**Project Report: Bankruptcy Prediction**

1. **INTRODUCTION**

*1.1 Project Overview*

The project aims to predict bankruptcy in Polish companies using advanced machine learning techniques. Leveraging a comprehensive dataset from the Emerging Markets Information Service (EMIS) spanning the years 2000-2013, the objective is to develop a sophisticated early warning system for companies facing the risk of bankruptcy.

*1.2 Purpose*

This project serves as a crucial tool for stakeholders, helping investors and financial analysts in identifying potential bankruptcies. By harnessing the power of machine learning algorithms, the system delivers accurate predictions based on an in-depth analysis of historical financial data.

1. **LITERATURE SURVEY**

*2.1 Existing Problem*

The landscape of financial instability and bankruptcy presents significant challenges for businesses and investors. The importance of early detection cannot be overstated, as it plays a pivotal role in mitigating risks and guiding strategic decision-making.

*2.2 References*

In the pursuit of addressing this critical issue, the project draws inspiration from seminal works in financial research:

Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. The Journal of Finance, 23(4), 589-609.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. Journal of Accounting Research, 109-131.

*2.3 Problem Statement Definition*

The project's primary challenge involves crafting a robust bankruptcy prediction model capable of empowering stakeholders to make well-informed decisions in a timely manner.

**Attribute Information:**

**X1** net profit / total assets

**X2** total liabilities / total assets

**X3** working capital / total assets

**X4** current assets / short-term liabilities

**X5** [(cash + short-term securities + receivables - short-term liabilities) / (operating expenses - depreciation)] \* 365

**X6** retained earnings / total assets.

**X7** EBIT / total assets

**X8** book value of equity / total liabilities

**X9** sales / total assets

**X10** equity / total assets

**X11** (gross profit + extraordinary items + financial expenses) / total assets

**X12** gross profit / short-term liabilities

**X13** (gross profit + depreciation) / sales

**X14** (gross profit + interest) / total assets

**X15** (total liabilities \* 365) / (gross profit + depreciation)

**X16** (gross profit + depreciation) / total liabilities

**X17** total assets / total liabilities

**X18** gross profit / total assets

**X19** gross profit / sales

**X20** (inventory \* 365) / sales

**X21** sales (n) / sales (n-1)

**X22** profit on operating activities / total assets

**X23** net profit / sales

**X24** gross profit (in 3 years) / total assets

**X25** (equity - share capital) / total assets

**X26** (net profit + depreciation) / total liabilities

**X27** profit on operating activities / financial expenses

**X28** working capital / fixed assets

**X29** logarithm of total assets

**X30** (total liabilities - cash) / sales

**X31** (gross profit + interest) / sales

**X32** (current liabilities \* 365) / cost of products sold

**X33** operating expenses / short-term liabilities

**X34** operating expenses / total liabilities

**X35** profit on sales / total assets

**X36** total sales / total assets

**X37** (current assets - inventories) / long-term liabilities

**X38** constant capital / total assets

**X39** profit on sales / sales

**X40** (current assets - inventory - receivables) / short-term liabilities

**X41** total liabilities / ((profit on operating activities + depreciation) \* (12/365))

**X42** profit on operating activities / sales

**X43** rotation receivables + inventory turnover in days

**X44** (receivables \* 365) / sales

**X45** net profit / inventory

**X46** (current assets - inventory) / short-term liabilities

**X47** (inventory \* 365) / cost of products sold

**X48** EBITDA (profit on operating activities - depreciation) / total assets

**X49** EBITDA (profit on operating activities - depreciation) / sales

**X50** current assets / total liabilities

**X51** short-term liabilities / total assets

**X52** (short-term liabilities \* 365) / cost of products sold)

**X53** equity / fixed assets

**X54** constant capital / fixed assets

**X55** working capital

**X56** (sales - cost of products sold) / sales

**X57** (current assets - inventory - short-term liabilities) / (sales - gross profit - depreciation)

**X58** total costs /total sales

**X59** long-term liabilities / equity

**X60** sales / inventory

**X61** sales / receivables

**X62** (short-term liabilities \*365) / sales

**X63** sales / short-term liabilities

**X64** sales / fixed assets

**Class -** 0 did not get bankrupt/ 1 - got bankrupt

1. **IDEATION & PROPOSED SOLUTION**

*3.1 Empathy Map Canvas*

An in-depth exploration via the empathy map revealed a pressing need for a tool that not only provides quick and accurate bankruptcy predictions but also empowers stakeholders to act promptly in response to emerging financial risks.

*3.2 Ideation & Brainstorming*

Brainstorming sessions centered around the exploration of diverse machine learning algorithms, with a particular emphasis on the implementation of Random Forest and Support Vector Machine (SVM) techniques to ensure accurate and reliable predictions.

1. **REQUIREMENT ANALYSIS**

*4.1 Functional Requirement*

The functional requirements include the precise prediction of bankruptcy probability and the development of a user-friendly interface, allowing stakeholders to easily input relevant financial attributes.

*4.2 Non-Functional Requirements*

To ensure the effectiveness of the system, the project has set ambitious non-functional requirements, aiming for a prediction accuracy of 98% and a seamless, user-friendly interface.

1. **PROJECT DESIGN**

*5.1 Data Flow Diagrams & User Stories*

Detailed user stories and comprehensive data flow diagrams have been designed to illustrate the intricate interaction between users and the sophisticated bankruptcy prediction model.

*5.2 Solution Architecture*

The project's solution architecture encompasses a dynamic web-based interface for user input, a robust backend server for intricate processing, and the incorporation of state-of-the-art machine learning models for accurate predictions.

**6. PROJECT PLANNING & SCHEDULING**

*6.1 Technical Architecture*

The technical architecture relies on Flask for the backend, HTML/CSS for the user interface, and the implementation of scikit-learn for the deployment of cutting-edge machine learning algorithms.

*6.2 Sprint Planning & Estimation*

Sprints have been meticulously planned, focusing on incremental development to ensure the timely delivery of features. Each sprint is crafted to address specific aspects of the project's evolution.

*6.3 Sprint Delivery Schedule*

The sprint schedule is as follows:

Sprint 1: User interface and input validation

Sprint 2: Model integration and testing

Sprint 3: Performance optimization and deployment

1. **CODING & SOLUTIONING**

*7.1 Feature 1*

The first feature involves the implementation of a Support Vector Machine (SVM)-based bankruptcy prediction, with a nuanced consideration of kernel functions and hyperparameters for optimal performance.

*7.2 Feature 2*

The second feature encompasses the implementation of a Random Forest-based bankruptcy prediction, emphasizing ensemble learning techniques to ensure robust and reliable predictions.

*7.3 Database Schema (if Applicable)*

As the project is focused on real-time prediction, it does not involve a traditional database schema.

**8. PERFORMANCE TESTING**

*8.1 Performance Metrics*

A comprehensive evaluation of the model's performance has been conducted, utilizing a range of metrics including accuracy, precision, recall, and F1-score to ensure the system's efficacy.

**9. RESULTS**

*9.1 Output Screenshots*

A collection of output screenshots from the web interface has been provided, showcasing the bankruptcy predictions for sample input data, and emphasizing the system's user-friendly design.

**10. ADVANTAGES & DISADVANTAGES**

A detailed analysis of the advantages of early bankruptcy prediction is provided, along with a candid exploration of potential limitations and constraints associated with the model.

**11. CONCLUSION**

The conclusion offers a succinct summary of the project, underscoring the critical importance of early bankruptcy prediction and affirming the effectiveness of the developed model.

**12. FUTURE SCOPE**

A forward-looking exploration of potential enhancements is presented, including considerations for integrating additional data sources and refining machine learning algorithms for continuous improvement.

**13. APPENDIX**

GitHub & Project Demo Link

<https://drive.google.com/file/d/1BzthCmY3IJslTMjTtYlmPfBuEfytqlvZ/view?usp=drive_link>