Contradicting the Hypothesis of Data Analytics with the Help of a Use-Case Related to Manufacturing Industry

Shweta .M. Devadiga

Department of Computer Science, Maharashtra, India.

---***---

Abstract - Manufacturing industry is now heavily relying on data analytics for its growth. Data analytics helps in analyzing the present data which further helps us in decision making. This research paper attempts to bring out a flow of how data analytics techniques can be used to find out trends in manufacturing industry and forecast future demand. Additionally, Covid-19 has pushed the industry towards online practices for most of its operations. This requirement has created need for automation and tracking at several levels. In the current research paper, these needs are addressed as well.

***Key Words*:** Data Analytics, Descriptive analytics, Diagnostic analytics, Predictive analytics, Prescriptive analytics, Manufacturing.

1. INTRODUCTION

We are now in the era of fourth Industrial Revolution which deals with industrial automation. Industries like chemical, manufacturing, food and beverage, textile are facing challenges in reducing waste such as material waste, time resources and reducing machine downtime. These industries are using technologies like IOT, Digital Twin, 3D printing, Cloud, AR/VR and many other automation tools.

Manufacturing industry uses Industrial IOT to prevent overstocking and understocking of inventory. In chemical industries, every plant wants to reach maximum level of performance. Automation can help reduce risks and improve productivity which in turn lowers operational costs. Cloud computing can be used in agricultural activities such as production planning [1][2], operational planning [1], various management strategies for indoor farming [2], storage monitoring [2] and reliable data sharing between farmers and consumers [2]. Textile industry uses AI for trend predictions, machine diagnosis and analyzing large chunk of data collected from manufacturing, purchase, marketing and logistics.

COVID-19 pandemic has disrupted operations of manufacturing industries on massive scale. It is forcing manufacturers to rethink on starting the production considering safety protocols and using new ways for manufacturing operations. Owing to this situation the demand for certain products such as sanitizers, digital devices, tablets, masks, PPE kits has grown up suddenly and industries is facing difficulties to meet the demand from several reasons such as unavailability of sufficient raw material, skilled labors and manufacturing process. In order to meet these demands certain industries have switched to automations and digitization of manufacturing processes [3].

Automation and digitization of manufacturing processes essentially involves data analytics. Data analytics involves collection of lots of data through various sensors inorder to track the process. Processing of this data involves cleaning, modeling and transforming data in order to lay out conclusions. Data analytics is classified into 4 types – Descriptive, Diagnostic, Predictive and Prescriptive. Descriptive analytics can be used to find out the trends in the data or to find out the current scenario in the industry. Diagnostic analytics helps us determine the reason behind the trend that is seen in the data. Predictive analytics predicts the future outcomes. Prescriptive analytics allows us to provide suggestions to the industry in order to improve the productivity.

In this research paper, for every case study relevant techniques of analytics are applied and the output of each technique is used as input to the other. These techniques if performed independently would not result into productive results. If in case, directly predictive analytics is applied then our study fails to identify the critical features and tend to use all the features which degrades the performance thereby giving us a faulty result. Similarly, prescriptive analytics cannot be applied directly to the case study. Hence, we need to first carry out diagnostic analytics as it gives us the root cause of the trend that is seen and then based on the results we move towards prescriptive. This research paper also focuses on Data Analytics methods used for the digitization and automations of the processes relevant to dealing with COVID -19 situations.

This research paper is organized into 4 sections. In the first section methodology is explained. In methodology there are 3 parts – Expenses, Employee and Domain. For these parts sample data is collected for analysis purpose. Data analytics methods are applied to data of each part and how digitization and automation can help carry out manufacturing processes are explained. Second section deals with implementation of machine learning algorithms in order to analyze data, predict the future and to manage manufacturing operations in covid-19 situation. Third part talks about the results that we get after implementation of algorithms on collected data. Finally we have conclusion summarizing the entire paper.

2. METHODOLOGY

In this study, Data Analytics techniques are applied for manufacturing industry. The data collected for this purpose is divided into three types of data– 1) Expenses 2) Employee 3) Domain. The data considered for this study is created in order to demonstrate how data analytics matters to manufacturing industry.

2.1 Data analytics

Data analytics is a process of cleaning, transforming and modelling collected data. This process helps the industry to take further decisions. Decisions in our day-to-day life are taken by analyzing our past or future results. For a business to keep growing all you need is analyze the data. If your business is not growing, then you need to look back at your mistakes and plan again. And even if it is growing, then you need to look forward towards making business grow more. Data analytics includes descriptive, diagnostic, predictive and prescriptive methods.

2.1.1 Descriptive analytics

This method is the starting point in data analytics as it tries to find out the trends in the past. It helps in identifying the areas of strengths and weaknesses in industry. Some common methods to represent descriptive analytics are bar charts, pie charts, histograms, line chart and others.

2.1.2 Diagnostic analytics

The next step after descriptive analytics is Diagnostic Analytics. In this type of analytics, past data is analyzed to drill down why these changes occurred or we try to find out the reason behind the trend that is seen.

2.1.3 Predictive analytics

This type of analytics helps in predicting the future. Predictive analytics makes use of dataset and certain features for prediction. The predicting feature will have a relation to other features in the dataset. The dataset used for analytics contains dependent and independent features. Independent features are used as input to the model and dependent feature predicts the output given by the model. This model is trained by providing historical/past data and prediction is done for given input. For this prediction, we use various Machine Learning or Deep Learning algorithms depending on our problem.

2.1.4 Prescriptive analytics

The purpose of prescriptive analytics is to prescribe or suggest actions to eliminate a future problem and ensure that we have a promising trend. This type of analytics takes a combination of what has happened, why it has happened and what will happen into consideration so that the industry is able to take best course of action. Prescriptive analytics models continuously analyze data and relationships to recommend an optimal solution.

2.2 Elements of manufacturing process

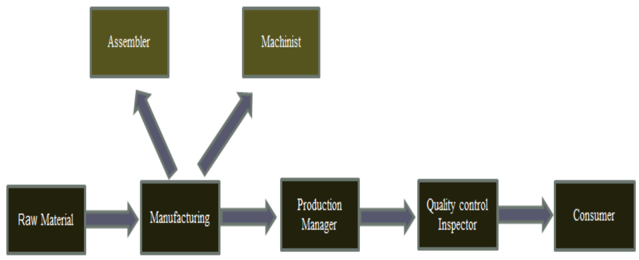

Fig -1: Processes in manufacturing industry.

The block diagram in Fig 1 presents how different processes are carried out in a manufacturing company.

Manufacturing process has elements – Assembler, Machinist, Production Manager and Quality Control Inspector. For a product to get manufactured, we require raw materials. These raw materials entering the company are transferred first to Assembler, where he collects the raw materials for a particular product and then passes it to the Machinist. He then performs certain operations on the raw materials collected and builds the product. The next task is to monitor the raw materials delivered to the industry, the number of products manufactured and so on. This task is done by production manager. Finally the product is manufactured and the quality control inspector examines the product. He checks whether there are any kind of defects, hazards to the product or the product is not manufactured as desired and appropriate action is taken by quality control inspector.

These elements are described in detail with the help of dataset for each element and different data analytics techniques are applied. Analysis of data is done for a span of 70 years (from 1950 to 2020). This data is present in the form of dataset. In expenses dataset, analysis of overall costs of the company on – raw materials, machinery, overhead, labor and profit earned is done for a particular year. In Employee dataset, analysis of details of employee like – age, salary, years of experience, group and expertise level is done so that we can cluster them into different domains. In domain dataset, analysis of data is done for different domains like – Assembler, Machinist/Operator, Production Manager and Quality Control Inspector. For domain dataset, features are same for all domains only values change. Finally for each case study the results are summarized.

2.3 Data analytics case studies

2.3.1 Case Study-1: Expenses data

Dataset:

Year	Employees	Working Days	Total hrs	Raw material(	Machinery cost(	Overhead cost	Labor cost	Total cost	Products_saled	Profit(in millions)
1950	2500	252	5040000	20	3	0.5	1.25	24.75	28	3.25
1951	2590	252	5221440	25	9	1	1.45	36.45	40	3.55
1952	2630	252	5302080	30	12	1.2	1.6	44.8	49	4.2
1953	2690	252	5423040	37	8	1.4	1.95	48.35	54	5.65
1954	2730	252	5503680	41	15	1.6	2	59.6	65	5.4
1955	2800	252	5644800	48	10	1.8	2.15	61.95	65	3.05
1956	2900	252	5846400	54	13	2.1	2.3	71.4	75	3.6
1957	3000	252	6048000	60	17	2.2	2.45	81.65	88	6.35
1958	3050	250	6100000	65	21	1.5	2.6	90.1	95	4.9
1959	4000	250	8000000	71	26	3	2.75	102.75	108	5.25
1960	4090	250	8180000	75	30	2	2.9	109.9	118	8.1
1961	4160	250	8320000	80	34	2	3	119	128.3	9.3
1962	4240	250	8480000	81	49	3	3.1	136.1	145	8.9
1963	4300	250	8600000	83	55	2.5	3.2	143.7	150	6.3

Fig -2: Dataset for Expenses.

The dataset in Fig 2 contains features like – Year, Number of Employees, Working Days, Total hours, Raw material(in million), Machinery cost(in million), Overhead cost(in million), Labor cost(in million), Total cost(in million), Products_saled(in million), Profit(in million). Here we will analyze the current condition of the company and perform analytics techniques.

Features:

Year - Specifies year.

Number of Employees - Number of employees present in the company during that year.

Working Days - Number of day's employees worked during that year in the company (calculated by excluding national holidays and some other holidays).

Total hours – Hours worked by 'x' number of employees. (Here 'x' is the number of employees during that year)

Raw material – Cost of total raw material that is bought into the company for manufacturing.

Machinery cost – Cost of machinery equipment or machines bought for manufacturing purpose.

Overhead cost – Cost of overhead expenses like electricity bills, maintenance, transportation, insurance, etc.

Labor cost – Cost of salary that is given to employees of company and the amount given to other labors who repairs the machines, or indirectly is part of the industry.

Total cost – This cost is calculated by adding raw material, machinery, overhead and labor cost together.

Products_saled – This is the amount at which the company has sold the product.

Profit – This is the amount the company has gained after selling the product.

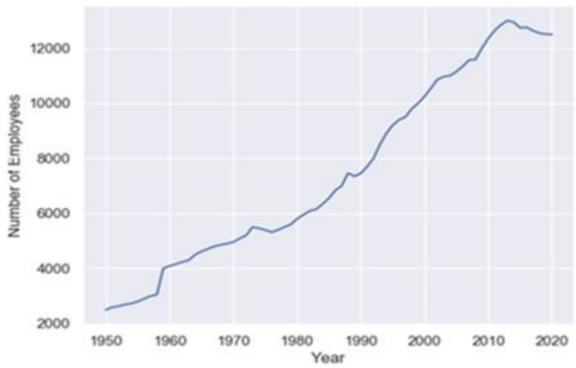

Chart -1: Year vs Number of employees

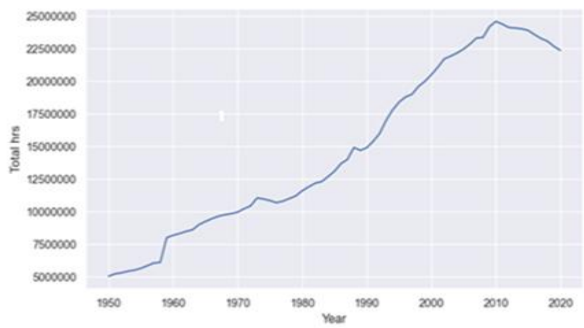

Chart -2: Year vs Total hrs

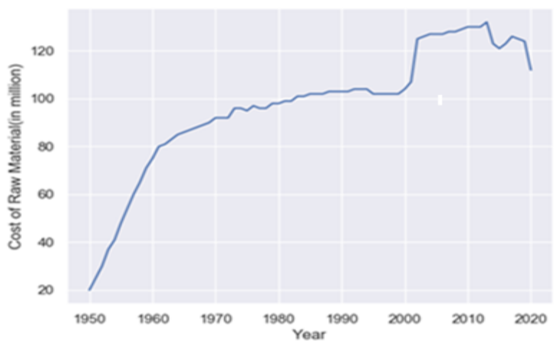

Chart -3: Year vs Cost of Raw Material.

Chart -4: Year vs Machinery cost

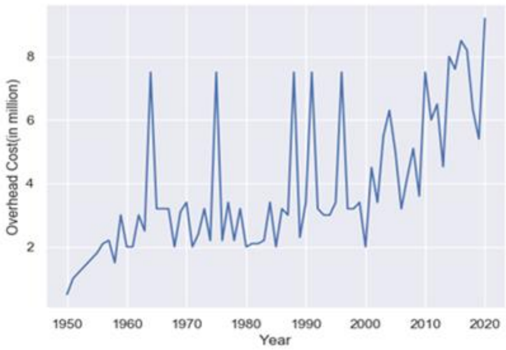

Chart -5: Year vs Overhead cost

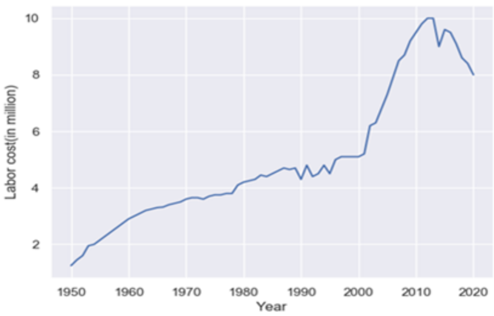

Chart -6: Year vs Labor cost

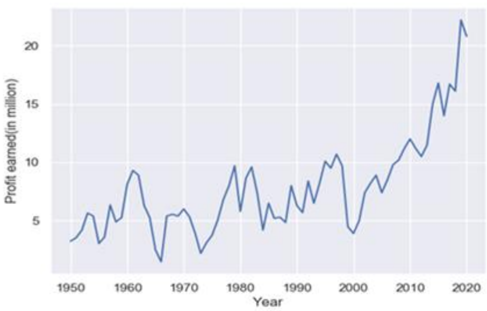

Chart -7: Year vs Profit

Descriptive analytics - The first graph (Year vs Number of Employees) in Chart-1 shows the count of employees for a given year. Here time period considered is from 1950 to 2020. This study assumes that the company started in the year 1950. The graph (Year vs Total hrs) in Chart-2 represents total hours spent by employees working in the company for a particular year. Working hours for each employee considered is 8 hours in a day. Therefore, total working hours is calculated by -

Total working hours = Number of Employees * 8 * Working Days

Chart-3 shows a trend line for the raw material that is bought by the company. Manufacturing of a product requires machines or machinery equipment in the company. Chart- 4 (Year vs Cost of Machinery) presents how much and when the company has invested on the machinery.

Chart-5 presents miscellaneous expenses that the company has spent every year. It also includes the cost spent on external or internal repairs and transportation. Chart-6 shows labor cost spent on the employees working in the company. This also includes the cost spent on labors who work for the company for internal/external repairs and delivering the products to the destination. Chart-7 shows the profit earned by the company every year by manufacturing the products.

Diagnostic analytics - In Chart-1, number of employees in the beginning is less because the company has just started. The rise of new technologies in every industrial revolution has led into hiring of employees. The company needs employees who are expert in handling the new technology and thus the number of employees will increase as seen in graph. Another reason for rise in number of employees is due to increase in production. Gradual increase-decrease is due to the change in needs of the company. In Chart-2, the feature "Total hours" will strongly depend on the feature Number of Employees. As the number of employees go on increasing hours will increase and vice-versa. This graph will show similarities to Chart-1.

At start, the cost spent to buy raw materials will be less as seen in Chart-3 (Year vs Cost of raw material) and the quantity of production will also be less for few years after 1950. When the company starts using new technologies which improves the quality of the products, the production will increase thereby increasing the amount of raw materials bought. In the graph at some points slight decrease is observed as well because, when the company buys raw materials for a year sometimes all those materials are not used for production. Some materials remain as it is and hence can be used for the next year for production. Initially, the company will buy less machinery (as we can see in Chart- 4) and then they will start investing money into machinery. In 1969, started 3rd Industrial Revolution which included new technologies making manufacturing process easy. The products manufactured with the help of these evolving technologies are of good quality than before which increases the production. The trend line decreases in this graph at some points because when a company once buys machines or machinery equipment, they use the same machines for longer period of time.

Overhead costs are the miscellaneous expenses that the company has to spend. When the company starts using new technologies, the maintenance cost, repairing cost, electricity cost and transportation cost goes on increasing as production increases which can be seen in Chart- 5. Labor cost depends on the number of employees working in the company. As the number goes on increasing labor cost will also increase (as seen Chart-6). Moreover, as the company starts using new technologies, the count of labors who work for the company for maintenance and repairing will go on increasing which again increases the labor cost.

In Chart-7, initially profit will be less because the company has just started and the quality of products will be satisfactory but not so good. As and when the company introduces new technologies, the quality of product will also increase thereby making more profit.

Prescriptive analytics - To maximize profit, to minimize the number of employees and working hours which in turn minimizes the labor cost, manufacturers need to practice Industry 4.0. In the era of Industry 4.0, currently companies use technologies like IOT, Digital twin, Data science, Big Data, Augmented reality, Virtual reality, Cyber-physical systems, 3D printing. More advanced technologies that a manufacturing industry can make use of are –

1) Ultrafast 3D printing: - In Industry 4.0 we have technology called 3D printing but 3D printing has limited applications. It is time consuming and expensive when a product is to be built layer by layer, for example, 3D printer used in Plastic manufacturing industry. However, the new technology called ultrafast 3D printing can be used which makes use of an ink jet head to deliver material, which it then fuses with an infrared lamp. Process is up to 100 times faster than the current technique [4].

2) Light based manufacture: - This technology can be used to manufacture electronic components in our smartphones and computers. This light based method relies on optical traps (devices using light to manipulate small objects in liquid). Currently, costly robots place and solder minute parts of circuitry. As electronic components get smaller and smaller, it becomes very difficult and time consuming process. Therefore, light based manufacture can provide a cheap and simple alternative [4].

3) Simulation: - Simulation gives us an idea beforehand regarding how parts would behave. This helps manufacturers reduce errors and cut costs [4].

4) Smart manufacturing: - In this technology, companies can use sensors to automate their machines or the manufacturing process [4]. Some of the sensors are given below –

a) Moisture sensor: - The quality of raw material that arrives at the company can be checked with many parameters, one of them is moisture. Moisture sensor can be used to measure the volumetric water content in industrial raw materials.

b) Vibration sensor: - This sensor can be used to monitor the jerks or vibrations rising from the internal and external defects. This leads to knowing of defects before occurrence. Vibration sensors give the machine condition, abnormal vibrations indicating problems in the machine. This can help detect early and repair the machine before it fails.

c) Motion sensor: - Motion sensor is used to detect movements of objects. Example – Detecting stalling of conveyors or seizing of bearings [5].

d) Temperature sensor: - Temperature sensor is used to monitor, control and display the temperature in its environment. Temperature sensors are used in industries to monitor and control temperature of machines which have higher workloads. Example - This sensor can be used to measure overheating of machinery which has workload in metal works, plastic and rubber industry [5].

e) Pressure sensor: - Pressure sensors can be used in manufacturing environments where high pressures are required to create a product. The generation of high pressures can be dangerous, so the levels must be measured to ensure that the manufacturing process remains safe. Pressure sensor measures the pressure applied and converts it into electrical signal, which can be read by the user. Example - Controlling the pressure of various gas or liquid in flow pipes and tanks [5].

f) Ultrasonic sensor: - Detects internal holes, cracks, presence of objects and measures the distance.

g) Barcode: - Barcode is used for product identification and tracking. It works fast and removes data entry errors.

h) Force sensor: - It used to measure various parameters like weight, torque and load. It also helps in recording torque of rotating machines like engine, gearbox, rotor, etc. Automated processes can take this data to provide the most optimal loads for each piece of machinery [5].

i) Vision sensors: - Used for quality control, monitoring purposes and completeness control. These are mostly used in industrial robots for handling systems, semiconductor manufacturing machines and assembly lines. Example - Inspection of supply chain and machinery for quality assurance [5].

j) Particle Sensors/Detectors: - These electronic devices sense dust and other airborne particulates and supply signals to inputs of control or display devices. Example - If rust occurs on machines or dust gets collected into some internal part [5].

This case study first applied Descriptive analytics which gave us a trend regarding the current condition of expenses in the company every year. Further applied Diagnostic analytics to understand the trend seen for which descriptive served as a base. Again, applied Prescriptive analytics to prescribe certain technologies that can be implemented inorder to improve and bring out a promising trend.

2.3.1 Case Study-2: Employee data

Dataset:

Employee	Age	Group	Year	Yrs of Experience	yrs_rate	skills_rate	Sale	manufacture_rate	Expertise	Salary
1	24	Trainee	2018	1.1	2	12	12	14	40	35000
2	25	Trainee	2016	1.3	2	12	12	14	40	35000
3	27	Trainee	2018	1.5	2	12	12	14	40	35000
4	26	Trainee	2017	2	2	12	12	14	40	35000
5	26	Trainee	2019	2.2	2	12	12	14	40	35000
6	27	Trainee	2015	2.9	8	19	18	15	60	45000
7	28	Trainee	2013	3	8	19	18	15	60	45000
8	27	Trainee	2017	3.2	8	19	18	15	60	4500
9	28	Trainee	2013	3.2	8	19	18	15	60	45000
10	29	Trainee	2015	3.7	8	19	18	15	60	45000
11	27	Trainee	2014	3.9	8	19	18	15	60	45000
12	28	Trainee	2015	4	8	19	18	15	60	45000
13	27	Trainee	2015	4	8	19	18	15	60	45000
14	28	Trainee	2015	4.1	8	19	18	15	60	45000

Fig -3: Dataset for Employees

Dataset in Fig-3 contains features like Employee id, Age, Group, Year, Years of Experience, yrs_rate, skills_rate, sale, manufacture_rate, Expertise, Salary. Here we will analyze the data of the employees that are working in the company.

Features:

Employee id – This is the id that is given to every employee of the company.

Age – This feature describes the age of each employee.

Group – Describes the different positions of employee in the company like Trainee, Assistant, and Experienced.

Year – Describes the year in which the employee entered the company.

Years of Experience – Describes the number of years the employee has experience in manufacturing domain.

Yrs_rate, skills_rate, manufacture_rate, sale – These features are used to calculate expertise level of each employee.

Expertise – Expertise level of each employee out of 100 working in manufacturing company.

Salary – Salary of each employee in the company.

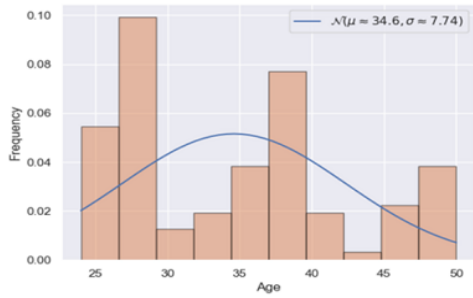

Chart -8: Normal distribution for Age

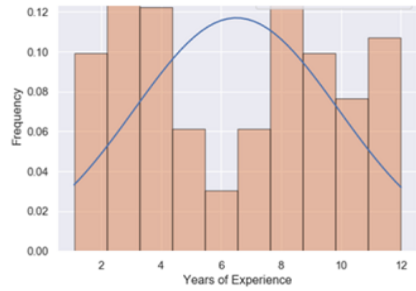

Chart -9: Normal distribution for Years of Experience

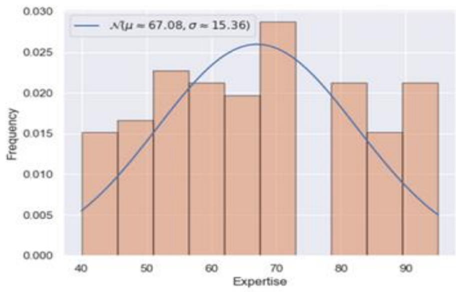

Chart -10: Normal distribution for Expertise.

Chart -11: Bar plot for Group

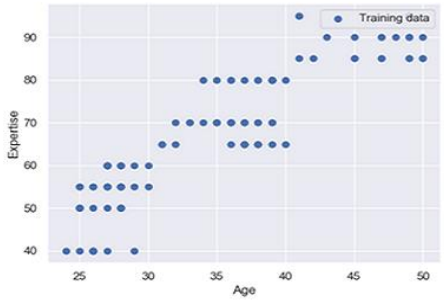

Chart -12: Training data for algorithm

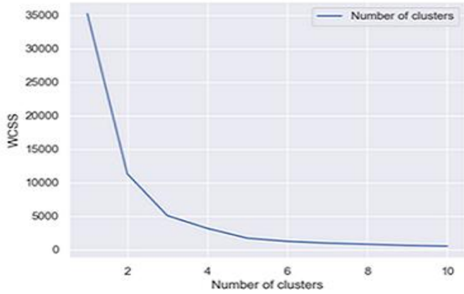

Chart -13: WCSS graph

Chart -14: Clustered graph

Descriptive analytics – Chart-8 shows the normal distribution of feature "Age" of employees working in the company. The mean value and the standard deviation are also mentioned in the graph. Mean value is 34 which means in this company most of the employees have age near to 34. The graph is asymmetric which means the number of employees having age less than 34 is not equal to number of employees having age greater than 34. The graph shows spread indicating the company has employees having variety of age groups.

Similarly in Chart-9 we have distribution of Years of experience. Here the mean value is 6 (out of 15) indicating company has most of the employees having working experience of 6 years. The graph shows symmetric behavior indicating the number of employees having experience less than 6 is equal to the number of employees having experience greater than 6.

Chart-10 shows the normal distribution of the feature "Expertise". Mean value is 67, which shows most of the employees have expertise level of 67 out of 100. The graph shows symmetric nature which means the number of employees on either side are equal. In Chart-11 we have a bar chart of the feature "Group". In this graph, the number of employees having position "Trainee" is more and then we have "Assistant" and finally "Experienced". The company recruits fresher's more, means the employees having age less than 30. Experienced employees are less in the company as compared to the other two categories.

Diagnostic analytics - Expertise level for most of employees is 67 out of 100 as seen in Chart-10 which means most of the employees in the company have good experience in manufacturing domain. The current generation is familiar and has worked on new technologies since we have Industry 4.0 from 2011. This also counts to the 67 rating of expertise level in the company.

In Chart-11, we have a bar plot indicating maximum count of Trainee (age less than 30). This again brings the picture of being familiar with new technologies. Having proper count of all 3 groups, the company is able to manage the processes very well because the seniors and managers will be guiding the trainee employees in building the products.

Predictive analytics - Due to the effect of Covid-19, data analytics has become a necessity. This technique analyzes data and splits employees into different domains under manufacturing. Then the company can decide which employees to allow to work from home and which to allow work physically. This technique uses clustering algorithm under machine learning. Under clustering we are using K – Means algorithm.

This technique first analyzes data using two features – Age and Expertise as shown in Chart-12. These features will cluster the data with the help of algorithm. Once we get the clusters we can conclude which cluster of employees to allow work from home. When using K- Mean's algorithm, we need to decide the number of "k". This number we can get using elbow method which plots a graph showing elbow point. The value that the elbow point represents on x-axis is the number of k.

Chart-13 shows the graph for elbow method, where see an elbow point at x = 3. Therefore the number of clusters that we will be using for our algorithm is 3. In Chart-14, we can see the results after applying K-Means algorithm. We have 3 clusters represented in three different colors. Color red represents the Assemblers, Machinist/Operators. Color green represents Production Manager and color blue represents Quality control inspector. Out of these 3 groups, Production Manager and Quality control inspector can do their tasks from home as well because their work is all about monitoring the production and monitoring the manufacturing process of product. But

Assemblers, Machinist/Operator need to be physically present. The company can have certain virtual techniques which can help Production Managers and Quality control inspectors to carry out their tasks so that the entire supply chain can be established.

This case study first applied Descriptive analytics technique which gave us the information of the overall employees. Finally, categorized the employees into 4 and accordingly allowed certain employees to be physically present and resume the functioning of the company.

2.3.1 Case Study-3: Domain data

Domain data includes the tasks assigned to the employees working in a particular domain. From our clustering results, we have 4 groups – Assembler, Machinist/Operator, Production Manager and Quality control inspector. Domain data includes 4 datasets for four different domains.

Dataset:

Task	Year	Employees	Automation(in %)	Time spent(Out of 100)
Assembler	1950	2500	3	94
Assembler	1951	2590	3	94
Assembler	1952	2630	4	94
Assembler	1953	2690	4	93
Assembler	1954	2730	4	93
Assembler	1955	2800	5	93
Assembler	1956	2900	5	93
Assembler	1957	3000	5	93
Assembler	1958	3050	5	93

Task	Year	Employees	Automation(in %)	Time
Machinist&Operator	1970	300	20	
Machinist&Operator	1971	400	20	
Machinist&Operator	1972	580	20	
Machinist&Operator	1973	700	20	
Machinist&Operator	1974	800	20	
Machinist&Operator	1975	900	20	
Machinist&Operator	1976	1000	23	27
Machinist&Operator	1977	1060	23	27
Machinist&Operator	1978	1100	23	27

Task	Year	Employees	Automation(in %)	Time Spent(out of 15)
ProductionManager	1990	500	10	14
ProductionManager	1991	800	10	14
ProductionManager	1992	1160	10	14
ProductionManager	1993	1400	10	14
ProductionManager	1994	1800	10	14
ProductionManager	1995	2180	13	14
ProductionManager	1996	2380	13	14
ProductionManager	1997	2590	13	13
ProductionManager	1998	2900	13	13

Task	Year	Employees	Automation(in %)	Time spent(out of 15)
Qualitycontrol	2000	400	10	
Qualitycontrol	2001	700	10	
Qualitycontrol	2002	1160	10	
Qualitycontrol	2003	1380	10	
Qualitycontrol	2004	1590	10	
Qualitycontrol	2005	1840	10	
Qualitycontrol	2006	1990	12	
Qualitycontrol	2007	2160	12	
Qualitycontrol	2008	2450	12	

Fig -4: Dataset for Domain

Datasets as shown in Fig-4 for each domain includes features like – Year, Employees, Automation, Time spent.

Features:

Year – Describes the details for that year.

Employees – The number of employees present in assembler domain. The sum of number for each domain equals the number of employees in the company mentioned in Expenses dataset.

Automation – Describes the percentage of automation the company has.

Time spent – The amount of time spent by those employees working for that particular year.

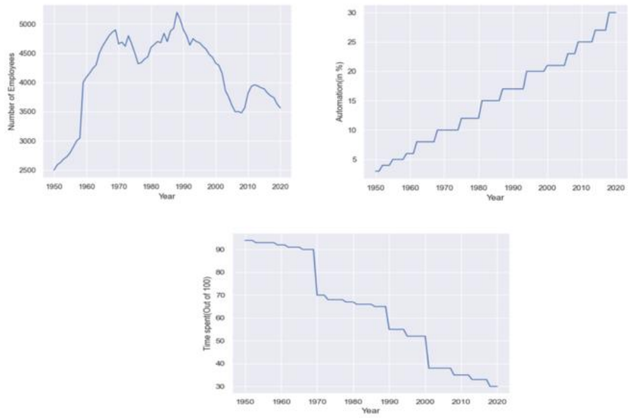

Chart -15: Graphs for Assembler domain

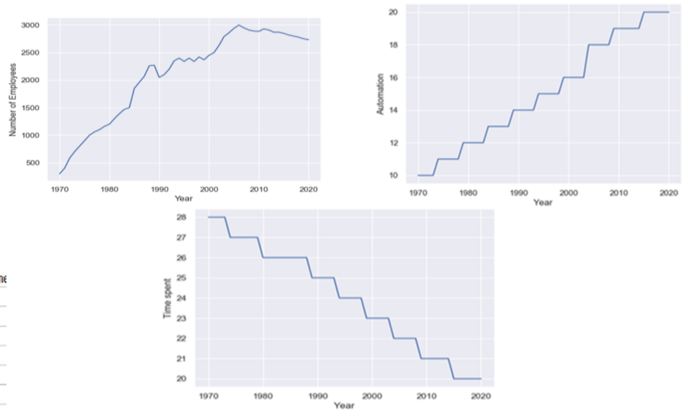

Chart -16: Graphs for Machinist/Operator domain

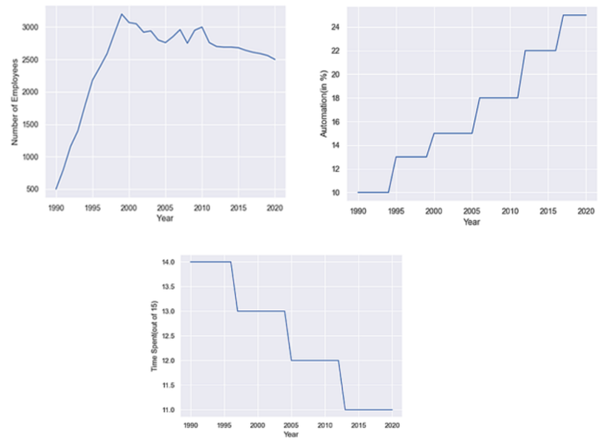

Chart -17: Graphs for Production Manager

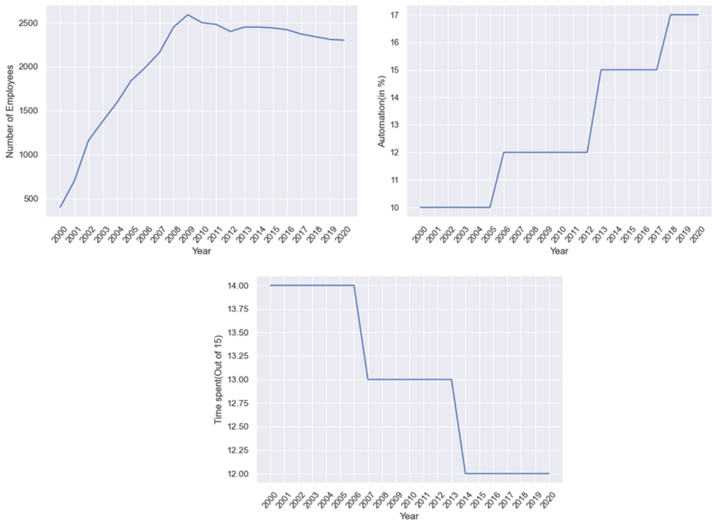

Chart -18: Graphs for Quality control inspector domain.

Descriptive analytics – The first plot (Year vs Number of employees) in Charts - 15, 16, 17 and 18 shows a trend for the number of employees present in each domain for a particular year. Plot 2 (Year vs Automation) for Charts - 15, 16, 17 and 18 presents the percentage of Automation during given year. Plot 3 (Year vs Time spent) for Charts 15, 16, 17 and 18 describes the amount of time spent by the employees in the company.

Diagnostic analytics - In Plot 1 (Year vs Number of employees) for Charts - 15, 16, 17 and 18, initially the number of employees is less as the company has just started. From 1950 onwards the company recruiting the employees would perform all the jobs – Assembling, Machinist, Production Manager and Quality control. Initially, the company has only assemblers performing all other tasks which increased their work load. Every industrial revolution surfaced new technologies which divided employees into different domains like – Assembler, Machinist, Production Manager and Quality control inspector. This means the company recruited employees for other domains or promoted the employees from one domain to other domain. That is why fall in the number of employees is seen. There is rise in the graph at some points because the company needs more employees in every domain as production goes on increasing.

In Plot 2 for Charts - 15, 16, 17, 18, feature “Automation” indicates the percentage of automation; company had each year for each domain. Different industrial revolutions gave rise to new technologies which lead to automating the manufacturing process. Automation and Time spent are dependent on each other because as automation in manufacturing process increases time spent by employees working on a product decreases as seen in plot 2 Charts - 15, 16, 17 and 18 when compared with plot 3 for Charts - 15, 16, 17 and 18.

Predictive analytics - Due to current covid-19 situation if in case, the company wants to reduce the number of employees working physically as well or wants to discontinue their job in some other situation then the company must have an idea beforehand regarding how it will affect the working of the company. Means the company needs to know the time required for the employees to spend in manufacturing.

Solution for this problem is to predict the time spent if number of employees and automation is given. This can be done by using random forest algorithm. This technique first trains the model by using training data and then predicts the result for given input. For training this technique uses 70% of data from dataset and for testing 30% of data. Thus the company can determine the amount of time ‘x’ number of employees need to spend in the company for manufacturing by providing the values for ‘x’ and percentage automation. Predictive analytics is applied for the other domains as well using same features and same algorithms only accuracy of the model is different in each domain.

Prescriptive analytics - Traditional methods in assembling included human intervention to gather different raw materials for a product to manufacture. As we are in Industry 4.0 phase, a lot of advancements are there which helps the manufacturing process. Assemblers can make use of technologies like – Simulation, AR/VR, Pick and place robots. This will help assemblers to identify the materials they need to collect for a particular product to get manufactured as they can have an idea beforehand by analyzing it on digital platform. Pick and place robots can help pick the materials and place it on the assembly line for manufacturing. This robot can have multiple hands as they can pick all the materials required in one go.

Machinist can use sensors to detect the dust that gets accumulated on the machines. Also we can have predictive maintenance techniques, by identifying possibility of failure of machines beforehand so that the company is able to take necessary actions. Companies can also make use of RADAR. Other sensors company can make use of are mentioned in Prescriptive analytics part for Expenses Data.

In case of Production Managers, current methods involve maintaining excel sheets to manage everything or he himself has to enter these things into the software so that the data can be stored into database. We can have technologies for this domain that sense the count of products manufactured every day and send data to database directly also when the products get delivered a signal can be sent to the manager and he can view it on android application or on website.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

Quality control inspectors can use image processing or a system which dumps data of the product which is to be manufactured again due to some defects present in it and the manager can easily recognize the product and manufacture it again. These technologies can be used to minimize the number of employees, to build good quality products in less time.

This case study first applied Descriptive analytics to find out the time spent by each employee and the percentage of automation the company has for each domain. This gave us a trend which we used further and applied diagnostic analytics to find out the reason behind ups and downs seen in the graphs. Further, predicted the amount of time spent each year by the employees in each domain when number of employees and percentage of automation is provided as input. This gives us a rough idea on the feature – time spent if the company increases/decreases the number of employees or automation. Moreover, the company can make use of certain technologies listed above for digitization of processes under each domain resulting into good quality of products.

3. RESULTS & DISCUSSIONS

This research paper focuses on techniques of Data analytics implemented on critical features in manufacturing company. The current crisis has devastated the supply chain in manufacturing industries which leads to analyzing how the supply chain can be set to working condition amid pandemic. This problem to get solved, this research paper considers 3 datasets- Expenses, Employee and Domain. By analyzing these datasets we can perform data analytics which will help the company grow. By analyzing Expenses data, we can have a picture of current condition of company. By analyzing Employee data, we can analyze the data of employees working in the company.

Also performing clustering algorithm on this data helped us find out which employees to allow to work from home. After analyzing different domain data, we can predict time the given number of employees needs to spend for manufacturing process. This analysis for domain is done for – Assembler, Machinist/Operator, Production Manager, Quality Control Inspector. For each domain data differs and so the feature time spent.

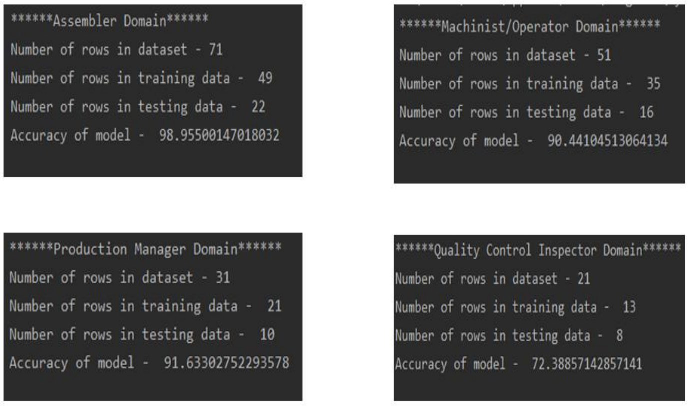

Fig – 5: Results of domain data

Fig-5 represents the results for random forest algorithm we used to predict time spent. We can see the number of rows we have in our dataset, the training and testing samples, the accuracy of model when we train our model and predict the testing results. We have also provided the technologies the company can use to carry out the tasks in the company effectively.

4. CONCLUSIONS

Data analytics is the technique with the help of which we can extract useful data from the dataset and perform various data analytics techniques. Under data analytics we have 4 techniques – Descriptive, Diagnostic, Predictive and Prescriptive. Descriptive analytics is used to describe the data with the help of line charts, bar charts, pie charts, histograms, normal distribution and many others. Diagnostic analytics tells us the reason behind the ups and downs in the graph. Predictive analytics uses machine learning and deep learning algorithms to predict the value of a feature for given input. Prescriptive analytics gives us suggestions regarding how management can be done to eliminate the problems caused previously. These techniques are applied to case studies of manufacturing industry. Here implementation of only those analytics techniques is done and only those features are taken into consideration which is critical. All the techniques of data analytics can be implemented but this research paper has demonstrated a flow which can be used to manage the different blocks of the company. Every analytics technique demonstrated acts as a base for the other. Therefore the hypothesis that all these 4 techniques should be carried out for a problem to be solved is false. These techniques are not processes that need to be implemented one after another. We need to analyze the problem, bring down the features responsible and then implement the techniques accordingly. The concept demonstrated in this study can be applied to data of any company.

This research paper also prescribes few technologies that can be used by the company to automate their processes. For each domain different technologies are suggested. By using these technologies the company is able to produce good quality, reliable products raising the production and increasing the profit.

5. REFERENCES

[1] Hori, M., Kawashima, E., & Yamazaki, T, (2010), "Application of cloud computing to agriculture and prospects in other fields", Fujitsu Sci. Tech. J, 46(4), 446–454.

[2] Biduaa K. R. & Dr. Patela, C. N, (2015), "Internet of Things and Cloud Computing for Agriculture in India", International Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 12.

[3] Paul, S. K. & Chowdhury, P.,(June 2020), "A production recovery plan in manufacturing supply chains for a high-demand item during COVID-19", International Journal of Physical Distribution & Logistics Management.

[4] Finch, S. (22nd November, 2017), "5 technologies now changing manufacturing", Retrieved February 21, 2021, from https://foundry4.com/5-technologies-now-changing-manufacturing

[5] Thomasnet.com (n.d.), "Sensors – A complete guide (Types, Applications and Suppliers)", Retrieved February 21, 2021, from https://www.thomasnet.com/articles/instruments-controls/sensors/#applications