A big data mining system is a software-based tool created to draw important conclusions and information from vast and intricate databases. In order to find patterns, trends, and relationships that may be utilized to improve business choices, streamline operations, and gain a competitive edge, modern data mining techniques and algorithms are employed to analyse massive amounts of organised, semi-structured, and unstructured data. The system may be implemented on-premises or in the cloud and typically includes tools for data gathering, preparation, analysis, and visualisation. Industries like banking, healthcare, retail, and telecommunications frequently use big data mining technologies to glean insightful information from their enormous datasets.

Summary of Lessons:

1. Three Major Characteristics of Big Data: Volume, Velocity, Variety. The volume of data generated by modern systems is immense and continues to rise rapidly, needing sophisticated tools and procedures to manage and analyze it. Big data is produced and processed quickly, frequently in real-time or very close to real-time. Structured, semi-structured, and unstructured data are only a few of the various types of big data.

2. Sources of Big Data: The three main categories of big data sources are social (human), machine (sensor), and transactional.

3. Database Workload: A database workload is a group of database operations or queries that a system or application does continuously. Examining a workload's properties, such as its frequency, duration, and resource usage patterns, is known as a workload analysis. By identifying performance bottlenecks and capacity restrictions, this research helps guide database tuning and optimization efforts.

4. Linear Regression Model: A linear regression model describes the relationship between a dependent variable, y, and one or more independent variables, X.

5. Three Phase Clustering Method: The Three Phase Clustering Method is a method for clustering related data points into clusters or groups, a step in the data analysis process. There are three main phases in the Three Phase Clustering Process. They are Preprocessing Phase, Similarity Calculation Phase, Clustering Phase.

6. Text Mining: The extraction of pertinent and meaningful information from text data is known as text mining. It involves analyzing unstructured text data, including emails, social media posts, news articles, and customer reviews, using natural language processing (NLP), machine learning, and data mining techniques.

7. Time Series Clustering: Time Series Clustering is a technique used to group similar time series data together based on their patterns or characteristics. Time series data is a sequence of observations taken at regular intervals over time, such as stock prices, weather measurements, or sensor readings.

8. Big Data Reduction Framework: The Big Data Reduction Framework is a collection of strategies and procedures for reducing the volume and complexity of big data while preserving crucial data. This framework is employed to handle the difficulties brought on by the enormous amount, speed, and variety of big data.

9. Multiple Classifier System: A Multiple Classifier System (MCS) is a machine learning technique that enhances the accuracy and resilience of a classification model by combining the predictions of various classifiers. An MCS involves the training of multiple distinct classifiers using the same data but varying feature sets, classification algorithms, or parameter values.

Nasir Uddin Ahmed
Student Id: S2015449

10. Deep Stacked Autoencoder Algorithms: A sort of deep learning architecture known as deep stacked autoencoder algorithms employs a number of stacked autoencoders to build a hierarchical representation of the input. A neural network that has been trained to rebuild its input data from a compressed representation of the data learned in a bottleneck layer is called an autoencoder. neural network.

11. Model Driven Deep Deterministic Learning: In order to understand a system's dynamics and forecast future outcomes, a machine learning technique called model-driven deep deterministic learning combines deep learning with dynamic system modeling. In this method, a deep neural network is trained using a model of the dynamic system to forecast the system's future state.