Professor Mingdi Xin

BANA 273 Machine Learning for Analytics

November 9, 2021

**Team 23:**

Kate Zagrebneva(ezagrebn@uci.edu)

Prashasti Sharma (prashass@uci.edu)

My Han Mai (han.mai@uci.edu)

Noor Zia (nzia1@uci.edu)

Samira Shafiei (shafieis@uci.edu)

Machine Learning Final Project Proposal

Loan Approval Predictions

Our team wants to use machine learning techniques to apply to loan data in order to predict the likelihood that a loan will be approved. We plan to use a dataset from Kaggle called “Loan Eligible Dataset”. The attributes include gender, education, marital status, employment, income, and more. We plan to use these attributes to predict whether the loan will be approved using Naiive Bayes and Decision Trees.

The dataset has 981 total rows and is split into train and test datasets. There are a total of 11 possible variables that can affect loan approval and that we are able to test. The variables are categorical and continuous.

Dataset: <https://www.kaggle.com/vikasukani/loan-eligible-dataset>

Challenge: Can we predict whether the loan will be approved based on the given attributes?

Variables:

| Loan ID | Applicant Income |
| --- | --- |
| Gender (F/M) | Co-Applicant Income |
| Married (Y/N) | Loan Amount |
| Dependents (0,1,2,..) | Loan Amount - Term (12- 480) |
| Education (Graduate/Not Graduate) | Credit History (Past loan history: 0,1,..) |
| Self-Employed (Y/N) | Property Area (Urban/Semi Urban) |