

# Machine Learning (UG): CSE343/ECE343

## Project Proposal

### Automated Financial Distress Predictor

Ashwin Sheoran: 2020288

Harsh Goyal: 2020562

Shivam Jindal: 2020125

Sourav Goyal: 2020341

## 1. Motivation

Banks play a crucial role in market economies. They decide who can get finance and on what terms and can make or break investment decisions. For markets and society to function, individuals and companies need access to credit.

Credit scoring algorithms, which guess the probability of default, are the methods banks use to determine whether or not a loan should be granted. This model helps to improve state of the art in credit scoring by predicting that someone will experience financial distress in the next two years.

## 2. Related Work

Credit Scorer is a widely addressed topic:

1. Automated Credit Scoring System for Financial Services in Developing Countries: How Automated Credit Scoring System is widely used in Developing countries

([https://www.researchgate.net/publication/328559774\\_Automated\\_Credit\\_Scoring\\_System\\_for\\_Financial\\_Services\\_in\\_Developing\\_Countries](https://www.researchgate.net/publication/328559774_Automated_Credit_Scoring_System_for_Financial_Services_in_Developing_Countries))

2. Financial Distress Prediction based on Multi-Layer Perceptron with Parameter Optimization: Pre-Bankruptcy period prediction and classification

([http://www.iaeng.org/IJCS/issues\\_v48/issue\\_3/IJCS\\_48\\_3\\_41.pdf](http://www.iaeng.org/IJCS/issues_v48/issue_3/IJCS_48_3_41.pdf))

3. Automated credit scoring process: How does it work? by Jason Bingham: Using automation to find credit score rather than just feeding values in a static function

(<https://lucidux.com/automated-credit-scoring-process-how-does-it-work/>)

## 3. Timeline

A Tentative 12-week timeline:

**Week 1-2:** Data Collection (including Scraping)

**Week 3:** Pre-processing and Data Visualization.

**Week 4:** Feature Extraction.

**Week 5:** Feature Analysis, Selection, Correlation, HeatMaps.

**Week 6:** Logistic Regression, Support Vector Machines.

**Week 7:** Decision Trees, Random Forest.

**Week 8-9:** Analysis and performance of models.

**Week 10:** Hyperparameter Tuning, Check for model Overfitting and Underfitting.

**Week 11:** Report Writing.

**Week 12:** Buffer

## 4. Individual Tasks

Tasks	Team Member /s
Data Collection	Ashwin, Harsh, Shivam, and Sourav
Pre-Processing and Data Visualization	Ashwin, Harsh, Shivam, and Sourav
Feature Extraction	Ashwin, Harsh, Shivam, and Sourav
Analysis of Features (EDA etc.)	Ashwin, Harsh, Shivam, and Sourav
Logistic Regression and Support Vector Machines	Ashwin and Sourav
Decision Trees, Random Forest	Harsh and Shivam
Analysis and Performance of models	Shivam and Sourav
Hyperparameter Tuning Check for Overfitting and underfitting models and selection of the best model	Ashwin and Harsh
Report Writing	Ashwin, Harsh, Shivam and Sourav

## 5. Final Outcome

The objective is to create a model which predicts, based on the probability, whether someone will experience financial distress in the next two years.

This will help in improving the credit scoring algorithms used by various banks, which leads to sustainable economic growth.

The project will also help us to identify various parameters essential for credit lending and which parameter is more important than others for checking eligibility.