10 Academy: Artificial Intelligence Mastery

Week 11 Challenge Document

Date: 26 Feb - 04 Mar 2025

# Time Series Forecasting for Portfolio Management Optimization

Business objective

**Guide Me in Finance (GMF) Investments** is a forward-thinking financial advisory firm that specializes in personalized portfolio management. GMF leverages cutting-edge technology and data-driven insights to provide clients with tailored investment strategies. By integrating advanced time series forecasting models, GMF aims to predict market trends, optimize asset allocation, and enhance portfolio performance. The company’s goal is to help clients achieve their financial objectives by minimizing risks and capitalizing on market opportunities.

At GMF Investments, financial analysts play a crucial role in interpreting complex financial data and providing actionable insights. By utilizing real-time financial data from sources like YFinance, GMF ensures its strategies are based on the latest market conditions, thereby maintaining a competitive edge.

## Situational Overview (Business Need)

As a **Financial Analyst** at GMF Investments, your objective is to apply time series forecasting to historical financial data to enhance portfolio management strategies. Your role involves analyzing data, building predictive models, and recommending portfolio adjustments based on forecasted trends.

You will:

* Utilize **YFinance** data to extract historical financial information such as stock prices, market indices, and other relevant financial metrics.
* Preprocess and analyze this data to identify trends and patterns.
* Develop and evaluate forecasting models to predict future market movements.
* Use the insights gained to recommend changes to client portfolios that aim to optimize returns while managing risks.

## Data

Use historical financial data for three key assets: Tesla (TSLA) Historical stock prices (Open, High, Low, Close), volume, and volatility., Vanguard Total Bond Market ETF (BND), and S&P 500 ETF (SPY). The data will be sourced from **YFinance** and cover the period from **January 1, 2015, to January 31, 2025**.

Fetch the data using the **YFinance** Python library **January 1, 2015, to January 31, 2025.**

Each dataset includes:

* Date: Trading day timestamp.
* Open, High, Low, Close: Daily price metrics, with Adj Close representing the adjusted close price to account for dividends and splits.
* Volume: The total number of shares/units traded each day.

Asset-Specific Descriptions

* TSLA: High-growth, high-risk stock in the consumer discretionary sector (Automobile Manufacturing).
* BND: A bond ETF tracking U.S. investment-grade bonds, providing stability and income.
* SPY: An ETF tracking the S&P 500 Index, offering broad U.S. market exposure.

Usage in Portfolio Analysis

* TSLA provides potential high returns with high volatility.
* BND contributes stability and low risk.
* SPY offers diversified, moderate-risk market exposure.

## 

## Expected Outcomes:

Skills:

* Competence in time series forecasting methods (ARIMA, SARIMA, LSTM).
* Experience in handling and analyzing real-world financial data using YFinance.
* Ability to develop, evaluate, and deploy predictive models.
* Skills in portfolio optimization using forecast insights.

Knowledge:

* In-depth understanding of financial market trends.
* Proficiency in data-driven decision-making for portfolio management.
* Insights into risk management and return optimization strategies.

## Team

Tutors:

* Mahlet
* Rediet
* Kerod
* Elias
* Emtinan
* Rehmet

## Key Dates

* Discussion on the case - Wednesday 26 Feb 2025. Use #all-week11 to pre-ask questions.
* Interim Solution - 20:00 UTC on Friday 28 Feb 2025.
* Final Submission - 20:00 UTC on Tuesday 04 March 2025

# Instructions

## Objectives:

The objective of this challenge is to equip trainees with the skills to preprocess financial data, develop time series forecasting models, analyze market trends, and optimize investment portfolios. Participants will gain hands-on experience in leveraging data-driven insights to enhance portfolio performance, minimize risks, and capitalize on market opportunities.

# Task 1: Preprocess and Explore the Data

Load, clean, and understand the data to prepare it for modeling.

* Extract historical financial data using YFinance for :
  1. TSLA provides potential high returns with high volatility.
  2. BND contributes stability and low risk.
  3. SPY offers diversified, moderate-risk market exposure.
* Data cleaning and Understanding.
  1. Check basic statistics to understand the distribution of the data.
  2. Ensure all columns have appropriate data types and check for missing values.
  3. Handle missing values by either filling, interpolating, or removing them.
  4. Normalize or scale the data if required, especially for machine learning models.
* Conduct Exploratory Data Analysis (EDA):
  1. Visualize the closing price over time to identify trends and patterns.
  2. Calculate and plot the daily percentage change to observe volatility.
  3. Analyze volatility by calculating rolling means and standard deviations to understand short-term trends and fluctuations.
  4. Perform outlier detection to identify significant anomalies.
     1. Analyze days with unusually high or low returns.
* Seasonality and Trends:
  1. Decompose the time series into trend, seasonal, and residual components (e.g., using statsmodels).
* Analyze Volatility
  1. Calculate rolling means and standard deviations to understand short-term trends and volatility.
* Document key insights like overall direction of Tesla’s stock price, Fluctuations in daily returns and their impact, VaR and Sharpe Ratio to assess potential losses and risk-adjusted returns.

## 

## Task 2: Develop Time Series Forecasting Models

This task involves building a time series forecasting model to predict Tesla's future stock prices. Below are the step-by-step instructions to develop, evaluate, and refine a forecasting model using common techniques such as ARIMA, SARIMA, or LSTM.

* You can choose between classical statistical models or deep learning models:
  + **ARIMA (AutoRegressive Integrated Moving Average)**: Good for univariate time series with no seasonality.
  + **SARIMA (Seasonal ARIMA)**: Extends ARIMA by considering seasonality.
  + **LSTM (Long Short-Term Memory)**: A type of recurrent neural network (RNN) well-suited for capturing long-term dependencies in time series data.
* Divide the dataset into training and testing sets to evaluate model performance.
* Train the Model you choose
* Use the model to forecast future stock prices and compare the predictions with the test set.
* **Optimize Model Parameters**:
* Use techniques like grid search or auto\_arima from pmdarima library to find the best (p, d, q) parameters.
* Calculate evaluation metrics to assess the model's performance.
  + **Mean Absolute Error (MAE)**
  + **Root Mean Squared Error (RMSE)**
  + **Mean Absolute Percentage Error (MAPE)**

### Task 3: Forecast Future Market Trends

In this task, you'll use the model developed in Task 2 to forecast Tesla's future stock prices. The goal is to generate future price predictions, analyze the results, and provide insights on potential trends and risks.

* Use the Trained Model for Forecasting
  + Depending on the model you chose (ARIMA, SARIMA, or LSTM), you’ll generate forecasts for 6-12 months.
* Forecast Analysis
  + Visualize the forecast alongside historical data.
  + The forecast should include confidence intervals to show the range within which the future prices are expected to lie.
* **Interpret the Results**

1. **Trend Analysis**:
   * Look for long-term trends (upward, downward, or stable).
   * Identify any patterns or anomalies in the forecast.
2. **Volatility and Risk**:
   * Discuss the level of uncertainty captured by the confidence intervals.
   * Highlight any periods where volatility is expected to increase.
3. **Market Opportunities and Risks**:
   * Based on the forecast, outline potential market opportunities (e.g., expected price increases) and risks (e.g., high volatility or expected declines).

### Task 4: Optimize Portfolio Based on Forecast

In this task, you'll use the forecasted data from Task 3 to make informed decisions about optimizing a sample investment portfolio. The objective is to adjust the portfolio to maximize returns while minimizing risks based on the predicted market trends.

* Use a simple portfolio with three assets:

1. **Tesla Stock (TSLA)** - A growth stock with higher risk.
2. **Vanguard Total Bond Market ETF (BND)** - A bond ETF for stability.
3. **S&P 500 ETF (SPY)** - An index fund for diversification.

* You’ve already forecasted Tesla’s future prices. Now, forecasts for BND and SPY.
* Combine the data into on dataframe df with columns TSLA, BND, and SPY, which contain daily closing prices for each asset.
* Compute the **annual return**, compound the average daily returns for each asset.
* Use **covariance matrix** helps you understand how asset returns move together.
* Define the portfolio weights and compute the weighted average return and risk (volatility).
* Use optimization to find the weights that maximize the Sharpe Ratio.
* Analyze Portfolio Risk and Return
  + Calculate the average of the portfolio returns.
  + Measure the standard deviation of portfolio returns to understand volatility.
  + Measure the potential loss in value of Tesla stock at a given confidence interval (Value at Risk - VaR).
  + Sharpe Ratio: This tells you the risk-adjusted return. Higher is better.
* Optimize the Portfolio:
  + Adjust allocations to maximize returns or minimize risks.
  + Increasing stable assets like BND if you expect higher volatility in Tesla.
* Visualize how the portfolio would perform based on the forecasted returns.
* Summarize the expected return, volatility, Sharpe Ratio, adjustments to asset allocation and reasons, and include cumulative return charts and risk-return analysis.

# Tutorials Schedule

In the following, the color **purple** indicates morning sessions, and **blue** indicates afternoon sessions.

## Wednesday

* Introduction to the challenge (Mahlet)
* Comparing time series modeling (Rediet)

## Thursday

* Time Series Forecasting and Portfolio Optimization (Kerod)
* Backtesting and Simulation for Trading Strategies (Elias)

## Friday

* Integrating robust risk analysis into portfolio management(Emitinan)

# Interim Submission

* Interim report - Covering task - 1
* Link to your GitHub code.

## Feedback

You may not receive detailed comments on your interim submission but will receive a grade.

# Final Submission

* A blog post entry (which you can submit for example to Medium publishing) or a pdf report.
* Link to your Github code, and make sure to screenshots demonstrating anything else you have done.

## Feedback

You will receive comments/feedback in addition to a grade.

# 

# References

**Data Science Workflow**

1. <https://www.datacamp.com/tutorial/arima>
2. <https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/>
3. <https://www.geeksforgeeks.org/time-series-analysis-and-forecasting/>
4. <https://www.datasciencewithmarco.com/blog/the-complete-guide-to-time-series-forecasting-using-sklearn-pandas-and-numpy>
5. <https://www.machinelearningplus.com/time-series/time-series-analysis-python/>

**Portfolio Optimization**

1. <https://miltonfmr.com/the-complete-guide-to-portfolio-optimization-in-r-part1/>
2. <https://www.machinelearningplus.com/machine-learning/portfolio-optimization-python-example/>
3. <https://builtin.com/data-science/portfolio-optimization-python>
4. <https://www.analyticsvidhya.com/blog/2021/04/portfolio-optimization-using-mpt-in-python/>
5. <https://github.com/robertmartin8/PyPortfolioOpt>
6. <https://blog.jetbrains.com/datalore/2024/01/26/portfolio-optimization-in-python-with-datalore-and-ai-assistant/>
7. <https://tradewithpython.com/building-an-optimal-portfolio-with-python>