Stock Price Prediction Using Recurrent Neural Networks

1. Introduction

This project investigates the use of Recurrent Neural Networks (RNNs) for stock price prediction, specifically for Apple Inc. (AAPL). For the stock predictions, I examined three models: LSTM, GRU, and an enhanced GRU model that has trading volume data and ReLU activations. For this project, I implemented the models in Python using TensorFlow and then evaluated each one using RMSE.

2. Dataset and Features

My dataset was sourced from Yahoo Finance via the `yfinance` API, covering AAPL stock data from January 1, 2010, to November 13, 2023. The LSTM and base GRU models use only the 'Close' price, while the enhanced GRU model uses both 'Close' and 'Volume.' All of the data was then normalized using Min-Max scaling.

3. Data Preparation

I processed the dataset into rolling sequences of 60 days as input, with the target being the 61st day's closing price. I chose to split the data into 80% training and 20% testing sets as a starting benchmark.

4. Model Architectures

Each Jupyter notebook constructs its respective model as outlined below:

• LSTM Model:
- Two stacked LSTM layers
- Dense layers with 25 and 1 units, respectively

• GRU Model:
- Identical to the LSTM model's structure
- Replaces LSTM layers with GRU layers

• GRU + Volume + ReLU Model:
- Input includes both 'Close' and 'Volume' features
- Uses ReLU activation in GRU and Dense layers
- Modified input shape: (60, 2)

5. Training Process

I trained each model in the same fashion with the following configuration:
- Optimizer: Adam (learning rate = 0.001)
- Loss Function: Mean Squared Error (MSE)
- Epochs: 100
- Batch Size: 32
- 10% validation split

Training and validation performance can be found in each notebook's training output.

6. Evaluation Metric

I used Root Mean Square Error (RMSE) to evaluate model performance:
RMSE = sqrt(mean((predicted - actual)^2))

The predictions and actual values are first inverse transformed from their normalized values to the USD scale before computing RMSE.

7. Results and Comparison

Each model's predictions were evaluated over the test set that was previously established as being 20% of the dataset. The approximate RMSE results were as follows:
- LSTM: ~3.62 USD
- GRU: ~2.82 USD
- GRU + Volume + ReLU: ~22.14 USD

From this training and testing, it appears that the enhanced GRU model demonstrates a much different outcome, but faster training. This is due to the model's ReLU activations and added volume context. With that said, the ReLU activation could be a Leaky ReLU.

8. Conclusion

This project allowed me to understand the potential of GRU-based RNNs in forecasting stock prices. From evaluating these models, it's clear that factors such as incorporating volume and using alternative activation functions like ReLU can improve predictive accuracy.