To,

IITD-AIA Foundation of Smart Manufacturing

Subject: Weekly Progress Report for Week 0

Dear sir, Following is the required progress report to the best of my knowledge considering relevant topics to be covered.

What's happening this week:

- Pre-processing
- Supervised Learning, Unsupervised Learning
- Naïve Bayes, K-Nearest Neighbor
- Linear Regression
- K-Means, PCA

(Level 2 got extended for me because of end-term examination and hence I was completing my Assignments. However, learning during that time was great for the project also.)

Weekly Progress:

June 01:

Started with Supervised Assignments.

1. Initiate Pre-Processing, dataset provided was processed in order to use it for training the model.
2. Learnt to use dictionary for mapping the string data with int data type. like 0 is mapped with on category X, category Y is mapped with 1.
3. Learnt how to divide the dataset to map according to ranges using cut, bins, like fare and age is categorized according to range.
4. done with processing of the data for training the models

June 02:

1. Went through KNN algorithm, basics behind the algorithm
2. Applied the algorithm using numpy and pandas.
3. Learnt various function of numpy and pandas like argmax, argsort
4. Data visualization:
    1. Error Rate
    2. Mapping the column data wrt answer to be predicted

June 03:

1. Introduced with new algorithm of the assignment Naïve Bayes, basic implementation
2. Applied the algorithm and learnt some other direct functions like exp, sqrt
3. Learnt seaborn libraries and applied heatmap in order to find correlation between the features and extract out most affecting features wrt charges.
4. Increased the efficiency of models using most correlated features by further processing the data and choosing limited features.

June 04:

1. Learnt Regression, Unsupervised Learning about K-means, PCA
2. Did pre-processing (this process become easy as learnt basic mapping, replacement on first day), applied Linear Regression on the same.
3. Applied pre-processing on Iris dataset, not much processing was required on that dataset Just processed categorical data.

    Next Day, I applied K-means and PCA.