Final Year Project – brief

What I have done:

* Researched traditional methods and indicators for predicting the stock market
  + Moving averages
  + Exponential smoothing
  + Arima (not a lot)
* Researched a range of model & methods for predicting time series (regression)
  + Random forest
  + Support vectors
  + Neural networks
    - One hidden layer
    - Two hidden layers
* Created Juypter notebook for each model
  + Takes apples stock from 2015-2023 as input
  + Sanitize and preprocess the data
  + Split data into training and testing groups
  + do a grid search to find the best hyper parameters for the model using MAE (mean absolute error)
  + Test the model on apple stock 01/01/2024 – 01/06/2024

Saved a few sources that have a good range of references I can explore

Possible Todo:

* Use different stocks
* Use different metrics to measure error
* Analyse and compare classification (Accuracy) (confusion matrix) (more important than accuracy)
* Expand inputs (feature engineering)
  + Moving averages
  + Indicators like (RSI)
  + Macroeconomic indicators (GDP growth, interest rates)
* Scrape web (reddit, X, the usual) and incorporate sentimental analysis
  + How to get data only available at time X?
  + Additional input that it wasn’t trained on? (post predicition)
* Expand prediction window

**January:**

Skeleton of report fleshed out with some details and a explanation of how much you’ve done and what you’re going to do

**Questions:**

* How many references? how many per technique – as many somewhere around 10-15
* Worried about scope – do I expand horizontally or vertically - Sentimental analysis could be quite complicated – do I have time?
* How much detail should I go into in the report about each method?
* Word count again (because I didn’t write it down) 10-12,000 words
* Content enough for a first?

Sources for lit review:

**TIME SERIES and pre ai/ml models**

* Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. **Book about arima and other models**
* Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. **Exponential smoothing**
* Hamilton, J. D. (1994). *Time Series Analysis*.
* Engle, R. F. (1982). *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation*. **ARCH model**
* Bollerslev, T. (1986). *Generalized Autoregressive Conditional Heteroskedasticity*. **ARCH to GARCH model**

**Random Forests**

* Breiman, L. (2001). *Random Forests*.
* Biau, G., & Scornet, E. (2016). *A Random Forest Guided Tour*.
* Khaidem, L., Saha, S., & Dey, S. R. (2016). *Predicting the Direction of Stock Market Prices Using Random Forest*.
* Wei, W., & Chaudhary, S. (2020). *A Random Forest Approach for Time Series Prediction of Cryptocurrency Prices*.

**Support vector machines**

* Cortes, C., & Vapnik, V. (1995). *Support-Vector Networks*.
* *Schölkopf, B., et al. (1999).Advances in Kernel Methods: Support Vector Learning*.
* Tay, F. E. H., & Cao, L. (2001). *Application of Support Vector Machines in Financial Time Series Forecasting*.
* Huang, W., Nakamori, Y., & Wang, S. Y. (2005). *Forecasting Stock Market Movement Direction with Support Vector Machine*.
* Kim, K. J. (2003). *Financial Time Series Forecasting Using Support Vector Machines*.

**NEURAL NETWORKS**

Kolarik and G. Rudorfer, “Time series forecasting using neural networks, department of applied computer science,” Vienna University of Economics and Business Administration, no. 1090, pp. 2–6, 1997.

* Kaastra and M. Boyd, “Designing a neural network for forecasting financial and economic time series,” Neurocomputing, vol. 10, pp. 215–236, 1996
* Chong, E., Han, C., & Park, F. C. (2017). *Deep Learning Networks for Stock Market Analysis and Prediction: Methodology, Data Representations, and Case Studies*.
* Fischer, T., & Krauss, C. (2018). *Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions*.
* Patel, J., et al. (2015). *Predicting Stock Market Index Using Fusion of Machine Learning Techniques*.
* Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). *Predicting Direction of Stock Price Index Movement Using Artificial Neural Networks and Support Vector Machines*.
* Qiu, M., & Song, Y. (2016). *Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model*.
* Rumelhart, D. E., et al. (1986). *Learning Representations by Back-Propagating Errors*. **Backpropagation**
* LeCun, Y., et al. (2015). *Deep Learning*. **Deep learning techniques**
* Hochreiter, S., & Schmidhuber, J. (1997). *Long Short-Term Memory*. **Long short term memory**

**Others**

* Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). *Predicting Stock Market Index Using Fusion of Machine Learning Techniques*
* Ballings, M., Van den Poel, D., Hespeels, N., & Gryp, R. (2015). *Evaluating Multiple Classifiers for Stock Price Direction Prediction*.

**NOTES**

Comparison mode – base line

* Use previous day
* Use lit review of similar models