**Multivariate Time Series Prediction for Stock Market Data**

資科一 資科一 資科二

曹昱維 鄭詠儒 謝政彥

110753201 110753126 109753207

**Introduction**

　　時間序列資料是由按照時間發生先後順序進行排列的數據點序列，以股票市場資料為例，從2000年至2020年為止，股票A的每日收盤股價就是一種時間序列資料。我們希望透過單一變數的時間序列資料(如股票A的每日收盤價)以及多變數的時間序列資料(如所有能源類股的每日收盤價)來訓練各種模型，來比較單變數與多變數對於預測效果的影響，也同時比較相同類型的變數下各種模型的預測效果。

　　股票市場以波動性、動態性和非線性著稱。由於政治、全球經濟狀況、突發事件、公司財務業績等多重（宏觀和微觀）因素，準確預測股價極具挑戰性。但是，所有這一切也意味著有大量數據可供尋找模式。因此，金融分析師、研究人員和數據科學家不斷探索分析技術來檢測股市趨勢。股票分析基本上可利用基本面分析與技術分析，本專案嘗試針對S&P 500股價，利用深度學習模型以及傳統的時間序列模型進行預測。

**Problem & Solution**

1. 透過單變數時間序列訓練ARIMA Model, RNN-based Model, Transformer-based Model來預測下一筆資料(many to one)，並比較同樣的變數類型下，各種模型的預測效果
2. 透過多變數時間序列訓練ARIMA Model, RNN-based Model, Transformer-based Model來預測下一筆資料(multivariate, many to one)，並比較同樣的變數類型下，各種模型的預測效果
3. 透過多變數時間序列訓練RNN-based Model, Transformer-based Model來預測接下來的數筆資料(multivariate, many to many)，並比較同樣的變數類型下，各種模型的預測效果

**Related Work**

1. Predicting Stock Prices Using Machine Learning: <https://neptune.ai/blog/predicting-stock-prices-using-machine-learning>
2. Prediction of price for ML with finance stats: <https://www.kaggle.com/hanseopark/prediction-of-price-for-ml-with-finance-stats/data>
3. Time-Series Forecasting: Predicting Stock Prices Using An LSTM Model: <https://towardsdatascience.com/lstm-time-series-forecasting-predicting-stock-prices-using-an-lstm-model-6223e9644a2f>
4. Berkshire Hathaway - Stock Time Series Analysis: <https://www.kaggle.com/kalilurrahman/berkshire-hathaway-stock-time-series-analysis>
5. A Transformer-based Framework for Multivariate Time Series Representation Learning: <https://dl.acm.org/doi/abs/10.1145/3447548.3467401>
6. A Multivariate Time Series Modeling and Forecasting Guide with Python Machine Learning Client for SAP HANA:

<https://blogs.sap.com/2021/05/06/a-multivariate-time-series-modeling-and-forecasting-guide-with-python-machine-learning-client-for-sap-hana/>