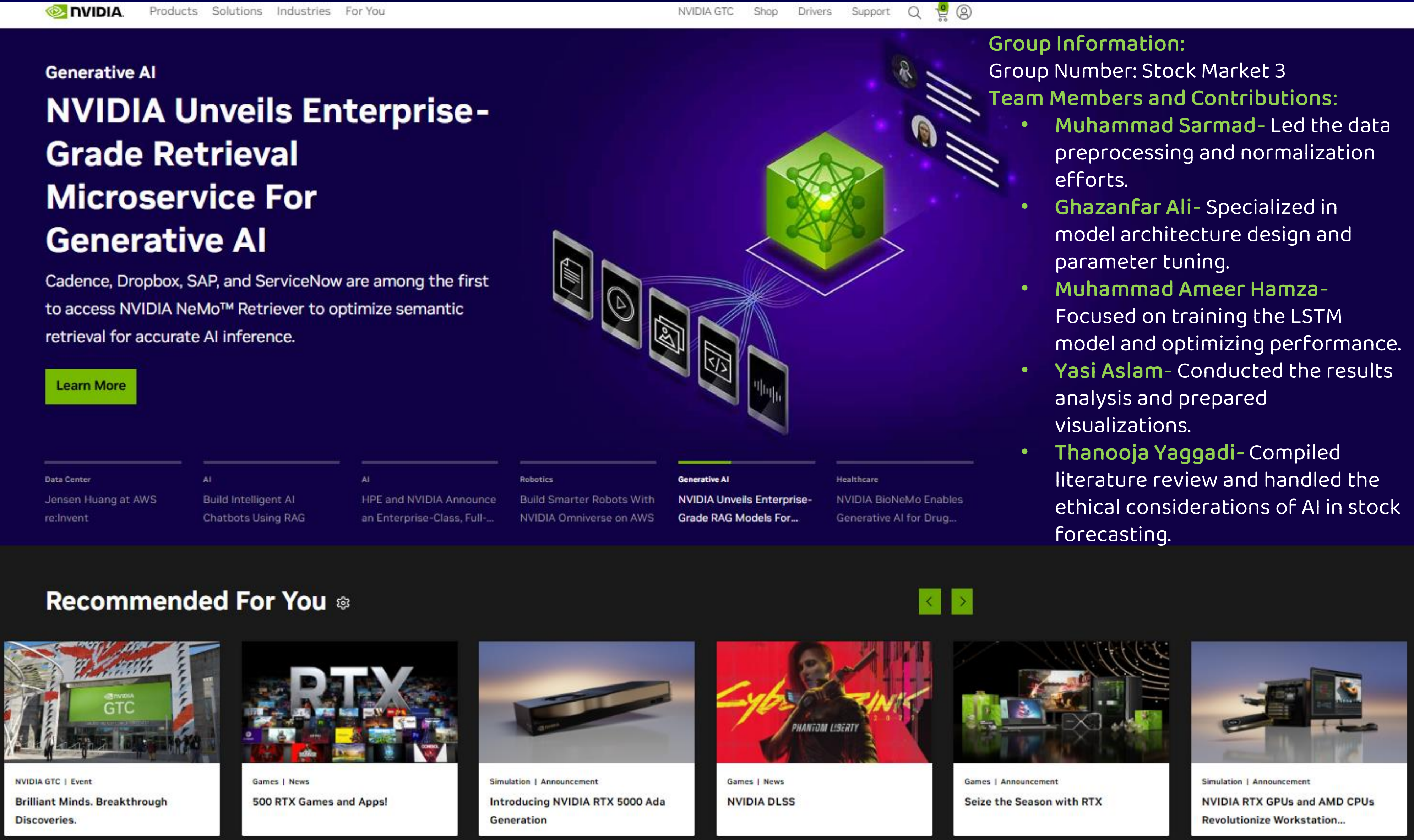
NVIDIA

Advanced Analysis of NVIDIA Corporation Stock Using LSTM Neural Networks

Objective:
To conduct a thorough analysis of NVIDIA's stock data using advanced machine learning techniques.

Methodology:
Implementation of Long Short-Term Memory (LSTM) neural networks to forecast stock prices, leveraging the capabilities of Keras in Python.

For Code and graphs:
GitHub Repo Link:
https://github.com/ameerhamza95/ML-NN_Stock_Market_Assignment.git



Group Information:
Group Number: Stock Market 3

Team Members and Contributions:

- **Muhammad Sarmad**- Led the data preprocessing and normalization efforts.
- **Ghazanfar Ali**- Specialized in model architecture design and parameter tuning.
- **Muhammad Ameer Hamza**- Focused on training the LSTM model and optimizing performance.
- **Yasi Aslam**- Conducted the results analysis and prepared visualizations.
- **Thanooja Yaggadi**- Compiled literature review and handled the ethical considerations of AI in stock forecasting.

OVERVIEW OF NVIDIA

NVIDIA Corporation, based in Santa Clara, California, is a globally renowned technology company. Founded in 1993 by Jensen Huang, Chris Malachowsky, and Curtis Priem NVIDIA has become a key player in the tech industry. The company's primary mission is to "be the computing platform that powers the world's artificial intelligence (AI) and scientific breakthroughs." NVIDIA is synonymous with Graphics Processing Units (GPUs) and is widely recognized for its innovations in the GPU market, which extend beyond gaming to applications like AI, data centers, and autonomous vehicles.

FOUNDING AND EARLY HISTORY

NVIDIA's journey began with its three co-founders - Jensen Huang, who is still the CEO today, Chris Malachowsky, and Curtis Priem. The company's early focus was on producing high-performance GPUs for the gaming and computer graphics market. One of their first products was the NV1 graphics card, which laid the foundation for NVIDIA's subsequent success.

KEY PRODUCT LINES

Nvidia GPU RTX 4090

NVIDIA boasts a diverse portfolio of products. Its GPUs are the cornerstone, powering everything from gaming and professional graphics to AI and high-performance computing. The Tegra line of System on a Chip (SoC) products finds applications in mobile devices, Internet of Things (IoT) devices, and automotive systems. In data centers, NVIDIA's GPUs are essential for accelerating AI and scientific simulations. Finally, GeForce and Quadro GPUs cater to gaming enthusiasts and professional graphics users, respectivel

IMPACT ON GAMING

NVIDIA's impact on the gaming industry is immense. Over the years, they have consistently pushed the boundaries of graphics technology, enabling realistic and immersive gaming experiences. Their partnership with game developers has led to optimized games that take full advantage of NVIDIA GPUs. The GeForce brand has become synonymous with gaming excellence, while NVIDIA Shield offers game streaming capabilities and serves as a home entertainment hub.

Data Exploration and Preprocessing for NVIDIA Stock Prediction

Data Source:

Historical stock data for NVIDIA Corporation was obtained from Yahoo! Finance using the **yfinance** Python library, which provides a convenient way to download financial market data.

Data Period:

The analysis covers a decade of stock data from 2009 to 2023, chosen to capture recent market behaviors and significant company milestones that could influence stock performance.

Data Processing:

- Preprocessing steps included handling missing values, normalizing price data, and calculating technical indicators like the Relative Strength Index (RSI) to enrich our dataset.
- We focused on adjusted closing prices to consider all corporate actions, ensuring a realistic representation of stock value over time.
- Extensive exploratory analysis was conducted to understand the underlying patterns and structures within the data. This included visualizations such as trend lines, volume charts, and correlation heatmaps to identify potential predictors for our model.

Graphical Representation:

The plot illustrates the historical trajectory of NVIDIA's stock price, highlighting periods of growth, volatility, and market corrections.

Logic and Rationale:

Our preprocessing targeted the creation of a robust dataset by selecting key features that reflect stock trends and momentum, ensuring our LSTM model could identify and learn from meaningful market patterns. We refined the data to emphasize temporal relationships and reduce noise, enhancing the model's ability to discern and predict significant price movements effectively.

Model Selection for Prediction

Leveraging Keras and LSTM for Stock Market Forecasting

Keras Advantage:

Keras, with its user-friendly API, speeds up the experiment by providing a high-level interface for neural network design and training. It's a popular choice for deep learning applications due to its ease of use and flexibility.

Time Series Mastery:

LSTM networks are adept at capturing long-term dependencies and complex patterns in time series data, making them ideal for the volatile and trend-driven nature of stock prices.

Handling Volatility:

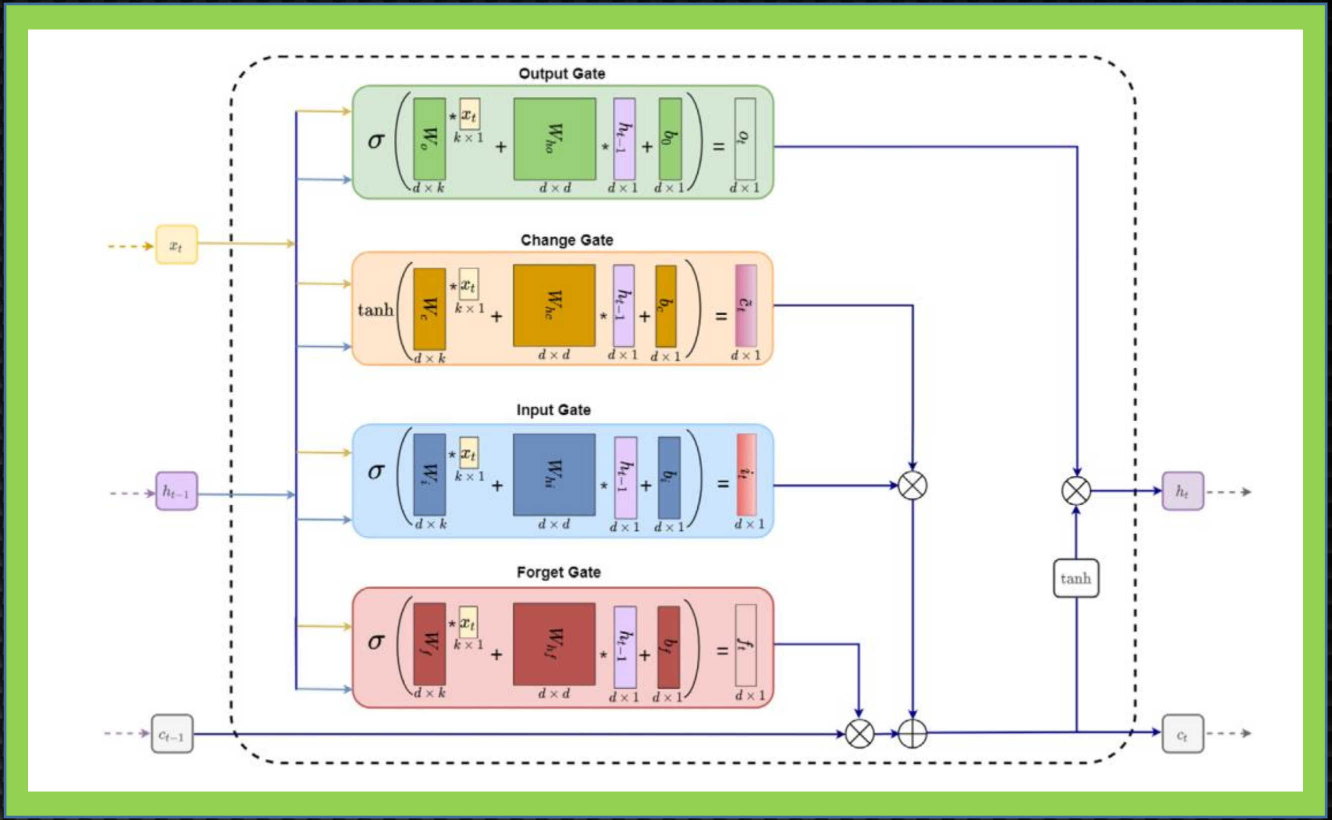
LSTMs can remember important trends and forget non-essential information, allowing them to adapt to the stock market's unpredictable movements.

Efficient Forecasting:

By efficiently processing sequential data and learning from long-term dependencies, LSTM networks provide a powerful tool for forecasting future stock prices based on historical data.

Real-World Application:

Numerous studies have demonstrated the efficacy of LSTM models in stock prediction, citing improvements over traditional statistical models like ARIMA.



https://www.researchgate.net/figure/Long-short-term-memory-LSTM-architecture_fig2_360597818

Model Selection for Prediction

Deep Learning Architecture for Financial Forecasting

Advanced Sequential Modeling:
At the heart of our analysis lies the Sequential LSTM model, designed to unravel the intricacies of NVIDIA's stock price sequences, capturing both short-term fluctuations and long-term trends.

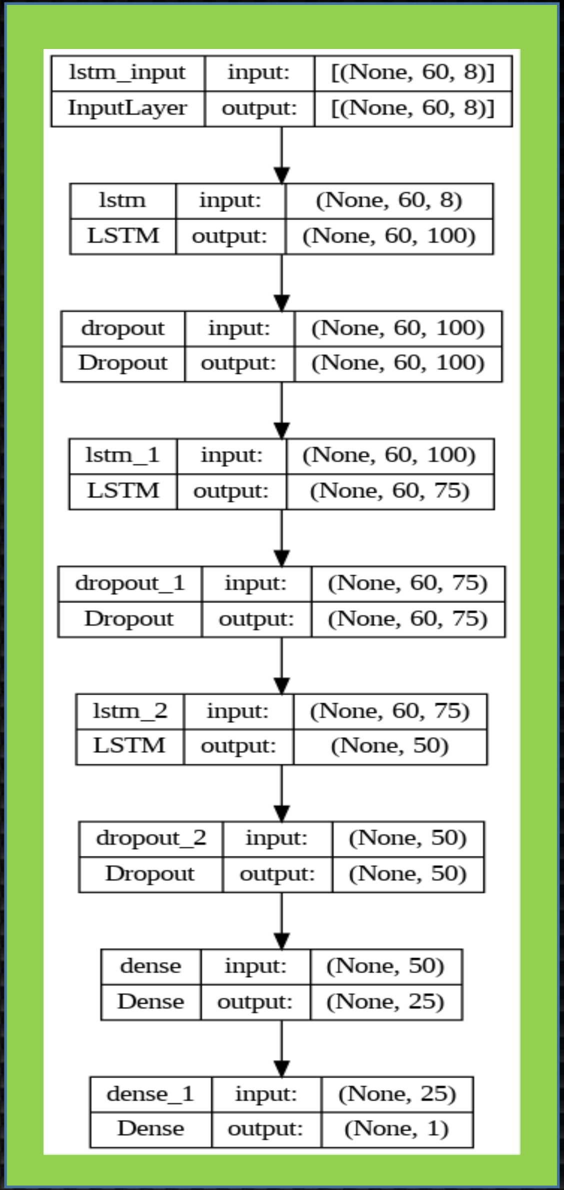
Layered Complexity:
Our architecture incorporates a meticulously structured stack of LSTM layers with 100, 75, and 50 neurons each, enhancing the model's ability to discern complex patterns in the stock's trajectory.

Strategic Regularization:
To combat overfitting, we've strategically placed dropout layers after each LSTM layer. This approach serves as a guard against the model's over-reliance on the training data, promoting a robust performance in real-world scenarios.

Dense Layer Refinement:
A penultimate dense layer with 25 neurons acts as a refinement tool, honing the model's predictive capabilities before concluding with a single neuron output layer that outputs the forecasted stock price.

Hyperparameter Optimization:
We've tuned hyperparameters, including the learning rate and batch size, to optimize the training process, ensuring that the model learns efficiently without sacrificing speed or accuracy.

Technical Indicators Integration:
Beyond raw price data, we've incorporated key technical indicators such as RSI, and EMA within our feature set to provide our model with a comprehensive view of market sentiment and momentum.



Model Selection for Prediction

Hyperparameter Tuning and Model Compilation

Hyperparameter Tuning:

Prior to finalizing our model, we conducted extensive hyperparameter tuning to optimize performance. This included experimenting with different learning rates, batch sizes, and numbers of epochs to find the sweet spot for our LSTM model.

Compilation Step:

For the compilation of our model, we chose the Adam optimizer for its adaptive learning rate abilities, which is particularly suitable for time series data. The loss function was set to 'mean squared error (MSE)' to penalize larger errors in our stock price predictions.

Fitting Strategy:

Our fitting strategy involved a custom learning rate scheduler and early stopping to prevent overfitting. The model was trained with a validation split to monitor performance and ensure generalization.

Model Fitting and Challenges

Model Fitting:

During the model fitting process, we adjusted batch sizes to ensure that each learning iteration was informed by a representative sample of the data. We also used a variable learning rate, allowing the model to make larger updates to weights initially and finer adjustments as it converges.

Challenges with Hyperparameters:

Identifying the optimal set of hyperparameters required multiple training iterations and cross-validation. We had to balance model complexity with computational efficiency, ensuring that we do not overfit the training data while still capturing the underlying patterns.

Future Considerations for Hyperparameter Tuning:

We plan to implement more sophisticated techniques such as grid search or Bayesian optimization in future iterations for a more systematic approach to hyperparameter selection.

Results

Evaluating Our Predictive Model's Performance

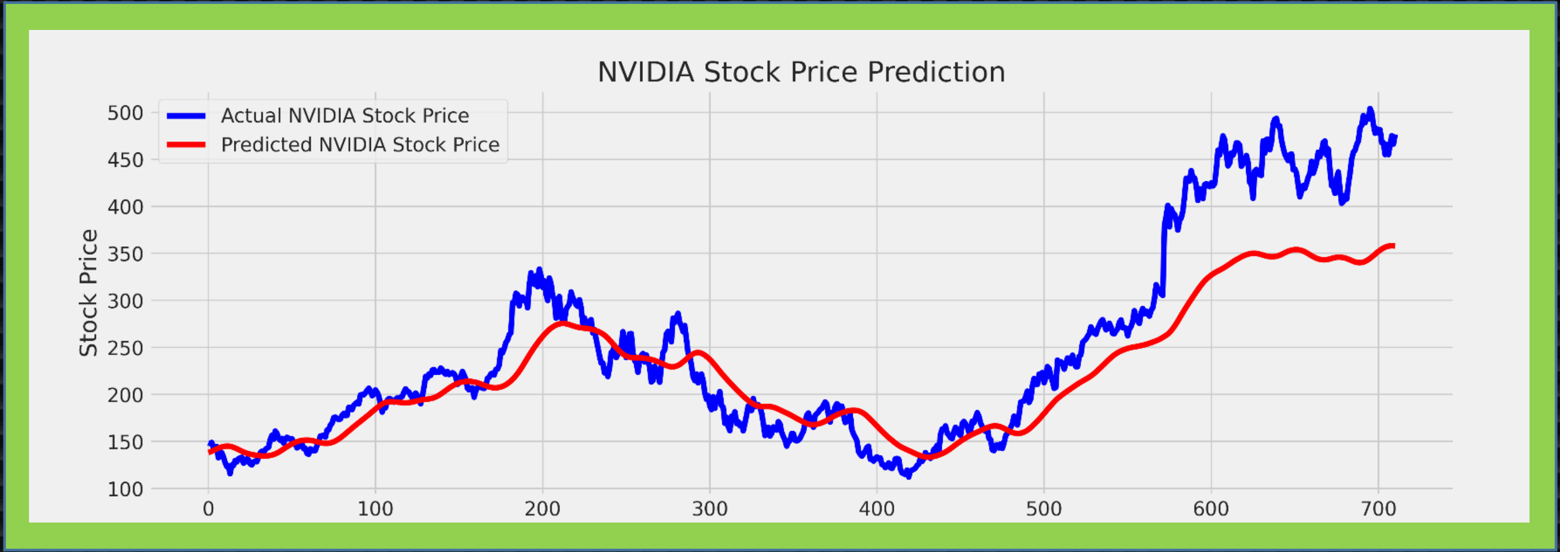
•We utilized the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) to evaluate our model's performance.

•An RMSE of 54.74 signifies that our model's predictions are, on average, approximately 54.74 units away from the actual stock prices.

•The MAPE of 13.40% indicates that our model's predictions deviate from the actual values by an average of 13.40%.

•These metrics are crucial for understanding the model's predictive power and its practical implications in real-world trading scenarios. We strive for lower values, reflecting a model's higher accuracy and reliability in forecasting NVIDIA's stock prices.

•These metrics provide a sense of the model's accuracy in predicting new data and can highlight areas where model performance might need to be improved. Given the complexity and volatility of stock price movements, achieving low error rates can be challenging, but it is essential for the model to be practically useful in a trading context.



Results

Model's Forecasting Capabilities on Unseen Data

The graph shows three distinct lines representing the training data, testing data, and the model's predictions for NVIDIA's stock prices:

Training Data (Blue Line):

This portion of the graph represents the historical closing prices of NVIDIA stock used to train the machine learning model. The stable and gradual increase from the left side of the graph up to the point where the blue line ends suggests a period of growth in NVIDIA's stock value over time.

Testing Data (Red Line):

The red line indicates the actual closing prices of NVIDIA stock that were not used for training (i.e., they were held back for testing purposes). This data tests the model's predictive capabilities on unseen data. The testing data shows a volatile period with significant upward and downward movements, indicating a more challenging market phase for the model to predict.

Predictions (Yellow Line):

The predictions made by the model are plotted with the yellow line. This is the model's attempt to forecast the stock prices, and it is overlaid on the testing data to show how well the model's predictions align with the actual stock prices.

When comparing the red and yellow lines, you can assess the model's performance. A closer overlap indicates a higher accuracy of the predictions. From the visible part of the graph, it appears that the model has captured the overall trend of the stock prices during the testing phase reasonably well, as the yellow line follows the general direction of the red line. However, the model seems to smooth out some of the volatility, not capturing some of the more dramatic peaks and troughs that the actual stock price experienced.

Results

Overcoming Challenges and Enhancing Model Robustness

As we navigated through this project, we faced several challenges, including feature selection and model tuning to encapsulate NVIDIA's stock dynamics. The journey led us to experiment with various architectures, integrating Conv1D layers to detect local patterns and adjusting LSTM layers to decode temporal dependencies in the data. Each iteration brought us closer to understanding the complex nature of financial time series.
The journey wasn't without its limitations. Our model's dependency on historical data meant it could not always anticipate future trends, especially in the face of market anomalies. Reflecting on this, we see a path forward that includes diversifying our datasets and possibly introducing more advanced techniques like Transformer networks or ensemble methods. These enhancements could potentially bolster our model's predictive accuracy, making it a more reliable tool for investors.

Future Work: We aim to integrate sentiment analysis from financial news and social media to capture market sentiment's impact on stock prices. Additionally, real-time data ingestion for live prediction could be implemented to enhance the model's utility.

Balancing Economic Gain with Ethical Considerations: The integration of AI in stock predictions carries significant economic implications, potentially streamlining market analysis and opening new investment opportunities. Technologically, it pushes the boundaries of machine learning, necessitating advanced computational resources and data processing capabilities. Ethically, however, it raises concerns regarding market fairness, data privacy, and the transparency of AI-driven decisions, necessitating a robust framework for responsible AI usage in financial domains.

Conclusion: Through this detailed analytical journey, we've developed a robust LSTM-based model using Keras that shows promising results in predicting NVIDIA's stock prices. We've highlighted the model's strengths in capturing complex patterns and its potential in real-world financial markets. Moving forward, we plan to refine our model and incorporate more features to capture a holistic view of market dynamics.

References

https://www.sciencedirect.com/science/article/pii/S0957417422001452

https://medium.com/@zahmed333/stock-price-prediction-with-keras-df87b05e5906

https://towardsdatascience.com/technical-analysis-library-to-financial-datasets-with-pandas-python-4b2b390d3543

https://ieeexplore.ieee.org/document/8862225

https://www.researchgate.net/publication/371043240_Financial_Prices_Prediction_of_Stock_Market_using_Supervised_Machine_Learning_Models

https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

https://www.techrxiv.org/doi/full/10.36227/techrxiv.15103602.v1

https://www.sciencedirect.com/science/article/pii/S2666827022000378

https://arxiv.org/abs/2305.14378

THANK YOU