Stock Price Prediction System

- **Due:** *Each Task will have its own due date. Please refer to the Task specification.*
- **Contributes** 50% of your final result
- **Individual Assignment – You will work in a project led by a teaching staff**

Summary

This project requires you to work in a project led by your teaching staff aiming to implement and demonstrate a stock price prediction system, **FinTech101**. Many people have high hopes for AI and machine learning (ML) technologies, believing that AI can do "magical" things including predicting the future. Surely, if AI is really smart, it should be able to tell you whether the share price of a company on the stock market will go up or down in the future. In this project, we will put this idea to the test by trying to predict the stock price of a company using the historical stock price of that company.
If you join this project, you will work under the supervision of your project leader, a friendly COS30018 teaching staff. Note that, even if there are many students who take this option, your project is completely independent of their projects. You will only work with the project leader and will be solely responsible for the outcome of the project.
The starting code for this project has been provided by an online resource as a single Python file. The code works and provides the basic function for this project. You will start from this basic code base. You will see that the given code base has many issues and limitations that require you to rectify. You will be given weekly tasks to allow you to work toward that goal.

The **FinTech101** system will download the historical data about the stock price of a specified company such as Amazon (AMZN), Facebook (FB), Commonwealth Bank of Australia (CBA.AX), etc. using an existing library (such as Yahoo Finance) and train a machine learning model to allow future stock price of the company to be predicted.

Project requirements

- The project will be done in Python. If you have zero experience with Python, this project is not suitable for you unless you are willing to learn the language in your own time and within one or two weeks. An easy way to test yourself against this requirement is to look at the Task 1 to see if you can handle it without any problem. If you already struggle with this Task, this project is not suitable for you.
- Source code maintained on Git based VCS (Github/Bitbucket/GitLab/…). You must provide read-only access to the tutor/lecturer
- Running illustrative demo of a working prototype
- Portfolio consisting of your weekly reports and a summary report. Detailed requirements of your weekly report will be included in the specification of the weekly tasks to be released on Canvas.
- The project summary report (8-10 pages) includes the following sections
    - Cover Page (with team details) and a Table of Contents (TOC)
    - Introduction
    - Overall system architecture
    - Implemented data processing techniques,
    - Experimented machine learning techniques,
    - Details of the extensions you have developed (in **Task 7**) **if applicable**
    - Scenarios/examples to demonstrate how the system works,
    - Some critical analysis of the implementation, and
    - Summary/Conclusion.

Option B – Marking Scheme:

How does the bonus marks (+3) work?
The bonus marks will be added to your total but the total mark for Option B cannot exceed 100 marks.

Task	Deadline	Mark
Task B.1: Set up your environment for the project including on your local machine and a github repository; run the given code bases.	11:59pm 18/08/2023 (Week 3)	10 (+3)
Task B.2: Data Processing 1: Load and process datasets with multiple features; use different methods to split the data into train/test data; store/load data locally for future uses.	11:59pm 25/08/2023 (Week 4)	10
Task B.3: Data Processing 2: Use different methods to visualise your data; store a visualisation of your data as an image for image processing.	11:59pm 1/09/2023 (Week 5)	10
Task B.4: Machine Learning 1: Use a more general way to construct the model. Try different DL networks: RNN/GRU/Simple FF. Experiment with different hyperparameter configurations.	11:59pm 8/09/2023 (Week 6)	10
Task B.5: Machine Learning 2: Solve more advanced prediction problems including multivariate prediction and multistep prediction.	11:59pm 22/09/2023 (Week 7)	15
Task B.6: Machine Learning 3: Work on ensemble approach/combining multiple approaches for improving prediction quality.	11:59pm 6/10/2023 (Week 9)	15
Task B.7: Simple Extension beyond B.1-6 -- Combine stock price data with other types of data (e.g., the number of times the company is mentioned on Twitter, etc.) – to be suggested and supervised by the project leader.	11:59pm 29/10/2023 (Sunday Week 12)	10
Task B.7: Comprehensive extension + Independent Research (approved by unit convenor) -- You initiate the extension by yourself and conduct an independent and outstanding research.	11:59pm 29/10/2023 (Sunday Week 12)	20
Total		100
You need to follow good programming practice (e.g., well-designed, well-structured codes with clear and helpful comments). Failure to do so get penalty.		**Up to -20**

Submission

You must email your complete portfolio of work to bvo@swin.edu.au and to your tutor by 11:59pm on 29/10/2023. Create a single zip file with your code and a working version of your system. Standard late penalties apply – 10% for each day late, more than 5 days late is 0%.

You must also provide to bvo@swin.edu.au and to your tutor read only access to your git repository before Week 3 (see the Task 1 requirements).