Artificial Intelligence – Week 6

Instructor: Teerapong Panboonyuen

Course Repository:
https://github.com/kaopanboonyuen/SC310005_ArtificialIntelligence_2025s1

🎯 Objective

This week, you will perform **time series forecasting** on 5 selected Thai stocks to predict their daily closing prices. Your goal is to build forecasting models that minimize the RMSE score on the test data.

You will:

- Understand and preprocess time series stock data
- Explore different forecasting techniques (e.g., Moving Average, machine learning, deep learning)
  Train and evaluate models to predict stock prices
- Visualize actual vs. forecasted stock prices
  Identify which method gives the best forecasting accuracy (lowest RMSE)

📚 Dataset Description

We will use a combined dataset of 5 Thai stocks updated through July 31, 2025:

📁 Dataset URL:

https://github.com/kaopanboonyuen/SC310005_ArtificialIntelligence_2025s1/raw/main/dataset/kku_homework_stock_dataset.csv

The dataset contains:

- Date: Trading date
- Close: Closing price of the stock
- Volume: Trading volume
- Stock: Stock symbol (one of AOT, BDMS, BAY, ESSO, HMPRO)

📝 Assignment Instructions

1. **Download & Load Data**
    - Load the dataset with pandas
    - Handle any missing values or anomalies
2. **Exploratory Data Analysis (EDA)**
    - Plot time series trends of each stock's closing price
    - Summarize key statistics and visualize volume vs price
3. **Model Development**
    - Implement forecasting techniques such as:
        - Moving Average (MA) or Exponential Smoothing
        - Machine learning models (e.g., Random Forest, XGBoost)
        - Deep learning models (e.g., LSTM, GRU) using PyTorch or TensorFlow
    - Train on historical data and forecast future prices

4. **Model Evaluation**
    - Evaluate your models using RMSE and MAE on the test set
    - Compare and report which model achieves the lowest RMSE for each stock
5. **Visualization**
    - Plot actual vs. predicted closing prices for all stocks
    - Include clear legends, titles, and axis labels

🔥 **Example Approaches**

- Use pandas rolling windows to compute moving averages
- Train LSTM models on scaled time series sequences
- Use sklearn's RandomForestRegressor for regression forecasting
- Tune hyperparameters for best performance

💡 **Deliverables**

📔 **Jupyter Notebook / Colab:**

- Data loading and cleaning code
- EDA plots and insights
- Forecasting model implementations
- Evaluation metrics and comparison

- Visualizations of forecasts vs actual prices
- Clear comments and explanations

📄 **Documentation:**

- Summary of each step taken
- Discussion on model performance and best approach
- Suggestions for future improvements

🚀 **Getting Started**

💻 **Setup Example:**

```
!wget
https://github.com/kaopanboonyuen/SC310005_ArtificialIntelligenc
e_2025s1/raw/main/dataset/kku_homework_stock_dataset.csv

import pandas as pd

df = pd.read_csv('kku_homework_stock_dataset.csv')
```

📦 **Recommended Libraries:**

python

CopyEdit

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt
```

```
import seaborn as sns

from sklearn.metrics import mean_squared_error,
mean_absolute_error

import torch

import torch.nn as nn
```

📅 Submission Deadline

To be announced in class. Submit your `.ipynb` notebook via LMS or GitHub Classroom.

Good luck! 🚀