Notes:

- Not all data is included in the zip file to keep file sizes reasonable, but everything required to run the various parts should be there.
- Hyperparameters-TSLA-intraday-SMA.csv file lists the different hyperparamter combinations tested for the simple moving average predictions of the TSLA stock. Parameters not listed in this file are set to the default values seen in the src\StockPrediction\nnconfig.py file.
- The models are setup to run on a GTX 1080, while they may run on a cpu this hasn't been tested and would most likely take a long time to run anyway. The RoBERTa models will crash if run on a GPU with less than 8GB of memory however the stock prediction models should work fine with less.
- A prediction model that used interday – daily stock prices was also developed but was not kept up to date so most likely will not run. By default, everything is set to use the intraday model/data.
- Adding CNNs to stock prediction model broke model saving/loading so models must be trained from scratch currently, the default parameters are set for the best performing hyperparameters. Saving/loading of sentiment analysis models does work.

Requirements:

- While models may work if run on a CPU running on a GPU is preferable, this project was tested using CUDA version 11.1.1
- Python version 3.8 required (pytorch doesn't support 3.9)

Installation Steps:

- Create venv inside root folder
- Run venv activate script
- Pip install -r requirements.txt

Configuration:

- Stock prediction models can be configured in the file: src\StockPrediction\nnconfig.py
    - By default this is set to run with the best known parameters however the use_default parameter of the method generate_param_dict in the same file can be set to False in order to tested different parameter combinations specified in the same method.
- Sentiment analysis models can be configured in the file: src\SentimentAnalysis\robert_config.py
    - By default this is set to load the best model for classifying TSLA posts.

Running the program:

python -m src.StockPrediction.stock_prediction.py

- All scripts should be run with the working directory set to the root folder i.e. wherever the contents of the zip file are placed.
- src\StockPrediction\stock_prediction.py runs the actual stock prediction model

- src\SentimentAnalysis\roberta_model_generation is used for creating/testing sentiment analysis models
- src\SentimentAnalysis\roberta_data_generation loads the best model and uses it to classify and save the data for the stock prediction model

Code also available on school gitlab: https://git-teaching.cs.bham.ac.uk/mod-ug-proj-2020/rxt711/-/tree/master