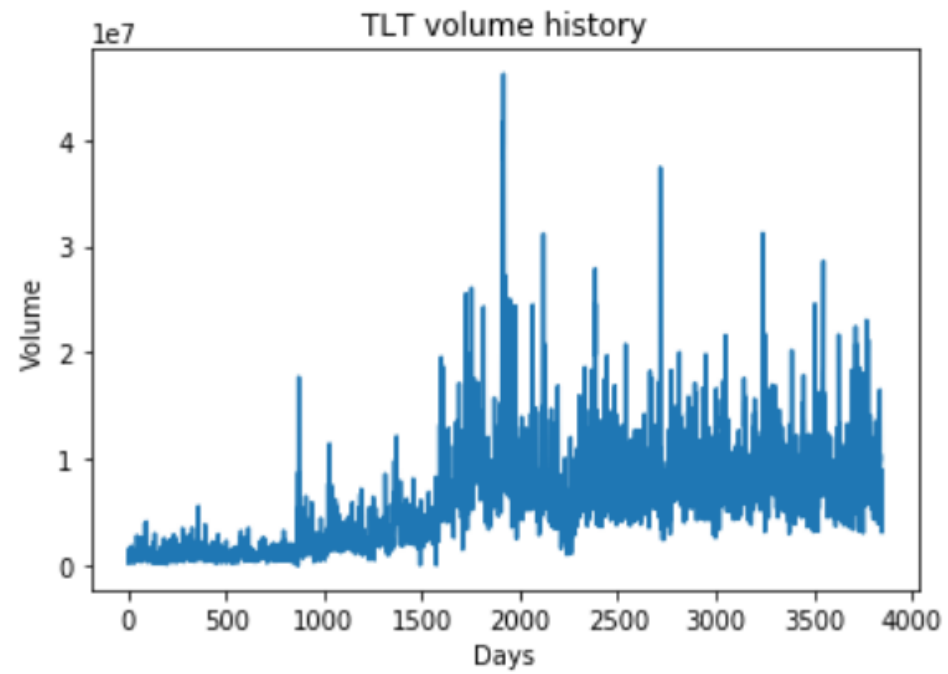
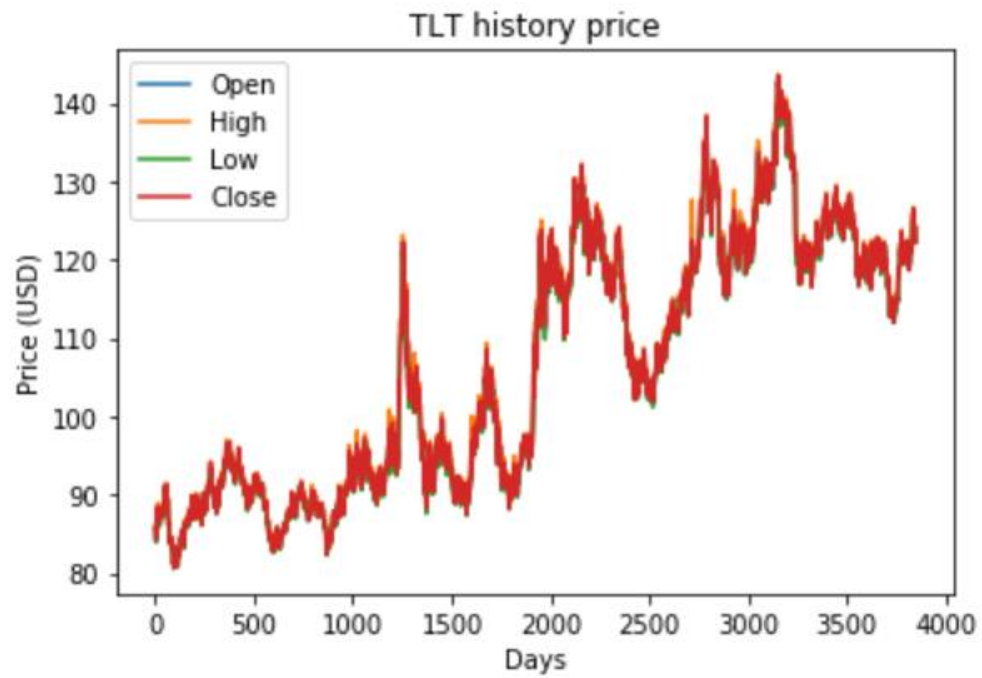


Prediction of TLT price with LSTM

Gan Luan, Wenlong Feng

Raw Data



Data Preprocessing

- ▶ New features
 - ▶ Date → day of the week (0~4) & month of the year (0~11)
 - ▶ Average open, high, low, and close price, respectively, in
 - ▶ Prior 90 trading days
 - ▶ Prior 180 trading days
 - ▶ Prior 360 trading days
 - ▶ 19 features in total
- ▶ Data normalization
 - ▶ MinMaxScaler → 0.0~1.0
 - ▶ Except day of the week & month of the year

Data Preprocessing

- ▶ Converting data to time series
 - ▶ X: 3-D matrix, [batch, Time_step, Features]
 - ▶ Y: 2-D matrix, [batch, Features]
 - ▶ Random shuffle
 - ▶ Train:Valid:Test = 8:1:1
 - ▶ Trim data, integral multiple of batch size

| | Date | Open | High | Low | Close | Volume | Week | Month |
|---|------------|-------|-------|-------|-------|----------|------|-------|
| 0 | 2002-07-26 | 82.67 | 82.80 | 82.42 | 82.51 | 316300.0 | 4 | 7 |
| 1 | 2002-07-29 | 82.06 | 82.16 | 81.32 | 81.42 | 8400.0 | 0 | 7 |
| 2 | 2002-07-30 | 81.75 | 81.90 | 81.52 | 81.52 | 6100.0 | 1 | 7 |
| 3 | 2002-07-31 | 81.95 | 82.80 | 81.90 | 82.53 | 29400.0 | 2 | 7 |
| 4 | 2002-08-01 | 82.54 | 83.02 | 82.54 | 83.00 | 25000.0 | 3 | 8 |
| 5 | 2002-08-02 | 83.16 | 84.10 | 82.88 | 83.85 | 52800.0 | 4 | 8 |

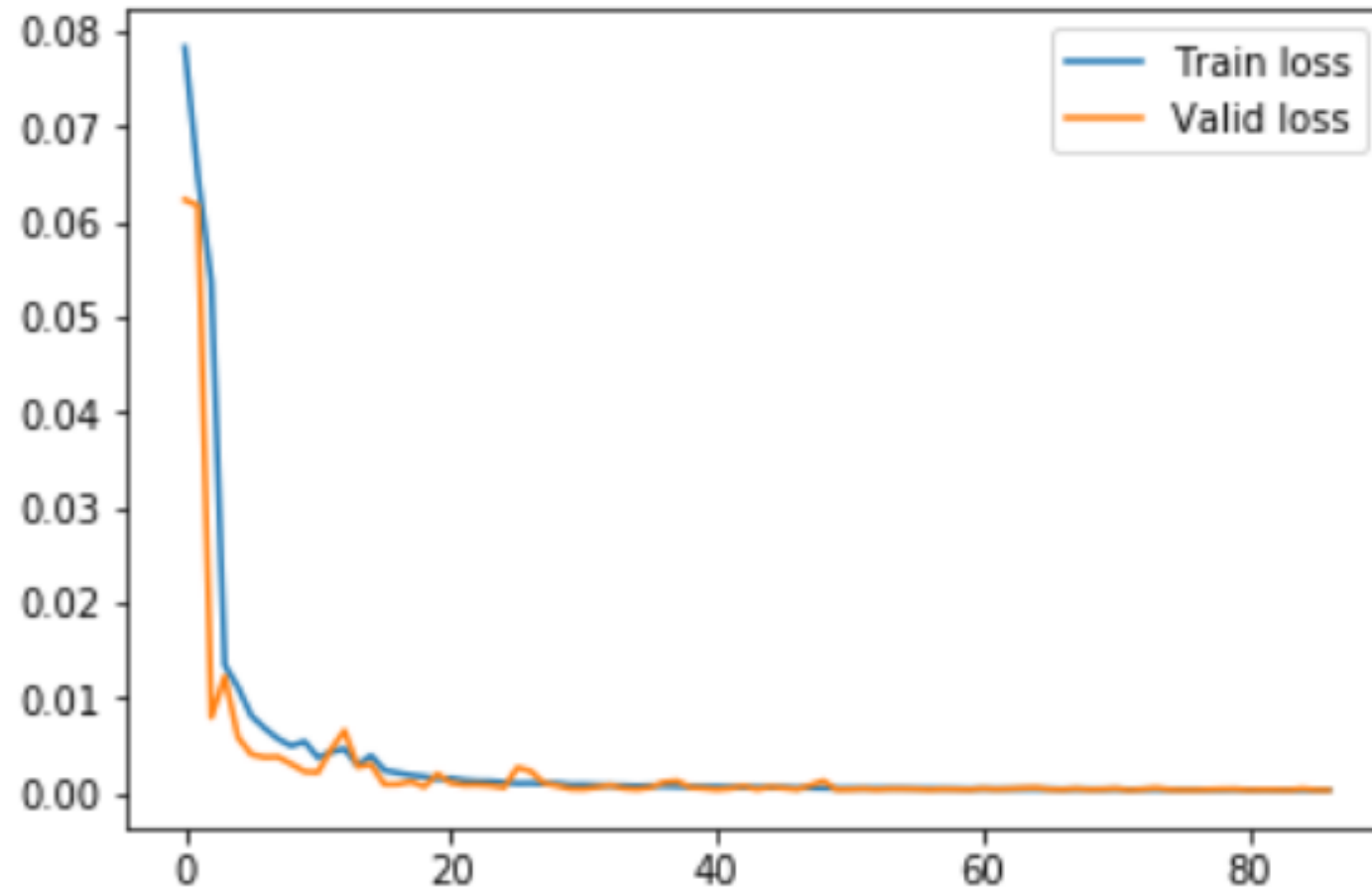
Model Architecture

| Layer (type) | Output Shape | Param # |
|------------------------------|----------------|---------|
| ===== | ===== | ===== |
| lstm_1 (LSTM) | (32, 60, 1000) | 4080000 |
| dropout_1 (Dropout) | (32, 60, 1000) | 0 |
| lstm_2 (LSTM) | (32, 60, 1000) | 8004000 |
| dropout_2 (Dropout) | (32, 60, 1000) | 0 |
| lstm_3 (LSTM) | (32, 1000) | 8004000 |
| dropout_3 (Dropout) | (32, 1000) | 0 |
| dense_1 (Dense) | (32, 200) | 200200 |
| dense_2 (Dense) | (32, 4) | 804 |
| ===== | ===== | ===== |
| Total params: 20,289,004 | | |
| Trainable params: 20,289,004 | | |
| Non-trainable params: 0 | | |

Hyperparameters

- ▶ Time_step: 30, 60
- ▶ Learning rate: 0.0001, 0.005, 0.001
- ▶ Batch size: 60, 32
- ▶ Epochs, patience: (100, 5), (100, 10), (100, 8), (200, 10), (200, 15),
- ▶ Optimizer: RMSprop, Adam
- ▶ LSTM layers: 3, 1
- ▶ Dropout: 0.2, 0.5

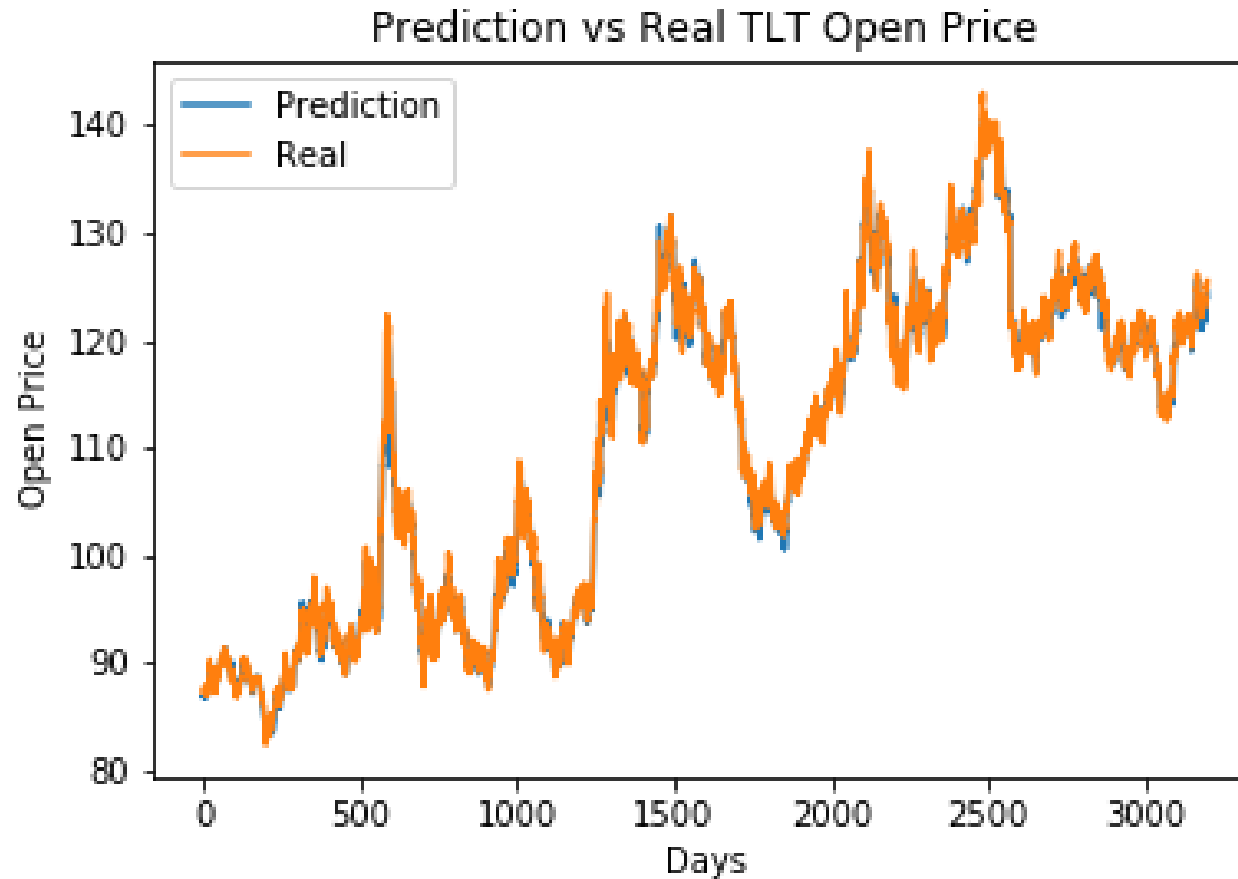
Tran loss vs. Valid loss



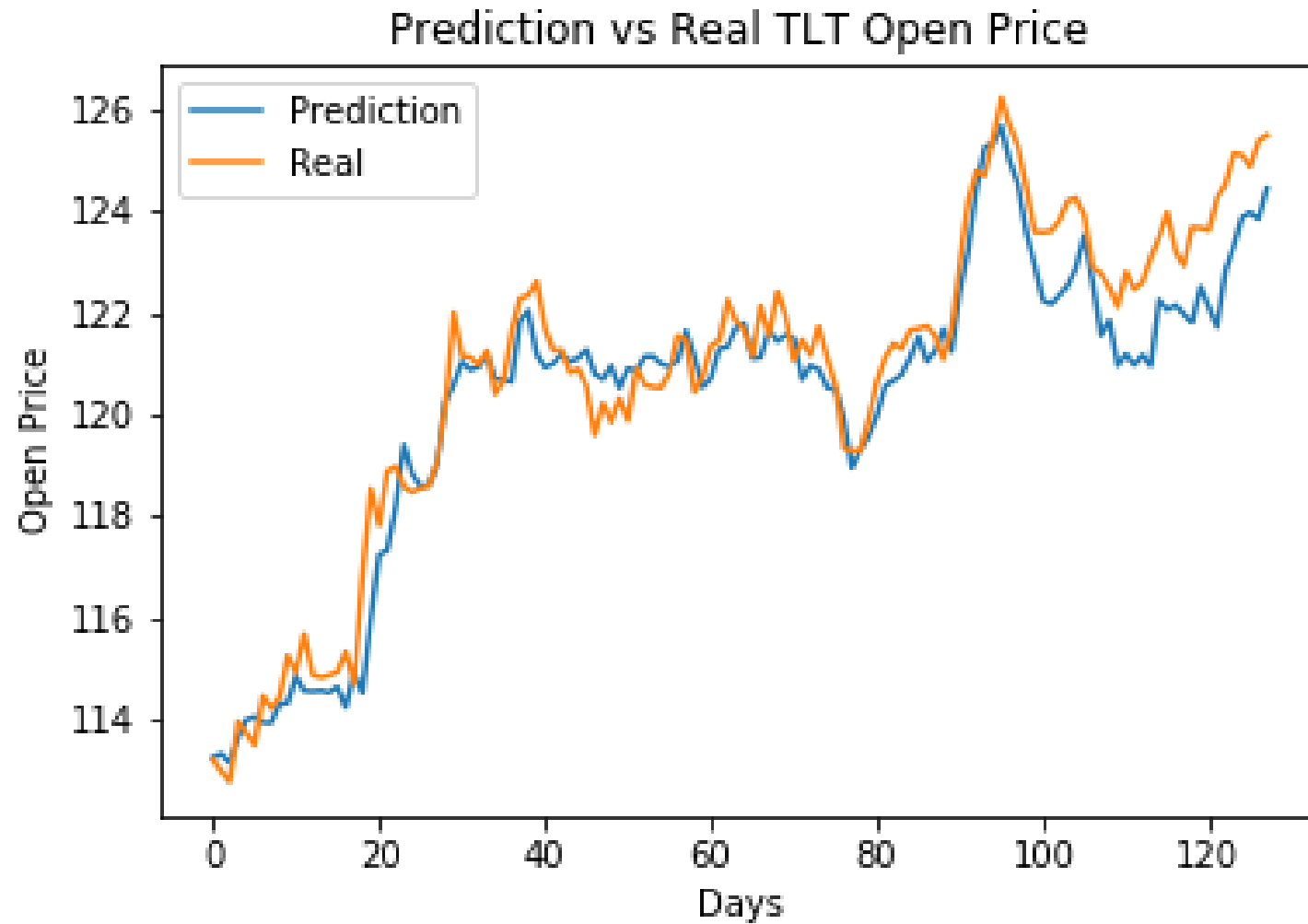
Model Performance Evaluation

Predict the TLT price for the past 3200 days, Mean square error is 0.905,

$$\frac{\sqrt{0.905}}{110} = 0.9\%$$



Model Performance Evaluation



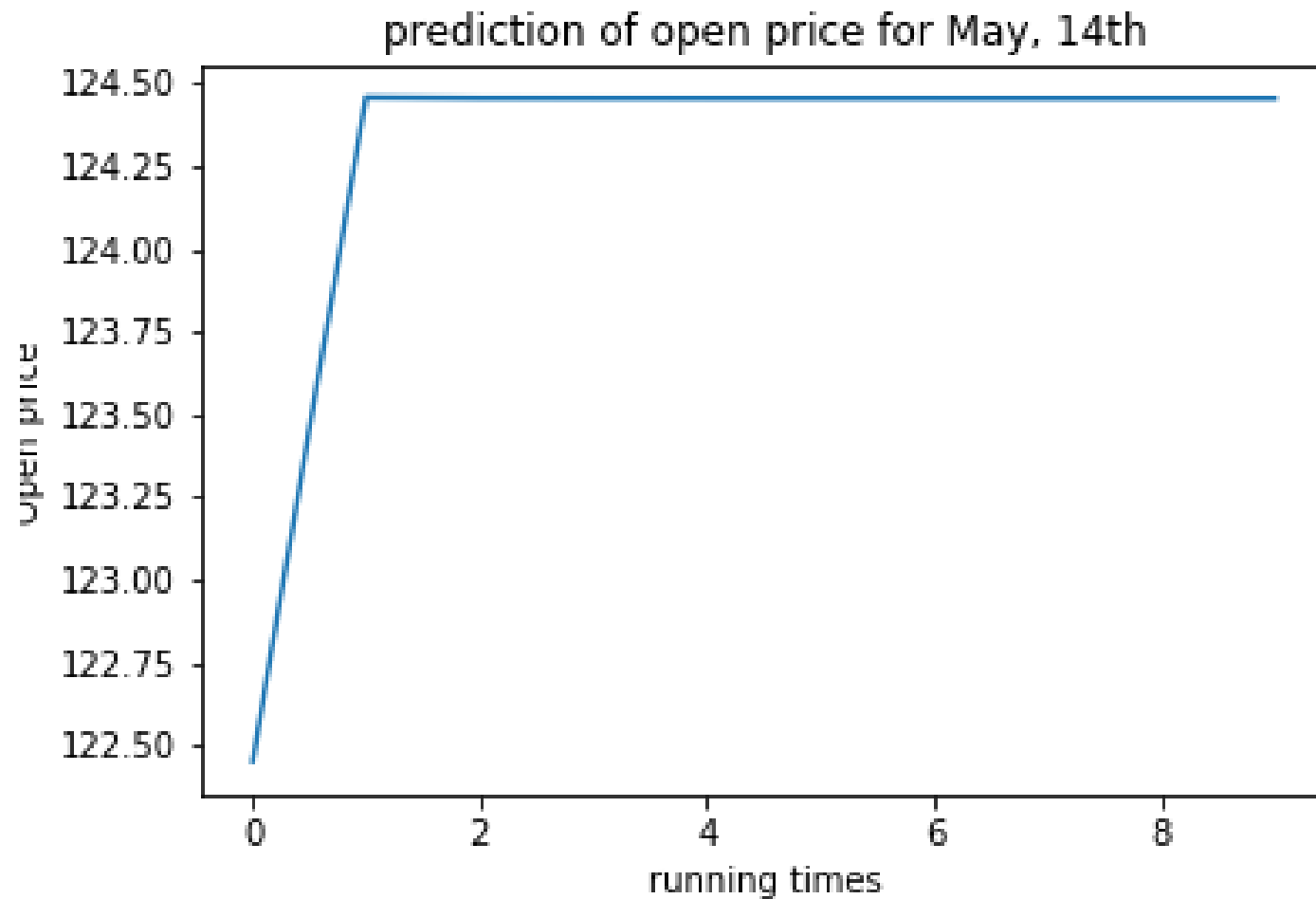
Limitations and future directions

- Performance at turning points
 - Adding more features
- Prediction for recent data
 - Overfitting
 - Updated the model with recent data
- Use of Sigmoid function as the final activation function
 - Use linear function instead

Limitations and future directions



```
yp = model.predict(xt)
```





Thank you!