**Module-1**

**Definition**

The predictive analytics refers to the use of statistics and modelling techniques to make predictions about future outcomes and performance. It looks at current and historical data patterns to determine if those patterns are likely to emerge again. This allows businesses and investors to adjust where they use their resources to take advantage of possible future events. It can also be used to improve operational efficiencies and reduce risk.

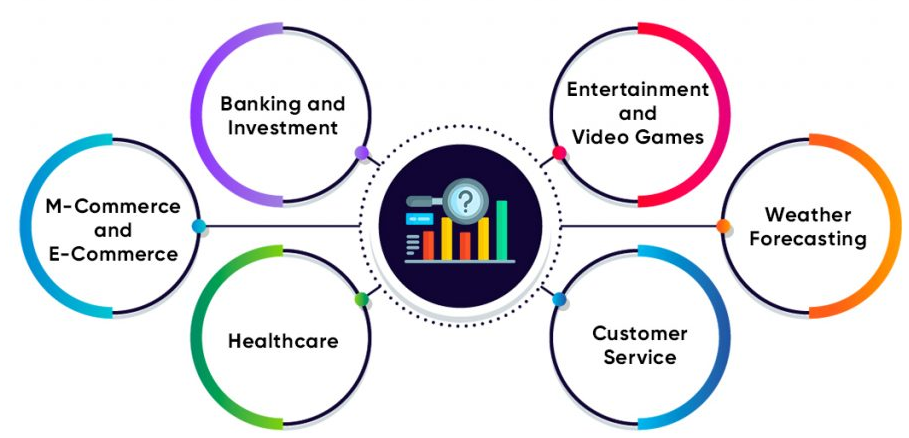
**Key points to remember**

* Predictive analytics uses statistics and modelling techniques to determine future performance.
* Industries and disciplines, such as insurance and marketing, use predictive techniques to make important decisions.
* Predictive models help make weather forecasts, develop video games, translate voice-to-text messages, customer service decisions, and develop investment portfolios

## **Understanding Predictive Analytics**

It draws on a series of techniques to make these determinations, including artificial intelligence (AI), data mining, machine learning, modelling, and statistics. Text analysis does the same, except for large blocks of text. Predictive models are used for all kinds of applications, including weather forecasts, creating video games, translating voice to text, customer service, and investment portfolio strategies. All of these applications use descriptive statistical models of existing data to make predictions about future data.

## **Uses of Predictive Analytics**



Predictive analytics is a decision-making tool in a variety of industries.

### **Forecasting**

Forecasting is essential in manufacturing because it ensures the optimal utilization of resources in a supply chain.

### **Credit**

Credit scoring makes extensive use of predictive analytics. When a consumer or business applies for credit, data on the applicant's credit history and the credit record of borrowers with similar characteristics are used to predict the risk that the applicant might fail to perform on any credit extended.

### **Underwriting**

Data and predictive analytics play an important role in underwriting. Insurance companies examine policy applicants to determine the likelihood of having to pay out for a future claim based on the current risk pool of similar policyholders, as well as past events that have resulted in payouts.

### **Marketing**

It includes how consumers have reacted to the overall economy when planning on a new campaign. They can use these shifts in demographics to determine if the current mix of products will entice consumers to make a purchase.

Active traders, meanwhile, look at a variety of metrics based on past events when deciding whether to buy or sell a security. Moving averages, bands, and [breakpoints](https://www.investopedia.com/terms/b/breakpoint.asp) are based on historical data and are used to forecast future price movements.

1. **Fraud Detection**

Financial services can use predictive analytics to examine transactions, trends, and patterns. If any of this activity appears irregular, an institution can investigate it for fraudulent activity.

### **Supply Chain**

Supply chain analytics is used to predict and manage inventory levels and pricing strategies. It use historical data and statistical models to forecast future supply chain performance, demand, and potential disruptions. This helps businesses proactively identify and address risks, optimize resources and processes, and improve decision-making.

### **Human Resources**

Human resources uses predictive analytics to improve various processes, such as forecasting future workforce needs and skills requirements or analysing employee data to identify factors that contribute to high turnover rates.

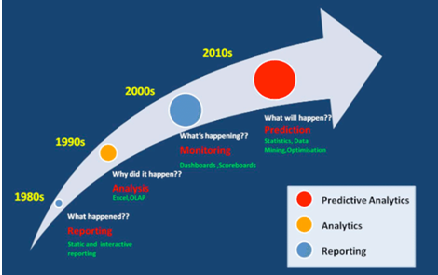
## **Benefits of Predictive Analytics**

Investors, financial professionals, and business leaders are able to use models to help reduce risk. For instance, an investor and their advisor can use certain models to help craft an investment portfolio with minimal risk to the investor by taking certain factors into consideration, such as age, capital, and goals. There is a significant impact to cost reduction when models are used. Businesses can determine the likelihood of success or failure of a product before it launches. Or they can set aside capital for production improvements by using predictive techniques before the manufacturing process begins.

## **What Is Predictive Analytics Good for?**

Predictive analytics is good for forecasting, risk management, customer behavior analytics, fraud detection, and operational optimization. Predictive analytics can help organizations improve decision-making, optimize processes, and increase efficiency and profitability. This branch of analytics is used to leverage data to forecast what may happen in the future.

**Evolution of Predictive Analytics**

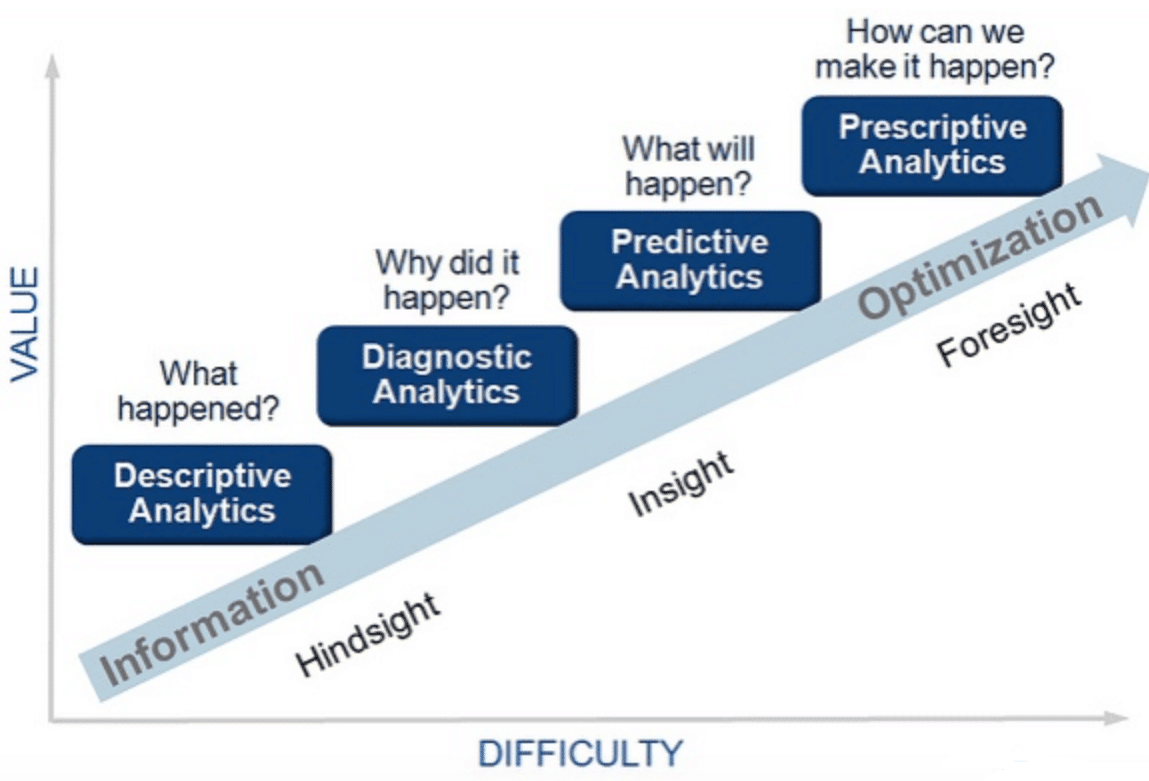


Analytics is the science of producing insight from the patterns in complex data to make better decisions. It relies on statistics, predictive modeling and computer science.

## **Different Types of Analytics**

Analytics can be split into 4 categories.

* **Descriptive Analytics (What happened?):** Showing what is actually happening based on given data; often usually via dashboards and reporting tools. The most basic analytics is descriptive analytics, which answers the question of what happened in the past with a given data set. Almost every organization uses this type of analytics, which involves arithmetic operations, mean, median, max, percentage, etc.
* **Diagnostic Analytics (Why did it happen?):** Analyzing past performance to determine not only what happened, but why it happened. It provides answers about why something happened in the past. Companies that use diagnostic analytics are more likely to understand the causal relationships between actions that can be considered as a deep insight into the root causes of the events. It involves techniques of data discovery, data mining, and correlation while using statistical expressions such as probability, likelihoods and distribution of outcomes for the investigation.
* **Predictive Analytics (What could happen?):** Describing what scenarios are likely to occur, often in a predictive forecast. Predictive analytics answers the question of what is going to happen in the future with data programming tools such as R, Phyton, IBM SPSS and RapidMiner. Predictive analytics provides and validates predictive models by using machine learning algorithms like random forests, learning, testing data. It expects the future will reflect a pattern similar to the precedent data. Organizations usually consult data scientists and machine learning experts to handle data for predictive analytics.
* **Prescriptive Analytics (What should we do?):** Making suggestions about what should be done and their basis. Analysts try to find out what action to take or how to take advantage of future opportunities with a given data set. Prescriptive analytics is not fail-proof, even though it uses advanced machine learning algorithms and business guidelines to analyze possible decisions and potential influences of those decisions.



## **Predictive Analytics Techniques**

Predictive analytics incorporates a combination of scientific methods and techniques as discussed below;

* **Data Mining:** In order to manage large amounts of data sets either structured or unstructured to recognize hidden patterns and relationships among variables provided, data mining is aimed to. Once identified, these relationships can be used to understand the behaviour of the event from which data is compiled.
* **Statistical Modelling:** In parallel to the data mining process, [statistical data models](https://www.analyticssteps.com/blogs/10-types-statistical-data-distribution-models) can be developed depending on the context of what needs to be anticipated using the same collected data as for data mining. Once the model builts, the new data is fed to models to predict future outcomes.
* **Machine learning:** ML can deploy iterative methods and techniques to identify patterns from large data sets and build models.

## **Why is Predictive Analytics Important?**

 Predictive analytics is leading to address complex business problems and discover potential opportunities, some common benefits include;

### **Fraud detection**

 With the growing concern of cybersecurity, conducting high performing behavioral analytics is demanding that examines all the suspicious behavior/activities across a network in a real-time to detect fraud actions, zero-day vulnerabilities and underlying threats.

### **Marketing campaigns optimization**

 Predictive analytics is beneficial in optimizing marketing campaigns and promotions’ events. To determine purchasing response of customers and publicizing cross-sell opportunities, predictive models help in businesses to attract, retain and increase valuable customers.

### **Minimization in risks**

Consider an example of reducing risks, Credit Scores, that is highly employed to recognize the likelihood of defaulters from the user's purchasing behaviour.

### **Improvements in business operations:**

Predictive analytics enables organizations to make smarter decisions by smoothing operations and functions more efficiently. For example, airlines are using predictive analytics to confirm a range of ticket pricing, hotels are using it to predict the number of guests on a particular day/hours to maximize space occupancy and increase revenue.

### **What is predictive modeling?**

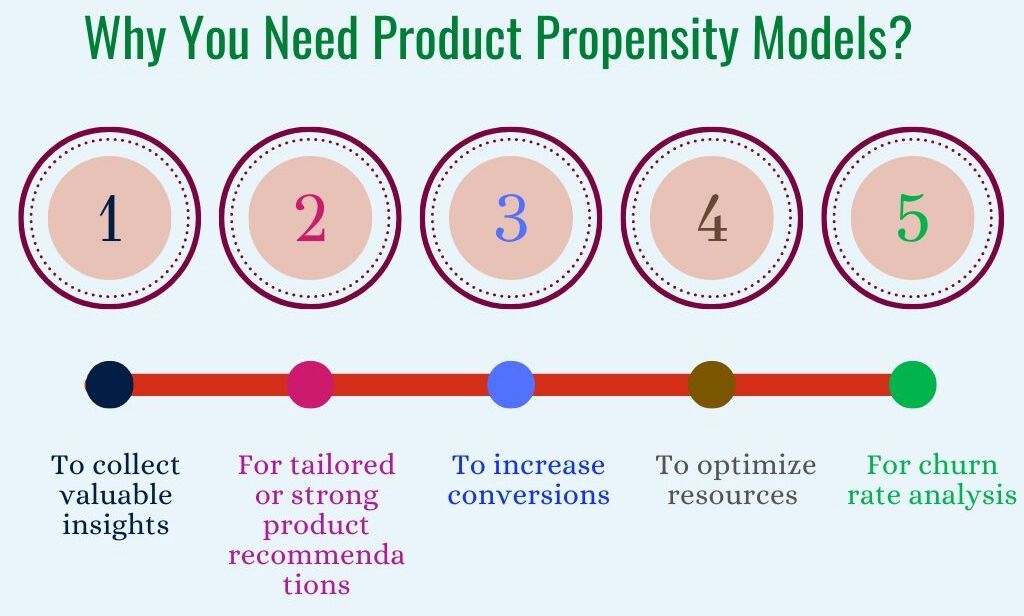
Predictive modeling is a mathematical process used to predict future events or outcomes by analyzing patterns in a given set of input data. It is a crucial component of [predictive analytics](https://www.techtarget.com/searchbusinessanalytics/definition/predictive-analytics), a type of data analytics which uses current and historical data to forecast activity, behavior and trends.

**Examples** : estimating the quality of a sales lead, the likelihood of spam or the probability someone will click a link or buy a product.

"Predictive modeling is a form of data mining that analyzes historical data with the goal of identifying trends or patterns and then using those insights to predict future outcomes," explained Donncha Carroll a partner in the revenue growth practice of Axiom Consulting Partners. "Essentially, it asks the question, 'have I seen this before' followed by, 'what typically comes after this pattern.'"

## **What is Propensity Modeling?**

Propensity Modeling is a statistical technique used to predict the chances of certain events happening in the future. With the increasing use of machine learning, companies can build robust propensity models and make accurate forecasts.

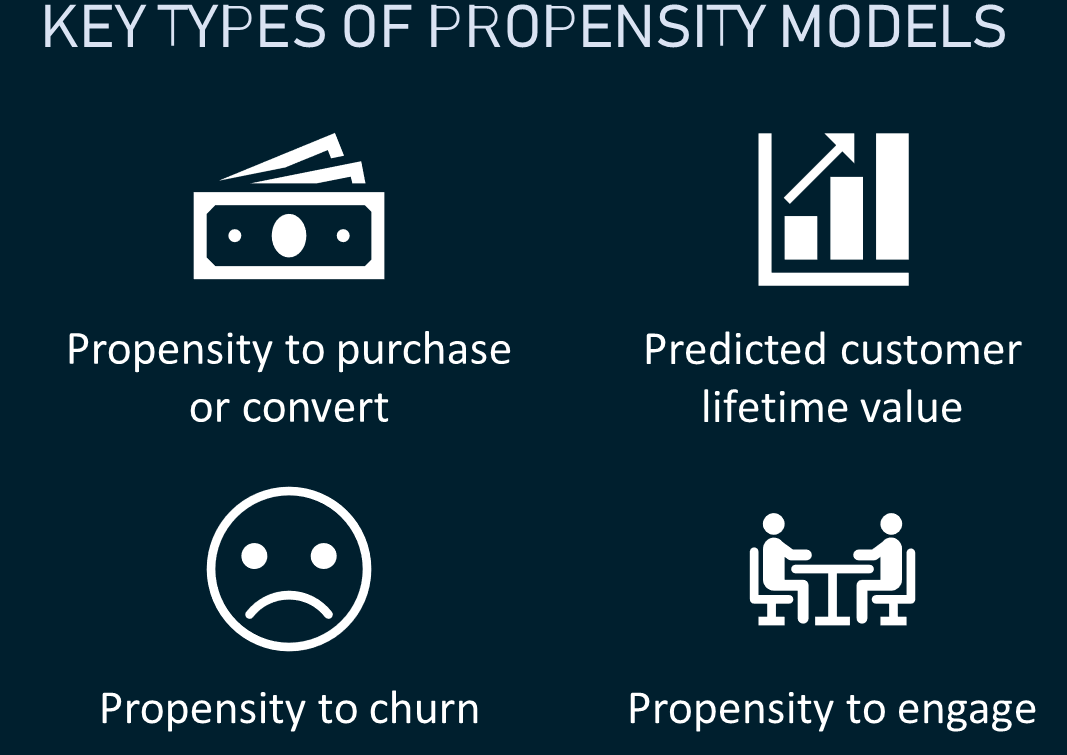


## **Benefits Of Propensity Modeling**

1. Such models can be used to predict customer behavior. By understanding the factors that influence customer behavior, businesses can use them to target their marketing and sales efforts in an effective manner.
2. Propensity modelling can help you make smarter decisions by throwing up insights that would otherwise not be available to an enterprise.
3. They can also be used to predict the value each customer brings in real time. By understanding which customers are most likely to make a purchase, businesses can gauge the value of those customers and allocate resources accordingly.
4. Such models can be used to optimize customer acquisition strategies. By figuring out which customers are most likely to become regulars, businesses can identify which campaigns are resonating.
5. These models can be used to optimize [customer retention strategies](https://www.expressanalytics.com/blog/why-retail-need-data-analytics/). By knowing in advance which customers are most likely to churn, businesses can identify the measures that can then stop them from leaving.
6. Propensity models can be used to predict the profitability of a given [customer segment](https://www.expressanalytics.com/blog/rfm-analysis-for-customer-segmentation/). By understanding which customers are most likely to generate revenue, businesses can optimize their marketing and sales efforts accordingly.
7. They can also be used to improve customer service and improve the levels of customer satisfaction.

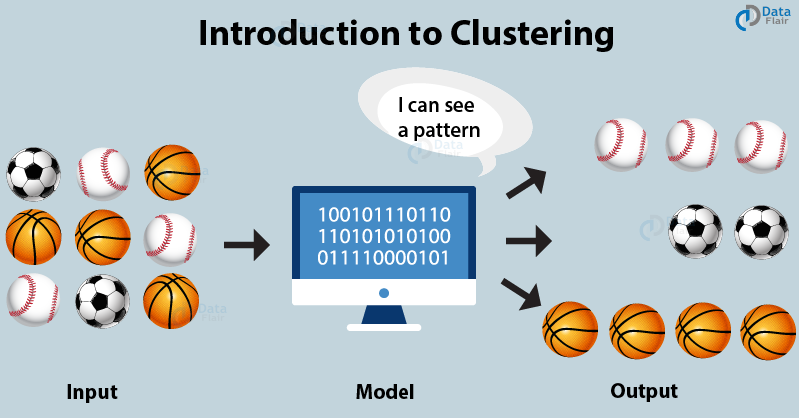
## **The Limitations Of Propensity Modeling**

Propensity models can be biased if the data used to create them is not representative of the population as a whole. Propensity models are only as good as the assumptions made around customer behavior. If these assumptions are inaccurate, the predictions will be as well. But despite such limitations, propensity modeling can be a valuable tool for businesses looking to better understand and predict customer behavior.

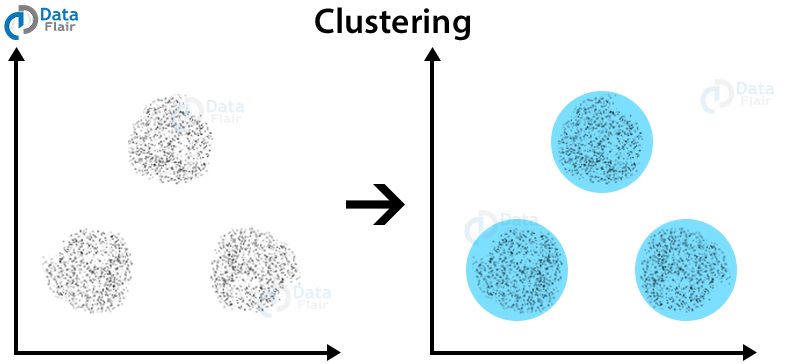


**Cluster modelling**

Clustering models **allow you to categorize records into a certain number of clusters**. This can help you identify natural groups in your data. Clustering models focus on identifying groups of similar records and labeling the records according to the group to which they belong.



**Clustering Example–** The data-points that are clustered together are in groups that hold similar data. Then we can further distinguish these clusters through the identification of three clusters as visualized below –

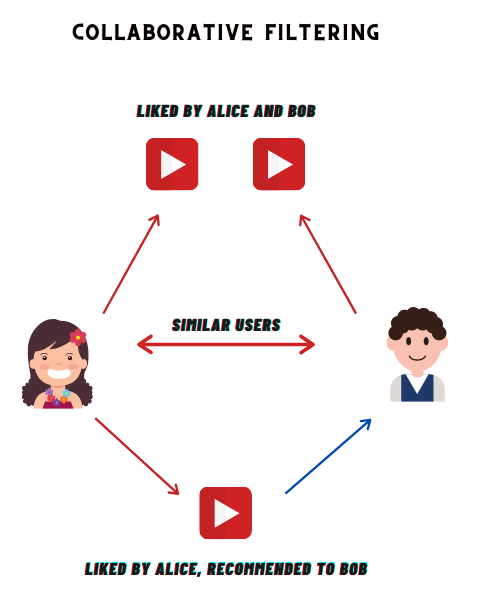


The goal of cluster analysis is to find clusters such that the observations within each cluster are quite similar to each other, while observations in different clusters are quite different from each other.

## **What Is Collaborative Filtering?**

Collaborative filtering is a technique that can filter out items that a user might like on the basis of reactions by similar users.

It works by searching a large group of people and finding a smaller set of users with tastes similar to a particular user. It looks at the items they like and combines them to create a ranked list of suggestions.



Amazon is known for its use of collaborative filtering, **matching products to users based on past purchases**. For example, the system can identify all of the products a customer and users with similar behaviors have purchased and/or positively rated.

**What are the two types of collaborative filtering?**

There are two classes of Collaborative Filtering:

* **User-based**, which measures the similarity between target users and other users.
* **Item-based**, which measures the similarity between the items that target users rate or interact with and other items.

