Finance and Economics Club,IIT Kharagpur
Certification Projects

1 .Machine Learning Bankruptcy Predictions:

Purpose:
To build a machine learning model that predicts whether a company is likely to go bankrupt based on financial indicators, helping in risk assessment and early warning systems.

Tools: Python, Jupyter Notebook,Scikit-learn,Pandas & NumPy ,Matplotlib/Seaborn (for data visualization)

What you'll learn:
- Data preprocessing and feature selection
- Classification algorithms (e.g., Logistic Regression, Random Forest, SVM)
- Model evaluation (accuracy, precision, recall, ROC-AUC)
- Interpreting financial data and extracting insights for decision-making

Dataset: https://www.kaggle.com/datasets/fedesoriano/company-bankruptcy-prediction

References:
- YouTube Course
  https://youtu.be/ax3XvWO0Pk4?si=HINB-4VcfEli1Von
  https://youtu.be/w5JE3iLyDtk?si=1S2EZrSO7NsuTVy2

2. Portfolio Optimization Tool

Purpose:
To optimize asset allocation using Modern Portfolio Theory and advanced models like Black-Litterman.

Tools:
 Python, `cvxpy`, `Riskfolio-Lib`, NumPy, pandas

What You'll Learn:
Risk-return trade-offs, Sharpe ratio, portfolio diversification, and simulations under various economic scenarios.

Dataset:
https://docs.google.com/spreadsheets/d/e/2PACX-1vRhjzfMq5pYYnwHaSp5agF7Ul3M2tSn2if4Z4nEPViOj7GBk8gZZQhyTZ7MP857FHgDND4qi7osjrMZ/pub?gid=0&single=true&output=csv

References:
- Youtube Playlist:
  https://youtube.com/playlist?list=PLcFcktZ0wnNnqefRpFMS1k9_VlhVw7bzc&si=uk6MNpw1RidN7sWL
- Medium Article:
  https://medium.com/@lorenzojcducv/an-introduction-to-portfolio-optimization-in-python-dcd32ea7b562

3. <u>ML for Asset Price Prediction</u>

Purpose:
To predict future prices of stocks or assets using machine learning techniques.

Tools:
ARIMA, XGBoost, LSTMs, Python, scikit-learn, TensorFlow/Keras

What You'll Learn:
Time series forecasting, handling financial data, building predictive models, and evaluating prediction accuracy.

Dataset: https://www.kaggle.com/datasets/akram24/google-stock-price-train

References:
- Youtube Video (Uses Deep learning):
  https://youtu.be/CbTU92pbDKw?si=OZnwcOy1q1S3KREy
- Youtube Video(uses Machine Learning):
  https://youtu.be/1O_BenficgE?si=EzmsJzMwAJ1R5wBv

4. <u>Market Risk Analyzer</u>

Purpose:
To measure financial risk using tools like Value at Risk (VaR), Expected Shortfall, and Stress Testing.

Tools:
Python, NumPy, pandas, historical S&P 500 or crypto data

What You'll Learn:
Parametric & non-parametric risk analysis, tail risk, and real-world risk application in portfolio management.

Dataset: https://www.kaggle.com/datasets/preethamgouda/financial-risk

References:
https://www.youtube.com/watch?v=C2DqhuytS5k
Articles & Blogs:
https://towardsdatascience.com/value-at-risk-var-explained-fa7455d96d53

5. <u>Credit Risk Assessment</u>

Purpose:
To build a model that predicts the likelihood of a borrower defaulting on a loan, helping financial institutions make informed lending decisions.

Tools:
Python, Pandas, scikit-learn, XGBoost, LightGBM, Jupyter Notebook, Matplotlib/Seaborn (for visualization)

What You'll Learn:

- Data cleaning and preprocessing (handling imbalanced data, missing values)
- Feature engineering and selection
- Classification algorithms (e.g., logistic regression, decision trees, XGBoost)
- Evaluation metrics (AUC, precision-recall, confusion matrix)
- Business understanding of credit scoring models

Dataset: https://www.kaggle.com/datasets/laotse/credit-risk-dataset

References: https://www.youtube.com/watch?v=Y7C5ThPONCw

Some Things to Remember:
- These references are just to get ideas of the projects.
- Your submissions for each project should be a jupyter notebook dedicated to it.
- Properly add comments to your codes.
- Assume reasonable values if not given (like free rate = 0)
- You can use extra datasets if your model needs that.
- In the end of each project you should add a comment of conclusions.