

Advanced machine learning algorithms for predictive analysis or anomaly detection in the big data.

Submitted by:

R.Vinisha

920121104057

B. E(CSE)

Bharath Niketen Engineering college

Introduction :

Big data primarily refers to data sets that are too large or complex to be dealt with by traditional data-processing application software. Data with many entries offer greater statistical power, while data with higher complexity may lead to a higher false discovery rate.

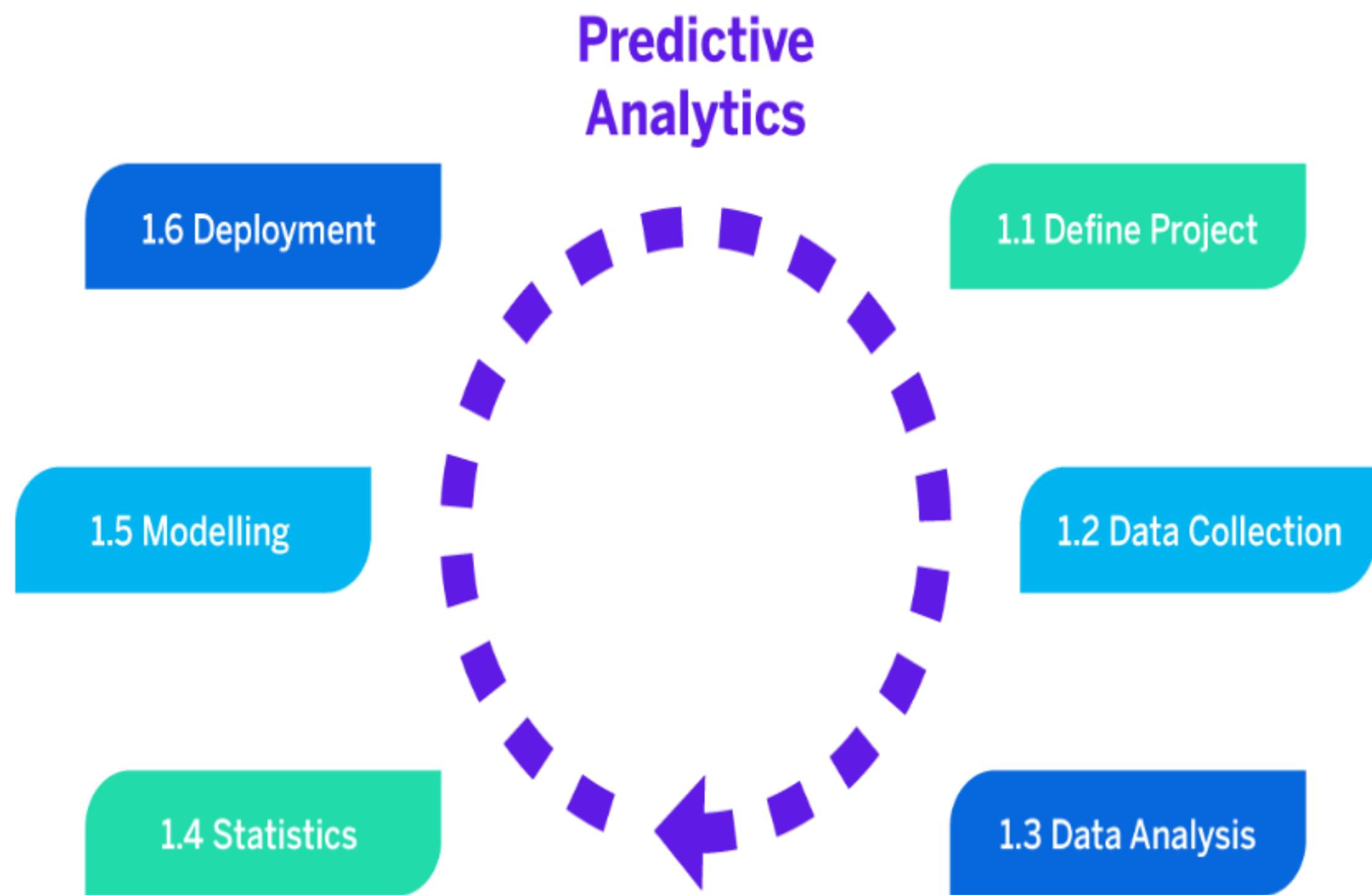


Predictive analysis :

Predictive analytics is the process of using data to forecast future outcomes. The process uses data analysis, machine learning, artificial intelligence, and statistical models to find patterns that might predict future behavior.

Predictive analytics models may be able to identify correlations between sensor readings. For example, if the temperature reading on a machine correlates to the length of time it runs on high power, those two combined readings may put the machine at risk of downtime.

Predictive Analytics Process



What is big data analysis?

Big data analytics is the often complex process of examining big data to uncover information -- such as hidden patterns, correlations, market trends and customer preferences -- that can help organizations make informed business decisions.



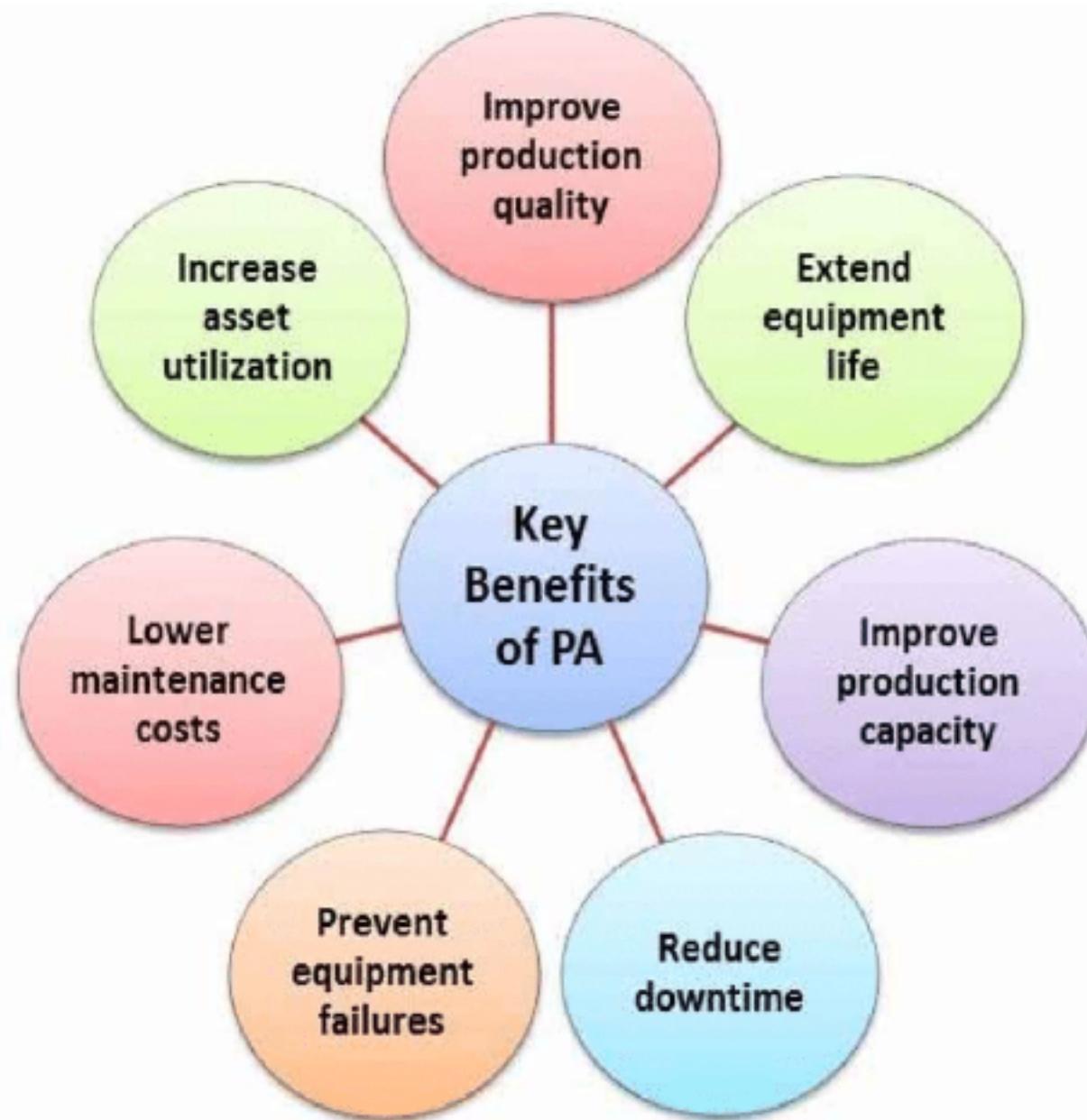
Big data analytics is a form of advanced analytics, which involve complex applications with elements such as predictive models, statistical algorithms and what-if analysis powered by analytics systems.

Impact of Big Data on Predictive Analytics:

Big Data stands for raw and voluminous sets of information that become valuable in studies and analyses. The more new technologies we have the more capacity of data these technologies accumulate. So, it is of great importance to make use of these large amounts of information by analyzing them. And, this type of analysis requires specific tools and automation as humans cannot process big data manually, because it's impossible. The process of automated big data analysis is called **big data analytics**.

Benefits of Predictive Analytics:

Predictive analytics helps in delivering enhanced customer experience.
Businesses can analyze customer data, preferences, and behavior to personalize their value proposition. This personalized approach can help companies to differentiate themselves from competitors and build stronger customer relationships.



Exploring the Benefits of Predictive Analytics:

In today's world, data has become an essential asset for businesses. However, collecting data alone is insufficient; it must be analyzed and turned into meaningful insights. This is where predictive analytics comes in.

Predictive analytics has been around for a long time, with roots dating back to the early 1800s. One of the earliest known examples of predictive analytics is the work of the English statistician Francis Galton, who used statistical techniques to predict the height of children based on the height of their parents. Since then,

predictive analytics has evolved significantly and is now a critical component of modern business intelligence.

Challenges of Predictive Analytics :

While predictive analytics can be a powerful tool for brands, it's essential to understand the challenges associated with using this technique.

According to a study by McKinsey & Company, many businesses struggle with these challenges when implementing predictive analytics. For example, the study found that only 19% of companies are very confident in the accuracy of their predictive models.

Development of Big Data Predictive Analytics Model for Disease Prediction using Machine learning Technique:

Now days, health prediction in modern life becomes very much essential. Big data analysis plays a crucial role to predict future status of health and

offers prominent health outcome to people. Heart disease is a prevalent disease cause's death around the world. A lot of research is going on predictive analytics using machine learning techniques to reveal better decision making. Big data analysis fosters great opportunities to predict future health status from health parameters and provide best outcomes. We used Big Data Predictive Analytics Model for Disease Prediction using Naive Bayes Technique (BPA-NB). It provides probabilistic classification based on Bayes' theorem with independence assumptions between the features. Naive Bayes approach suitable for huge data sets especially for big data. The Naive Bayes approach train the heart disease data taken from UCI machine learning repository. Then, it was making predictions on the test data to predict the classification. The results reveal that the proposed BPA-NB scheme provides better accuracy about 97.12% to predict the disease rate. The proposed BPA-NB scheme used Hadoop-spark as big data computing tool to obtain significant insight on healthcare data. The experiments are done to predict different patients' future health condition. It takes the training dataset to estimate the health parameters necessary for classification. The results show the early disease detection to figure out future health of patients.

What is predictive analytics? An enterprise guide

Predictive analytics is a form of advanced analytics that uses current and historical data to forecast activity, behavior and trends. It involves applying statistical analysis techniques, data queries and machine learning algorithms to data sets to create predictive models that place a numerical value -- or score -- on the likelihood of a particular action or event happening.

Predictive analytics is a key discipline in the field of data analytics, an umbrella term for the use of quantitative methods and expert knowledge to derive meaning from data and answer fundamental questions about a business, the weather, healthcare, scientific research and other areas of inquiry. In the context of businesses, the main focus here, that process is often referred to as business analytics.

Conclusion:

The availability of Big Data, low-cost commodity hardware, and new information management and analytic software have produced a unique moment in the history of data analysis. The convergence of these trends means that we have the capabilities required to analyze astonishing data sets quickly and cost-effectively for the first time in history. These capabilities are neither theoretical nor trivial. They represent a genuine leap forward and a clear opportunity to realize enormous gains in terms of efficiency, productivity, revenue, and profitability.