**CS 513 A/B – KDD PROJECT PROPOSAL: (PROJECT NAME)**

(Sample) CS 513 A - KDD Project Proposal: Loan Default Prediction

**Project Group No: (As per canvas group allotment)**

Problem Statement: (Description of the problem statement)

Loan defaulting occurs when a borrower fails to pay back a debt according to the initial arrangements. This can occur for banks as well as for certain peer-to-peer lending clubs. The concept of a lending club enables borrowers to obtain a loan and investors to purchase a loan based on the notes provided. During such instances, there is always a high risk of borrowers defaulting their loans. Therefore, we want to predict using his previous finances if a certain borrower is a defaulter or not. Given, a set of features **we want to predict the target variable as, 1 which is defaulter or 0 which is non-defaulter**.

Dataset: (Description)

The dataset comprises of 34 features in form of columns, out of which we may opt to use the essential features only, during implementation. We can use PCA and correlation for feature reduction.

Source of Dataset: (Link to dataset)

Implementation Strategy and algorithms used: (List different models)

We have decided to implement and compare 8 different models among four different group members. We have chosen a few models from our course and few from outside the course. The following are the models selected by us:

1. Decision Trees
2. AdaBoost Classifier
3. Multilayer Perceptron
4. Bagging along with Random Forest
5. K-nearest neighbor with Grid Search CV
6. Logistic Regression with Grid Search CV
7. Random Forest with Randomized Search CV
8. Support Vector Machine with Grid Search CV

Model metrics and Evaluation: (List evaluation metrics like: AUC-ROC, Confusion matrix , F1, Recall, Precision)

Evaluation of different models used in project.

Team Members: Group 1

1. Member 1 Name
2. Member 2 Name
3. Member 3 Name
4. Member 4 Name