

CSP 571 – Data Preparation and Analysis

Spring 2023

Project – Plan and Detail

Topic: Predicting and analyzing credit card frauds

Project Group:

Pradyothan Govrineni (Leader)

Sohaib Syed

Uttam Kotadiya

Harsh Gordhan Dungrani

	Task Name	Due Date	Assigned To
Data Gathering	Find available datasets	2/5/2023	All
	Finalize the datasets	2/10/2023	All
Data Preparation/ Cleaning/ Processing	Import datasets and explore the size, variable types, structure, summary of the datasets	2/24/2023	Uttam, Harsh
	Handling variable types	3/20/2023	Harsh, Sohaib
	Handling variables with Null values	3/20/2023	All
	Feature Engineering: Variable Importance based on fitting decision tree models etc	4/4/2023	Pradyo, Uttam
	Feature Engineering: Boruta	4/4/2023	Harsh, Sohaib
	Performing sampling techniques to tackle class imbalance	4/9/2023	All
	Define final dataframe	4/11/2023	
Data Exploration/ Visualization	EDA plots of features in original dataframe	4/4/2023	Uttam, Pradyo
	Correlation maps to understand the collinearity between the different variables of data	2/24/2023	Sohaib
	Histogram plots for variables with a certain amount of null values	3/20/2023	Pradyo, Sohaib
	Plots of decision trees, feature importance plots	4/4/2023	Harsh, Sohaib
	EDA plots of features in final dataframe	4/13/2023	Harsh, Sohaib
	ROC Curve, Confusion matrices of fitted models	4/15/2023	All
Model Evaluation	Fit different ML models on preprocessed data	3/28/2023	Harsh, Pradyo
	Fit different ML models on processed data	4/13/2023	All
	Evaluate fitted models on performance metrics	4/17/2023	All
	Select the best model and tune it further if possible	4/20/2023	All
Project Submissions	Final Project Report	4/26/2023	All
	Final Project Pre-recorded presentation	4/28/2023	All