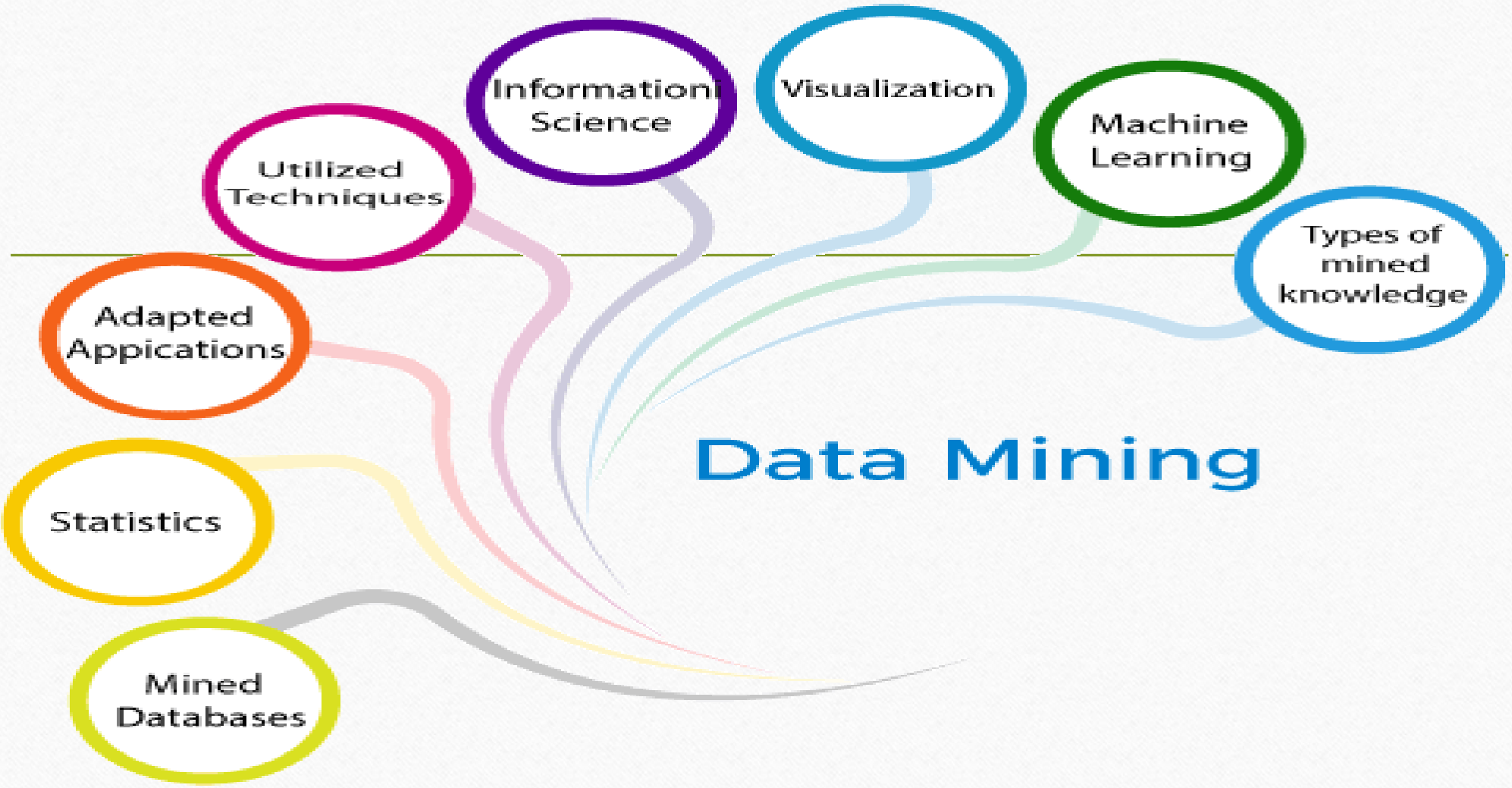
Classification Of Data Mining System

Presented By: Akash Pratap Singh


Information Science
Visualization
Machine Learning
Utilized Techniques
Types of mined knowledge
Adapted Appications
Data Mining
Statistics
Mined Databases

Classification Of Data Mining System

- Classification based on the mined databases
- Classification based on the type of mined knowledge
- Classification based on Information Science.
- Classification based on utilized techniques
- Classification based on adapted applications
- Classification based on statistics
- Classification based on Machine Learning
- Classification based on visualization

Classification Based on the Mined Databases

- A data mining system can be classified based on the types of databases that have been mined. A database system can be further segmented based on distinct principles, such as data models, types of data, etc., which further assist in classifying a data mining system..

- For example, if we want to classify a database based on the data model, we need to select either relational, transactional, object-relational or data warehouse mining systems.

Classification Based on the type of Knowledge Mined

- **Characterization:** This refers to summarizing data of class under study. This class under study is called as Target Class.
- **Discrimination:** It refers to the mapping or classification of a class with some predefined group or class.
- **Association Analysis:** This process refers to the process of uncovering the relationship among data and determining association rules.
- **Correlation Analysis:** It is a kind of additional analysis performed to uncover interesting statistical correlations between associated- attribute-value pairs or between two item sets to analyze that if they have positive, negative or no effect on each other.
- **Classification:** It predicts the class of objects whose class label is unknown. Its objective is to find a derived model that describes and distinguishes data classes or concepts. The Derived Model is based on the analysis set of training data i.e. the data object whose class label is well known.
- **Prediction**: It is used to predict missing or unavailable numerical data values rather than class labels. Regression Analysis is generally used for prediction. Prediction can also be used for identification of distribution trends based on available data.
- **Outlier Analysis**: Outliers may be defined as the data objects that do not comply with the general behavior or model of the data available.
- **Evolution Analysis** Evolution analysis refers to the description and model regularities or trends for objects whose behavior changes over time.

Classification Based on the Techniques Utilized

- A data mining system can also be classified based on the type of techniques that are being incorporated. These techniques can be assessed based on the involvement of user interaction involved or the methods of analysis employed.

Classification Based on the Applications Adapted

- Finance
- Telecommunications
- Stock Markets
- E-mail
- DNA

Classification Based on the Visualization

- The visualization approach to data mining is based on an assumption that human beings are very good at perceiving structure in visual forms. The basic idea is to present the data in some visual form, allowing the human to gain insight from the data, draw conclusions, and directly interact with the data.

Classification Based on the Machine Learning

- Binary Classification.
- Multiclass Classification.
- Multi-Label Classification.
- Imbalanced Classification.

Classification Based on the Machine Learning

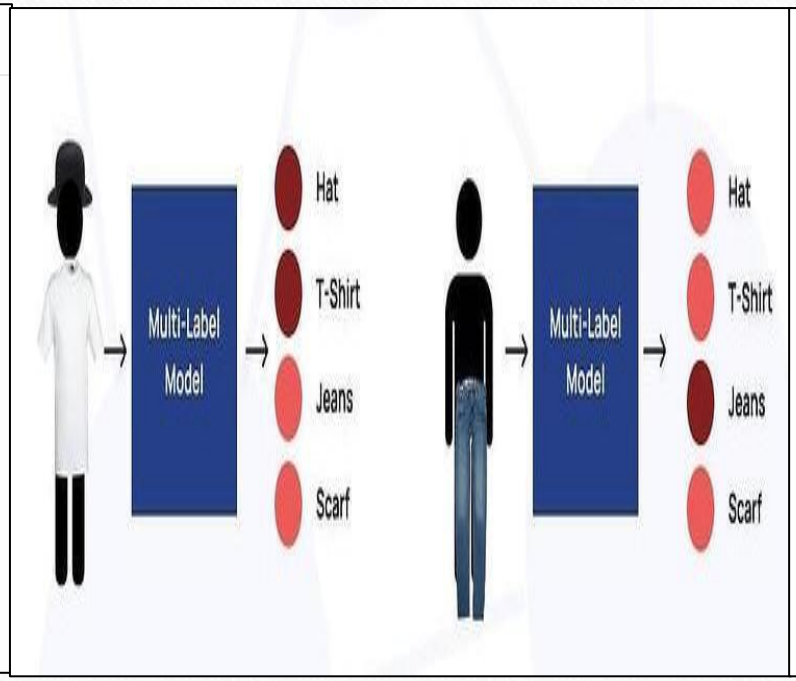
Three Type of Classification Tasks

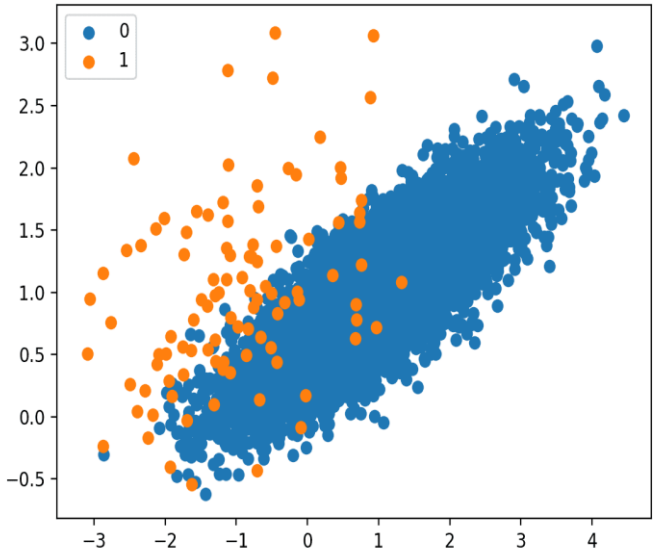
Binary Classification

- Spam
- Not spam

Multiclass Classification

- Dog
- Cat
- Horse
- Fish
- Bird
- ...

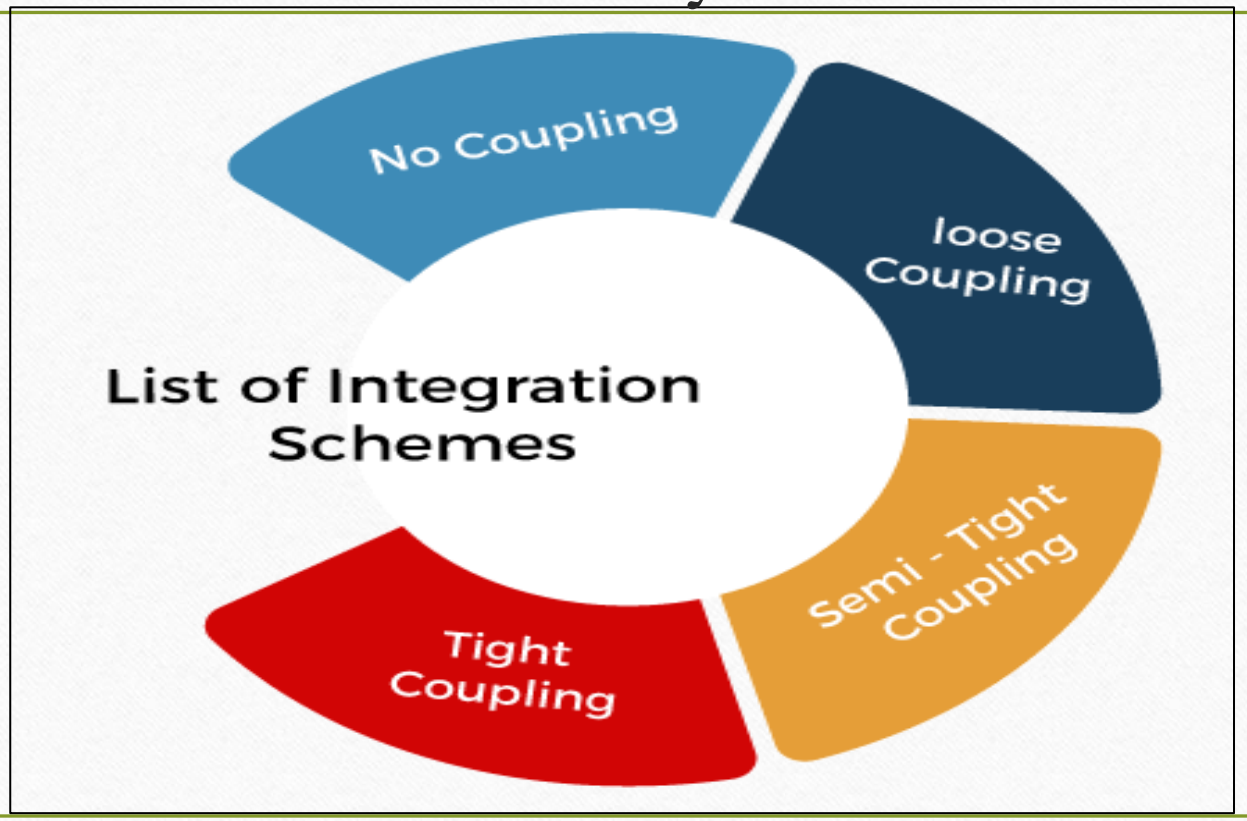




Classification Based on the Statistics

- Nominal
- Odinal
- Continious
- Discrete

Integration schemes of Database and Data warehouse systems



- **No Coupling**

In no coupling schema, the data mining system does not use any database or data warehouse system functions.

- **Loose Coupling**

In loose coupling, data mining utilizes some of the database or data warehouse system functionalities. It mainly fetches the data from the data repository managed by these systems and then performs data mining. The results are kept either in the file or any designated place in the database or data warehouse.

- **Semi-Tight Coupling**

In semi-tight coupling, data mining is linked to either the DB or DW system and provides an efficient implementation of data mining primitives within the database.

- **Tight Coupling**

A data mining system can be effortlessly combined with a database or data warehouse system in tight coupling.

Some Useful Links

- https://machinelearningmastery.com/types-of-classification-in-machine-learning/
- https://www.javatpoint.com/classification-of-data-mining-systems
- https://www.tutorialspoint.com/data_mining/dm_tasks.htm
- https://www.tutorialspoint.com/data_mining/dm_tasks.htm
- https://www.geeksforgeeks.org/classification-of-data-mining-systems/
- https://www.brainkart.com/article/Classification-of-Data-Mining-Systems_8309/
- https://www.analyticsvidhya.com/blog/2021/07/demystifying-the-difference-between-multi-class-and-multi-label-classification-problem-statements-in-deep-learning/

Thank you

Adapt it with your needs and it will capture all the audience attention.