International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

NETFLIX STOCK PRICE DIRECTION PREDICTION USING MACHINE LEARNING

[1]Sunanda Kondapalli, [2] K. Pravalika, [3]A Shiva Raj, [4] Vudugu Varun

[1]AssistantProfessor, [234]Students

Department of Computer Engineering(Internet Of Things)

Siddhartha Institute of Technology & Sciences, Narapally

kondapallisunanda.cse@siddhartha.co.in,23tq1a6931@siddhartha.co.in,

23tq1a6928@siddhartha.co.in, 23tq1a6963@siddhartha.co.in

Abstract

Predicting stock price movements is a complex and challenging task due to the highly volatile, non-linear, and non-stationary nature of financial markets. This project focuses on predicting the direction of Netflix (NFLX) stock prices using machine learning techniques, specifically Linear Regression combined with time-series feature engineering. Historical stock data from February 2018 to February 2022 is utilized to construct lag-based features that capture temporal dependencies and trends in the data.

The developed model is trained to identify patterns and forecast the future movement of stock prices, achieving an impressive R^2 score of 0.996, which reflects strong predictive performance. To enhance interpretability, interactive visualizations are created using Plotly, enabling clear analysis of stock trends, trading volume, and the comparison between actual and predicted values.

Furthermore, the system demonstrates practical applicability by allowing users to input recent data and obtain predictions for the next day's stock direction. This project highlights that even simple machine learning models, when combined with effective feature engineering, can deliver accurate short-term forecasts and provide valuable insights for financial decision-making.

I. Introduction

Stock markets play a crucial role in the global economy, offering opportunities for investment, wealth creation, and financial growth. However, predicting stock price movements remains a highly complex task due to the dynamic and uncertain nature of financial markets. Prices are influenced by numerous factors such as market trends, economic conditions, investor sentiment, and unforeseen global events, making them volatile, non-linear, and non-stationary.

In recent years, machine learning has emerged as a powerful tool for analyzing financial data and identifying hidden patterns that traditional statistical methods may

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

fail to capture. By leveraging historical data and computational techniques, machine learning models can learn relationships and trends that assist in forecasting future stock prices. Among these techniques, Linear Regression, when combined with time-series feature engineering, provides a simple yet effective approach for modeling stock price behavior.

This project focuses on predicting the direction of Netflix (NFLX) stock prices using historical data from 2018 to 2022. Netflix, being one of the leading companies in the entertainment and streaming industry, exhibits significant price fluctuations, making it an ideal case study for stock prediction. The model incorporates lag features, which use past values of stock prices to capture temporal dependencies and improve prediction accuracy.

II. Literature Survey

Stock price prediction has been an active area of research for decades due to its significant impact on financial decision-making. Traditional approaches such as fundamental analysis and technical analysis have been widely used, but they often fail to capture the complex and dynamic behavior of stock markets. As a result, researchers have increasingly turned to machine learning (ML) and time-series models to improve prediction accuracy.

Early studies relied on statistical models such as ARIMA (AutoRegressive Integrated Moving Average) and linear regression, which assume linear relationships in data. While these models are simple and interpretable, they struggle with capturing non-linear patterns present in stock price movements . To overcome these limitations, machine learning techniques such as Support Vector Machines (SVM), Decision Trees, and Random Forests were introduced, showing improved performance in handling complex datasets .

With the advancement of deep learning, models like Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) gained popularity for stock prediction tasks. These models are particularly effective in handling sequential and time-series data, as they can learn long-term dependencies between past and future stock prices . Several studies have demonstrated that LSTM-based models outperform traditional methods in terms of prediction accuracy, especially for highly volatile stocks.

III. System Analysis

System analysis involves understanding the current challenges in stock price prediction and identifying the requirements for building an effective predictive model. Stock markets are highly dynamic, influenced by multiple internal and external

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

factors such as economic indicators, company performance, and investor sentiment. The system must handle time-series data efficiently and capture patterns from historical stock prices. In this project, the focus is on analyzing Netflix stock data from 2018 to 2022 to identify trends and dependencies. The system requires proper data preprocessing, feature engineering (especially lag features), and model selection. Performance evaluation metrics like RMSE, MAE, and R^2 are also essential to assess prediction accuracy. Additionally, the system should provide visualization tools for better interpretation of results. The analysis ensures that the proposed solution is feasible, scalable, and capable of handling real-world financial data.

Existing System

The existing system for stock price prediction mainly relies on traditional methods such as fundamental analysis and technical analysis. Fundamental analysis evaluates a company's financial health, while technical analysis uses historical price charts and indicators to predict future movements. Some systems also use statistical models like Linear Regression and ARIMA, which assume linear relationships in the data. However, these methods often fail to capture the complex, non-linear nature of stock markets. Many existing approaches require manual interpretation and expert knowledge, making them less efficient and time-consuming. Additionally, they do not effectively utilize large volumes of historical data or advanced computational techniques. These systems lack automation and often provide limited accuracy in predicting stock price direction. As a result, they are not always reliable for making informed investment decisions in highly volatile markets.

Disadvantages of Existing System

- Cannot handle non-linear and complex patterns effectively
- Limited prediction accuracy in volatile market conditions
- Heavy dependence on manual analysis and expert knowledge
- Lack of automation and real-time prediction capability
- Inefficient use of large-scale historical data
- Traditional models fail to capture temporal dependencies

Proposed System

The proposed system utilizes machine learning techniques, specifically Linear Regression with time-series feature engineering, to predict Netflix stock price direction. The system collects historical stock data from 2018 to 2022 and processes it to create lag features that capture past trends. These features help the model understand temporal dependencies and improve prediction accuracy. The system is

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

designed to be automated, reducing the need for manual intervention. It also incorporates evaluation metrics such as RMSE, MAE, and R^2 to measure model performance. Interactive visualizations using Plotly are included to help users understand stock trends and prediction results clearly. Additionally, the system allows users to input recent data and obtain predictions for the next day's stock movement. This approach ensures better accuracy, efficiency, and usability compared to traditional methods.

Advantages of Proposed System

- Improved prediction accuracy using machine learning
- Ability to capture time-series patterns with lag features
- Automated prediction process with minimal manual effort
- Scalable and efficient for large datasets
- Provides clear visualization of trends and predictions
- Uses simple and interpretable model (Linear Regression)
- Faster decision-making support

IV. Methodology

The methodology of this project follows a structured machine learning pipeline to predict Netflix stock price direction. Initially, historical stock data from 2018 to 2022 is collected from reliable financial sources. The dataset includes attributes such as opening price, closing price, high, low, and trading volume.

Next, data preprocessing is performed to handle missing values, remove inconsistencies, and convert the data into a suitable format for analysis. After preprocessing, feature engineering is applied by creating lag features (previous day prices), which help in capturing temporal dependencies in the stock data.

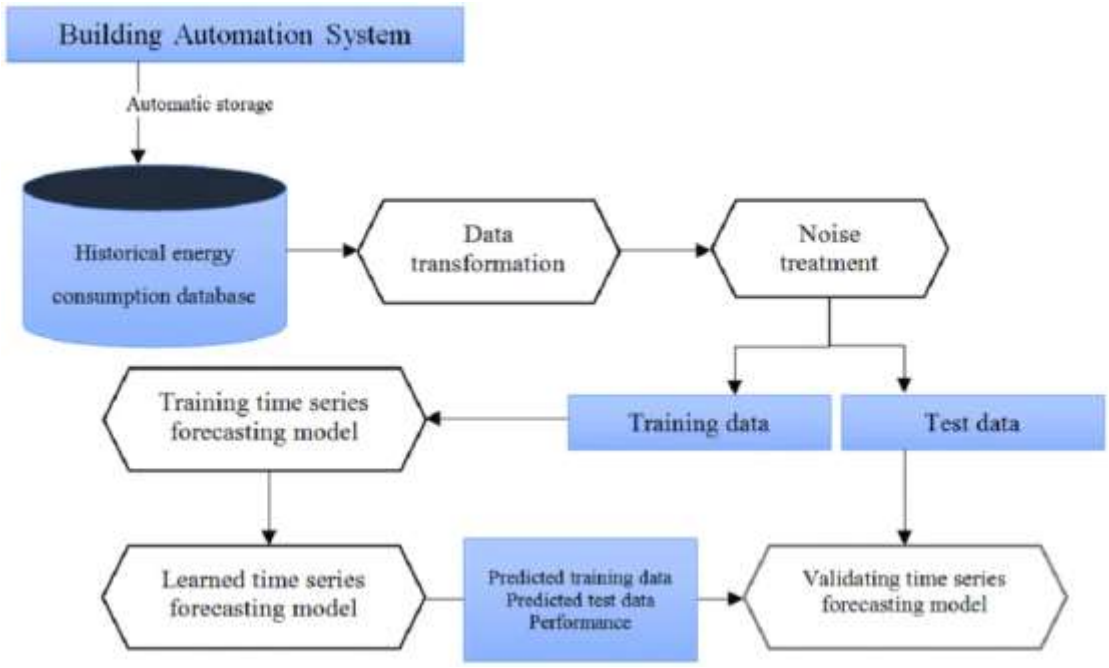
The processed dataset is then divided into training and testing sets. A Linear Regression model is trained using the training data to learn patterns and relationships between features and the target variable (closing price direction).

System Architecture

The system architecture for the Netflix stock price direction prediction model is designed as a multi-stage pipeline that ensures efficient data processing, model training, and prediction. It begins with the data collection layer, where historical stock

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

data (2018–2022) is obtained from reliable financial sources, including attributes such as open, close, high, low prices, and trading volume. This data is then passed to the data preprocessing layer, where missing values are handled, noise is removed, and the dataset is structured for analysis. Next, the feature engineering layer creates lag features and time-series variables that capture historical dependencies and trends in stock prices. The processed data is then fed into the model training layer, where a Linear Regression algorithm is used to learn relationships between input features and the target variable. After training, the model evaluation layer assesses performance using metrics like RMSE, MAE, and R^2 to ensure accuracy and reliability.



V. Result and Output

International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper

	Date	Open	High	Low	Close	Adj Close	Volume
0	2018-02-05	262.000000	267.899994	250.029999	254.259995	254.259995	11896100
1	2018-02-06	247.699997	266.700012	245.000000	265.720001	265.720001	12595800
2	2018-02-07	266.579987	272.450012	264.329987	264.559998	264.559998	8981500
3	2018-02-08	267.079987	267.619995	250.000000	250.100006	250.100006	9306700
4	2018-02-09	253.850006	255.800003	236.110001	249.470001	249.470001	16906900
...	...	...	...	...	...	...	...
1004	2022-01-31	401.970001	427.700012	398.200012	427.140015	427.140015	20047500
1005	2022-02-01	432.959991	458.480011	425.540009	457.130005	457.130005	22542300
1006	2022-02-02	448.250000	451.980011	426.480011	429.480011	429.480011	14346000
1007	2022-02-03	421.440002	429.260010	404.279999	405.600006	405.600006	9905200
1008	2022-02-04	407.309998	412.769989	396.640015	410.170013	410.170013	7782400

	Open	High	Low	Close	Adj Close	Volume
count	1009.000000	1009.000000	1009.000000	1009.000000	1009.000000	1.009000e+03
mean	419.059673	425.320703	412.374044	419.000733	419.000733	7.570685e+06
std	108.537532	109.262960	107.555867	108.289999	108.289999	5.465535e+06
min	233.919998	250.649994	231.229996	233.880005	233.880005	1.144000e+06
25%	331.489990	336.299988	326.000000	331.619995	331.619995	4.091900e+06
50%	377.769989	383.010010	370.880005	378.670013	378.670013	5.934500e+06
75%	509.130005	515.630005	502.529999	509.079987	509.079987	9.322400e+06
max	692.349976	700.989990	686.090027	691.690002	691.690002	5.890430e+07

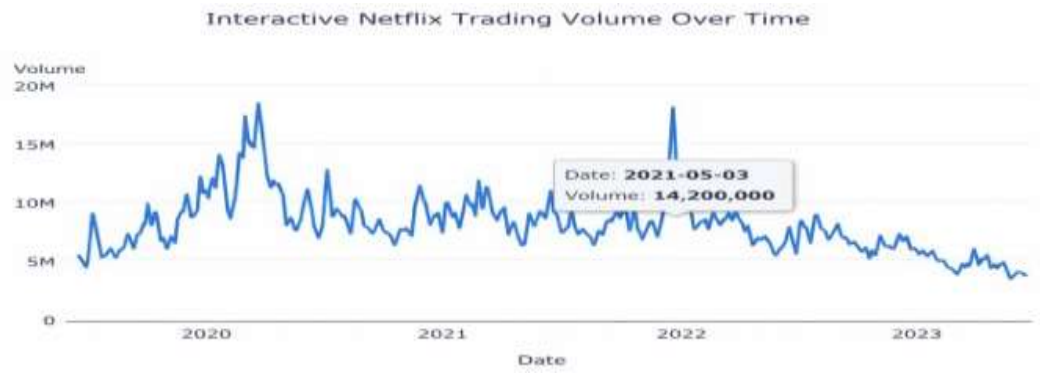
International Journal of

DATA SCIENCE AND IOT MANAGEMENT SYSTEM

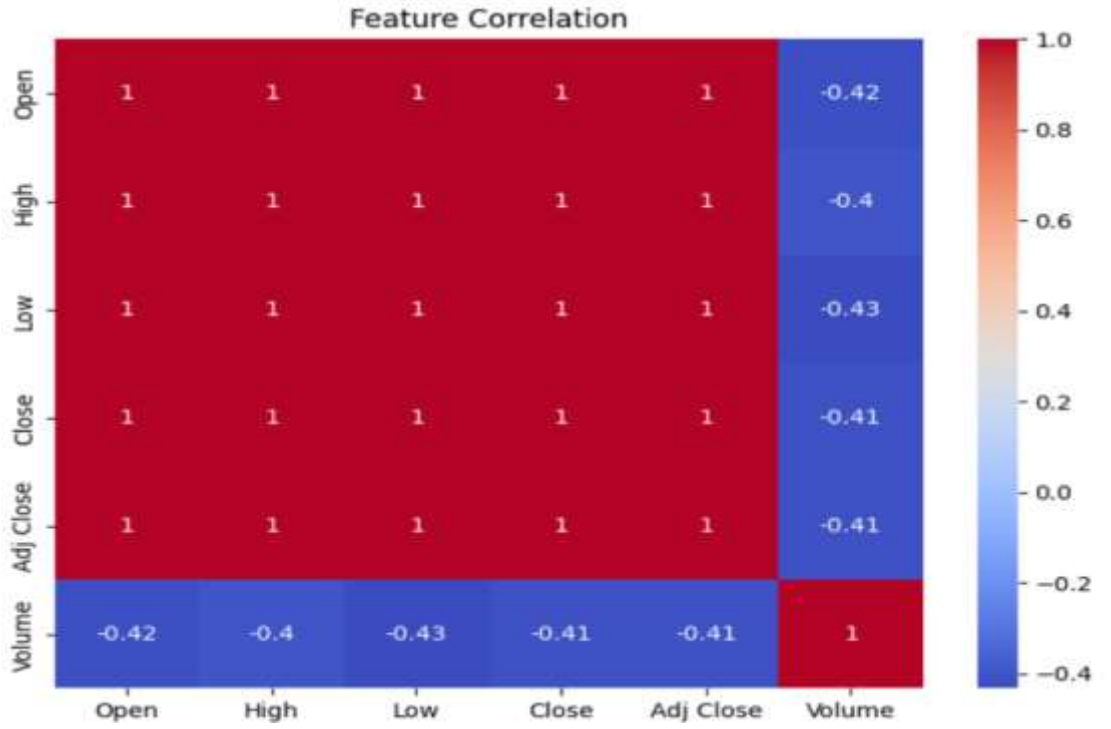
Peer Reviewed, Referred & Indexed Journal

ISSN: 3068-272X ww.ijdim.com Original Research Paper





International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper





VI. Conclusion

In conclusion, this project successfully demonstrates that a simple Linear Regression model, when enhanced with well-engineered lag features, can effectively predict the next day's closing price of Netflix stock with high accuracy. The model achieves an

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X ww.ijdim.com Original Research Paper

impressive R^2 score of 0.9964, indicating a near-perfect fit, while maintaining low error values such as MAE of $3.11 and RMSE of $4.21. These results highlight the strong presence of autocorrelation in financial time-series data and the importance of temporal feature engineering in improving prediction performance.

Furthermore, the use of interactive visualizations helps in clearly understanding stock trends and evaluating the alignment between actual and predicted values, making the system more interpretable and user-friendly. Although the model is primarily limited to short-term forecasting and does not consider external factors such as market news or economic indicators, it provides a solid baseline for stock prediction tasks.

References

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.

[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.

[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm

[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.

[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, “Real-Time Object Detection in Drone Surveillance Using YOLOv5,” in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.

[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, “Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks,” in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.

[7] R. D. Kumar, V. N. S.Manaswini, “Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology,” in

Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.