Empleo, Jericho B.

BSIT 3-5

Elective 2 – Data Mining

Read the statement/s carefully. Identify and answer the given questions

1. Explain the fundamental concept of data mining and why do we need to mine data?

    - The fundamental concept of data mining is the process of discovering patterns from a large amount of data using machine learning. The goal of data mining is to get information derive from data. We need to mine data because it can help us to analyze raw data into valuable information that can be used in decision making.

2. Discuss how data mining became part of the natural evolution of information technology

    - Data mining became part of the natural evolution of information technology because of the increasing amount data from the internet where the data has been generated where it is now difficult to analyze. Data Mining provides a way to extract knowledge from the availability of data. Using statistical and machine learning algorithms to discover patterns for decision making. It is a computing technology where big data technologies like spark and NoSQL databases are used for data mining.

3. Explain the importance of data mining and how it powers business values.

    - The importance of data mining and how it powers business values provides the business industry insights that are useful by having a trend and relationship that is powerful for making a decision that drives the business growth. Data mining helps the business value by having a better customer experience, increased efficiency, improved product development and more. As data continues to grow data mining is an important asset to the company.

4. In determining some major issues in data mining, choose in which part do you think it needs attention. How can they be addressed?

- Data mining privacy concerns are related to the usage and potential misuse of personal information. The likelihood of violating people's privacy rights rises as data mining tools get more sophisticated. This is particularly valid when dealing with private data like financial or health records. Organizations may solve these issues and guarantee that they are employing data mining tools in an ethical and responsible manner by using data anonymization, data sharing agreements, data access controls, and legal compliance.

5. Explain why it makes sense to model the data warehouse in a multidimensional data structure called data cube.

- A data cube is a three-dimensional or higher-dimensional array that stores data in cells, with each cell representing a unique combination of dimensions. The benefits of using a data cube include flexibility, performance, improved decision-making, and scalability.

6. How can you deal with different kinds of data to easily choose correct tools and algorithms in mining data?

- Understanding the different forms of data, preparing it, selecting the right tools and algorithms, analyzing the outcomes, and fine-tuning the process are all necessary when working with diverse types of data. These procedures will help you select the best tools and algorithms for data mining and produce insightful results.

7. Enumerate the goals of data mining and its application and determine in which area can we still apply those goals in our generation now. (Should not be existing)

- The goals of data mining and its application is to create a pattern with enormous number of datasets by using methods such as statistical and computational. In our generation, we still can apply our goals like data summarization where we can reduce large number of datasets into smaller to manage to have a better analysis. Also, clustering of data since there are different data we can specify and separate the similarities and differences of each.

8. Search at least two (2) frameworks can you use to effectively mine different data sets.

 Keras: A library network for data mining. It is an API for deep learning models. Where it can be used in different types of data like images, text, and more.

TensorFlow: Data mining applications frequently employ the open-source machine learning framework TensorFlow. A complete collection of tools, including deep neural networks, are provided for creating and refining machine learning models. Text, pictures, and time-series data are just a few of the many forms of data that TensorFlow is made to handle.