

Case Study: Predicting Company Bankruptcy Using Financial Data

Objective: To perform an in-depth analysis of a financial dataset to predict the likelihood of a company going bankrupt. The analysis involves data preprocessing, exploratory data analysis (EDA), hypothesis testing, feature engineering and selection, and applying machine learning techniques for classification.

Tasks:

1. **Data Understanding and Preprocessing:**
 - Load and inspect the dataset.
 - Handle missing values appropriately.
 - Detect and handle outliers.
2. **Exploratory Data Analysis (EDA):**
 - Generate descriptive statistics.
 - Visualize feature distributions and target variable.
 - Analyze relationships between features and the target variable.
3. **Hypothesis Testing:**
 - Perform hypothesis testing to identify significant features influencing bankruptcy.
4. **Feature Engineering and Selection:**
 - Create new features to enhance predictive power.
 - Use dimensionality reduction techniques if necessary.
 - Select relevant features using methods like RFE or feature importance.
5. **Modeling:**
 - Split the data into training and testing sets.
 - Apply Logistic Regression for classification.
 - Evaluate model performance using appropriate metrics.
6. **Model Interpretation and Insights:**
 - Interpret model coefficients to understand feature impact.
 - Summarize key insights and discuss implications for financial risk management.
 - Provide actionable recommendations.

Deliverables:

- Detailed report of the analysis.
- Python code used.
- Supporting visualizations and tables.
- Presentation summarizing findings and recommendations.

Dataset:

- Provided financial dataset with 96 columns (95 features and 1 target variable).
- Target variable: `Bankrupt?` (1 for bankruptcy, 0 for non-bankruptcy).