

Hello students, I Ms. Likita Shetty welcome you all to the module 02 of the paper Data Mining.

In this module we are going to cover the topics from Unit 1 i.e. Data Mining Overview which includes Tasks of Data Mining, Which kinds of applications are targeted? That is your Business Intelligence and Web Search Engines, and What are the benefits of data mining?

The outline for the today's session is:

We will have the tasks of data mining, which are the tasks that are needed in data mining.

Second, which kinds of applications are targeted? That is your Business Intelligence and Web Search Engines and third benefits of data mining.

At the end of the session, the learning outcomes will be:

The student will be able to define and describe the tasks of data mining.

Understand the kinds of applications that are targeted in data mining basically with respect to Business Intelligence and Web Search Engines. Thirdly understand the benefits of data mining.

So let's start with the first topic, that is tasks of data mining.

The first one is Anomaly Detection, which is also called as Outlier or Change or Deviation Detection.

In this, the identification of unusual data records that might be interesting or data errors that require further investigation.

Like for example, suppose we consider a data set which contains many objects that do not comply with the general behavior or the model of the data.

This data objects will be called as Outliers.

The analysis of Outlier data is referred to as Outlier Analysis or Anomaly Mining.

The second task is Association Rule Mining, which is also called as Dependency Modeling.

It searches for relationships between your variables.

For example, if you consider a supermarket, it might gather data on customer purchasing habits.

Using Association Rule Learning the supermarket can determine which products are frequently brought together and use this information for marketing purposes. This is also referred to as Market Basket Analysis.

The third task is Clustering.

In this it is a task of discovering groups and structures in the data that are in some way or another similar without using the known structures in the data.

Now cluster analysis can be performed on the company's customer data to identify the homogeneous subpopulation of customers.

This clusters may represent individual target groups for marketing.

Now next is Classification.

Classification is the task of generalizing known structure to apply to the new data.

Basically, for example, if we consider an email program that might attempt to classify an email as a legitimate or as spam.

So here it tries to classify the email either as a legitimate email or a spam.

Next task is Predictive Modeling.

It is also a commonly used statistical technique to predict the future behavior.

Predictive modeling solutions are a form of data mining technology that works by analyzing historical and current data and generating a model to help predict future outcomes.

So, in predictive modeling it considers your previous historical data and your current data and tends to give the output for your future data.

Next is summarization.

It provides a more compact representation of the dataset, including visualization and report generation.

Next topic is which kinds of applications are targeted ??

As a highly application driven discipline, Data mining has great successes in many applications.

In this we have two highly successful and popular application examples of data mining which our Business Intelligence and Web Search Engines.

The first one i.e. Business Intelligence.

Business Intelligence is very critical for businesses to acquire a better understanding of the commercial context of their organization, such as their customers, the market, supply, resources, and competitors.

Now, How important is this Business Intelligence??

Without data mining, many businesses may not be able to perform effective marketing analysis, compare the customer feedback based on the similar products, discover the strengths and weaknesses of their competitors, retain highly valuable customers and make smart business decisions.

Data mining is the core of your Business Intelligence.

Even Classification and Prediction techniques becomes the core of your predictive analysis in Business Intelligence, for which there are many applications in analyzing markets, supplies, and sales.

Moreover, clustering also plays a central role in customer relationship management wherein group of customers based on their similarities are grouped into one.

Next application is your Web Search Engine.

A Web Search Engine is a specialized computer server that searches for information on the web.

The search results of a user query are often returned as a list which is also called as hits.

The hits may consist of web page, web pages, images or any other types of files.

Now web search engines are essentially very large data mining applications.

Here various data mining techniques are used in all aspects of search engines, ranging from crawling, for example deciding which pages should be crawled and their crawling frequencies.

Next the indexing.

What is indexing?

Indexing is selecting the pages which has to be indexed and deciding to which extent the index should be constructed.

Next

Searching that is deciding how the pages should be ranked and which advertisement should be added and how the search results can be personalized or made context aware.

Next topic is Benefits of data mining.

Seeing into the benefits of data mining, it helps us in identifying the fraudulent transactions based on user

behavior and pattern of the data which will help in the banks as well as financial institution to issue loan, credit card etc.based on the user behavior.

It helps in getting customers who can buy the product based on the relevant advertising campaign based on their previous purchase behavior, as well as the search patterns on Google.

Data mining techniques such as machine learning helps in increasing the sales of a business.

This will benefit the customers, advertisers and marketing companies.

It also helps in improving the layout of retail and other grocery stores based on customer feedback and previous purchases, so this will help the retail stores to keep the most sold items at the right place to have the highest attention of customers.

It also helps in obtaining the desired search results of queries posed to E-Commerce websites such as Amazon, Snapdeal, Walmart, Flipkart, eBay etc.

Even search engines like Google, Yahoo, Bing, ask.com etc.

It has been also used in many different areas or domains i.e. bioinformatics, medicine, genetics, education,agriculture, law enforcement, e-marketing, electrical power engineering etc.

These are the references.

Thank you.