Data Mining Assignment 1

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

1. What the data is.

In YouTube, when watching videos regarding tech news regarding mobiles, in recommendations it displays some irrelevant video which aren't of my interests.

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

I think it should provide all sorts of information regarding user response and it should also maintain some privacy of its user information.

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

I think Clustering would be relevant because it makes useful cluster of objects which have similar characteristics. So, when I watch videos of my interests it shows only relevant and similar characteristics of videos.

4. Name one type of data mining that you think would not be relevant, and describe briefly why not.  
For each, illustrate with an example, e.g., if you think clustering is relevant, describe what you think a likely cluster might contain and what the real-world meaning would be.

Regression helps to deal with the numerical data, in real world experience for my example mentioned above it doesn’t need to measure any data regarding my interests**.**

Write one to two pages of 11 point single-spaced typeset text - you aren't writing a paper, but it isn't short answer either.

Most of the time gathering of information about certain elements should depend on user interests for recommending similar characteristics. In YouTube, when I was watching videos like tech news, in recommendations it displays some irrelevant video which aren't of my interests and with irrelevant information. It is a known fact that data mining collects information about people using some market-based techniques and information technology. It should not violate the privacy of the user data so that misuse of information through the mining process is reduced. In my point of view Clustering would be relevant because it makes useful cluster of objects which have similar characteristics.