ISSN No: - 2456 –2165

Impact of Machine Learning on Manufacturing Industries

Vignesh Manoj Varier
Sir M.Visvesvaraya Institute of Technology
Vigneshvarier@Gmail.Com

Dhruv Sharma
Sir M.Visvesvaraya Institute of Technology
Dhruv95@Live.Com

Srinath Karmungi
Sir M.Visvesvaraya Institute of Technology
Srinath.Karmungi@Gmail.Com

Abstract— This paper emphasizes on developing new methods, evaluating the approaches and conveying the results for the problems faced in manufacturing industry ingeniously using machine learning.As the manufacturers want to augment their profit with least capital investment, a model has been proposed in which machine learning is incorporated in production.Linear regression algorithm has been utilized for predicting the outcome for the given input and gradient descent for optimization of result. It can help maximize the production and increase profit for the company by predicting the outcome so that the company can take the required action or anticipate the problem beforehand.

Keywords—Machine Learning, Linear Regression, Correlation, Gradient Descent, Scatter Diagram

I. INTRODUCTION

In worldwide market, cost of production is something that the manufacturers have to compete for, as cost of raw materials will be almost similar worldwide. The challenge faced by the various industries is the effective analysis of data, and on the other hand making big changes or developments via investments in corporate structure or new machinery. Machine learning offers a propitious way for manufacturers to face both these problems as they are in excellent position to employ learning techniques with their massive resource of registered production data.

Machine learning is a subfield of data science that evolved from study of pattern recognition and computational learning theory. It uses algorithms that generalize and perform the assigned task very effectively. Nowadays machine learning is being widely used in many manufacturing industries for various purposes such as spam filters, fraud deduction, drug designing, data prediction, speech recognition, diagnosis of diseases etc. This paper emphasizes on designing a model which uses machine learning in a manufacturing industry to maximize its efficiency. A computer learns based on previously provided sample data called as training sets to predict output for any input and its efficiency improves with experience.

Machine learning systems consist of three major parts: Model, Parameters and Learner. Model is the system that makes predictions or identifications, parameters are the factors used by the model to form its decisions and learner adjusts the model by analyzing differences in predictions versus actual outcome. Now the question which usually arises is, “Can we invest the least and create acontingency model, supported by data?” This is an important question, since managers often like to know if there is some scenario in which operational cost is low and it is supported by data; this method would be able to help find such scenarios directly.

A. *Current Scenario in Production*

Manufacturing can be defined as process of converting raw materials, components or parts into finished goods that meet the customer’s expectations or specifications. It consists of a man-machine system with division of labor in a large scale production.

One of the precarious scenarios is manufacturing market leadership. The need to include adaptability in machinery and fabrication is now important. Machines used for manufacturing currently can only be used for core production without the prospect of improvising. With the varying information, machines are not able to extemporize and harness changes. This is resulting in an increased manufacturing lead time for production.

The change in environment which contributes to increase in expenditure is something that the machine is not able to deal with nowadays. The superiority of products also deteriorates as complication in process selectionrises.

So few of the main challenges faced by a manufacturing industry are:

- Lack of adaptability
- Lack of skilled labor
- Delivery responsiveness
- Big data management
- Remaining cost competitive
- Optimum utilization of resources
- Overheating of tool

ISSN No: - 2456 –2165

B. *Machine Learning Concepts*

Machine learning is a field in the cross-section between computer science, statistics, artificial intelligence and mathematical optimization. Tom Mitchell in his book Machine learning defined it as "A computer is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

There machine learning tasks are basically of two types:

- Supervised Learning
- Unsupervised Learning

In supervised learning the algorithm is given a set of training data in which it is told the "correct answer" for each example. In unsupervised learning no special appellation is given, the computer itself should find a structure for the given data.
One of the main problems faced in machine learning is overfitting and underfitting. In general, it is the irregularity where error diminutions in training set but in the test set error surges. It occurs when model is particularly convoluted. The simplest means of avoiding overfitting is to split the data into two sets: validation and training set.A lot of data can also be used so that accuracy of algorithm increases.

Using cross validation is a standard method in applied machine learning for estimating model accuracy on unseen data. If you have the data, using a validation dataset is also an excellent practice.

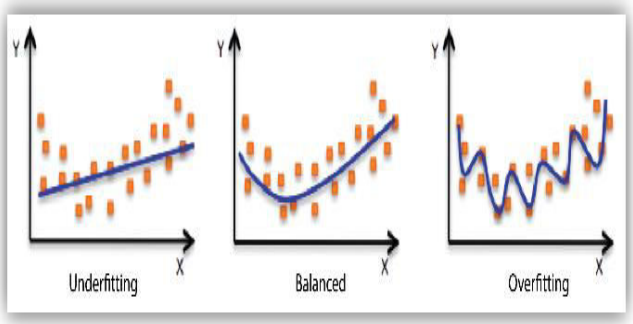

Fig. 1: Picture from Google Images

II. HISTORY

The advance of machine learning plays a pivotal role in the development of artificial intelligence. Arthur Samuel an IBM scientist in 1952 made the game of checkers to create the first learning program. Each time he played the game it improved in a 'supervised learning mode'. Later in 1957 the Perceptron model was built which was a type of neural network which can be used to solve intricate problems.

In the 1990's people began to apply machine learning in data mining, adaptive software, web applications, text learning and language learning. Finally the new epoch brought about a whole range of wide opportunities and machine learning was used where adaptive programming was required. These programs were capable of identifying patterns, learning from experience and theorizing new information from data.

III. PROPOSED MODEL

A concept which can be used for formulation of a machine learning algorithm is the use of linear regression. Linear regression in statistics is an approach for modeling the relationship between a scalar dependent variable Y and one or more explanatory variables (or independent variables) denoted X. Regression and correlation analysis determines both the nature and strength of relationship between variables.

Regression involves the improvement of an estimating equation which is a formula that relates its known variables and the unknown variables. On obtaining a pattern or a trend, we use correlation analysis to determine degree to which the variables are related.

Now machine learning concept can be used in predicting the required outcome in following areas:

- Cost of raw materials
- Required capital investment and resources
- Performance of different branches
- Tool wear and overheating
- Delivery responsiveness
- Big data management
- Customer reviews
- Market trend
- Bottleneck in production line
- Site selection
- Capacity Planning
- Product failure

Also machine learning can be used in trying to achieve the following tasks in an industry:

- Computer vision
- Lean manufacturing

So over time the machine can foretell outcomes very accurately. As a part of the process, we must decide what data must be kept for amenability reasons, what data can be discarded of and what data should be kept and examined in order to improve current business processes or provide a business with aneconomical advantage. This process requires careful data classification so that finally, the profit

can be maximized. So we can use machine learning in following topics in the stated way.

1. Cost of Raw Materials

Once the cost of raw materials for the previously purchased stock and similarly collecting data from previous years and feeding it into the algorithm can help us know how much capital investment will be required in future for buying the raw materials. So this helps us plan accordingly and not face a financial crisis all of a sudden. For predicting the cost of raw materials using linear regression, the cost of raw material can be plotted on y-axis and the respective years on x-axis. Now using regression technique the cost of raw material can be predicted to certain amount of accuracy.

If they realize that the cost of materials is going to hike, they can buy and store a certain amount of raw material at present moment and store it in the inventory for future use. On a large scale this can save a lot of money.

2. Performance of Different Branches

Machine learning can also help realize a company which branch or branches of it is not performing up to the mark. They can accordingly close down or shift their branch based on the prediction. The sales data of previous years is gathered and then using the algorithm, it can predict the future sales possible from that particular branch. Now if this is showing negative or minimal profit means the company take the necessary action. The sales of a certain branch can be plotted on y-axis and the respective year on x-axis. Then regression technique can be used to foretell the future sales from the considered branch of the company.

3. Required Capital Investment and Resources

Machines have been the focal point for optimum utilization of resources. For example in an automobile manufacturing industry the plant in Karnataka requires more than one in Tamil Nadu, since the sales have been more in Karnataka. So now gathering previous year data and feed it to algorithm as before it can tell which region will require the particular resource more and which will require comparatively less, so that there is no accumulation of resources at one place and the investment is made at the right states. The resources are appropriately distributed to the different plants of the company.

4. Tool Wear and Overheating

Tool wear and overheating of the workpiece has been a major problem in the industry. The application of adequate amount of lubricants and timing play a vital role in maintaining the temperature in the required limits. Now collecting this data from previous experiences and using algorithm we can help predict the amount and the correct timing of application of lubricants. This helps increase the tool life and machine life. For predicting the moment at which tool will overheat using linear regression, the temperature of tool can be plotted on y-axis and the point of tool failure on x-axis. Now using regression technique the point at which tool overheats can be predicted to certain amount of accuracy.

5. Delivery Responsiveness

Delivery responsiveness is how fast goods are manufactured and delivered to customer. A company should have the least lead time possible. So once all the orders are received one needs to know how to achieve minimal lead time and get the goods delivered fast. Again using the algorithm and providing input as the previous orders, it will tell how many working shifts, labourers and delivery time required for the given set of orders.

Machine learning can also be used to know how long will it take for the product dispatched from warehouse will take to reach the customer. The exact timing can be predicted and the customer can stay more informed.. For predicting the time taken for delivery using linear regression, the time taken can be plotted on y-axis and the respective warehouse distances on x-axis. Now using regression technique the cost of raw material can be predicted to certain amount of accuracy.

6. Big Data Management

Big data management is the organization, administration and controlof large volumes of both structured and unstructured data. The goal of big data management is to ensure a great level of data quality and convenience for business intelligence and big data analytics applications. A structured data is fairly easier to organize and analyse whereas an unstructured data is very tedious and time consuming to be organized and analysed by individuals or group of individuals. This is where machine learning can step in and help speed up the process and do the obligatory work in lesser amount of time.Unsupervised learning methodscan be used to make sense of unstructured data. The total time taken to analyse through this concept method will be comparatively less than manual analysing of data.

7. Customer Reviews

The company can use machine learning to find out how have the audience responded to their product, what are its key advantages and disadvantages and get a review of product so that later they can improve on it.

The algorithm can be used to read innumerable posts which are trending on social communication websites such as twitter and Facebook and filter the comment as positive or negative review. It reads all the data from such websites and groups similar data collected, it basically looks for keywords and clusters them together. The best suitable machine learning algorithm for this would be k-nearest neighbours.

8. Market Trend

It can also be used to know the trend in the market or trend in the society or amongst the youngsters, which will help the company to know the customer requirements and develop a product accordingly. They will be able to know what and how their product should be so that it can run in market say in terms of appearance, working etc. The key point here is that the accuracy of this prediction increases with experience or increase in trial data.

9. Bottleneck in Production Line

It can also help predict where a bottleneck situation might occur during production process. For predicting the possible place where bottleneck would occur in production line using linear regression, the number of products being handled by machine can be plotted on y-axis and the places where bottleneck has occurred previously on x-axis. Now using regression technique the cost of raw material can be predicted and as a result it improves throughput by strengthening the weakest link in the manufacturing process.

10. Site Selection

One of the main decisions to be made while starting an industry is deciding the location of the industry. The following points have to be kept in mind while deciding the site for the factory:

- To abate shipping costs, both for raw materials coming into the plant and for finished goods going out, managers often want to localise plants close to suppliers, customers, or both.
- Locate in areas with bounteous numbers of skilled workers.
- Locations where managers and workers families can enjoy living with compulsory necessities around.
- Costs for resources and other expenses—land,labour, construction, utilities, and taxes—are low.
- Encouraging business climate—one in which, for example, local governments might offer financial incentives (such as tax breaks) to entice them to do business in their locales.

So all this constraints have to be kept in mind and the most optimum solution has to be found. Gradient descent can be used to find an optimum solution complying with the constraints.

11. Capacity Planning

Capacity planning is testing task in hand due to the need for very accurate forecasting of data. Unfortunately, dwindling to balance capacity and projected demand can be seriously detrimental to your bottom line. If the capacity is set too low (and so produce less than one should), company won't be able to meet demand and one will lose sales and customers or ifone sets capacity too high (and turn out more units than required), company will waste resources and inflate operating costs. So it needs to be planned as precisely as possible.

For a basic level of capacity planning using linear regression, the demands of various yearscan be plotted on y-axis and the respective years on x-axis. Now using regression technique the cost of raw material can be forecasted with satisfactory amount of accuracy.

12. Product Failure

It can also be used to predict when a structure or product or machine might fail, so that the company can avoid a crisis scenario during the usage or handling of product.For predicting the point of product failure using linear regression, the point of failures of previous versions of products can be plotted on y-axis and the stress or force exposed to on x-axis. Now using linear regression technique the cost of raw material can be predicted to certain amount of accuracy.

Now following concepts can be achieved in an industry by the stated ways.

i. Lean Manufacturing

Lean manufacturing can be defined as the philosophy that continuously shortens the time between customer order and shipment by the elimination of anything that increases the cost of production and time of delivery. Lean manufacturing can be integrated into a business philosophy having the following objectives in mind.

- Elimination of waste
- Shorter lead time
- Increased efficiency

Lean manufacturing can be converted into a business strategy and can be installed in the organisation. Elimination of waste enables the operation to be steered swiftly and efficiently.A40% gain in overall productivity can be achieved .Shorter lead times can be obtained with careful evaluation and thoughtful of the process. The goal of lean manufacturing is to achieve solitary piece flow in every operation possible. Single piece flow can be described as a methodology of production in which batch sizes are supplanted by working on one product at a time.

The benefits of lean manufacturing include amended quality and fewer defects, reduced inventory enhancement of overall manufacturing flexibility.

ISSN No: - 2456 –2165

Overproduction is the accumulation of unsalable inventories in the hands of businesses. So by using this concept and minimize overproduction to a negligible amount. The waiting time between diverse processes during manufacturing can also be reduced, since the time taken by each process can be predicted quite accurately. So the labourers can be appointed and machines can be allocated and made availableaccordingly.

ii. Computer vision

Computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images .Computer vision helps in analysing and understanding digital images from the actual world, extract relevant data and apply it to make suitable decisions .As a scientific field computer vision is the theory behind artificial systems that extract information from images, the image data can take many forms such video, audio and multidimensional images. The sub- domain of computer vision includes video tracking, object recognition.

The various application of computer include

- Controlling events with respect to surrounding.
- Organizing
- A better computer human interaction.
- Automatic inspection

The arena of artificial intelligence benefits the most with the application of computer vision. Robots can work effectively and efficiently without the need for human interaction. Computer vision is a field that has seen rapid improvement since the development of faster processors that enable it to concurrently synthesis information and act accordingly.

Now this concept can be utilized for material handling system and for visual inspection of a product.

In order to formulate an efficient material handling system, one needs to follow the principles of planning, ergonomics, space utilization, life cycle cost and automation. So machine learning can be effectively used to create a good material handling system. The robotic arms can be used to pick and place the products at the required places. Therefore we can use object recognition techniques to automate the process for faster and efficient transporting of products from one place to another.

Now even visual inspection can be done using essential inspecting setups. The data received after scanning the product can be used to find out if product is decamped. Now using machine learning provides the extra benefit of finding or foretelling when a product might actually be defected.

IV. ARCHITECTURE

A. Estimating Relation Using Regression Line

The most basic equation used to represent the relation between two variables is given below

$$Y=a+bX$$

Where,

Y is the dependent variable

X is the independent variable

a is the y intercept

b is the slope of the straight line

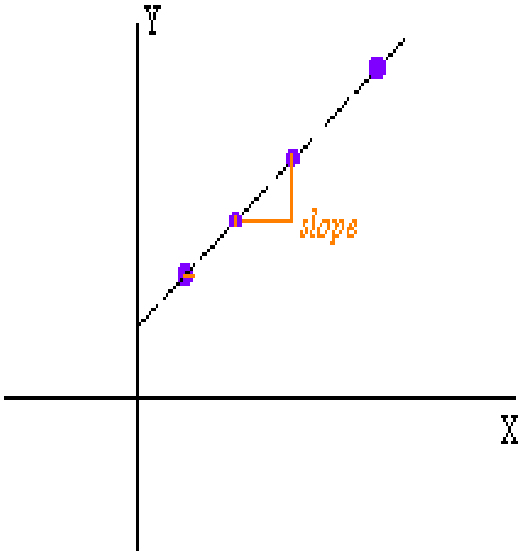

Figure 2: Relationship between X and Y

Now the estimating line is given by $\hat{Y}=a+bX$

$\hat{Y}$ is the individual value of the estimated point

The slope of the best fitting regression line

$$b = \frac{(\sum XY - (n\bar{X}\bar{Y}))}{\sum X^2 - nX^2}$$

$\bar{X}$and$\bar{Y}$ is the mean of the values of the individual variables

n is the number of data points

After plotting the corresponding points in X-Y axis it will look as shown in Figure2 and is called a scatter diagram which shows how values from data are scattered.

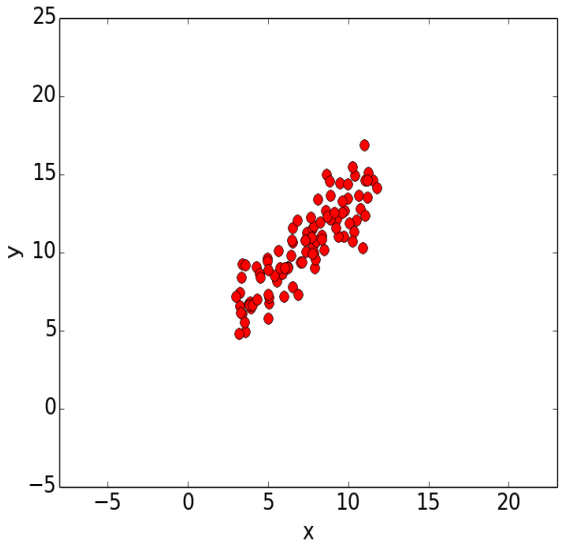

Figure 3: Scatter Diagram

Now the linear regression algorithm tries to find the best fitting line for the data. Using the formula the slope and intercept of the line is figured and it is used to draw the regression line as shown in Figure 3. The required output can now be predicted by interpolating from the regression line for the opposite input. It is quite accurate and its accuracy increases with more the number of training sets provided. Training a linear regression model is usually much faster thanany other models and it is simple yet effective.

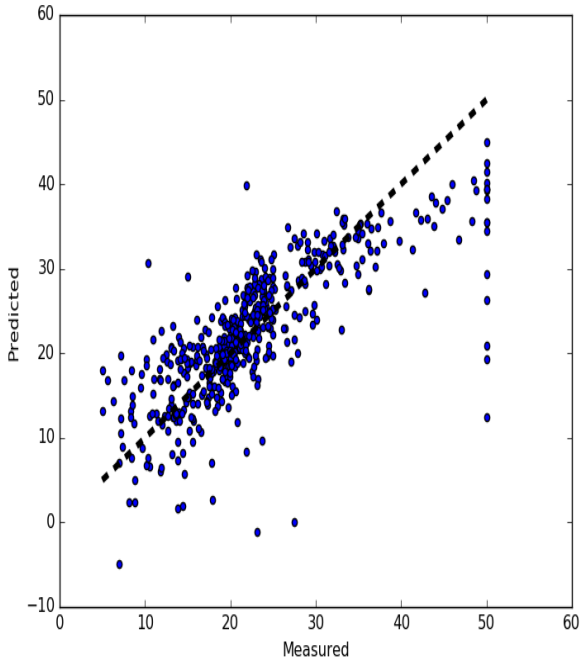

Figure 4: Regression Line for the Given Set (Black Line Indicates the Regression Line)

B. Standard Error of Estimate

The standard error of the estimate measures the variability of the observed values around the regression line.
The standard error is obtained using the formula:

$$S_e = \sqrt{(\sum Y^2 - a\sum Y - b\sum XY)/\sqrt{n-2}}$$

The larger the value of standard error,the greater the smattering of points around the negative line .Conversely if standard error is zero then we expect the estimating equation to be a perfect estimator of the dependent variable. In that case all data points would lie directly on the regression line and no points would be scattered around it.

C. Gradient Descent

Another method used for linear regression in machine learning algorithms is gradient descent. It is basically the method of curtailing the square of error of a function. It starts with a set of values and then moves iteratively towards a value which minimizes the error function. Our basic requirement is to make sure J(θ) decreases with every iteration.

$$\theta := \theta - \alpha \frac{\partial}{\partial\theta} J(\theta)$$

Where α is the learning rate andJ(θ) is called as the cost function and defined as

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}[h_\theta(x^{(i)}) - y^{(i)}]^2$$

wherem is the total number of training examples
and $h_\theta(x^{(i)})$ is the hypothesis function and is defined as

$$h_\theta(x^{(i)}) = \theta_0 + \theta_1(x^{(i)})$$

Where θ_0, θ_1 are the y intercept and slope for the linear regression respectively.

The learning rate α needs to be chosen aptly because if we choose a very small learning rate then too many computations will have to be made to reach the minimum of the given hypothesis function which is time consuming. But if learning rate is too large then it might overshoot the minimum point and gradient descent can diverge. Gradient descent can converge to local minimum even with a fixed learning rate. As we approach local minimum it will automatically take smaller steps.

Input training data set
Preparation of Scatter diagram
Formation of hypothesis function
Preparation of Cost function J(θ)
Calculate slope at “x” position
If slope is negative, move right
If slope is positive, move left
Repeat above until slope = = 0
Output

Table: 4 Basic Algorithms for Gradient Regression

V. CONCLUSION

Efficient production system is an utmost requirement for any manufacturing industry. The aim of this proposed approach is to use machine learning for the fabrication of products. A machine learning algorithm will run at the industryor plant server and its function is to do sustain and accomplish production with minimal error. Initially, previously recorded data are poised and then it is fed into the algorithm which helps predict the outcome and help plan consequently. Comparative studies envisage that the accuracy of this model is around 85% over an eclectic range of input data and results are exciting.

VI. ACKNOWLEDGMENT

We would like to thank our college faculty for providing the opportunity to undertake this project .We would also like to thank all the authors and researchers for making information accessible and available to us .The project helped us to discover and learn facts about the various applications of machine learning and its potential in the manufacturing industry.

REFERENCES

[1]. Statistics for engineers by Douglas Montgomery
[2]. Statistics for engineers by Miller & Miller
[3]. http://openclassroom.stanford.edu
[4]. https://en.wikipedia.org/wiki/Regression_analysis
[5]. https://en.wikipedia.org/wiki/Machine_learning
[6]. Crafting papers on machine learning –Pat Langley
[7]. Machine learning that matters – Kiri.L.Wagstaff
[8]. Machine learning with operational cost – Theja Tulandhula and Synthia Rudin