

Identify a problem from your own experience that you think would be amenable to data mining. Describe:

1. What the data is.

The data will be about the students and the topics they can answer in the examination. If each question came with a different topic then we can easily say that if the student read about the topic then he will answer the question. These stats will help the student and as well as the mentor to keep track of the student and help him with the topics he didn't understand.

2. What type of benefit you might hope to get from data mining.

With the help of data cleaning and data transformation steps, we can clean the data and make datasets ready for analysis. Later we can use plots to get better knowledge and patterns on the student's subjects which can affect the decision making and operations.

3. What type of data mining (classification, clustering, etc.) you think would be relevant

Clustering would be a kind type since it divides the data into similar patterns. If similar students have a problem with the same topic, then they will be present in a cluster and they can be taught together with the topics by their mentor.

4. Name one type of data mining that you think would not be relevant and describe briefly why not

Regression since it is numerical and usually used to measure future data and make certain predictions. It will not be helpful to the data we have.