**CS 677 Final Project: LSTMs on SPY Data**

Rahul Basu and Michael Lan

**TABLE OF CONTENTS**

1. Introduction to Models
2. Model Architecture
3. Raw Output of Actual vs Model Prices
4. Evaluation of Actual vs Model Prices
5. Comparing lookback sizes for Model A’s Low + Close

**1) Introduction to Models**

TIME SERIES TRAINING DATA:

SPY is an Exchange-traded fund (ETF) that reflects the performance of the S&P 500 Index.

The S&P 500 acts as major benchmarks of the U.S. equity market and indicates the financial health and stability of the economy. It is composed of five hundred selected stocks trading in the U.S., spanning over 20 separate industry groupings.

MODEL A:

Uses 4 separate models, each trained on a specific type of stock price. For instance, one model is trained only one open prices and outputs only open prices. This model is trained from 11/22/2013 to 4/16/2019. Open+High models used lookback of 100, while Low+Close models used lookback of 150. These are the results posted on Moodle.

MODEL B:

After this was submitted on 5/8, we realized we may have misunderstood the instructions, so we trained a new model that takes in all 5 time series (including volume). This new model is trained on ALL data up to 5/7 and uses lookback of 100. Its predictions are given in Results B and outputs all 5 types (open, high, low, close, vol). As shown in our reports, Results B were closer to the actual data.

**2) Model Architecture**

For all models, we trained using the Adam optimizer, with mean squared error as Loss and using MSE as evaluation metric. Validation data (split using 10% of training data) was initially used to select hyperparameters, but was later not used when we decided that the original model using 50 units in 4 LSTM layers with a final fully connected dense layer worked well.

MODEL A TRAINING

During training, models were trained on various subsets of training data. Evaluation was done by testing the model on 2 types of test data: from 4/17 to 5/3, and on all data (including the training data). The Root Mean Square Error was taken for these evaluations.

Lookback: sliding window used to obtain the past observations used for predictions

We compared various models using different combinations of training data subsets and lookback. These training data subsets included:

* All data up to 4/16
* Observation #4000 to #5348 (so 11/22/2013 to 4/16/2019)

We tested these in combination with various lookbacks values: 30, 60, 90, 100, 120, 150.

We found that using observations starting from 11/22/2013 with lookback of 100 worked well. We found a lookback of 150 worked better for Low and Close prices when evaluated on 4/17 to 5/3 data.

Dropout was used at first, but it was later not used as it was found that dropout made the result trends too smooth, which did not reflect the jagged nature of the actual price trends.

MODEL A Architecture:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

lstm\_1 (LSTM) (None, 100, 50) 10400

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_2 (LSTM) (None, 100, 50) 20200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_3 (LSTM) (None, 100, 50) 20200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_4 (LSTM) (None, 50) 20200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 1) 51

=================================================================

Total params: 71,051

Trainable params: 71,051

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

MODEL B TRAINING

We tested on all data up to 5/7. Since this took a long time to train, we only trained this once.

MODEL B Architecture:

Layer (type) Output Shape Param #

=================================================================

lstm\_1 (LSTM) (None, 100, 50) 11200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_2 (LSTM) (None, 100, 50) 20200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_3 (LSTM) (None, 100, 50) 20200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

lstm\_4 (LSTM) (None, 50) 20200

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 5) 255

=================================================================

Total params: 72,055

Trainable params: 72,055

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**3) Raw Output of Actual vs Model Prices**

ACTUAL

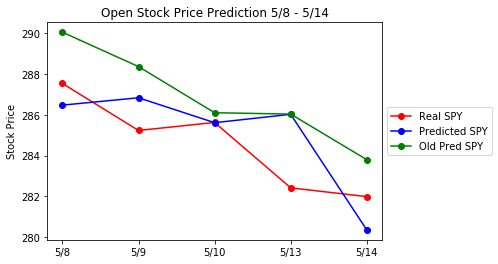
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Open | High | Low | Close |
| 5/8 | 287.53 | 289.43 | 286.87 | 287.53 |
| 5/9 | 285.23 | 287.33 | 283.3 | 286.66 |
| 5/10 | 285.62 | 288.94 | 282.3 | 288.1 |
| 5/13 | 282.42 | 283.49 | 279.93 | 280.86 |
| 5/14 | 281.99 | 285.1 | 281.85 | 283.4 |

RESULTS A

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Open | High | Low | Close |
| 5/8 | 290.04 | 293.25 | 285.62 | 295.36 |
| 5/9 | 288.36 | 290.22 | 287.09 | 292.02 |
| 5/10 | 288 | 289.3 | 283.45 | 290.52 |
| 5/13 | 286.03 | 289.39 | 281.72 | 291.15 |
| 5/14 | 288.27 | 287.37 | 279.74 | 289.2 |

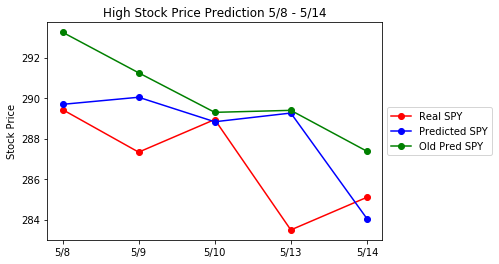
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RESULTS B | |  |  |  |
|  | Open | High | Low | Close |
| 5/8 | 286.47 | 289.69 | 283.92 | 286.01 |
| 5/9 | 286.83 | 290.04 | 284.85 | 286.92 |
| 5/10 | 285.61 | 288.83 | 283.23 | 285.4 |
| 5/13 | 286.02 | 289.26 | 283.94 | 286.05 |
| 5/14 | 279.98 | 283.69 | 276.8 | 279.47 |

**4) Evaluation of Actual vs Model Prices**



5-output (blue) RMSE for Open: 1.96

1-output (green) RMSE Open: 3.47



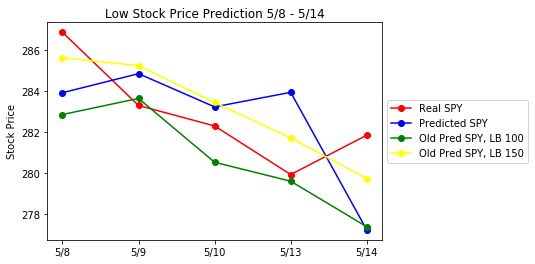
5-output (blue) RMSE for High: 2.89

1-output (green) RMSE High: 3.74

**5) Comparing lookback sizes for Model A’s Low + Close**

When evaluated on 5/8 to 5/14 data, we found that a model trained with a lookback window of 100 was actually better for Close, but a model trained with a lookback window of 150 was slightly better for Low. Both were trained on the same subset of data with the same model architecture + hyperparameters.

|  |  |  |
| --- | --- | --- |
|  | 100 | 150 |
| 8-May | 282.86 | 285.62 |
| 9-May | 283.64 | 287.09 |
| 10-May | 280.53 | 283.45 |
| 13-May | 279.61 | 281.72 |
| 14-May | 277.38 | 279.74 |

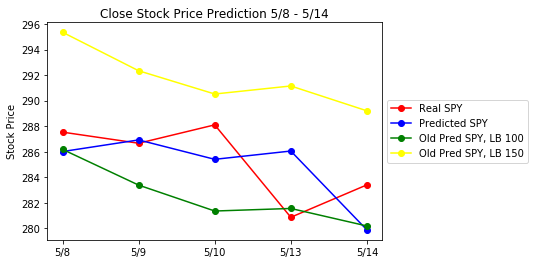


BLUE: 5-output RMSE for Low: 3.14

GREEN: 1-output, Lookback 100, RMSE for Low: 2.81

YELLOW: 1-output, Lookback 150, RMSE for Low: 1.69

|  |  |  |
| --- | --- | --- |
|  | 100 | 150 |
| 8-May | 286.18 | 295.36 |
| 9-May | 283.38 | 292.02 |
| 10-May | 281.35 | 290.52 |
| 13-May | 281.55 | 291.15 |
| 14-May | 280.18 | 289.2 |



BLUE: 5-output RMSE for Close: 3.13

GREEN: 1-output, Lookback 100, RMSE for Close: 3.71

YELLOW: 1-output, Lookback 150, RMSE for Close: 6.91