|  |
| --- |
| **INTERNSHIP RESEARCH WORK**    Project Topic: Predicting Default Risk of Lending Club Loans    Introduction  Using the dataset provided by Lending club loan we will maximize expected returns from Lending Club Investment by predict the default risk of loans. We will evaluate different machine learning models for better accuracy and prediction. Finally, we will deploy it.    Dataset  We will be using Lending club loans dataset from year 2007-2018 for our dataset.  LINK: <https://www.kaggle.com/wordsforthewise/lending-club/download>    Further Pipeline   * We will be using different **EDA** techniques to visualize the data and gather information from dataset. * Further, we will be using different **pre-processing** techniques to clean and transform the data. * Then we use the data to **train** the different models using different algorithms for classification. * **Evaluation phase** where we check different metrics and confusion matrix * We will be **tuning parameters** for further increasing the accuracy, F1 Score and handling if model is overfitting. * Then, we will be **predicting** using the best model for further insight gathering. * Then we will be using the best model for **deploying** the model on a cloud platform like *Heroku* using *Flask*. |