Neural Networks (Classification Problem)

Course-End Project Problem Statement

Course-End Project: House Loan Data Analysis

Problem Statement:

For safe and secure lending experience, it's important to analyze the past data. In this project, you have to build a deep learning model to predict the chance of default for future loans using the historical data. As you will see, this dataset is highly imbalanced and includes a lot of features that make this problem more challenging.

Objective:

Create a model that predicts whether or not an applicant will be able to repay a loan using historical data.

Domain: Finance

Analysis to be done: Perform data preprocessing and build a deep learning prediction model.

Steps to complete:

1. Perform data quality check by checking for missing values if any
2. Balance and then split the data into a test-train split
3. Understand what factors contributed most to employee turnover by EDA
4. Determine whether grouping features together into a single feature helps Specifically, look at time spent at company and number of projects (which both correlate with tenure, but together can capture a notion of “engagement”)
5. Train classification models (SVMs, Decision Trees, Random Forest, and an NN with 2 hidden layers, each with 12 nodes per layer) using k-folds. And evaluate which model is the best performer
6. Train the best model using all available data in the train set
7. Test the trained model against the test set and output accuracy, precision, recall, and f1 score