AMAZON STOCK MARKET ANALYSIS

PRESENTED BY GROUP 11
OLAJIDE STEPHEN WALE-ONI (22013012)
ONYEDIKACHI ONWUACHUKE (22021169)
ERNEST UMENSOFOR (22)
OWOBU CHRISTIAN (22017251)
FRIDAY OVBIRORO (22034665)

INTRODUCTION

OVERVIEW:
Amazon.com, Inc., widely recognized as Amazon was founded by Jeff Bezos on July 5, 1994, is an American multinational technology company focusing on e-commerce, cloud computing, online advertising, digital streaming, and artificial intelligence.

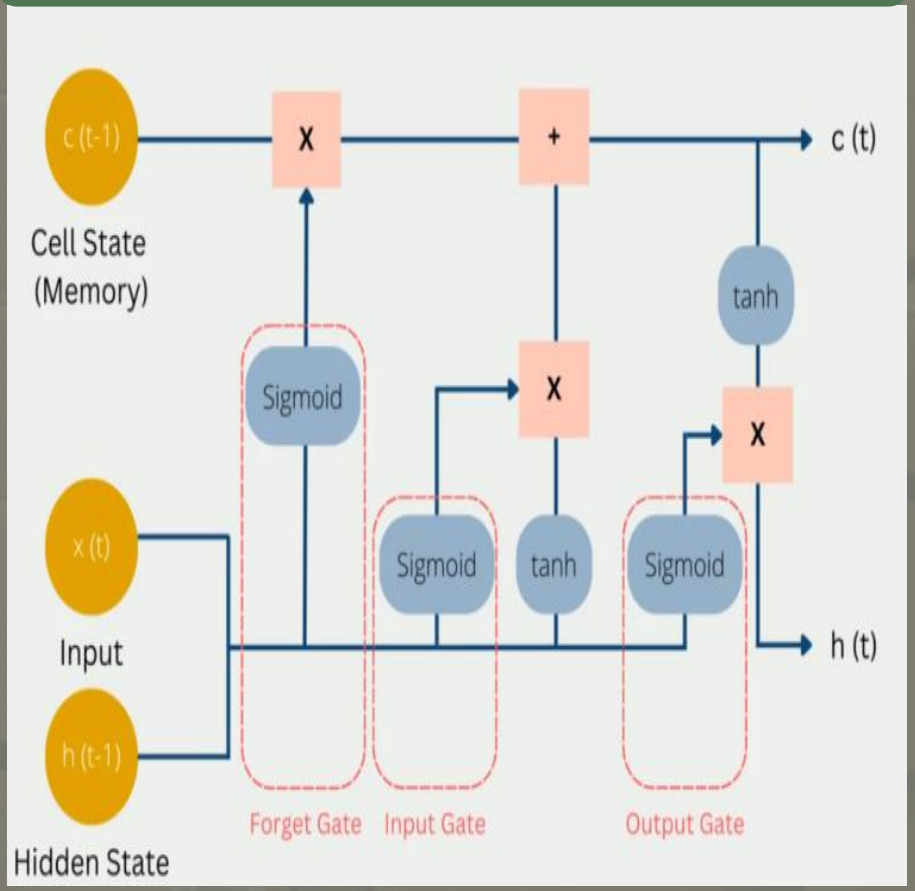
Amazon, initially an online marketplace for books quickly diversified its product range to include music, video games, consumer electronics, home improvement items, software, games, and toys. Amazon's impact spans industries, reflecting its commitment to innovation, customer satisfaction, and continuous global expansion.

AMAZON STOCK

- IPO and Early Growth (1997-2000): Amazon went public in 1997, and the IPO price was $18.00. It saw a significant growth during the dot-com boom, but later faced a drop in stock price during the dot-com bubble burst.
- Diversification (Early 2000s): Amazon expanded from books into various e-commerce categories.
- Steady Growth (Mid-2000s to Early 2010s): Amazon recovered and grew steadily, introducing Amazon Prime and Kindle.
- Tech Giant (2010s): Became a tech giant with e-commerce dominance, AWS, and new ventures.
- Pandemic Boost (2020): COVID-19 led to a surge in online shopping, benefiting Amazon.
- Continued Expansion (2021-2022): Amazon expanded into new markets and announced a stock split in 2022.
- Amazon's stock history has been marked by periods of volatility, as well as sustained growth, and it has become one of the most valuable and influential companies globally.

LSTM Model

How LSTM model works



LSTM Model: LSTMs which means Long ShorTerm Memory (LSTM) Networks are a special kind of Recurrent Neural Network (RNN) capable of earning long-term dependencies. They are particularly well-suited for time series data
Due to their ability to remember past information and use it to influence current predictions. This makes them ideal for tasks like sequence prediction, which is a common requirement in stock market analysis.

Key Features of LSTM

Memory Cells: LSTMs contain memory cells that can maintain information in memory for long periods. This is crucial for understanding patterns over time.

Gates: They have three types of gates (input, forget, and output) that regulate the flow of information into and out of the cell, helping the model to keep or discard data.

Long-Term Dependencies Handling: Traditional RNNs struggle with long-term dependencies due to the vanishing gradient problem. LSTMs are designed to address this by maintaining information over long sequences, making them effective for tasks like language modeling and time-series prediction.

SUITABILITY OF LSTM MODEL

When it comes to analyzing stock data, such as Amazon's, LSTMs can be quite effective due to their ability to understand and predict patterns over time. Stock prices are influenced by a variety of factors that unfold over time, making the temporal modeling capabilities of LSTMs particularly valuable.

Temporal Data Handling: LSTMs are well-suited for time series data like stock prices due to their ability to remember information over long time intervals.

Pattern Recognition: They can recognize patterns in historical stock data, which is key for predicting future stock movements.

Volatility Management: Stock data is often volatile and non-linear; LSTMs can handle such complexities better than many other algorithms.

Feature Integration: LSTMs can integrate multiple features (like opening price, closing price, volume, etc.) to make comprehensive predictions.

Conclusion: LSTMs are a powerful tool for modeling time series data like stock prices, their effectiveness in predicting stock market trends, specifically for companies like Amazon, depends on a variety of factors including the quality of the data, the model's configuration, and the inherent unpredictability of the stock market.

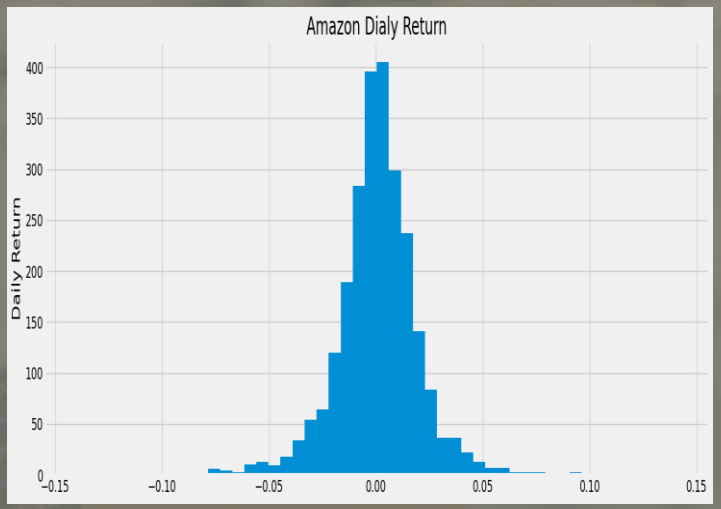
DATA EXPLORATION

1.Data Acquisition:

The historical stock data for Amazon is downloaded using the yfinance library for a specified date range (from 2013-2023).

2.Data set summary

Basic data exploration is conducted using print(stock_data), stock_data.describe(), stock_data.info(), and stock_data.columns. Visualization of the stock data is performed using matplotlib.pyplot, specifically plotting the opening and closing prices, and the histogram of daily returns.



Visualizing the daily return percentages of Amazon's stock, providing an insight into how much the stock's price is changing each day. This kind of analysis is useful in finance to assess the volatility of a stock. A more volatile stock will show wider swings in daily returns, whereas a less volatile stock will show more consistent, smaller changes.

AMAZON STOCK ANALYSIS

PRE-PROCESSING TECHNIQUES

1.Data Normalization:
The MinMaxScaler from sklearn preprocessing is employed to scale the closing prices of the stock. This normalization process transforms the ata to fit within a specified range (0, 1), Which is a common practice for neural Network inputs to improve training Efficiency and convergence.

2.Data Splitting:
The data is split into training and testing datasets, considering the sequential nature of time series data.

3.Time Series Data Structure Creation:
A time series data structure is created using the TimeseriesGenerator from Tensor flow keras pre-processing sequence. This involves defining a look back period, which determines how many previous time steps are used to predict the next time step. The data is split into training and testing sets. The training set consists of 80% of the data, and the testing set includes the remaining data plus the look-back period.

4.Reshaping for LSTM Input: The LSTM in Tensor Flow/Keras requires the input to be in a specific shape (samples, time steps, features). This is inherently handled by the TimeseriesGenerator, which creates sequences of these time stamped data points.

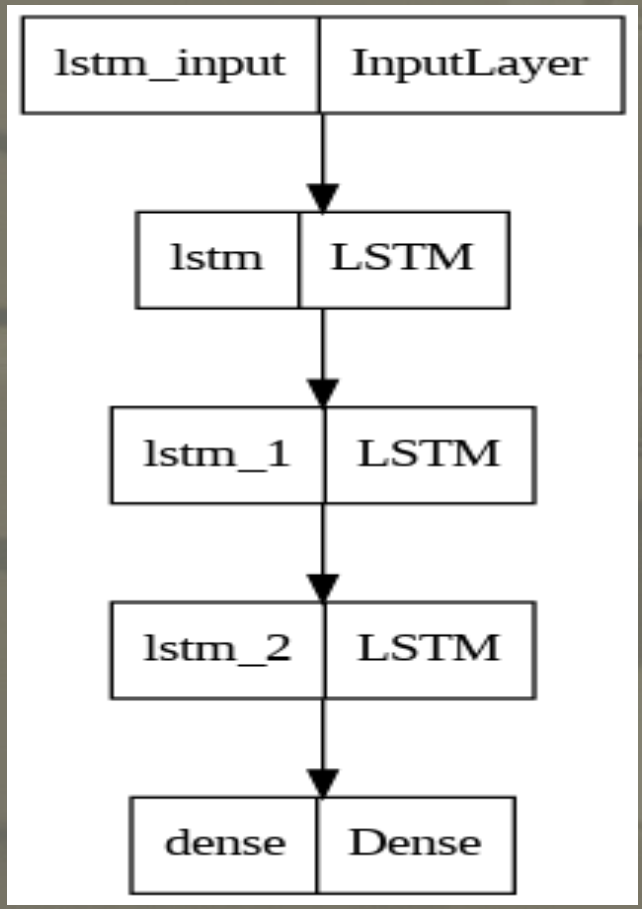
AMAZON STOCK ANALYSIS

LSTM ARCHITECTURE AND PARAMETERS.

The LSTM model in your code for predicting Amazon's stock price has the following key components:

- **Sequential Model**: A linear stack of layers in TensorFlow's Keras framework.
- **LSTM Layers**:
  - First Layer: 50 units, returns full sequence.
  - Second Layer: 100 units, returns full sequence.
  - Third Layer: 100 units, returns last time step output.
- **Dense Layer**: A single unit for final output prediction.
- **Parameters**:
  - Optimizer: 'adam', suitable for large datasets.
  - Loss Function: 'mean square error', common for regression tasks.
- **Training**: Conducted for 86 epochs,where each epoch represents one complete pass through the training data.
- **Input Shape**: Defined by look back (60), the model uses the past 60 data points for prediction.

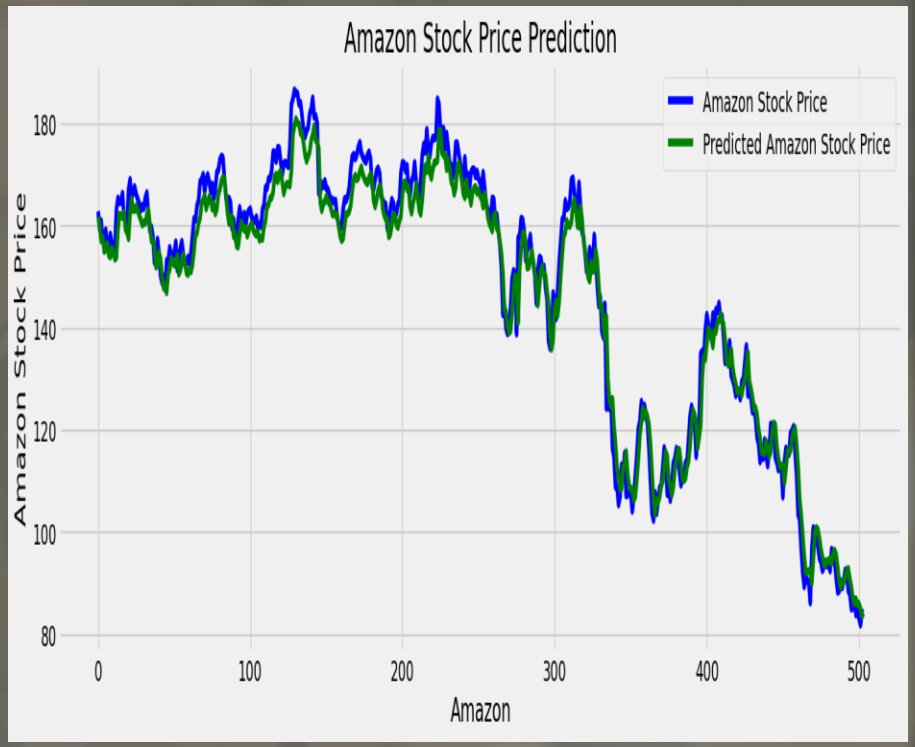
LSTM MODEL LAYOUT



RESULTS AND PREDICTIONS

- **Trend Analysis**: The model's training on historical Amazon stock data shows a long-term upward trend followed by recent declines.
- **Volatility**: The daily return histogram indicates typical financial return distribution, with most changes being close to zero.
- **Model Structure**: The sequential LSTM model comprises multiple layers designed to understand time-series data.

AMAZON STOCK PREDICTION



RESULTS AND PREDICTIONS

- **Future Forecasts**: The model forecasts the next 30 days' stock prices, extending the historical trend into the near future.
- **Performance Evaluation**: Loss plots indicate learning but with signs of overfitting, highlighting the need for possibly more nuanced model tuning.
- **Predictive Accuracy**: Plots comparing actual and predicted prices demonstrate the model's effectiveness in mimicking historical trends.
- R-squared (R^2) score of approximately 0.98 in the code indicates excellent model performance, with the model explaining about 98% of the variance in Amazon's stock
- The Mean Absolute Error (MAE) is approximately 3.476, indicating that, on average, the model's predictions differ from the actual stock prices by this amount.

FORCASTING FOR 30 DAYS

LIMITATIONS AND IMPROVMENTS OF THE MODEL

LIMITATIONS

- **Overfitting**: Risk of fitting noise, not just the underlying trend.
- **Limited Data**: Reliance on historical prices ignores broader market factors.
- **Feature Set**: Current features may not fully capture market complexities.
- **Unpredictability**: Inability to foresee sudden market shifts.
- **Simple Architecture**: May not grasp complex stock patterns.
- **Evaluation Depth**: Metrics used may not reflect true predictive power.

IMPROVEMENTS

- **Regularization**: Apply techniques to prevent over fitting.
- **Data Diversity**: Integrate wider data sources for prediction.
- **Enhanced Features**: Include technical indicators for depth.
- **External Factors**: Consider economic indicators in modeling.
- **Advanced Models**: Experiment with more sophisticated neural networks.
- **Optimize Hyper parameters**: Search for the best model settings.
- **Robust Testing**: Backtest against various market scenarios.
- **Update Mechanism**: Incorporate real-time data updates.

REFRENCES AND LINK

- LINK TO GOOGLE COLAB CODE (Group 11 Stock Market Analysis)
- https://colab.research.google.com/drive/1dAbDel7NuixnpsTTssxKP6fkyB2QAu_T

REFRENCES

- https://www.kaggle.com/code/abdallahwagih/tesla-stock-forecasting-lstm/comments
- https://pypi.org/project/yfinance/
- https://www.kaggle.com/code/matteoaries/tesla-stock-rnn-vs-lstm
- https://en.wikipedia.org/wiki/Jeff_Bezos

THANK YOU!!!