11/24/2024

Antonczak, Marcin

S00183460

**Amazon stock price analysis**

Data Analytics and Visualisations (COMP09014)

Table of Contents

[1 Introduction 2](#_Toc183384810)

[2 Amazon stock price dynamics 2](#_Toc183384811)

[2.1 Historical performance 2](#_Toc183384812)

[3 Existing tools for stock analysis 3](#_Toc183384813)

[4 Machine learning approach for stock price predictions 3](#_Toc183384814)

[4.1 Overview 3](#_Toc183384815)

[4.2 Gathering data 4](#_Toc183384816)

[4.3 Data preprocessing 5](#_Toc183384817)

[4.4 Feature selection 5](#_Toc183384818)

[4.5 Choosing the model 6](#_Toc183384819)

[4.6 Model training 6](#_Toc183384820)

[4.7 Model evaluation 6](#_Toc183384821)

[5 Conclusion 7](#_Toc183384822)

[6 References 7](#_Toc183384823)

# Introduction

Amazon Inc., established in 1994 and headquartered in Seattle, Washington, is a global leader in online retail, technology, and media services. Its’ diverse product range includes electronic devices, streaming platform membership, advertising programs and IT services related to storage and analytics. Amazon also enables third-party sellers, authors, and content creators to publish and sell products and services. (Investing.com, 2007-2024)

Amazon is the company investors and customers pay attention to. It hugely influences various markets and is unafraid to take risks in new endeavours. As the company continues to grow, understanding its stock price dynamics has become increasingly important for investors and analysts who rely on data-driven tools to make informed decisions. (Johnson, 2024)

Data visualisation has become an essential tool for analysing complex financial data. It allows organisations to transform large datasets into intuitive visuals, enabling efficient analysis, precise decision-making, and time-saving processes. By doing so, businesses can gain deeper insights into market trends and predict outcomes more effectively. These visualisations are especially useful for financial analysis. Moreover, dashboards that integrate machine learning models offer enhanced predictive capabilities, making them valuable. (BLUENTTECH, 2023)

This literature review explores existing approaches to analysing Amazon's stock price, identifies gaps in current models, and examines how machine learning techniques could enhance prediction accuracy, thus contributing to more effective investment strategies.

# Amazon stock price dynamics

## Historical performance

The historical performance of Amazon's stock reflects its journey from an online bookstore to a global tech giant. After its initial public offering (IPO) in 1997, Amazon's early stock performance faced volatility as the company focused on growth over immediate profitability. By the early 2000s, Amazon reached profitability for the first time, with net profits of $35 million in 2003. This achievement followed the introduction of diverse product categories and innovative services.

A pivotal driver of Amazon’s sustained stock growth has been the launch and expansion of Amazon Prime in 2005, which boosted revenue streams by creating a loyal customer base. The subscription service became a cornerstone of Amazon’s growth strategy, particularly during the 2010s, with Prime-related spending outpacing that of non-members.

Amazon has achieved remarkable success and continues to seek growth. It is constantly working to improve its services as technology evolves. One of the company’s recent innovations includes drone and autonomous car deliveries. While the success of these programs remains uncertain, it is clear that Amazon is dedicated to exploring new technologies to make its delivery process even more efficient. The company’s efforts suggest a continued focus on staying ahead in technology while expanding its influence across industries. (Hopkins, 2023)

# Existing tools for stock analysis

There are several online platforms offering stock analysis. While these platforms have useful features, they also have significant limitations:

* Yahoo Finance and Google Finance -   
  Both Yahoo Finance and Google Finance provide free and convenient access to stock prices and basic charts. While they are excellent for real-time price updates, they lack advanced features for deeper technical analysis.
* StockCharts.com - delivers a comprehensive dashboard for displaying historical stock prices. However, the free version restricts users to viewing charts for only up to six months.

The worst thing is that none of these platforms offer machine learning-based price predictions. (Sham, 2020)

# Machine learning approach for stock price predictions

## Overview

Using machine learning algorithms and computational tools, we can discover hidden connections and patterns in large datasets. Machine learning models are capable of analysing a significant amount of historical data, selecting the most relevant features, and generating predictions based on those learned patterns. The main objective of machine learning is to develop models that can learn from existing data and predict outcomes for new or future data. This process involves two main steps: training and inference.

* Training: The algorithm is fed a dataset that includes both input data and the desired output (target). It learns the relationship between inputs and outputs, adjusting its parameters to reduce errors.
* Inference: Once trained, the model can analyse new data and make predictions, classify data into categories, or detect irregularities.

Machine learning techniques can be classified based on their purposes, such as regression (predicting continuous values), classification (categorizing data), or clustering (grouping similar items). In the field of stock price prediction, regression models are often employed. These models aim to forecast continuous outcomes, such as stock prices. By examining historical data and identifying trends, regression models can provide insights into future price movements. (Blanco, 2024)

## Gathering data

Collecting accurate and relevant data is a critical foundation for building effective machine learning models that aim to predict stock prices. Various data sources contribute to stock price predictions:

1. Historical Stock Data: This includes past records of stock prices (open, close, high, low, and trading volume).
2. Fundamental Data: This encompasses a company's financial metrics, such as revenue, earnings, debt, and assets, providing insights into its overall health and performance.
3. Market and Sentiment Data: Data about market dynamics, economic indicators, and sentiment from news or social media can enhance model accuracy by reflecting broader influences on stock prices.

Methods to Gather Data:

* Financial Databases: Resources like Bloomberg, Yahoo Finance, and Alpha Vantage offer access to historical prices, company fundamentals, and other market information.
* Web Scraping: Automated methods can extract information from financial news websites or social media platforms, aiding sentiment analysis.
* APIs: various services provide programmable access to a wide range of financial data for analysis.

## Data preprocessing

Data preprocessing is an essential stage in preparing data for analysis and machine learning model training. It involves cleaning, transforming, and structuring the data to ensure it meets the quality standards required for use with machine learning algorithms.

The main steps of data preprocessing include:

1. Data Cleaning: This step involves addressing missing values, outliers, and inconsistencies.
2. Data Transformation: This step focuses on converting data into a suitable format for analysis. It may involve scaling numerical features so they fall within a similar range (through normalization or standardization), encoding categorical variables into numerical formats (using methods like one-hot encoding), or applying mathematical transformations to improve the data's distribution properties.
3. Feature Selection: Feature selection identifies the most relevant and impactful features for the model.
4. Data Splitting: Before training the machine learning model, the dataset is divided into training, validation, and testing sets.

Careful execution of these preprocessing steps ensures the data is clean, relevant, and ready for modelling. Skipping or neglecting preprocessing can lead to inaccurate predictions, poor model performance, or biased results.

## Feature selection

Feature engineering is a vital phase in the machine learning process where new features are created or existing ones are transformed to enhance the model's predictive accuracy. By extracting relevant information from data, it plays a significant role in improving stock price prediction accuracy. The selection of features should be driven by domain knowledge, problem context, and the availability of relevant data. It is important to choose features that logically influence stock price movements.

## Choosing the model

Selecting the right machine learning model is essential for making accurate stock price predictions. Different algorithms offer distinct strengths and limitations, and the model selection depends on several factors, such as the complexity of the problem, dataset size and quality, interpretability, and available computational resources. Below are some widely used models for stock price forecasting:

* Linear Regression: This is a straightforward model that assumes a direct linear relationship between the input variables and the target. While effective for identifying simple trends and patterns, it may not perform well with more intricate data relationships.
* Support Vector Machines (SVM): SVM is a flexible model capable of handling both linear and non-linear relationships. It works by finding an optimal hyperplane to separate data into distinct categories or predict continuous values.

The choice of model should be guided by the problem’s complexity, the dataset's characteristics, and other practical considerations. of a model’s effectiveness.

## Model training

Training the machine learning model is a key phase in predicting stock prices. This step involves inputting the preprocessed data into the model and adjusting its internal parameters to recognize the underlying patterns and relationships within the data. The goal is to optimize the model by selecting the most effective combination of features. Throughout the training, it’s essential to monitor the model’s performance to prevent overfitting, as it can result in poor performance on new data.

## Model evaluation

Evaluating the trained model's performance is essential for understanding its accuracy and dependability in forecasting stock prices. This process helps assess the effectiveness and pinpoints areas where improvements can be made. It’s crucial to avoid evaluating the model only on training data, as this can provide misleading information on its performance in real-world situations. Instead, testing the model on unseen data offers a more accurate evaluation.

If the model's results are unsatisfactory, several strategies can be explored to enhance its predictive power. These may include tweaking hyperparameters, adding new features, experimenting with different algorithms, or expanding the dataset. Through consistent assessment and refinement, a stock price prediction model can be improved to deliver more precise and trustworthy forecasts. (Blanco, 2024)

# Conclusion

In summary, the literature reveals significant advancements and challenges in the application of machine learning for stock price prediction and analysis. Traditional platforms like Yahoo Finance, and StockCharts.com provide foundational tools for stock research but often lack advanced predictive features. These limitations highlight the need for more dynamic, integrated dashboards.

Machine learning introduces transformative possibilities in stock price forecasting. However, the inherent complexities of financial markets pose challenges, such as sensitivity to unpredictable shocks and difficulties in quantifying qualitative factors like geopolitical events or management decisions.

Despite these challenges, there is a potential for combining machine learning with advanced dashboards to create comprehensive tools for stock analysis. These systems could integrate historical data, technical indicators, and predictive insights into a unified interface, providing users with deeper, actionable insights into market dynamics.

# References

Blanco, N., 2024. *How To Use Machine Learning To Predict Stock Prices.* [Online]   
Available at: https://robots.net/fintech/how-to-use-machine-learning-to-predict-stock-prices/  
[Accessed 15 November 2024].

BLUENTTECH, 2023. *Data Visualization in Finance: Telling Better Stories and Helping Make Even Better Decisions.* [Online]   
Available at: https://www.bluent.com/blog/data-visualization-in-finance/  
[Accessed 15 November 2024].

Hopkins, C., 2023. *The History of Amazon and its Rise to Success.* [Online]   
Available at: https://sites.lsa.umich.edu/mje/2023/05/01/the-history-of-amazon-and-its-rise-to-success/  
[Accessed 15 November 2024].

Investing.com, 2007-2024. *Amazon.com Inc (AMZN).* [Online]   
Available at: https://www.investing.com/equities/amazon-com-inc-company-profile  
[Accessed 15 November 2024].

Johnson, J. N., 2024. *Amazon.com (AMZN) Stock Price, News & Analysis.* [Online]   
Available at: https://www.marketbeat.com/stocks/NASDAQ/AMZN/  
[Accessed 15 November 2024].

Sham, J., 2020. *Stock Price Dashboard built with Plotly Dash — Part I.* [Online]   
Available at: https://medium.com/@jjsham/building-stock-price-dashboard-with-plotly-dash-part-i-b165b3edfdd6  
[Accessed 15 November 2024].