EDA Lending Case Study

- **Prepared by: Sandeep B, Nishanth A.V:**

Photo by Tierra Mallorca on Unsplash

Introduction

- **Definition and Importance of EDA in Lending:** EDA lending, or Exploratory Data Analysis lending, involves analyzing customer and loan data to make informed approval decisions.
- **Objectives of the Case Study:** Understand patterns in borrower attributes and loan performance, and develop predictive insights to improve loan approval decisions.

Photo by Stephen Dawson on Unsplash

Background

- **Evolution and Challenges in the Lending Industry:** The lending industry faces challenges like increasing default rates and changing customer preferences, which impact loan approval processes.
- **Key Factors Influencing Loan Defaults:** Factors such as credit history, debt-to-income ratio, loan amount, interest rate, and loan terms significantly influence loan defaults.

Photo by Melinda Gimpel on Unsplash

Problem Statement

- **Identifying Patterns in Consumer and Loan Attributes:** Explore specific characteristics of borrowers and loan terms linked to defaults to improve loan approval decisions.

- **Developing Predictive Models for Loan Defaults:** Utilize insights from EDA to create models that assess new loan applicants' likelihood of defaulting, enabling informed decisions.

Photo by Jason Briscoe on Unsplash

Data Extraction & Cleaning

- **Data Sources and Extraction Process:** Data extracted from loan records (2007-2011), converted to CSV format, and imported for EDA.
- **Cleaning Procedures:** Handled missing values, standardized numerical variables, checked for outliers, and removed irrelevant columns to ensure data quality.

Photo by Museums Victoria on Unsplash

Exploratory Data Analysis (EDA)

- **Univariate Analysis:** Examines single variables to identify patterns and trends. Key variables analyzed include loan status, grade, and verification status.
- **Segmented Univariate Analysis:** Divides data into segments based on categorical variables, allowing for comparison across different groups, such as interest rates by grade.
- **Bivariate Analysis:** Explores relationships between two variables, such as loan status and verification status, to understand correlations and patterns.
- **Derived Metrics:** New metrics created from existing data to uncover hidden patterns, such as extracting year and month from issue dates.

Photo by Luke Chesser on Unsplash

Results

- **Key Findings from EDA:** Identified patterns in borrower attributes and loan terms associated with defaults. Developed predictive models for assessing new loan applicants.
- **Implications for Loan Approval Decisions:** Insights can help financial institutions make more informed loan approval decisions, reducing risk and enhancing profitability.

Photo by Sigmund on Unsplash

Conclusion

- **Summary of Insights:** EDA lending provides critical insights into borrower behavior and loan performance, enabling better loan approval decisions.
- **Recommendations for Financial Institutions:** Implement EDA techniques to identify high-risk borrowers, refine loan approval criteria, and guide marketing campaigns.

Photo by Sincerely Media on Unsplash

Future Implementations

- **Machine Learning for Loan Default Prediction:** Implementing machine learning models to analyze borrower and loan data. Algorithms such as logistic regression, decision trees, and random forests will be used to predict the likelihood of loan default.

- **Data Preprocessing and Feature Engineering:** Cleaning and preparing data, extracting relevant features, and engineering new features to improve model performance.

- **Model Training and Validation:** Using training and validation datasets to develop, test, and refine models. Employing cross-validation techniques to ensure robustness.

- **Deployment and Monitoring:** Deploying the model into the loan approval process. Continuously monitoring performance and updating the model with new data to maintain accuracy.

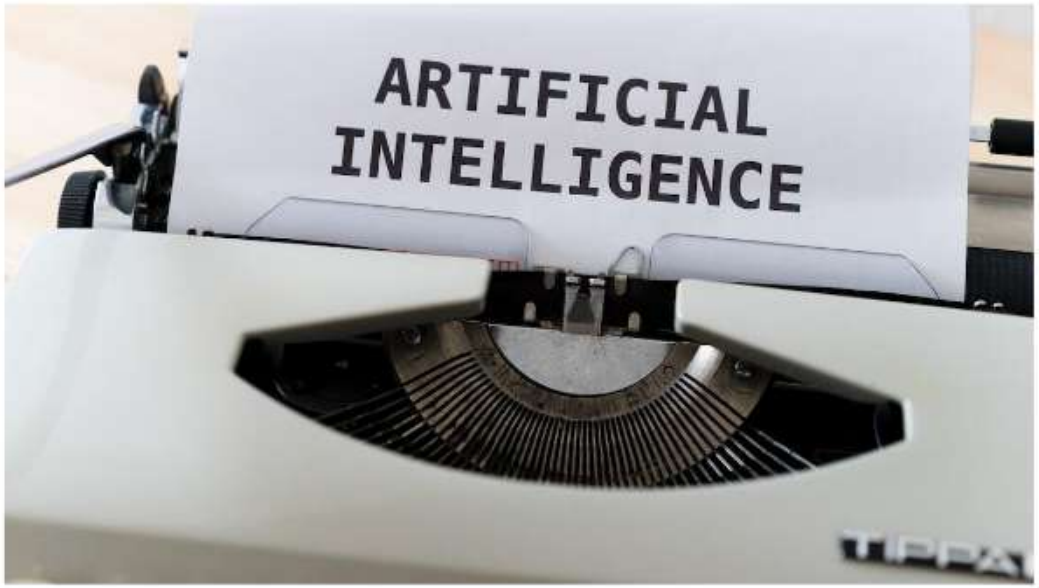

Photo by Markus Winkler on Unsplash