Stock Pulse

Category: Data Science Dominion

Where data science meets financial foresight.

Team Members

- Ishna Fatima
- Mehran Jamali
- Muhammad Ayan
- Sarim Hussain
- Syed Muhammad Yasir

Mentor:

Waqar Aamir

Acknowledgement:

We would like to extend our heartfelt gratitude to everyone who contributed to the success of the "**Stock Pulse**" project. This endeavor would not have been possible without the support and guidance of many individuals and organizations.

First and foremost, we would like to express our deepest thanks to my Mentor & advisor **Sir Waqar Aamir**. His expertise and guidance has been invaluable throughout the project. His encouragement and commitment to excellence inspired us to push the boundaries of our research and deepen our understanding of **Data Science**.

We would also like to acknowledge our team members and peers who collaborated with us during various stages of this project. Their diverse perspectives and collaborative spirit greatly enriched our experience and helped us overcome complex challenges. The brainstorming sessions and discussions were instrumental in shaping the direction of this project.

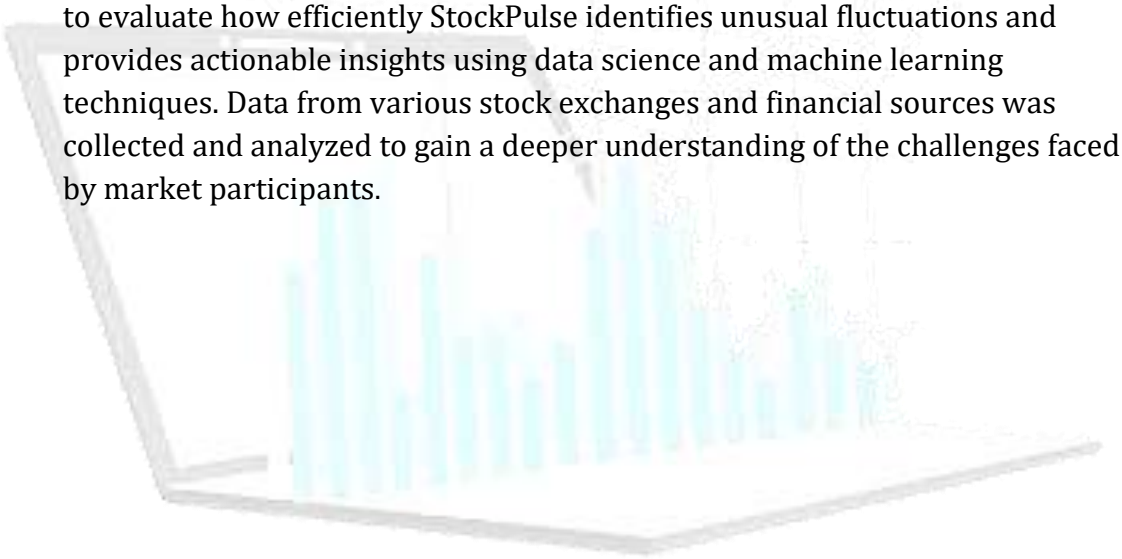
Lastly, we hope that "**Stock Pulse**" will make a meaningful contribution to the field and inspire future research and innovation. Thank you all for your invaluable contributions and support.

Contents

Chapter 1: INTRODUCTION....4
1.1 Motivation....5
1.2 Problem statement....5
Chapter 2: Data Collection....6
2.1 Introduction....6
2.2 Secondary Data:....7
Chapter 3: Scope of Project....8
Chapter 4: Project Architecture....15
Functional Requirements:....9
5. Software Requirements:....12
Design and Develop By:....11
Task Allocation:.... **Error! Bookmark not defined.**
Chapter 6 : Design Specifications:....14
Chapter 7: Conclusion and Future Work:....29
7.1 Introduction....29
7.2 Conclusion....29
7.3 Future Work....30

Abstract:

Financial market stability is a critical concern for investors, analysts, and regulatory bodies in today's rapidly evolving economic landscape. This study explores the effectiveness of **StockPulse** in safeguarding market integrity by detecting anomalies in stock price movements and trading patterns, assessing its ability to prevent financial risks and fraudulent activities. Many stakeholders face challenges such as limited access to advanced monitoring tools and the increasing complexity of high-frequency trading, making it difficult to maintain market transparency. The primary aim of this research is to evaluate how efficiently StockPulse identifies unusual fluctuations and provides actionable insights using data science and machine learning techniques. Data from various stock exchanges and financial sources was collected and analyzed to gain a deeper understanding of the challenges faced by market participants.



Chapter 1: INTRODUCTION

1.1 Motivation

Financial time series data, such as stock prices and returns, are vital indicators of a company's performance and broader market dynamics. These datasets are often characterized by volatility, noise, and non-linear trends that are hard to interpret using traditional methods. Anomalies such as sudden spikes or dips can signal critical events like insider trading, economic crises, cyber-attacks, or system failures. Manual inspection is impractical, and traditional statistics often fail. Hence, a data-driven anomaly detection system is essential for modern financial infrastructure

1.2 Problem statement

- Financial markets produce large volumes of volatile and noisy time series data.
- Anomalies (sudden spikes, dips, or irregular patterns) may indicate:
- **Traditional methods** (statistical models, manual monitoring) are:
- Lack of automated anomaly detection leads to:
- Need for a **scalable, data science–driven solution** to:

Chapter 2: Data Collection

2.1 Introduction

StockPulse is an advanced anomaly detection system designed to identify unusual patterns in financial time series data, particularly stock prices. In today's dynamic markets, stock data is characterized by high volatility, noise, and sudden fluctuations that may indicate critical events such as insider trading, economic instability, cyber-attacks, or technical errors. Traditional statistical methods often fail to capture these hidden irregularities, and manual monitoring is impractical due to the sheer volume and speed of financial data.

By leveraging Data Science and Machine Learning techniques, **StockPulse** provides a scalable solution for analyzing stock price movements, detecting anomalies, and delivering actionable insights in near real-time. The system not only helps investors and analysts recognize potential risks and opportunities but also supports regulatory bodies in maintaining market integrity. With interactive dashboards and reporting features, Stock Pulse empowers stakeholders to make informed, data-driven decisions in a rapidly evolving financial environment.

2.2 Secondary Data:

StockPulse relies on secondary data sources to analyze stock price anomalies. Secondary data refers to information already collected and published by reliable financial platforms and institutions. These datasets provide historical and real-time insights into market behavior, which are crucial for training and testing anomaly detection models.

Sources of Secondary Data.

- **Financial APIs**:
  - Yahoo Finance
  - Google Finance
- **Stock Exchange Data**:
  - NYSE
  - LSE (London Stock Exchange)
- **Historical Data Archives**:
  - CSV datasets are available from Kaggle, Bloomberg terminals, or official exchange websites.
- **Research & Publications**:
  - Academic datasets on financial anomalies
  - Market reports and whitepapers on stock price trends
  -

❖ Types of Secondary Data Collected

- Date and time of trade
- Opening, closing, high, and low prices
- Adjusted closing price
- Trading volume
- Market indices for comparison

Chapter 3: Scope of Project

The objective of this project is to develop a powerful application that can accurately detect and analyze anomalies in financial stock market data. By utilizing large datasets that include historical stock prices, trading volumes, and indices, the project will apply machine learning algorithms to identify unusual spikes, dips, or patterns that deviate from expected market behavior. This will enable early detection of irregular activities such as market manipulation, insider trading, or sudden economic shifts.

The scope includes creating a scalable data pipeline for continuous data ingestion, allowing integration of both real-time and historical stock market data from reliable financial APIs. Advanced preprocessing and feature engineering techniques will support this pipeline, ensuring that the data is cleaned, normalized, and optimized for accurate anomaly detection. This will help the models deliver reliable results across different stock exchanges and financial instruments.

The project also aims to build an intuitive user interface that is user-friendly for financial analysts, investors, and regulatory authorities. This interface will enable users to easily upload datasets or connect to live market feeds, run anomaly detection models, and view real-time results through interactive dashboards. The system will also provide reports with clear insights and visualizations, making it easier for users to interpret anomalies and take informed actions.

Additionally, the project will focus on adaptability and scalability, ensuring that StockPulse can be applied to different markets, asset classes, and trading environments. By integrating macroeconomic indicators and external market factors, the system can offer more precise anomaly detection and contextual insights. Continuous updates and improvements will be made to enhance the underlying machine learning models, keeping them effective against evolving market conditions and emerging financial patterns.

Functional Requirements:

The **StockPulse** system will provide the following core functionalities to ensure accurate and efficient anomaly detection in financial time series data:

1. Data Ingestion and Management
    - The system shall collect historical and real-time stock market data from APIs (Yahoo Finance, Alpha Vantage, Quandl).
    - The system shall allow users to import data in structured formats (CSV, Excel).
    - The system shall store and organize collected data chronologically in a database for easy access.
2. Data Preprocessing
    - The system shall clean raw data by handling missing values, duplicate records, and inconsistencies.
    - The system shall normalize and standardize stock features (open, high, low, close, adjusted close, volume) for uniform analysis.
    - The system shall generate derived features such as returns, moving averages, and volatility indicators.
3. Feature Selection and Engineering
    - The system shall provide statistical tools to identify and select the most relevant features.
    - The system shall reduce dimensionality and noise to improve algorithm efficiency.
4. Anomaly Detection Module
    - The system shall implement machine learning algorithms (Isolation Forest, Support Vector Machines, etc.) for detecting anomalies in time series data.
    - The system shall allow configuration of algorithm parameters for fine-tuning performance.
    - The system shall process both training and testing datasets to validate anomaly detection results.
5. Model Training and Evaluation

- The system shall train models using historical datasets to learn normal stock behavior.
- The system shall evaluate models with performance metrics such as precision, recall, F1-score, and accuracy.
- The system shall enable visualization of flagged anomalies for user confirmation.

6. Visualization and Dashboard
    - The system shall generate interactive dashboards using Tableau/Plotly/Streamlit.
    - Dashboards shall display stock trends, anomalies, and risk indicators in a user-friendly format.
    - The system shall provide graphical reports highlighting anomalies in price and volume data.
7. Reporting and Export
    - The system shall allow users to export anomaly detection results in CSV, Excel, or PDF formats.
    - The system shall generate detailed reports containing detected anomalies, their timestamps, and confidence scores.
8. User Interface (UI) and Interaction
    - The system shall provide a simple interface for uploading stock data, initiating analysis, and viewing results.
    - The system shall support real-time user interaction with charts and anomaly reports.
    - The system shall ensure easy navigation for both technical and non-technical users.

Design and Develop By:

SR No	Student Name	Student ID
1	Ishna Fatima	Student1404649
2	Mehran Jamali	Student1432322
3	Syed M Yasir	Student1389546
4	M Ayan	Student1366065
5	Sarim Hassan	Student1437725

5. Software Requirements:

5.1 Technologies to be used:

Library	Description
Numpy	Provides support for large, multi-dimensional arrays and matrices, along with a wide variety of mathematical functions to operate on these arrays.
pandas	Offers data structures and tools for efficient manipulation and analysis of structured data, particularly useful for data cleaning and preparation.
Pillow	An image-processing library that adds support for opening, manipulating, and saving image files in various formats.
requests	Used for sending HTTP requests in Python, making it easy to interact with APIs and retrieve data over the web.
tensor flow	A deep learning library that provides tools for building machine learning models, including neural networks for tasks like image recognition.
opencv-python	Provides functions to process images and videos, useful for tasks like object detection and computer vision.
matplotlib	A plotting library used for creating static, interactive, and animated visualizations in Python.

Task Allocation:

Presentation Activity			
Sr no	**Task**	**Team**	**Status**
1	Finding Dataset	Sarim Hassan	Complete
2	Dataset Preprocessing	Syed M Yasir	Complete
3	Training Model	Sarim Hassan	Complete
4	Frontend	M Ayan	Complete
5	Database	Mehran Jamali	Complete
6	Visualization	Syed M Yasir	Complete
7	Documentation	Ishna Fatima	Complete

Chapter 6 : Design Specifications:

Dataset Used in our Project

We used **historical stock market datasets** in our project **StockPulse**, obtained from reliable financial data providers.

Dataset Sources:

- Yahoo Finance API
- Alpha Vantage API
- Quandl
- Kaggle datasets containing historical stock data

About Dataset:

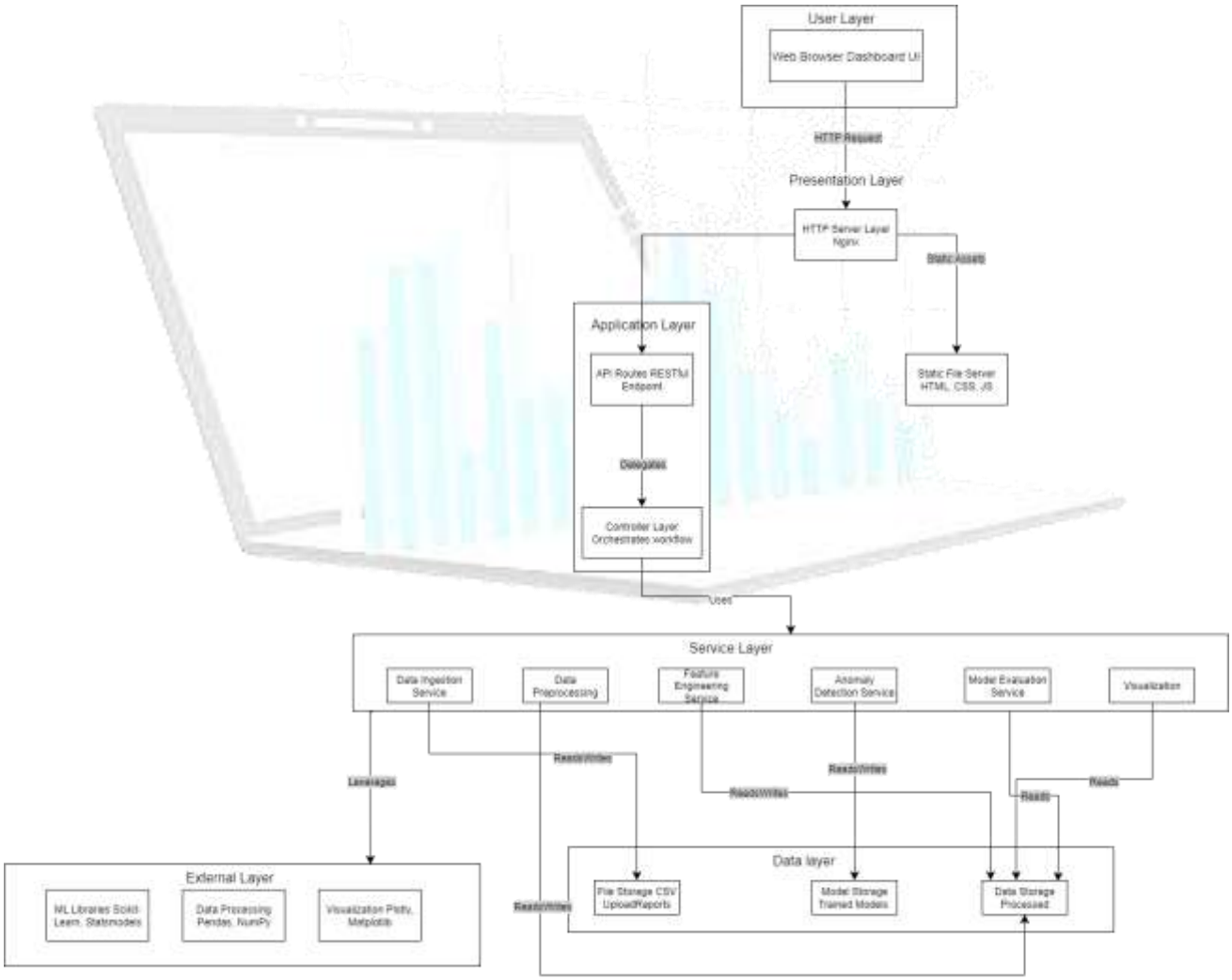
The dataset consists of daily stock price information for multiple companies, including:

- Date
- Open, High, Low, Close, Adjusted Close
- Volume

For preprocessing, the dataset was split into training (80%) and testing (20%) sets. Derived features such as **returns, moving averages, and volatility indicators** were generated to enhance anomaly detection. The training dataset was used to model normal behavior, while the testing dataset contained potential anomalies for validation.

Chapter 4: Project Architecture

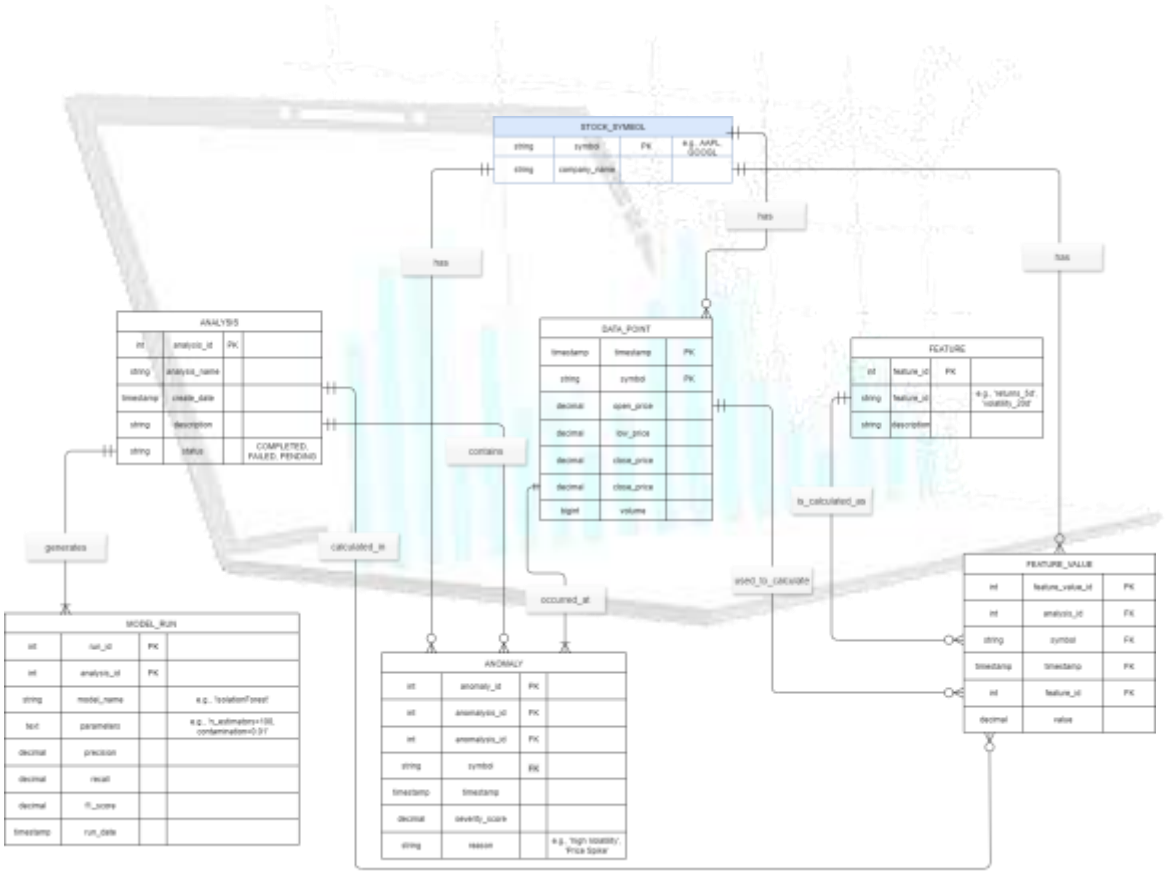
The architecture of StockPulse consists of multiple interconnected components designed to ensure seamless anomaly detection in stock market data.



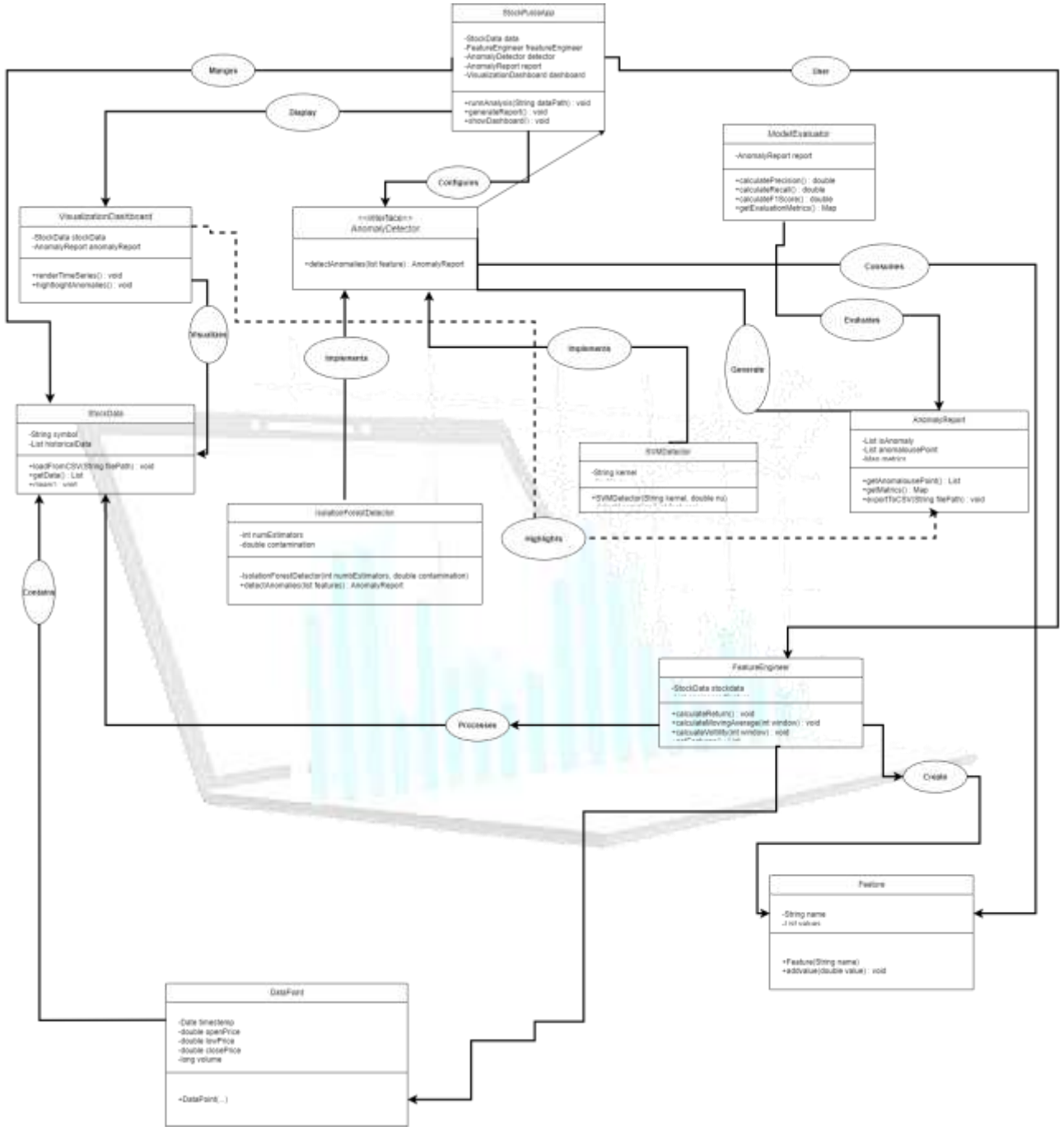
Entity-Relationship Diagram (ERD)

- The ERD for **StockPulse** outlines how different entities in the system are related to each other. It helps in understanding how stock market data, anomaly results, and users interact within the application.

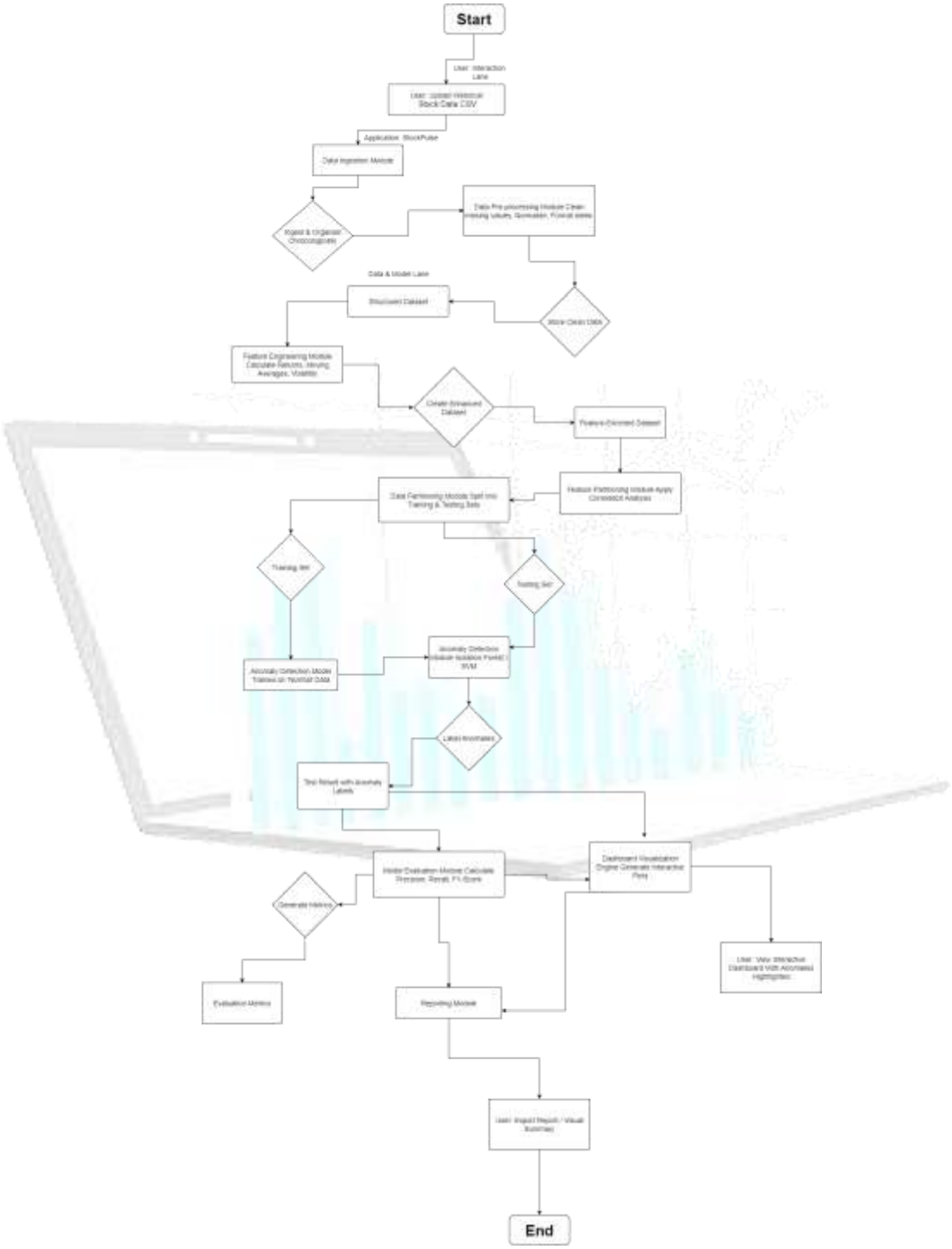
4.2 ERD Diagram



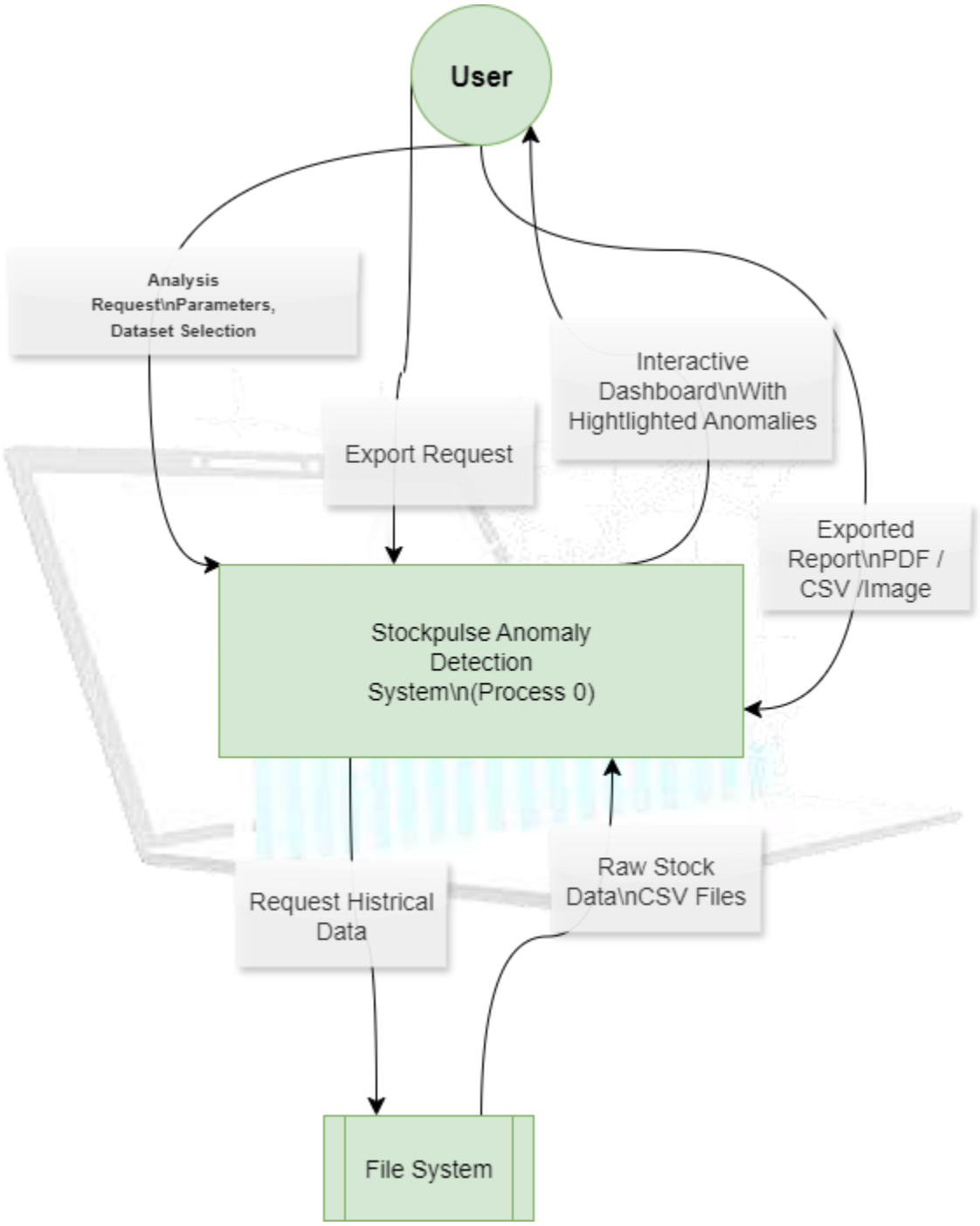
UML Class Diagram



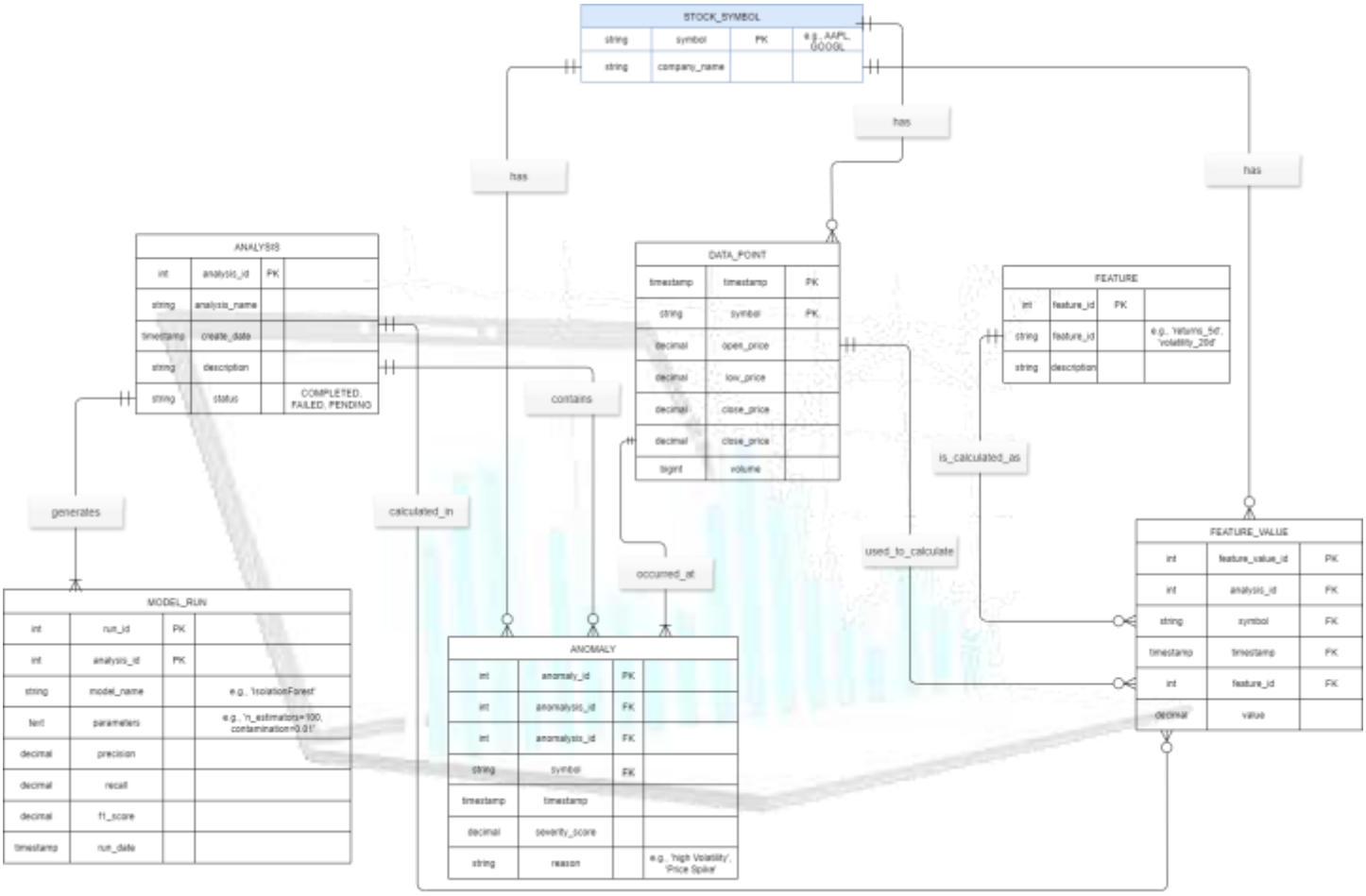
FLOWCHART DIAGRAM



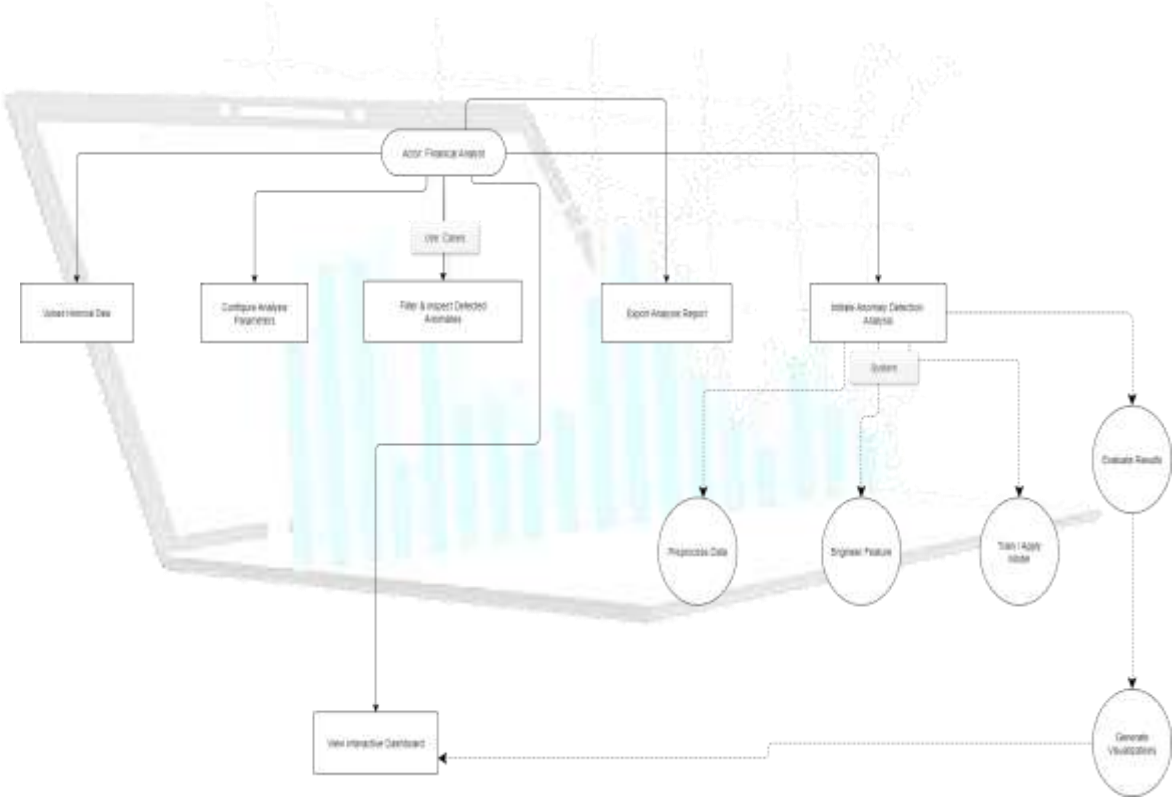
DFD Level 0 DIAGRAM



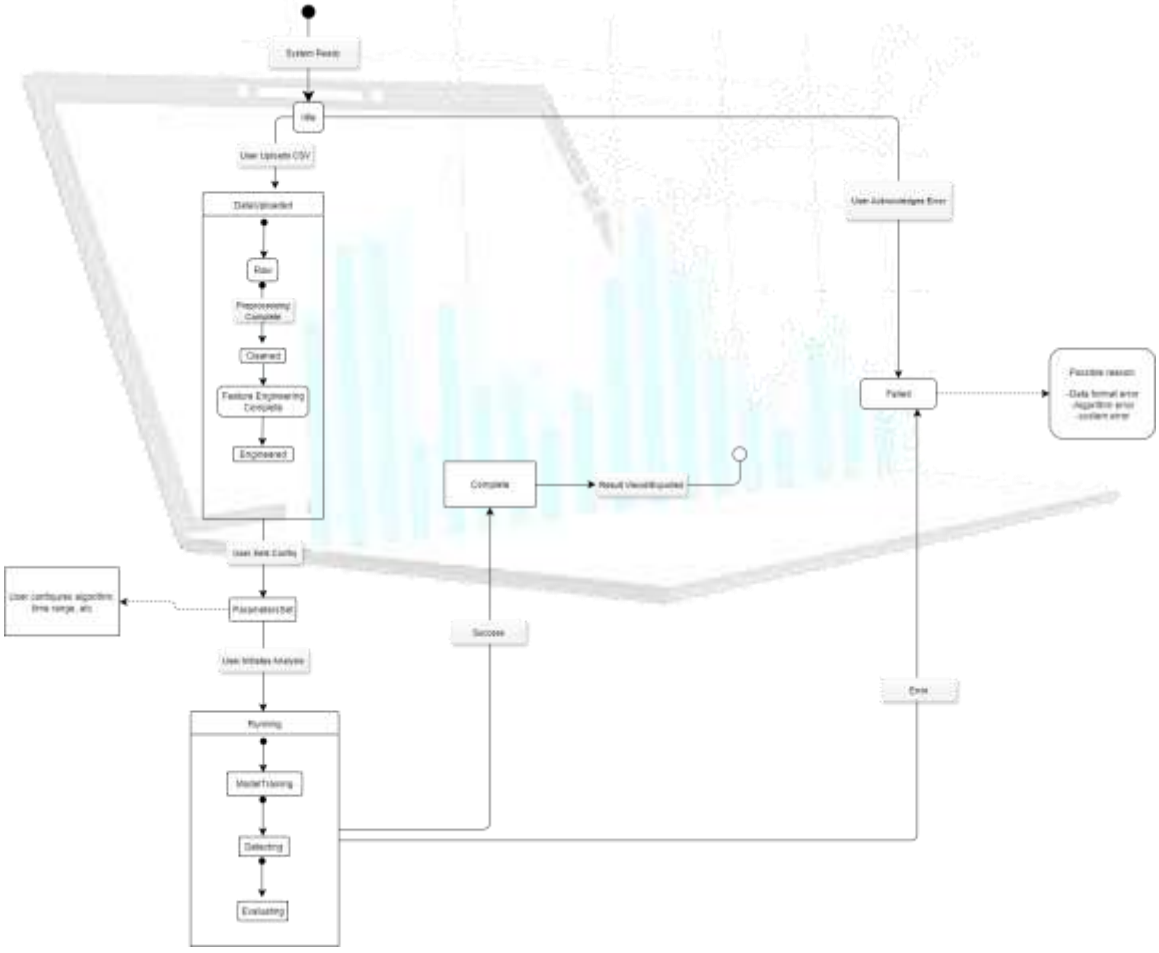
DFD Level 1 DIAGRAM



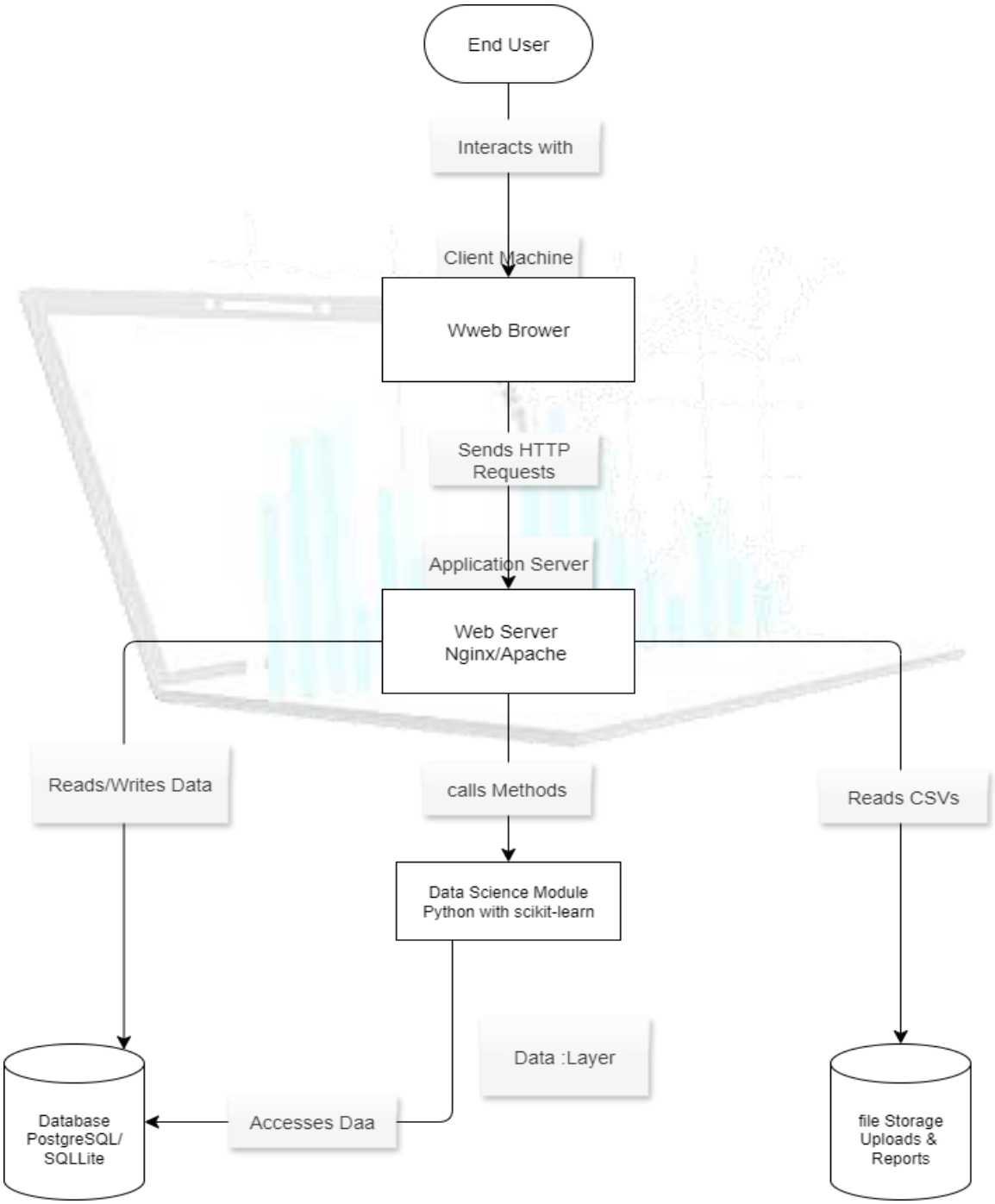
USE CASE DIAGRAM



STATE DIAGRAM



DEPLOMENT DIAGRAM

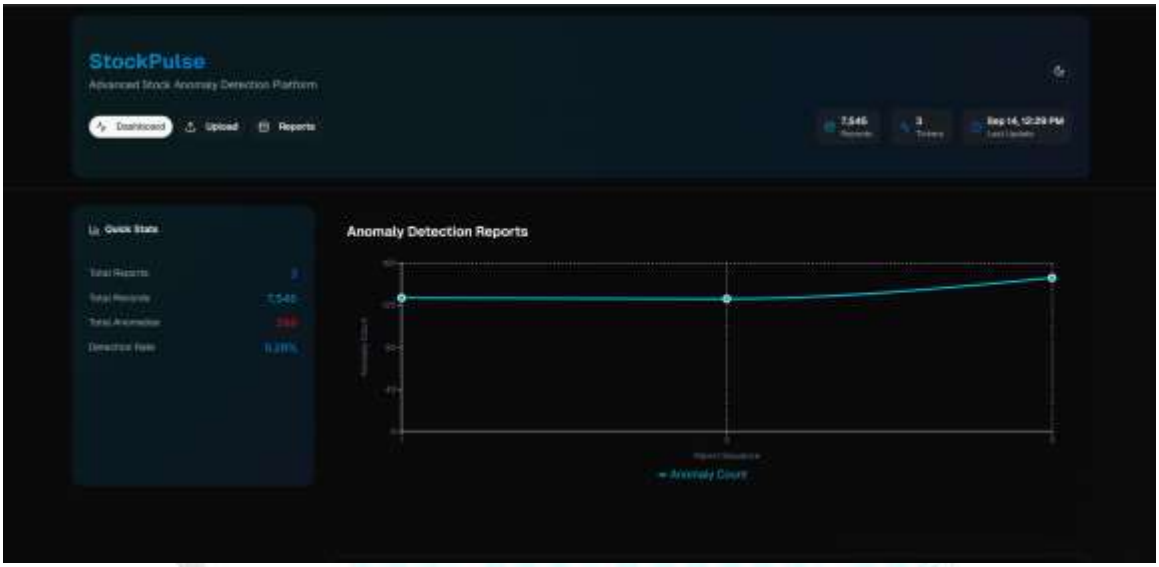


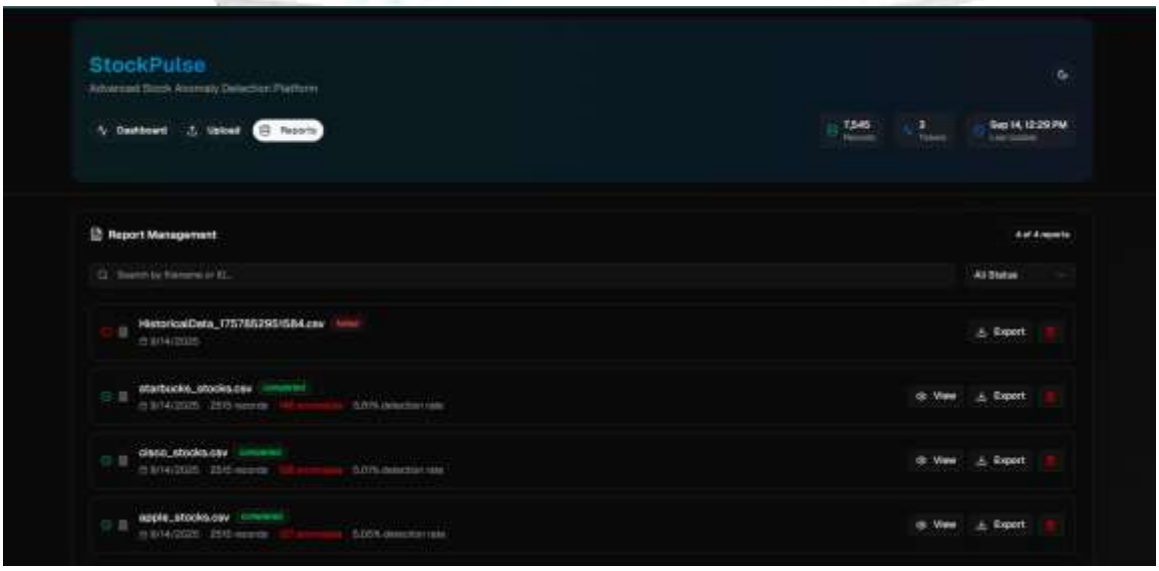
BLOG URL

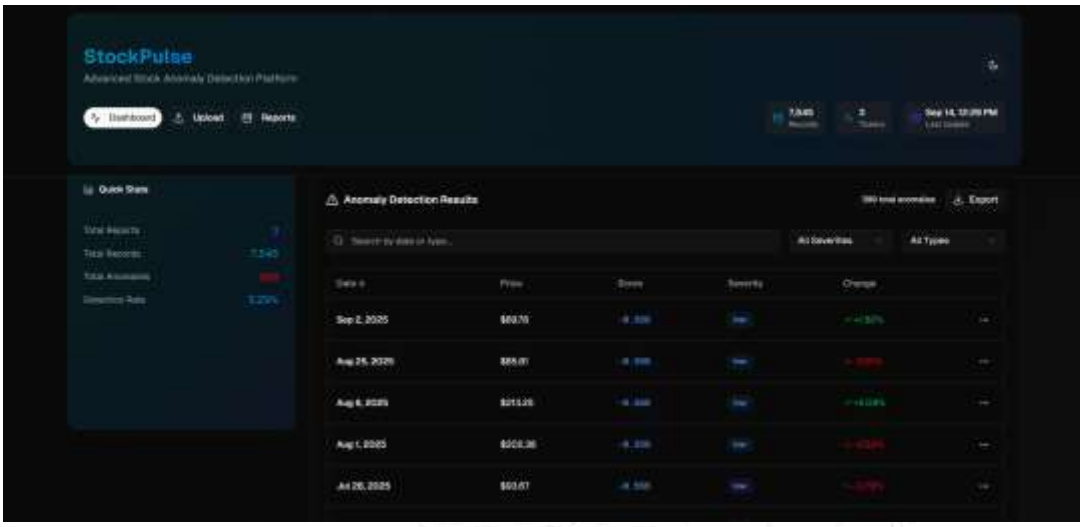
https://medium.com/@aptech.waqar00/stockpulse-redefining-stock-price-anomaly-detection-with-data-science-bae93c51c192

SCREENSHOTS OF UI

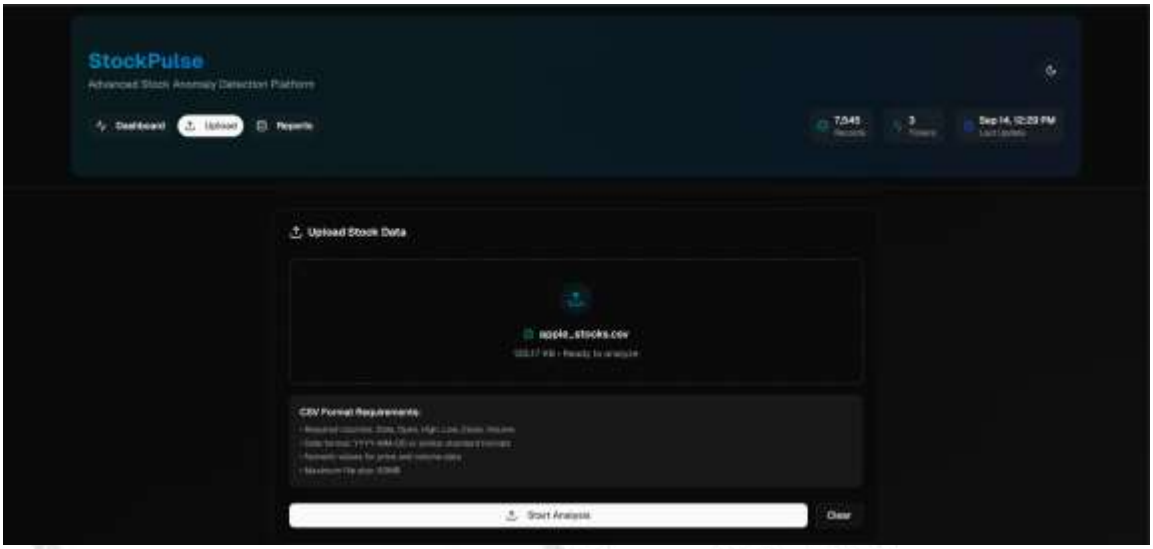
Dashboard View

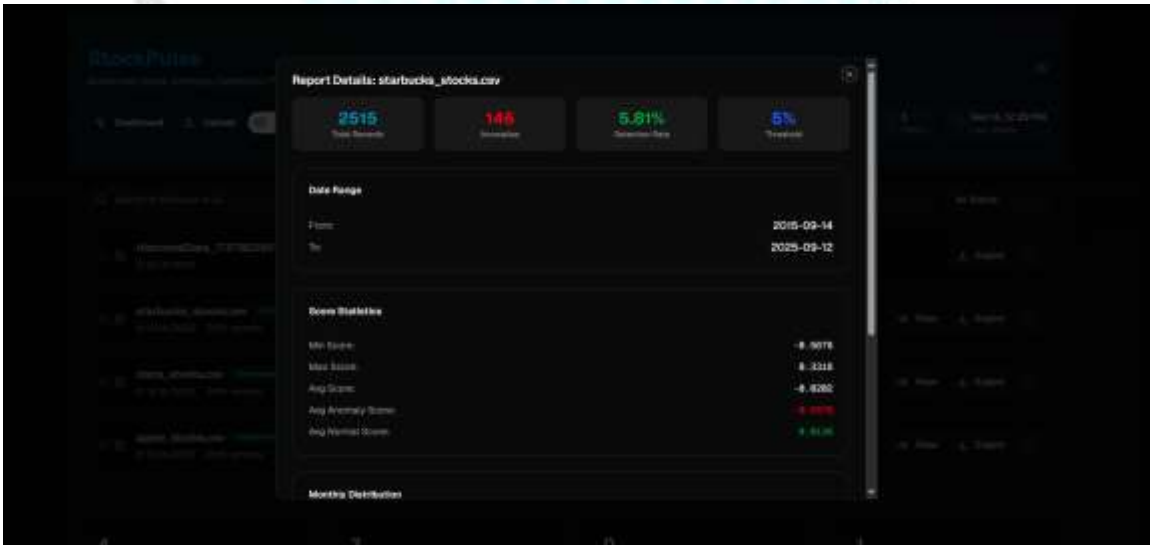




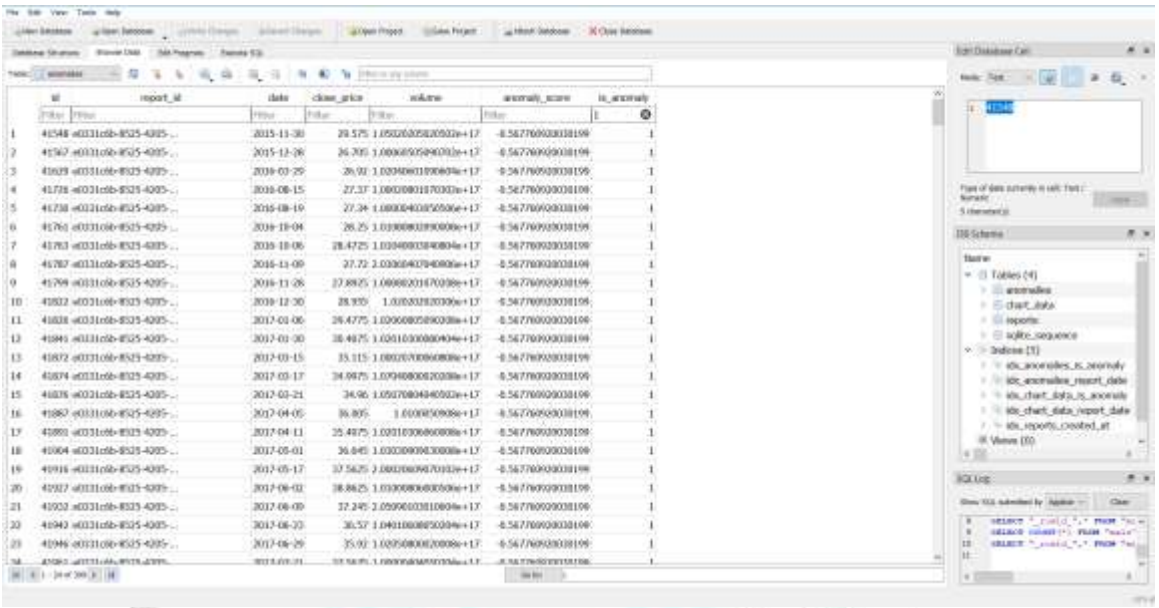

StockPulse
Dashboard
Upload
Reports
Quick Stats
Anomaly Detection Results
Export

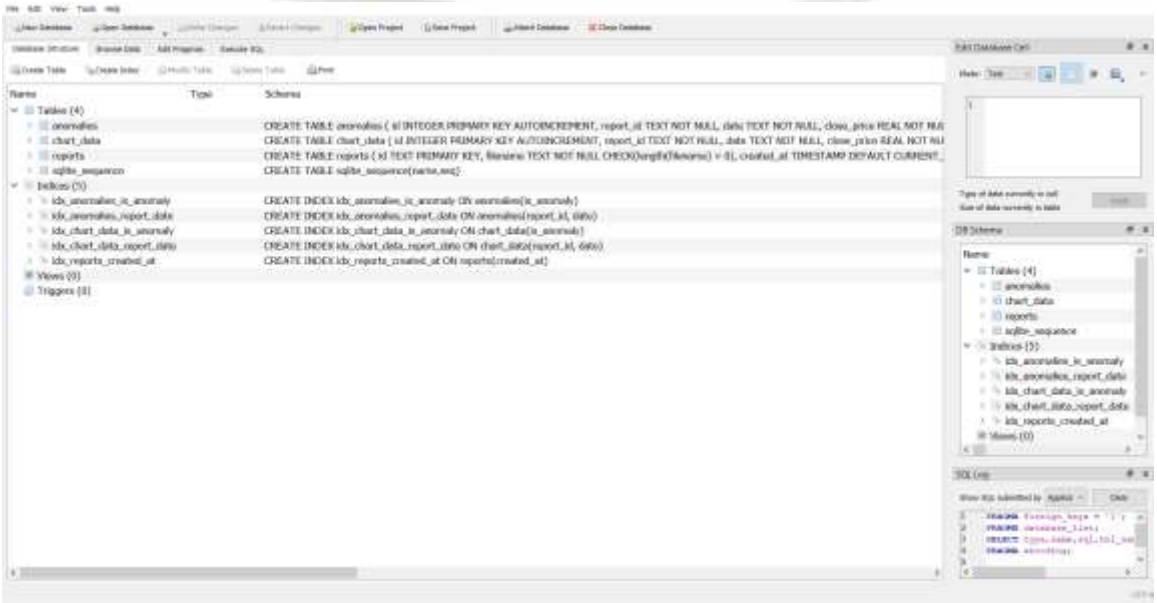

Score Statistics
Monthly Distribution


StockPulse
Upload Stock Data
apple_stocks.csv
CSV Format Requirements
Start Analysis
Clear


Report Details: starbucks_stocks.csv
2515
145
5.81%
5%
Date Range
2015-09-14
2025-09-12

SCREENSHOTS OF DATABASE





Chapter 7: Conclusion and Future Work:

7.1 Introduction

This chapter provides the overall summary of this research survey and also the future works that can be done to enhance the proposed findings.

7.2 Conclusion

The **StockPulse** project demonstrates the potential of data science and machine learning in addressing one of the most critical challenges in financial markets—detecting anomalies in stock price movements. By leveraging historical and real-time stock data, the system applies advanced preprocessing techniques, feature engineering, and anomaly detection algorithms to identify irregularities such as sudden spikes, dips, or unusual patterns.

Through interactive dashboards and automated reporting, **StockPulse** equips analysts, investors, and regulatory authorities with actionable insights that can help reduce risks, detect potential fraud or manipulation, and maintain overall market integrity. By combining scalable architecture, intuitive visualization tools, and user-friendly interfaces, **StockPulse** highlights how technology can make financial analysis more efficient, transparent, and reliable.

In conclusion, **StockPulse** represents a step toward intelligent anomaly detection systems that empower stakeholders to make data-driven decisions in increasingly complex and fast-moving financial environments.

The journey is just beginning, and with continued dedication and collaboration, we can make significant strides toward a healthier planet.

7.3 Future Work

As the Stock Pulse system continues to evolve several areas of improvement and expansion can further enhance its effectiveness:

- Integration of Advanced AI Models
  Incorporating deep learning models such as LSTM (Long Short-Term Memory) networks and Transformers for time series analysis can improve detection of complex patterns and rare anomalies.
- Real-Time Monitoring and Alerts
  Adding real-time anomaly detection with instant alerts would allow investors and regulators to respond quickly to suspicious market behavior.
- Mobile Application Development
  Building a mobile app version of Stock Pulse would make anomaly detection insights more accessible to investors and traders on the go.
- Expanded Data Sources
  Integrating additional financial indicators such as economic data, social media sentiment, or news feeds could provide more context for anomalies.
- Collaboration with Financial Institutions
  Partnering with stock exchanges, banks, and regulatory bodies would help validate the system and enable large-scale adoption in the financial industry.

This documentation concludes the project details, findings, and outcomes.
We hope this serves as a valuable resource for future reference.

"Thank you for reviewing this documentation.".