Self-evaluation report

1.Ranjan Gsk 2..Kailash Kolluru

Bits Pilani, Hyderabad, India

1.f20160245@hyderabad.bits-pilani.ac.in

2.f20160210@hyderabad.bits-pilani.ac.in

Team achievements:

Developed the skill to understand what all information can be understood from the given dataset. The pre-processing techniques required before starting working on the dataset. Different visualizing skills to understand the nature of the data. The dataset chosen by our team tells the number of dug wells which have been closed and the effect they caused on irrigation. We could cluster the results of all such states for each of the factor. This result will help us understand the clusters of states which are least and highest effected for each factor. This information can be used to develop the facilities in each state. The clustering has been done with the help of two clustering algorithms K-Means and DBSCAN.

Improvements which can be made:

Couldn't find any correlation among the factors given. Finding the correlation among any of the factors might have helped us to understand certain patterns which will be common to different states. Couldn't find any association rule with any other dataset. The confidence for the rules we tried were not significant. To reduce the sparsity of the dataset we had to group all the dataset so that we get a combined dataset which will be for each state. This will reduce the precision of the results.

Individual team contributions:

Since the team had only two members almost all the work has been done in coordination.