Steps in the Predictive Analytics process

The predictive analytics process involves defining a goal or objective, collecting and cleaning massive amounts of data, and then building predictive models using sophisticated predictive
algorithms and techniques. This traditionally complex process is becoming more automated and more accessible to the average business user thanks to new AI technologies, but companies may still need IT to help in certain steps or to build certain models.

The steps in the predictive analytics process.

1. **Define your project's objectives.** What is the desired outcome? What problem are you trying to solve? The first step is to define your project's objectives, deliverables, scope, and data required.
2. **Collect your data.** Gather all the data you need in one place. Include different types of current and historical data from a variety of sources – from transactional systems and sensors to call center logs – for more in- depth results.
3. **Clean and prepare your data.** Clean, prepare, and integrate your data to get it ready for analysis. Remove outliers and identifying missing information to improve the quality of your predictive data set.
4. **Build and test your model.** Build your predictive model, train it on your data set, and test it to ensure its
accuracy. It may take multiple iterations to generate an error-free
model.
5. **Deploy your model.** Deploy your predictive model and put it to work on new data. Get results and reports – and automate decision-making based on the output.
6. **Monitor and refine your model.** Regularly monitor your model to review its performance and ensure it's providing the expected results. Refine and optimize your model as needed.

What is Data Mining?

Data mining refers to a process of analyzing data from different contexts and summarizing it into useful information. The information gathered from data mining could include customer patterns, purchase patterns, transaction times, customer demand, and the relationship between the sold items. It is a powerful technology with great potential to assist companies in targeting the most significant information in the data set they have gathered about the customer behaviors and potential of the customers.

These are the given steps involved in the process of data mining

1. Business Understandings
2. Data Selection
3. Data Preparation
4. Modelling
5. Evaluation
6. Deployment

Application of data mining

1. Financial Analysis
2. Biological Data Analysis
3. Market Analysis
4. Retail Industry
5. Manufacturing Engineering
6. Criminal investigation

Difference between data mining and data exploration:

Data mining and data exploration are the two main techniques used in data science to extract data from various sources.

Business users and an increasing number of citizen data scientists—individuals who have no formal training in data science or analytics but whose jobs depend on being able to see patterns and trends in data—perform the wide process of data exploration. This diverse group benefits from the use of visualization tools to more effectively export and analyze a range of metrics and data sources.

Data specialists typically carry out the specific procedure known as "data mining." To sort through incredibly massive data sets, detect patterns, and predict future trends, data analysts build association rules and criteria.

Data exploration is typically carried out first to evaluate the relationships between variables. Data mining then starts. Data models are developed through this approach to glean more information from the data.

How does predictive analytics work?

Data scientists use predictive models to identify correlations between different elements in selected datasets. Once data collection is complete, a statistical model is formulated, trained, and modified to generate predictions.

The workflow for building predictive analytics frameworks follows five basic steps:

1. **Define the problem**: A prediction starts with a good thesis and set of requirements. For instance, can a predictive analytics model detect fraud? Determine optimal inventory levels for the holiday shopping season? Identify potential flood levels from severe weather? A distinct problem to solve will help determine what method of predictive analytics should be used.
2. **Acquire and organize data**: An organization may have decades of data to draw upon, or a continual flood of data from customer interactions. Before predictive analytics models can be developed, data flows must be identified, and then datasets can be organized in a repository such as data warehouse like bigquery .