

# Forecasting Stock Returns

Barbara Aryeley Aryee  
MSc in Financial Technology

## Introduction

The forecasting of stock returns is an important research subject for both academics in finance and machine learning experts, despite the challenges involved in getting data as well as the nature of this data which is very noisy and time varying, this due to the opportunities it possesses. The prediction of future return of stocks is a known difficult tasks Despite this difficulty interest in this research area is growing rapidly especially in the use of machine learning to forecast stock returns. Previous research have shown that there is a link between analyst recommendations and stock returns.

Aim- This project aimed to look at the ability to forecast stock returns using machine learning techniques from various forecasting metrics by equity research analyst. The focus is on metrics such as forecasted earnings, revenue and recommendations over various time periods. This paper provides a link between the literature on analyst recommendations and the predictability of



## Data

The data used for this study was obtained from Thomson Reuters Eikon and Datastream. The data contained all the stocks in the S&p500 stocks from the year 2012 to 2017. In order to perform the analysis two sets of data were collected. One set of data was collected which described the stock price and the other set which described the analyst recommendations.

For the recommendation dataset, the details obtained include the stock code, the recommendation mean, the number of analyst, the number of recommendations, the date of recommendation and date of stock transaction. The recommendation data collected were on a weekly basis. And for the data on the stock price, the stock code, the stock name, the price, the unadjusted price and the total return index were collected.

## Approach

The models are built using the Python programming language with the ScikitLearn open source machine learning library.

Machine learning is a very powerful tool when it comes to prediction. In building the machine learning model, the data was split into three components, the training data, validation data and testing data. Based on previous research, three techniques were selected to build the predictive model. These are

- Random Forest
- Multilayer Perceptron
- Linear Regression

