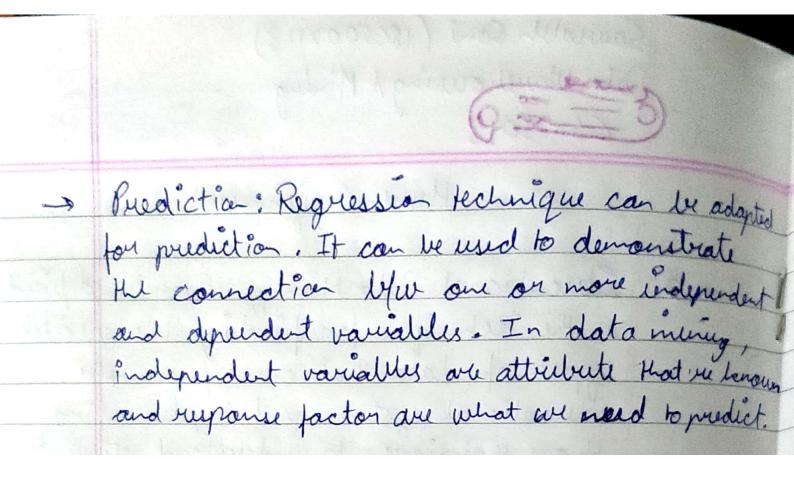Saurabh Soni (18CS002113)
Data Warehousing & Mining
CS-3013

## Home-Assignment

**Ans. 2** Educational Data Mining can be defined as the technique for finding the specific types of data that come from the educational system and implementing those techniques to understand students and the system better.

Approaches of Data Mining in educational data:

→ Clustering: It refers to the process of identification and classification of objects into different groups, the segmentation of a data set into clusters so that the data in each subset share some common characteristics of similar classes of objects.

→ Classification: It refers to process of describing data relationships & express values for future observation. There're various types of classification, namely, Decision Tree-Based Methods, Memory-Based Reasoning, Rule-Based Methods, Naive Bayes and Bayesian Belief Networks, Neural Network.

→ Prediction: Regression technique can be adopted for prediction. It can be used to demonstrate the connection b/w one or more independent and dependent variables. In data mining, independent variables are attribute that are known and response factor are what we need to predict.

Ans.2 Data Mining is the practice of automatically searching large stores of data to discover patterns and trends that go beyond simple analysis.

Types of

i) Pictorial Data Mining :- It's based on analysis of multidimensional relations through transforming the initial data into artificial pictures.

ii) Text Mining :- It's an AI technology that uses NLP to transform free (unstructured) text in documents & database into normalized, structured data suitable for analysis or to drive machine learning algorithm.

iii) Social Media Mining :- It's process of obtaining big data from user generated content on social media sites & mobile apps in order to extract pattern, from conclusions about users, & act on info.

iv) Web Mining :-In this we extract info. from Web documents & services.

v) Audio Mining :-Audio Mining embraces speech recognition, keyword, phonetic or transcription technologies to extract insights from pre-recorded voice streams.