

*Datawarehousing and Analytics*  
*WS 15/16*

*Assignment-1*

*Team 18*

*Krishna Damarla (2988256 )*

*Mathumathi Anbalagan (2991285)*

*Shalini Chellathurai( 2927550)*

## Task1

- c) The MSHA offers various services to former patients. Hence, it should be easy to determine which of these services should be offered to which patient or group of patients. This suggestion should be based on the services offered to patients in the past.

- *a justification for the data mining technique you would use*

In this case Prediction can be used. To determine with services should be offered to which group of patients Regression approach can be used.

- *List of attributes the data mining algorithm would need and the semantics of these attributes*

Patient\_ID : To know the drugs prescribed to the patients, the patient ID should be known.

Doctor\_ID : To know the number of Doctors giving treatment to the patients

Procedures : To know about the treatment of patient with time, duration of stay.

These values are obtained from FCWH

Schedule : To know about the schedule of treatments obtained from MSIC

Treatments : To know about protocols and results of MRI scan, CT scan, blood counts, etc.

- *a description of the patterns that could be derived*

Here response variable is taken as one attribute, 'y' is modeled as a linear function of other attributes i.e., predictor variables  $x_1, \dots, x_n$ .

$$y = a_0 + a_1x_1 + \dots + a_nx_n$$

- *examples how these patterns could be used to address the mentioned issue and for future decisions.*

In this case the set of services which should be offered is taken as response variable and the patient or group of patients are taken as predictor variable.