International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Role of Machine Learning in Manufacturing Sector

Vimanyu Chopra, Devinder Priyadarshi

Abstract: ***This paper discusses the basic concept Machine learning and its techniques, algorithms as well the impact of Machine Learning in Manufacturing processes and Industrial Production. There has been an unprecedented increase in the data available in the last couple of decades. This has enabled machine learning to be applied in various fields. Machine learning is field of study which enables the computer system to learn automatically as well as improve from experience and perform various tasks without explicit instructions. The primary aim of machine learning is to allow the computers learn automatically without human assistance or intervention and adjust actions accordingly It is being employed in many fields. Manufacturing one area where machine learning is very useful. This paper discusses the various areas where machine learning can improve the process of manufacturing like predictive maintenance, process optimisation, quality control, scheduling of resources among others. This can be done by employing various machine learning techniques and algorithms using concepts such as deep learning, neural networks, supervised, unsupervised and reinforcement learning. The relationship of how machines and humans can co-exist and work together to improve the efficiency of production is also discussed. Industry 4.0 or fourth revolution that has occurred in manufacturing which deals with advent of automation in manufacturing industry and its significance is discussed. We take a look at the various benefits and applications of machine learning in the field of manufacturing engineering. This paper also discusses the various challenges and future scope of employing machine learning in the manufacturing.***

Keywords: Machine learning, Industry 4.0, Manufacturing, Automation.

I. INTRODUCTION

Machine learning as the name suggest is the ability of the machine to learn by themselves. It deals with the programs which learn over time to improve and adapt the performance based on experience and past results without explicit instructions from humans. Machine learning is being employed in various fields ranging from image processing, predictive analysis, military to medicinal field. It has revolutionised the manufacturing industry over the last few decades. There are many techniques and approaches which enable the use of machine learning across various fields. This paper first discusses these techniques employed in manufacturing industry and then we take a look at the various aspects of industry that machine learning has helped including- predictive maintenance, quality control, optimization and scheduling, supply chain forecasting as well as the concept of industry 4.0 and the symbiotic relationship between machines and human beings needed for greater efficiency and better performance. The various challenges for employing machine learning in the field and the future scope of the technology is also discussed.

Manuscript published on November 30, 2019.
* Correspondence Author
Vimanyu Chopra*, UG Scholar, DAV Institute of Engineering & Technology, Jalandhar, India. Email: vimanyuchopra2@gmail.com
Devinder Priyadarshi, Faculty, DAV Institute of Engineering & Technology, Jalandhar, India. Email: priyadarshidevinder@gmail.com

II. TECHNIQUES

Machine learning uses many algorithms and tools for its implementation. Some of the types of machine learning along with their algorithms are:

A. Supervised Learning

Supervised machine learning algorithms use previous instances which are used to make prediction for future instances. The data is divided into training and test dataset [1]. A function is inferred from the training data which can be used to make predictions about output values. This approach after sufficient training can be used to provide targets for new inputs ad can also compare the predicted output to the intended value of output which is useful in finding errors so that the model can be modified accordingly. This approach when training intends to find some patterns in training dataset which can be used in prediction or classification [1]. Some of the most popular algorithms in supervised learning are:

- Decision Trees
- Naive Bayes

Decision Trees

Decision trees (Fig 1) are the type of trees which based on the values group the attributes together. Decision trees rely heavily on pre-determined classification [1]. Decision trees are used in classification to identify the attributes which can be used to extract the most information required for solving the classification puzzle.

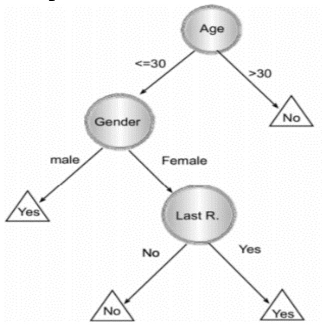
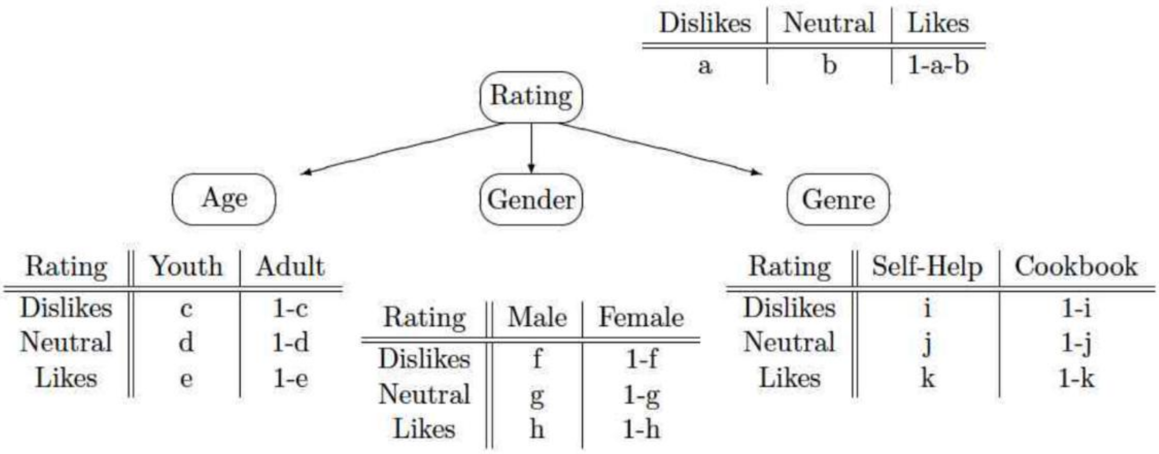


Fig 1 - Decision Tree [19]

Naive Bayes

Naive bayes is a classifier which is based on bayes theorem of

$$P(c|x) = \frac{P(x|c)}{P(x)} P(c).$$



conditional probability.

Fig 2. Example of Bayesian Network [18]

- When predictor (x, attributes) of class (c,target) is given then posterior probability is donated by P(c|x)
- The class prior probability is denoted by P(c)
- P(x|c) is the likelihood which is the probability of predictor given class. The probability of predictor which is the likelihood is denoted by P(x|c)
- The predictor prior probability is denoted by P(x)

It forms bayes network which is a group of trees based on probability of things happening [2]. It is simple to apply and is widely used for classification and clustering large datasets. Disadvantage of using these classifiers is the requirement of these predictors to be independent. It is used for real time prediction, text classification and sentiment analysis, recommendation systems.

B. Unsupervised Learning

These techniques are applied to the dataset which has neither labelled nor classified Unsupervised learning is followed when input variables are known and we have no idea about output variables. This approach is used to uncover unknown patterns and find structures which might be hidden in the datasets. The most common technique used is clustering in which unorganized data is split into similarity groups. Clustering aims to form groups or “clusters” of similar data based on metrics such as probabilistic distance [2]. Some popular unsupervised learning models are:

- **K means clustering**
- **Principal Component Analysis**

K means clustering

K-means clustering aims to form clusters which are groups of data points grouped together because of similarities. The first step is to define the number of clusters k and the centre of these clusters which can be chosen randomly. These centroids are placed as far away from each other as possible. The data points belonging to a data set are then associated to the nearest centre and keeping the centroid as small as possible. At the end of this stage k is re-calculated and the same data points are allocated again to the new k centres [2]. This process is repeated iteratively till there is no change in the position of the centroid.

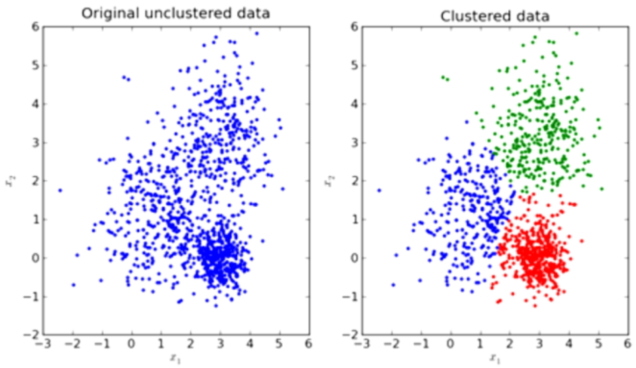


Fig 3. Clustered Data [20]

Principal Component Analysis

Multivariate Analysis generally involves a large number of correlated values in the data. Principal Component Analysis can be used to reduce the large amount of variables to a smaller set while still containing the important information. This mathematical procedure converts correlated variables to smaller sets known as principal components. For example a 2d graph can be represented as a 1d graph while still containing the important information with the help of PCA [2].

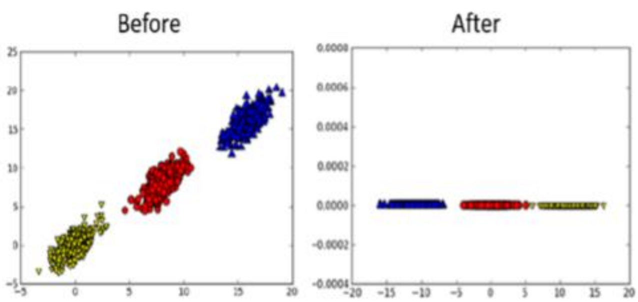


Fig 4. Data visualization before and after applying PCA [1]

C. Semi Supervised Learning

This mode of learning involves the power of both above-mentioned techniques - supervised and unsupervised learning [1]. This technique is utilised when a large amount of unlabelled data is present along with a small amount of labelled data in the dataset.

This labelled data is used for training while also improving accuracy. Some categories of semi-supervised learning are:

- Generative Models
- Self-Training
- Transductive SVM

D. Neural Networks

Neural networks are derived from biological neural networks. Neural network is not a machine learning algorithm rather it's a framework of many machine learning algorithms which work together to process the data. It works identically to a biological neuron which consists of dendrites which receive electrical signals, soma which process these signals and axon which take output to terminals [1]. Neural networks contain three layers and the data flows unidirectionally through these layers. The layers are:

- Input layer
- Hidden layer
- Output layer

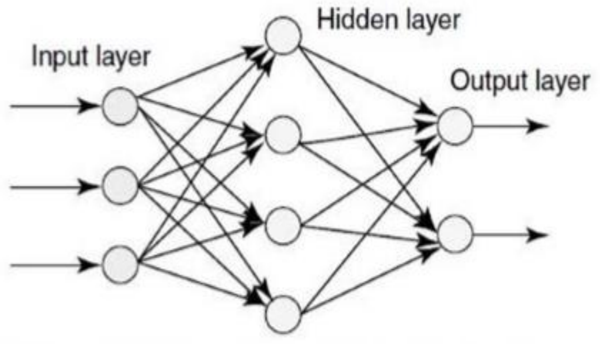


Fig 5 Artificial Neural Network [1]

The input layer takes the input which is processed by the hidden layer and finally output layer is responsible for sends the calculated output [1]. They can follow all three modes of learning and can thus be further classified as:

- Supervised Neural Network
- Unsupervised Neural Network
- Reinforcement Neural Network

Neural networks are popular because of their varied applicability as they are used for various tasks like computer vision, speech and handwriting recognition, robotics etc. [3]. Their drawback is the difficulty in understanding the outputs along with time taken to train the model are being addressed with the advancement of technology.

E. Reinforcement learning

Reinforcement learning employs suitable action which is required to achieve the maximum rewards[2]. It is used to find the best path or behaviour which is required in a particular environment or situation. Unlike supervised leaning the user has no idea about the outcome and thus this model learns from experience to find the best course of action to be taken for a positive result and the next decision to be taken is based on previous one[4]. It employs delayed error and trial and error search to get the ideal behaviour in a situation for maximum efficiency.

III. APPLICATIONS

Machine learning is being widely utilized in various fields like healthcare, agriculture, education, automotive industry etc. Some of the most notable uses of machine learning include:

- **Computer Vision**
  Autonomous vehicles and self-driving cars utilise machine learning algorithms for extracting high level understanding from images and videos for making just like humans do by analysing these for making decisions while driving.

- **Virtual Assistants**
  Virtual assistants like Siri, Alexa, Google Assistant make use of machine learning for natural language processing while also collecting data and refining the search results for a particular user according to their searching habits.

- **Medical Diagnosis and monitoring**
  Various tools based on machine learning are being worked upon which are intended to provide basic healthcare diagnoses based on historical data and patterns. These tools form the diagnosis based on the symptoms of the patient as well as information like gender, age and previous history of illness in the patient. This is especially useful in areas where medical facilities are not available. Machine learning is also being utilized in MRI scanning for identifying abnormalities and anomalies in the scan as well as in ICU for monitoring the patients.

- **Product Recommendations**
  Many online stores and e-commerce websites nowadays used machine learning models to determine the most suitable product recommendations according to each individual user based on their search and purchase history, increasing the likelihood of the customer to buy the recommended product.

- **Facial, Speech and Handwriting recognition**
  These applications are possible because of machine learning models which are trained extensively for these purposed and are employed widely in various applications and products.

Machine learning is being incorporated in manufacturing and production industry as well. The four major applications of machine learning in the field of manufacturing are:

- Predictive Maintenance
- Supply Chain Forecasting
- Quality Control
- Dynamic Scheduling

These techniques are discussed further in the following section.

Applications in Manufacturing

The various applications of machine learning in the manufacturing industry are:

A. Predictive Maintenance

In any major industry the reliability of machines and equipment is very important as the unexpected breakdown of these machines result in economic loss as well as halts the production. The reactive maintenance cost is significantly higher than predictive maintenance cost [5].These breakdowns can be avoided with help of predictive maintenance which continuously monitors the machines and equipment as well as improves the system reliability. A significant amount of time is spent on locating source of the problem and with help of automated supervision system this time can be reduced and the production process can start again. Automation of industry has brought about new technologies. The machines are equipped with various sensors which monitor the status of machine and the process being performed by it in real time. With data from these sensors we apply the machine learning algorithms to perform predictive maintenance to determine the state of machine and predict its failure in the future [6].

Generally, the data is collected form each individual devices and the data needs to be synchronised to get the information about the whole system. Machine learning methods require data such as the fault history, maintenance history as well as the data about condition of the machine. These techniques try to predict the failures due to degradation of equipment, some unexpected breakdowns that might occur due to anomalies [7].

Methods for detecting faults are:

- Modelling and parameter estimation approach
- Discrete observation approach
- Finite automata state machine
- PLC-based logic diagram
- Simulation
- Event-based time-window approach [5]

The fault diagnosis is done based on the training the machine learning models. These models also predict the lowest downtime for maintenance by studying the peak hours and the work hours of the manufacturing industry. The data collected is firstly pre-processed where it is made ready to be input into the model. The information collected form the equipment is sorted into records and classified into categories such as maintenance records, fault history, machine conditions etc which is followed by the designing of functions and sequence aggregation. The predictive maintenance tools mainly follow the following approaches

- **Classification**

  Classification is used to calculate the probability of whether the machines or the equipment will fail over the course of a time period. This is determined with the help of data acquired from the system[3]. Thus, the time for maintenance is also determined which is required to avoid the downtime and breakdown of the machine. For this approach two examples need to be identified which are positive and negative. Each of these examples describes the operating conditions of the asset in question with the help of historical records. The positive example denotes the errors while the negative ones signify regular operation. The objective is to determine the probability of each upcoming example i.e. whether it would fail or work normally within the next time period.

- **Regression**

  Regression is generally used when data exists in a range and used to predict response for continuous values. In predictive maintenance this method is used to determine the Remaining Useful Life (RUL) of the equipment in question which is the time it would function properly before the next failure [7]. The remaining life of each of example as a continuous number is found. Linear regression is the most common method used due to its speed and simplicity.

A proactive maintenance approach is being considered which will enable us to monitor the machine adaptively and issue a warning before a breakdown can occur. This approach aims to eliminate failure equipment and increase mean-time-before-failure to essentially infinity. This will prolong the life of the machine to unlimited length [5].

B. Supply Chain Demand Forecasting

Supply chain management is the management of the goods present in the system. It is very important in the industry and includes storage of raw materials, inventory and includes all the processes which are needed to get the final product. Collaboration among the partners in a firm and sharing information is an important part in reducing demand errors but this approach is not always feasible as the full collaboration between stakeholders at all times is unfeasible [9]. Thus, supply chain forecasting is introduced to improve the savings. Some of the critical issues required that permit supply chain collaboration to be successful are listed below:

- Business interests should be similar
- Management of relationship in the long-term
- Unwillingness to share information
- The convoluted design and management of large-scale chains
- The workers of supply chain should be competent [9]

The machine learning algorithms can perform this forecasting better than humans as the system is very complex. Some of the traditional forecasting methods are listed below.

Naive forecasting

Naïve method is the simples and is generally used as a baseline to which other models are compared to. This method takes the smallest value for future estimates [8].

Moving average

This method uses the average of some previous entries or historical data entries for prediction[8].

Multiple Linear Regression

This method uses previous demand changes to predict the upcoming changes in demand by treating the previous changes as independent variables. [8]The more advanced methods act as universal approximators can approximate any function to an arbitrary accuracy.

These approximators can learn any function and thus all machine learning estimators
fall under this category of universal approximators

Neural Networks

There are a number of neural networks which can be employed but the most popular is feed forward error back-propagation type of networks. In these networks layers of elements (neurons) are used. The output signal from one layer is transferred to the neurons of next layer. Thus, the data moves layer to layer and additional layers called hidden layers can be added between the input and output layers to increase the performance of these neural networks [8].

Support Vector Machines

These machines are a kind of universal approximators which follow the structural risk minimisation principle as opposed to empirical risk minimization principle followed by other methods[8]. SVM algorithms are used to reduce true errors on a random dataset by projecting the data into a higher dataspace or hyperplane and increasing margins between classes and reducing error margins for regression.

C. Quality Control

In today's industry the competition is fierce and the company wants to deliver the best possible product to the customer, but manufacturing process is prone to errors and a few out of millions of products can be prone to errors. This can affect the product and brand image in a negative way [10]. To reduce this and to impose quality control machine learning algorithms and techniques are being applied as the manufacturing processes are becoming more and more complex and the conventional quality control techniques are not feasible[11]. The processes in a production line are linked to one another and these processes being independent can lead to variations and errors in a product [11]. Quality control is achieved by error detection as well as predictive quality control. The data collected is used to predict the pattern of failures and find the exact factors which influence these supply chain problems by analysing large sets of data collected by the company every day and also forecasting the supply demand with accuracy .The patterns in previous data can be used to predict the anomalies and detect the errors in the given dataset like when system detects too much variation in same product it triggers a response[11]. This also reduces the probability of occurrence of bullwhip effect which is the phenomenon in which creates false supply and demand fluctuations.

Quality assessment of products is one of the most important and time-consuming tasks in the manufacturing industry, now this intensive task of testing the products for errors is performed by these algorithms. The quality of product can be improved by optimizing the assembly line as well as predictive maintenance of the machines. One such example of quality control using machine learning is the application of supervised learning and clustering to the process states as well as deep learning.

D. Dynamic Scheduling

Scheduling is a very important part of production process as when there are resources being shared by multiple entities it is very important to schedule these resources efficiently so as to improve the performance of the system [12]. Dispatching rules are used to schedule tasks in a flexible manufacturing system (FMS) and it is very difficult to change these dispatch rules according to changing situations by traditional methods, thus machine learning algorithms provide a way in which dispatch rules can be changed dynamically according to the state of the system [13].The approaches to solve problems in FMS scheduling can be categorized as

- The analytical approach.
- The heuristic approach.
- The simulation-based approach.
- The artificial intelligence-based approach

The first approach views the flexible manufacturing system scheduling problem as an optimization model having certain constraints. To resolve this model an appropriate model is then chosen.

Heuristic approach is generally used to solve NP-complete type problems [13].

Simulation based approach uses simulation for testing dispatching rules and choosing the most efficient one.

Artificial intelligence-based approach uses techniques to dynamically change dispatching rules based on previous data.

The machine learning model can be trained with the data collected from previous runs of the production and techniques like inductive learning, case based learning or neural networks can be applied to make the best possible decision regarding the dispatch rules and the scheduling of resources in the production process to increase efficiency [13].

IV. INDUSTRY 4.0

Machine learning and AI has brought about a revolution in the manufacturing industry where the traditional methods of production are being replaced by the new methods where smarter production is taking place as the machines are able to take decisions and improve automatically by continuous learning process, this revolution is generally known as Industry 4.0.As the industry now uses IOT and sensors which collect a huge amount of data every day ,this data is used to find patterns and make predictions to improve the process of production. This phenomenon uses various machine learning techniques like – reinforcement learning, neural networks, supervised learning among others [14].

One Important aspect of industry 4.0 is the symbiotic relationship between humans and machines. The machines are computationally faster they provide an advantage over humans in terms of some processes like decision making

Table- I: Comparison of humans and machines [15]

Factor		Humans	Machines
Disparity of Quality in	Mechanical Jobs	The differences between individuals are high	Very low
		Training and employee satisfaction can decrease Quality variation and improve quality	Inadequate maintenance can result in degradation during its lifetime
		Problem solving abilities yield vast differences between individuals	Moderate based upon factors such as data quality, precision of algorithm etc.
	Decision Making	Interests (like personal, institutional) may impact this process.	It can be ameliorated by using large datasets to train the machines.
		Decision Complexity and other factors may affect the decision-making process	These factors do not affect machines
Variation in Performance	While a task is being performed/executed	High	Low
		High likelihood of dissatisfaction and weariness from work	Negligible

during ramp up process. Humans on the other hand while being at a disadvantage in some scenarios like iterative processes are sometimes more efficient in making decisions based on experience and in situations with incomplete data. So, there is a need for a human-machine relationship in which both the entities work together to provide maximum efficiency. It has been observed that a human guided exploration in ramp up processes was more efficient than the traditional machine learning approaches receiving close to optimal behaviour by combining best of each of their abilities [16].

One such example is where a human designs work plans which are product specific and a sequencer in a factory which is responsible for scheduling of tasks and it obtains its knowledge of task plans and the order of jobs from information of these work plans. Humans are not able to keep up with the changes required due to situational changes. In a human-machine relationship both can benefit as the humans can gain an understanding from the sequencer recommendations/suggestions and decisions and on the other hand when humans overrule the decision the machine can learn decision making behaviour [15].

Another example might be the sequencing of the tasks to be performed during production. The sequence of tasks is determined by a sequencer which assigns times to different tasks. Generally, the rule of first-in first-out is followed but the sequencer also changes the sequence dynamically if some changes occur in the production process. The sequencer utilises the data based on the work plan prepared by a human worker. This provides us with best of both worlds as humans cannot change sequence of tasks dynamically as quickly as the machines. Thus, with the guiding input from humans the sequencer performs at peak efficiency increasing resulting in a smooth production process.

Thus, the respective capabilities of humans and machines are mutually beneficial to one another as humans are better at analysing and countering unexpected errors and new situations whereas machines are better at iterative tasks and pattern recognition.

V. CHALLENGES

Machine learning apart from manufacturing industry is now being employed in almost every major filed like healthcare, education, military among others. With the advent of Industry 4.0 the role of machine learning algorithms is increasing day by day as they provide accurate predictions and improv the efficiency of production processes in addition to improving the cost savings of companies.
For a machine learning technique to be viable to be used in the industry, the following properties are required -

- Ability to handle various data types (text, numbers, etc)
- Ability to handle noise, and fuzzy data
- Processing should be done in real-time
- Ability to handle high dimensional and large data sets
- Easy to comprehend results should be generated
- Simple implementation [3]

Implementation of these machine learning models in the industry faces many challenges. Some of the challenges which are widely accepted are:

- Using modern technologies in manufacturing.
- Increasing emphasis on producing high-value added commodities
- Using artificial intelligence, advance knowledge
- Products to be manufactured and the process to be employed should be feasible.
- Improvement in services, products.
- Increased cooperation between researchers and manufacturing industry regarding latest technologies.
- Modern manufacturing management standards [17].

A major challenge in implementing these techniques is the gathering of data or specifically usable data as the data generally acquired contains large amount of redundant data which can affect the output. Another challenge is the

Table- II: Suitability of ML techniques in process of manufacturing [17]

Manufacturing requirement	Ability (Theoretical) of Machine Learning to meet the requirements
The system should have the potential to reduce complicated results and present accurate and clear advice to users	Machine learning enables automatic recognition of the existing patterns an thereby deduce algorithms for upcoming events. This enhances the decision-making ability of the process owners.
The system should have the potential to dynamically change according to environmental conditions keeping in check the effort and cost involved	As ML is based on AI, it thus has the ability to adjust/modify dynamically according to the requirement without involving the system operator.
By studying the results, the system should have the potential to improve the pre-existing knowledge	By seeking out patterns in data, machine learning are used to derive new information or even knowledge.
Even without special requirements at the beginning the system should have the potential to use the manufacturing data available to perform the tasks.	Machine learning techniques employ already stored data to convert it into information which can be used in various applications like making predictions in a system
The inter and intra relations should be found and correlated by the system	Detection of irregularities and patterns which define data is the main aim of some machine learning approaches
Ability of handling high-dimensional problems and datasets	Some problems like over-fitting need to be taken care of even in machine learning approaches which have the ability to handle data of high dimensionality

pre-processing which can have problems like unbalanced or incomplete/incorrect data set values and adding or replacing values can influence the results. Another hurdle is to select the best machine learning algorithm suited for the situation among a large number of such algorithms as there are a number of factors which need to be addressed while choosing the algorithm which would be best suited to the problem. General approach for selection of algorithm consists of three steps-

1. The data is analysed to see whether it is labelled or unlabeled to choose between the best approach – supervised, unsupervised or reinforcement learning
2. Checking the applicability of current algorithms based on various factors like data types, amount of data available etc.
3. Investigation of algorithms in similar cases to identify the best algorithm for the given problem.

Nowadays a combination or hybrid technique is also used which use more than one algorithm to tackle the problem with better accuracy and results than a single algorithm.
Security of data is also a major concern as most data in this age is being stored on cloud platforms making this data vulnerable and it is possible to gain remote access to these machines if appropriate security measures are not taken [14]. Lastly, the interpretation of results acquired from machine learning techniques is very difficult along with difficulties in finding the talent for such jobs as this is a relatively new field and finding qualified individuals is a difficult [17].

VI. Suitability of Machine Learning Algorithms

Many machine learning algorithms handle large amounts of data and one of the problems they might face while handling data with high-dimensionality is over-fitting. This problem rarely occurs due to power of modern algorithms but there exist techniques to reduce dimensions should the problem occur. In techniques like Support Vector Machines the need for reducing dimensions is reduced as dimensionality is not a practical problem. This in turn means that we can be more liberal towards information which may seem irrelevant and include it as it may become relevant under certain circumstances down the road [17].

VII. ADVANTAGES

Machine learning approaches are used to solve optimization problems in intelligent manufacturing system as they are competent enough to tackle NP complete problems. Most of the problems encountered in manufacturing are data rich, machine learning helps to improve understanding of data by extracting implicit relationships from large data-sets due to its capability of handling data of high dimensions and multi-variate datasets. Machine learning is universally believed to improve utilization of resources and reduces cycle time and scrap in some of the NP-hard problems. In complex processes like manufacturing of semiconductors it provides tools for continuous quality control and management of production process.

As mentioned earlier, machine learning algorithms are equipped to handle problems and data of high dimensions but to varying degrees depending upon the algorithms or technique chosen and cannot be generalised. Some algorithms like Distributed Hierarchical Decision Tree are known to handle the high-dimensional data and problems better than other and are thus preferred in manufacturing process as large amount of data encountered is of high dimensionality.
Various tools and programs are now available nowadays for application of machine learning algorithms and has made process easier offering features like adjustment of parameters easily in order to improve classification performance.
We prefer of machine learning algorithms due to their ability of identifying implicit relationships in large and chaotic data-sets by using Supervised, Unsupervised or Reinforcement learning depending upon data-set which was not possible until a few years ago. Thus, the huge amount of data collected over the decades can now be utilised for improving performance and production process.

Another advantage of machine learning algorithms is their ability to adapt dynamically to the situation. Manufacturing and production processes are very uncertain and problems can arise at any time for instance breaking down of a machine or bottleneck in some machines. Machine learning allows the production process to run smoothly by techniques like continuous monitoring of machines, dynamically changing sequence of jobs scheduled adapting to change in circumstances in production or scheduling repair and maintenance of machines during off hours or when the machine is not being utilised. This enables production line to run smoothly without any interruption leading to an increased production rate which may not be achievable by traditional methods and practices. [17]

VIII. CURRENT TRENDS AND FUTURE OF MACHINE LEARNING IN MANUFACTURING

Machine learning is experiencing rapid progress in various fields including manufacturing and industrial production. The two most important directions taken nowadays in the field of machine learning are:

Scaling up of machine learning algorithms

During the last couple of decades there has been a drastic increase in the amount of data collected in every field including the manufacturing industry. There is large amount of data collected during the production process in every step from production planning, design to maintenance. Scaling up of machine learning algorithms is thus required to accommodate and process these large data-sets in the best way possible. Data mining which is an field which uses these algorithms has highlighted this demand as it requires processing millions of training examples and handling terabytes of data [3].

This scaling of algorithms can be done by two approaches

- By producing more efficient algorithms
- By partitioning of data

In the first approach the aim is to generate machine learning algorithms which are more efficient or improve the performance of current algorithms to improve the processing time and performance of machine learning model. This includes techniques like optimizing search an representation, finding approximate result instead of an exact one.

In the second approach the data set is broken up into subsets which avoid running the algorithms on very large sets of data. Generally, data is divided into subsets with common features to make processing simple. This approach also aims to avoid memory management issues [3].

Learning Multiple Models

This is another area currently being heavily researched. This approach involves increasing accuracy in supervised learning. This is achieved through using ensemble of classifiers or multiple models. These refer to combining classifiers in some fashion rather than using them individually. It is believed that ensemble of classifiers are an improvement over individual classifiers in terms of accuracy, provided that the individual classifiers are themselves diversified and accurate. This means that the individual classifiers should be accurate meaning they should provide results which are an improvement over guessing and thus they should which should result in varying errors when encountering new data points.

There are three approaches for generating multiple classifiers

Bagging

In bagging technique, a replication is generated of the training set by sampling with replacement from the training instances [3] with each training set having same size as that of the original data. Examples may appear multiple times or not appear at all. These types of training sets are defined as bootstrap replicate of original training set with technique being called bootstrap aggregating. A training classifier is derived from each replication of training set and all the classifiers generated then use voting scheme to classify each instance in test case.

Boosting

In this technique boosting maintains for each instance in the training set a weight which signifies its importance instead of drawing bootstrap samples from the original instance successively. The higher the weight of an instance, the more it influences the learned classifier. The weights are adjusted during each round according to the performance of corresponding classifier. A weighted voting scheme where every component classifier contributes with different strength based on their accuracy of training is used to construct the final classifier.

Error-Correcting Output Coding (ECOC)

This method was devised to solve multiple two-class problem as a means of handling multiclass problems. In Error-Correcting Output Coding each binary classification function corresponding to a unique partitioning of class is encoded by a set of output bits of classes. Class predictions are then generated by combining all functions which is derived from algorithms that use ECOC [3].

Machine learning has been employed by many Multi-National Companies in their production plants around the world. However, the machine learning field has a lot of scope for improvement as it is a relatively new field. There are many areas which need further research to improve and optimize the techniques we employ for better results. Many of these are already being worked upon. Some of these areas are:

- Utilizing existing knowledge
- Learning in structural domains
- Learning with continuous variables
- Constructive / meta learning
- Integration of machine learning with existing engineering software
- Exploratory systems [21]

IX. CONCLUSION

Machine learning has become a very important part of the manufacturing industry in recent times and its importance will increase further in the future. It has brought about the industry 4.0 revolution and is changing the face of production process with techniques like neural networks.

The manufacturing process has become more efficient and the production line more streamlined with machine learning various applications like predictive maintenance, scheduling, quality control etc. The symbiotic relationship between humans and machines is very important as it results in near optimal performance. The development of machine learning algorithms is advancing at a rapid rate and along with the increase in computational power, the challenges like interpretation of output and data and pre-processing can be solved in the future leading to industry-wide implementation of machine learning techniques for improved efficiency and production.

REFERENCES

1. A. Dey, "Machine learning algorithms: a review.", International Journal of Computer Science and Information Technologies, vol .7, no. 3, 2016, pp. 1174-1179.
2. T. O. Ayodele, "Types of machine learning algorithms.", New advances in machine learning. IntechOpen, 2010, pp.19-48
3. D. T. Pham, A. A. Afify, "Machine-Learning Techniques and Their Applications in Manufacturing." Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, vol. 219, no. 5, May 2005, pp. 395–412.
4. L.Monostori, A.Markus, H.Van Brussel, E.Westkämpfer, "Machine learning approaches to manufacturing." CIRP annals, vol. 45, no. 2, 1996, pp. 675-712.
5. J.Lee, "Machine performance monitoring and proactive maintenance in computer-integrated manufacturing: review and perspective.", International Journal of computer integrated manufacturing, vol. 8, no. 5, 1995, pp. 370-380.
6. A. Kanawaday and A. Sane, "Machine learning for predictive maintenance of industrial machines using IoT sensor data," *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, 2017, pp. 87-90.
7. G.A. Susto, J. Wan, S. Pampuri, M. Zanon, A.B. Johnston, P.G. O'Hara, S. McLoone,"An adaptive machine learning decision system for flexible predictive maintenance," *2014 IEEE International Conference on Automation Science and Engineering (CASE)*, Taipei, 2014, pp. 806-811.
8. R. Carbonneau, K. Laframboise, R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting". European Journal of Operational Research, vol.184, no. 3, 2008, pp.1140-1154.
9. R. Carbonneau, R. Vahidov, K. Laframboise. "Machine Learning-Based Demand forecasting in supply chains." International Journal of Intelligent Information Technologies (IJIIT), vol. 3, no. 4, 2007, pp. 40-57.
10. C. A. Escobar, R. Morales-Menendez, "Machine learning techniques for quality control in high conformance manufacturing environment". Advances in Mechanical Engineering, vol. 10, no. 2, 2018, DOI: 10.1177/1687814018755519.
11. T. Wuest, C. Irgens, K.D. Thoben, "An approach to monitoring quality in manufacturing using supervised machine learning on product state data.", Journal of Intelligent Manufacturing, vol. 25, no. 5, 2014, pp. 1167-1180.
12. P. Priore, D. de la Fuente García, I.F. Quesada, " Manufacturing Systems Scheduling through Machine Learning", NC, 1998, pp. 914-917.
13. P. Priore, D. de la Fuente, A. Gomez, J. Puente, "A review of machine learning in dynamic scheduling of flexible manufacturing systems", Artificial Intelligence for Engineering Design, Analysis and Manufacturing, vol. 15, no. 3, 2001, pp.251-263.
14. B. Bajic, I. Cosic, M. Lazarevic, N. Sremcev, A. Rikalovic, "Machine Learning Techniques for Smart Manufacturing: Applications and Challenges in Industry 4.0", Department of Industrial Engineering and Management Novi Sad, Serbia, 2018, pp.29.
15. F. Ansari, S. Erol, W. Sihn, "Rethinking Human-Machine Learning in Industry 4.0: How Does the Paradigm Shift Treat the Role of Human Learning?" Procedia Manufacturing, vol. 23, 2018, pp.117-122.
16. S. Doltsinis, P. Ferreira, N. Lohse, "A symbiotic human–machine learning approach for production ramp-up." IEEE Transactions on Human-Machine Systems, vol. 48, no. 3, 2017, pp.229-240.
17. T. Wuest, D. Weimer, C. Irgens, K.D. Thoben, "Machine learning in manufacturing: advantages, challenges, and applications." Production & Manufacturing Research, vol. 4, no. 1, 2016, pp. 23-45.
18. R. Gangula, G. Nandam, C. Sudha, S. Rekha, "Usage of Machine Learning Algorithms in Data Mining", International Journal of Recent Technology and Engineering (IJRTE), vol.8, no. 1C2, 2019, pp. 411-414
19. L. Rokach and O. Maimon, "Top-down induction of decision trees classifiers - a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 35, no. 4, 2005, pp. 476-487.
20. http://pypr.sourceforge.net/kmeans.html (Accessed on oct. 12, 2019)
21. B.L. Whitehall, S.CY. Lu, "Machine learning in engineering automation—The present and the future.", Computers in Industry, vol. 17, no. 2-3, 1991, pp. 91-100.

AUTHORS PROFILE

Vimanyu Chopra is an undergraduate scholar at DAV Institute of Engineering and Technology, Jalandhar, Punjab, India.

Devinder Priyadarshi is Assistant Professor in Department of Mechanical Engineering at DAV Institute of Engineering and Technology, Jalandhar, Punjab, India.