

Data Mining Assignment 1

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

Covid-19

1. What the data is?

Number of persons effected, number of persons having symptoms, number of persons may affect, number of persons recovered, and number of persons died.

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

Knowledge about cure, predictions based on the different kinds of people which is vectorized accordingly and based on their health issues.

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

Regression which deals with numerical data i.e., recovery rate, death rate, total cases in different locations and prediction about the number of persons are going to effect.

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

Classification cannot be applied because there will be same treatment for different persons with different symptoms and also, the time lapse for the symptoms will be different so we cannot do grouping.

5. 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.

The present situation is totally based on Covid-19 pandemic, for this situation the data will be number of persons effected, no of persons having symptoms, number of persons may affect, number of persons recovered, and number of persons died. By learning data mining concepts people got the knowledge about cure, predictions based on the different vectored people and their health issues. For this situation, according to the datamining techniques we can choose regression to get the data of recovery rate, death rate, total cases in different locations etc. and can predict about the number of persons going to be effect. In my view the classification cannot be applied because there will be same treatment for different persons with different symptoms. In addition to that the time lapse is also considered for the symptoms to be shown will be different for this reason we cannot do grouping.