9/15/2025

# Team Name

# SFC - DataWeavers

# Team Members:

|  |
| --- |
| Mr. MUHAMMAD USMAN |
| Mr. AMIR KHAN |
| Mr. MUHAMMAD IBTESAM ARIF |
| Mr. YOUSUF YOUHANA |

StockPulse

Contents

[Background and Necessity for the Application: 3](#_Toc177867325)

[Proposed Solution: 4](#_Toc177867326)

[Proposed of the Document: 5](#_Toc177867327)

[Scope Of Project: 7](#_Toc177867328)

[Contraints: 8](#_Toc177867329)

[Functional Requirements: 10](#_Toc177867330)

[Non-Functional Requirements: 12](#_Toc177867331)

[Interface Requirements: 13](#_Toc177867332)

[Hardware: 13](#_Toc177867332)

[Software: 13](#_Toc177867332)

[Project Deliverables: 13](#_Toc177867332)

[Future Enhancements: 13](#_Toc177867332)

# 

# 1. Background and Necessity for the Application:

Stock prices and financial time series are highly volatile and complex, making anomalies—such as sudden spikes, drops, or irregular patterns—hard to detect manually. These anomalies may indicate market manipulation, system errors, or economic shifts. Detecting them is crucial for analysts, regulators, and investors to manage risk and maintain market integrity. Traditional methods often fail to capture complex patterns, creating the need for a **data science-driven solution** that can identify anomalies efficiently and provide actionable insights for informed decision-making.

# 2. Proposed Solution:

The **StockPulse** application aims to detect anomalies in stock price data by identifying unusual patterns, spikes, or dips that deviate from expected trends. The solution involves:

1. **Data Collection:** Gather historical and real-time stock data (e.g., via Yahoo Finance or Alpha Vantage), including Open, High, Low, Close, and Volume.
2. **Exploratory Data Analysis (EDA):** Visualize trends, volatility, and seasonality using plots and moving averages to detect early signs of anomalies.
3. **Data Preprocessing:** Clean missing values, standardize data, and ensure chronological order for accurate analysis.
4. **Feature Engineering & Selection:** Create derived metrics like returns, moving averages, and momentum indicators; select relevant features to reduce noise.
5. **Anomaly Detection:** Apply unsupervised algorithms such as Isolation Forest or SVM to identify outliers without labeled data.
6. **Model Evaluation:** Use precision, recall, F1-score, and visual inspection to assess detection accuracy.
7. **Visualization & Reporting:** Present results via an interactive dashboard, highlighting anomalies for actionable insights.

# 3. Purpose of the Document:

This document defines the **objectives, features, and constraints** of the StockPulse application. It serves as a reference for **stakeholders and developers**, detailing the functional and non-functional requirements, system design, and workflow for detecting anomalies in stock price data. The goal is to ensure a clear understanding of the project scope, implementation, and expected outcomes

# 4. Scope of the Project:

The scope of the StockPulse project is to develop a **data science-driven application for detecting anomalies in financial time series data**, specifically focusing on stock price movements. The key objectives and boundaries of the project include:

1. **Data Handling:** Collect historical and real-time stock data from reliable sources (e.g., Yahoo Finance, Alpha Vantage), including features such as Date, Open, High, Low, Close, Adjusted Close, and Volume.
2. **Data Analysis and Preprocessing:**
   * Explore trends, volatility, and seasonality using Exploratory Data Analysis (EDA).
   * Clean and preprocess data by handling missing values, formatting dates, sorting chronologically, and standardizing numerical features.
3. **Feature Engineering and Selection:**
   * Generate derived features such as returns, moving averages, momentum, and volatility indicators to provide richer input for anomaly detection.
   * Use statistical techniques to select relevant features, reducing noise and improving model efficiency.
4. **Anomaly Detection:**
   * Apply unsupervised machine learning algorithms such as Isolation Forest or Support Vector Machine (SVM) to identify unusual stock behaviors.
   * Evaluate model performance using metrics like precision, recall, and F1-score, supplemented by visual inspection of flagged anomalies.
5. **Visualization and User Interaction:**
   * Develop an intuitive dashboard to display trends, time series plots, and detected anomalies.
   * Provide a simple user interface to upload datasets, initiate analysis, and interact with visual outputs.
6. **Project Boundaries and Limitations:**
   * The system is designed for **offline analysis**; it does not include real-time anomaly detection or automated trading features.
   * Focus is on providing **insights and alerts** rather than executing trades or making investment decisions.
7. **Deliverables:**
   * Reports, visual dashboards, and anomaly detection results that support **data-driven decision-making**.

# 5.Contraints:

The StockPulse project is subject to the following constraints that define its operational and technical limitations:

1. **Data Assumptions:**
   * Financial time series data is inherently **volatile and non-stationary**, meaning assumptions of data stability may not hold.
   * Sudden market events or irregularities may affect the consistency of the data.
2. **Analysis Scope:**
   * The system is designed for **offline analysis** only. Real-time detection or automated trading functionality is **not included**.
   * Anomalies detected are intended for **informational and decision-support purposes**, not for executing trades automatically.
3. **Computational Resources:**
   * The algorithms must run efficiently within the constraints of typical hardware (e.g., 8 GB RAM, Intel i5/i7 processor).
   * Processing extremely large datasets may require additional resources or optimizations.
4. **Visualization Limitations:**
   * The dashboard provides basic visualization and interaction features.
   * Advanced or highly interactive real-time dashboards are **out of scope** for this project.
5. **Algorithmic Constraints:**
   * Unsupervised anomaly detection methods may produce **false positives or negatives** due to the complexity of stock market patterns.
   * Parameter tuning is necessary for optimal performance, but perfect detection accuracy cannot be guaranteed.

# 6. Functional Requirements:

The StockPulse application must provide the following functionalities to effectively detect anomalies in stock price data:

1. **Data Ingestion and Management:**
   * Import historical stock market data (CSV or standard formats).
   * Organize data chronologically for accurate time series analysis.
2. **Data Preprocessing:**
   * Handle missing values, inconsistencies, and formatting issues.
   * Normalize or standardize data to ensure uniform scaling for analysis.
3. **Feature Engineering:**
   * Compute derived metrics such as returns, moving averages, volatility indicators, and momentum measures.
   * Enrich the dataset to improve anomaly detection accuracy.
4. **Feature Selection:**
   * Identify and retain relevant features using statistical techniques (e.g., correlation analysis, variance thresholding).
   * Reduce noise and dimensionality for efficient model performance.
5. **Data Partitioning:**
   * Split datasets into training and testing sets.
   * Use the training set to learn normal behavior patterns and the testing set to evaluate anomalies.
6. **Anomaly Detection Module:**
   * Implement unsupervised algorithms such as Isolation Forest or Support Vector Machine (SVM).
   * Allow configuration and tuning of algorithm parameters for optimal detection.
7. **Model Evaluation:**
   * Evaluate performance using metrics like **precision, recall, and F1-score**.
   * Enable visual inspection of flagged anomalies for validation.
8. **User Interface (UI) and Interaction:**
   * Provide a simple interface for uploading datasets, running analysis, and viewing results.
   * Enable easy navigation and interaction with visual outputs.
9. **Dashboard Visualization:**
   * Generate interactive dashboards showing key trends, time series plots, and detected anomalies.
   * Provide clear, actionable insights for users.
10. **Export and Reporting:**
    * Export anomaly reports and visual summaries for documentation or further analysis.

# 7. Non-Functional Requirements:

The StockPulse application must fulfill the following non-functional requirements to ensure reliability, performance, and usability:

1. **Usability:**
   * The system should have a clear, intuitive interface that allows users to easily upload datasets, run analyses, and interpret results.
2. **Scalability:**
   * The system should efficiently handle large volumes of financial time series data without significant performance degradation.
   * It should be capable of scaling to accommodate growing datasets and additional features.
3. **Accuracy and Reliability:**
   * The application must provide consistent and dependable results in anomaly detection.
   * False positives and false negatives should be minimized to ensure trustworthy insights.
4. **Compatibility:**
   * The application must run on common operating systems (Windows, macOS, Linux) and support standard data formats such as CSV.
   * It should integrate smoothly with external visualization tools like Tableau for generating dashboards.
5. **Maintainability:**
   * The system should be easy to update, maintain, and extend, allowing improvements in algorithms or addition of new features.
6. **Performance:**
   * Analysis and visualization processes should be completed in a reasonable time, even for large datasets, ensuring timely insights.

# 8. Interface Requirements:

The StockPulse application requires both **hardware and software interfaces** to operate effectively.

### **8.1 Hardware Requirements**

* **Processor:** Intel Core i5/i7 or higher
* **RAM:** 8 GB or higher
* **Storage:** 500 GB Hard Disk
* **Display:** Color SVGA monitor
* **Input Devices:** Mouse and keyboard

### **8.2 Software Requirements**

* **Data Storage:** CSV files
* **Backend:** Apache Spark or Apache Hive; Flask or Django for web services
* **Database:** MongoDB or MySQL
* **Programming / IDE:** Python 3.11+ or R; Jupyter Notebook, Anaconda 23.1+, or Google Colab
* **Libraries:** Pandas, NumPy, scikit-learn, TensorFlow, PyTorch, OpenCV, Matplotlib, Seaborn
* **Visualization Tools:** Tableau Desktop or other interactive visualization platfor

# 9. Project Deliverables:

The StockPulse project will produce the following deliverables:

1. **Source Code:**
   * Complete implementation of the project including Jupyter Notebook (.ipynb) or Google Colab files.
   * All scripts, modules, and configuration files uploaded to a **public GitHub repository** with proper access.
2. **Project Documentation:**
   * Problem definition and objectives
   * Design specifications and architecture
   * User flow diagrams or journey maps
   * Test data used and sample results
   * Installation and execution instructions
   * Assumptions, limitations, and notes in a ReadMe file
3. **Visualization Dashboards:**
   * Interactive dashboards displaying stock trends and detected anomalies
   * Charts and graphs for clear interpretation of results
4. **Blog Publication:**
   * A detailed blog (2000 words) explaining the project, methodology, and outcomes
   * Published on a free platform (Blogger, Tumblr, Ghost, etc.) with a public link
5. **Demonstration Video:**
   * A mandatory video (.mp4) showcasing the working application and all functionalities
6. **Reports and Exported Results:**
   * Exported anomaly reports and visual summaries for analysis and documentation purposes

# 10. Future Enhancements:

In future versions, StockPulse can be enhanced to support **real-time anomaly detection** with instant notifications via email or SMS. Integration with trading platforms could allow automated or semi-automated responses to unusual stock behavior. Advanced machine learning models, additional data sources such as news or social sentiment, and more interactive dashboards can improve detection accuracy and usability. A mobile application and personalized user settings would further enable on-the-go monitoring and a tailored experience for investors and analysts.