StockSensei - Intelligent Stock Market Forecasting Tool

TEAM MEMBERS:

1. Kanak Jaiswal
2. Mansi Mantri
3. Muskaan Mahajan
4. Niraj Mohabey
5. Vaibhav Chavan

PROJECT FUNCTION:

StockSensei is a sophisticated forecasting platform designed to generate daily, short-term stock market predictions for up to 30 days. By utilizing time series analysis models—ARIMA, SARIMA, and Prophet—StockSensei allows users to view forecasted stock trends with a high degree of accuracy. The tool also enables users to compare model performance, assisting in selecting the optimal model for specific stocks. With a user-centric interface and automated visualizations, *StockSensei* makes advanced stock market analysis accessible to investors of all experience levels.

PURPOSE:

The volatility of stock markets makes accurate prediction challenging yet invaluable. *StockSensei* provides actionable insights into stock trends, enabling informed investment decisions by leveraging statistically sound forecasting models.

TARGET AUDIENCE:

- **Individual Investors** seeking to enhance decision-making with accessible forecasting tools.
- **Financial Analysts** who benefit from model-based forecasts for market analysis.
- **Data Science Professionals and Researchers** interested in applying time series analysis to financial data.

EXISTING TOOLS:

While platforms like Bloomberg Terminal and Yahoo Finance offer financial insights, they often lack customizable, multi-model forecasting options. These tools may be complex or generalized, presenting barriers to non-specialist users.

DIFFERENTIATORS FOR STOCKSENSEI:

- **Integrated Model Comparison**: *StockSensei* is distinctive in offering ARIMA, SARIMA, and Prophet models within a single environment, with error metrics such as RMSE and MAE to enable effective model evaluation.
- **User-Friendly Interface**: Designed with accessibility in mind, *StockSensei* minimizes technical complexity, providing a seamless experience for those with limited financial modelling knowledge.
- **Automated Analysis and Visualization**: Through automated error calculations and intuitive visualizations, *StockSensei* provides fast, clear insights that reduce the need for complex setups.

DEVELOPMENT CHALLENGES

The primary challenges in developing *StockSensei* include:

- **Data Quality and Processing**: Ensuring access to reliable, high-quality stock data is essential. Addressing missing values and data anomalies is crucial to maintain model integrity.
- **Model Optimization**: Fine-tuning parameters for ARIMA and SARIMA across various stocks demands extensive calibration and testing.
- **Interactive Visualizations**: Implementing interactive visuals with Plotly while ensuring ease of use requires efficient, user-friendly integration.

DEVELOPMENT PLAN

DATA SOURCE: *StockSensei* will rely on historical stock data stored in a PostgreSQL database, structured to accommodate data such as date, closing price, and stock symbol. This setup will facilitate efficient data management and retrieval, streamlining the forecasting process.

CORE ALGORITHMS:

- **ARIMA** and **SARIMA** models will be implemented to capture the dependencies in time-series data, facilitated by Python's *statsmodels* library.
- **Prophet Model**: Facebook's Prophet model will address seasonal trends in stock data, adding robustness to the predictions.

IMPLEMENTATION STACK:

- **SQLAlchemy** for database interactions,
- **Pandas and NumPy** for data manipulation,
- **Statsmodels and Prophet libraries** for time series modelling,
- **Plotly** for generating interactive visualizations.

CURRENT RESOURCES

StockSensei will utilize a combination of open-source resources:

- **Prophet** for seasonal forecasting,
- **Statsmodels** for time series implementation,
- **Plotly** for dynamic graphing capabilities,
- **SQLAlchemy** for efficient database management.

DEMONSTRATING EFFECTIVENESS

To illustrate *StockSensei*'s utility, it will be tested on a representative set of stocks. Each model (ARIMA, SARIMA, Prophet) will produce forecasts that are validated against real market data, with RMSE and MAE metrics highlighting each model's accuracy. A user-friendly interface will allow stock selection, historical data visualization, and interactive forecasting displays, showcasing *StockSensei*'s practicality for investment analysis.

In providing a sophisticated yet accessible tool for stock forecasting, *StockSensei* bridges a gap in the financial technology space, supporting investors and analysts with robust, model-driven insights.

LITERATURE REVIEW

The paper "An Introductory Study on Time Series Modeling and Forecasting" offers an in-depth look at various methods for predicting stock market trends, with a focus on how machine learning models can improve forecast accuracy. It reviews key studies on regression techniques, neural networks, and hybrid models, covering methods like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Long Short-Term Memory (LSTM) networks, and Artificial Neural Networks (ANNs), and assesses each model's strengths and weaknesses in the context of financial forecasting. The study also highlights the usefulness of traditional models like ARIMA, GARCH, and ARCH in handling volatility, while emphasizing that machine learning techniques show promising results in enhancing prediction accuracy, especially in the complex, ever-changing landscape of financial data. The paper “Bitcoin Forecasting Using ARIMA and PROPHET” explores recent developments in deep learning for financial time series forecasting, analysing models like Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and hybrid approaches for their effectiveness in stock market prediction. By comparing deep learning models with traditional forecasting techniques, the paper demonstrates significant improvements in accuracy and efficiency, particularly in volatile markets. It also evaluates models such as Support Vector Machines (SVM), K-Nearest Neighbour’s (KNN), and Artificial Neural Networks (ANN), finding that while ANN achieved an accuracy of 88%, SVM outperformed it with 96.15% accuracy, making SVM the top performer for stock price predictions. The study further highlights the limitations of LSTM and Random Forest for this application. The paper “An Introductory Study on Time Series Modelling and Forecasting” explores recent advancements in deep learning techniques for financial time series forecasting, assessing models such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and hybrid approaches. It focuses on the effectiveness of these models in capturing complex patterns in stock market trends, especially in volatile environments. Through a comparison with traditional forecasting methods, the study demonstrates that deep learning models, particularly LSTM and CNN, offer improvements in accuracy and adaptability, making them more effective for financial forecasting in dynamic markets. The study compares ARIMA, Prophet, and deep learning models (LSTM and CNN-LSTM) for forecasting weekly wholesale food prices from 2013 to early 2020. Prophet provided the quickest setup but showed the lowest accuracy, with a mean absolute percentage error (MAPE) of 3.54%, close to the no-change model’s 3.77%. ARIMA achieved moderate accuracy with a MAPE of 2.49%. The CNN-LSTM model, while computationally intensive, attained the best accuracy with an average MAPE of 2.28% across products, reflecting its suitability for precise forecasts. These results highlight CNN-LSTM’s potential for high-accuracy applications and Prophet’s utility for faster, preliminary predictions, supporting more automated pricing strategies. The research paper “Stock price prediction using the RNN model” investigates product prediction using Recurrent Neural Networks (RNN) to identify product trends that change over time. RNN is particularly suitable for sequential data because it can use historical data to predict future outcomes. The learned data preprocessing includes normalizing the data value to the range [0,1] for good modelling. The researchers previously generated input and output data using historical data of Apple stock (AAPL) with time steps of 5 and 10 days, respectively, which is important for predicting the next day’s stock price. The RNN model was implemented with two hidden layers of 50 and 100 nodes and optimized using Adams algorithm with mean square error (MSE) as a function of failure. The results show that the model performs better at smaller time steps, with lower mean absolute error (MAE) and root mean square error (RMSE) of time step = 5 compared to time step = 10. Short-term stock price accuracy decreases as the data progresses, meaning that models need to be further developed for long-term forecasts.

REFERENCES

1. Ratnadip Adhikari, R. K. Agrawal. (2013). **An Introductory Study on Time Series Modelling and Forecasting**.
2. IJRASET. Prof. Gowrishankar B S, Pooja V, Girichandra R Karaveermath, Guruprasad M, Swaroop R Hosamani (2021). **Stock Market Prediction using Machine Learning.**
3. IEEE Xplore. Işil Yenidoğan, Aykut Çayir, Ozan Kozan, Tuğçe Dağ, Çiğdem Arslan (2018). **Bitcoin Forecasting Using ARIMA and PROPHET.**
4. Forecasting 2021, 3(3), 644-662, Menculini, L., Marini, A., Proietti, M., Garinei, A., Bozza, A., Moretti, C., & Marconi, M. (2021). **Comparing Prophet and Deep Learning to ARIMA in Forecasting Wholesale Food Prices.**
5. International Conference on Applied Physics and Computing (ICAPC 2020), Yongqiong Zhu, **Stock price prediction using the RNN model.**