**Introduction:**

%========= Introduction

\section{Introduction}~\label{sec:introduction}

The stock market is a marketplace where investors may trade assets such as stocks, bonds, investment trusts, and trading funds. This all happens through buyers and sellers. These buyers and sellers can be anyone, can be individuals known as retail investors or can be large organization investors such as managers, insurance companies or could be even banks. In order to purchase or sell securities, one must first understand where they are all listed. The stock exchange, such as the New York Stock Exchange (NYSE) in the United States, Australian Securities Exchange, is where securities are listed. Basically, all the buys and sell orders, which are made by the buyers and sellers, go through traders or brokers who manage to negotiate the best possible price for buyers and sellers.

There are few stock indexes which represents the whole group of stock or their particular sector, S\&P 500 is one of them. They are used to evaluate the performance of the individual stock; they are used as a benchmark for the evaluation. To achieve the best possible price for buyers and sellers, few researchers believed that by anticipating the future price of the stock index they might modify the way trading is generally done.

\subsection{Classification Metrics:}~\label{subsec:classification\_metrics}

For a long time, the problem of predicting the stock market has been a concern for the investors, as the market remains unpredictable. Dynamic, high volatile and non-linear data remains the main reason for the problem, which results in trading difficulties for ordinary individuals. Retail investors traded stocks based on their gut instincts or subjective judgements and they lacked appropriate decision-making aids. Researchers kept dwelling on the problem and been trying hard to find a perfect decision-making model to predict the data trend, which helps the investors to make right decision and preparations for the future. The Efficient Market Hypothesis (EMH) believes that these prediction of market values are not achievable. Any pattern, even if its illogical, which is found during examine of historical data would not persist and would not provide investors any extra returns (Malkiel, Burton, G. 2003.).

Few academics disagree on EMH, as the progress made demonstrates that stock market prediction is conceivable, which has become a major topic in the world of finance. It all began with the use of technical and fundamental analysis, which were the two-primary thought in the aspect of financial market analysis. Technical analysis attempts to anticipate the future price changes using the security's price movements. On the other hand, economic and financial issues that drive a company were examined in the fundamental analysis. The price were also forecasted using linear models such as ARMA and ARIMA. GARCH is another model which was used to make predictions and also used in a variety of disciplines. The fundamental issue with these models is that they are unable to perform when non-linear data is used in the prediction process, which is the case with stock market data. According to EMH, both technical and fundamental investigations are meaningless since stock's price comprises all the market information.

Since the rise of artificial intelligence, one of the most important things people have realized is its capacity to do predictive analysis, and they feel it is the ideal match for the most complex field, which is stock market. Not only that, a significant amount of comprehensive historical data from many corporations and stock exchanges assisted algorithms in uncovering hidden patterns in the data. These algorithms are capable of processing data and predicting the near future changes in the stock price.

The advancement of machine learning has resulted in a few models that have assisted in the prediction process. (2) Linear regression, SVR, DTR are few models which were used in the prediction. Every model has a unique method of learning and forecasting. The features and respective labels are being memorized by the classifier. It recalls the combinations of features and the labels that corresponds to them, in our instance, it’s the stock price in near future. Then it continues to figure out what patterns the features are using to create their individual labels. Even though the prediction accuracy of these models were good, advanced models better than machine learning was needed to extract the deep information from the time series data.

To overcome this challenge and to deal with dynamic data, neural networks came into picture. But the problem is, (3) traditional neural networks have shallow neural network, since it only has one or two hidden layers, which minimize the analysis accuracy. To overcome the traditional neural network problem, Deep learning algorithms are used to identify patterns and correlations in data through a self-learning process. On the other hand, deep learning models are gaining popularity, and are being employed in a variety of domains such as image recognition, natural language processing, and more. These development in the field of deep learning has made researchers to apply deep learning in the field of finance, and more precisely in stock market prediction. These models outperformed linear and machine learning models, because of its ability to detect hidden patterns and learn high-level features from the data. Recurrent Neural Networks (RNNs) are the most widely used Deep Learning approach for Time Series Forecasting because of its accurate prediction on the time series data in a variety of problems. Even the input size increases, the model's size does not rise. RNNs vary from feedforward and CNN based on their "memory," which allows them to influence current input and output based on prior inputs. The inputs and outputs of traditional deep neural networks are assumed to be independent of one another.

These novel prediction models have advanced rapidly in recent years, even the changes could be seen in input data. Apart from market data, data such as twitter data, sentiment analysis and investor’s trading data provide new source of input which helps with predicting and improving the model. Diverse data sources has diverted researches point of view on different factors which can be the reason for market fluctuation. Even few studies have considered multi-dimensional historical data, where the input is no longer single index as most of the studies. Different combination of algorithms is even used with the prediction models, to increase the data processing techniques, to improvise the model performance and to enhance the model accuracy. These new models and other combinations of algorithms are being developed in order to forecast the stock market trend and price with great accuracy.

This Survey is more about that, recent trends and developments which have been made in deep learning models in the field of stock market prediction. This will help researchers, enthusiast, and even new beginners to be in update with the progress. We review the most recent advancements made in applying deep learning approaches to stock market forecasting, especially those which appear from 2020. We have also explained the algorithms and techniques which are used to enhance the model performance and forecast stock market trends.

\subsection{Structure}~\label{subsec:structure}

Section~\ref{sec:overview} describes the overview of the survey. Section~\ref{sec:dataset} presents the description about data set used. Section~\ref{sec:processing} describes the approach and data processing techniques in detail. Section~\ref{sec:technical\_indicators} Describes the technical indicators used. Section~\ref{sec:predictionmodels} Briefly presents the prediction models. Section~\ref{sec:comparative\_analysis} shows the models which are used to for comparative analysis. Section~\ref{sec:evaluation} discusses evaluation matrix which is used to validity the model accuracy and Section~\ref{sec:conclusion} concludes the report.

**Aim & Objective:**

we summarize the latest progress of deep learning techniques for stock market prediction

especially those which appears/papers published after 2020

Based on our summary of the surveyed papers, we try to point out some future research directions in this survey, which would help the readers to choose their next movement.

We summarize the latest progress of applying deep learning techniques to stock market prediction, especially those which only appear in the past three years. 2. We give a general workflow for stock market prediction, based on which the previous studies can be easily classified and summarized. And the future studies can refer to the previous work in each step of the workflow. 3. We pay a special attention to implementation and reproducibility, which is often neglected in similar surveys. 4. We point out several future directions, some of which are on-going and help the readers to catch up with the research frontiers. 5. Last but not least, an open GitHub repository on this topic is created 1 , where relevant studies will be collected and updated continuously

# Malkiel, Burton, G. 2003. "The Efficient Market Hypothesis and Its Critics ." Journal of Economic Perspectives, 17 (1): 59-82.

# DOI: 10.1257/089533003321164958

# Stock Market Analysis using Supervised Machine Learning

# <need to find paper>

# <need to find paper>