***TO DO: GIVE MINI INFO ABOUT PROJECT YOU WANT TO DO BELOW***

Option1:

Stocks

* time series analysis
* Portfolio selection, signaling (NN with time series)
  + Hedge funding (different types of securities);
  + Value trader: based on fair market value of company;
  + Concern: size of available datasets for NN;
  + Multi-layer convolutional NN;

Historical NASDAQ data(10 years): <https://www.nasdaq.com/market-activity/quotes/historical>

S&P 500: <https://www.kaggle.com/camnugent/sandp500>

Option2:

* <https://data-flair.training/blogs/machine-learning-project-ideas/>

Ritu: <https://drive.google.com/drive/folders/19Mk6ceEfqqKXwGC-HXZ2AueBs-X6qtO4?usp=sharing>

Master folder: <https://drive.google.com/drive/folders/1qH2no8msJnZzVftmM7b03DCYRqIF-uup>

**Oct 31st meeting:**

Pick up a few companies, and predict their stock prices in the next period:

* 5 different algorithms we apply to the datasets to predict the stock price
  + **Linear regression**
  + Neural Networks - MLP(Weight initialization, loss function, Optimizer, Activation function)
  + **SVRs**
  + **Random forest regressor**
  + Logistic Regression
* Feel free to look up algorithms that you want to implement

<https://www.nasdaq.com/market-activity/quotes/historical>

(FAANG)

NEXT MEETING: ASSIGNING PARTS TO INDIVIDUALS - EVERYONE MUST BE PRESENT PLS

**FIND WHAT AFFECTS STOCK PRICES: WILL USE THESE AS PREDICTORS FOR STOCK PRICES**

* **GAS PRICES**
* **NASDAQ Done**
* **S&P 500 index - Done**
* **GDP growth rate**
* **inflation rate - Done**
* **interest rate - Done**
* **Money supply**

[**https://www.researchgate.net/publication/338913676\_ANALYSIS\_OF\_INTERNAL\_AND\_EXTERNAL\_FACTORS\_AFFECTING\_TO\_THE\_STOCK\_MARKET\_PRICE\_OF\_NON-\_FINANCIAL\_JOINT\_STOCK\_COMPANIES\_LISTED\_ON\_THE\_VIETNAM'S\_STOCK\_MARKET**](https://www.researchgate.net/publication/338913676_ANALYSIS_OF_INTERNAL_AND_EXTERNAL_FACTORS_AFFECTING_TO_THE_STOCK_MARKET_PRICE_OF_NON-_FINANCIAL_JOINT_STOCK_COMPANIES_LISTED_ON_THE_VIETNAM'S_STOCK_MARKET)

**IMF:** [**https://datahelp.imf.org/knowledgebase/articles/729291-table-how-to-select-countries-for-which-i-want-da**](https://datahelp.imf.org/knowledgebase/articles/729291-table-how-to-select-countries-for-which-i-want-da)

**Yield curves:**

[**https://www.treasury.gov/resource-center/data-chart-center/interest-rates/pages/TextView.aspx?data=yieldAll**](https://www.treasury.gov/resource-center/data-chart-center/interest-rates/pages/TextView.aspx?data=yieldAll)

**Nov-17th**

**Decided IBM stock (Keep the latest 20 years)**

1. **got 59 years data of IBM**
2. **Crude oil 60 year price data**

* **Related data** 
  + **commodities - mineral/raw materials/ crude oil price**
  + **GPD growth rate - Quarterly**

| **Day 1 - 10** | **7% - GDP** |
| --- | --- |
| **Day 2 -11** | **7% - GDP** |
| **Day 3 - 13** | **7% - GDP** |
|  |  |

**What we need to do**

1. **Data cleaning**
   1. **Null values**
   2. **Missing values**
   3. **Outliers**
2. **Variable selection (Correlation/ SHAP)**
3. **Linear regression**
4. **Random forest Regressor**
5. **XGBoosting**
6. **Comparing performance - MSE, ROC-AUC, R2**