

---

# COSE474-2024F: Final Project Proposal

## “Predicting Future Nvidia Stock Price using Pre-trained Foundational Model ”

---

Muhammad Faiz Chan

### 1. Introduction

Nvidia stock has been rising to become the second largest company in the United States by market capitalization. In the S&P 500 index, it is also ranked second in the amount of weights it contributed to the index. Thus, a lot of people invest and trade with this stock possibly for long term or short term gains. If someone invested 1,000 US Dollars into the stock 5 years ago with the dividends reinvested, the person would have around 28,000 US Dollars by today which makes predicting its future prices crucial in profit making or minimizing losses.

### 2. Problem definition & Challenges

The core objective of this project is to harness the power of pre-trained deep learning models, which are typically used in natural language processing (NLP), to predict Nvidia's stock prices. The aim is to model complex dependencies in both time series financial data and external textual data like financial news and social media. Pre-trained models are well-suited for learning patterns, extracting deep contextual features, and can potentially outperform traditional statistical and shallow machine learning methods in understanding financial markets.

#### Challenges

- External Events: there are a lot of factors that could affect the stock market and one of them are events that are happening in the world such as financial crises or the troubles within the company
- Outliers: because of the same reason as the above, there will be outliers in the dataset as well which resulted in people selling or buying the stock at a fast rate
- Huge Number of Features: Predicting stock prices also involved in understanding the company behind it and what products it sells which could drive the stock prices up or down. Thus, feeding a number of this features in the model would be challenging.

### 3. Related Works

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, 270(2), 654-669.

### 4. Datasets

Datasets about daily Nvidia stock prices from kaggle that are a few months before so that we can compare the accuracy of the model against the real thing.

### 5. State-of-the-art methods and baselines

In this section, we will establish baseline models such as ARIMA, LSTM, and Random Forest, and compare them to the latest state-of-the-art methods, including TFT, Informer, and FinBERT.

### 6. Schedule & Roles

- Acquire the dataset and study the model
- Figure out the coding that is needed to get the desired output
- Test the model using a smaller dataset and optimize it
- Fine tune the model to get a better performance and accuracy
- Document the task and finalize the project

### References