5 MARKS

20 December 2023 21:09

Q How Hierarchal clustering helps in data mining? Discuss key issues

Ans

Hierarchal clustering is a method of cluster analysis that create a hierarchal representation of a clusters in a dataset
The method starts by treating each data point as a seprate clusters and iteratively combines the closest clusters until a stopping criterion is reached

The result of hierarchical clustering is a tree like structure called dendrogram which illustrate the hierarchal relationship among the clusters.

Advantages

- The ability to handle non convex clusters and clusters of different sizes and densities.
- Ability to handle missing data and noisy data.
- ability to reveal the hierarchical structure of the data, which can be useful for understanding the relationship among the clusters.

Drawbacks

- The need for a criterion to stop the clustering process and determine the final number of clusters.
- The computational cost and memory requirements of the method can be high especially for large datasets
- The algorithm can never be changed or deleted once it was done previously

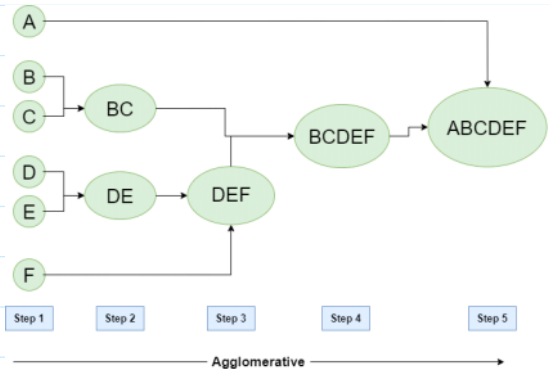
Types of Hierarchical clustering

- Agglomerative clustering
- Divisive clustering

Agglomerative clustering

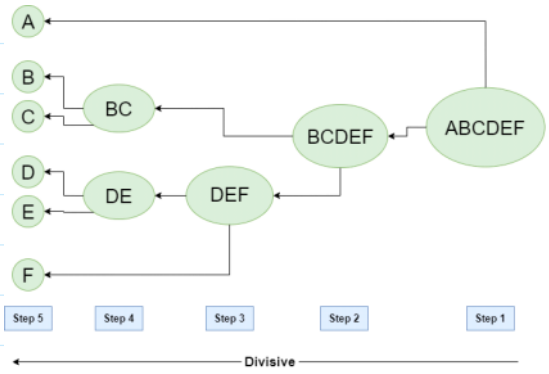
Initially consider every data point as an **individual** Cluster and at every step, merge the nearest

Initially consider every data point as an **individual** Cluster and at every step, merge the nearest pairs of the cluster. (It is a bottom-up method). At first, every dataset is considered an individual entity or cluster. At every iteration, the clusters merge with different clusters until one cluster is formed.



Divisive Hierarchical

We can say that Divisive Hierarchical clustering is precisely the **opposite** of Agglomerative Hierarchical clustering. In Divisive Hierarchical clustering, we take into account all of the data points as a single cluster and in every iteration, we separate the data points from the clusters which aren't comparable. In the end, we are left with N clusters.



Q4 Discuss the need of data cleaning?

Ans

Ensuring Data Accuracy and Reliability:

- Missing Values: Incomplete information can skew results and misrepresent trends. Data cleaning involves identifying and addressing missing values through methods like imputation or dropping incomplete entries.

- Inconsistent Formats: Different units, spellings, and data types can hinder analysis and create errors. Data cleaning standardizes formats, ensuring consistency and accuracy.

- Outliers: Extreme values or anomalies can distort findings. Data cleaning identifies and

handles outliers effectively, preventing them from skewing results.

2. Improving Model Performance:

- Garbage In, Garbage Out: Dirty data leads to flawed models and unreliable predictions. Cleaning improves data quality, leading to robust models that generate accurate and insightful results.
- Reduced Training Time: Efficient algorithms thrive on clean data. Data cleaning optimizes data, improving processing speed and reducing training time for machine learning models.
- Enhanced Interpretability: Cleaner data leads to clearer patterns and relationships. This, in turn, makes models easier to understand and interpret, providing valuable insights into the underlying data.

3. Saving Time and Resources:

- Avoiding Downstream Errors: Dirty data can propagate errors throughout the analysis pipeline, costing time and resources to fix later. Cleaning upfront avoids these costly rework loops.
- Improved Data Sharing and Collaboration: Clean data is easier to share and collaborate on across teams and organizations. This promotes efficient knowledge sharing and facilitates research and development.
- Streamlining Future Analysis: By investing in data cleaning now, you create a foundation for future analyses. Clean data is readily available for subsequent projects, minimizing preparation time and maximizing efficiency.

Q Explain Apriori algorithm for frequent item set using candidate generation

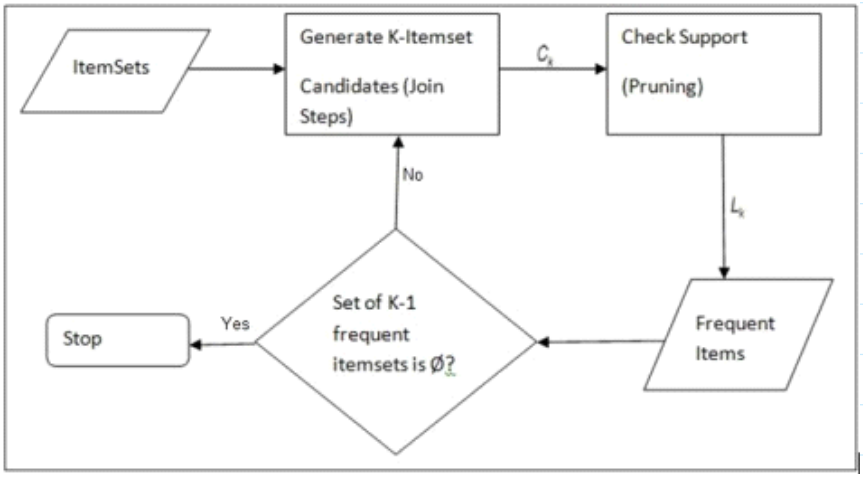
Ans

The apriori algorithm is an algorithm that is generally used in the field of data mining & association rule learning. It is

The apriori algorithm is an algorithm that is generally used in the field of data mining & association rule learning. It is used to identify frequent itemsets in a dataset & generate an association rule based on the itemset

Suppose You have a data base about the items a customer purchases from the store The apriori algorithm help to uncover intresting relationships & patterns in this data it does that by finding the sets of items that occur together, frequently

foreg
↳ The algorithm would discover that when a customer buys bread, they often end up buying butter & eggs as well This indicates a strong association b/w these items
These associations help business to make decisions to improve sales, customer satisfaction etc.



Setting the Stage:

- You start by defining the minimum support threshold - the percentage of transactions containing an itemset to be considered "frequent."

- You also provide the initial set of frequent itemsets, containing individual items with support exceeding the threshold.

2. Candidate Generation:

- This is the heart of Apriori. Here's the iterative process:
- In the first iteration, you combine each pair of frequent items from the previous stage to

- This is the heart of Apriori. Here's the iterative process:
- In the first iteration, you combine each pair of frequent items from the previous stage to generate candidate itemsets of size 2 (pairs).
- In subsequent iterations, you combine frequent itemsets of size k-1 to generate candidate itemsets of size k (e.g., combining triplets if you're in the third iteration).
- Important Rule: During each generation, a crucial step involves pruning: any candidate itemset containing a non-frequent subset is discarded. This significantly reduces the search space and prevents processing unnecessary itemsets.

3. Support Counting:

- Once you have generated the candidate itemsets for a specific size, you scan the entire transaction database to count the support of each candidate.

- Any candidate exceeding the minimum support threshold becomes a new frequent itemset.

4. Iteration and Termination:

- The process continues with iterating candidate generation, support counting, and adding new frequent itemsets until no new itemsets of size k+1 can be generated from frequent itemsets of size k. This signals that you've discovered all frequent itemsets within the defined support threshold.

---

06 what is the need to maintain hierarchy in web structure? Illustrate what are different rules for web mining.

Ans

Maintaining hierarchy in web structure is essential for several reasons:

1. Organized Information: Hierarchy helps organize information in a structured manner, making it easier for users to navigate and find what they are looking for. A clear hierarchy ensures that content is logically arranged, improving the user experience.
2. Navigation: Hierarchy provides a navigation system that guides users through different levels of information. It helps users understand where they are on a website and how to move to other sections or pages.
3. User Experience: A well-defined hierarchy enhances the overall user experience. Users can quickly grasp the structure of the website, leading to a more efficient and enjoyable interaction.
4. Search Engine Optimization (SEO): Search engines use website hierarchy to understand the

grasp the structure of the website, leading to a more efficient and enjoyable interaction.

4. Search Engine Optimization (SEO): Search engines use website hierarchy to understand the organization of content. A clear hierarchy with meaningful headings and subheadings can improve a site's SEO, making it more likely to appear in relevant search results.
5. Content Prioritization: Hierarchy allows for the prioritization of content. Important information can be placed prominently at higher levels, while more detailed or specific content can be nested at lower levels.
6. Consistency: A consistent hierarchy across a website provides a uniform structure, which helps in brand consistency and recognition. Users become familiar with the layout, making it easier for them to navigate different sections.
7. Scalability: As websites grow and evolve, maintaining a hierarchical structure makes it easier to scale and add new content. New pages or sections can be integrated seamlessly into the existing structure.

Web Mining:

web mining is the process of data mining techniques to automatically discover and extract informations from web documents and services. The main purpose of web mining is discovering useful information from www and its usage pattern.

Application of web mining:

— Personalized Marketing

web mining can be used to analyze customer behavior on websites and social media platforms This info can be used to create personalized marketing campaigns that target customers based on their preferences

— E commerse

web mining can be used to analyse customer behaviours on E commerse websites enhance user experiance and increase sales. by recommending things or product on their preferences.

— Search engine optimization

web mining can be used to analyse search engine query and search engine result pages. (SERPs) This information can be used to improve the visibility of websites in search engine result

— fraud detection

- fraud detection

- Sentimental Analysis

- web content analysis

web mining can be used to analyze web content and extract valuable information such as keywords, topic and themes

- customer services

How web mining work:

Raw Data $\xrightarrow[\text{tools}]{\text{Mining}}$ Patterns $\xrightarrow[\text{\& visualization}]{\text{Representation}}$ Knowledge.

There are various rules of web usage mining which are as follows –

Preprocessing – The web usage log is not in a format that is accessible by mining applications. For some data to be used in a mining application, the data can be required to be reformatted and cleansed. There are some issues specifically related to the use of weblogs. There are some steps included in the processing phase include cleansing, user identification, session identification, path completion, and formatting.

Data structure – There are several unique data structures have been proposed to keep track of patterns identified during the web usage mining process. A basic data structure that is used is called a tree. A tree is a rooted tree, where each path from the root to a leaf represents a sequence. Trees can save strings for pattern matching applications. The only problem with trees is space requirements.

Pattern discovery – The most common data mining technique used on clickstream data is that of uncovering traversal patterns. A traversal pattern is a group of pages inspected by a user in a session. The other type of pattern may be uncovered by web usage mining. Patterns are found using different combinations which are used to discover different features and for different purposes.

Pattern analysis – When patterns are discovered, they must be analyzed to determine how that information can be used. Some of the patterns can be deleted and not determined to be of interest.

Q what is data mining and explain the knowledge discovery process

Ans

The process of extract information to identify patterns, terends and useful data that would allows the business to take data driven decisions from huge set of data is called data mining.

In other words, we can say that data mining is the process of investigating hidden patterns of information to various perspectives for categorization into useful data, which is collected and assembled in particular area such as data warehouse, effective analysis, data mining algo, helping decision making and other data requirements to ~~method~~ eventually cost cutting and generating revenue

Data mining is the act of automatically searching for large stores of information to find trends and patterns that go beyond simple analysis procedures.
Data mining utilizes complex mathematical algorithms for data segments and evaluates the probability of future events

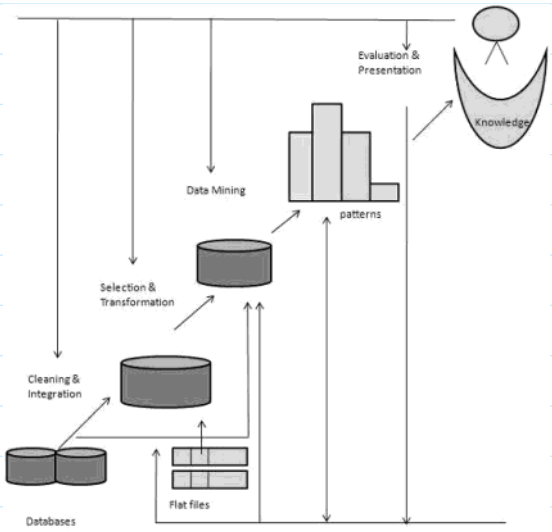
Types of data mining

Data mining can be performd the following type of data.

- Relational data base
- Object oriented data base
- Data ware house
- Transactional data base

KNOWLEDGE DISCOVERY PROCESS

KDD is a process that involves the extraction of useful, previously unknown and potentially valuable information foram large datasets The KDD process is an iterative process and it requires multiple iterations to extract accurate knowledge from the data.



Data Cleaning

Data cleaning is defined as removal of noisy and irrelevant data from collection.

1. Cleaning in case of **Missing values**.
2. Cleaning **noisy** data, where noise is a random or variance error.
3. Cleaning with **Data discrepancy detection** and **Data transformation tools**.

Data Integration

Data integration is defined as heterogeneous data from multiple sources combined in a common source(DataWarehouse). Data integration using **Data Migration tools, Data Synchronization tools and ETL**(Extract-Load-Transformation) process.

Data Selection

Data selection is defined as the process where data relevant to the analysis is decided and retrieved from the data collection. For this we can use **Neural network, Decision Trees, Naive bayes, Clustering**, and **Regression** methods.

Data Transformation

Data Transformation is defined as the process of transforming data into appropriate form required by mining procedure. Data Transformation is a two step process:

1. **Data Mapping**: Assigning elements from source base to destination to capture transformations.
2. **Code generation**: Creation of the actual transformation program.

Data Mining

Data mining is defined as techniques that are applied to extract patterns potentially useful. It transforms task relevant data into **patterns, and** decides purpose of model using **classification** or **characterization**.

Pattern Evaluation

Pattern Evaluation is defined as identifying strictly increasing patterns representing knowledge based on given measures. It find **interestingness score** of each pattern, and uses **summarization** and **Visualization** to make data understandable by user.

Knowledge Representation

This involves presenting the results in a way that is meaningful and can be used to make decisions.

Q what is classification

Ans The classification algorithm is a supervised learning technique that is used to identify the category of new observations on the basis of traning data

In classification a program learns from the given datasets or observations and then classifies new observation into no of classes or groups such as:

- → Yes or No
- → 0 or 1
- → Spam or Not Spam
- → Cat or dog.

Classes can be called as target s / label s or categories

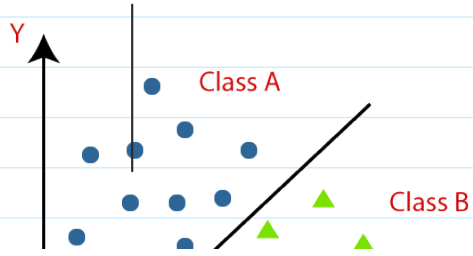
unlike regression the output variable of classification is a categorical not a value. such as "Green" or "Blue", "fruit" or "animal",

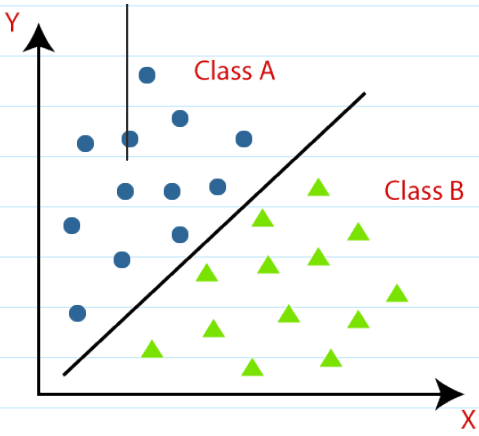
In classification algo, a discrete output function (y) is mapped to input variable (x)

$$y = f(x)$$

The Best eg of classification is email spam detector

The main goal of the Classification algorithm is to identify the category of a given data set and these algorithm are mainly used to predict the output for the categorical data.





There are two type of classification.

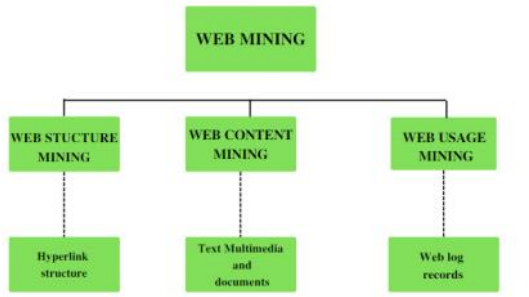
- **Binary Classifier:** If the classification problem has only two possible outcomes, then it is called as Binary Classifier.
  Examples: YES or NO, MALE or FEMALE, SPAM or NOT SPAM, CAT or DOG, etc.

- **Multi-class Classifier:** If a classification problem has more than two outcomes, then it is called as Multi-class Classifier.
  Example: Classifications of types of crops, Classification of types of music.

Q what is web content mining? write its different steps?

Ans

Web mining can be broadly divided into three different types of techniques of mining: Web Content Mining, Web Structure Mining, and Web Usage Mining. These are explained as following below.



Categories of Web Mining

1. **Web Content Mining:** Web content mining is the application of extracting useful information from the content of the web documents. Web content consist of several types of data – text, image, audio, video etc. Content data is the group of facts that a web page is designed. It can provide effective and interesting patterns about user needs. Text documents are related to text mining, machine learning and natural language processing. This mining is also known as text mining. This type of mining performs scanning and mining of the text, images and groups of web pages according to the content of the input.
2. **Web Structure Mining:** Web structure mining is the application of discovering structure

information from the web. The structure of the web graph consists of web pages as nodes, and hyperlinks as edges connecting related pages. Structure mining basically shows the structured summary of a particular website. It identifies relationship between web pages linked by information or direct link connection. To determine the connection between two commercial websites, Web structure mining can be very useful.

3. **Web Usage Mining:** Web usage mining is the application of identifying or discovering interesting usage patterns from large data sets. And these patterns enable you to understand the user behaviors or something like that. In web usage mining, user access data on the web and collect data in form of logs. So, Web usage mining is also called log mining.

Process of Web Mining

The process of web mining typically involves the following steps -

- **Data collection** -
  Web data is collected from various sources, including web pages, databases, and APIs.
- **Data pre-processing** -
  The collected data is pre-processed to remove irrelevant information, such as advertisements and duplicate content.
- **Data integration** -
  The pre-processed data is integrated and transformed into a structured format for analysis.
- **Pattern discovery** -
  Web mining techniques are applied to identify patterns, trends, and relationships.
- **Evaluation** -
  The discovered patterns are evaluated to determine their significance and usefulness.
- **Visualization** -
  The analysis results are visualized through graphs, charts, and other visualizations.

Points	Data Mining	Web Mining
Definition	Data Mining is the process that attempts to discover pattern and hidden knowledge in large data sets in any system.	Web Mining is the process of data mining techniques to automatically discover and extract information from web documents.
Application	Data Mining is very useful for web page analysis.	Web Mining is very useful for a particular website and e-service.
Target Users	Data scientist and data engineers.	Data scientists along with data analysts.
Access	Data Mining access data privately.	Web Mining access data publicly.
Structure	In Data Mining get the information from explicit structure.	In Web Mining get the information from structured, unstructured and semi-structured web pages.
Problem Type	Clustering, classification, regression, prediction, optimization and control.	Web content mining, Web structure mining.
Tools	It includes tools like machine learning algorithms.	Special tools for web mining are Scrapy, PageRank and Apache logs.

Skills	It includes approaches for data cleansing, machine learning algorithms. Statistics and probability.	It includes application level knowledge, data engineering with mathematical modules like statistics and probability.

Q Explain Naive Bayes classification ?

Ans

Naïve Bayes Classifier Algorithm

- Naïve Bayes algorithm is a supervised learning algorithm, which is based on **Bayes theorem** and used for solving classification problems.
- It is mainly used in *text classification* that includes a high-dimensional training dataset.
- Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.
- **It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.**
- Some popular examples of Naïve Bayes Algorithm are **spam filtration, Sentimental analysis, and classifying articles.**

Why is it called Naïve Bayes?

The Naïve Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as:

- **Naïve**: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other.
- **Bayes**: It is called Bayes because it depends on the principle of Bayes' Theorem.

Bayes' Theorem:

- Bayes' theorem is also known as **Bayes' Rule** or **Bayes' law**, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability.
- The formula for Bayes' theorem is given as:

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)}$$

Where,

P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B.

P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

P(A) is Prior Probability: Probability of hypothesis before observing the evidence.

P(B) is Marginal Probability: Probability of Evidence.

Advantages of Naïve Bayes Classifier:

- Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
- It can be used for Binary as well as Multi-class Classifications.

- It performs well in Multi-class predictions as compared to the other Algorithms.
- It is the most popular choice for **text classification problems**.

Disadvantages of Naïve Bayes Classifier:

- Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.

Applications of Naïve Bayes Classifier:

- It is used for **Credit Scoring**.
- It is used in **medical data classification**.
- It can be used in **real-time predictions** because Naïve Bayes Classifier is an eager learner.
- It is used in Text classification such as **Spam filtering** and **Sentiment analysis**.

Q what are the reasons that make OLAP better than OLTP? Discuss two scenarios where OLAP can be used to extract information

Ans

Online Analytical Processing (OLAP)

Online Analytical Processing (OLAP) consists of a type of software tool that is used for data analysis for business decisions. OLAP provides an environment to get insights from the database retrieved from multiple database systems at one time.

Any type of Data Warehouse System is an OLAP system. The uses of the OLAP System are described below.

- Spotify analyzed songs by users to come up with a personalized homepage of their songs and playlist.
- Netflix movie recommendation system.

OLAP

Benefits of OLAP Services

- OLAP services help in keeping consistency and calculation.
- We can store planning, analysis, and budgeting for business analytics within one platform.
- OLAP services help in handling large volumes of data, which helps in enterprise-level business applications.
- OLAP services help in applying security restrictions for data protection.
- OLAP services provide a multidimensional view of data, which helps in applying operations on data in various ways.

Online Transaction Processing (OLTP)

Online transaction processing provides transaction-oriented applications in a 3-tier

architecture. OLTP administers the day-to-day transactions of an organization.

OLTP Examples

An example considered for OLTP System is ATM Center a person who authenticates first will receive the amount first and the condition is that the amount to be withdrawn must be present in the ATM. The uses of the OLTP System are described below.

- ATM center is an OLTP application.
- OLTP handles the ACID properties during data transactions via the application.
- It's also used for Online banking, Online airline ticket booking, sending a text message, add a book to the shopping cart.

OLTP vs OLAP

Benefits of OLTP Services

- OLTP services allow users to read, write and delete data operations quickly.
- OLTP services help in increasing users and transactions which helps in real-time access to data.
- OLTP services help to provide better security by applying multiple security features.
- OLTP services help in making better decision making by providing accurate data or current data.
- OLTP Services provide Data Integrity, Consistency, and High Availability to the data.

APPLICATIONS OF OLAP:

Business Reporting for sales: The Business Reporting gives an overview of the sales activity in the sales activities within an organization. It shows the trends in the sales over a certain time period. It also analyzes the different steps for sales and sales executive performance. These reports can be used to analyze the sales data and assess the situation to make the best decisions to undertake.

Marketing: Industries like digital marketing, health care, eCommerce, and finance uses OLAP in their marketing.Example: Market Basket Analysis is a technique that gives the careful study of purchases done by a customer in a supermarket. This concept identifies the pattern of frequent purchase items by customers. This analysis can help to promote deals, offers, sale by the companies and data mining techniques helps to achieve this analysis task.

Q what is the application of DHP? Does it affect performance of algorithm? Justify your answer with example.

Ans

Direct Hashing and Pruning: Hashing & Pruning is very popular association rule mining

technique to improve the performance of traditional Apriori algorithm. Hashing technique uses hash function to reduce the size of candidate item set.

Direct Hashing & Pruning (DHP),Perfect Hashing &Pruning (PHP) are the basic hashing algorithms. Many algorithms have been also proposed by researchers like Perfect Hashing Scheme (PHS), Sorting-Indexing andTrimming (SIT),HMFS etc.

THP arranges the item sets into vertical format and then hashed the transactions id (TID) of candidatek item sets into hash table bucket corresponding to that item set.

Pros & Cons of DHP algorithm:

1. DHP uses simple hash function to reduce the size of candidate item set.
2. Size of hash table is small which requires less memory to store.
3. There is collision problem in DHP algorithm.
4. More database scans are required to count the support of collided item .

"Direct hashing and pruning" (DHP) is a technique specifically used in the context of association rule mining, particularly related to the Apriori algorithm for identifying frequent itemsets within a transactional dataset.

Application of DHP:

- Reduce the number of candidate itemsets generated: Instead of generating all possible combinations of items to search for frequent sets, DHP uses hashing functions to estimate the support of potential itemsets before actually generating them. This significantly reduces the computational overhead and processing time involved, especially for large datasets.

- Improve efficiency of support counting: By pre-filtering potential frequent itemsets through hashing, DHP avoids unnecessary database scans to count their actual support. This further optimizes the algorithm and reduces overall runtime.

Impact on Performance:

DHP positively affects the performance of the Apriori algorithm in several ways:

- Faster processing: Due to the reduced candidate generation and efficient support counting, DHP can significantly decrease the overall runtime compared to the traditional Apriori approach. This is especially noticeable for large datasets with numerous potential

itemsets.

- Reduced memory usage: By minimizing the number of generated and processed itemsets, DHP lowers the memory footprint of the algorithm, making it more suitable for resource-constrained environments.

Example:

Imagine a transactional dataset with thousands of items and millions of transactions. To find frequent itemsets using traditional Apriori, the algorithm would need to generate and count the support for all possible combinations of items (resulting in an exponential number as the itemset size increases). This exhaustive process could take hours or even days for large datasets.

DHP, on the other hand, utilizes hashing functions to estimate the support of itemsets before formally generating them. This pre-filtering dramatically reduces the number of candidates that need to be checked, potentially enabling the algorithm to identify frequent itemsets within minutes or hours, depending on the dataset size and complexity.

Q.

5. The distance between some Indian cities are given below, Apply the clustering algorithm to make three clusters. Indicate intermediate steps.

	Bathinda	Patiala	Delhi	Amritsar	Mathura
Bathinda	0	190	400	250	460
Patiala	190	0	240	225	300
Delhi	400	240	0	450	60
Amritsar	250	225	450	0	510
Mathura	460	300	60	510	0

To apply clustering to the given distance matrix, we can use hierarchical clustering. In hierarchical clustering, the algorithm starts with each data point as a separate cluster and then iteratively merges the closest clusters until the desired number of clusters is reached. The process involves creating a dendrogram that represents the hierarchy of cluster merges. I'll walk you through the intermediate steps:
Step 1: Create a distance matrix based on the given distances.

plaintext Copy code

```
|           | Bathinda | Patiala | Delhi | Amritsar | Mathura |
|-----------|----------|---------|-------|----------|---------|
| Bathinda  |    0     |   190   |  400  |   250    |   460   |
| Patiala   |   190    |    0    |  240  |   225    |   300   |
| Delhi     |   400    |   240   |   0   |   450    |    60   |
| Amritsar  |   250    |   225   |  450  |    0     |   510   |
| Mathura   |   460    |   300   |   60  |   510    |    0    |
```

Step 2: Start with each city as a separate cluster.

```plaintext
Clusters: [Bathinda], [Patiala], [Delhi], [Amritsar], [Mathura]
```

Step 3: Identify the closest pair of clusters and merge them. Update the distance matrix.

```plaintext
Clusters: [Bathinda, Patiala], [Delhi], [Amritsar], [Mathura]

Updated Distance Matrix:
|           | Bathinda/Patiala | Delhi | Amritsar | Mathura |
|-----------|------------------|-------|----------|---------|
| Delhi     |       240        |   0   |   450    |    60   |
| Amritsar  |       225        |  450  |    0     |   510   |
| Mathura   |       300        |   60  |   510    |    0    |
```

Step 4: Repeat Steps 3 until the desired number of clusters is reached.

```plaintext
Clusters: [Bathinda, Patiala, Delhi], [Amritsar], [Mathura]

Updated Distance Matrix:
|           | Bathinda/Patiala/Delhi | Amritsar | Mathura |
|-----------|------------------------|----------|---------|
| Amritsar  |          225           |    0     |   510   |
| Mathura   |          300           |   510    |    0    |
```

Step 5: Repeat Steps 3 once more.

```plaintext
Clusters: [Bathinda, Patiala, Delhi], [Amritsar, Mathura]

Updated Distance Matrix:
|               | Bathinda/Patiala/Delhi | Amritsar/Mathura |
|---------------|------------------------|------------------|
| Amritsar/Mathura|          225          |       510        |
```

Now, we have three clusters: [Bathinda, Patiala, Delhi], [Amritsar, Mathura]. The dendrogram would represent the hierarchy of cluster merges.