Task 7: Dimensionality Reduction & Stock Price Prediction

This project consists of two key objectives:

1. **Dimensionality Reduction** – Using **Principal Component Analysis (PCA)** for visualization.
2. **Stock Price Prediction** – Using **Time Series Forecasting with ARIMA**.

Part 1: Dimensionality Reduction

Objective:

Reduce high-dimensional data to **2D** for visualization.

Project Steps:

1. **Dataset Selection:**
    - Choose a dataset with **high-dimensional features** (e.g., **Iris dataset**).
2. **Tasks to Perform:**
    - **Apply PCA** to reduce the dataset to **two principal components**.
    - **Visualise the reduced data** using a **scatter plot**.

Deliverables:

✅ **Reduced dataset (2D representation).**
✅ **Scatter plot showing reduced dimensions.**

Part 2: Stock Price Prediction Using Time Series Forecasting

Objective:

Predict **future stock prices** based on historical data.

Project Steps:

1. Dataset Selection:

- Dataset Name: **stock_prices.csv**
- Columns:
    - **Date** (Timestamp)
    - **Open** (Opening price)
    - **Close** (Closing price)
    - **Volume** (Trade volume)

2. Tasks to Perform:

1. Load and Preprocess the Dataset

✅ Convert the **Date** column to **DateTime format**.
✅ Handle **missing values** (if any).
✅ Set **Date** as the index for time-series analysis.

2. Exploratory Data Analysis (EDA)

✅ Plot the **time series of Close prices** to observe trends.
✅ Analyze **seasonality, trends, and noise** in the data.

3. Feature Engineering

✅ Create **lag features** (previous day's close price as a feature).
✅ Perform **rolling window calculations** (moving averages, etc.).

4. Model Training

✅ Train an **ARIMA (AutoRegressive Integrated Moving Average) model** for forecasting.
✅ Tune ARIMA **(p, d, q) parameters** for better accuracy.

5. Model Evaluation & Visualization

✅ Compare **actual vs. predicted** stock prices.
✅ Plot the **forecast vs. real stock prices** for visualisation.
✅ Analyze **forecasting errors (MAE, RMSE, MAPE)**.

Deliverables:

✅ **Trained ARIMA model for stock forecasting.**
✅ **Time-series plots comparing predictions vs. actual prices.**
✅ **Insights on stock trends, seasonality, and forecast accuracy.**

Improvements Over the Previous Version

✔ Clearer project objectives & structure.
✔ More detailed steps for both PCA and Stock Forecasting.
✔ Added data preprocessing, feature engineering, and error analysis.
✔ Better deliverable clarity for actionable insights.

Deadline Compliance

- **Restriction**: **Submit the project within 7 days** from the start date.
- **Reason**: Meeting deadlines is crucial in the real-world software development environment. This restriction helps students practice **time management** and **task prioritization**. In professional settings, tight deadlines are often the norm, and learning to meet them without compromising quality is an essential skill.
- **Learning Outcome**: Students will learn to manage their time effectively, complete projects under pressure, and **deliver results on time**, which are all important skills in the workplace.