Project Documentation: Financial Stock Portfolio Risk and Return Analysis

Introduction

This project analyzes the risk and returns of a financial stock portfolio using various data science techniques. The key steps include fetching financial data, performing asset selection using K-means clustering, and optimizing the portfolio for maximum returns while minimizing risk. The analysis is performed on stocks from the NIFTY50 and S&P500 indexes.

Setup and Installation

The project begins with the installation and import of necessary Python libraries:

- numpy for numerical computations.
- pandas for data manipulation.
- yfinance for fetching financial data from Yahoo Finance.
- matplotlib for data visualization.

Data Fetching

The historical stock prices for selected stocks and indexes are fetched using the yfinance library. This includes:

- Individual stock data (e.g., Apple Inc. from 2016 to 2019).
- Multiple stock data (Nifty 50 constituents from 2019 to 2023).
- Index data (NIFTY50 from 2019 to 2023).

Data Visualization

The project visualizes the historical adjusted close prices for the fetched stocks using matplotlib, providing a clear view of the price trends over time.

Data Preprocessing

The fetched data is saved into CSV files for further analysis. The project reads stock price data from these CSV files and processes it to extract relevant asset labels and prices.

Computing Daily Returns

A function is defined to compute daily returns for the stocks. This is done by calculating the percentage change in stock prices from one day to the next. The daily returns provide a basis for further analysis and modeling.

K-means Clustering for Asset Selection

The project employs K-means clustering to identify optimal asset selections for the portfolio. This clustering technique groups stocks into clusters based on their historical returns and volatilities, helping to identify stocks with similar performance characteristics.

Risk-Return Graphs

Risk-return graphs are created for different portfolio types. These graphs plot the risk (standard deviation of returns) against the return (mean of returns) for various portfolios, providing a visual representation of the trade-off between risk and return.

Portfolio Optimization

The project performs portfolio optimization to recommend optimal weightings for assets. This involves:

- Maximizing returns while minimizing risk using mathematical optimization techniques.
- Applying the optimization to both the NIFTY50 and S&P500 indexes and their constituent stocks.

ARIMA Model for Stock Price Forecasting

To forecast future stock prices, the project integrates the ARIMA model:

- **Data Preparation:** Historical stock price data is fetched and split into training and testing sets.
- **Model Fitting:** The ARIMA model is fitted to the training data, and future prices are forecasted.
- **Performance Evaluation:** The forecasted prices are compared with actual prices using evaluation metrics.
- **Visualization:** Actual vs. forecasted prices are plotted to assess model accuracy.
- **Portfolio Integration:** Forecasted prices are used to simulate future returns and adjust the portfolio optimization strategy.

Conclusion

The project successfully analyzes the risk and returns of a financial stock portfolio. By employing techniques such as K-means clustering, optimization, and ARIMA modeling, it provides valuable insights and recommendations for asset selection, portfolio weightings, and future price forecasting. This comprehensive analysis aids in making informed investment decisions that balance the trade-off between risk and return.