Financial Health Monitoring System

Predictive Analytics for Corporate Financial Risk Assessment

Broke Blusters

MUSI CALORI
TEAM LEAD

SHANE MWANGI
NOTEBOOK
CRISP DM

MICHAEL MAINA
NOTEBOOK
TRACKING TOOL
DEPLOYMENT

IVY CHELAGAT
NONTECHNICAL
PRESENTATION

HASSAN ALI
README

Problem Statement

- Investors often rely on outdated or incomplete financial information
- A data-driven approach is needed to evaluate financial health accurately
- Our project analyzes key financial statements to assess performance and risk clearly.

Why It Matters

Investors need clear insights to identify stable and profitable opportunities.

Lenders require reliable financial indicators to evaluate credit risk.

Business leaders depend on data-driven analysis to plan and manage growth effectively.

Project Objectives

MAIN:
To build a **data analysis and scoring system** that evaluates a company's financial health using real-world financial data.

SPECIFIC:

- Collect and clean financial data from Yahoo Finance.
- Analyze key financial metrics like revenue and profitability trends.
- Build a financial health scoring model.
- Visualize insights with clear charts and dashboards.
- Provide data-driven recommendations for decision-makers.

Data Uderstanding

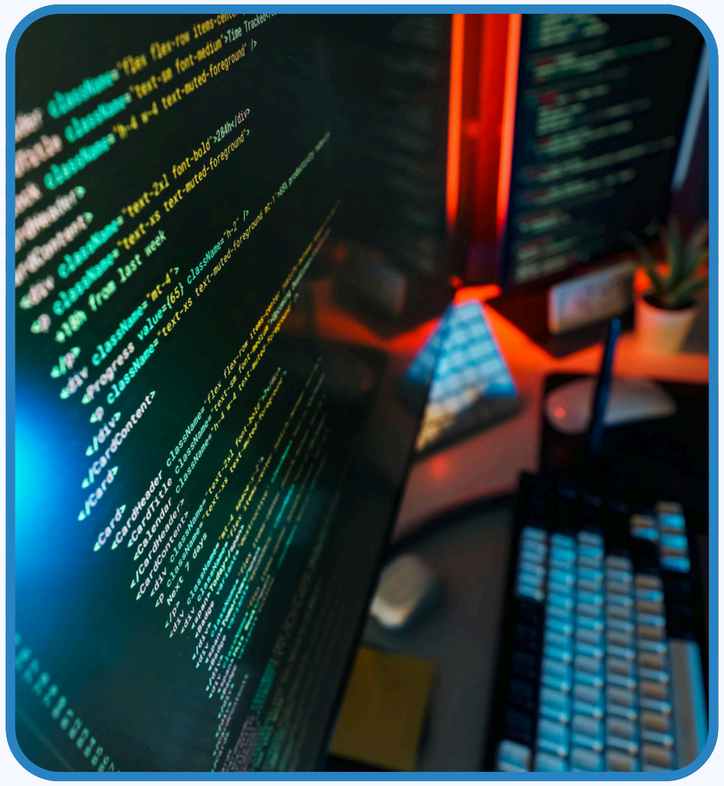
Data Sources:

- Yahoo Finance API and related financial datasets

Datasets Used:

- Income Statements
- Balance Sheets
- Cash Flow Statements
- Stock Price History

Combined S&P 500 and NASDAQ tickers.
Filtered for companies with Market Cap ≥ $100M for data relevance.
Final Dataset: 495 companies, covering 99.5% of the total market cap



Data Pipeline & Engineering

- **Extraction:** Fetched data for 495+ tickers using yfinance with caching for efficiency.
- **Cleaning & Standardization:**
  - Handled missing values.
  - Scaled financial figures to **Billions** for consistency.
  - Removed duplicates.
- **Feature Engineering:** Calculated critical financial ratios.
- **Merging:** Created a unified master dataset for analysis.

Feature Engineering & Key Ratios

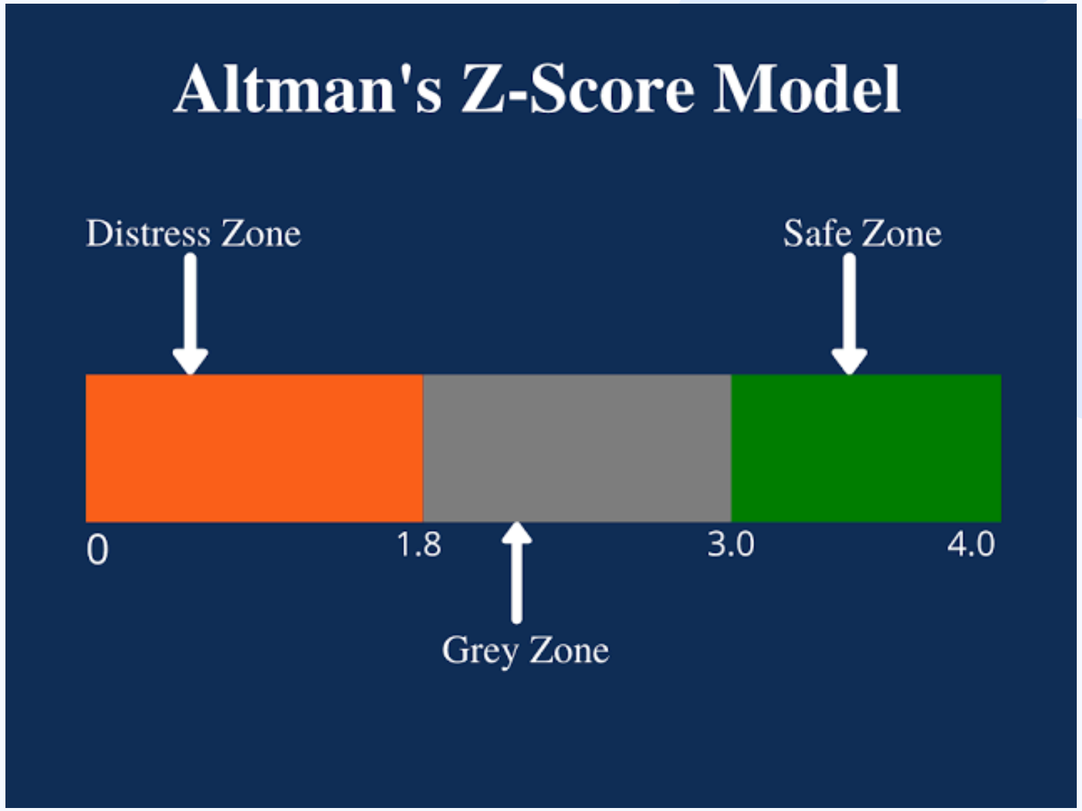
We engineered a suite of financial ratios to power our model:

- **Profitability:**
  - Gross Margin, Operating Margin, Net Margin
  - Return on Assets (ROA), Return on Equity (ROE)
- **Leverage & Liquidity:**
  - Debt to Equity Ratio, Current Ratio
- **Cash Flow:**
  - Free Cash Flow (FCF), FCF Yield, CapEx Ratio
- **Risk Score:**
  - **Altman Z-Score** (A composite risk indicator)

Altman Z-Score Methodology

ORIGINAL ALTMAN Z-SCORE (1968):
Z = 1.2A + 1.4B + 3.3C + 0.6D + 1.0E

Where:
A = Working Capital / Total Assets
B = Retained Earnings / Total Assets
C = EBIT / Total Assets
D = Market Value Equity / Total Liabilities
E = Sales / Total Assets



Model Approach

Models Used

- Random Forest – baseline classification.
- XGBoost – improved handling of complex patterns and imbalance.

Features

- Financial ratios + Altman Z-Score.

Target

- Financial Health Category: Safe / Grey / Distress

Model Evaluation

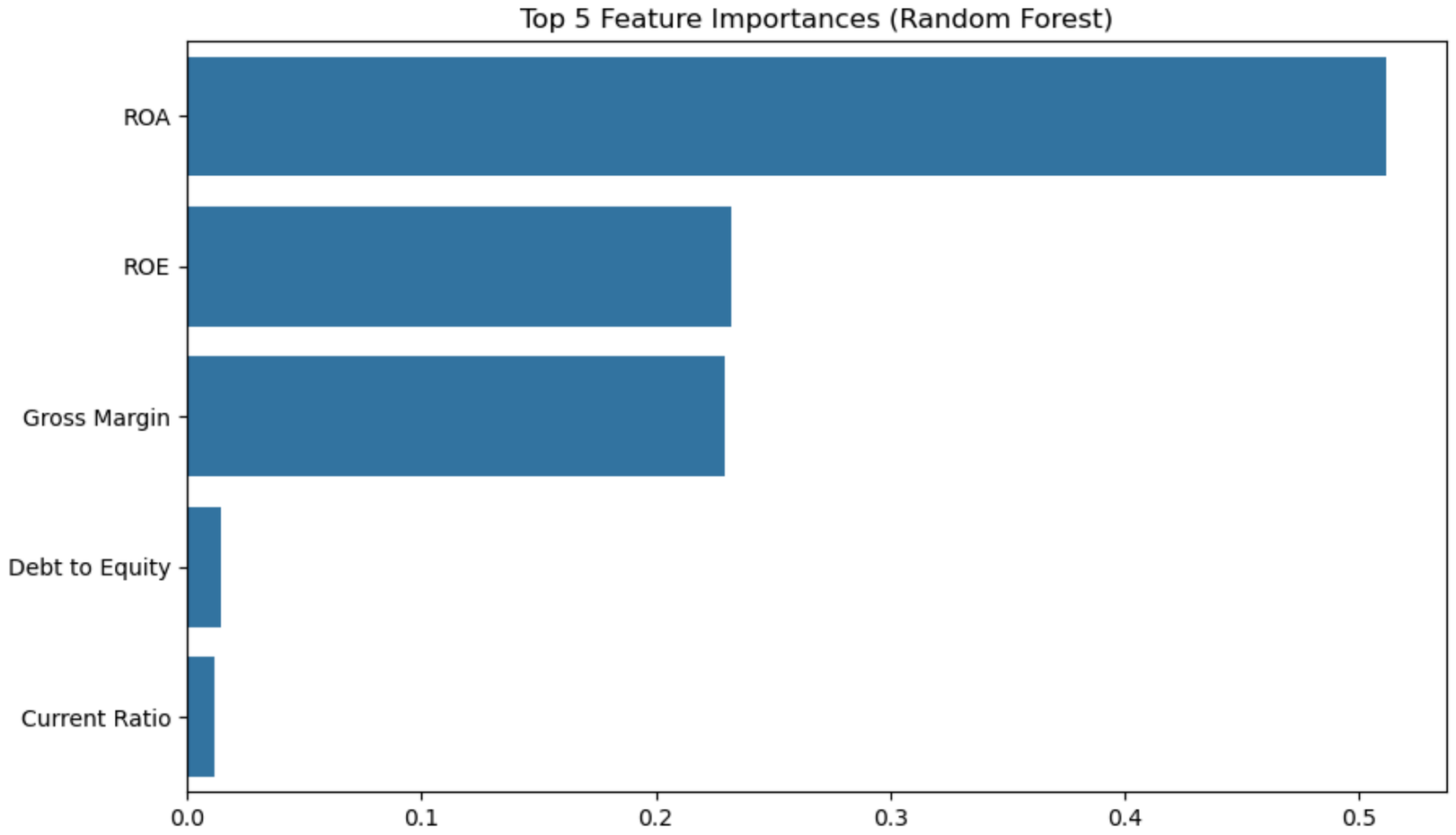
Evaluation Metrics Used

- Accuracy: Overall correctness of the model.
  Random Forest = 83%
  Xgboost = 83%
- Precision & Recall: To measure how well the model identifies each financial health class correctly.
- Confusion Matrix: Shows where the model makes most classification errors.

Key Result

- Performance varied by class due to class imbalance (fewer “Safe” companies in the dataset), affecting precision & recall especially in the Safe category

Feature Importance



Profitability and liquidity ratios are the strongest signals of financial health.

Companies that generate steady earnings and manage short-term cash obligations tend to score higher.

Metrics of Success

- Balanced Recall Across Classes: The model should correctly identify companies in all risk categories, not just the majority class.
- Focus on Distress Detection: The most important goal is minimizing false negatives so that financially distressed companies are not misclassified as safe.
- Desired Outcome: Improved and more even recall scores for Safe, Grey, and Distress categories to support reliable decision-making.

Conclusion

- **Scalable Data Pipeline**: Successfully built a system that collects, cleans, and processes financial data for many companies efficiently.
- **Financial Risk Indicator (Z-Score)**: Average Z-Score ≈ 1.3, suggesting most companies in the sample show signs of potential financial distress.
- **Data Quality Check**: A small number of negative values appeared in profit-related fields, which is expected for loss-making firms and confirms the importance of our cleaning steps.

Recommendations

1. **For Safe Companies**: Continue current financial practices and consider strategic reinvestment for growth (e.g., new markets or product expansion).
2. **For Grey-Zone Companies**: Review costs and efficiency, and strengthen cash reserves to improve financial stability.
3. **For Distressed Companies**: Consider debt restructuring and streamline operations by reducing or removing underperforming activities.
4. **For Investors & Lenders**: Use the financial health score as a screening tool, prioritize Safe, monitor Grey and apply caution or risk controls for Distressed companies.

Next Steps

- Link scores to actionable guidance: Translate each risk level (Safe, Grey, Distress) into practical business advisory messages to support real-world financial decisions.
- Add model explainability: Use SHAP to clearly show which financial features drive each prediction
- Create dashboards: Build visualizations that display Z-Score trends and financial health ratings in an easy-to-read format.

Deployment

To make the financial health analysis accessible and interactive, we deployed the model using Streamlit

- **Users can upload company financial data directly into the app.**
- **The system calculates key financial ratios and assigns a Financial Health Score (Safe / Grey / Distress).**
- **The app visualizes results through clear indicators for easy interpretation**

Thank You Very Much