Stock Market Analysis: Statistical Indicators And Future Projections

Abstract—From the beginning of time it has been man’s common goal to make his life easier. The prevailing notion in society is that wealth brings comfort and luxury, so it is not surprising that there has been so much work done on ways to predict the markets. Various technical, fundamental, and statistical indicators have been proposed and used with varying results. However, no one technique or combination of techniques has been successful enough to consistently beat the market. Traditional time series forecasting analyses past data and projects estimates of future data values. Basically, this method attempts to model a nonlinear function by a recurrence relation derived from past values. The recurrence relation can then be used to predict new values in the time series, which hopefully will be good approximations of the actual values.

Keywords — Statistical indicators, Stock market, Time series forecasting, Future projections, Recurrence relation, Predict.

I. Introduction

Forecasting is a data science task that is central to many activities within an organization. For instance, large organizations like Facebook must engage in capacity planning to efficiently allocate scarce resources and goal setting in order to measure performance relative to a baseline. Producing high quality forecasts is not an easy problem for either machines or for most analysts. We have observed two main themes in the practice of creating a variety of business forecasts: Completely automatic forecasting techniques can be brittle and they are often too inflexible to incorporate useful assumptions or heuristics. Analysts who can produce high quality forecasts are quite rare because forecasting is a specialized data science skill requiring substantial experience.

Stock market Predictions: Determining the future value of a company stock or other financial instrument traded on an exchange either by fundamental or technical analysis or time series analysis. The successful prediction of a stock's future price could yield significant profit.

Prophet: Fortunately, the Core Data Science team at Facebook recently published a new method called Prophet, which enables data analysts and developers alike to perform forecasting at scale in Python.

Using Prophet for Stock market Predictions: Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data.

II. Overview

2.1 Objective

A mobile app with built-in Artificial Intelligence will project the returns based on the user's selection of types of investments and the risk he/she is willing to take. Historic correlation data and projections are calculated. User will be able to change all the parameters that affect the final outcome and see how interest rate and inflation rate affects the final return. There is no such app which calculates the projections of all these following sectors:

- Fixed Deposit/Recurring Deposit/Public Provident Fund/Sukanya Samriddhi Yojana/Senior Citizens Savings Scheme
- Commodities: Gold /Silver/Crude Oil
- Nifty 50 shares

Returns are projected in the form of a graph. Updates are given for every 15 days according to the current interest rate and inflation rate.

2.3 Methodology And Software Usage

We decided to use Python. Other statistical programming languages such a R provide automated ways to solve this issue, but those have yet to be officially ported over to Python. Pandas library is imported and is used to perform

data analytics with back-end source written in python. Here, the data is a python dictionary with a key, value pair. Prophet is used for time series analysis. The input to prophet is always a dataframe with two columns. The major dependency that Prophet has is pystan. PyStan has its own installation instructions. We need to install pystan with pip before usinf pip to install fbprophet. PyStan provides an interface for statistican modeling and high-performance statistical computation.

Prophet is an additive model with the following components:

y(t) = g(t) + s(t) + h(t) + ??

g(t) models trend, which describes long-term increase or decrease in the data. Prophet incorporates two trend models, a saturating growth model and a piecewise linear model, depending on the type of forecasting problem.

- s(t) models seasonality with Fourier series, which describes how data is affected by seasonal factors such as the time of the year (e.g. more searches for eggnog during the winter holidays)
- h(t) models the effects of holidays or large events that impact business time series (e.g. new product launch, Black Friday, Superbowl, etc.)
- ?? represents an irreducible error term

2.5 Dataset

For stock market analysis, we directly imported the dataset from Zen securities pvt ltd, a privately held company. Historical data of 37 years is collected and imported into a Excel sheet consisting of 2048 records.

III. Implementation

3.1 Pre-Requisites

- Time series analysis can be performed on either a local desktop or a remote server. For papers with less than six authors: To change the default, adjust the template as follows.
- Working with large datasets can be memory intensive, so in either case, the computer will need at least 2GB of memory to perform some of the calculations
- For this, we'll be using Jupyter Notebook to work with the data. If you do not have it already, you install and set up Jupyter Notebook for Python 3.

- Step 1 — Pull Dataset and Install Packages
- Step 2 — Import Packages and Load Data
- Step 3 — Time Series Forecasting with Prophet

3.2 IInstalling Pystan:

Before installing and using fbprophet we need to install PyStan. PyStan and the required packages are installed using pip by the following command.

pip install pystan

3.3 Installing Prophet

To get started with Prophet, we installed it using the following command:

pip install fbprophet

3.4 Loading Csv Data In Python With Pystan

Pandas is a powerful data analysis Python library that is built on top of numpy which is yet another library that let's you create 2d and even 3d arrays of data in Python. The pandas main object is called a dataframe. A dataframe is basically a 2d numpy array with rows and columns, that also has labels for columns and rows.

We create dataframes out of input data format CSV.

fig

ds is the date and y is the price of the a particular sector on that particular day.

Before we can do any analysis with this data, we need to log transform the 'y' variable to a try to convert non-stationary data to stationary. This also converts trends to more linear trends. This isn't always a perfect way to handle time-series data, but it works often enough that it can be tried initially without much worry.

3.5 Modeling

Modeling can be done by:

- Instantiating the prophet to the model.
- Fit the model with the data frame

3.6 Forecasting

With Prophet, some future time data has been built by using make_future_dataframe method with a parameter period. The value of period depends upon the number of days, months or years we predict the values. Forecasting is done using predict command.

IV. Results:

The output is given in the form of a graph. A graph is plotted using matplotlib. It is a plotting library in python. It is used for converting numerical data into a graph. This is imported using the command:

import matplotlib.pyplot as plt

The graph is obtained by using forecast_data command

As we have selected 10 traditional sectors. The output is generated for those sectors in the form of a graph

The seasonality and the trend components of a sector are determined using data_forecast command

Fig: Example output of gold sector

Fig: Example output of PPF sector

V. Advantages

- One doesn’t require much prior knowledge or experience of forecasting time series data since it automatically finds seasonal trends beneath the data and offers a set of ‘easy to understand’ parameters.
- it allows non statisticians to start using it and get reasonably good results that are often equal or sometimes even better than the ones produced by the experts.

VI. Disadvantages

- Organizations should never rely 100 percent on any forecasting models
- It is not possible to accurately forecast the future
- Data may be flawed

VII. Conclusion And Future Work

Since stock market data are highly time-variant and are normally in a nonlinear pattern, predicting the future price of a stock is highly challenging. Prediction provides knowledgeable information regarding the current status of the stock price movement. With some historical data, you can use forecasting tools to predict into the future a specific metric. Prophet is designed for analyzing time series with daily observations that display patterns on different time scales. Thus this can be utilized in decision making for customers in finalizing whether to buy or sell the particular shares of a given stock. Many researchers have been carried out for predicting stock market price using various data mining techniques. The past data of the selected stock will be used for building and training the models. The results from the model will be used for comparison with the real data to ascertain the accuracy of the model.

For future work, using traditional time series analysis to forecast stock market prices will be a continuous area of research as researchers and investors strive to outperform the market with the ultimate goal of bettering their returns. It is unlikely that new theoretical ideas will come out of this applied work.

References

1. Idrees, Sheikh Mohammad, M. Afshar Alam, and Parul Agarwal. 'A Prediction Approach for Stock Market Volatility Based on Time Series Data.' IEEE Access 7 (2019): 17287-17298.
2. Picasso, Andrea, et al. 'Technical Analysis and Sentiment Embeddings for Market Trend Prediction.' Expert Systems with Applications (2019).
3. Somarajan, Siddarth, et al. 'Modelling and Analysis of Volatility in Time Series Data.' Soft Computing and Signal Processing. Springer, Singapore, 2019. 609- 618.
4. Nyoni, Thabani. 'Time series modeling and forecasting of the consumer price index in Belgium.' (2019).
5. Abraham, Cerene Mariam, M. Sudheep Elayidom, and T. Santhanakrishnan. 'Analysis and Design of an Efficient Temporal Data Mining Model for the Indian Stock Market.' Emerging Technologies in Data Mining and Information Security. Springer, Singapore, 2019. 615-628.
6. Kannan, K. Senthamarai, et al. 'Financial stock market forecast using data mining techniques.' Proceedings of the International Multiconference of Engineers and computer scientists. Vol. 1. 2010.
7. Khedr, Ayman E., and Nagwa Yaseen. 'Predicting stock market behavior using data mining technique and news sentiment analysis.' International Journal of Intelligent Systems and Applications 9.7 (2017): 22.
8. Shen, Shunrong, Haomiao Jiang, and Tongda Zhang. 'Stock market forecasting using machine learning algorithms.' Department of Electrical Engineering, Stanford University, Stanford, CA (2012): 1-5.
9. Choudhry, Rohit, and Kumkum Garg. 'A hybrid machine learning system for stock market forecasting.' World Academy of Science, Engineering and Technology 39.3 (2008): 315-318.