Brazilian Stock Index Forecasting

Khanh Tran – T00711160 and Roberto Sanches – T00  
*Department of Mathematics and Statistics and Department of Computing Science*  
*Thompson Rivers University*Kamloops, Canada  
[trank22@mytru.ca](mailto:trank22@mytru.ca)

***Abstract***—The.

**Keywords**—Stock Forecasting, Volatility Modeling, Times Series Analysis, Machine Learning

**1. Introduction**

The Brazilian stock market is a dynamic and complex financial environment, represented by the Bovespa Index (BVSP), . Investors and analysts often employ various forecasting techniques to gain insights into future market movements. In this study, we explore the application of different forecasting methods, with a particular focus on mean and volatility forecasting. We compare the results obtained from traditional time series models and advanced models.

**2. Background**

Stock

**3. Data**

This study uses the Brazilian Stock Index, available on the Yahoo Finance website [1]. Historical BVSP data was downloaded from Yahoo Finance by using the "quantmod" library, for the period from January 1, 2017, to December 31, 2022. The adjusted closing prices are extracted and used to calculate daily log returns. The data is then split into training and testing sets. The training dataset covers data from January 1, 2017, up to October 31, 2022, and the remaining is a testing set.

**4. Method**

This methodology section consists of seven main sub-items, a data cleaning and processing item, and six regression model items. All data processing and model building codes were implemented in R and can be found by this [GitHub link](https://github.com/RobertoCurti/DASC6510-01-Project-GroupG/tree/main).

**4.1 Data Processing**

The adjusted closing prices are chosen to account for corporate actions and other adjustments. For volatility forecasting, the data is transformed into daily log returns, a common practice in financial time series analysis, to stabilize variance and make the series approximately stationary.

**4.2 Exploratory Data Analysis (EDA)**

Before model fitting, an EDA is conducted to understand the temporal patterns in BVSP returns. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots are created to identify potential autoregressive and moving average components. These visualizations guide the selection of appropriate models by revealing the time dependencies present in the data.

**4.3 Models**

Regarding volatility forecasting, various models are fitted to the training data using the fable package in R. The models include: Naïve, Mean, Drift, ETS, ARIMA, NNETAR, and Prophet. The Naïve model assumes future values are equal to the last observed value. The mean model assumes future values are equal to the mean of historical values. The Drift model assumes a linear trend in the data. ETS (Error, Trend, Seasonal model is used to captures the error, trend, and seasonality in the data. ARIMA (AutoRegressive Integrated Moving Average is a parametric model combining autoregressive and moving average components. NNETAR (Neural Network Time Series) utilizes a neural network for time series forecasting. The Prophet is a forecasting tool developed by Facebook, capable of handling daily observations with seasonality.

In this study, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is also employed to capture volatility dynamics. The GARCH model is specified with a zero-mean assumption and a squared GARCH term. Parameters are estimated, and the resulting model is used to forecast volatility for the test period.

The forecasted volatility from the GARCH model is compared with the volatility forecasts from the traditional time series models. Visualizations are used to qualitatively assess the consistency and accuracy of volatility predictions. Additionally, accuracy metrics specific to volatility forecasting, such as Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), are computed and compared.

**5. Results & Discussions**

The results from each forecasting model, including both return and volatility predictions, are presented and discussed. Emphasis is placed on the trade-offs between simplicity and accuracy, and insights are drawn regarding the models' ability to capture the inherent complexities of the Brazilian stock market.

**6. Conclusions**

This study

**References**

[1] Yahoo Finance (2023). *BVSP Index* [*Dataset*](https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques/data). Retrieved 2023, December 03 from <https://finance.yahoo.com/quote/%5EBVSP/>

[2]