Implementation and Analysis of a Machine Learning Approach to long-term Value Investing

Minimize Risk while Maximizing Cash Flow through Stock Picking Based on Fundamental Company Data

Bachelor Exposé

Submitted at the

IMC Fachhochschule Krems

(University of Applied Sciences)

Bachelor Programme Informatics

by

Daniel Netzl

for the award of academic degree
Bachelor of Science in Engineering (BSc)

under the supervision of
Dr. Rubén Ruiz Torrubiano

Declaration of honour

I declare on my word of honour that I have written this Bachelor Exposé on my own and that I have not used any sources or resources other than stated and that I have marked those passages and/or ideas that were either verbally or textually extracted from sources. This also applies to drawings, sketches, graphic representations as well as to sources from the internet. The Bachelor Exposé has not been submitted in this or similar form for assessment at any other domestic or foreign post-secondary educational institution and has not been published elsewhere. The present Bachelor Exposé complies with the version submitted electronically.

Daniel Netzl
15.06.2022

Table of Contents

List of Figures

1

Introduction

1.1   Motivation

When it comes to investing in the financial market, there are numerous approaches and tactics to consider. Investors strive for long-term success, but they frequently fail for a variety of reasons. Younger investors, in particular, have a tendency to underestimate the risk and end up accepting a significant financial loss. That has been more common in recent years, with a significant portion of generation Z following self-proclaimed "gurus" and their "expert" opinions and recommendations on numerous internet platforms. Following those incoherent investing schemes is tantamount to speculation or outright gambling. The promise of quick and cheap returns attracts investors, who fall prey to Wall Street's countless fads. In this paper, the author has two objectives. The first half of the paper describes the common mistakes made by investors and the challenges they confront, particularly in the present era of easy access to financial markets. In the second half of the paper the author provides his approach to asset management, as well as an algorithm that may aid him in executing it, in the hopes of assisting investors in recognizing and so avoiding these losing methods [1].

1.2   Problem Definition

It's terrifying to see how many naïve and ingenious investors have had horrible financial outcomes. If this paper and its algorithm succeed in their approach, the author will be overjoyed if he can persuade even a few of the readers to avoid risky investment selections in favor of sensible ones that will safeguard and keep their

hard-earned cash. Investors are frequently their own worst adversaries. On the one hand, when price trends are rising, investors are more likely to speculate and follow their emotional greed, placing high-risk bets based on optimistic expectations and ignoring related danger. When prices are declining, on the other hand, emotions again play a huge role. Fear of loss causes the investors to concentrate solely on the prices continuing to fall, rather than on the underlying data of the companies. Regardless of the current market scenario, many people are looking for a winning recipe. Reality, however, does not follow any mathematical equations.

Younger investors, in particular, are more likely to acquire their financial advice from dubious sources, such as influencers who claim to have had amazing success on Wall Street and know exactly what they are doing. Due to the ease of access to financial markets and the availability of super-cheap transactions provided by online brokers, a significant portion of Generation Z is perceived to be significantly involved in extremely speculative high-frequency trades. This effect has been particularly noticeable in recent years, when market prices have only shown one direction. The S&P 500, for example, climbed by over 98 percent between May 2017 and January 2022. That's nearly a 20 percent annual increase. The NASDAQ 100 hit its interim peak around the same time, gaining roughly 190 percent, or 38 percent annually, in the same time frame. It goes without saying that many new investors were enticed by the supposedly easy and extraordinary gains. However, as this paper is written, those new investors are experiencing their first baisse, revealing their expertise to be nothing more than riding a wave together with the rest of the market. It is critical to understand what one is doing and to have a clear approach during such times.

The strength of such speculative investors should not be underestimated. As can be observed in the case of the Gamestop stock, a downward-pointing company's stock price has risen by over a thousand percent in half a year, only to plummet by half immediately after (but still remain at a high level). There have been multiple instances where private investors have banded together on social media, particularly on the website reddit.com, to artificially inflate prices to unheard-of highs, enticing a large number of naive investors and leaving the vast majority of them with irreversible losses. Many individual and institutional investors overlook or deliberately disregard core corporate principles, perceiving stocks as nothing more than pieces of paper to be traded back and forth.

Investors must ultimately choose their preferred methods. Either they take a seemingly simple way that provides the comfort of consensus, or they take a path

that involves emotional responses fueled by greed and fear and guided by short-term thinking [1].

Most people are unwilling to make the commitment required by the alternative. Those methods, which include value investing, involve fundamental analysis, which treats equities as fractional ownership of the underlying company they represent [1].

It is critical to distinguish between speculation and investing. Anyone who buys and sells stocks nowadays is referred to be an investor. Nonetheless, the vast majority makes no attempt to justify their investing decision. Most of the time, no evaluations are performed, and stocks are bought and sold when markets rise and fall. The recent trend of the stock price is frequently used as a buying criterion. If the stock outperformed the market, it gets purchased. If any analysis is conducted, they frequently include a review of long-term past growth that is expected to continue. Also, "investors" may select companies that have not yet produced spectacular outcomes but are expected to do so in the future. Growth stocks and assets from the technological or health-care sectors are common in these companies. "Investors" hope to benefit from enormous future results [2].

1.3 Research Objectives and Questions

The purpose of this bachelor's thesis is to create an algorithm that aids in the making of sound financial decisions. When a selection outperforms its benchmark, the Vanguard FTSE All-World High Dividend Yield Index, in the long run, it is regarded good. During the implementation phase, the author will look for ways to automate the above-mentioned value investing investment technique. The algorithm's foundation will be fundamental company data and its result will be one of three signals: "Buy," "Hold," or "Sell." Please keep in mind that the algorithm only sends out signals based on the data it receives and the machine learning model that was trained on that data. The final product is not a professional investment advice. The thesis will be regarded successful if the algorithm can consistently exceed its benchmark. The author will see the Efficient Market Theorem [3] refuted in this scenario.

If the target is not accomplished, the author will adhere to the passive investment technique and invest in index funds on a regular basis. In this sense, the author accepts average returns and is unable to disprove the Efficient Market Theorem.

On the road to developing the algorithm the author will examine appropriate machine learning approaches for creating the basis for value investing. Additionally, important features needed for the prediction of the intrinsic value of a company will be determined. This way the author hopes to uncover undervalued companies whose stock prices will increase to a higher extent than the benchmark index.

The machine learning model will utilize backpropagation on historical data to update weights on various features in order to determine the importance of each for predicting the intrinsic value.

It is critical that the model works well over a long period of time, i.e., continuously throughout several years. Short-term success is typically based on luck and cannot be replicated. Technical analysis for speculative short-term stock movement predictions will not be covered in the thesis. The model's results will be assessed on a yearly basis. The above-mentioned problem statement and goals allow the formulation of the following research questions:

- RQ1: Can a machine learning model based on qualitative and quantitative fundamental company data reliably and accurately predict the intrinsic value of a company?

During the implementation, a list of predetermined stocks will be tracked in order to answer the first research question. Each company's qualitative and quantitative data are collected in the hopes of identifying aspects that have a significant impact on the stocks intrinsic worth. A machine learning model is trained for this process. After taking into account a margin of safety, the model compares the stock's suggested price to the current price and issue a "Buy," "Hold," or "Sell" signal.

- RQ2: Can the identified undervalued stocks be used to consistently beat the market and thus disprove the efficient market theorem?

The second phase of the empirical study seeks to beat the market using the model's output. To do so, the identified undervalued stocks, i.e., stocks that emit a "Buy" signal, are purchased at a specific moment in the past. The stock picks are evaluated, and their returns are compared to the benchmark index after each of the next ten years.

From those research questions the following hypotheses can be derived:

- H0: A machine learning model based on qualitative and quantitative fundamental company data cannot reliably and accurately predict the intrinsic value of a company.

- H0: The identified undervalued stocks cannot be used to consistently beat the market and thus the author cannot disprove the efficient market theorem.

- H1: A machine learning model based on qualitative and quantitative fundamental company data can reliably and accurately predict the intrinsic value of a company.

- H1: The identified undervalued stocks can be used to consistently beat the market and thus the efficient market theorem can be disproved.

The following risks are anticipated throughout the thesis:

- What qualitative measures should be used to calculate intrinsic value? And how can they be retrieved in bulk?

- How can a machine learning technology be used to uncover patterns of undervalued stocks?

- How to appropriately assess the model's success within a given time frame and update the model accordingly?

2

Related Work

According to the author's research, very little has been done in the field of study. The majority of present research is based on high-frequency trading and short-term price fluctuations. Only a few strategies concentrate on long-term value investment. The difficulty in determining and translating qualitative company evaluation metrics for machines appears to be one of the reasons for the scarcity of related work.

2.1 State of the Art

Nonetheless, the author wishes to draw attention to the work done by Amel-Zadeh et al. [4]. This paper explored the capabilities of various machine learning algorithms to forecast abnormal stock returns around earnings releases solely based on financial statement data. Similar to this study, Amel-Zadeh et al. [4] focused their study on a neural network architecture and a Random Forest regression and benchmarked it to the OLS linear regression and Lasso regression. As can be seen in the results, the non-linear methods were able to correctly anticipate the direction of various market reactions to earnings announcements. The models created in that study will be used as reference for this paper [4].

It is vital to emphasize that the goal of this study is not only to estimate the magnitude of immediate price reactions to earnings announcements, but also to determine a company's fair worth. As a result, the current study differs significantly from the previous one. In addition, qualitative data sources will be searched in the hopes of uncovering more important features that can be used to evaluate a stock's intrinsic value.

3

Background

The author will advocate one specific investing technique for the remainder of the paper: the value-investment philosophy. This idea encapsulates the technique of investing in assets that trade at a significant discount to their intrinsic value. The strategy has been used for a long time, with investors experiencing minimal risk and good rewards [1]. To achieve investment success, it is of utmost importance to know where others go wrong and deliberately choose a path to avoid those pitfalls. The thesis will mainly be built upon the most honorable representatives of value investing, including Benjamin Graham, David Dodd, and Seth Klarman.

Security Analysis [5], written by Benjamin Graham and David Dodd more than fifty years ago, is widely considered as the bible of value investing. For generations of value investors, that single work has paved the road. Graham's most recent book, The Intelligent Investor [2], is a less scholarly account of the value-investing process. Warren Buffett, the chairman of Berkshire Hathaway, Inc., and a Graham student, is widely recognized as the most successful value investor today [1] Seth Klarman published the most recent book this thesis' methods are based on. With Margin of Safety [1] Klarman emphasizes the necessity of avoiding typical blunders. By describing his approach to value investing, he demonstrates that success in the financial markets requires a defined strategy backed by patience, ambition, and hard effort.

Investors face two distinct dangers in their search for the most promising stocks. They could be wrong about the company's future progress. And even if they are correct, the present market price may already reflect the anticipated development. Insofar as they are predictable, a company's near-future results are often already taken into account. By making a judgment based on those criteria, one is likely to discover that others have already done so. To summarize, in order to obtain

above-average results, one must adhere to policies that are essentially sound and promising, even if they are unpopular on Wall Street [2]. Value investing aims to identify stocks that have been overlooked and are consequently undervalued. However, it is not so straightforward, since the process requires a lot of patience. Selling an overrated and overly popular issue takes boldness and endurance. The theory is sound, and while successful application is not impossible, mastering it is a difficult art [2]. Even more so nowadays when stock prices are adopted in a fraction of a second.

Even yet, the concept of value investing is unlikely to turn anyone into a profitable value investor. Hard work and tight discipline are required for value investment. Only a small percentage of people are willing and able to devote the necessary time and effort, and only a small percentage of people have the right mindset to be successful in the long run. Because those virtues are becoming increasingly rare as the modern environment becomes more dynamic, the algorithm under investigation tries to aid in making value investing accessible to a wider audience.

Naturally, this paper will not present a foolproof investment method. It will not guarantee any profits in advance, but it will highlight the personal risk that everyone must analyze before making investing decisions. The presented theory, as well as the algorithm developed, do not offer any financial advice or suggestions. The algorithm's signals are nothing more than the results of various calculations that, according to the inventor, might be utilized to aid in the discovery of undervalued companies. It is entirely up to the readers to decide how they will use the material.

4

Methodology

The author's research methodologies are briefly detailed in the next section. These include the data gathering process, the programming language(s) utilized, an implementation plan, a description of the trained machine learning models employed, and an evaluation strategy.

4.1 Data

One of the most crucial questions is for which companies the model makes predictions. Because the value investing method only considers stable companies with a good business model, consistent returns, and low volatility, the focus will be on about 300 companies that have been in the market for a long time and have a history of dividend payouts. Being able to pay dividends on a consistent basis, and even increasing payments in most years, is seen as additional evidence of stability. Still, one should not be fooled by a company's payout ratio, as a high payout and continual dividend increase could be a ruse to appease shareholders.

In a Microsoft Excel file, all the companies are listed with their ticker symbol. This file will be used to access extra data from various online APIs. YahooFinance and Finviz.com are two of the sites that will be used to retrieve financial statement data. Additional web services will be researched to discover if any trustworthy qualitative data sources exist. Automatically retrieving bulk qualitative company data might be a challenge, though.

The author purposefully includes a huge number of variables from the balance sheet, income statement, and cash flow statement in the prediction models without applying any economic logic to the selection. This allows the models to learn from the data over the whole available time frame which factors and variable com-

binations perform best for predicting the intrinsic value around earnings announcements over a one-year period.

4.2 Implementation Plan

Python and R will be the primary programming languages used throughout the pipeline. While R will be utilized for the first two phases of the project, namely data collection and processing, Python will be used mostly for the modeling and assessment processes.

The data gathering phase is the initial step in the algorithm's implementation. The author will endeavor to collect as much financial data as possible during this time. The APIs YahooFinance and Finviz.com are the most common data providers. Data for the balance sheet, income statement, and cash flow statement will be retrieved and saved in a data frame for each firm and period. Sources for qualitative corporate data will also be sought and researched. If a reliable source is discovered, it will be included to the data collection. One of the most important jobs at that phase will be to locate such sources.

Data is prepared for further calculations and modeling during the processing phase. In the modeling phase the below mentioned models are applied and tuned to generalize well. As one can observe, the implementation plan heavily correlates with the famous Cross Industry Standard Process for Data Mining (CRISP-DM) [6].
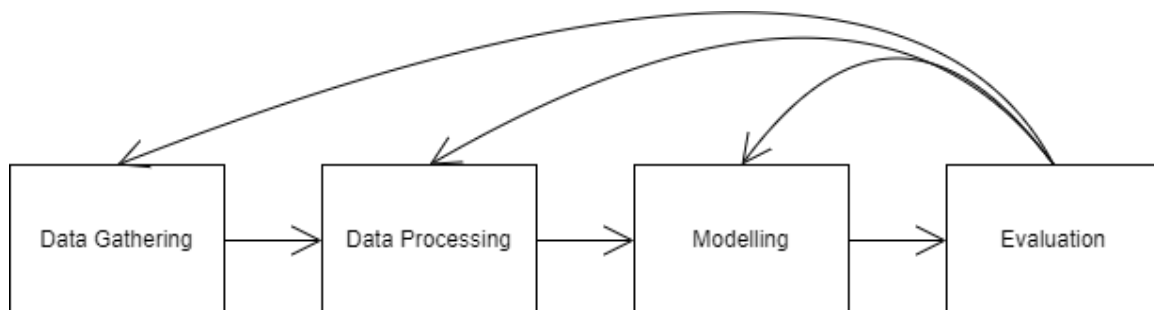

Figure 4.1: Implementation Plan

4.2.1 Machine Learning Models

The machine learning algorithms used in the study include a variety of models that have become increasingly popular in recent years. A Neural Network and a Ran-

dom Forest model for regression are among them. Both Neural Networks [7] [8] and Random Forest [9] have heavily been studied in the field of finance for several years now. To benchmark these two models, simpler linear models are used. These include an Ordinary Least Square (OLS) and a Least Absolute Shrinkage and Selection Operator (Lasso) linear regression. As all the models require some dependent variable y, the author will make use of two different approaches.

First, static equations will be employed to determine a company's fair worth using multiple ways. Following that, the dependent variable y will be the average of all the outputs. Those equations include the Discounted Cashflow (DCF) technique, the Graham Stock Valuation Formula (short: Graham Formula), and the Earnings Power Value formula (EPV).

The other approach will involve a self-created program which will label the data entries automatically based on the development of the stock prize a year after. If the stock rose higher than the benchmark, the label will be "Buy". If the price rose significantly less, whereas different thresholds will be tried out, or even decreased over the period of one year, the label will be "Sell". Everything in between will be labeled as "Hold". These labels will be used to validate the output of the machine learning algorithm.

Before introducing the machine learning models applied during the study, a short introduction on the methods to estimate the fair value of an issue will be provided.

Discounted Cash Flow

The Discounted Cash Flow (DCF) valuation is one of the most popular formulae among value investors, as it is considered one of the best methods for estimating the intrinsic value. One significant advantage of the DCF valuation over earnings-based values is that earnings can be readily manipulated to boost reputation or reduce tax burden, whereas cash flows are just the leftovers that can be used for a variety of purposes. The DCF calculates the investment's value using future cash flows, or the amount of money the investment will earn in the future. Future cash flows are forecasted and discounted to the present. The DCF yields a value that should be lower than the asset's current market value in order for it to be considered a prospective investment [10]. One disadvantage of the DCF is that it is based on a variety of assumptions, including the cash flow growth rate, discount rate, and terminal rate, all of which will be utilized to establish the DCF's final value

[10]. Looking at past cash flows and continuing the trend into the future is one technique to assess the growth rate. However, keep in mind that no "ideal" rate exists because no one can predict the future [10]. The discount rate is calculated using the weighted average cost of capital (WACC). The WACC is in its essence nothing more then the weighted average of a company's cost of debt and cost of equity [10]. The implementation of DCF and hence also WACC will be provided in a later chapter, as there might be the need to so adjustments to fit the purpose of this paper.

Benjamin Graham Number

The Graham Number, like any other stock valuation technique, will include assumptions, resulting in a range of possible values. The following is the original Graham Formula from Security Analysis [5]:

$$V = EPS * (8.5 + 2g)$$

where V is the intrinsic value, EPS the trailing twelve-month EPS, 8.5 the Price-to-Earnings (PE) ratio of a no-growth investment and g being the growth rate for the next seven to ten years. Even yet, the intrinsic worth should not be determined only on a single twelve-month period, which is why the author uses Old School Value's [11] technique. The issue with the Graham Number is that it is heavily reliant on the PE, which can be anything one considers to be correct. Nonetheless, according to Old School Value [11], any value between 7 and 8.5 is a reasonable fit for no-growth businesses. Alternatively, the analyst's five-year growth rate projections might be used [11]. In addition, the $2g$ multiplier will be decreased to 1 in order to reduce the growth rate's aggressiveness and produce results that are more in line with the present market situation [11]. The function that is built to implement the Graham Number computation will be given later.

Earnings Power Value

Three factors must be considered when calculating the Earnings Power Value (EPV):

1. The worth of assets that a rival must have in order to acquire the same market value as the firm being studied.

2. The value of earnings power is calculated using current financial data.

3. Whether or not growth is important, as it is frequently overlooked.

The EPV necessitates making changes to the balance sheet and income statement. The result of the EPV calculations is the amount a competitor must spend to compete with the company under investigation. Although it is critical to compare the EPV with other valuation approaches, it is still a useful tool to have in any investor's toolbox [12].

All the above mentioned metrics will be averaged and decreased by the margin of safety, which equals at least the average yearly return of the benchmark index, to get an estimation of the intrinsic value based on above-mentioned static calculations.

During the study the following machine learning techniques will be employed.

Neural Network

A Neural Network is nothing more than a collection of mathematical equations linked together. Input data in form of a vector traverse the network of equations to generate a vector of outputs [13].
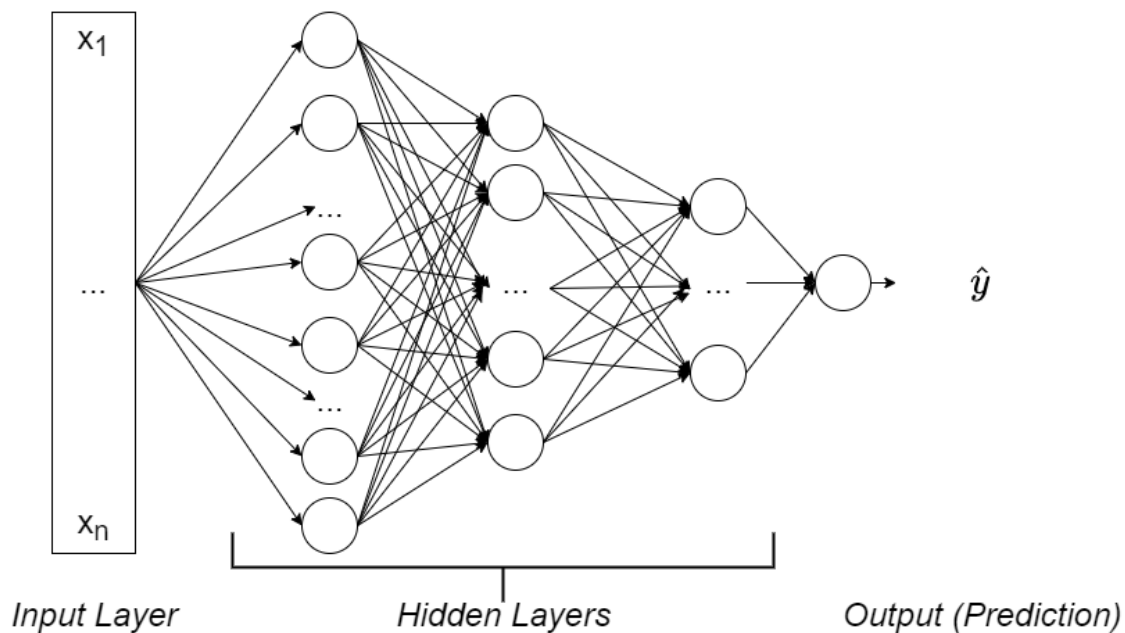

Figure 4.2: Neural Network Architecture (created by the author)

The input layer, one or more hidden layers, and an output layer are the three layers that make up the structure. The independent variables make up the input layer. Those variables are denoted as x_1 to x_n in the graph below, and they comprise of each company's quantitative and qualitative data. As a result, n denotes the number of companies that are being investigated. Please keep in mind that, despite the fact that the data set currently contains 408 companies, a lack of data may cause several of them to be removed again [13].

The hidden layers are made up of one or more hidden nodes or neurons. Both a linear and an activation function can be found in these neurons. The linear function is simply a line of best fit, but the activation function controls whether or not a single neuron's linear function is used. As a result, each node determines which nodes in the adjacent layer are activated until it reaches an output. In terms of concept, that is the essence of a neural network [13].
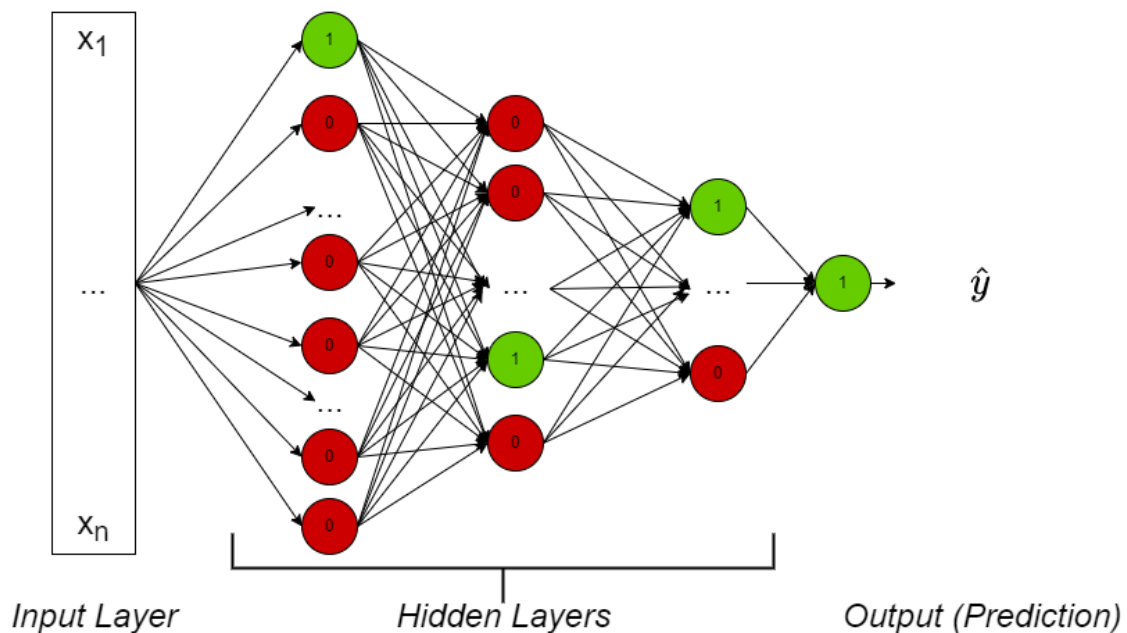

Figure 4.3: Neural Network Architecture with Activation (created by the author)

The activation of neurons can be seen in the figure 1.2. When a neuron turns green, it signifies it is active and allows data to pass through it.

In this thesis, the author will first construct a simple neural network architecture before considering the implementation of more complex structures such as Recurrent Neural Networks (RNN) or sequence architectures such as Long-Short-

Term-Memory (LSTM) in the event that the basic structure fails to comprehend the data.

Several hyperparameters will be tested during the implementation process, which means that before the network starts training, alternative activation functions and different numbers of layers and neurons will be defined. This will ensure that the best neural network design is utilized to learn from the data.

Random Forest

In essence, the Random Forest algorithm is a collection of random decision trees. Each decision tree is random since it is constructed using a random sample of the original data (with replacement) and a subset of features is randomly selected at each tree node to yield the optimum split [14].

Bootstrap aggregation, or bagging, is the term for this method of fitting a decision tree on different bootstrap samples of the training data set. Bagging is useful for a wide range of issues, and it has led to the development of powerful approaches like Random Forest. On each of those samples, or bags, a decision tree will be trained, and the predictions will be averaged in the end. Because each decision tree's training set is a little different, the prediction and prediction error are less correlated, which means that ensemble techniques like Random Forest generalize better than individual decision trees. Furthermore, the bagging technique does not tend to overfit the training data set and can be scaled until performance plateaus [15].

Note that Random Forest can be used for regression as well as for classification problems. In this paper, it will be used for estimating the fair value of a stock given several input features. The output will thus be numerical, meaning that the model will be used for regression purposes.

Ordinary Least Squares

Ordinary Least Squares (OLS) refers to a Simple Linear Regression model with the formula

$$y_i = \alpha + \beta * x_i + \epsilon_i$$

where ϵ_i describes the error term, and alpha and beta are the parameters of the regression. The latter describes the variation of the dependent variables when the independent variables have a unitary variation. The parameter alpha, on the other
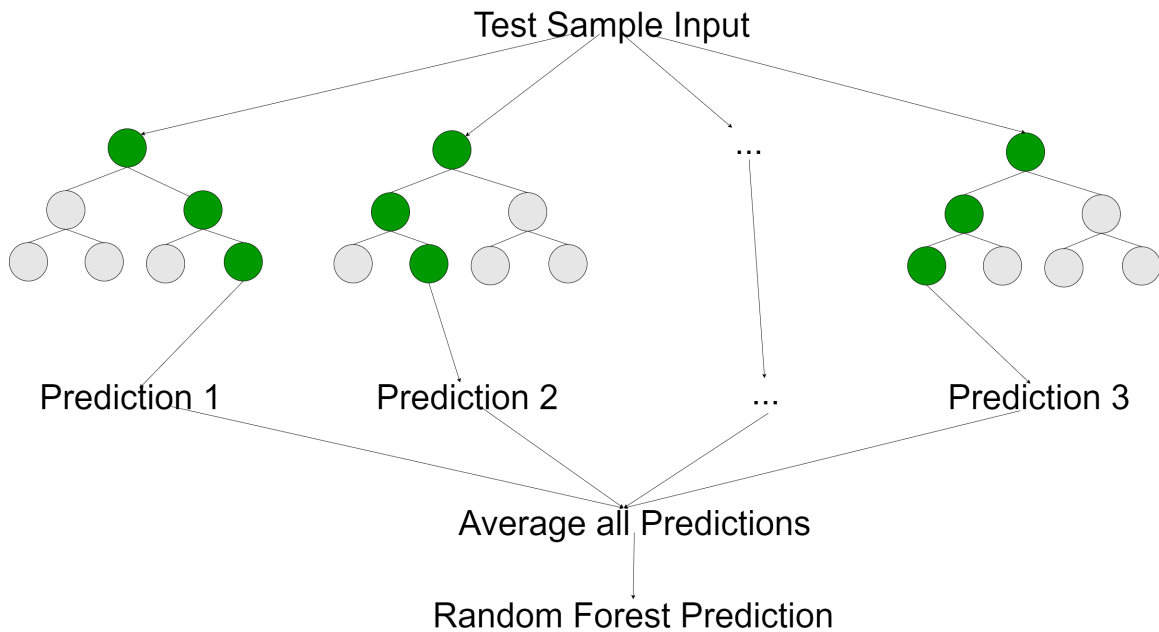
Figure 4.4: Random Forest is based on several decision trees fitted on several random samples (created by the author).

hand, represents the value of the dependent variable, when the independent one is equal to zero [16].

The goal is to determine the alpha and beta parameters that minimize the error term. Because negative and positive errors punish the model in the same way, OLS refers to the error terms being squared [16].

Lasso Regression

The Least Absolute Shrinkage and Selection Operator (Lasso) is a regularization approach that prevents regression models from overfitting. Regularization is the process of punishing the best model in order to acquire one that is more generalizable. It also uses shrinkage, which is the process of compressing data values towards the mean, to limit the influence of independent variables. For extremely multicollinear features, the lasso approach is an excellent fit. It also performs feature selection, making it an excellent fit if the data set contains a large number of features [17].

4.2.2   Training and Test Set

The training set will contain the quantitative and qualitative fundamental data of each company at the end of each year. This way the time dimension is deliberately left out as (1) that simplifies the models and (2) there is nothing worth remembering from the past. The models analyze current data only and try to estimate a stock's fair price. For training purposes 90 percent of the available data will be used, while the remaining 10 percent are used to test the models.

4.3   Evaluation Plan

A one-year sliding window is used to create the training and test sample for our models. This pane navigates through all the companies in our sample period's records history. Let P_t be the price of an issue at any point in time when the algorithm issues a "Buy" signal. Then P_{t+1} represents the price after one year of holding the paper, emulating the purchase of a stock. The return of the stock is denoted by

$$R_{s,t} = (P_{t+1} - P_t)/P_t$$

The result is in continuation compared to the return $R_{i,t}$ of the benchmark index in that exact same time frame. If $R_{s,t_1} > R_{i,t_1}$ the prediction is considered successful for t_1. Returns are evaluated until t_n, where n signifies the last possible complete, i.e., the most recent time frame. The sliding window continues for every year thereafter and after each year, and in total after t_n, the returns are compared. Once the stock's signal switches to "Sell," the algorithm will cease predicting and calculate $R_{s,t}$ and $R_{i,t}$ to complete the evaluation early.

5

Conclusion

The research is being carried out in the hopes of establishing an automated method of investing defensively and accurately identifying inexpensive equities. If successful, this research might be a breakthrough in the field of automated value investing. As very few strategies appear to be successful when applied to financial market predictions the job to be done is regarded difficult to complete. Nonetheless, the author is hopeful that his methods will produce acceptable outcomes and will be more than delighted with the study's conclusion if it attracts attention and serves as a foundation for future research in the field.

Bibliography

[1] S. Klarman, *Margin of Safety: Risk-averse Value Investing Strategies for the Thoughtful Investor*. Harper Collins, 1991.

[2] B. Graham, *The Intelligent Investor*. Harper & Brothers, 1973.

[3] L. Downey, "Efficient market hypothesis (emh)." Investopedia, 2021. [Online]. Available: https://www.investopedia.com/terms/e/efficientmarkethypothesis.asp

[4] A. Amel-Zadeh, J.-P. Calliess, D. Kaiser, and S. Roberts, "Machine learning-based financial statement analysis." University of Oxford, 2020. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3520684

[5] B. Graham and D. Dodd, *Security Analysis*. McGraw-Hill Education - Europe, 1940.

[6] D. S. P. Alliance, "Efficient market hypothesis (emh)." Data Science Process Alliance, 2022. [Online]. Available: https://www.datascience-pm.com/crisp-dm-2

[7] B. K. Wong and Y. Selvi, "Neural network applications in finance: A review and analysis of literature (1990–1996)," 1998, pp. 129–139. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378720698000500

[8] H. F. Calvo-Pardo, T. Mancini, and J. Olmo, "Neural network models for empirical finance." Multidisciplinary Digital Publishing Institute, 2020, p. 265. [Online]. Available: https://www.mdpi.com/1911-8074/13/11/265

[9] L. Kheidem, S. Saha, and S. R. Dey, "Predicting the direction of stock market prices using random forest," 2016. [Online]. Available: http://arxiv.org/abs/1605.00003

[10] D. Ahern, "Explaining the dcf valuation model with a simple example," 2020. [Online]. Available: https://einvestingforbeginners.com/dcf-valuation

[11] O. S. Value, "Benjamin graham formula & stock valuation," 2022. [Online]. Available: https://www.oldschoolvalue.com/stock-valuation/benjamin-graham-formula

[12] ——, "Earnings power value (epv) stock valuation how-to," 2019. [Online]. Available: https://www.oldschoolvalue.com/stock-analysis/earnings-power-value-epv

[13] T. Shin, "A beginner-friendly explanation of how neural networks work," 2022. [Online]. Available: https://towardsdatascience.com/a-beginner-friendly-explanation-of-how-neural-networks-work-55064db60df4

[14] H. Deng, "An introduction to random forest," 2018. [Online]. Available: https://towardsdatascience.com/random-forest-3a55c3aca46d

[15] J. Brownlee, "Essence of bootstrap aggregation ensembles," 2021. [Online]. Available: https://machinelearningmastery.com/essence-of-bootstrap-aggregation-ensembles

[16] V. Alto, "Understanding the ols method for simple linear regression," 2019. [Online]. Available: https://towardsdatascience.com/understanding-the-ols-method-for-simple-linear-regression-e0a4e8f692cc

[17] G. L. Team, "A complete understanding of lasso regression," 2021. [Online]. Available: https://www.mygreatlearning.com/blog/understanding-of-lasso-regression