Stock Price Prediction Application Report

Yuting Zhou, Jimin Chen

The main goal of this application is to provide an opportunity for anyone to construct a regression model with selected variables. By feeding the data processing mechanism with files in acceptable formats, our software allows the user to create their regression model based on selected variables to predict the underlying stock price. In the code, we use Alphabet Inc. (Google) as an example. As shown in the dashboard, our software automatically generates two scatter plots and associated R square scores first to illustrate how the user-generated model fits into the situation. Then, by making use of functions in the sidebar, the algorithm recommends a stock price to the user.

The data in the example directly come from Google's quarterly reports. Those can be easily retrieved by anyone and have a sufficient number of data for us to construct the model. We choose that specific data set with two intentions: firstly, many data from the financial reports, such as dividends and cash flows, serve as determinant factors of the stock price in many traditional models and investors do follow models in reality. Those determinant data perform as the "checking points" of our software: if our model predicts well through using those data, then that indicates good effectiveness of our model. Additionally, there are also data that are not traditionally considered determinant by many models. Those data grant the model potential to detect unrevealed but actually detrimental factors or a combination of factors.

In order to make those data available to be further processed, we first need to gather the financial data from all the quarterly reports. That process can be done manually, but we

download the data from Yahoo finance, where the financial data have been put together. By merging the balance sheet, income statement, and shareholder's equity together, a database is constructed. However, since Exchange does not operate during the weekend, stock prices may not be available at the date financial data report on. We believe that the disparity in date may result in a butterfly effect on the effectiveness of the model, thus those financial dates without a corresponding available stock price are dropped from our analysis.

In the dashboard, users are allowed to choose from all available variables. Once selected, variables are fed to the regression model, which will be trained by the corresponding historical data. Two scatter charts are drawn to give an intuitive idea of how the model performs on the data the model was given and the data the model was not given. Observing the charts, any user, even without any statistical background, is able to determine the effectiveness of the model shown by the difference between the predicted value and the actual value. Two R square scores are also generated for advanced users to determine the performance of the model on the given and ungiven data sets. The user then is allowed to enter an expected growth rate for each variable selected. Based on the growth rate and the last released report, the model will generate a recommended stock price for the user to make investment decisions.

In conclusion, this application can predict Google's stock price, or any price if the application is fed with correctly formatted variables, using personalized linear regression models. Users can explore and find the relationships between the stock price and quarterly report data. It's helpful for the people who need a tool for Google stock price prediction when they have information about the expected change of this company's financial data.

Get access to the application:

Dashboard

Deepnote

Github