**Problem Statement**

**About Company**: Sunrise Financial provides customers with different types of loans. Customers first apply for the loan and then the company validates for the eligibility of the loan.

**Problem:** Sunrise Financial wants to explore various features of borrowers like LoanId, CustomerId, current loan amount, Loan term, credit score, annual income, years in current job,

homeownership, the purpose of the loan, Monthly debt, years of credit history, months since the last delinquent, number of open accounts, number of credit problems, current credit balance,

maximum open credit, bankruptcies, and Tax liens and forecast if the loans will be paid or not.

**Data Set**

| Columns | Definition | Example data |
| --- | --- | --- |
| Loan ID | Loan id | f738779f-c726-40dc-92cf-689d73af533d |
| Customer ID | Customer id | ded0b3c3-6bf4-4091-8726-47039f2c1b90 |
| Loan Status | Status of Loan | Fully Paid |
| Current Loan Amount | Current Loan Amount | 611314 |
| Term | Short/long Loan term | Short Term |
| Credit Score | Credit score | 747 |
| Annual Income | Annual Income | 2074116 |
| Years in current job | Years in current job | 10+ years |
| Home Ownership | Home Morgage/ rent/ own | Home Mortgage |
| Purpose | Purpose of taking a loan | Debt Consolidation |
| Monthly Debt | Monthly Debt | 42000.83 |
| Years of Credit History | Years of Credit History | 21.8 |
| Months since the last delinquent | A delinquent mortgage is a home loan where the borrower is late in one or more required payments | NA |
| Number of Open Accounts | Number of Open Accounts | 9 |
| Number of Credit Problems | Number of Credit Problems | 0 |
| Current Credit Balance | Current Credit Balance | 621908 |
| Maximum Open Credit | Maximum Open Credit | 1058970 |
| Bankruptcies | Bankruptcies | 0 |
| Tax Liens | A tax lien is a claim the government makes on a property when the owner fails to pay the property taxes. · | 0 |

**Source of Dataset:** <https://www.kaggle.com/zaurbegiev/my-dataset/activity>

For this problem, we have two files: train and test.

* The train file will be used for training the model. It contains all the independent variables and the target variable (Loan Status).
* The test file contains all the independent variables, but not the target variable. We will apply the model to predict the target variable for the test data.

**Stakeholders:** Sunrise Financials,Regulatory bodies, policymakers, Investors.

The stakeholders want the loan paid.

**Data Science approach being used:** Supervised - Classification. Supervised since we want to build a statistical model for predicting if the loan will be paid based on the dataset values. Classification, since we want a qualitative response from the model.

**Use of my results to Client for Loan Forecasting:** Logistic Regression (Model not finalized. Still studying options.)

**Deliverables:** The following are the list of required deliverables:

* All Jupyter notebooks developed as part of this proposal typically, one per major project phase (e.g., data wrangling, EDA, etc.)
* A written final report.
* A presentation slide deck.