**Ismael Ben Daoud Dissertation**

**Abstract**

-

**1. Introduction**

* 1. **Research background**

The main area of study proposed is Machine Learning. Machine learning is a branch of artificial intelligence which focuses on the use of data and algorithms to imitate the way that a user can learn and improve overall accuracy [1] Machine learning is classified into three types of categories: supervised, unsupervised, and reinforcement learning [2] Machine learning has the benefit of being able to evaluate massive amounts of data and uncover unique trends and patterns by comprehending different sorts of behaviours to assist cater to the proper outcomes. Machine learning has the issue of requiring a large dataset to train on, which might affect a delay in the process while waiting for new set of data to be created. As stated by [3] machine learning aims to create models that mimic and generalize data. These models learn how to differentiate between different items in order to reach the intended results.

The researcher will concentrate on the return-on-investment ROI in this study by using supervised learning, also known as supervised machine learning, which is a subclass of artificial intelligence and machine learning. It is distinguished by the use of labelled datasets to train algorithms to properly categorize data or predict outcomes.

The researcher aims to forecast the forex exchange market with LSTM and Neural Networks by creating a bot using Python script and generate stock data from a historical trend. This would help better understand the forex market and provide a more refined experience for future trading. The research method will be quantitative, as the data will be collected from existing forex data to monitor and analyse how the market is changing and developing creating a detailed outlook and unique perspective on how the market is moving over a period. The experiment will make use of past datasets to analyse the factors and types of markets that can be used for predictions. The experiment will be developed using Python to deploy a bot and use it to scrape data for market insight which can be made to establish different datasets per each market.

**1.2 Research purpose**

The key objective of this research is how machine learning maximises profits by implementing machine learning to improve decision-making from predictions using past data and current exchange rates. Many machine learning methods are employed in sentiment analysis for sentiment categorization. More individuals are expressing their feelings and thoughts on the internet and through various social media platforms. This has resulted in a rise in the number of user-created phrases carrying sentiment information, making it hard for humans to read and analyse them all. Thus, automatic analysis of opinions expressed on various web platforms is becoming progressively important for making effective decisions. However, it is also difficult because extracting emotions is a complex activity that requires understanding the content and capturing hidden sentiments in written text, which necessitates the study of new methods.

Also, more educated decisions could be made to boost the predicted number, returned by the system. When it comes to the actual implementation, the predictions would also be checked by industry experts in this field, to give feedback on the accuracy on the implemented solution.

Furthermore, this study can be used as a reference by other researchers conducting similar studies, in other words to obtain the relevant and sustainable information about the use of machine learning techniques or it can be also used as a reading material for any person who is interested in this area of study.

**1.2.1 Hypothesis and research questions**

The Hypothesis of this research is:

*“The possibility of maximising profitability by implementing machine learning to improve decision making from predictions applying past data and current exchange rates.”*

This study will attempt to prove or disapprove key factors in forex prediction. This study will seek to confirm or disprove important forex forecast factors. The primary goal of this research is to improve market value forecasting. Experiments will be conducted using online surveys and experiments to analyse the market’s relevance using machine learning algorithms. The findings will be executed and compared to previous studies.

The research questions in this study are:

RQ1. Are movements in rates of Forex exchange predictable when taking into consideration past datasets with exchange rates?

RQ2. Can Machine learning contribute to an optimal decision when dealing with Forex exchange?

These research questions aim to study key factors in forex market prediction. This section will present a detailed idea about how the proposed study will be conducted. This section will include the target participants and a sample size which will include data collection methods and data analysis. The target participants for this project are people who are knowledgeable and work in this specific type of field and a sample size will be of 20 participants. For the data collection methods will include an online survey created by google form which will be used to evaluate the data sets. For the experiment section the use of python programming language will be implemented to construct a machine learning algorithm to forecast particular foreign exchange markets. In the data analysis section descriptive statistics data will be used to characterize the data sets properties. In order to address the method hypotheses and research questions, an evaluation of existing research on the subject and review of related techniques adopted by different markets will take place to study their implementation and effectiveness determining how different markets are created, understanding how they work and how it can be utilized, how generated content can be implemented. From the research conducted a combination of various techniques will be analysed and used to develop a prototype of the proposed project. The prototype will focus more on the forecasting of the market. Will also provide a detailed description of the data being used and how it was cleaned and split into training and testing models, additionally, gives a better understanding of the variables present within the research. Lastly, A detailed description of the implementation developed will be provided, in addition to the algorithms utilized. Further to this, metrics and accuracy measures are listed.

**1.2.2 Significance of this research**

The key objective of this research is how machine learning maximises profits, by implementing machine learning to improve decision making from predictions using past data and current exchange rates. According to Lasod and Pawar, [4] many machine learning methods are employed in sentiment analysis for sentiment categorization. More individuals are expressing their feelings and thoughts on the internet and through various social media platforms. This has resulted in a rise in the amount of user created phrases carrying sentiment information, making it hard for humans to read and analyse them all. Thus, automatic analysis of opinions expressed on various web platforms is becoming progressively important for making effective decisions. However, it is also difficult because extracting emotions is a complex activity that requires understanding the content and capturing hidden sentiments in written text, which necessitates the study of new methods [5].

Also, more educated decisions could be made to boost the predicted number, returned by the system. When it comes to the actual implementation, the predictions would also be checked by industry experts in this field, to give feedback on the accuracy on the implemented solution.

Furthermore, this study can be used as a reference by other researchers conducting similar studies, in other words to obtain the relevant and sustainable information about the use of machine learning techniques or it can be also used as a reading material for any person who is interested in this area of study.

This research can contribute to investigating the potential of various machine learning techniques within forex exchange and to detect whether exchange rates may be used to forecast stock movements. An additional contribute is to learn whether machine learning could potentially contribute to effective decision making within a forex exchange.

**1.3 Importance of research**

This research is important for several reasons. Firstly, the researcher has a great interest in both software and business aspects. As a business analytics student:

It is a great opportunity to develop a project-based on both software sector and business sector combined.

Using different techniques of analysing the effects of forecasting on the market.

Over the years a huge difference has been seen in the forex market, so this project can show different types of development and present figures of how a market is doing.

Secondly, Machine Learning is an emerging technology within the Artificial Intelligence field. The researcher has an interest into understanding how machine learning techniques work, the requirements of making a good machine learning system, their methods and uses.

Machine Learning has been applied in many sectors even in Bioinformatics and for various applications like weather predictions. There is few research on the study of Machine Learning when it comes to predicting potential trading but not in a great depth. Therefore, this will allow the researcher to delve in more profundity into this.

Furthermore, this study can be used as a reference by other researchers conducting similar studies, in other words to obtain the relevant and sustainable information about the use of machine learning techniques or it can be also used as a reading material for any person who is interested in this area of study.

**1.4 Research Boundaries**

The context of this research is can Machine learning be used to predict the outcome of the forex exchange market. This research will cover how Machine learning has the benefit of being able to evaluate massive amounts of data and uncover unique trends and patterns by comprehending different sorts of behaviours to assist cater to the proper outcomes and it will not cover other data but only forex exchange data. Reinforcement machine learning techniques will not be used in this project. In this study a deep learning-based model integrated with sentiment analysis will be implemented for forex prediction. The prototype will be trained to predict the market prices of several markets based on the dataset given. Sentiment analysis will be amalgamated into the dataset consisting of financial news headlines. Additionally, deep learning will be used to compare and construct various techniques in forex price predictions.

**1.5 Research Outline**

This research will contain the following chapters:

**1. Introduction:**

This chapter has given an overview of the research topic chosen. This chapter included Research Background, Research Purpose, Hypothesis and Research Questions, Significance of this Research, Importance of Research, Research Boundaries and Research Outline and Conclusion.

**2. Literature Review:**

This chapter will give an overview of the literature that is related to the research topic chosen. This chapter will include introduction, forex, dealing stocks, types of trading markets, types of forex analysis, machine learning, supervised learning, unsupervised learning, neural networks, deep learning networks, recurrent neural networks, long short-term memory networks, and natural language processing.

**3. Research Methodology:**

This chapter will discuss the methodology used for the research topic chosen. This chapter will include: Research Strategy, Data Collection Methods and Tools, Prototype, Pilot Testing, Errors and Ethical Considerations and Conclusion.

**4. Analysis of Results and Discussion:**

This chapter will contain the data analysis methods used throughout this research and the Analysis of the data gathered. This chapter will include the Data Analysis Method, Analysis and Discussion of the Online Survey, Analysis and Discussion of the Experiment, Analysis and Discussion in relation to the Literature and Analysis and Discussion in relation to the Hypothesis and Research and Conclusion.

**5. Conclusions and Recommendations:**

This chapter will give an overview of the recommendations and limitations of this research. This chapter will include Recommendations, Limitations, Future Work and Concluding Remarks.

**1.6 Conclusion**

This chapter has discussed the research background, research purpose, hypothesis and research questions, the significance of this research, the importance of this research, research boundaries and the research outline. The next chapter reads the key literature used to build a solid basis for this research will be outlined. This chapter will include: introduction, forex, dealing stocks, types of trading markets, types of forex analysis, machine learning, supervised learning, unsupervised learning, neural networks, deep learning networks, recurrent neural networks, long short-term memory networks, and natural language processing.

**Chapter 2**

**Literature Review**

This chapter will give an overview of the literature that is related to the research topic chosen. This chapter will include introduction, forex, dealing stocks, types of trading markets, types of forex analysis, machine learning, supervised learning, unsupervised learning, neural networks, deep learning networks, recurrent neural networks, long short-term memory networks, natural language processing natural language processing (NLP).

Diagram

Description automatically generated

Figure 2.1: Literature Map

**2.1 Introduction**

Due to the volatility of the stock market, researchers find it challenging to find a solution to forecast price movement. Before the age of computers, people traded stocks and commodities primarily on intuition. As the level of trading and investing grew, people searched for methods and tools that would increase their gains whilst minimizing their risks. Therefore, traditional stock price prediction methods including a statistical approach has been implemented such as statistics, technical analysis, fundamental analysis, and linear regression, which are all used to attempt to predict and benefit from the market’s direction. However, none of these techniques has proven to be consistently correct prediction tool that is desired, and many analysts argue about the usefulness of many of the approaches. However, machine learning approaches are proven to be more accurate and efficient in predicting price movement than traditional methods [6] and [7]

Research on stock prediction is typically based on three primary data sources: financial news, historical stock information and social media. A hybrid approach proves to be more efficient because all these factors affect the volatility of the market. Multiple techniques frequently are used due to the ability to maintain a large amount of data. Long-short-term memory (LSTM) is a type of artificial recurrent neural network (RNN) [8] [9]. This section will outline different types of Machine learning techniques and types of techniques used by researchers to predict stock prices and market movements.

Diagram

Description automatically generated

**Figure 1: Process of Forex Market Prediction**

**()**

**2.2 Forex**

According to Investor.gov [10], A stock is a type of security that signifies ownership in a corporation and represents a claim on the part of the corporations’ assets and earnings, which are also called “equities”. Stock is also known as shares or equity. Owning shares in a company gives the owner the right to vote in the shareholder meetings, receive dividends and the right to sell the shares. If one owns more shares, then they have more voting power and hence will have indirectly control of a company. However, for a stock trader, the primary concern is the value of the stock.

Stock Trading is done at stock exchanges where traders buy and sell stocks. In modern days, most trades are done through electronic communications instead of a physical trading building. As this allows the traders to be immediate. There are many stock exchanges around the world such as London Stock Exchange (XLON), Shanghai Stock Exchange (SSE), Securities Dealers Automated Quotations (NASDAQ) and the New York Stock Exchange (NYSE). Due to the nature of the modern markets, many of the exchanges are linked digitally, and hence they can have an impact on each other’s performance [11].

**2.2.3 Dealing Stocks**

As mentioned above, for most stock traders the most crucial aspect of stocks is the ability to sell and buy to make profits. A company’s stock price is initially set based on the company’s value, revenue, and some other factors. After that its value changes based on the availability and demand of the stock. When a large portion of the available stock is purchased, then the stock’s value goes up after that purchase [12]. Similarly, if a large portion of the stock is sold and available, then its value goes down. In stock markets, there is always an upward trend due to inflation and other forces because new money goes into the market every day. Hence making a profit does not automatically imply a good trading system as one can merely buy well fairing stocks, and over a long period of time, they lend to make a profit. Typically, it is essential for a trader to understand what a company’s stock should be valued at. Therefore, having this information will allow the trader to buy those stocks if its present value is below what it should be and sell it if its value is above the actual value [13].

**2.2.4 Types of Trading Markets**

There are typically two types of trading techniques, long term, and short term. As mentioned above, making a long-term investment in settled forex will give profit over a longer period of time. This approach requires less time and effort from the trader and returns/profits over time tend to be more than from bonds and other investments. In comparison to zineb [14] these techniques are very promising when it comes to forecasting fields because of its ability to memorize data.

In short-term trading, the trader aims to profit from small changes in stock prices. These changes can be over as little as a few seconds, called intra-day stock trading, or over a day to a few days, which called swing trading [15]. Short-term trading can be very time consuming, especially when done manually and the trader must spend a lot of effort and time. As, the trader has to track the latest stock changes and news to earn the best possible profit [16]. For this thesis, the focus will be a trading approach. Since it is so time consuming, having a machine learning algorithm predicts the price movements that can be beneficial for a trader. [17]

**2.3 Types of Forex Analysis**

The concept of stock volatility is well understood. In comparison [18], said that this volatility is caused by the numerous forces that influence stock prices, and it is because of these numerous factors that predicting stock prices is so difficult. Despite extensive research and a wide variety of articles and dissertation papers on the subject, there is no complete technique or algorithm for predicting accurate stock value. According to some studies, predicting stock prices accurately is nearly impossible. Even so, if we attempted to forecast stock market prices, some of the numerous factors that could affect stock prices include historical price trends, politics, psychological feelings toward a stock, global economics, and a company's financial condition (Introduction to the foreign exchange market Violeta [19]. Each of these elements can be segmented further. This demonstrates the diverse nature of the factors that influence stock prices. It is simply not possible for a single algorithm to comprehend and evaluate the significance of each of these considerations on the value and market rate of a particular investment. However, it is possible to obtain a prediction within a specific error range by considering only a subset of the numerous factors that can influence stock prices. People use various methods to analyse stock price movements due to the numerous factors. These are broadly divided into two types: Fundamental and Technical Analysis as stated by [20].

**2.4 Machine Learning**

According to [21], Machine learning is the study of how intelligent human behaviour can be imitated with the use of machines. Humans and machines have similarities: both use memory and electrical signals to transmit and retrieve data, and to reach conclusions based on data given. Machine learning which is a sub field of AI tries to achieve human actions such as understanding written text, recognition of visual scenes and performing actions based on sentiment. Just like humans before classifying visual objects, or text analysing, the machine needs to learn by being supplied a data set on that specific scenario.

Most of the modern machine learning research focuses on creating various algorithms that can be used in specific applications. Using machine learning a model is created from data. This model can then be used to make predictions. The main difference between regular computer code and machine learning code is that in a machine learning code the rules are not hard coded, the code creates a model which has its own set of rules [22] A prime requirement for machine learning is having access to a large amount of data. Generally, the better the data the better the machine learning algorithm. The algorithm tries to find a correlation between the input and the expected output. This is accomplished by separating the data into two groups – training data and testing data. The model is generated with the training data, and its performance is tested on the testing data. A significant part of machine learning is getting the right data. The data must be relevant and of the correct size. More about this is discussed in the later part of this thesis. [23]

Diagram

Description automatically generated

Figure 2: Advantages and Disadvantages of Machine learning [5]

**2.4.1 Supervised Learning**

According to [24] Supervised learning algorithms analyse the training data and create a function, and this function can be used for mapping new examples. An ideal case scenario can produce fantastic results with supervised learning algorithms. It can give insights into instances which were unexpected and help improve the overall functioning of an application. For instance, in stock markets, if the right data, in the right conditions are applied to a supervised learning algorithm then it should be able to give a very high accuracy for the output because, theoretically, it can create patterns that were previously unknown to man.

Supervised learning algorithms can be used where we have labelled data, meaning there is an explicit dataset consisting of an input and expected output data. Expected output, also known as target data, is the output for a given input data which the machine learning algorithm should learn. Data can be generated using various means depending on the application and the datatype. For finances, input data is usually the stock values, and the target data is usually a value over a particular time jump. If we are trying to predict stock value over one day, then the expected output would be the close values of stock after the next day [25].

**2.4.2 Unsupervised Learning**

Similarly [26] mentioned that unsupervised learning algorithms are used where there is training data which does not contain any information about the desired output. The main purpose of unsupervised learning is to understand the data by getting the basic structure of the data. Unlike supervised learning, with unsupervised learning, there is no correct answer as there is no output for the training data to learn from. The algorithms learn from just the input data and discover some exciting structures from the data.

**2.5 Neural Networks**

In comparison [27], states that the ability of neural networks to extract patterns from complex data is what makes them so accessible. Neural networks are made up of a large number of interconnected elements called neurons that work in parallel to solve a problem. This early approach to neural networks was heavily based on the neural setup of the human brain, an idea that is still loosely followed to this day. Since their introduction, neural networks have been popular on and off, owing to a lack of hardware capability to handle neural networks for complex data. However, there has been a surge in interest in neural networks in recent years, accompanied by a flood of research papers. [27]

A picture containing shape

Description automatically generated

Figure 3: Artificial neural network architecture. [28]

Figure 3 is an example of a simple neural network architecture. IN this type of network, information never is being fed backwards, when calculating the output for the best possible outcome. Neural networks consist of the following components.

**2.6 Deep Learning Networks**

Although a deep neural network can have significant drawbacks, such as disappearing gradients, it has recently been used in a number of applications, including speech and picture recognition. The shallow neural network in Figure 3 is a neural network with only one hidden layer. There are several hidden layers in a deep neural network. A deep neural network can perform significantly better than a shallow neural network in many complicated circumstances if it has been properly trained. [29]

According to [30] the image-based and sequence-based fields of use for deep neural networks are both common. Recurrent Neural Network, or its variant, is particularly helpful in sequence-based applications, and Convolutional Neural Network, or one of its variants, can be utilized in image-based applications. It can also perform well in text- and speech-based applications. Given that this thesis makes use of a specific type of recurrent neural network, the topic is covered in more detail in the following chapter.

**2.6.1 Recurrent Neural Networks (RNN)**

Recurrent neural networks are one of the few varieties of neural networks that have feedback, in comparison to [31] The inputs are thought to be independent of one another in a conventional neural network. A sort of neural network called a recurrent neural network (RNN) is designed to learn from sequential and time-varying patterns. Due to the non-linear nature of the stock market, RNN provides an excellent method for preserving historical data and making precise predictions.

Data may not always be independent for many real-world [31] applications, especially when it comes to sequential data like music or text. The network can comprehend the context of a sentence or a paragraph from textual data or a series of tones from a musical recording that is agreeable to human ears thanks to feedback. As a "memory," the RNN's feedback loop enables for the gathering of data regarding specific calculations that have already been made.

Diagram

Description automatically generated

Figure 4: An architecture of Recurrent neural network

[32]

As it can be noticed in the following Figure 4 is distinguishing an input layer, which is responsible for the normalisation of the meaning most often it is re-scaling the input data to a given range, which is responsible for the singles processing and output later, as there are at least as many neurons in it as there are outputs from the network [32]. The terms in Figure 4 are:

*Y(x)* = the vector of feedback outputs at the discrete-time instant specified by (x),

*Y1(k)* = the outputs at the discrete-time instant specified by (.), I = 1, n,

*Du* = maximal level of input delay,

*Dy* = maximal level of feedback outputs delay

**2.7 Long Short-Term Memory Network**

Long Short-Term Memory units are a special type of units within a Recurrent Neural Networks. An RNN consisting of Long Short-Term Memory units is called a Long Short-Term Memory network and will be referred to as LSTM henceforth in this thesis. The vanishing gradient issue can be overcome using this approach. LSTM was first proposed by German researchers Sepp Hoch Reiter and Juergen Schmid Huber in 1997. [33]

An LSTM model has the capacity to learn and analyse past data. This unique unit, often referred to as a memory cell, offers a suitable answer due to the complexity of the data structure seen in the stock market. Since LSTM uses special units in addition to conventional units and helps to store information for a long time, it is an enhancement over RNN. [33]

The main difference between a regular RNN and an LSTM is the ability of the LSTM to distinguish between when the memory needs to be cleared and when data needs to be read from memory. In an RNN, incremental memory is always used, which makes it impractical for long-term dependencies.

Diagram

Description automatically generated

Figure 5: Architecture structure of Long Short-Term Memory Unit [34]

Figure 5 depicts an LSTM unit that employs gating, which is essentially just component-wise multiplication. It also shows the forget gate (*f*), input gate (*i*) and output gate (*o*). The stock's historical data is stored in the memory call (*x*). The input gate determines whether the input is written to the cell, the output gate determines whether the cell's output should be read, and the forget gate determines whether the previous cell value should be reset.

To ensure national economic stability, it is critical to predict stock prices in a timely and accurate manner. However, because of market volatility, stock price prediction has problematic characteristics, as wang highlights [35]. The beauty of this network is that it employs the gating approach, which teaches the network when to turn the switches on and off. As a result, it will determine which parts of the information are required and which are not.

**2.8 Natural Language Processing**

[36] mentions that sentiment analysis is an NLP subtype that studies human emotions, opinions, and attitudes towards something more. Sentiment analysis is a classification process that can be divided into three categories: documentation, sentences, and feature level. The primary objective is to classify sentiment shared in each sentence, whereas document-level sentiment analysis labels sentiment on the entire data set, which is typically expressed as a positive or negative sentiment. The feature level attempts to categorize sentiment using specific parts of the data. [36]

**2.9 Sentiment Analysis**

Sentiment analysis is used to discover users’ opinion, preferences, opinions, level of satisfaction and so on. [37] (Low Cheng Kuan\*, 1Maizatul Akmar Ismail, 1Tasnim M. A. Zayet, and 2Shuhaida Mohamed Shuhidan.) Sentiment analysis also involves detecting, extracting, and classifying opinions, emotions and attributes related to a topic from a text data. For instance, comparison with products and descriptions or movie reviews where users freely express opinions, news is objective without using explicitly positive or negative vocabulary, which makes its difficult for sentiment analysis. [38] (Jae Jung Han & Hyun-jung Kim)

An approach that has been explored for prediction within FOREX and stock exchange was by utilizing sentiment analysis. [39] Sreyash Urlam, Bijit Ghosh, Dr. A. Suresh (2021) indicated that sentiment analysis is a mechanism of analysing, recognizing, and categorizing by computing a set of text such as sentence, paragraphs or even pages in order to evaluate if the essence of the text is positive, negative or neutral. It's being exploited in several societies relating to commercial advertisements and research for better understanding the moods of the people which not only helps them in targeting the right audience but also in understanding what changes need to be made based on the result.

**3.0 Conclusion**

This chapter provided a summary of the forex exchange market and machine learning methods associated with forex exchange price prediction. This section has discussed several market conditions that affect the financial markets, as well as the various ways sentiment analysis can analyse, process, and recognize a statement. According to the research, forex price prediction is challenging to implement because market prices fluctuate constantly. A hybrid approach of historical stock information, financial news, and social media is the best technique for researchers. Lastly, further research demonstrates that LSTM outperforms other machine learning techniques in terms of performance. The following chapter focuses on the research methodology and the implementation of the model used to predict forex movement. This chapter gave an overview of the literature that is related to the research topic chosen. This chapter will include: introduction, forex, dealing stocks, types of trading markets, types of forex analysis, machine learning, supervised learning, unsupervised learning, neural networks, deep learning networks, recurrent neural networks, long short-term memory networks, and natural language processing. In the next chapter, the Research Methodology will be discussed which includes: the research strategy, data collection methods and tools, prototype, pilot testing, errors, and ethical Considerations.

**Chapter 3**

**Research Methodology**

This chapter will describe the adopted approach by this study. The main objectives of the research are outlined. This chapter provides a detailed description of the data being assessed. Furthermore, how it was cleaned and split into training and testing models. To give a better understanding of the variables present within the research. A detailed description of the implementation developed is provided, in addition to the algorithms utilized.

This chapter also discusses the method used and the implementation of machine learning techniques for stock prediction using Long Short-Term Memory (LSTM) and Neural Networking processing language (NNL) analysis to obtain the results. This chapter will include Research Strategy, Research method, Data Collection Methods and Tools, Prototype, Pilot Testing, Errors, and Ethical Considerations.

**3.1 Research Strategy**

Deductive reasoning is the process from one or more statements to reach a logical conclusion. As stated by [39], there is less uncertainty in this reasoning because a conclusion is reached reductively by applying general rules which hold over the entirety of a closed domain of discourse which narrows the range under consideration until only the conclusion is reamed. By applying deductive reasoning, the following conclusion was achieved.

To meet the requirements of the hypothesis, applied research was adopted. This involves the use of machine learning algorithms onto the historical datasets. For the data web scraping from yahoo finance was used within this study. Within the research itself, a Quantitative approach was being utilised in terms of analysing the data and the results. Within this research, there are an independent variable and a depended variable being presented.

Diagram

Description automatically generated

**Figure 6 – Variable Framework**

In this study, the independent variables consist of the closing Price of the historical datasets. Whereas the depended variable will be the predicted price values.

**3.1.1 Research Method**

This study focuses on machine learning techniques for stock prediction using neural network long short term memory analysis. As outlined in Chapter 2.2.3, market prediction has several complex characteristics. Since external factors have an effect on stock movement, several techniques are combined in the creation of the models. As described in previous chapters Neural networks are made up of a large number of interconnected elements called neurons that work in parallel to solve a problem thus, the use of historical data with a period of 5 years are included in this research. Stock prediction studies used features from historical data such as: Open, High, Low, Volume and Close [40]. This study will use the same features but with the use of close instead of adjusted close and compare the final result and accuracy with that of adjusted close since it considers factors such as dividends in addition to the closing price and will be able to compare the proper result with that of the method used through the machine.

Diagram, schematic

Description automatically generated

Fig 6: Data Flow Diagram: Overview

As seen in Figure 6, a general prototype overview is illustrated. The first section focuses on data gathering and producing an LSTM on the gathered data. For the research compiled in 2.7 LSTM will be used as an analysis tool since it provides a better option than other tools for larger data sets. Than the model will be configured and processed into visualization the findings.

References

[1] I. Education, "What is Machine Learning?", *Ibm.com*, 2020. [Online]. Available: https://www.ibm.com/cloud/learn/machine-learning. [Accessed: 20- May- 2022].

[2] T. Anwar, "Different Branches of Machine Learning | Artificial Intelligence", *Bleed AI*, 2021. [Online]. Available: https://bleedai.com/different-branches-of-machine-learning-artificial-intelligence-part-3-4-episode-5-cvfe/#:~:text=Machine%20Learning%20can%20be%20further,Unsupervised%20Learning%2C%20and%20Reinforcement%20Learning. [Accessed: 18- May- 2022].

[3] J. WILPON, D. THOMSON, S. BANGALORE, P. HAFFNER and M. JOHNSTON, *THE FUNDAMENTALS OF MACHINE LEARNING*. Interactions, pp. 8-11.

[4] A. Lasod and R. Pawar, "Sentiment Analysis Using Machine Learning Techniques", *INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN TECHNOLOGY*, vol. 6, no. 7, pp. 153-157, 2019.

[5] F. Jemai, M. Hayouni and S. Baccar, *Sentiment Analysis Using Machine Learning Algorithms*. ResearchGate, 2021.

[6] R. Dase, D. Pawar and D. Daspute, "Methodologies for Prediction of Stock Market: An Artificial Neural Network", *International Journal of Statistika and Mathematika*, vol. 1, no. 1, pp. 08-15, 2011. [Accessed 15 October 2022].

[7] K. Reddy, "Stock Market Prediction using Machine Learning Techniques", *International Research Journal of Engineering and Technology (IRJET)*, vol. 7, no. 6, pp. 6780-6785, 2020. [Accessed 2022].

[8] Z. Iqbal, "Efficient Machine Learning Techniques for Stock Market Prediction", *Journal of Engineering Research and Applications*, vol. 3, no. 6, pp. 855.867, 2013. [Accessed 2022].

[9] R. Batra and S. Daudpota, "Integrating StockTwits with Sentiment Analysis for better Prediction of Stock Price Movement", *2018 International Conference on Computing, Mathematics and Engineering Technologies – iCoMET 2018*, 2018. [Accessed 2022].

[10] "Stocks | Investor.gov", *Investor.gov*, 2022. [Online]. Available: https://www.investor.gov/introduction-investing/investing-basics/investment-products/stocks. [Accessed: 2022].

[11] A. Hayes, "Stocks: What They Are, Main Types, How They Differ From Bonds", *Stocks: What They Are, Main Types, How They Differ From Bonds*, 2022. [Online]. Available: https://www.investopedia.com/terms/s/stock.asp. [Accessed: 2022].

[12] D. Duxbury and S. Yao, "Are investors consistent in their trading strategies? An examination of individual investor-level data", *International Review of Financial Analysis*, pp. 77-87, 2017. [Accessed 2022].

[13] G. Di Crescenzo, *Privacy for the Stock Market*. 445 South Street, Morristown, NJ, 07960: Telcordia Technologies, Inc., 2002, pp. 269-288.

[14] Z. Lanbouri and S. Achchab, "A new approach for Trading based on Long-Short Term memory Ensemble technique", *IJCSI International Journal of Computer Science Issues*, vol. 16, no. 3, pp. 1694-0814, 2019. [Accessed 2022].

[15] J. Yu, *Simple Profits from Swing Trading*, 2nd ed. John Wiley & Sons, Incorporated, 2013, pp. 1-82. [Accessed 2022].

[16] G. CESPA and X. VIVES, "The Beauty Contest and Short-Term Trading", *The Journal of Finance*, vol. 70, no. 5, pp. 2099-2154, 2015. Available: 10.1111/jofi.12279 [Accessed 2022].

[17] R. A. Kamble, "Short- and Long-Term Stock Trend Prediction using Decision Tree", *International Conference on Intelligent Computing and Control Systems ICICCS 2017*, pp. 1371-1375, 2017. [Accessed 2022].

[18] H. Stoll and R. Whaley, "Stock Market Structure and Volatility", *The Review of Financial Studies*, vol. 3, no. 1, pp. 37-71, 1990. [Accessed 2022].

[19] V. Gaucan, "Introduction to the foreign exchange market", *Journal of Knowledge Management, Economics and Information Technology*, vol. 1, no. 1, pp. 1-14, 2010. Available: https://mpra.ub.uni-muenchen.de/28078/. [Accessed 2022].

[20] M. Ismail, M. Md Noorani, M. Ismail, F. Abdul Razak and M. Alias, "Predicting next day direction of stock price movement using machine learning methods with persistent homology: Evidence from Kuala Lumpur Stock Exchange", *Applied Soft Computing*, vol. 93, p. 106422, 2020. Available: 10.1016/j.asoc.2020.106422 [Accessed 2022].

[21] L. Saes, *Unit test generation using machine learning*, 1st ed. Universiteit van Amsterdam: Faculteit der Natuurwetenschappen, Wiskunde en Informatica Master Software Engineering, 2018, p. 56.

[22] B. Roychoudhury, S. Nath Bhattacharya, and M. Bhattacharya, "Applicability of Machine Learning Tools in Foreign Exchange Market", *International Journal of Digital Content Technology and its Applications (JDCTA)*, vol. 9, no. 4, 2015. [Accessed 2022].

[23] F. KAMALOV and I. GURRIB, *MACHINE LEARNING BASED FORECASTING OF SIGNIFICANT DAILY RETURNS IN FOREIGN EXCHANGE MARKETS*, 1st ed. 2020, pp. 1-29.

[24] M. Gupta, "A Comparative Study on Supervised Machine Learning Algorithm", *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 1, pp. 1023-1028, 2022. Available: 10.22214/ijraset.2022.39980 [Accessed 2022].

[25] V. Nasteski, *An overview of the supervised machine learning methods*, 1st ed. University "St. Kliment Ohridski" - Bitola: ResearchGate, 2017.

[26] M. Usama et al., "Unsupervised Machine Learning for Networking: Techniques, Applications and Research Challenges", *IEEE Access*, vol. 7, pp. 65579-65615, 2019. Available: 10.1109/access.2019.2916648 [Accessed 2022].

[27] P. Kinderman et al., *LEARNING HOW TO EXPLAIN NEURAL NETWORKS: PATTERNNET AND PATTERNATTRIBUTION*, 1st ed. 2017.

[28] F. Bre, J. Gimenez and V. Fachinotti, "Prediction of wind pressure coefficients on building surfaces using artificial neural networks", *Energy and Buildings*, vol. 158, pp. 1429-1441, 2018. Available: 10.1016/j.enbuild.2017.11.045 [Accessed 2022].

[29] H. Larochelle, Y. Bengio, J. Louradour and P. Lamblin, "Exploring Strategies for Training Deep Neural Networks", *Journal of Machine Learning Research 1*, pp. 1-40, 2009. [Accessed 2022].

[30] R. Miikkulainen et al., "Evolving Deep Neural Networks", *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, pp. 293-312, 2019. Available: 10.1016/b978-0-12-815480-9.00015-3 [Accessed 2022].

[31] A. Yoshihara, K. Fujikawa, K. Seki and K. Uehara, "Predicting Stock Market Trends by Recurrent Deep Neural Networks", *Lecture Notes in Computer Science*, pp. 759-769, 2014. Available: 10.1007/978-3-319-13560-1\_60 [Accessed 2022].

[32] K. Laddach, R. Łangowski, T. Rutkowski and B. Puchalski, "An automatic selection of optimal recurrent neural network architecture for processes dynamics modelling purposes", *Applied Soft Computing*, vol. 116, p. 108375, 2022. Available: 10.1016/j.asoc.2021.108375 [Accessed 2022].

[33] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. Available: 10.1162/neco.1997.9.8.1735 [Accessed 2022].

[34] F. Altché and L. F. Arnaud de, "An LSTM Network for Highway Trajectory Prediction," ArXiv.Org, 2018. Available: https://ejournals.um.edu.mt/login?url=https://www-proquest-com.ejournals.um.edu.mt/working-papers/lstm-network-highway-trajectory-prediction/docview/2071280241/se-2.

[35] B. Wang, H. Huang, and X. Wang, "A novel text mining approach to financial time series forecasting", *Neurocomputing*, vol. 83, pp. 136-145, 2012. Available: 10.1016/j.neucom.2011.12.013 [Accessed 2022].

[36] K. Chowdhary, *Fundamentals of artificial intelligence*. New Delhi: Springer (India) Private Limited, 2020, pp. 603-649.

[37] L. Cheng Kuan, M. Akmar Ismail, T. M. A. Zayet, and S. Mohamed Shuhidan, “Prediction of Malaysian Stock Market Movement Using Sentiment Analysis,” *Journal of Physics: Conference Series*, vol. 1339, no. 1, p. 012017, Dec. 2019, doi: 10.1088/1742-6596/1339/1/012017.

[38] J. J. Han and H. Kim, “Prediction of Investor-Specific Trading Trends in South Korean Stock Markets Using a BiLSTM Prediction Model Based on Sentiment Analysis of Financial News Articles,” *Journal of Behavioral Finance*, vol. 4, no. 1, pp. 1–13, Oct. 2021, doi: 10.1080/15427560.2021.1995735.

[39] S. Urlam, B. Ghosh, and A. Suresh, “Stock Market Prediction Using LSTM and Sentiment Analysis,” *Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 11, pp. 4653–4658, 2021.