Team ID: 39
Members -

1) Sundeep Ammisetti – F2017A7PS1218H
2) Pranav Saikiran Sista – F2017A7PS1225H
3) Satish Kumpatla – F2017AAPS0284H
4) Nakul Gupta – F2017A7PS0211H

Contributions -

Pranav -

- FuzzyC-means on data set for different attributes in the dataset
- Hopkin's statistic.
- Completed the task of filling in NaN values in the dataset.
- Sampling
- Cleaning
- Outlier Analysis
- Helped to make the final Presentation.
- Helped in writing a report.

Nakul:-

- Sampling
- Cleaning
- Help with DBSCAN clustering on-network dataset for a specific month, year, service provider some other attribute
- Evaluate the appropriate DBSCAN parameter(eps) for better visualization
- ELBOW method for cluster stability and its validation
- Help in preparation of PROJECT Report
- Evaluated the clustering techniques to pick the better one (distance matrix) and chose appropriate DBSCAN parameters
- Merged the datasets corresponding to various states

Sundeep -

- Data Cleaning
- Binning(Numerosity Reduction) of Data set
- Aggregation
- Merged the datasets of different months and years
- Stratified Sampling
- Outlier Analysis using Z-Score for outlier detection
- Helped to make final Presentation

Satish:-

- Performed DBSCAN Algorithm on Sampled cleaned and normalized dataset.
- Visualized the results of the DBSCAN Algorithm for unique sub-datasets of the dataset.
- Visualized the dataset after pre-processing to rectify mistakes notified by Mam.
- Helped in writing the report and Presentation
- Helped in finding parameters for DBSCAN i.e. epsilon and min points.
- Hopkin's statistic.
- Stratified sampling of the dataset with 5% of unique data points.
- Cleaning of the dataset by taking the average of all the samples in unique points by filling NAN values.
- Merging datasets of different months and years to create new features.
- Visualized the plots after sampling, cleaning, and merged the dataset.

Achievements/Reflections -
Our main goal is to find good relationships that were hidden in the dataset. Most of the analysis gave the existing relationship between Network providers and network technology and hope the other inferences made could be of some help to people.We found how data speeds and signal strength vary among different states in different months and in different years, this could help companies where to improve their service.This project also helped us improve our Python programming skills.