IBM Specialised Models: Time Series Analysis and Survival Analysis: Tesla Share Price Prediction

This project seeks to use various LSTM models in order to attempt to predict the future share prices of the company Tesla Inc. Share price prediction is invaluable in the world of finance as it can often determine whether one invests in a company or not. Using recurring neural networks, we hope to show how this can be done computationally using deep learning methods.

1 Introduction

The data used is taken from yahoo finance via Pandas web reader as describes the (adjusted) share price of Tesla Inc between 2014 and 2020.

The main objective of this analysis is to determine the viability of using LSTM's in share price prediction and to accurately predict the short term change in Tesla's share price.

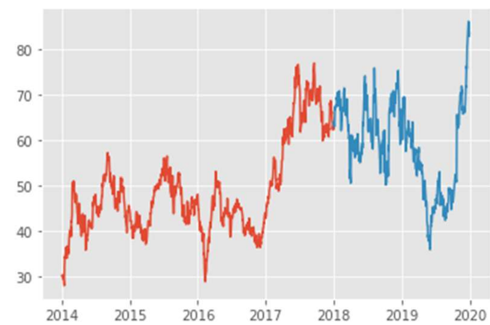

Figure 1: Tesla share price (adjusted) split into a training set (red) and a test set (blue)

2 Data Exploration and Cleaning

Though the data was for the daily adjusted close, due to the nature of financial share price reporting, the data was not periodic as weekend and holiday data was not recorded. In order to fix this, the data was reindexed in order to make it daily periodic and missing data points, using linear interpolation to fill missing values.

A single timeseries is used showing Tesla's share price in 10s of USD with the dataset ranging from 2014 to 2020. As it is a single timeseries, no other features were included and this series also compromised the target variable – ie future values of the timeseries.

As the aim was to show that future predictions could be made, the training data was taken from 2014 to 2018 while the testing data was for 2018 to 2020.

In order to fit the data with the recurring neural networks, it was necessary to scale it with a standard scaler.

3 LSTM Models

For all 3 models, only 10 input steps were used and the LSTM layers each had 5 units. To show how the model performed, the model was run every 100 periods, using the output as input to the next sequence iteratively for 40 future steps.

3.1 Simple LSTM

Long short term memory (LSTM) is a type of layer in a recurring neural network that stops the tendency to 'forget' earlier information in a sequence. The predictions from this model can be seen in Figure 2.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 50)                10400
_________________________________________________________________
dense_1 (Dense)              (None, 1)                 51
=================================================================
Total params: 10,451
Trainable params: 10,451
Non-trainable params: 0
_________________________________________________________________
```

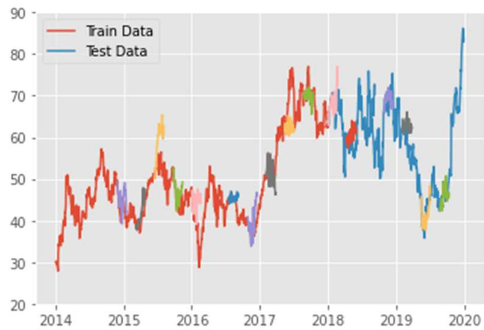

Figure 2: Predictions from the simple LSTM

3.2 Bidirectional LSTM

It is noted that sometimes by using a bidirectional LSTM layer that outputs may improve results. Here a single bidirectional layer was used with hyperparameters similar to that used in the previous simple LSTM. The results can be seen in Figure 3.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
bidirectional (Bidirectional (None, 100)               20800
_________________________________________________________________
dense_4 (Dense)              (None, 1)                 101
=================================================================
Total params: 20,901
Trainable params: 20,901
Non-trainable params: 0
_________________________________________________________________
```

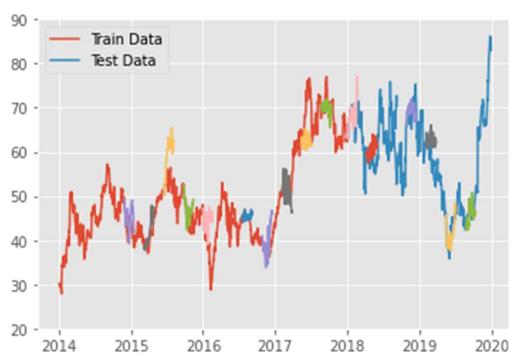

Figure 3: Predictions from the Bidirectional LSTM

3.3 Multilayer LSTM

For more complex series, it is noted that more complex patterns can be modelled using multiple layers within the model – taking the full output from every cell in the first LSTM layer and feeding it into the second. The results for our two-layer LSTM model can be seen in Figure 4.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_4 (LSTM)                (None, 10, 50)            10400
_________________________________________________________________
lstm_5 (LSTM)                (None, 50)                20200
_________________________________________________________________
dense_3 (Dense)              (None, 1)                 51
=================================================================
Total params: 30,651
Trainable params: 30,651
Non-trainable params: 0
_________________________________________________________________
```

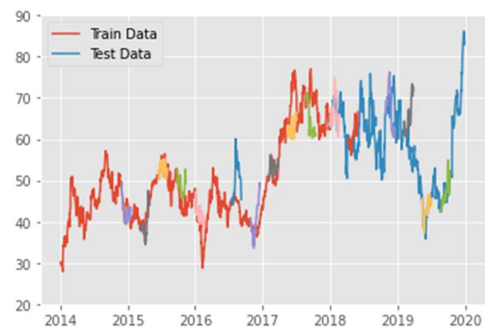

Figure 4: Predictions from the two-layer LSTM

4 Key Findings

From the results shown in Figures 2,3 and 4 it is clear that the multilayer LSTM and bidirectional LSTM perform much better than the simple LSTM. This is likely due to the complex nature of share prices meaning the more complicated models can better fit underlying patterns. Though the models all performed fairly well for a 40-day prediction, longer predictions were often flawed due to the cumulative nature of the error – an erroneous result will be used as an input to the next iteration of model prediction.

5 Possible Flaws

LSTMs are prone to overfitting, a key concern when modelling our data. Possible ways of countering this include using a validation set to know when to stop training the model. Additionally, due to the large number of parameters of LSTMs, they often need to be trained using GPUs or TPUs and can take much longer to train than alternative methods – a key reason for the shorter input sequence that we used.

6 Next Steps

For future steps, it is interesting to consider using the percentage growth/difference as oppose to absolute share price to make predictions. Additionally, using longer input sequences may improve the accuracy of our models. By far the most interesting future possibility is that of doing multi-feature prediction – one could combine multiple stocks with various macroeconomic indicators in order to make better predictions for a wider range of stocks.