**Progress Report**

1. **Principal Investigator**

Akhil Reddy Buchala - [buchalaa1@newpaltz.edu](mailto:buchalaa1@newpaltz.edu)

Rohan Reddy Yeruva - [yeruvar1@newpaltz.edu](mailto:yeruvar1@newpaltz.edu)

Chandra Sekhar Pokuri - [pokuric1@newpaltz.edu](mailto:pokuric1@newpaltz.edu)

1. **Title of Project**

Customer Churn Prediction with AWS EMR

1. **Project Progress Summary**

We have established the problem to be dealt with, Customer Churn prediction. The data is ready to be analysed. We have created the AWS EMR cluster. Created the notebook in the cluster.

1. **Activities**

Meeting 1: Deciding about the problem statement

Meeting 2: Selecting all the references to be used for the project

Meeting 3: Finding the appropriate data

Meeting 4: Creating EMR cluster and Notebook

1. **Remaining Work**

Importing the data from S3 bucket to EMR notebook.

Data pre-processing: Data cleaning, Data preparing

ML models: Creating various models like Logistic Regression, Random Forest, etc.

Comparing the results

1. **Difficulties**

Explain what kind of difficulties you have encountered and what solutions used for these difficulties.

1. **Significant Changes to Proposal (if any)**

If you have made a significant changes to the submitted proposal.