BIG DATA IN MANUFACTURING

A. INTRODUCTION

Big data has applications in just about every industry – retail, healthcare, financial services, government. Any organization that can assimilate data to answer nagging questions about their operations can benefit from big data. All big data projects start with a viable use case. Here are four sample big data use cases for the manufacturing industry.

Analyzing big data use cases in the manufacturing industry can reduce processing flaws, improve production quality, increase efficiency, and save time and money. Tata Consultancy Services asked manufacturers to rate the following big data benefits on a scale of one to five:

- Product quality and defects tracking – 3.37
- Supply planning – 3.34
- Manufacturing process defect tracking – 3.32
- Supplier, components, and parts defect tracking – 3.11
- Supplier performance data to inform contract negotiations – 3.08
- Output forecasting – 3.03
- Increasing energy efficiency – 2.97
- Testing and simulation of new manufacturing processes – 2.88
- Support for mass-customization of manufacturing – 2.75

The range of big data use cases in the manufacturing industry is limited only by available data and imagination.

Why Big Data Use Cases in the Manufacturing Industry?

Before looking at some specific big data use cases in the manufacturing industry, let's address the role use cases play in big data analytics.

A big data use case provides a focus for analytics, providing parameters for the types of data that can be of value and determining how to model that data using Hadoop analytics. For example, answering a question such as “where is the next big market for my product” is harder to answer than “who is likely to buy more product in the United States.” Here are four sample big data use cases in the manufacturing industry:

1. Improving Manufacturing Processes

McKinsey and Company offers a big data use case in pharmaceutical manufacturing. A biopharmaceutical company was using live, genetically engineered cells and tracking 200 variables to track the purity of its manufacturing process for vaccines and blood components. However, two batches of the same substance manufactured using identical processes showed a yield variation from 50 to 100 percent. The inconsistency in capacity and quality could attract regulatory attention.

The project team segmented its manufacturing processes into clusters of activity. Using big data analytics the team assessed process interdependencies and identified nine parameters that had a direct impact on vaccine yield. By modifying target processes the company was able to increase vaccine production by 50 percent resulting in savings between $5 and $10 million annually.

2. Custom Product

Tata Consultancy Services cites the case of a $2 billion company that generates most of its revenue by manufacturing products to order.

Using big data analytics this company was able to analyze the behavior of repeat customers. The outcome is critical to understanding how to deliver goods in a timely and profitable manner.

3. Better Quality Assurance

Intel has been harnessing big data for its processor manufacturing for some time. The chipmaker has to test every chip that comes off its production line. That normally means running each chip through 19,000 tests.

Using big data for predictive analytics Intel was able to significantly reduce the number of tests required for quality assurance. Starting at the wafer level, Intel analyzed data from the manufacturing process to cut down test time and focus on specific tests.

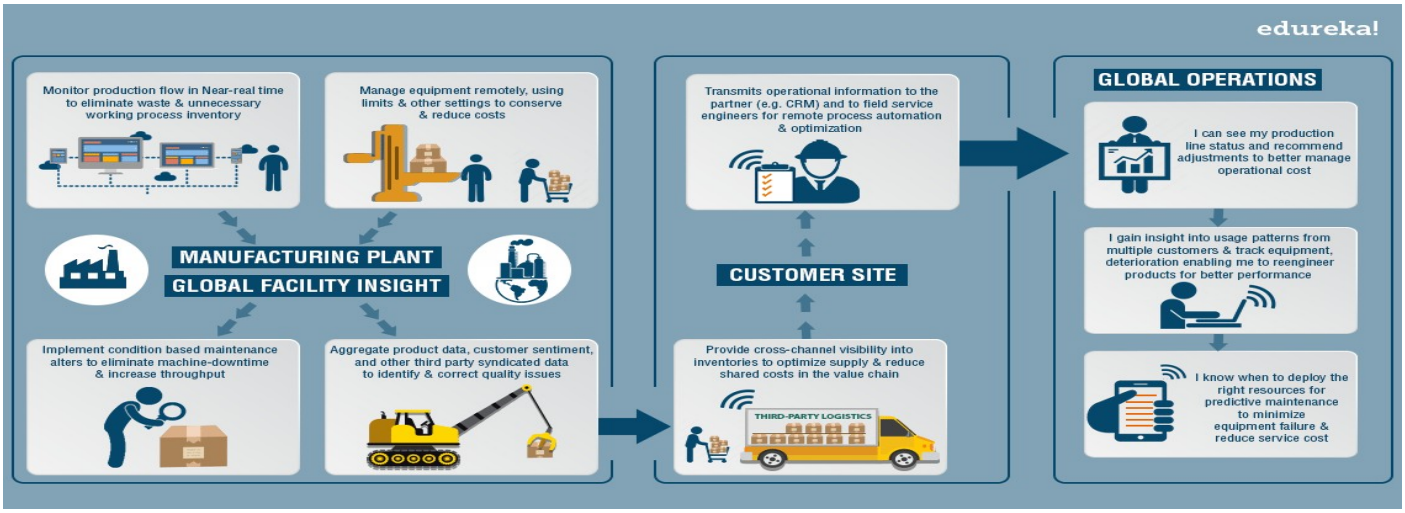
The result was a savings of $3 million in manufacturing costs for a single line of Intel Core processors. By expanding big data use in its chip manufacturing, the company expects to save an additional $30 million.

4. Managing Supply Chain Risk

One manufacturer is using big data to reduce risk in delivery of raw materials, no matter what happens in the supply chain.

Using big data analytics, the company has overlaid potential delays on a map, analyzing weather statistics for tornadoes, earthquakes, hurricanes, etc. Predictive analytics allow the company to calculate the probabilities of delays. The company uses the analytics findings to identify backup suppliers and develop contingency plans to make sure production isn’t interrupted by natural disaster.

B. METHODOLOGY TO IMPLEMENT



Long before any analysis can happen, you have to start aggregating data. In some cases, it's not a problem at all: you just ***deploy/add sensors*** on your manufacturing equipment, ***prepare data storing facilities*** and enjoy the flow of 'freshly-cut' data.

But in other cases, such as if your production cycle is months- or even years-long, it can prove difficult because you may lack the info on how your production process parameters influence output. And without knowing it, it's all really a shot in the dark. But don't get upset: there are ways to fight it. For example, try not to concentrate on the entire manufacturing cycle at once. Rather, focus on one part of your manufacturing process (say, inoculation in cheese production), gather data about it, analyze it and see how you can improve it.

C. TOOLS USED

- APACHE HADOOP
  Apache Hadoop is a software framework employed for clustered file system and handling of big data. It processes datasets of big data by means of the MapReduce programming model.

- CASSANDRA
  Apache Cassandra is free of cost and open-source distributed NoSQL DBMS constructed to manage huge volumes of data spread across numerous commodity servers, delivering high availability. It employs CQL (Cassandra Structure Language) to interact with the database.

- KNIME
  KNIME stands for Konstanz Information Miner which is an open source tool that is used for Enterprise reporting, integration, research, CRM, data mining, data analytics, text mining, and business intelligence. It supports Linux, OS X, and Windows operating systems.

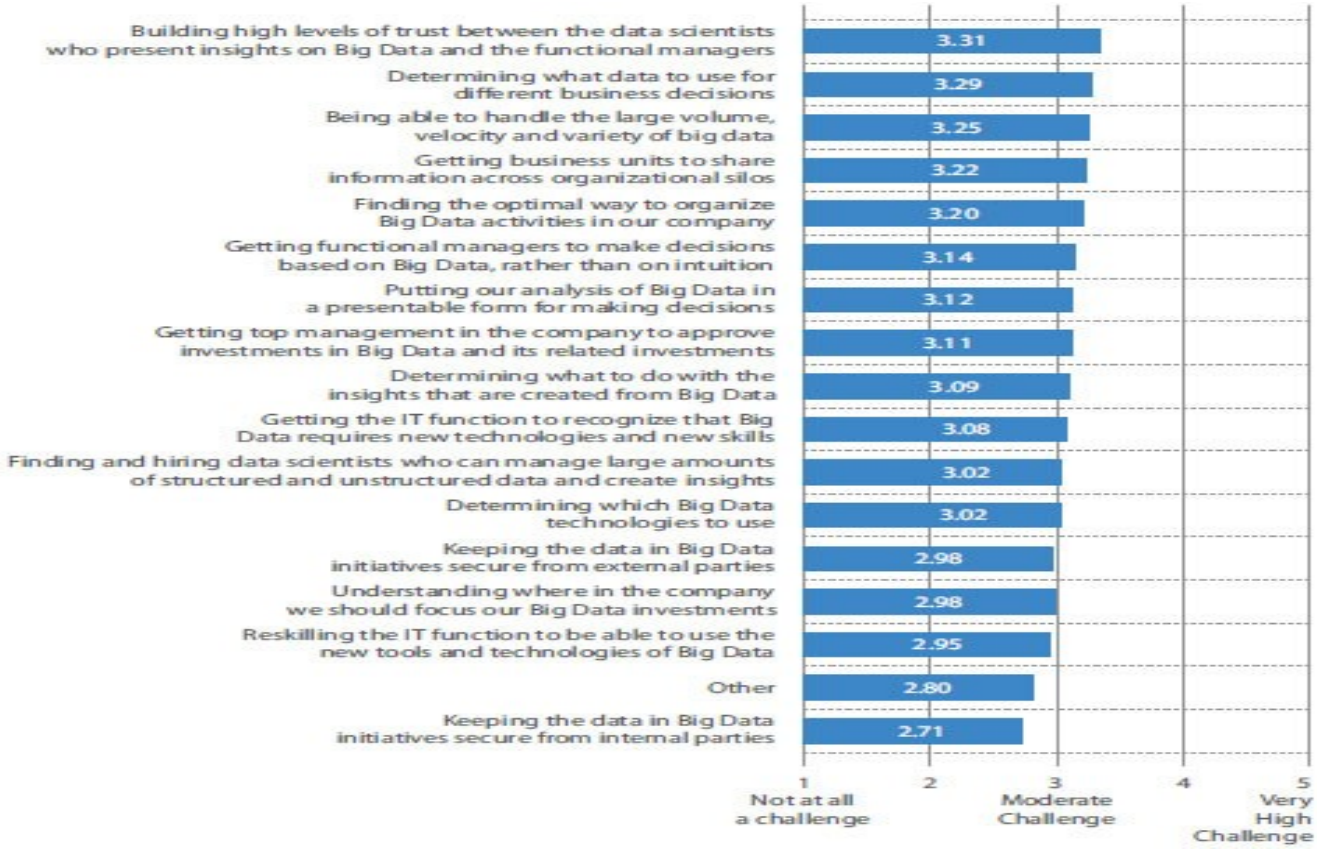
D. DATASETS USED

1. http://www.kdnuggets.com/datasets/index.html

2. http://archive.ics.uci.edu/ml/datasets.html?sort=nameUp&view=list

3. https://cloud.google.com/bigquery/public-data/

E. CHALLANGES



F. CONCLUSION

To remain competitive, manufacturers must embrace Big Data. Manufacturers that make the most of thecustomer, product, and equipment data they capture stand to improve their ability to innovate, please their customers, and bring more profitable products and services to market more quickly.