Analyzing Trading Markets

San Jose State University

Master of Science in Data Analytics

Aarathi Srinivas Nadathur – 013704405

Amar Ippili – 013728975

Sameer Rajput – 013802841

Shubh Johri - 013750672

Abstract

Big data plays an important and crucial role within the business for creating higher predictions over business information that is gathered from around the world. Finance and trading is a sector, where the big data technologies like Hadoop, NoSQL are creating its mark in predictions from financial data by the data analysts. The movement in the stock market depends upon the losses and capital gains, and its considered to be erratic and unpredictable. But, patterns allow the predictions of some movements and these patterns can be studied by stock market analysis by applying data mining techniques. Hence, we are worked on introducing various techniques to be able to analyze and predict the trading and financial markets.

Contents

[Abstract 2](#_Toc532491068)

[Analyzing Trading Markets 5](#_Toc532491069)

[Introduction 5](#_Toc532491070)

[Techniques for Analyzing Trading Market 6](#_Toc532491071)

[To Find a Trading Idea 7](#_Toc532491072)

[Finding and Turning Right Parameters 7](#_Toc532491073)

[Crossover 8](#_Toc532491074)

[Mutation 8](#_Toc532491075)

[Selections 8](#_Toc532491076)

[Disposing the Model when it becomes Obsolete 8](#_Toc532491077)

[Sentiment Analysis 9](#_Toc532491078)

[Unsupervised Machine Learning 9](#_Toc532491079)

[Supervised Machine Learning 10](#_Toc532491080)

[Job Opportunities 12](#_Toc532491081)

[B-Stock Solutions 12](#_Toc532491082)

[NASDAQ 13](#_Toc532491083)

[Virtu Financial 13](#_Toc532491084)

[RBC Bank 14](#_Toc532491085)

[Capitalogix 14](#_Toc532491086)

[Bloomberg 15](#_Toc532491087)

[LoanDepot 15](#_Toc532491088)

[Conclusion 16](#_Toc532491089)

[References 17](#_Toc532491090)

Analyzing Trading Markets

# Introduction

Data Science is claimed as the sexiest job of the 21st century. Huge amounts of data availability and the rise of data mining and machine learning has led to a conversion of simplest of tasks being done by computers with utmost precision. Over the last few years, data science has revolutionized trading and financial practices of buying and selling financial assets for making a profit. What started off at the floor is now dominated largely by computers. Algorithms are built day by day to understand and predict the trading market. It is making possible rapid split-second decisions faster than humans could ever manage. It is so rapid that many go on to compare the rate at which these decisions are made by the algorithms and the speed of light. Today, machine learning and data mining are an integral part of trading activities and will likely to grow multi-fold. Hence, it makes it important to discuss what Machine learning and data mining are and how do they come into the picture of analyzing trading markets.

Data Mining generally deals with using statistics and other programming languages to find hidden patterns in the historical stock market data to get an idea of how the stock market might behave. It tries to build an intuition about what is really going on in the data. On the other hand, machine learning uses Data Mining technique to build models based on the historical data to be able to predict future outcomes.

We will discuss many examples of the application of data science in trading or more commonly called as quantitative trading. Some of the forms of quantitative trading include High Frequency Trading (HFT), statistical arbitrage, algorithmic trading, etc. We will extensively discuss genetic algorithms to forecast financial markets and neural networks for forecasting profits. Also, with the recent fame the term ‘data science’ has received, we have included a section which describes what a data scientist does, and the job opportunities offered in the industry.

# Techniques for Analyzing Trading Market

For stock exchange analysis both regular data and historical data of specific stock exchange are required for predictions. There are various techniques used for analyzing the unstructured data like stock exchange reviews and historical statistic of economic information. The information gathered over stock message boards are large assets for around 71 percent organizations and uses the historical time series data for accurate predictions of stock market (Kavitha, 2015). It gains new insights of investors as well as financial organizations. Big data technologies help to get a meaningful information from these types of data in financial market. The unstructured data is collected from message board, they are classified based on the user’s sentiment to predict the correct results by integrating the historical data. Financial volatility models are used for analyzing the historical data.

One can be successful in the business of Stock Market and Trading by harnessing the power of Neural networks. There is still a lot of unexplored potential in Neural Networks that can be utilized for Trade and Stock Analytics. But as some would have thought Neural Networks aren’t that magic solution for all your problems where with a click of a button you find results. It takes a lot of time to prepare a neural network on your trading idea and deploy it after which if it were successful, you might expect an increase in at most 10% efficiency in your trading decisions. Sometimes traders make the mistake of following the simplest path by forecasting a price a few bars ahead and basing your trading system on this forecast. Other traders forecast [price change](https://www.investopedia.com/terms/p/price-change.asp) or percentage of the price change. This approach seldom yields better results than forecasting the price directly. Both the simplistic approaches fail to uncover and gainfully exploit most of the important longer-term interdependencies and, as a result, the model quickly becomes obsolete as the global driving forces change.

So, to build a successful Neural Network, we need to follow the 3-step rule (DIMA VONKO, 2009):

1. To Find a Trading Idea
2. Finding and Tuning right parameters of your Model
3. Disposing the Model when it becomes Obsolete

## To Find a Trading Idea

A trader should understand that his or her neural network is not intended for inventing trading ideas and concepts. It is intended for providing the most trustworthy and precise information possible on how effective your trading idea or concept is. Therefore, you should come up with an original trading idea and clearly define the purpose of this idea and what you expect to achieve by employing it. This is the most important stage in the network preparation cycle.

## Finding and Turning Right Parameters

The overall quality of a model can be improved by modifying the data set used by finding the right parameters and tuning them. One of the best approaches to find the right parameters and tune them is by using Genetic Algorithms. So, what is a Genetic Algorithm?

Genetic algorithms are problem-solving methods that mimic the process of natural evolution. These are commonly used as optimizers that adjust parameters to minimize or maximize some feedback measure, which can then be used independently or in the construction of a Neural Network.

These Genetic Algorithms follow a step by step evolutionary model. This evolution of data can be divided in 3 phases:

### Crossover

Like in the process of reproduction where characteristics of a parent is passed to an off-spring, few parameters are selected from a random population where these parameters tend to impact the results.

### Mutation

This is like the biological mutation which are used to maintain the genetic diversity from one generation to other by introducing random small changes. Similarly, small random changes are applied on different parameters in the random population such that they might end up providing the mutated parameters with desirable results.

### Selections

This is the stage where after conducting several Crossover’s and Mutation’s we end up with a Random Population with an increased potential in obtaining the results.

Thus, by using Genetic Algorithms we will be able to find and optimize the right parameter for our Neural Network model. (JUSTIN KUEPPER, 2018)

## Disposing the Model when it becomes Obsolete

Like and Statistical model, a neural-network based model has a life span and cannot be used indefinitely. The longevity of a model's life span depends on the market situation and on how long the market interdependencies reflected in it remain topical. However, sooner or later any model becomes obsolete. When this happens, you can either retrain the model using completely new data, add some new data to the existing data set and train the model again, or simply retire the model altogether.

One can experience success with Neural Networks only by formulating a solid trading strategy. After all a neural network can only validate the effectiveness of your strategy. By following the above steps, one can be successful in yielding results through Neural Networks in Stock and Trading markets.

Below are some other models that can be used with predictive analytics of big data in financial market for improved and better predictions:

## Sentiment Analysis

It is a process of identifying user’s outlook from the feedbacks and reviews from social media. These feedbacks are in unstructured format and can be handled by big data tools, which apply the machine learning algorithm for analyzing sentiments. In financial analysis, it is the feedbacks and reviews on stock exchange dashboard from investors, analysts or by any organization. The sentiments are categorized into positive, negative and neutral. There are two main tasks, first one is to identify the product features from the comments of reviewers and second is to classify the comments as positive, negative or neutral. These tasks can be done by using machine learning technologies with supervised and unsupervised algorithms. (Kavitha, 2015)

### Unsupervised Machine Learning

These algorithms infer pattern from a dataset without reference to known, or labeled, outcomes. Unsupervised machine learning methods cannot be directly applied to a regression or a classification problem because there is no clue what values for the output data might be, making impossible to train the algorithm. It can be used for discovering the underlying structure of data. Unsupervised machine learning techniques include clustering, anomaly detection, association mining, latent variable model and lexicon-based approach for sentiment analysis.

#### Clustering

It allows to automatically split the dataset into groups according to the similarity. Sometimes, cluster analysis overestimates the similarity between the groups and doesn’t treat data points as individuals.

#### Anomaly Detection

It can automatically discover unusual data points in the dataset. It is useful in identifying an outlier caused by a human error, while entering a data.

#### Association Mining

It identifies set of items that occurred frequently in the dataset. Data analyst use it for stock market analysis as it allows them to discover stocks that are purchased at the same time.

#### Latent Variable Model

It is used for data processing, like reducing the number of features in a dataset and breaking the dataset into multiple components.

#### Lexicon Based Approach

It is a practical and easy approach for sentiment analysis of data without a requirement of training. It’s mainly projected to perform task using opinion words in a phrase. Opinion words are those words which express positive or negative sentiments. The count of positive and negative words is used to determine the feature of the product.

### Supervised Machine Learning

These algorithms uncover patterns, relationships and insights from a labeled training dataset. It’s a dataset that contains known value for the target variable for each record. As machine learning algorithm already know the correct answers for a problem during training, it can learn how the rest of the features relate to the target, enable to uncover insights and to make predictions about future outcomes based on historical data.

#### Naïve Bayes Algorithm

It is a classification technique which generates Bayesian Networks for a given dataset based on Bayes theorem. It is easy to build and useful for large datasets and known to outperform highly sophisticated classification techniques. First, convert the dataset into frequency table and create likelihood table by calculating the probabilities of the events. Calculate the posterior probability using the Naïve Bayesian equation. The class with the highest posterior probability is the result of prediction.

#### Support Vector Machine

It is a supervised machine learning algorithm which is used for classification and regression problems. It uses a technique called kernel trick to transform the data and based on the transformations, it finds an optimal boundary between the possible outputs. In this algorithm, we plot each data item as a point in N-dimensional space, where N is number of features, with value of each feature being the value of a coordinate. Then, we perform classification by finding the hyperplane that differentiate the two classes.

(Ray, 2017)

Support Vectors are coordinates of individual observation. Support Vector Machine is a frontier which segregates the two classes (hyper-plane/line), as shown in above figure.

# Job Opportunities

Technology has influenced everything it has touched so far including the landscape of finance and trading. Quantitative trading is where trading opportunities are identified using elements of quantitative aptitude such as mathematics, statistics, coding, and historical data.

Skills required for a Quantitative Trader: math and statistics, statistical programming tools like Python, R, etc.

(Financial Analyst Salaries in San Jose, CA, 2018)

## B-Stock Solutions

B-Stock Solutions Inc. was found in 2008, it operates an online marketplace for overstock and excess inventory. The company provides solutions for enterprises and small businesses to meet inventory situations, recovery rates, and customer returns. The company focuses on marketplace technology, auction strategy, demand generation, marketplace management, and logistics and support for sellers. Its client base includes various industries.

#### Job Position: Marketing Analyst

This key role will be a central resource to a diverse marketing team, to help develop both a view to our customer and business data and to dig deeper to analyze specific needs and trends to better understand how to continually drive both buyers and sellers within our network of marketplaces.

#### Role

1. Use data from multiple sources to identify emerging markets, forecast marketing/sales trends and highlight opportunities for new initiatives.
2. Run statistical analysis and create predictive models.
3. Assist with development of an internal analytics platform. (B-Stock Solutions, n.d.)

## NASDAQ

Nasdaq is a leading provider of trading, clearing, exchange technology, listing, information and public company services across six continents. Through its diverse portfolio of solutions, Nasdaq enables customers to plan, optimize and execute their business vision with confidence, using proven technologies that provide transparency and insight for navigating today's global capital markets.

#### Job Position: Research Analyst

The position requires strong research, writing, and Analytic programming experience (e.g., SQL, R, Python). Analysts become experts in quantitative analysis, macroeconomics, data science, financial products, market microstructure, financial regulation, and exchange competition.

## Virtu Financial

Virtu Financial is one of the largest high-frequency trading and market making firms. It provides two-sided quotations and trades in equities, commodities, currencies, options, fixed income, and other securities on over 200 exchanges, markets, and dark pools. Virtu uses proprietary high-frequency trading strategies to trade large volumes of securities.

#### Job Position: Trading Operations Analyst

Work closely with software engineers to enhance and further develop the firm’s trading strategies. Trading Operations Analysts not only become experts in trading technology and its interaction with the market, but they also play a crucial role in the evolution of market microstructure and the driving forces behind liquidity.

## RBC Bank

RBC Capital Markets is a premier investment bank that provides a focused set of products and services to corporations, institutional investors and governments around the world. They deliver the expertise and execution required to raise capital, access markets, mitigate risk and acquire or dispose of assets. With a strong capital base and consistent financial performance, RBC is among a small group of highly rated global banks.

#### Role

Assist in creating and monitoring complex algorithms in a wide variety of asset classes. Use quantitative analysis to continuously improve strategies and technology. Obtain experience in a wide array of asset classes, including ETFs, equities, options, futures, FX, bonds, and commodities. Assist in developing innovative businesses and trading strategies, conducting important market research, and evaluating models and technologies.

## Capitalogix

It is a financial technology platform that uses adaptive A.I. to maximize trading performance. We track thousands of systems to find what’s working in real-time, all the time, using machine learning, high-performance computing, and statistics. While most A.I. traders focus on using technology to make thousands of transactions in a short period of time, our technology performs millions of observations and only acts on the opportunities determined most likely to have an edge.

#### Job Position: ANSI SQL Database Analyst/Programmer Developer

Maintain existing applications, as well as coding new features and new applications.

## Bloomberg

Bloomberg, the global business and financial information and news leader, gives influential decision makers a critical edge by connecting them to a dynamic network of information, people, and ideas. The company’s strength – delivering data, news, and analytics through innovative technology, quickly and accurately.

As an Equity Research Data Analyst will be responsible for ensuring the consistency, quality and timely delivery of special industry data for a portfolio of companies across a sector, working closely with experienced Bloomberg Intelligence (BI) Research Analysts to learn the ins and outs of conducting investment research while ensuring that we have the best in class industry and company specific data coverage on the market. (Bloomberg, n.d.)

## LoanDepot

LoanDepot, America's lender, matches borrowers through technology and high-touch customer care with the credit they need to fuel their lives. As a fast-growing national consumer lender, the loanDepot platform is disrupting finance by dissolving the lines between mortgage and non-mortgage credit.

#### Job Position: Capital Market Analyst

Capital Market Analysts, function such as trading & pricing and multiple operational groups you will leverage data comprehension valuation skills and financial analysis to gain high visibility throughout a leading financial services company. Skills required: Strong technical, financial, analytical, business and interpersonal skills. Experience with Excel, SQL, and other statistical/data analysis programing languages such as SAS & Python, and development of financial models will help generate quantitative analytics that drive profits to the organization.

# Conclusion

Depending on supply and demand, stock prices change moment by moment. There are many factors that influence prices, the most important of which are expectations on earnings. However, there is no consensus as to why stock prices move in the same way. Every investor has their own strategy, such as particular fundamentals they tend to be fond of and level of risk they are willing to accept. These types of analysis are becoming outdated and more efficient investment tools have become evident to help increase performance. Technical algorithms that can simultaneously analyze many stocks and determine quantifiable target forecasts are becoming increasingly popular with investors as an improved strategy to optimize returns and mitigate prospective risk. In conjunction with reasonable risk management and fundamental analysis, advanced Big Data predictive analytics can potentially help to improve portfolio performance for retail and professional investors.

References

(n.d.). Retrieved from Bloomberg: https://www.bloomberg.com/company/

*B-Stock Solutions*. (n.d.). Retrieved from Linkedin: https://www.linkedin.com/company/bstocksolutions/

DIMA VONKO. (2009, February 18). *Neural Networks: Forecasting Profits*. Retrieved from Investopedia: https://www.investopedia.com/articles/trading/06/neuralnetworks.asp

*Financial Analyst Salaries in San Jose, CA*. (2018, December 8). Retrieved from Indeed: https://www.indeed.com/salaries/Financial-Analyst-Salaries,-San-Jose-CA?from=serpsalaryblock

JUSTIN KUEPPER. (2018, December 07). *Using Genetic Algorithms to Forecast Financial Markets*. Retrieved from Investopedia: https://www.investopedia.com/articles/financial-theory/11/using-genetic-algorithms-forecast-financial-markets.asp

Kavitha, R. V. (2015). BIG DATA ANALYTICS IN FINANCIAL MARKET. *IJRET: International Journal of Research in Engineering and Technology*.

Ray, S. (2017, September 13). *Understanding Support Vector Machine algorithm from examples*. Retrieved from https://www.analyticsvidhya.com: https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/