector machines', *Neurocomputing*. Elsevier B.V., 151(P1), pp. 296–303. doi: 10.1016/j.neucom.2014.09.036.

Uddin, S. *et al.* (2019) 'Comparing different supervised machine learning algorithms for disease prediction', *BMC Medical Informatics and Decision Making*. BioMed Central Ltd, 19(1), pp. 1–16. doi: 10.1186/s12911-019-1004-8.

Wang, J. *et al.* (2018) 'Deep learning for smart manufacturing: Methods and applications', *Journal of Manufacturing Systems*. Elsevier B.V., 48, pp. 144–156. doi: 10.1016/j.jmsy.2018.01.003.

Wen, J. *et al.* (2012) 'Systematic literature review of machine learning based software development effort estimation models', *Information and Software Technology*. Elsevier B.V., pp. 41–59. doi: 10.1016/j.infsof.2011.09.002.

Zeng, J. *et al.* (2016) 'A visual weld edge recognition method based on light and shadow feature construction using directional lighting', *Journal of Manufacturing Processes*. Elsevier Ltd, 24, pp. 19–30. doi:

10.1016/j.jmapro.2016.07.002.

## Biographies

**Mohamed Gamal** is a Production team leader with +10 years of experience in managing & following up all activities, people, information, and resources involved in moving a product from supplier to customer, production management, process improvement, and quality assurance. In parallel to his career, he has done his master degree, he is currently a PhD researcher-Minia University, and Lecturer assistant in production Engineering Department, Nahda University Beni-Sweif, his research interests include anomalies detection in smart manufacturing, what can be best utilized in promoting the growth and the success of both his company and him.

**Ahamed Donkol** was born in Qena, Egypt in 1988. He received the B.Sc., M.Sc., and PhD degrees in Electrical Engi- neering, Assuit University, Egypt, in 2010, 2015, and 2018, respectively. He is Lecturer at Department of Communication and Computer Engineering, Faculty of Engineering, Nahda University in Beni-Sweif, Egypt. His research interests include wave- let analysis of digital streams, image processing, artificial intelligence, machine learning, deep learning.

**Ahmed Shaban** received his BSc degree in industrial engineering (IE) in 2006 with honour degree from the Industrial Engineering Department at Fayoum University, Egypt. He obtained his MSc degree in mechanical design and production (IE specialty) from the Mechanical Design and Production Department at Cairo University in 2010. He received his PhD degree in IE from Sapienza University of Rome, Italy in 2014. He is an Associate Professor in the Mechanical Engineering Department at Fayoum University, where he is involved in research and teaching activities relating to IE & Operations Research. His current research interests involve supply chain modelling, modelling and optimization, simulation, healthcare systems and energy systems. He has published numerous research papers in high quality journals. He is a reviewer for many international journals. He is also appointed as an editorial board member for a number of journals. He is the PI and Co-PI for a number of research project.

**Francesco Costantino** is currently associate professor at Sapienza University of Rome (Italy). He held the chair of Smart Factory, the chair of Industrial Systems Engineering, and the chair of Operations Management. He is member of the Industrial Production Engineering Ph.D. Collegium. He is director of the Course of Higher Education "Technologies, organizations, individuals and behaviors in the fourth industrial revolution for health and safety at work". He is national expert in the ISO working groups ISO/IEC JWG21 “Smart Manufacturing Reference Model” and ISO/TC184/SC5/SG6 “Mass Customization”. His research activity mainly concerns supply chain management, maintenance and logistics, analysis of financial systems for industrial investment, integrated approaches to compliance & risk management, resilience engineering, application of AI technologies into production systems. The results achieved are highlighted both by publications in highly impacted scientific journals, also in collaboration with major international research groups, and by participating in several funded research projects.

**Giulio Di Gravio** is associate professor at Sapienza University of Rome (Italy), holding the chair of Quality Management, and Operations Management. Currently, he is the scientific coordinator of Sapienza’s PhD Programme in Industrial and Management Engineering. His research activity mainly concerns supply chain management, maintenance and logistics, techno-economic optimization of industrial plants, integrated approaches to compliance & risk management, resilience engineering and management. The results achieved through his research are highlighted both by publications in high-impact scientific journals, and by participating or coordinating several research projects. He collaborates actively with both national and international research institutions, enterprises, and universities.

**Riccardo Patriarca** is a post-doc researcher and adjunct professor at Sapienza University of Rome (Italy) of Industrial Systems Engineering. His research activity mainly concerns resilience and risk management for complex socio-technical systems, human factors and ergonomics, maintenance and logistics, as well as supply chain management at large. He acts as a reviewer for several international peer-reviewed journals, and he is a member of the editorial board of Elsevier’s Safety Science and Reliability Engineering and System Safety, among others. His research has been published widely in international scientific journals, and conference proceedings, often in collaboration with national and international scholars.