Combination of ARIMA and GARCH Model for Time Series Analysis and Its Application in Stock Price Prediction

Introduction: Time series analysis has been a major area of research for finding out the meaningful statistics and other characteristics of data to carry out extrapolation of data. Time series analysis finds great applications in the field of finance. A very important use-case is the prediction of stock prices. Various statistical models like regression, exponential smoothing, ARMA, GARCH etc. have been used to model stock prices in the past and they have shown good results.

Project Idea: The project has two parts. The first part is to fit an ARIMA model for carrying out stock price prediction. The second part is taking the residues or errors from the ARIMA model and fitting a GARCH model on the residues to predict the volatility in the returns. Finally combine these two models by subtracting the regions of high volatility (output of GARCH model) from the predictions of ARIMA model to obtain a more robust model for stock forecasting.

My Contribution: I have taken references from two papers (Stock Price Prediction Using the ARIMA Model) which talks about implementing ARIMA model for predicting stock prices and (Forecasting Stock Market Volatility using GARCH Models) which talks about implementing GARCH model for predicting stock market volatility and making buy/sell decisions based on predicted future volatility. My contribution in this project would be to combine these two ideas, that is to use the standard ARIMA model for first making a prediction of stock prices and then using the errors of ARIMA model to implement a GARCH model to predict the volatility in the errors and thus reducing the overall error by combining the two models and make a better forecast.

Dataset: The stock data obtained from National Stock Exchange (NSE) India are used to develop stock forecasting models. The data is the price history and trading volumes of the fifty stocks in the index NIFTY 50 downloaded from Kaggle. All datasets are at a day-level with pricing and trading values split across .csv files for each stock along with a metadata file with some macro-information about the stocks itself. The data spans from 1st January, 2000 to 30th April, 2021.

References:

Paper 1: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7046047

Paper 2:
https://www.researchgate.net/publication/305803327_Forecasting_Stock_Market_Volatility_using_GARCH_Models_Evidence_from_the_Indian_Stock_Market

https://www.sciencedirect.com/science/article/pii/S2210832714000258