# Capstone Project Submission

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| 1. **Aadarsh Pandey**   E-mail- [aadarshpandey086@gmail.com](mailto:aadarshpandey086@gmail.com)   * Logistic Regression * Data Visualization * Framework of Project * Model Building * Technical Documentation * PPT Presentation  1. **Ankita Hanamshet**   E-mail- [ankii7546@gmail.com](mailto:ankii7546@gmail.com)   * Framework of project * Data Analysis * Data Visualization * Model Building * Technical Documentation * PPT Presentation  1. **Darpan Agrawal**   E-mail- [agarwaldarpan421@gmail.com](mailto:agarwaldarpan421@gmail.com)   * Framework of project * Data Analysis * Feature Engineering * Model Building * Technical Documentation * PPT Presentation  1. **Vandana Pattnaik**   Email- [vandanapattnaik2588@gmail.com](mailto:vandanapattnaik2588@gmail.com)   * Approach towards plan * Data Analysis * Data Visualization * Model Building * Technical Documentation * PPT Presentation  1. **Vinay Kulkarni**   E-mail- [kulkarnivinayv@gmail.com](mailto:kulkarnivinayv@gmail.com)   * Approach towards plan * Data Analysis * Feature Engineering * Technical Documentation * PPT Presentation |
| **Please paste the GitHub Repo link.** |
| GitHub link :- https://github.com/AadarshPandey29/Credit-Card-Default-Prediction |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| It was quite an interesting project which helped us to polish our skill. In this project we didn’t have to handle missing values since we weren’t having any noise in the data. Having no noise made our lot of work simple, but we still faced some problem in finding the collinearity and it was also a bit tough for getting our best model after apply multiple models to it and faced lot of error in that.  Our 1st approach was to get the proper and useful information from the data and then checking for the null and duplicate values or in short checking for the noise in the data. We also did the feature engineering which is separating the numerical feature and categorical feature.  2nd Approach was to perform EDA on the dataset and get useful insights from it and it was pretty helpful to us.  3rd Approach was to train and test our data and also to build model and we tried many different models like Logistic Regression, Random Forest, Xg boost and At the last we have the ROC and AUC Curve on all Models we get a lot of useful insights from that. |