PUSL3190 Computing Project

Stock Market Trend Prediction System using Machine Learning

Name: Nethmi Matharaarachchi
Degree Programme: BSc (Hons) Software Engineering
Affiliated University: University of Plymouth, UK
Local University: NSBM Green University, Sri Lanka
Submission Deadline: 31st October 2025 – 4.00 PM (Sri Lankan Time)
Submission Mode: PDF Report Submission
File Name: 10953053_Proposal.pdf

Chapter 01 – Problem Statement

The stock market is among the most vivacious and powerful factors of the world economy, determining not only the market but also business strategies, investment choices, and sometimes even the government's decisions. The capacity to predict market trends with high accuracy can deliver great insights that facilitate investors in risk reduction and rational decision-making. However, on the other hand, stock price movements are unpredictable and erratic, and are affected by many external factors like political unrest, natural disasters, and the moods of the investors.

In Sri Lanka, the Colombo Stock Exchange (CSE) acts as the main trading center for the country, and it connects various industries' companies to investors. The number of participants in the market has grown, but still, the majority of local investors make their decisions based on manual chart analysis, gut feeling, or media commentary. This way of thinking can lead to making mistakes and letting emotions take control, which in turn leads to either losing money or not being able to take advantage of a forthcoming opportunity.

The international stock trading and forecasting systems available today—Yahoo Finance, TradingView, Investing.com—are all great but none of them are specifically adapted to the conditions of the Sri Lankan market. They are often lacking on the data integration side from the CSE, require a subscription that is expensive, and not all investors require it, and lastly, they do not provide a proper multi-period analysis such as daily, weekly, monthly, and yearly visualizations which are relevant to the local investor.

 This means that there is an urgent requirement for a system that is intelligent, data-driven, and oriented toward the local market to identify the past CSE market data, analyze them up and down, and above all use advanced machine learning algorithms to tell the direction of the future price trends. The project is focused on creating a gap between the Stock Market Trend Prediction System using Machine Learning and the market itself, which gathers historical data from the Colombo Stock Exchange through Alpha Vantage API and EODHD API, and provides interactive daily, weekly, monthly, and yearly visualizations.

Chapter 02 – Project Description

2.1 Overview

The Stock Market Trend Prediction System is an analytical platform that is web-based, and it is specially designed to carry out the prediction and visualization of the performance of the Sri Lankan stocks listed on the Colombo Stock Exchange. The system will make use of

the machine learning methods and the historical financial data to predict the future price movements and the trend reversals, as well as the volatility of the stocks.

Alpha Vantage API and EODHD API will be the means through which the collection of data will be done, and they will provide structured access to the CSE market data which includes OHLC prices and trading volumes. This data will be collected and then, it will be processed to create time-series forecasts and multi-level analyses that will cover daily, weekly, monthly, and yearly perspectives for a one-year period.

The application will provide:

- Interactive visualizations of historical and predicted stock trends. Comparative charts display short-term and long-term market behavior.
- Analytical indicators such as moving averages, volatility index, and trading signals.
- A clean, intuitive dashboard allows users to filter data and generate personalized reports.

This system will not only make investment decision-making easier but also facilitate academic research, financial education, and policy analysis in Sri Lanka.

2.2 Project Objectives

Main Objective:

The primary goal of this project is to create and implement an advanced web-based predictive analytics system based on machine learning and historical datasets of the Colombo Stock Exchange with a very high level of accuracy.

 The system will be able to produce accurate predictions of market trends using different time intervals, thanks to the combination of trustworthy APIs like Alpha Vantage for their strong intraday and real-time functionality and EODHD for their vast end-of-day historical data. Among the different temporal scales, the system will provide very precise daily predictions for short-term trading, weekly forecasts for momentum assessment, monthly predictions for portfolio rebalancing, and yearly estimates for strategic planning.

Specific Objectives:

The specific objectives outline a well-organized way to achieve the main goal of the project, which is to develop a robust stock prediction system for the Colombo Stock Exchange (CSE). Each objective is a step in the building of the project, from data foundation to deployment, which assures both theoretical and practical aspects of technology.

1. Data Gathering and Preprocessing: During this first stage, the focus is on getting historical datasets that are comprehensive and that stretch far back in time CSE daily OHLC pricing for the most detailed insights, weekly aggregates for detecting trends, monthly summarization for seasonal patterns, and yearly overviews for the identification of long-term cycles through the Alpha Vantage API (for real-time efficiency) and the EODHD API (for archival depth). Preprocessing involves the detection and removal of anomalies, filling of gaps, and feature scaling to produce a corpus that is trustworthy while at the same time reducing the biases that are usually found in unstable markets.

2. Model Creation and Development: The renowned machine learning models are here: among them are Linear Regression, which serves the purpose of drawing simple trend lines; Random Forest, which takes care of the intonation among the features and feature interactions; and LSTM, which precisely talks about the capturing of temporal nonlinearities in time-series data. The models' training through cross-validation is a continual process that ensures they are in tune with CSE's peculiar volatility caused by, for example, shifts in economic policy, and thus produce forecasts that have a high degree of accuracy.

3. Interactive Dashboard Presentation: The use of an interactive dashboard will provide a visual representation of the data through a user-driven interface. This process can be facilitated by tools like Plotly. The historical charts that are depicted would have the predictions on top of them, thus allowing a seamless exploration of data. There are filters that can be used to customize the data presentation (for instance, zooming in on a view that is weekly) and thus the data which was complex is turned into easy stories and the users are helped to navigate through the data confidently.

4. Generation of Interpretable Insights: The models create effective inputs like RSI-based buy/sell alerts or Bollinger-derived volatility gauges, which are then linked to users through natural language summaries and confidence bands that make the interface less algorithmic and more user friendly.

5. Performance Analysis: A comprehensive assessment utilizes RMSE for error magnitude, MAE for absolute deviations, and R^2 for variance explanation, comparing against baselines to improve accuracy and reliability.

6. Scalable Platform Development: The last goal creates a web application that is cloud-hosted, responsive, and has a user-friendly navigation system, accessibility features, and a modular design. This not only provides access to students and investors but also encourages financial inclusion in Sri Lanka by making the access to such investments easier.

2.3 Project Keywords

- Machine Learning
- Stock Market Prediction
- Time-Series Analysis
- Colombo Stock Exchange
- Data Visualization

Chapter 03 - Research Gap

The global financial landscape has been changed dramatically and for the better with machine learning, but Sri Lankan-specific implementations are still not at all researched. The majority of the available tools and studies are primarily focused on the high-volume developed markets and thus do not address the particular features of the emerging markets like the CSE which has lower liquidity, restricted data granularity, and special trading patterns.

Existing Research and Systems:

1. The research of Rahul et al. (2022) presented a model based on LSTM for the prediction of NASDAQ stocks which was merely accurate due to the ensemble of huge datasets that were eliminated for the CSE market.
2. Kumara et al. (2021) used linear regression for the prediction of Sri Lankan stocks, however, the use of scarce feature selection led to the prediction accuracy being quite low.
3. Senanayake and Perera (2020) employed ARIMA models to interpret CSE data but could not deal with the situation when sudden market shocks occurred.
4. TradingView and Yahoo Finance, for example, are international platforms that provide sophisticated visualizations, but they are not connected to the CSE locally and cannot be adjusted according to one-year interval trend studies.

Research Limitations Identified:

- No multi-frequency trend analysis (from daily to yearly) has ever been attempted.
- The connection between CSE and international APIs is too weak.
- Very little visual interpretability for small or emerging markets has been achieved.

Proposed Contribution:

This project presents a new hybrid approach that consists of:

- CSE historical data integration through Alpha Vantage and EODHD APIs.
- Multi-level time-series forecasting (daily, weekly, monthly, yearly).
- An advanced visualization dashboard for both historical and predicted trends.

By fusing machine learning prediction with interactive visual analytics, the system creates a significant breakthrough in the local, available, and smart stock forecasting tools for the Sri Lankan financial market.

Chapter 04 – Requirements Analysis

Hardware Requirement Analysis

- Laptop (16 GB RAM, i5 processor)

Software Requirement Analysis

Category	Technologies
Programming Language	Python, JavaScript
ML Libraries	Scikit-learn, TensorFlow/Keras, Pandas, NumPy
Backend Framework	Flask / Django
Frontend Framework	React.js / Vue.js
Database	MySQL / PostgreSQL
Visualization	Plotly, Chart.js, Matplotlib
APIs	Alpha Vantage API, EODHD API
Deployment	Render, Heroku, Firebase
Version Control	Git & GitHub

Knowledge and Skills Required

- Machine Learning: Time-series forecasting, regression models, feature engineering.
- Web Development: REST API integration, frontend/backend communication, deployment.
- Data Science: Data cleaning, normalization, and exploratory analysis.
- Finance: Understanding stock metrics (Open, High, Low, Close, Volume).

Data Description

The dataset will contain historical financial data for selected CSE-listed companies over a one-year period, extracted from Alpha Vantage and EODHD APIs.

Attributes include:

- Date, Open, High, Low, Close prices
- Adjusted close
- Trading volume

The system will aggregate and analyze this data at daily, weekly, monthly, and yearly intervals, generating visual trend charts for each time scale.

Chapter 05 – Finance

The project is conducted in a virtual environment; there are no costs related to physical aspects. The collection of all data is done through APIs and cloud tools, spending just 3,000 LKR for access of the dataset.

Chapter 06 – External Organizations

Data Sources

- Colombo Stock Exchange (CSE), Primary financial and historical trading data.
- Alpha Vantage API: Provides global and local sets of financial data with JSON access.
- EODHD API: Provides end-of-day and intraday historical data of CSE-listed companies.

Support and Collaboration

- Supervising Lecturer: Ms. Nimesha Hewawasama
- Department: Software Engineering, NSBM Green University

- External Consultants: Mr. Pasindu Oshadha, Senior Software Engineer in IGT 01 Lanka (PVT) Ltd

Chapter 07 – Timeline

Phase	Activity	Duration	Timeline
1	Project Approval and Proposal Preparation	2 weeks	October 2025
2	Data Collection	1 weeks	November 2025
3	Data Preprocessing & Feature Engineering	3 weeks	November 2025
4	Model Development (Regression, Random Forest, LSTM)	4 weeks	December 2025
5	System Implementation (Web Dashboard + Backend API)	4 weeks	January 2026
6	Chart Visualisation (Daily Yearly Analysis)	3 weeks	February 2026
7	Testing, Evaluation & Performance Tuning	3 weeks	Fed – Mar 2026
8	Final Report Writing	2 weeks	April 2026

References

1. Kumara, A. et al. (2021). *Predicting Colombo Stock Exchange Movements Using Linear Regression.* University of Colombo Journal of Computing.

2. Senanayake, R. & Perera, D. (2020). *ARIMA-Based Forecasting for CSE Market Data.* IEEE Conference on Computational Intelligence.

3. Rahul, S. et al. (2022). *Stock Market Prediction Using LSTM Networks.* International Journal of Machine Learning Applications.

4. Colombo Stock Exchange. (2024). *Official Data Portal.* Available at: https://www.cse.lk/

5. Alpha Vantage. (2024). *Global Stock Market API.* Available at: https://www.alphavantage.co/

6. EOD Historical Data (EODHD). (2024). *Market Data API for Financial Research.* Available at: https://eodhd.com/

7. Yahoo Finance. (2024). *Historical Data and Charts.* Available at: https://finance.yahoo.com/

8. J. Brown. (2023). *Machine Learning in Financial Forecasting.* Springer Publications.

9. World Bank. (2024). *Sri Lanka Economic Update.* Retrieved from https://www.worldbank.org/