

Executive Summary - **FORECASTING STOCK VOLATILITY**

AIMS

The purpose of our project is to forecast volatility and value at risk of assets using predictive modeling.

DATASET

Preprocessed daily closing returns of twenty stocks among the biotech, industrial, tech, and medical industries from January 5, 2010 to December 30, 2022.

Methods

Baseline: The baseline forecast of an asset's volatility is the average of the variances of index movement over a predetermined time window.

Random Forest Regression: An ensemble learning algorithm that fits a regression model by averaging over a large number of decision trees.

GARCH(Generalized Autoregressive Conditional Heteroskedasticity): GARCH is an autoregressive model used to predict an asset's Value at Risk, a common measurement used in measuring a portfolio's overall risk.

LSTM (Long Short-Term Memory): LSTM are recurrent neural network models widely used in time series analysis, are used to predict an asset's volatility, and are effective in capturing long-term dependencies in sequential data.

Data Visualization: Charts give explicit visual performance/comparison of Baseline, Random Forest, GARCH, and LSTM models.

Data Measurements: Statistical measurements are analyzed to draw conclusions about the performance of models.

Conclusion

Random Forest was useful for measuring the early volatility trend. GARCH was especially effective at measuring Value at Risk around extreme fluctuations in volatility. LSTM was able to successfully predict some assets' volatility movements over an extended period of time, e.g. two months.