PORTFOLIO MANAGEMENT FOR PROFITABLE STOCK TRADING IN EMERGING MARKET USING BOT TECHNIQUES: A SURVEY

Omosola Jacob OLABODE[1], Dayo Reuben AREMU[2], Rotimi Oluwasegun FOLARANMI[3], Ebunayo Rachael JIMOH[4], Felix Olukayode AYEPEKU[5]
[1,3,5]Department of Mathematical and Computing, Thomas Adewumi University, Oko, Kwara State, Nigeria
[2]Department of Computer Science, University of Ilorin, Ilorin, Nigeria
[4]Department of Computer Science, Crown-Hill University, Eyenkorin, Ilorin, State, Nigeria
[1]omosolaolabode2@gmail.com, [2]draremu2006@gmail.com, [3]rotimifolaranmi@gmail.com, [4]jimohebunayo@yahoo.com

Keywords: Market, Portfolio, Stock Trading, Emerging Market, Bot Techniques

Abstract: *Predicting stock market prices has always been a challenge, due to the number of variables involved in the prediction process and the volatile nature of the market. Added to this fact is the peculiar nature of emerging markets, signposted by economic underdevelopment. This has aroused the attention and interest of researchers, and a sizable number have beamed searchlight on the application of bot methods for stock trading, the result of which has shown great potential. This paper is premised on the survey of stock market operations and bot trading methods. Discussion centers on achievements recorded in research, using bot approaches to portfolio management for profitable stock trading. Finally, opportunities for further research are presented, with some known constraints and challenges.*

1. INTRODUCTION

Stock is a form of investment in which an investor owns a part of a company's assets. The investor is known as a shareholder or stockholder of the company and is entitled to the company's assets and earnings to the tune of that proportionate ownership. Other names used for stocks are "shares" or "equity". Stocks are traded majorly on the floor of the exchange; however, there are also private buying and selling as well. Stocks form the foundation of nearly all portfolios. Stock trade is being regulated by government laws, and the buying and selling transactions that take place on the stock exchange must conform to these laws so that investors would be protected from fraudulent practices. In the long run, stocks have been found to outperform most other investments.

Shares are part of financial assets that make up the financial market. Other financial assets, also called securities, are bonds, commodities, currencies and diverse kinds of funds. According to [1], a financial market is where stocks and other derivatives are bought and sold. Stocks or shares are investment instruments in which, traditionally, stockbrokers trade on behalf of investors. Trading in stocks simply means buying and selling shares with the sole aim of maximizing profit. It is one vehicle through which corporate businesses are run worldwide. Finance managers are always researching new ideas and technologies to increase profits [2].

Stocks are commonly categorized by type of economic activity such as sectors, industry, and regions [3]. Sectors are the sections of the economy that group companies together with similar products and services, while industry is a subclass of sectors. For instance, banking industry is a subclass of the financial sector, just as insurance industry is also a subclass of financial sector. Regions highlight the level of development of markets, so there are developed markets, such as the ones from developed countries of USA, Britain and Australia, and emerging markets, from developing countries such as Nigeria.

According to [4], many inter-related factors such as political, economic, psychological, and company-related variables affect the performance of stock markets. To achieve high profit investing in stocks, investors have used technical, fundamental, and sentiment analysis to analyze the financial markets in the process.

The expectation of investors for purchasing security is to increase their return [3]. Return is the amount made, after some time, which comes from the investment, and which could either be positive or negative. Positive returns signify profit while negative returns express loss. The possibility of the return not ending up with the desired expectation is called risk, [5]. So, the seemingly simple task of buying and selling is therefore not as simple as it seems. It comes with a certain number of risks. It is therefore a challenging task to trade successfully in the stock market. This is due to the interplay of very complex market dynamics that determine the price of a stock, [6].

A portfolio could be defined as a grouping of financial assets like shares, currencies, commodities, bonds, cash and counterpart funds such as mutual exchange-traded funds. Non-publicly tradable securities such as art, real estate, etc. could also be part of a portfolio. Investors could hold portfolios directly, or they could be managed by professional financial or money managers, on behalf of investors. Ideally, investors are expected to construct an investment portfolio taking into consideration original investment objectives and risk tolerance. Depending on one's objectives, an investor can have multiple portfolios, each meeting a particular purpose.

Bot techniques are the application of computer programs and algorithms to solve a particular task. Investment managers are always in search of new ideas and technologies to increase the chance of better returns on investments. Lately, following development and success recorded in other fields, money managers have been paying close attention to machine learning methods [2]. Software applications are designed solely for financial market operations [5]. Machine learning techniques are now beginning to attract the attention of fund managers, as part of resources used for management functions [2].

Many types of research have been carried out to look for an effective model to trade in stock. For many years, traditional statistical prediction techniques were the main ones available. These techniques include a statistical method known as linear regression (LR) and auto-regressive integrated moving average (ARIMA) methods [6]. Recently, however, machine learning techniques have been attracting a lot of interest within the community of researchers, primarily because the traditional statistical learning algorithms have not been able to adequately address the intricacies associated with the non-stationary and non-linear nature of the stock markets data. According to [7], there are other popular learning methods also, like Genetic Algorithms (GA), Evolutionary Learning (EL), etc., but these are not that very efficient.

This study seeks to surveys the novel idea of using the peculiar economic environment of emerging markets, to investigate bot techniques as possible good candidates for building profitable stock portfolio management trading strategies. Particular attention is being paid to building trading tensors, using data from these markets, and also to leverage the trade-off between the strengths and weaknesses of Policy Optimization and Q-Learning, in training the agents.

The remaining part of this paper comprises section B which presents the literature review, section C gives the summary and discussion of findings, and section D which concludes the paper. Section E gives the list of references.

2. LITERATURE REVIEW

Traditionally on the floor of the exchange, traders carry out trades by finding buyers and sellers through stockbrokers. The process involved in consummating a trade on the floor of the exchange can take up to two or three days to settle completely. The development in technology today has however made the process enable brokers and large institutions to be trading electronically, in real-time, using bot techniques.

The exchange executes a trade in two basic ways. One, manually, on the floor of the exchange and the other one, online trading. Recently, a strong attempt has been made in the financial sector to engage more in electronic trading, and less on the trading floors. This effort has however not gone without some resistance. Some major advanced markets, such as NASDAQ, however, trade only electronically. However, the New York Stock Exchange (NYSE), still operates a human-based manual system. It is not yet clear how far a manually operated system such as that of NYSE will be able to continue to satisfactorily service its, numerous customers, with the very small percentage being handled on the electronic online platform, as compared to NASDAQ, its rival that operates completely on the electronic platform.

Executing a simple trade on a typical floor of exchange can take the following process pattern: First, the number of shares of a particular company that an investor would like to buy or sell at the market is communicated to the investor's Broker, who in turn carries out the instructions of the investor by directing its order department to forward the clerk on the trading floor of the exchange. Next, a signal is sent from the trading floor to one of the traders on the floors of the exchange who looks for another trader on the floor willing to sell a number of that particular company's shares.

The process is not as complicated as it looks here, because the traders on the exchange have complete information about which traders handle markets in particular stock markets, thus facilitating agreement on a price and hence completing the transaction. The process goes back up the line, culminating in the broker to the investor calling to inform him of the final result of the transaction. The length of time the process takes will depend largely on the type of stock and the market. A few days later, a confirmation notice in the mail is received by the investor. The aforementioned narrative represents an example of a very simple trade; considerable detail is involved in the trading of complex portfolios of stocks.

Online electronic markets on the other hand, use a wide network of computers to execute trades by connecting buyers and sellers, instead of using human brokers, as used in the case of manual exchange floor trading. This system is robust, efficient and fast, though it lacks the existing atmosphere of the NYSE floor. Big institutions such as insurance, mutual funds, banks and so on, prefer the electronic market to that manual exchange floor. Electronic trading facilitates almost instant confirmations of trades, especially for the investor. It also facilitates further control of online investing by ensuring closeness to the market. However, since individuals don't have access to the electronic market, a broker is still needed to handle trades on behalf of investors, while the system finds a buyer or seller, depending on the investor's request. These days, electronic market apps are now available on android devices or iPhones, through which brokers could easily place trades.

It is however important for investors to have an idea of how these things work because most of the time, all the details of the transactions will be hidden from the investor, and in case something goes wrong, the investor will not be totally in the dark, but at least have an idea of what is going on to know the correct decision to take.

2.1 Stock Portfolio Management

A stock portfolio is a collection of stock investments. The main constituent parts of portfolio are generally considered to be stocks and bonds, though portfolio could be built and grown by combining many different types of assets such as real estate, gold, paintings, collections and so on. Diversification is one key concept in portfolio management, in which critical factors to consider when assembling and adjusting an investment portfolio are the investment objectives and a person's level of risk tolerance. A key concept in portfolio management is diversification. The critical factors to consider when assembling and adjusting an investment portfolio are the investment objectives and a person's tolerance for risk.

Portfolio management is the art of constructing and monitoring a group of investments to meet the financial objectives and risk tolerance of an investor.

Portfolio management requires monitoring investment over time, involving investments such as stocks, bonds, funds, etc. According to [3], the mathematical framework for portfolio construction is the most accepted in modern portfolio theory, and this field was pioneered by Nobel winner by name Harry Markowitz, in 1952, in his paper titled "Portfolio Selection".

Three sets of rules, follow the Winner, follow the Loser, and uniform constant rebalanced portfolios (UCRP) defines portfolio management strategies. These are the rules which fund managers follow for optimum asset allocation in a portfolio. The Follow the Winner rule works by transferring portfolio weights from assets that are under-performing to those that are performing well.

The follow the Loser approach on the other hand works on the assumption that the under-performing assets will bounce back and outperform the other assets in the subsequent periods. Therefore, portfolio weights are moved from the assets that perform well to the under-performing ones.

The UCRP approach proposes that the budget be equally distributed amongst all the assets in the portfolio without making changes throughout the trading period. In this way transaction cost incurred by the trading, an agent is eliminated.

2.2 Techniques of Stock Portfolio Management

Recent studies in stock analysis and trading use approach that falls under the following five categories: pattern recognition, statistical approach, sentiment analysis, hybrid approach, and machine learning (ML). Presented below are a summary of some of these techniques which have shown to have great potential in the field of stock analysis, and thus gained increasing popularity in recent time:

2.2.1 Pattern Recognition

Pattern recognition and machine learning techniques are much alike. Concerning stock analysis, however, the two are used differently. Pattern recognition is primarily used in detecting patterns and data trends [8], machine learning on the other hand, behaves differently. Patterns, as it relates to stock markets, are sequences that occur repeatedly displayed in OHLC candlestick charts that stock market operators have been using as signs for buying and selling [9]. Two of these pattern recognition methods, which have gained popularity amongst researchers, are Perceptually Important Points (PIP) and Template Matching TM. Perceptual Important Point methods involve reducing the number of data points (i.e. time-series dimensions) by keeping the important points. Template Matching on the other hand works by matching a given stock pattern with a pictographic image for object identification [10].

2.2.2 Statistical Approach

Statistical techniques were the methods being used to analyze and predict stocks, far before machine learning began to gain popularity. There are quite a several studies that successfully employed the use of statistical techniques. One of them used a group of statistical approaches that uses historical stock data as input. These techniques are known as ARIMA (Auto-Regressive Integrated Moving Average), ARMA (Auto-Regressive Moving Average), GARCH (Generalized Autoregressive Conditional Heteroscedastic) volatility, and STAR (Smooth Transition Autoregressive) model [11].

Stock market analysis widely employs the techniques of the ARIMA model [12]. ARMA comprises AR models and MA models. AR models address the momentum and mean reversion effects often observed in trading markets, while MA models, work to capture the shock effect observed in time series data. One limitation of the ARMA model is that it does not take into consideration the unstable nature of historical time series data, a key phenomenon in almost all financial historical data. The ARIMA, on the other hand, as a natural extension to the class of ARMA models, reduces a non-stationary series to a stationary series. A group of statistical models, proposed by [11], that use multiple input variables are Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and regression algorithms. Stock market analysis and prediction have been used to test these statistical techniques.

2.2.3 Sentiment Analysis Approach

Most often, divergence in the price of a stock and the true value of a company's share is usually caused by short-term market fluctuations, which are driven by sentiments. However, over long periods, the fundamentals of a company ultimately cause the value and market price of its shares to converge. Sentiments are an important part of stock markets, and the idea of analyzing sentiments based on various data sources can give insights into how stock markets react to different kinds of news in the immediate and medium-term. Hence a novel approach-sentiment analysis has emerged which gauges the sentiment from data sources or sentiment behind the news to identify its impact on the markets.

An SVM derivative model was proposed by [14] based on three different textual representations, namely the Bag of Words (BoW) model, the Noun Phrases model, and the Named Entities model, to study the effect of breaking news on stock prices within 20 minutes after the release. First, news data was gathered and stored in a database using the three different textual representations. Next, the authors fetched the closing prices of the respective stocks for the last 60 minutes and used Support Vector Regression to predict the price for the next 20 minutes based on the price and sentiment analysis. The experiments performed, using the model, significantly did better than the simple linear regression model in terms of closeness, directional accuracy and simulated trading. The authors also stated that the noun phrases method performed much better when compared to the bag of words and named entities models.

Sentiments analysis based on various data sources can give insights into how stock markets react to different kinds of news in the immediate and medium-term. The authors of [15] performed sentiment analysis and implemented a model, based on Twitter data. The authors used 2-layer NN (N-gram and Word2vec) to analyze the causes of differences between the sentiments tweeted. The study concluded that the accuracy of the model is directly proportional to the amount of available data, meaning that more data would result in better accurate results.

2.2.4 Hybrid Approach

The combination of two or more approaches forms a hybrid model. As an example, a hybrid model is formed when statistical and pattern recognition or statistical and machine learning approaches are combined and applied to analyze and or predict the price of a stock. An interesting hybrid model was proposed by [16] in which Hierarchical Hidden Markov Models (HHMM) were combined with a supervised learning technique, using a decision tree, to predict sensitive index trend of Bombay Stock Exchange, using its historical closing prices, dividends and earnings. Relevant features were selected from the extracted features, using a decision tree, after which prediction was carried out using a set-based classifier. The predicted values were evaluated using HHMM to generate the final prediction. The final prediction resulted in an accuracy of 92.1%.

Various indices have been subjected to a hybrid approach to predict stock price. In one such study, [17] used a hybrid approach to predict the daily stock price of S&P 400 and S&P 500. This study shows that the hybrid models did not do better than its constituent models, while the BPNN model gave very accurate predictions for the daily predictions. The study conclusively showed that most hybrid models would perform better for long-term predictions, but suffered from the instability of the market when applied to daily predictions.

2.2.5 Bot Approach

With the development in technology over the years, the methods employed in the management and optimization of portfolios have been evolving and becoming more sophisticated. In the last few years, several studies have been undertaken to show diverse ways of optimizing portfolios, with bot techniques becoming increasingly popular over the traditional methods.

Machine Learning (ML) bot techniques use models and algorithms to solve problems. It originates from pattern recognition, computational statistics and mathematical optimization. ML models, rather than being explicitly specified, aims at producing a function from data and iteratively work on the data to adapt it by minimizing errors. The quality of results got from the use of this method is therefore dependent on the data it processes and learns from.

According to [18], ML has been used extensively for analyzing the stock market, because of its capacity to predict financial markets. Stock price direction predictions have been carried out using several different algorithms. Recently, better high-performing algorithms such as neural networks have replaced simpler ones, such as the single decision tree and naïve Bayes [19]. Deep Artificial Neural Networks (DANNs) have become a very popular analysis tool in the financial market, with analysis of multivariate nonlinear, data-driven, and easy-to-generalize characteristics. Deep non-linear neural networks are beginning to attract the attention of researchers in the area of historical time series prediction [20]. ANN and Support Vector Regression (SVR) are two widely used machine learning algorithms for predicting stock price and stock market index values [21]. A literature survey of supervised, unsupervised and reinforcement machine learning methods applied in stock market analysis is hereby presented as follows:

(i) Supervised Learning Techniques

Supervised learning techniques like Support Vector Machine (SVM) and Decision Trees can learn to predict stock market prices and trends based on historical data and provide meaningful analysis of historical prices. A subclass of Recurrent Neural Networks (RNN), known as Echo State Networks (ESN),was implemented by [22] to predict S&P 500 stock prices using price, moving averages, and volume as features. The technique outperforms the Kalman Filter technique with a meagre test error of 0.0027. To generalize and validate their result, the authors examined the algorithm on 50 other stocks and reported that their results performed well against the state-of-the-art techniques

(ii) Unsupervised Machine Learning

Unsupervised learning tries to group data into clusters of similarity. It throws up clusters of identical relationships in a disorganized dataset like that of stock markets. A group of researchers in [23] experimented to draw a comparison between the supervised technique, SVM, and unsupervised technique, K-means, by performing Principal Component Analysis (PCA), to reduce the features. The outcome of this experiment shows that the two techniques have similar performance, when the models were tested on S&P 500 data, with SVM achieving 89.1% and K-means 85.6% accuracy respectively. The study also stated how different distance measures for clustering affect the prediction accuracy and that the best performance was when the Canberra distance metric was used.

Another clustering method that tries to group stocks into the best trend and momentum characteristics at a given time was proposed by [24]. The experiment was performed using five-year historical data of stocks listed on the Thailand Stock Exchange (SET). The result showed that, in the long run, the proposed method can outperform the market.

(iii) Reinforcement Learning

Reinforcement learning (RL), a neural network learning method, is concerned with how a software agent takes actions in an environment. It refers to learning how to control a system (environment), to maximize some numerical value, which represents a long-term objective (discounted cumulative reward signal), [25]. Reinforcement learning methods are ways by which the agent learns behaviors to achieve its goal. As a sub-field of machine learning, it can be applied to a gamut of domains including Robotics, Medicine, and Finance [3]. There have been many successful applications of RL to the above domains. Inspired by these successes, this study proposes to experiment with the idea of creating an agent which will be smart enough to understand the dynamics of the stock market and help traders to maximize their profits with minimal supervision and inputs. The proposed portfolio management agent will interact with the stock market environment, taking action as new states are presented to it. And with each action taken, the environment will send feedback to the agent in terms of the amount of wealth made or lost.

Beyond the agent and the environment, one can identify four main sub-elements of a reinforcement learning system: a policy, a reward signal, a value function, and, optionally, a model of the environment [28]. A good understanding of related terms to the subject of reinforcement learning is a prerequisite to the success of this research. A brief explanation of each of these related terms is therefore given below.

According to [3], a state defines the totality of what makes up an environment at a particular point in time, which is often denoted as st. And an observation, often denoted as ot, gives the information available from which the agent can learn. It often happens occasionally that both the state and observation will be the same. For example, the state in a game of chess is the location of each piece on the board. This information is available for both players whether the player is a human or an artificial agent.

The behavior of the learning agent at a given time is termed a policy. Roughly speaking, it could be described as a mapping of state to actions to be taken when in those states, in an environment. It could also be likened to what Psychologists refer to as a set of stimulus-response rules or associations. Policies vary in complexities and type, while in some cases the policy may be a simple function or lookup table, in some others it may involve extensive computation such as a search process. The policy is the core of a reinforcement learning agent in the sense that it alone is sufficient to determine behavior. In general, policies may be stochastic, specifying probabilities for each action.

Put in another way, a policy could be defined as a rule that is followed by an agent to take action on the environment. This is often denoted with μ for deterministic policies or π when dealing with stochastic policies. In practice, policies are parameterized (a neural network for instance) so it is common to subscript them with the letter θ:

$$a_t = \mu_\theta(o_t) \tag{1}$$

$$a_t = \mu_\theta(.|o_t) \quad (2)$$

Being the actual decision-maker in RL framework, the policy is often used interchangeably with agent terms in phrases such as: "The agent/policy tries to maximize the reward. Deterministic policies directly predict the action that should be taken by the agent, while in general, a stochastic policy predicts the parameters for another probabilistic model from which actions can be sampled. In a sense, deterministic policies are easier to understand, a simple feed-forward neural network can be a deterministic policy, the same input the same output. Stochastic policies require a bit more care, to be understood and trained. There are two operations to be taken into account when talking about stochastic policies. First, actions can be sampled from a given policy, and last, likelihoods of particular actions are computed. This study will later provide examples of how the two types of policies are used in various algorithms.

Easily enough, a trajectory is a sequence of action-state pairs:

$$\tau = (s_0, a_0, ..., s_t, a_t) \quad (3)$$

The states evolve according to the laws of the environment and often the dependency on the action is tied only to the most recent action taken by the agent

$$S_t = f(s_t, a_t) \quad (4)$$

$$s_{t+1} = P(\cdot|s_t, a_t) \quad (5)$$

The cumulative reward is what the agent seeks to maximize at the end of each trajectory. The reward function, a real number, measures how well or bad is the performance of an agent. Rewards can be expressed as a function of the explicit dependency of the current state, the action taken by the agent, and the future state of the environment:

$$r_t = R(s_t, a_t, s_{t+1}) \quad (6)$$

Reward, is the only feedback the agent has from the environment, therefore care should be taken while choosing the reward function, as an inappropriate choice of the reward function can lead to poor performance of the system. The situation here presented resembles that of choosing numerical state representations to convey the information needed by the agents. A poorly formed reward function might guide the agent to learn sub-optimally or make the system obtain more rewards rather than solve the problem. It should also be clear from the rewards the terminal state of the environment.

The goal of a reinforcement learning problem is defined by a reward signal. A single number called the reward is sent by the environment to the reinforcement learning agent on each time step. The sole objective of the agent is maximization of the reward received in the long run. What the reward signal thus defines are the good and bad events for the agents. Reward signals could be thought of as being analogous to the experiences of pleasure or pain, in a biological system. They are the immediate and defining features of the problem faced by the agent. They are the primary basis for changing the policy; if an action selected by the policy resulted in a low reward, the policy could be changed to select some other action in that situation in the future. Generally speaking, reward signals may be probabilistic functions of the state of the environment and the actions taken.

What is good in an immediate sense is indicated by the reward signal, while what is good in the long run is specified by a value function. The aggregate amount an agent is expected to accumulate over the future is, roughly speaking, measured by the value of a state. While rewards measure the immediate desirability of environmental states, values on the other hand indicate the long-term desirability of states, taking into consideration the states that are likely to follow and the rewards in those states. A state, for example, might always give a low immediate reward but still have a high value if it is regularly followed by other states that yield high rewards. The converse could also be true. For example, a state might always yield a high immediate reward but still have a low value because it is regularly followed by other states that yield low rewards.

This could be illustrated by using human analogy thus: high rewards are like pleasure while low rewards are pains, whereas values refer to a more refined judgment of how pleased or displeased we are that the environment is in a particular state.

A direct corollary of the above is that, primarily, without rewards, there could be no value, as value is the aggregation of many rewards. However, it is the values that are of utmost importance while making and evaluating decisions as action choices are made based on value judgments. Therefore, actions that bring about states of highest value are sought and not the highest reward because these actions obtain the greatest amount of reward in the long run. Unfortunately, it is much harder to determine values than it is to determine rewards. Rewards are given directly by the environment, but values must be estimated and re-estimated from the sequences of observations an agent makes over its entire lifetime. The most important component of almost all reinforcement learning algorithms we consider is a method for efficiently estimating values. The central role of value estimation is arguably the most important thing that has been learned about reinforcement learning over the last six decades.

Finally, the model of the environment, which mimics the behaviour of the environment, or more generally, that allows inferences to be made about how the environment behaves, is the fourth element of some reinforcement learning systems. For example, given a state and action pair, the model might predict the resultant next state and reward. Models are used for planning, by which we mean any way of deciding on a course of action by considering possible future situations before they are experienced. Methods for solving reinforcement learning problems that use models and planning are called model-based methods, as opposed to simpler model-free methods that are explicitly trial-and-error learners—viewed as almost the opposite of planning.

To gauge the performance of the assets they are interested in financial analysts and portfolio managers use technical indicators. There are many indicators available in the financial domain. However, only a few of them will be considered here in this review study.

The Average Daily Return is a technical indicator that measures the simple mathematical average of a series of returns generated over a period, which is calculated the same way a simple average is calculated for any set of numbers. All the numbers are added together, and then the sum is divided by the count of the numbers in the set. The formula to calculate Average Daily Return is given as:

$$\text{Av. Return} = \frac{\text{Sum of Returns}}{\text{Number of Returns}} \qquad (7)$$

The Sharpe Ratio, another technical indicator, is calculated by subtracting the risk-free rate from the return of the portfolio and dividing that result by the standard deviation of the portfolio's excess return. The ratio describes how much excess return you receive for the extra volatility you endure for holding a riskier asset. The Sharpe Ratio can be calculated as follows:

$$S_{a.} \frac{\lfloor R_a - R_b \rfloor}{r_a} = \frac{E\lfloor R_a - R_b \rfloor}{\sqrt{\mathrm{var}\lfloor R_a - R_b \rfloor}} \qquad (8)$$

where Ra is the asset return, Rb is the risk-free rate, E[Ra-Rb] is the expected value of the excess of the asset return over the benchmark return, and √var [Ra-Rb] is the standard deviation of the asset excess return.

2.3 Reinforcement Learning Techniques

Many studies have already used Reinforcement Learning in trading stock, portfolio management and portfolio optimization [6]. A study titled "Reinforcement Learning Application for Portfolio Optimization in the Stock Market", [3] explored the usage of reinforcement learning algorithms for portfolio management in the stock market. RL agents are trained to trade in a stock exchange, using portfolio returns as rewards for RL optimization. A set of 68 stock tickers in the Frankfurt exchange market was selected, and two RL methods were applied, namely Advantage Actor-Critic (A2C) and Proximal Policy Optimization (PPO). Their performance was compared against three commonly traded ETFs to assess the algorithm's ability to generate returns compared to real-life investments. Both algorithms were able to achieve positive returns in a year of testing (5.4% and 9.3% for A2C and PPO respectively). A European ETF (VGK, Vanguard FTSE Europe Index Fund) for the same period, reported 9.0% returns, as well as a healthy risk-to-returns ratio.

In [6] it was explored how to optimally distribute a fixed set of stock assets from a given set of stocks in a portfolio to maximize the long-term wealth of the Deep Learning Trading agent using Reinforcement Learning. For the problem, the learning agent directly interacts with the trading environment, thus enabling optimization of trading results, through the application of model-free Reinforcement Learning Algorithms. The study focused on Policy Gradient and Actor-Critic Methods, a class of state-of-art techniques that constructed an estimate of the optimal policy for the control problem by iteratively improving a parametric policy. A comparative analysis of the Reinforcement Learning based portfolio optimization strategy vs the most traditional

"Follow the Winner", "Follow the Looser", and "Uniformly Balanced" strategies was performed, and found that Reinforcement Learning based agents either far outperform all the other strategies, or behave as good as the best of them.

2.4 Deep Reinforcement Learning

In "Deep Reinforcement Learning in Portfolio Management", [26], applied the methodology of deep reinforcement learning to the problem with the help of Deterministic Policy Gradient (DPG). Models using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) were designed and developed. The performances were evaluated and compared with several benchmarks and baselines.

The strategies of reinforcement learning (RL) were interpreted by [27] and compared to that given by academic portfolio advice. RL is highly suited for the volatile environment of portfolio management, compared to other machine learning methods in use, chiefly because it is an approximate dynamic programming (DP) model. RL compares favorably with other models in performance ratings, with the possibility of achieving the same average terminal wealth of 33% in portfolio value over five years, as a model with low risk-aversion, reducing the standard deviation of terminal wealth by 30% and having a lower turnover than CP model by three percent.

Three state-of-art continuous reinforcement learning algorithms, Deep Deterministic Policy Gradient (DDPG), Proximal Policy Optimization (PPO) and Policy Gradient (PG) were implemented by [28] in portfolio management. All the three algorithms are widely used in game-playing and robot control. What's more, PPO has appealing theoretical properties which are hopefully potential in portfolio management. The performances of the algorithms were presented under different settings, including different learning rates, objective functions, and feature combinations, to provide insights for parameter tuning, feature selection and data preparation. Experiments were also conducted using China Stock Exchange market and show that PG is more desirable in financial market than DDPG and PPO, although both of them are more advanced

Portfolio Value is the total monetary value of the portfolio obtained by multiplying the weights of the assets by the daily prices. Volume is the number of shares or contracts traded in a security or an entire market during a given time. For every buyer, there is a seller, and each transaction contributes to the count of total volume. That is, when buyers and sellers agree to make a transaction at a certain price, it is considered one transaction. If only five transactions occur in a day, the volume for the day is five.

Absolute asset prices are not directly useful for an investor. On the other hand, price changes over time are of great importance, since they reflect the investment profit and loss, or more compactly, its return. The percentage change in asset price from time (t-1) to time t is called the simple return of the asset.

Transaction costs are expenses incurred when buying or selling a good or service. Transaction costs represent the labor required to bring a good or service to market, giving rise to entire industries dedicated to facilitating exchanges. In a financial sense, transaction costs include brokers' commissions and spreads, which are the differences between the price the dealer paid for a security and the price the buyer pays.

The main goal of any portfolio optimization algorithm is to maximize its value over time. The return of investment is often the most important metric regarding portfolio performance, and often the returns are related to the risk taken. High-risk investments are expected to yield larger returns whereas low-risk investments are expected to achieve smaller but consistent returns.

According to [3], returns are more appealing for modelling than raw price behavior, and in many cases, researchers and investors are not so interested in knowing the exact market prices, but in understanding the trends and the changes happening as time passes. One of the advantages of using returns is that data now is on the same scale, this is a desired feature for machine learning and in particular neural networks. The definition of simple return given for a single asset can be extended to portfolio return, given a portfolio vector and the portfolio constituents' vector of returns at a given time

Portfolio Management Strategies is a set of rules that are followed by portfolio managers for optimal allocation of assets in a portfolio. Follow the Winner is a portfolio management strategy approach characterized by transferring portfolio weights from the under-performing assets (experts) to the outperforming ones.

The Follow the Loser is another portfolio management strategy approach that assumes that the under-performing assets will revert and outperform others in the subsequent periods. Thus, their common behavior is to move portfolio weights from the outperforming assets to the underperforming assets.

The UCRP is a portfolio management strategy approach that suggests the wealth be equally distributed between the chosen assets in a portfolio without making any kind of changes throughout the trading period. This helps avoid the transaction costs incurred by the trading agent.

3. SUMMARY, DISCUSSION AND CONCLUSION

Stock price prediction analysis continues to be a challenging task and thus has been attracting the interest of researchers. These challenging problems increase as more data becomes more readily available. The problem areas include acquiring, processing and extracting knowledge, to examine the effect on the stock price. Some of these challenging problems include, but are not limited to, algorithm trading, issue of live testing, self-defeating nature of bot trading, short-term prediction, and sentiment analysis on company reports. On the issue of live testing, the claim being made in most literature that the proposed bot methods can be used for stock trading to make profits in the stock market maybe is too farfetched a claim to make, in that all such methods may work well during back testing, as they are mostly carried out in controlled environments, but may behave otherwise when used to trade in real life, as a lot of factors such as noise, price instability, etc. are the order of the day in real life trade.

Bot stock trading techniques have caused the way stock markets function used to be, as most trading features are being generated by computer algorithms rather than a human being that used to be responsible. The obvious advantages and benefits of bot trading like reduced cost, non-dependence on sentiments, etc., are equally matched by the challenges it poses for the retail investors who cannot acquire what it takes to build such systems. It is not uncommon today for the market to overreact with panic selling as a result of these systems. And with new computer algorithms flooding the markets, almost on daily basis, evaluating the effectiveness of these systems brings out another kind of problem. Another aspect of this area of research in stock market operations is its nature to cause more problems than solve them ultimately, in that if a model can be used to generate high profits, making it available to all market operators will automatically render the technique self-defeating. Thus, the research or methodologies behind efficient algorithms deployed for stock trading in the market are usually kept very confidential and secret, and never published.

These days, due to effect of the rising influence of social media, sentiment analyses are attracting great attention, with the source of data coming from news or Twitter. However, data from social media can be very unstable, unreliable and addition difficult to analyse, as there are multiple sources of news, most of which are fake. Luckily, there is a good alternative to these sources of social media news in the quarterly or annual reports of companies filed, which can be used for stock prediction. If studied carefully and correctly, these annual and quarterly reports can give clear information about the status of the company, which may assist in the efforts at knowing the future trend of the stock

A very good and unique environment for stock trading and investing, in which devices connected to the Internet can be used to carry out trades, is provided by financial markets. Stock trading over the Internet has drastically changed the way investors buy and sell shares, as the financial markets have rapidly become a robust interconnected global marketplace. These technological advancements have thrown up new opportunities, as the techniques of data science offer many advantages, which are also accompanied by a new set of challenges.

This survey proposes a scientific classification of bot techniques used for analysis and prediction of the stock market, presenting in the process a detailed literature review of algorithms and methods used for these predictions and analysis. Challenges being encountered in this area that requires close attention are discussed, providing opportunities for future development and research direction. Stock markets today, unlike the traditional manual systems are built around different technologies combining, communication with one another to facilitate more robust investment decisions. Such technologies include machine learning, big data, expert systems, etc. Access to the market through the Internet has rendered the market to be prone to the whims and caprices of the customers and hence to malicious attacks. This presents a very good opportunity for further research into making the market highly robust to meet the challenges of the future. This survey has conclusively affirmed that there exists a very good opportunity for research in the direction of exploring various algorithms to find out whether they are good enough, not just to predict stock prices, but to provide a robust profitable stock trading strategy in the long run. Reinforcement ML methods may ultimately likely prove to be more useful as trading strategies and not just price prediction model

.

4. REFERENCES

[1] M. Moukalled, E. Wassim, J. Muhamad, (2019). Automated Stock Price Prediction Using Machine Learning W19-6403

[2] A. Huertas, (2020) A Reinforcement Learning Application for Portfolio Optimization in the Stock Market: [Online; accessed November 2020]

[3] Vulpoiu (2018), M.Sc. Thesis: University of Twente Faculty of Electrical Engineering, Mathematics & Computer Science.

[4] Shalom, O. S., Jannach, D. J., Konstan, J. A. (2019). Towards More Impactful Recommender Systems Research

[5] Kanwar, N. (2018) Deep Reinforcement Learning-based Portfolio Management: Presented to Faculty of the Graduate School of The University of Texas at Arlington

[6] Nair, B. B., Mohands, V. P., Nayanar N., Teja, E. S. R, Vigneshwari, S., and Teja, K.V.N.S(2015). A Stock Trading Recommender System Based on Temporal Association Rule Mining

[7] Nguyen, T. N. and Yoon, S., (2019). A Novel Approach to Short-Term Stock Price Movement Prediction using Transfer Learning

[8] Quang-Vinh, D., (2019). Reinforcement Learning in Stock Trading. [hal-02306522]

[9] Lakshmana, P., Maguluri, and Ragupathy,P. (2020). An Efficient Market Trend Prediction using the Real-Time Stock Technical Data and Stock Social Media Data

[10] Mandar, G., Deshpande, Ashwin, P., Nishi, T. (2017). Intelligent Recommender in Stock Market

[11] X. Zhong and D. Enke (2017). Forecasting Daily Stock Market Return Using Dimensionality Reduction. Elsevier Ltd: http://www.sciencedirect.com/science/article/pii/s0957416305115

[12] M. Hiransha, E. A. Gopalakrishan, V. K. Menon, K. P. Soman (2018). NSE Stock Market Prediction Using Deep Learning Models. International Conference on Computational Intelligence and Data Science (ICCIDS 2018): Procedia Computer Science Volume 132, 2018, Pages 1351-1362

[13] Yang, Y., Li, and Yang, Y. (2015). An Efficient Stock Recommendation Model

[14] Schumaker, R. P. and Maida, N. (2018). Analysis of Stock Price Movement Following Financial News Analysis of Stock Price Movement Following Financial News Article Release

[15] V. S. Pagolu; K. N. Reddy; G. Panda; B. Majhi (2016). Sentiment Analysis of Twitter Data for Predicting Stock Market Movements. https://ieeexplore.ieee.org/xpl/conhome/7948575/proceeding

[16] Y. Xu and S. B. Cohen (2018). Stock Movement Prediction from Tweets and Historical Prices. Association of Computational Linguistics (Volume 1: Long Papers). Pages 1970-179

[17] Rasekhschaffe, K. C.,Jones, R. C. (2019). Machine Learning for Stock Selection. Financial Analysts Journal, vol. 75, no. 3, Available at SSRN: https://ssrn.com/abstract=3330946

[18] Creighton and Zikemin (2012)

[19] Robert, P. S. and Nick, M. (2018). Analysis of Stock Price Movement Following Financial News Article Release

[20] Binoy B. N., Mohandas, V. P., Nikhil N., Teja, E. S. R., Vigneshwari, S., and Teja, K.V.N.S. (2015). A Stock Trading Recommender System Based on Temporal Association Rule Mining

[21] Bruce Yang (2020). Deep Reinforcement Learning for Automated Stock Trading

[22] Dietmar, J., Oren Sar, S., Joseph, A. K. (2019). Towards More Impactful Recommender Systems Research

[23] Weijs, L. (2018) Reinforcement learning in Portfolio Management and its interpretation. ttps://laurenswe.github.io

[24] Fouzan, A. Q., Asem, T., Mustafa A. (2016). Factors Affecting the Market Stock Price - The Case of the Insurance Companies Listed in Amman Stock Exchange

[25] R. Vulpoiu, (2018). Machine Learning Applications in Financial Advisory https://essay.utwente.nl/76709/1/Vulpoiu_MA_EEMCS.pdf [Online; accessed November 2020]

[26] Zhang, D., Mudinas, A., Mudinas, Levene, L. (2019). Market Trend Prediction using Sentiment Analysis:Lesson Learned and Paths Forward

[27] Jinho L., Raehyun K., Seok-Won Y., and Jaewoo K. (2020) MAPS: Multi-agent Reinforcement

Learning-based Portfolio Management System: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20). Special Track on AI in FinTech

[28] Richard S. Sutton and Andrew G. Barto, Reinforcement learning: An introduction", Second Edition, MIT Press, 2019