Neural Network Stock Return Prediction & Portfolio Optimization: A Case Study of Stocks Listed on the Nairobi Securities Exchange

Radeny, Victor Graham
119368

Submitted in partial fulfillment of the requirements for the Bachelor of Business Science in Actuarial Science at Strathmore University

Strathmore Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

August, 2022

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the Research Project contains no material previously published or written by another person except where due reference is made in the Research Project itself.

..…………….. [*Name of Candidate*]
…………….. [*Signature*]
…………….. [*Date*]

This Research Project has been submitted for examination with my approval as the Supervisor.

..…………….. [*Name of Supervisor*]
…………….. [*Signature*]
…………….. [*Date*]

Strathmore Institute of Mathematical Sciences
Strathmore University

List of Abbreviations

ADF Augmented Dickey-Fuller

ANN Artificial Neural Networks

ARIMA Autoregressive Integrated Moving Average

LSTM Long-Short Term Memory

ML Machine Learning

MSE Mean Squared Error

Multi-Portfolio A portfolio comprising of multivariate LSTM-predicted stock returns

MV Mean-Variance

NASI Nairobi All Share Index

NSE Nairobi Securities Exchange

ReLU Rectified Linear Unit

RFE Recursive Feature Elimination

RMSE Root Mean Squared Error

SCOM Safaricom

Uni-Portfolio A portfolio comprising of univariate LSTM-predicted stock returns

Abstract

Investment in assets and other financial instruments is critical for the growth of any economy. To maximize returns while minimizing risk, financial investors must accurately predict the level of expected returns prior to selecting a portfolio to invest in. This study aims to shed light on stock prediction and portfolio optimization by comparing the performance of a univariate and multivariate LSTM model in predicting closing stock returns and the subsequent portfolios constructed from each model's return predictions.

The research findings indicate that the multivariate LSTM approach outperforms the univariate LSTM model in forecasting stock returns, as evidenced by lower root mean squared error per predicted stock. Furthermore, the study examines the impact of predicted returns on portfolio performance and finds that the choice of LSTM architecture can have a significant impact on portfolio performance. Empirical analysis conducted demonstrates that the optimal portfolio selection outcomes between univariate and multivariate LSTM predicted returns are quite different. Based on the accuracy of stock return predictions, this paper recommends the multivariate LSTM for asset preselection due to its ability to better capture stock return variability.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background to Study

1.1.1 Stock Market Prediction

A stock or equity is a type of investment security that represents the ownership of a fraction of the issuing company. Units of these stocks are known as shares. The prediction of stock market values and trends has for a while been a keen problem of interest. The dynamic behavior of stock prices often driven by the law of demand and supply has attracted the attention of numerous investors who seek to gain huge profits from making the right investments at the right time. Hellström (1998) asserts that most proficient traders tend to purchase securities during periods of market optimism, colloquially referred to as a "bullish" market, and sell them during periods of market pessimism, commonly known as a "bearish" market.. These market participants try to enhance the profitability of their investments by anticipating the stock price as closely as possible. A method otherwise known as prediction. However, in the stock market, uncertainty is a prevalent characteristic due to the random-like movement and non-linear nature of the stock price. Uncertainty is undesirable for existing investors albeit impossible to eliminate. To solve this issue, investors have attempted to reduce uncertainty levels inherent in the stock market through the development and application of stock market forecasting models.

1.1.2 Asset Selection for Portfolio Optimization

Stock market forecasting and portfolio optimization often go hand in hand. A portfolio is defined as a collection of financial investments and assets. In the construction of an optimal portfolio, or in portfolio optimization, the investor must determine the best combination of assets that would provide a high level of return for a low level of risk within a given time horizon. Doing so however, especially in the stock market, can prove a challenging task due to stated uncertainty; also known as risk. In an effort reduce inherent risk, financial experts and associated researchers have dedicated decades of effort into the development of pricing models and optimization techniques that predict and measure the stock market return; such as the classical financial theories: Markowitz's legendary Modern Portfolio Theory (MPT) where a portfolio's risk is explained by correlation between individual assets, the single-factor Capital Asset Pricing Model (CAPM) where stock returns are a function of systematic risk—and the multifactor Five-Factor Asset Pricing Model developed by Fama and French (2014) where stock returns are primarily based on size of firm, book-to-market ratio, the market portfolio return less the risk-free rate of return, profitability and investment. The expected return on a particular asset is a significant factor in the portfolio optimization process, which therefore means that a careful selection of investment assets is paramount for effective portfolio management (Guerard Jr et al., 2015, as cited in Wang et al., 2019). The framework of asset selection must include a quantitative assessment of the uncertainty accompanied with obtaining the asset's expected return (Freitas et al., 2009) . In the context of the stock market, construction of an efficient portfolio is primarily dependent on the future performance of stock returns (Chen et al., 2021) and therefore, so is the development of an accurate forecasting model.

1.1.3 Deep Learning in Finance

In recent years, the exponential increase in volumes of data has encouraged the development of advanced computing systems and algorithms combined with mathematical and statistical methods to birth deep learning. Deep learning is part of the broader domains of machine learning and artificial intelligence (AI) where computing systems simulate the behavior of the human brain allowing them to learn and draw inferences from large amounts of data, also known as big data.

Deep learning has of late shown great potential in the financial world incorporating stock prediction theory into the portfolio selection and optimization process. For instance, as determined in a comparative study of deep learning methods by Gu et al. (2018), tree-based algorithms and neural network designs greatly outperform traditional regressive forecasting approaches, delivering out-of-sample equal-weighted long-short portfolios with a Sharpe ratio that doubles that of a portfolio produced from a benchmark panel regression. Additionally, Sen et al., (2021) compares Long-Short Term Memory (LSTM) predicted returns of nine mean-variance portfolios each per sector of the Indian stock market with actual returns after a hold-out period of five months; finding very close correspondence between both. The stock selection criteria with highest predictions on stock returns, is a complex task. However, LSTM networks and other advanced algorithms have shed light on prediction efficiency delivering accurate return predictions and optimal portfolios that can outperform the general market index (Ta et al., 2020). In this paper, we put forward the use of LSTM for financial forecasting.

1.1.4 Stock Market Participants

Ideally, within any market, there are two main methods market participants apply to analyze and predict the price movement of securities: technical and fundamental analysis. Technical analysts observe a security's historical price movements and trends to determine indicators of future market performance. Fundamental analysts on the other hand observe the economic and financial factors that influence the security's intrinsic price. Technical analysts are more of traders rather than investors; who purchase a share at a low price and sell at a higher price within a short time horizon for profit. Fundamental analysts can be regarded as investors. Their underlying conviction is that in the short-term, the stock price is not aligned with its intrinsic value; however, in the long run, it will correct itself. They therefore buy and hold stocks for a long period of time. The inherent nature and methodological approach applied in technical analysis make it more useful for short-term trading and market timing while fundamental analysis is better suited for long-term investments (Petrusheva et al., 2016). In this study, we shall consider the approach of the technical analyst, proposing a model framework to aid the short-term trader in stock anticipation and diversification.

1.2 Problem Statement

Stock price trading and anticipation has been a common topic among stock market participants within any economy including Kenya's. The Nairobi Securities Exchange (NSE), a renowned trading institution that facilitates the buying and selling of shares of multiple listed companies, has witnessed unprecedented growth in recent years due to the ever-increasing trading activity within the Kenyan financial market. The Central Bank of Kenya in its Monthly Economic Indicators report lists the NSE as one of the capital market indicators with a market capitalization of shareholders wealth currently at Ksh.2,006 billion and equities turnover (traded shares value) at Ksh.10,788 million. To put it into perspective, Kenya's banking sector's total balance sheet for the same period stands at Ksh.5,110 billion; meaning that the NSE's capitalization is almost half the total assets held by Kenyan banks (Central Bank of Kenya, 2022 May). Stock trading is therefore big business within the Kenyan economy with the NSE controlling a significant portion of the national capital.

Established by an Act of Parliament, Cap 485 on December 1989, the Capital Markets Authority (CMA) was entrusted with the responsibility of supervising, licensing, and regulating the operations of market participants and intermediaries, including the NSE and the Central Depository and Settlement Corporation (CDSC). A capital market refers to a financial system where buyers and sellers indulge in trading of securities such as bonds and stocks. As part of its regulatory functions, the CMA oversees the issuance of capital markets products (shares) with one of its core objectives being to facilitate the development, diversification, and uptake of capital market products for efficient market development. (Capital Markets Authority, 2019)

The CDSC, another key player within the Kenyan stock market was established by the CMA to set up a system of handling and settling securities listed on the NSE. The created system enables investors to register securities in electronic form and update the registry whenever any sale or purchase of securities takes place. Shareholders maintain a registry of their stock portfolios using a Central Depository System (CDS) account. (Central Depository & Settlement Corporation, 2022). A system though which transactions are swiftly and easily carried out and investors are made aware of all transactions taking place, increasing investor confidence. It is the functions of the CMA and CDSC to ensure a robust, streamlined, and competitive stock market infrastructure that has contributed to the increase in the number of shareholders and stock trading within the Kenyan capital market.

With higher market participation and confidence, there is creation of liquidity due to increased spending and long-term commitment of capital. The importance of liquid equity markets is that they reduce the severity of risk within investments making them more attractive leading to even more investment (Levine, 1996). In fact, investors tend to possess a greater risk tolerance when confident, allocating more of their investment in small-cap stocks that have higher risk (Meier, 2018). The increased sales and earnings of company shares enables corporations to use the newly acquired capital to fund their activities boosting national economic growth and Gross Domestic Product (GDP).

To diversify their investments, Masese (2017) states Mean-Variance (MV) optimization as the most used portfolio optimization model by shareholders for asset selection in the Kenyan equities market. The purpose of portfolio optimization is simply to maximize expected return subject to risk. These shareholders have applied various optimization techniques to solve the problems of uncertainty and unpredictability of assets. However, classical optimization methods often use historical prices and returns as model inputs which has limited influence on future stock behavior therefore obtaining inaccurate estimates on expected returns (Ma et al., 2020). Given the substantial amount of trade money involved, improper investment could potentially lead to significant losses for investors. This lack of guaranteed returns can cause reluctance among potential investors to participate in the stock market hindering economic activity and growth. A suitable prediction and optimization model is therefore necessary.

To address the above problem, this research study will create an optimal portfolio via MV optimization using returns predicted by a multivariate LSTM model with technical indicators as inputs. The proposed model will be compared to a univariate LSTM + MV model consisting of only stock returns inputs via the Sharpe ratio metric.

We implement the LSTM neural network framework owing to its high-level predictive capabilities on time series data outperforming other deep learning methods in prediction accuracy and computation time (Lara-Benitez et al., 2020).

1.3 Objectives of the study

The primary objective of this study is to compare the portfolio selection results of a portfolio consisting of LSTM-predicted returns with multivariate inputs and a portfolio consisting of predicted returns with only stock return as input.

The specific objectives are:

1. To apply a univariate and multivariate LSTM framework in obtaining stock return predictions
2. To compare the prediction performance of both neural network models
3. To optimize a portfolio of predicted stock returns using the mean-variance optimization technique.
4. To compare the adequacy of our proposed multivariate LSTM + MV with a univariate LSTM + MV model.

1.4 Significance of the Study

The informed selection and optimized allocation of assets based on risk and return factors has been a crucial yet challenging issue in the financial investment area. Application of complex portfolio optimization techniques to aid the investment decision-making process is only feasible with high-quality asset input (Wang et al., 2019). To anticipate future price, Wanjawa (2014) states that stock market participants currently rely on solely technical and fundamental analysis, or rather, non-AI methods, for stock selection which are often arbitrary and short-sighted owing to their limited ability to analyze raw data. Thanks to deep learning, the integration of sophisticated stock price prediction models in portfolio construction can greatly improve asset selection (Ma et al., 2020). The results of this study may therefore be of great significance particularly to the short-term investors of the stock market who wish to maximize returns for a lower level of portfolio-inherent risk by making better investment decisions thus ultimately reducing losses. Additionally, shareholders assured of suitable return on investment are more inclined to participate and invest in the stock market.

The capital market system relies heavily on investor confidence and market integrity. One of the primary aims of securities regulators is to encourage greater participation of investors in financial markets (Ross, 2022). This paper proposes deep learning and portfolio optimization to enhance shareholders' investment return stability increasing investor confidence and market participation ultimately contributing to national economic growth. This study is therefore conducted keeping in mind the core objective of the regulator to affirm sound market development via technological advances (Capital Markets Authority, 2019).

Moreover, as of the present moment, there exists scarce research on the utilization of deep learning techniques for stock prediction in Kenya, and to the best of my knowledge and research, no studies have yet applied such methods for portfolio optimization. This study will therefore add to the existing literature on the application of deep learning in the field of finance within the Kenyan market.

Chapter 2

Literature Review

2.1 Theoretical Review

2.1.1 Portfolio Theory

Portfolio optimization is the process of selecting the best combination of assets that would yield a maximum return for a certain level of risk. The foundations of portfolio optimization were set by Markowitz (1959) in his doctoral dissertation where he proposed the intuitive notion that portfolio selection should be based on a risk-return assessment of an investor's entire portfolio of assets, rather than individual asset analysis as was the case prior. This is the basis of a school of thought Markowitz developed known as Modern Portfolio Theory (MPT). He assumed that investors are risk-averse; or rather, in general, investors would prefer to choose a portfolio of assets with a higher 'likely return' and lower uncertainty. Any portfolio that met this criterion was named an efficient portfolio. The risk-return relationship of assets in an efficient portfolio was graphically represented by the efficient frontier. The mean-variance model was then developed which accounted for the expected return given by the weighted assets in the portfolio and the variance or standard deviation (risk) which is calculated as the correlation of a pair of assets in the portfolio. The higher the correlation between two such assets, the higher the risk.

Based on Markowitz' mean-variance optimization technique, Sharpe (1964) developed the Capital Asset Pricing Model (CAPM) which establishes that the expected return of an asset or portfolio equals to the expected return on a risk-free asset plus the market risk premium multiplied by some beta (β) value representing the risk of a particular asset. Sharpe challenges the Markowitz model by demonstrating that not all risks should affect asset returns; rather, for each individual asset in a portfolio that has specific risk, through the combination of all investments, or what is otherwise known as diversification, an investor's risk exposure can be reduced to the systematic risk of the market portfolio.

However, as time progressed and more advanced portfolio selection models were introduced, CAPM was faced with heavy criticism and was generally considered impractical on its own as it oversimplifies the complexities of financial markets to systematic risk alone.

2.1.2 Stock Market Prediction Theories

Hellström (1998) puts forward two major tenets to be considered in financial forecasting: the random walk process and the efficient market hypothesis.

2.1.2.1 The Random Walk Process

The random walk theory was first proposed by Malkiel (1973) where he described it as a process which future steps cannot be predicted based on past history. In the context of the stock market, the theory posits that changes in stock price are identically distributed and independent of each other. Therefore, past stock price movements cannot be used to forecast future stock price movement.

A random walk process is normally defined as:

$$y(t) = y(t-1) + a(t)$$

Where $y(t)$ and $y(t-1)$ is the stock price at time t and $t-1$ respectively and $a(t)$ is the adjustment or error term. $a(t)$ has zero mean with its values being independent of each other. In other words, the change $\Delta y(t) = y(t) - y(t-1)$ being $a(t)$ is independent of previous changes $\Delta y(t-1), y(t-2)$ .... and so on. $a(t)$ denotes the true impact of all privately and publicly available information on stocks which should have no influence on the changes in stock price as per the efficient market hypothesis (Malkiel, 2003).

2.1.2.2 Efficient Market Hypothesis

The origin of the efficient market hypothesis is closely related to the work of acclaimed economist Fama (1970) who described an efficient market as one whose security prices at any time fully reflect all available information. As new information enters the market, any imbalance is instantly detected and rectified through an accurate adjustment in the stock price (Hellström, 1998). This means that all investors immediately react to any form of informational advantages available in the market therefore eliminating any possible profit opportunities from any trading strategies. Therefore, from the available information, it is impossible to predict future stock price. The efficient market hypothesis places all investors within a level financial playing field where even uninformed market participants (with no knowledge of any information) purchasing a diversified portfolio of stocks can achieve a generous rate of return as that of expert investors and traders (Malkiel, 2003).

Fama (1970) mentions three forms of the efficient market hypothesis based on the varying degree of information available within a competent financial marketplace: the weak form where only historical price or past information is used, the semi-strong form which considers all publicly available information—including past information—and lastly the strong form which includes all information; both publicly and privately available.

2.1.3 Stock Market Prediction

Wanjawa (2014) states four methods of stock prediction: Technical analysis, Fundamental analysis, Time series analysis and Machine learning.

2.1.3.1 Technical Analysis

Technical analysis is the use of the technical data such as price, price movement and volume to forecast future pricing of a stock. It involves the use of charts, graphs and tables to visualize historical data for investors to analyze market movements (Petrusheva et al., 2016; Huang et al., 2011; Pring, 1991). Technical analysts use pattern-based signals known as technical indicators to identify trends which should then inform a direction of future prices. Deng et al., (2011) mentions some of the popular technical indicators including the Rate of Change (ROC), the Moving Average Convergence-Divergence (MACD) and Bias. ROC shows the difference between the current price and lagged version of the current price, MACD is an oscillator showing the momentum or speed at which the stock price changes and Bias checks the difference between the stock prices and the moving average. These indicators are calculated and their values are used to determine future trends to guide stock prediction.

2.1.3.2 Fundamental Analysis

Thomsett (1998) defines fundamental analysis as the study of financial information to forecast profits, supply, demand, management ability, industry strength and other factors influencing the stock's market value and growth. In short, it looks at financial and economic elements that affect the stock's value. These elements are referred to as fundamentals. Technical and Fundamental analysis are closely related as both can be said to examine inherent company performance. While Technical analysis looks at looks at the internal operations and financial situation of the company, Fundamental analysis concentrates on external factors such as the political, economic, and general business environment in which the company operates (Wanjawa, 2014).

2.1.3.3 Time Series Analysis

A time series is defined as a sequence of data points that occur in some successive order over a certain time period. The stock market is an example of a time series where the stock price is akin to a data point that occurs periodically, i.e.- daily, weekly, monthly. Popular time series prediction models include linear regression, Autoregression (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA) or the Box & Jenkins Model, Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) (Zhuge et al., 2017). These methods are referred to as statistical techniques. Over time, with the development of more sophisticated soft-computing-based prediction models within deep learning and machine learning, researchers have begun to move away from statistical techniques. In fact, there is a growing consensus among researchers to refer to time series approaches as 'traditional' in contrast to soft-computing techniques within the context of time series prediction (Adebiyi et al., 2014; Zhuge et al., 2017; Wen et al., 2019; Ma, 2020).

2.1.3.4 Machine Learning

Machine Learning (ML) refers to an algorithm's methods of using computing systems and software applications draw inferences from data without any explicit programming (Moghar & Hamiche, 2020). It is a subset of artificial intelligence (AI) where computer systems can perform tasks normally done by humans. Janiesch et al. (2021) list popular machine learning algorithms which consist of regression models, decision trees such as Random Forest, Bayesian methods, Support Vector Machines (SVM) and k-nearest neighbors. ML shows best applicability in tasks related to high-dimensional data such as regression, classification, and clustering.

From machine learning comes its subset known as deep learning where computer systems can determine patterns from data using structures that loosely model human brain. Deep learning models consist of the family of neural networks which apply bio-inspired algorithms to process data. They include: Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs) from which LSTM is based, Convolutional Neural Networks (CNNs) and Autoencoders (Janiesch et al., 2021).

As per the efficient market hypothesis, investors react immediately to available information. Malkiel (2003) therefore states that no form of technical and/or fundamental analysis can enable an investor to achieve a profit any greater than that obtained from holding a portfolio consisting of randomly selected assets. Additionally, because time series analysis methods for stock market prediction are increasingly becoming obsolete, this study proposes deep learning—or more specifically, LSTM—for financial forecasting.

2.2 Empirical Review

Stock price is essentially a form of time series. Various researchers have applied time series designs to forecast future stock price. The most common model applied being ARIMA which was first introduced by Box and Jenkins in 1970. Adebiyi et al., (2014) implemented an ARIMA model to forecast Nokia stock prices from the New York Stock Exchange and Zenith stock prices from the Nigeria Stock Exchange. For both sets of predictions they found suitable accuracy with some instances of very closely corresponding predicted and actual prices.

In a later study, Adebiyi et al., (2014) compared the predictive capabilities of ANNs—a soft-computing technique—and the ARIMA model—a statistical technique. Both models were implemented on daily New York Stock Exchange (NYSE) historical closing prices. In their study, despite both models exhibiting low forecast error, they found that ANN model had better forecast accuracy than the ARIMA (1,0,0) model used. On comparing the prediction results with the actual stock price data, the predictive pattern of the ARIMA process was found to be linear—revealing only the trend of the data—while ANN was able to capture the underlying fluctuations of the stock prices and even accurately predicting some of the actual values. Nevertheless, the study highlighted that the difference in performance was not statistically significant.

Ma (2020) highlights the drawbacks of ANNs for stock price prediction. Despite their broad use and applications, they are unable to measure the continuity of evolving price trends. He proposes the use of LSTM of which its feedback connection makes it easier to find development trends via backpropagation of historical prices and current price. The study compares the predictive ability of LSTM, ANN and ARIMA (1,0,0) models. Results found that there was no obvious difference between the prediction accuracy of ANN and ARIMA, whereas the LSTM model surpasses both showing better performance. Furthermore, the study highlights the ability of LSTM to distinguish between temporary price surges and long-term trend reversals via the inclusion of technical indicators as model inputs; a property that is not available in ANN and ARIMA models. Technical indicators were divided into four basic types: trading volume, trend, momentum, and volatility. On-Balance Volume (OBV) represented trading volume, MACD and Fibonacci retracements indicated trend and Relative Strength Index (RSI) characterized momentum.

For LSTM networks, multivariate inputs are extremely important for optimal stock prediction accuracy. In fact, Ma (2020) states that the addition of technical indicators is compulsory, enabling users to distinguish real price trends from market anomalies. Du et al., (2019) compares a univariate method based on only closing stock price input and a multivariate method based on low, high, open, close, and trading volume as inputs for LSTM stock price prediction. The multi-feature method was found to have lower average absolute error of 0.033 as compared to 0.155 for the single-feature method.

Wang et al. (2019) extends price prediction into portfolio optimization where they propose a mixed method consisting of LSTM stock return prediction for asset selection and

Markowitz MV optimization technique. The mixed model was applied to 21 randomly chosen stocks from the UK Stock Exchange 100 Index (FTSE). The LSTM network was applied to predict future returns of the sample stocks after which varying k stocks were chosen to build the MV model on the basis of having the highest return predictions. In their work, they compared the proposed model with Support Vector Machines (SVM), Random Forest, deep neural networks and the ARIMA model each combined with MV framework. For LSTM, the study applied a multivariate approach performing feature selection via Recursive Feature Elimination (RFE). Features selected as inputs include several lagged observations about return, return measures based on open, close, high, low prices and volume, as well as five technical indicators: RSI, Momentum, True Range, Average True Range and Parabolic SAR. Results found that LSTM has the most accurate predictions and the conglomerate LSTM + MV model with $k = 10$ selected stocks outperformed the other benchmark strategies in terms of having a higher Sharpe ratio per triennium, higher cumulative return per year and highest average of the average return to the risk per month of each triennium. After the inclusion of transaction costs, in terms of return metrics, the LSTM + MV still had better performance throughout and higher Sharpe ratio for majority of the analyzed triennia.

To determine the most efficient portfolio based on deep learning predictions, Ta et al. (2020) applied multiple optimization techniques— equal-weighted modelling (EQ), Monte Carlo simulation modelling (MCS) and MV optimization modelling (MVO)—to the proposed LSTM prediction model. The constructed portfolios were compared to other portfolio-based prediction models—Gated Recurrent Units, linear regression, and support vector regression (SVR). The results reveal that LSTM achieved a better predictive accuracy as well as exhibiting the highest expected returns and Sharpe ratio on each of the LSTM-based constructed portfolios. Portfolio weights from the LSTM + MVO model were the most optimal as the LSTM + MCS model required more computational resources with each simulated scenario (although it should be noted that it had very similar portfolio weights with LSTM + MVO) and the LSTM + EQ model obtained a considerably lower Sharpe ratio.

2.3 Research Gap

We have seen that LSTM is superior to other deep learning techniques in both financial time series forecasting and in combination with portfolio optimization methodologies. Research on Kenyan stock prediction via deep learning is very limited with some existing researches applying ANNs (Wanjawa, 2014; Kihoro & Okango, 2014) and the ARIMA model (Ochieng', 2016) for forecasting which are not very efficient for financial time series. Also from the existing Kenyan literature, there is no research on deep learning methods for portfolio optimization. This study is conducted to fill these gaps proposing a multi-feature LSTM + MV optimization model for asset selection.

Chapter 3

Methodology

3.1 Research Design

This study will adopt a quantitative research design as it is focused on the collection, analysis, and prediction of stock data for the construction of an optimal portfolio. An exploratory design will as well be adopted since the study seeks to assess the adequacy of our proposed model to a univariate LSTM and mean-variance model.

3.2 Population and Sampling

The population of this study will be based on the 63 listed securities on the Nairobi Securities Exchange (NSE). The securities selected must contain full information on the prices and exhibit superior excess return to risk when compared to the risk-free asset rate. The risk-free rate used will be in accordance with the monthly interest rate on the 91-day Treasury Bills issued by the Central Bank of Kenya.

The data used will be the daily historical prices of the stocks listed on the NSE All Share Index (NASI). It is from this data that we will predict future stock returns for implementation to the MV optimization model. An optimal portfolio will be constructed from predicted returns of chosen assets out of the obtained sample.

The interest rate on the 91-day Treasury Bills issued by the Central Bank of Kenya is used as a proxy of the risk-free rate.

3.3 Data Source & Description

Secondary data on the daily historical price lists of the stocks trading on the NSE will be obtained from the NSE and the Treasury Bill interest rate will be obtained from the Central Bank of Kenya website.

3.4 Neural Networks: LSTM Structure

Artificial Neural Networks (ANN) are computational models that are based on mathematical representations of interconnected units or nodes that mimic the synapses and neurons of the biological brain. There exist various architectural variations of the ANN model however this study will focus on the most used model network adapted specifically for time series or sequential data: Long-Short Term Memory (LSTM).

Just like any other neural network design LSTM networks are composed of node layers: an input layer, several hidden layers and an output layer each with its own number of nodes or units (neurons) as is the case with majority of other neural network models. Their distinguishing feature is the memory cells contained within the hidden layers; responsible for retention of previous data in the sequence and removal (forgetting) of unimportant data.

Each memory cell contains three gates: input gate, output gate and the forget gate. Each gate is tasked with regulating information via a non-linear sigmoid and *tanh* activation function. Information within the memory cells is contained in the cell state C_t which is the long-term memory of the network and hidden state h_t containing input data from the previous memory cell and output data to be conveyed to the next cell.

The forget gate determines the data that should be discarded from the cell state (the network's long-term memory). The hidden state from the previous time step h_{t-1} and the current input data x_t are fed into the cell state via the forget gate. Through sigmoid activation, the forget gate generates a vector of weights w with values between 0 and 1 where values closer to one are deemed relevant and values closer to zero are removed.

These weight values are pointwise multiplied—multiplication in terms of relevance—with the cell state. The equation representing this process is given as:

$$f_t = sigmoid\ (w_f[h_{t-1}, x_t] + b_f)$$

The input gate comes next which is responsible for deciding which data can be added to the cell state. The inputs and process are the same as for the forget gate except that it applies a $tanh$ activation function. $tanh$ values lie within the range [-1,1]. The possibility of negative values enables the network to reduce the impact of a component within the cell state. A new memory update vector of weights is generated, informing the model on how to update the cell state as shown in the following equation:

$$\hat{C}_t = tanh\ (w_c[h_{t-1}, x_t] + b_c)$$

Still within the input gate, a sigmoid activation function is applied acting as a filter (F) to identify the components of the 'new memory network' that are worth retaining.

$$F_t = sigmoid\ (w_F[h_{t-1}, x_t] + b_F)$$

After the above processes our new cell state should be represented by the following equation:

$$C_t = f_t * C_{t-1} + F_t * \hat{C}_t$$

We then have the output gate which decides on the data from the memory cell that can be outputted to the next memory cell. This process is particularly important otherwise the network would output literally everything it has ever learned from the data. The sigmoid activation function is used to filter the output. This is given by the equation:

$$O_t = sigmoid(w_o[h_{t-1}, x_t] + b_o)$$

The cell state C_t is passed through the $tanh$ activation function to scale the values into the interval [-1,1].

Lastly, the scaled cell state is multiplied by the filtered output to obtain the new hidden state h_t that is then passed on to the next memory cell. This is given by the equation:

$$h_t = O_t * tanh$$

The bias b is added to each of the products to enhance the flexibility of the model in fitting the data.

LSTM network architectures are built from a class of neural networks known as Recurrent Neural Networks (RNNs). However, they are much more powerful such that they are capable of learning long-term dependencies. Normally in RNNs, as more layers using certain activation functions are added, small weights generated by the activation function are multiplied recurrently through several time steps decreasing the gradients of the loss function asymptotically to zero. A decreasing gradient means that the change in output given the change in input will be much smaller and up to the point when there is no change at which model is not learning anymore. This is known as the vanishing gradient problem, one of the major downsides to RNNs. LSTMs solve this problem via the use of the forget gate, enabling the network to discard smaller weighted elements from its memory cell therefore making it less likely that the gradient will vanish after iterative multiplication.

3.5 Building the Model

Building the LSTM model will be implemented using TensorFlow Keras in Python. An open-source deep learning library.

3.5.1 Input Variable Selection

As was done by past studies (Wang et al., 2019; Zelenka, 2019) the LSTM model shall be applied to forecast historical return. Therefore, the first set of inputs to be implemented are the historical returns $R_t, R_{t-1}, R_{t-2}, \ldots.$ based on the closing stock price. The simple return R_t will be computed as:

$$R_t = \frac{P_t}{P_{t-1}}$$

This study will apply a multivariate input approach for the LSTM network. First the return measures based on open, close, high, low prices and volume shall be added as inputs to the LSTM model. Previous studies have used several lagged variables of return as inputs to predict future stock return as well as technical indicators (Wang et al., 2019; Paiva et al., 2019). This paper will apply the same approach. Furthermore, we shall make feature selection on the lagged return variables and technical indicators by recursive feature elimination (RFE). To be precise, RFE works by recursively removing unwanted features and building a model on the most significant features that remain. It uses the model accuracy to identify features that contribute the most to predicting the target feature (the next period's return R_{t+1}). RFE will be applied via the Random Forest classification algorithm owing to its high feature selection accuracy compared to other classification methods (Chen et al., 2020). Therefore, our final dataset should consist of historical data as well as the selected lagged return and technical indicator variables which shall be fed as input to the LSTM model to forecast future return.

3.5.2 Model Topology

The data will then be separated into training data which will be used to construct the model and test data for accuracy evaluation. The training and test data will then be converted into a three-dimensional NumPy array as it is the required format the LSTM model can work with. An 80/20% respective train-test ratio.

The LSTM network architecture is then constructed where relevant layers imported from Keras will be added to the model; main ones being LSTM which is responsible for the model's 'memory', Dropout which discards irrelevant inputs during the training to prevent overfitting and Dense which is the output layer.

The model will then be compiled using the Adaptive moment estimation (Adam) optimizer as recommended by previous studies (Kingma & Ba, 2014; Wang et al., 2019; Zelenka, 2019; Kirci & Baydogmus, 2022). The optimizer then trains the model to ultimately reduce a specified loss or error with each epoch. The number of epochs defines how many times the algorithm will process and train the data, therefore greater accuracy is expected with more epochs (Moghar & Hamiche, 2020).

Deciding the specific hyperparameters for the model is the most important part in building the network architecture. A random search method will be performed to dynamically establish the optimal combination of hyperparameters (Wang et al., 2019). The random search method has been established to produce more efficient models than a grid search and manual search (Bergstra & Bengio, 2012). The model with the highest accuracy shall be applied in prediction. The random search samples the following hyperparameters: number of layers, number of hidden units (neurons) per layer, the activation function, learning rate, the decay rate, the number of epochs, and batch size. The loss function applied during training will be the mean squared error (MSE) which has been found to produce the most accurate predictions in conjunction with the Adam optimizer (Kirci & Baydogmus, 2022).

The MSE is represented by the following equation:

$$MSE = \frac{1}{N} \sum_{t=1}^{N} \left(R_t - \hat{R}_t\right)^2$$

Where N is the number of training pairs, R_t and $\hat{R}_t$ represents the actual and predicted returns at time t respectively.

3.5.3 Prediction Performance

Finally, as a performance metric for the proposed LSTM model this study shall apply the root mean squared error (RMSE) which has been recommended for use when evaluating time series data (Shcherbakov et al., 2013).

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(P_t - \hat{P}_t\right)^2}$$

3.6 Test for Financial Time Series

Prior to prediction, this study will apply stationary checks to the stock return data. For any time-series, stationarity is a fundamental prerequisite. Stationarity is where the statistical properties of a time-series process (mean, variance, covariance) remain unchanged with time. To clarify, it is not the values of the times series that does not change, but it is the overall behavior of the data that remains constant without any significant trend. For statistical time-series techniques, without stationarity, the analysis and interpretation of a time series is quite problematic. The statistical properties of the data or *how* the data changes are ever altering and volatile making the process unpredictable. It may be easier to think of stationarity in terms of a target. Forecasting a non-stationary time series can be akin to trying to aim at a constantly moving target which is difficult while the converse applies to stationary data. For LSTM networks, while it is not a strict requirement for the data to be stationary owing the model's robust architecture (Preeti et al., 2019) using stationary time-series data can certainly improve prediction performance (Livieris et al., 2021).

3.6.1 The Augmented Dickey-Fuller (ADF) Test

Numerous studies have applied various statistical checks for stationarity within a financial time series prior to prediction with the most common method being the Augmented Dickey-Fuller test (ADF) (Qian & Chen., 2019; Wu et al., 2018; Manurung et al., 2018). The ADF test is a unit root test for stationarity. A unit root is a characteristic of a time series that makes it non-stationary. For a simple discrete time-stochastic process $(Y_t, t = 1)$, consider an autoregressive process of order 1:

$$Y_t = \alpha_1 Y_{t-1} + e_t$$

Where Y_t is the value of a time series at time t and e_t is a serially uncorrelated and stationary error term with zero mean and constant variance σ^2. For convenience we will assume $Y_0 = 0$. The above equation will have a unit root when $\alpha = 1$ which implies that the moments of the stochastic time series process depend on t. When the moments of a process depend on t so will its statistical properties (mean and variance) which makes the process non-stationary.

3.7 The Mean-Variance Optimization Model

Using the return prediction output, we shall compute our expected return and variance which will be used as inputs to the MV optimizer oriented to maximize the Sharpe ratio.

The optimization procedure works by allocating weights

$$w_1, w_2, \dots, w_k$$

to a portfolio consisting of k risky assets with the aim of minimizing volatility (variance) for a given level of return subject to certain constraints as shown in the equation below:

The objective of MV optimization is:

$$\boldsymbol{Min\ w' \sum w}$$

$$\text{s.t. } \boldsymbol{w'\,\bar{R} \geq R,}$$

$$\boldsymbol{w'1 = 1,}$$

$$\boldsymbol{w_i \geq 0}$$

Where $\bar{R}$ is the return on an asset, R is the minimal rate of return required by an investor and Σ is the covariance matrix representing risk. The argument $w'1 = 1$ is to ensure the assigned weights add up to one and $w_i \geq 0$ means no short selling is allowed.

3.7.1 Expected Return & Risk of a Portfolio

The widely used return formula is:

$$\bar{R} = \frac{1}{N} \sum_{t=1}^{N} R'_t$$

Where $\bar{R}$ is the mean expected return of an asset, N is the number of periods and R'_t is the realized return per period.

The realized excess returns per period will be computed as:

$$R'_t = \log R_t$$

Where R_t represents the predicted returns from the LSTM model. We apply log returns due to their wide use in financial theory and also due to their applicability in statistical concepts such as volatility and correlation computations (Estrada, 2011) as well to ensure the stability of returns from different stocks.

Given any set of k assets each with assigned weightage that describe how the investment portfolio is split, the general formula of expected return from a portfolio will be:

$$R = E(R_p) = \sum_{i=1}^{k} w_i\, E(\bar{R}_i)$$

Subject to the constraint:

$$\sum_{i=1}^{k} w_i = 1$$

Where:

- k represents the number of stocks
- w_i represents the proportion of weights invested in security i
- $E(\bar{R}_i)$ is the mean expected return per stock.
- $R = E(R_p)$ represents the expected return of the portfolio

Markowitz defined the variance as the risk measure in a portfolio which will be computed as:

$$V = \sigma_p^2 = \sum_{i,j=1}^{k} w_i w_j P_{ij}$$

P_{ij} represents the covariance of stock returns i and j which is defined as:

$$P_{ij} = \frac{1}{N}\sum_{t=1}^{N}(R_{it} - \bar{R}_i)(R_{jt} - \bar{R}_j)$$

Where R_{it} and R_{jt} represent the realized return of stocks i and j respectively at time t, $\bar{R}_i$ and $\bar{R}_j$ are the mean expected returns of stocks i and j and N is the number of observed past times.

The covariance measures the relationship between the direction of movement of two asset prices. It is a crucial factor in obtaining a portfolio with overall risk V that is lower than the risk its individual assets. In short, lower correlation between assets means lower portfolio risk (Markowitz, 1959).

With the fundamental measures defined, the goal of the MV optimization model is to minimize the portfolio risk V given a desired portfolio return R subject to:

1.

$$\sum_{i=1}^{k} w_i = 1$$

2.

$$R = E(R_p) = \sum_{i=1}^{k} w_i E(R_i)$$

and $w_i > 0$ for $i = 1,2, \dots , n$

3.7.2 Portfolio Performance Measurement

The Sharpe ratio of a portfolio is defined as:

$$Sharpe\ Ratio = \frac{E(R_p - R_f)}{\sigma_p}$$

R_p is the expected portfolio return, R_f is the risk-free rate and σ_p is the portfolio's standard deviation; also known as the portfolio risk.

The goal of the MV optimization model will be to maximize the Sharpe ratio. From a portfolio consisting of k number of assets, a suitable Python algorithm will first generate k random asset weights after which 50000 portfolios will be constructed each with varying asset weights which from a statistical standpoint, should cover all possible weightage scenarios (Wang et al., 2019; Chen et al., 2021). All 50000 portfolios will be assessed via grid search with regards to MV optimization to find the portfolio with the most optimal combination of weights such that the Sharpe ratio is maximized.

The selected asset weights and Sharpe ratio of the proposed multivariate LSTM + MV model will be compared to a univariate LSTM + MV model consisting of only stock returns as inputs.

Chapter 4

Data Analysis & Discussion

4.1 Data Overview

The data that was used for analysis in this study was the daily stock prices of the listed companies in the NSE from the period 1 January 2013 to 1 November 2022. Data for companies that had been delisted or ceased to trade within this period was considered incomplete and therefore not collected. For the remaining 55 out of 63 listed companies, the data was checked to ensure there were no null values and had available daily information for all days. Data for companies that did not have full information on the prices for all trading days within the sample period and/or had data inconsistencies was filtered out from the collected sample leaving daily stock prices for 40 listed companies that proceeded to the analysis stage.

In this study, the almost 10-year historical price data for 31 stocks are part of the experimental data set for analysis. Related studies have been conducted by selecting at least 20 stocks for prediction prior to portfolio construction. Wang et al. (2019) randomly selects 21 stocks from the Financial Times Stock Exchange 100 Index (FTSE 100) as sample data for prediction, which was considered sufficiently large for asset preselection before forming an optimal portfolio for individual investors. Furthermore, Chen et al. (2021) randomly selects 24 candidate stocks from the Shanghai Stock Exchange 50 Index (SSE50). Chaweewanchon & Chaysiri (2022) as well choose 25 stocks at random from the Stock Exchange of Thailand 50 (SET50). In the same vein, this study randomly selects 25 stocks as sample data for LSTM prediction. The names of the selected sample stocks are: "ABSA Bank Kenya Plc." (ABSA), "Bamburi Cement Plc" (BAMB), "CIC Insurance Group" (CIC), "Crown Paints Kenya Plc" (CRWN), "E.A.Cables Plc" (CABL), "East African Breweries Ltd" (EABL), "Equity Group Holdings" (EQTY), "Flame Tree Group Holdings Ltd" (FTGH), "HF Group Ltd" (HFCK), "Jubilee Holdings Ltd" (JUB), "KCB

Group Ltd" (KCB), "KenGen Ltd" (KEGN), "Kenya Power & Lighting Co. Ltd" (KPLC), "Liberty Kenya Holdings Ltd" (LBTY), "National Bank of Kenya Ltd." (NBK), "Nairobi Securities Exchange" (NSE), "NCBA Group Plc" (NCBA), "Nation Media Group" (NMG), "Safaricom Plc" (SCOM), "Sameer Africa Plc" (SMER), "Sasini Plc" (SASN), "Stanbic Holdings Plc." (SBIC), "The Co-operative Bank of Kenya Ltd" (COOP), "Total Kenya Ltd" (TOTL), and "Trans-Century Ltd" (TCL).

Table 4.1 shows the descriptive statistics of the close prices for the 25 selected stocks' closing prices. As seen, JUB has both the highest mean (354.3316) and standard deviation (98.468644). While FTGH has the lowest mean (3.509423) and SMER has the lowest standard deviation (1.494497).

Table 4. 1: Descriptive statistics for sample stocks

	Mean	Standard Deviation	Minimum	Maximum
ABSA	12.29544	3.006155	7.15	7.15
BAMB	126.6754	65.345291	19.9	225
CABL	7.05761	5.715002	0.81	17.65
CIC	4.507321	2.273836	1.87	11.9
COOP	13.83299	2.346903	8.12	19.95
DTK	138.5787	57.399297	47.75	250.00
EABL	233.2006	59.301747	110	423.00
EQTY	13.83299	2.346903	8.12	19.95
FTGH	3.590423	2.223255	0.78	10.82
HFCK	13.35536	10.541584	2.86	41.73
JUB	354.3316	98.468644	144.63	535.00
KCB	42.78013	7.685632	23.00	65.00
KEGN	7.501685	3.073584	3.36	16.83
KPLC	8.128452	5.744012	1.24	20.00
LBTY	12.96203	5.093243	5.00	27.75
NCBA	32.39007	9.360971	16.74	65.68
NMG	114.6638	95.644855	9.04	307.27
NSE	13.36052	4.343966	6.58	23.75
SASN	18.48846	3.744064	10.9	31.25
SBIC	90.4186	19.077892	41.25	148.00
SCOM	23.04076	9.63313	5.1	44.95
SMER	3.818224	1.494497	1.72	9.1.0
TCL	9.856147	10.154617	1.00	36.75
TOTL	23.51898	4.421009	13.1	35.75
TPSE	27.22033	11.320635	11.5	55.5.0

4.2 Model Analysis

For both univariate and multivariate predictions, the selected stocks' daily close prices were converted to daily returns via the following equation:

$$R_t = \frac{P_t}{P_{t-1}}$$

LSTM was then used to predict the closing stock returns for the next 30 days which would then be utilized as inputs in Markowitz' mean-variance optimization model to construct a suitable portfolio of stocks.

4.2.1 ADF Stationarity Tests

The ADF test checks for any statistical evidence of a unit root. The null hypothesis confirms the existence of a unit root while the alternate hypothesis confirms that the time-series is stationary (no unit root is present) at a specified level of significance. If the ADF statistic is lesser than the critical value, the null hypothesis is rejected. Conversely, if the ADF statistic is greater than the critical value, we fail to reject the null hypothesis. The ADF test was applied test to all 25 companies' stock returns to confirm stationarity of the data at the 1%, 5% and 10% confidence level. Stationarity was found in closing stock returns for all 25 companies. Taking SCOM (Safaricom) stock return data as an example:

Data: SCOM daily historical closing returns

ADF Statistic: -16.379, p-value= 2.7602

Critical Values:

1%: -3.433028

5%: -2.862723

10%: -2.567400

Reject null hypothesis for alternative hypothesis: time-series is stationary.

4.2.2 Univariate LSTM Return Prediction

For univariate LSTM predictions, the stock returns computed from daily historical closing prices for the study period were solely used as inputs into the neural network. The input closing returns were first normalized within the range (0,1) and the data was split into 80% for training and 20% for testing.

To select the most optimal combination of the LSTM model's hyperparameters for prediction, a random search method via Keras tuner was enforced. The random search sampled the following hyperparameters: (1) The number of LSTM layers between two or three; we found that including too many layers caused overfitting which diminishes prediction accuracy, (2) The number of units or neurons per LSTM layer ranging from 10 to 200 units, (3) The activation function applied on the output dense layer between sigmoid and Rectified Linear Unit (ReLU).

Finally, the specified topology of the LSTM network is confirmed. Three LSTM layers were set each proceeded by a Dropout layer. The model performance was seen to increase with a decrease in the Dropout rate, hence, a relatively low dropout rate of 0.1 was set. For the first, second and third LSTM layers we set 150, 70, and 60 units respectively. For the output Dense layer, we set 1 unit and a ReLU activation function. The optimal time steps length of 72 was set as recommended by Wang et al. (2019), a batch size of 32 and 100 epochs were set which should be a substantial number of times the model will process the data to improve prediction performance (Moghar & Hamiche., 2020). Lastly, mean squared error was applied as a loss function for training. With each recursive epoch, the mean squared error should decrease.

To monitor the training process, callbacks were implemented. In machine learning, a callback is an object that can perform actions at various stages of training, for instance at the start or end of an epoch. The study applied an "early stopping" callback which would halt the training process if there was no significant change in loss after each epoch. The study also applied a "reduce learning rate on plateau" callback which would half the model's learning rate during training when the rate of decrease of loss per epoch reduces. Both callbacks were set to act 10 epochs after the loss behavior is observed.

4.2.3 Multivariate LSTM Return Prediction

4.2.3.1 Feature Generation and Recursive Feature Elimination (RFE)

For multivariate LSTM predictions, various features were used as inputs into the model to forecast future returns. The inclusion of multiple features into an LSTM neural network can significantly improve prediction performance. This paper generates multiple technical indicators and return calculations as recommended by various studies (Wang et al., 2019; Ma, 2020; Paiva et al., 2019; Oriani & Coelho, 2016; Tanaka-Yamawaki & Tokuoka, 2007). Furthermore, the use of lagged versions of the variables can allow varying amounts of recent history to be brought into the forecasting model therefore improving model results (Bouktif et al., 2018). Therefore, in this light, this study applied various combinations of lagged versions of returns lagged to a maximum of 40 days. In total, the study sampled 167 features from which 25 were selected using RFE via a Random Forest regressor of 500 decision trees. The selected input variables are summarized in Table 4.2 below:

Table 4. 2: Selected input features summary

Attribute	Details	Attribute	Details
1.	$R_1 = \frac{Close\ Price_t}{Close\ Price_{t-1}}$	15.	$R_{15} = \left(\frac{Low\ Price_t}{Low\ Price_{t-1}}\right)_{t-40}$
2.	$R_2 = (R_1)_{t-1}$	16.	$R_{16} = \left(\frac{High\ Price_t}{High\ Price_{t-1}}\right)_{t-40}$
3.	$R_3 = (R_1)_{t-6}$	17.	$R_{17} = \left(\frac{Open\ Price_t}{Open\ Price_{t-1}}\right)_{t-40}$
4.	$R_4 = (R_1)_{t-30}$	18.	Triple Exponential Moving Average (Medium). Length= 75
5.	$R_5 = (R_1)_{t-40}$	19.	Kaufman Adaptive Moving Average (Fast). Length= 21
6.	$R_6 = \frac{Low\ Price_t}{Open\ Price_t}$	20.	Relative Strength Index. Period= 14
7.	$R_7 = \frac{High\ Price_{t-30}}{Open\ Price_{t-30}}$	21.	Average True Range (high, low and close price)
8.	$R_8 = \frac{Low\ Price_{t-30}}{Open\ Price_{t-30}}$	22.	Bollinger Bands. (Lower band)
9.	$R_9 = \frac{High\ Price_{t-30}}{Close\ Price_{t-30}}$	23.	Bollinger Bands (Bandwidth)
10.	$R_{10} = \frac{Low\ Price_{t-30}}{Close\ Price_{t-30}}$	24.	Bollinger Bands (Percentage)
11.	$R_{11} = \frac{Low\ Price_{t-35}}{Close\ Price_{t-35}}$	25.	Moving Average Convergence Divergence
12.	$R_{12} = \frac{Low\ Price_{t-39}}{Close\ Price_{t-40}}$		
13.	$R_{13} = \left(\frac{Low\ Price_t}{Low\ Price_{t-1}}\right)_{t-20}$		
14.	$R_{14} = \left(\frac{Low\ Price_t}{Low\ Price_{t-1}}\right)_{t-30}$		

A brief explanation of each indicator feature is as following:

1. *Simple Return*

This is represented as $R_t = \frac{P_t}{P_{t-1}}$ where P_t is the price of a stock at time t and P_{t-1} is the price of a stock at time $t-1$.

2. *Triple Exponential Moving Average (TEMA)*

This is a trend-type indicator that which applies multiple exponential moving average (EMA) calculations to subtract out any lag and create a trend following indicator that reacts rapidly to price changes. An EMA is a moving average, which places greater weight on more recent data points. The medium-term TEMA of 75 days was chosen amongst the short-term TEMA of 21 days and a short-term TEMA of 150 days.

3. *Kaufman Adaptive Moving Average (KAMA)*

This trend indicator reflects volatility. When volatility is low, the indicator will remain near the current market price and conversely when volatility is high, it will lag behind the market price. A fast moving KAMA was selected by the RandomForest regressor amongst the medium and long-term KAMA of 75 and 150 days respectively.

4. *Relative Strength Index (RSI)*

This is a momentum indicator that measures the speed and magnitude of a security's price change. It is very effective in assessing the overvalued or undervalued condition of an asset. According to the parameters of research of Wang et al. (2019), this paper set the period as 14.

5. *Average True Range (ATR)*

While True Range (TR) measures the maximum change in the price of the day compared to the previous day, ATR is the average of TR over a specified number of days. It reflects market volatility by decomposing the entire range of the price of an asset for a period.

6. *Bollinger Bands*

Bollinger bands are trend-based indicators that are defined by a set of standard deviations plotted above (upper band) and below (lower band) the Simple Moving Average of a security’s price. The standard deviation represents the stock’s volatility where a tight range indicates low volatility while a large range indicates high volatility. The percentage indicator quantifies the security’s price relative to the upper and lower bands or in other words, how close the stock's current price is to the bands.

4.2.3.2 Multivariate LSTM Architecture

The input dataset consisted of 25 sets of features as columns and a few-thousand days of daily feature data as rows. It may be easier to think of these 25 features as independent variables and the daily closing stock return as the dependent or target variable part of a multi-regression econometrical model. As was done before for the univariate LSTM, the data was split into a testing and training set of 80% and 20% respectively then normalized within the range (0, 1). A random search was performed on the same hyperparameters and within the same ranges. The model topology determined is as follows. Three LSTM layers were used each with 120, 170 and 50 units for the first second and third layer respectively. Each LSTM layer was followed by a Dropout layer with 0.1 set as the dropout rate. The output Dense layer had a single unit and a ReLU activation function. Just as before, 72-time steps, 100 epochs, a batch size of 32 and a mean squared error loss function was applied. The same callbacks were utilized to monitor the training process.

4.3 Findings

A univariate and multivariate LSTM neural network was applied to forecast stock return for each the chosen 25 companies. For the number of days of future predictions, this study will forecast returns for 30 days which is consistent with the methodology applied by Banik et al. (2022) who employs an LSTM-enforced Decision Support System to assist traders in accurately analyzing and forecasting future stock values.

In the discussion of each model's prediction results we will focus on SCOM (Safaricom) returns data.

4.3.1 Univariate LSTM Return Predictions

The below figure 4.1 shows the historical closing stock return data for SCOM over the 10- year period:

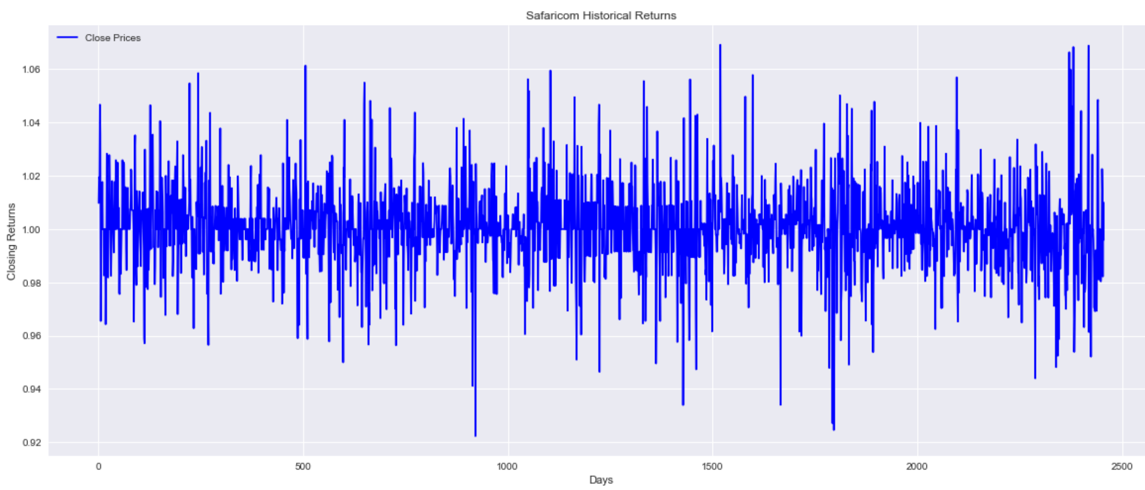

Figure 4. 1 SCOM historical returns plot

The above plot shows that the return data is highly volatile. However, it satisfies both stationarity and volatility clustering requirements where large shocks (changes) tend to be followed by large shocks and small shocks tend to be followed by small shocks. Upon training and prediction via a univariate LSTM we see the following prediction results:

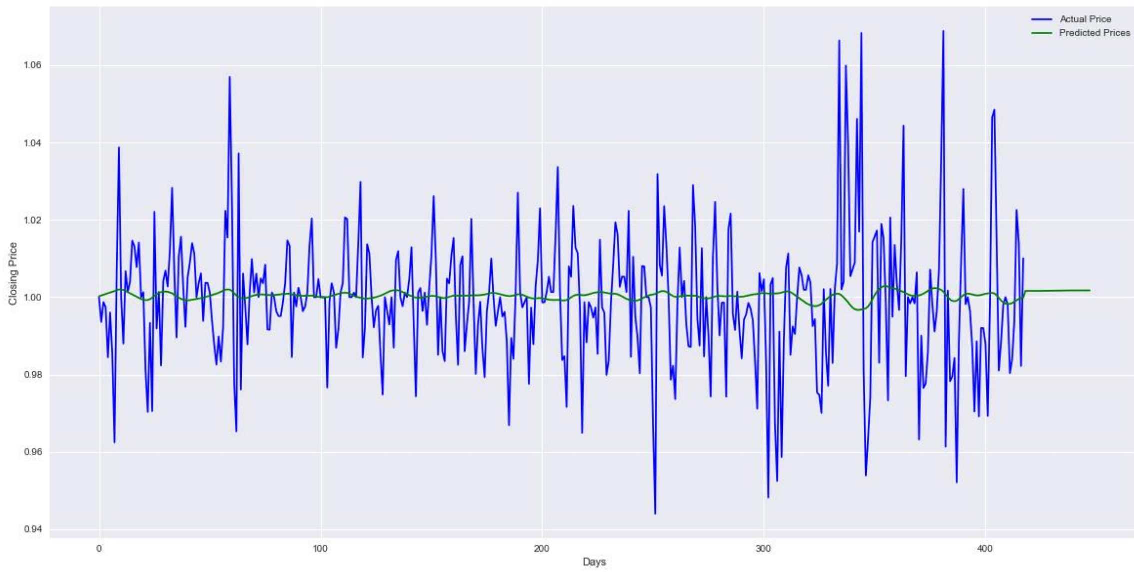

Figure 4. 2 Forecast of SCOM closing returns using a univariate LSTM

The model's predictions are very close to the mean with some points where the model attempts to capture the volatility clusters of the returns. Looking at the 30 day forecasted returns, the model seems to have maintained the predicted values at the mean. The root mean squared error for the above predictions is 0.0166. Hansson (2017) discovered similar results with an LSTM model upon forecasting return data for various benchmark indices. In his study, he found the LSTM model's prediction output on the returns was better than traditional time series models that could not at all capture the variability of the data. However, despite this, the above results are wanting.

While LSTM may have the ability to capture the volatility clusters of the data, it cannot be assumed that better capturing of the variance necessarily leads to better forecast results. In fact, Hansson (2017) in his study, states that such predictions from a financial perspective are of little use, only useful when analyzing market directions rather than actual stock values via a binary classification method. From the above Figure 4.2, the simple straight line at the end of the graph represents trend which implies that in the upcoming 30 days, the stock return will remain constant. Kuber et al. (2022) found similar results when predicting future stock price finding a downward trend line that represented future price predictions. As we cannot get any clear idea on the actual stock return for the upcoming 30 days, this paper recommends a multivariate approach for forecasting actual stock behavior.

4.3.2 Multivariate LSTM Return Predictions

This study proposes the use of a multivariate LSTM model for return prediction and asset preselection. By visually investigating the model's predictions, we see some interesting results:

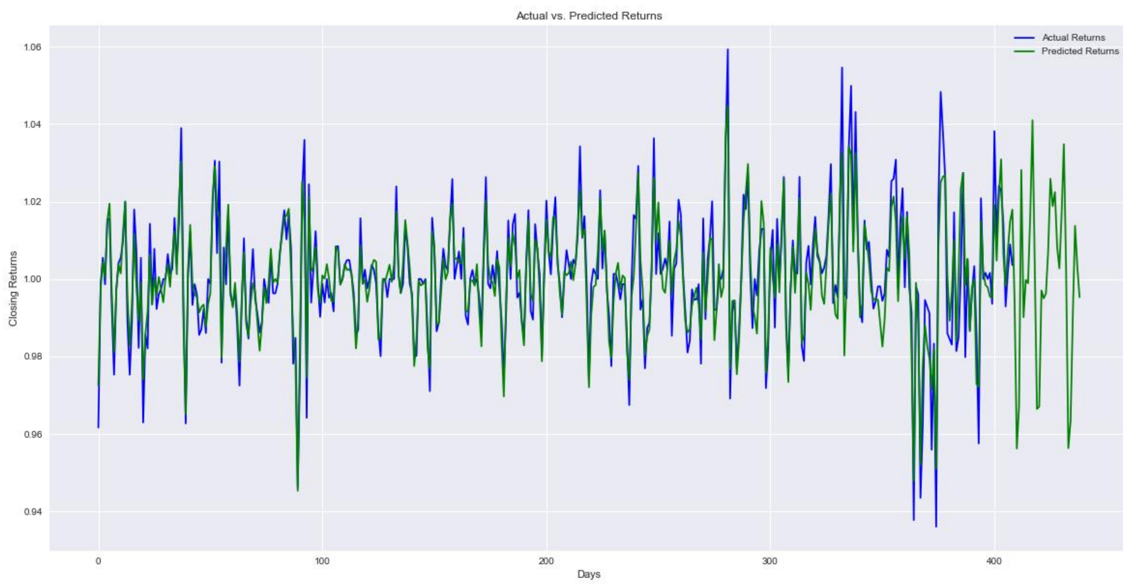

Figure 4. 3: Forecast of SCOM closing returns using a multivariate LSTM

The proposed model has achieved an outstanding level of accuracy in capturing the movement of returns and volatility clusters. The level of correspondence between the predicted and actual returns is notably high, resulting in a significantly lower root mean squared error of 0.0066. Moreover, the 30-day return predictions of the proposed model have exhibited consistency with the historical stock returns behavior, a feature that was not observed in the univariate LSTM predictions. The above figure 4.3 illustrates the actual stock return values, providing a precise picture of the future market behavior of assets, as opposed to a simplified trendline.

4.3.3 Model Performance Comparison

4.3.3.1 Model Training & Loss Function

For both models, the Adam optimization function was implemented in conjunction with a mean squared error loss function during training. The Adam optimizer is tasked with training the model, reducing the loss function to a minimum with each epoch. 100 epochs were set for each model. Despite both models having the same training architecture, they each had some differences during the actual training. Figure 4 & 5 illustrates the loss function against the number of epochs for the univariate and multivariate LSTM respectively.

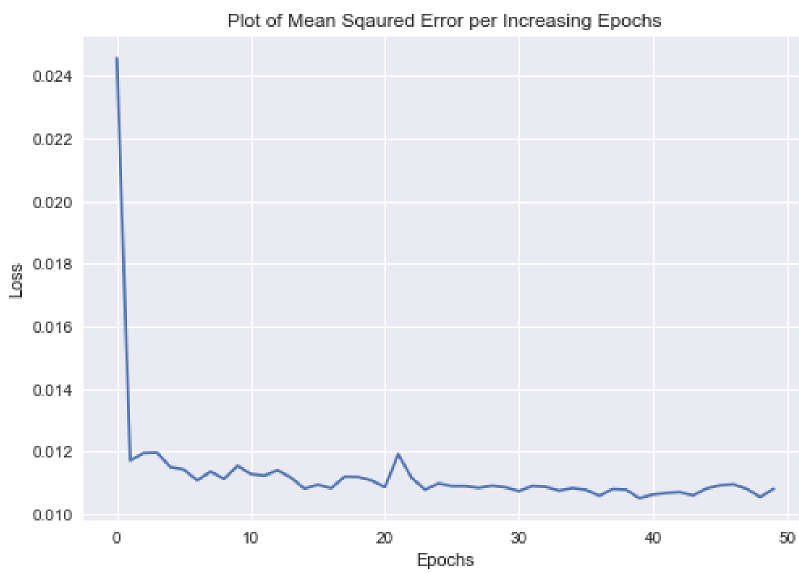

Figure 4. 4: Univariate LSTM mean squared error loss function for SCOM data

At the first epoch there is a sharp decrease of the mean squared error from 0.0246 to around 0.0117. As the number of training epochs increased, the loss function exhibited a gradual decrease, starting at 0.0117 during the second epoch and eventually reaching 0.0108 by the fiftieth epoch. At this point, the early stopping callback was triggered, leading to the termination of model training. With the muti-input LSTM model we can see different results:

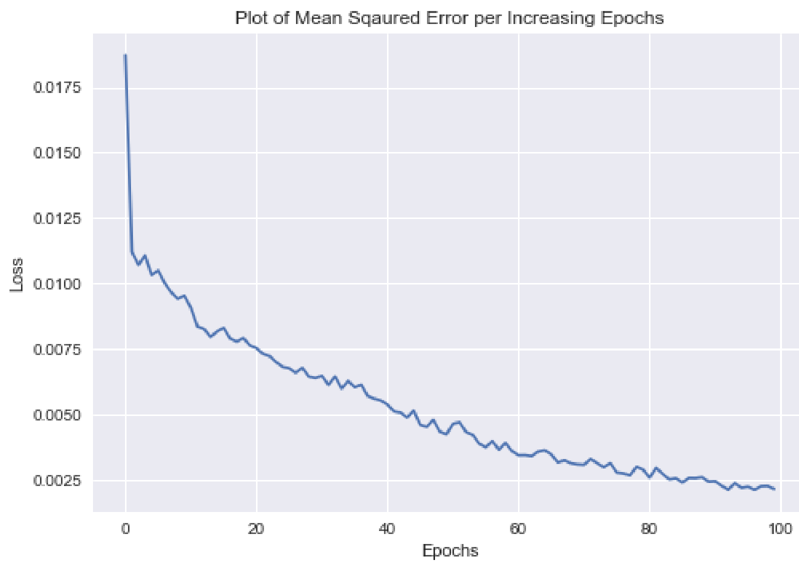

Figure 4. 5: Multivariate LSTM mean squared error loss function for SCOM data

Here we see the robustness of the multivariate LSTM in learning new data. The findings show that even at the first epoch, the multivariate LSTM model already outperforms its univariate counterpart in terms of the initial mean squared error loss function, with a value of 0.0125. By the tenth epoch the loss is at 0.0085 already lower than the univariate model's minimum loss of 0.0108 at the fiftieth epoch when training was halted. The loss decrease for the multi-feature model is more pronounced than for the single feature model. The mean square error gradually decreases to 0.0022 at the one hundredth epoch when training stops. It is interesting to note that no callbacks were triggered during training meaning that with more epochs the loss function can be decreased even further or rather, model accuracy can be further improved.

4.3.3.2 Model Prediction Accuracy

This study applies root mean squared error (RMSE) to evaluate predictive accuracy for both models. Table 4.3 summarizes the RMSE loss results for each of the 25 stocks analyzed:

Table 4. 3: Root mean squared error summary

	RMSE	
	Univariate LSTM	**Multivariate LSTM**
ABSA	0.0136	0.0116
BAMB	0.0169	0.0185
CIC	0.0271	0.0202
COOP	0.0135	0.0097
DTK	0.0179	0.0178
EABL	0.0183	0.0106
EAC	0.0317	0.0248
EQTY	0.0149	0.0097
FTGH	0.0330	0.0349
HFCK	0.0303	0.0230
JUB	0.0227	0.0210
KCB	0.0126	0.0077
KEGN	0.0190	0.0186
KPLC	0.0228	0.0193
LBTY	0.0414	0.0244
NCBA	0.0199	0.0119
NMG	0.0312	0.0255
NSE	0.0269	0.0256
SASN	0.0325	0.0207
SBIC	0.0291	0.0066
SCOM	0.0166	0.0066
SMER	0.0422	0.0164
TCL	0.0411	0.0369
TOTL	0.0243	0.0200
TPSE	0.0374	0.0314
Average RMSE	**0.0255**	**0.0190**

On balance, the multivariate LSTM model has a lower RMSE than the univariate LSTM as well as on average. For the single feature model, SMER had the greatest RMSE at 0.0422 while KCB had the least prediction error at 0.0126. For the multi-feature model, TCL had the greatest prediction error at 0.0369 while KCB and SCOM had the lowest at 0.0066.

4.3.4 Asset Preselection

In this stage, we analyze the characteristics of portfolios containing k number of assets. While it is appropriate to hold numerous stocks as a portfolio to enhance diversification, previous studies postulate that it is more realistic for individual investors to select at least

5-10 stocks in the construction of an optimal portfolio (Wang et al., 2019; Chaweewanchon & Chaysiri, 2022; Chen et al., 2021; Tanaka et al., 2000; Almahdi & Yang, 2017). Wang et al. (2019) assumes that an individual investor holding less than or equal to 10 assets is realistic in a financial market. They compare the portfolio performance of LSTM-predicted returns in a mean-variance optimization model over $k \in \{4, 5, 6, 7, 8, 9, 10\}$ assets; where stock return predictions were ranked and those stocks with the highest predictions were chosen. They found the most optimal portfolio at $k = 10$. This study applies a similar approach choosing the top 10 stocks with the highest average predicted returns for each model to proceed to the portfolio optimization stage. The selected assets and their average predicted returns for the 30 days period are summarized in tables 4.4 & 4.5 below.

Table 4. 4: 30-day average univariate LSTM-predicted stock returns

Univariate LSTM	
TOTL	1.003755
NSE	1.002778
SCOM	1.001956
KEGN	1.001469
EQTY	1.001131
DTK	1.000699
JUB	1.000525
SBIC	1.000427
TPSE	1.000103
CIC	0.99962

Table 4. 5: 30-day average multivariate LSTM-predicted stock returns

Multivariate LSTM	
NMG	1.008389
HFCK	1.007671
TPSE	1.007008
SMER	1.006715
FTGK	1.005086
TOTL	1.004934
SASN	1.004046
EQTY	1.003621
CIC	1.002723
ABSA	1.002697

4.3.5 Mean-Variance Portfolio Optimization

This research assumes that the optimization technique represented by Markowitz' mean-variance optimization model is well-suited to enhance the Sharpe ratio of a portfolio of assets. For simplicity, the study will refer to the portfolio consisting of returns predicted by the single feature LSTM as a 'uni-portfolio' and likewise a portfolio consisting of multi-feature LSTM-predicted returns as a 'multi-portfolio'.

Prior to the portfolio optimization process, all returns for selected stock assets were 'logarithmized' via a natural log function to ensure stability of returns from different stocks. The log return plots for assets within the uni-portfolio and multi-portfolio are illustrated in Figure 4.6 & 4.7 respectively below.

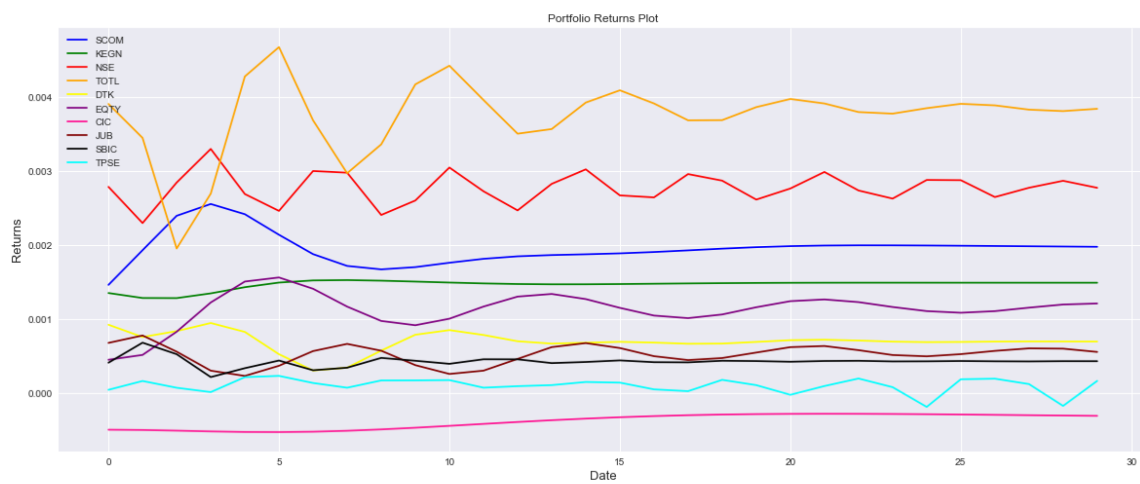

Figure 4. 6: Uni-portfolio returns plot

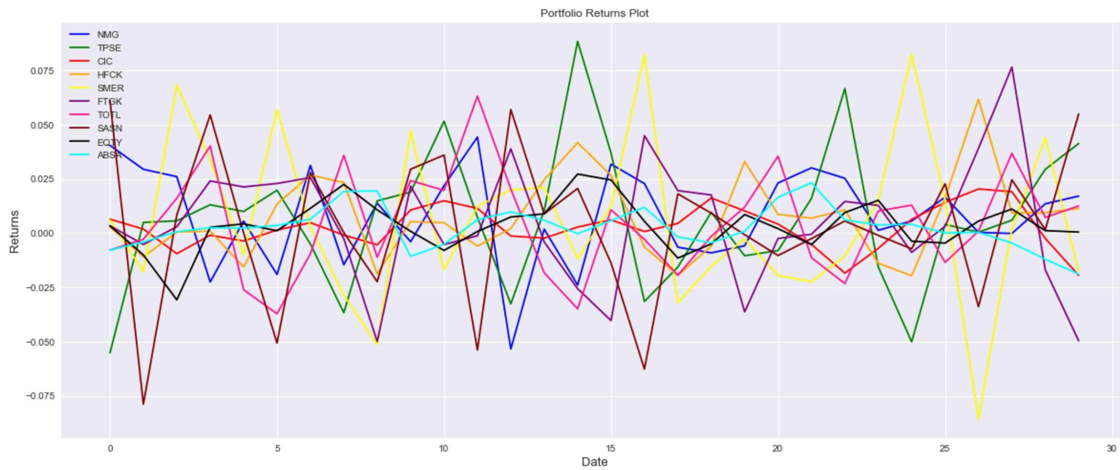

Figure 4. 7: Multi-portfolio returns plot

The next step is determining the optimal combination of portfolio weights that maximizes the Sharpe ratio. 50,000 portfolio simulations are repeated each with randomized weights that yield various portfolio returns, risks and Sharpe ratios within a frontier. A 91-day treasury bill risk free rate of 9.581% was applied as per the Central Bank of Kenya (Central Bank of Kenya, 2022). Following the research conducted by Wang et al. (2019), which examines the annual risk-return characteristics of various machine learning-based portfolio types, this study endeavors to continue in a similar vein by comparing the annualized portfolio performance of a uni-portfolio to a proposed multi-portfolio.

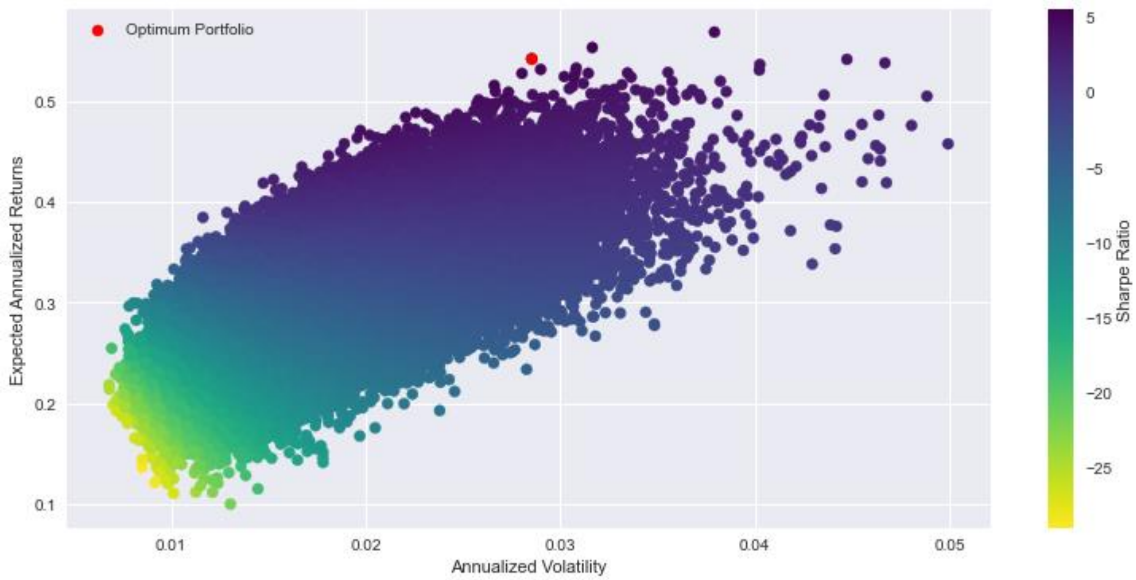

Figure 4. 8: Uni-portfolio efficient frontier

The above Figure 4.8 shows the efficient frontier of the 50,000 uni-portfolios constructed. The most optimum portfolio that had the highest Sharpe ratio is marked in red. The constructed optimum portfolio had a Sharpe ratio of 5.592, annualized expected return of 0.543 and annualized volatility of 0.0285.

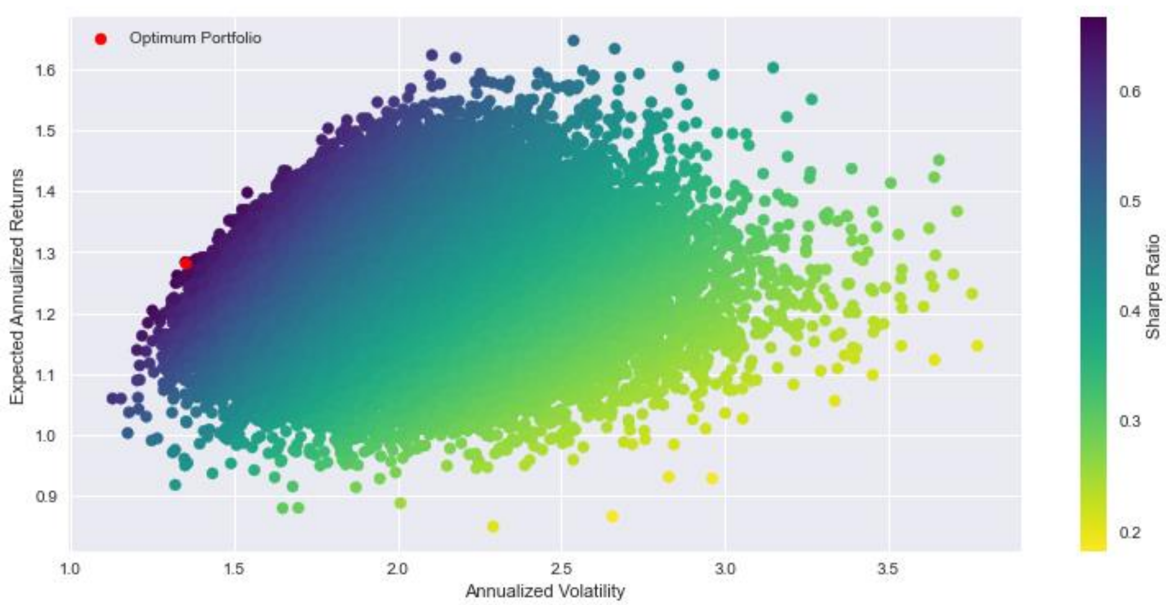

Figure 4. 9: Multi-portfolio efficient frontier

Figure 4.9 shows the efficient frontier consisting of various random-weighted multi-portfolios. The optimal portfolio marked in red had a lower annualized Sharpe ratio compared to its univariate-based counterpart of 0.667 though had a higher annualized expected return of 1.283 and higher annualized volatility of 1.351.

4.4 Discussion of Findings

4.4.1 Model Prediction Performance

LSTM is applied to forecast future stock returns. The experimental LSTM neural network applies two input methods, namely, a univariate input feature of solely stock return and a multifactor approach that incorporates various input characteristics based on stock returns and technical indicators. Twenty of these input features were iteratively selected through RFE via a RandomForest regressor. Between both input configurations, each LSTM architecture was kept constant with the only differences being the hyperparameters and array shapes. Both models were applied to predict closing returns of 25 randomly selected stocks listed on the NASI.

The prediction results found in this study are consistent with the research of Du et al. (2019) who found greater prediction accuracy a lower mean squared error loss function for a multi-input LSTM compared to a single-feature LSTM model of Apple stock price. The study under consideration solely analyzed the stock price of Apple, which the authors recognize as a potential limitation since the price of an individual stock may not reflect the overall forecast effect of the stock market. Building upon their study, this paper expands their analysis by applying LSTM for return prediction to a sample of 25 selected stocks. The results indicate that the multivariate LSTM model exhibits superior prediction performance across the sample of stocks, as evidenced by lower mean squared error loss functions during training, and lower root mean squared error forecast errors for each stock.

These findings show that the multivariate LSTM approach is a more effective means of forecasting stock returns as multiple variables have more impact on the outcome. The observed improvement in prediction accuracy highlights the potential usefulness of multivariate LSTM models in financial forecasting and investment decision-making.

While the multivariate LSTM demonstrates superior performance in financial forecasting, this does not mean ruling out the single-feature LSTM entirely. While the univariate LSTM fails to accurately capture stock return values, it can still effectively capture their trends, indicating that the model may be better suited for classification tasks and trend prediction, rather than direct value forecasting (Kuber et al., 2022).

4.4.2 Portfolio Performance

The main objective of this study was to compare the portfolio performance of the proposed portfolio consisting of multivariate LSTM-predicted returns with a portfolio consisting of predicted returns from a single input LSTM model. A mean-variance portfolio optimization technique was applied to determine the optimum allocation among 10 assets with the highest predicted returns.

Before portfolio optimization, the significant difference in the stock prediction results between both LSTM architectures should be noted. Specifically, the univariate LSTM return predictions fail to capture the volatility that is commonly observed in stock returns, and instead cluster closely around the mean value. In contrast, the multivariate LSTM return predictions more accurately reflect the volatility clusters and variability observed in stock returns. As each set of stock returns is utilized as inputs into a mean-variance optimizer, the observed difference in stock return predictions translates into a difference in portfolio performance, suggesting that the choice of LSTM architecture can have a significant impact on the performance of portfolio performance.

The empirical analysis conducted in this study has shown that the uni-portfolio exhibits superior performance, as evidenced by a Sharpe ratio of 5.592, in comparison to the proposed multi-portfolio which has a Sharpe ratio of 0.667. However, the latter portfolio provides more realistic results. The financial market is characterized by the volatility inherent in stocks, which is measured by the degree of market price fluctuations. This volatility is fundamental to the theory and practice of asset pricing, asset preselection, asset allocation, and risk management (Andersen et al., 2001). Therefore, the returns predicted by the univariate LSTM may not accurately reflect the dynamics of financial markets.

The present study has demonstrated that the optimal portfolio selection outcomes between a univariate and multivariate LSTM are quite different. While achieving excess returns to risk is imperative for any portfolio, the accuracy of returns predictions is a crucial aspect to be considered. This notion is substantiated by Lin et al. (2006) who in their study compare the portfolio selection results of returns predicted by a proposed Elman neural network and those predicted by vector autoregression. Their research found that both portfolios had differing risk-return characteristics. As the Elman network had more accurate return predictions, they concluded that the portfolio of Elman-predicted returns should be deemed the most optimal dynamic portfolio selection. Correspondingly, this study found the multivariate LSTM to have more accurate return predictions that captures stock return variability and therefore recommends a multi-feature long-short-term memory neural network for asset preselection.

Chapter 5

Conclusion & Recommendations

5.1 Conclusion

This paper has proposed an investment decision framework termed "multivariate LSTM + MV" which combines a multivariate LSTM model and Markowitz' mean-variance optimization technique for optimal asset selection. The results indicate that the multivariate LSTM approach outperforms the single-feature LSTM model in terms of prediction accuracy therefore substantiating previous studies that including input features can significantly improve model forecasting performance (Wang et al., 2019; Paiva et al., 2019; Du et al., 2019; Ma, 2020; Kuber et al., 2022).

This study notes that the strengths and weaknesses of each model approach may depend on the specific financial application. In applications where trend prediction is more important than actual value forecasting, univariate models may be more appropriate. Conversely, in applications where multiple inputs are critical for accurate predictions, multivariate models may provide an advantage.

Regarding portfolio performance, the empirical analysis shows that the optimal portfolio selection outcomes between a univariate and multivariate LSTM are quite different. Although the uni-portfolio exhibited superior performance, the multivariate LSTM-based portfolio provides more realistic results, capturing stock return variability more accurately. As noted by Lin et al. (2006), the accuracy of the underlying machine learning model is fundamental in determining the efficacy of dynamic portfolio selection. Therefore, this study advocates for the adoption of a multi-feature long-short-term memory neural network, which has been shown to produce more accurate predictions for asset preselection.

5.2 Limitations

In the undertaking of this study, the following limitations were noted:

1. Due to the time-consuming and computationally intensive nature of hyperparameter tuning and recursive feature elimination, it was not feasible to perform fine-tuning for each stock data. Therefore, this process was only conducted at the beginning of analysis and in cases where abnormal model performance was observed.
2. This study used 8 technical indicator variables and 17 lagged versions of return to predict future stock return; however, additional external factors such as government policies, interest rates and public events as well as macroeconomic variables such as dividend yields, book-to-market ratios and earning-price ratios can have an impact on the financial market and can also be considered as inputs into the models (Dai et al., 2021; Wang et al., 2019). Furthermore, in the stock market, the behavior and opinions of investors can significantly affect the fluctuations of stock markets. As a result, it is crucial to take investor sentiment data into account when making stock predictions (Jin et al., 2019).
3. This study did not account for transaction costs such brokerage costs and taxes which can have a significant impact on portfolio performance, especially in the short term. Their effects on portfolio optimization can be considered for further study.

5.3 Recommendations

The limitations noted in section 5.2 give rise to the following research recommendations:

1. Future research can incorporate additional macroeconomic, market-related and investor sentiment variables to improve the model's robustness to capturing stock return variability and to provide a more comprehensive view of the market.
2. Future works could include the effect of transaction costs into the portfolio optimization process by considering factors such as brokerage fees, taxes, and bid-

ask spreads. These factors can significantly influence portfolio performance especially in the short term. Therefore, it is essential to include them in portfolio optimization models to ensure that the returns are not overly optimistic and that the portfolios selected are feasible in practice

3. Lastly, further research can be conducted to compare LSTM (multivariate) with other advanced deep learning algorithms for financial forecasting. For instance, Dave et al. (2020) propose a hybrid ARIMA-LSTM for forecasting Indonesia's exports. They found that the hybrid approach surpasses the stand-alone LSTM and ARIMA in terms of prediction accuracy. Therefore, implementing such a model in stock prediction can represent a promising research avenue.

References

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014.

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Stock price prediction using the ARIMA model. *UKSim-AMSS 16th international conference on computer modelling and simulation* (pp. 106-112). IEEE.

Almahdi, S., & Yang, S. Y. (2017). An adaptive portfolio trading system: A risk-return portfolio optimization using recurrent reinforcement learning with expected maximum drawdown. *Expert Systems with Applications*, 87, 267-279.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). *The distribution of realized stock return volatility. Journal of financial economics*, *61(1)*, 43-76.

Banik, S., Sharma, N., Mangla, M., Mohanty, S. N., & Shitharth, S. (2022). LSTM based decision support system for swing trading in stock market. *Knowledge-Based Systems, 239*, 107994.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. A. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies, 11(7)*, 1636.

Capital Markets Authority. (2019). Retrieved 22 August 2022, from https://www.cma.or.ke/index.php/about-us/who-we-are

Central Bank of Kenya. (2022). *Treasury Bills*. Retrieved December 26, 2022, from https://www.centralbank.go.ke/bills-bonds/treasury-bills/

Central Bank of Kenya. (2022, May). Monthly Economic Indicators, May 2022. Retrieved August 22, 2022, from https://www.centralbank.go.ke/monthly-economic-indicators/

Central Depository and Settlement Corporation Limited (CDSC). (2022). About Us. Retrieved August 22, 2022, from http://www.cdsckenya.com/about-us/company-profile/

Chaweewanchon, A., & Chaysiri, R. (2022). Markowitz Mean-Variance Portfolio Optimization with Predictive Stock Selection Using Machine Learning. *International Journal of Financial Studies, 10(3)*, 64.

Chen, R. C., Dewi, C., Huang, S. W., & Caraka, R. E. (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, 7(1), 1-26.

Chen, W., Zhang, H., Mehlawat, M. K., & Jia, L. (2021). Mean–variance portfolio optimization using machine learning-based stock price prediction. *Applied Soft Computing*, 100, 106943.

Dai, Z., Zhu, H., & Kang, J. (2021). New technical indicators and stock returns predictability. *International Review of Economics & Finance*, *71*, 127-142.

Dave, E., Leonardo, A., Jeanice, M., & Hanafiah, N. (2021). Forecasting Indonesia exports using a hybrid model ARIMA-LSTM. *Procedia Computer Science*, *179*, 480-487.

Deng, S., Mitsubuchi, T., Shioda, K., Shimada, T., & Sakurai, A. (2011, December). Combining technical analysis with sentiment analysis for stock price prediction. *IEEE ninth international conference on dependable, autonomic and secure computing* (pp. 800-807). IEEE.

Du, J., Liu, Q., Chen, K., & Wang, J. (2019, March). Forecasting stock prices in two ways based on LSTM neural network. *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (pp. 1083-1086). IEEE.

Estrada, J. (2011). Returns. The Essential Financial Toolkit (pp. 1-13). Palgrave Macmillan, London.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417.

Fama, E. F., & French, K. R. (2014). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Freitas, F. D., De Souza, A. F., & De Almeida, A. R. (2009). Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, 72(10-12), 2155-2170.

Gu, S., Kelly, B. T., & Xiu, D. (2018). Empirical asset pricing via machine learning. *Chicago Booth Research Paper*, (18-04), 2018-09.

Guerard Jr, J. B., Markowitza, H., & Xu, G. (2020). Earnings forecasting in a global stock selection model and efficient portfolio construction and management. In *HANDBOOK OF APPLIED INVESTMENT RESEARCH* (pp. 75-85).

Hansson, M. (2017). On stock return prediction with LSTM networks.

Hellström, T. (1998). A random walk through the stock market (Doctoral dissertation, Univ.).

Huang, C., Chen, P., & Pan, W. (2011). Using Multi-Stage Data Mining Technique to Build Forecast Model for Taiwan Stocks. *Neural Computing and Applications*.

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685-695.

Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. *Neural Computing and Applications, 32*, 9713-9729.

Kihoro, J. M., & Okango, E. L. (2014). Stock market price prediction using artificial neural network: an application to the Kenyan equity bank share prices. *Journal of Agriculture, Science and Technology*, 16(1), 161–172. https://www.ajol.info/index.php/jagst/article/view/112890

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kirci, B. K., & Baydogmus, G. K. (2022, June). The Effect of Loss and Optimization Functions on Bitcoin Rate Prediction in LSTM. *International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-4). IEEE.

Kuber, V., Yadav, D., & Yadav, A. K. (2022). Univariate and Multivariate LSTM Model for Short-Term Stock Market Prediction. *arXiv preprint arXiv:2205.06673.*

Lara-Benítez, P., Carranza-García, M., & Riquelme, J. C. (2020). An Experimental Review on Deep Learning Architectures For Time Series Forecasting. *International Journal of Neural Systems*, 31(03), 2130001.

Levine, R. (1996). Stock Markets: A Spur to Economic Growth. *Finance & Development*, 7-10.

Lin, C. M., Huang, J. J., Gen, M., & Tzeng, G. H. (2006). Recurrent neural network for dynamic portfolio selection. *Applied Mathematics and Computation, 175(2)*, 1139-1146.

Livieris, I. E., Stavroyiannis, S., Iliadis, L., & Pintelas, P. (2021). Smoothing and stationarity enforcement framework for deep learning time-series forecasting. *Neural Computing and Applications, 33(20)*, 14021-14035.

Ma, Q. (2020). Comparison of ARIMA, ANN and LSTM for stock price prediction. *In E3S Web of Conferences* (Vol. 218, p. 01026). EDP Sciences.

Ma, Y., Han, R., & Wang, W. (2020). Portfolio optimization with return prediction using deep learning and machine learning. *Expert Systems with Applications*, 165, 113973.

Malkiel, B. G. (1973). A Random Walk Down Wall Street. Norton & Co.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59-82.

Manurung, A. H., Budiharto, W., & Prabowo, H. (2018). Algorithm and modeling of stock prices forecasting based on long short-term memory (LSTM). *ICIC Express Letters, 12(12)*, 1277-1283.

Markowitz, H. (1959). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Masese, J.M. (2017). Portfolio Optimization in the Kenyan Stock Market: A Comparison between Mean-Variance Optimization and Threshold Accepting. SU+ Digital Repository.

Meier, C. (2018). Aggregate Investor Confidence in the Stock Market. *Journal of Behavioral Finance*, 421-433.

Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. *Procedia Computer Science*, 170, 1168-1173.

Ochieng', S. A. (2016). Forecasting Equity Prices For Selected Companies At The Nairobi Securities Exchange (dissertation). SU+ Digital Repository.

Oriani, F. B., & Coelho, G. P. (2016, December). Evaluating the impact of technical indicators on stock forecasting. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1-8). IEEE.

Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. *Expert Systems with Applications*, 115, 635-655.

Petrusheva, N., & Jordanoski, I. (2016). Comparative analysis between the fundamental and technical analysis of stocks. *Journal of Process Management and New Technologies*, 4(2), 26-31.

Preeti, Bala, R., & Singh, R. P. (2019, July). Financial and non-stationary time series forecasting using LSTM recurrent neural network for short and long horizon. *10th international conference on computing, communication, and networking technologies (ICCCNT)* (pp. 1-7). IEEE.

Pring, M. J. (2002). Technical analysis explained: The successful investor's guide to spotting investment trends and turning points. *McGraw-Hill Professional*.

Qian, F., & Chen, X. (2019, April). Stock prediction based on LSTM under different stability. *IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* (pp. 483-486). IEEE.

Ross, V. (2022). The Major Challenges Facing Securities Regulators. *Paris: European Securities & Markets Authority*.

Sen, J., Dutta, A., & Mehtab, S. (2021, October). Stock portfolio optimization using a deep learning LSTM model. *IEEE Mysore Sub Section International Conference (MysuruCon)* (pp. 263-271). IEEE.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425-442.

Shcherbakov, M. V., Brebels, A., Shcherbakova, N. L., Tyukov, A. P., Janovsky, T. A., & Kamaev, V. A. E. (2013). A survey of forecast error measures. *World Applied Sciences Journal*, 24(24), 171-176.

Ta, V. D., Liu, C. M., & Tadesse, D. A. (2020). Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading. *Applied Sciences*, 10(2), 437.

Tanaka, H., Guo, P., & Türksen, I. B. (2000). Portfolio selection based on fuzzy probabilities and possibility distributions. *Fuzzy sets and systems*, *111(3)*, 387-397.

Tanaka-Yamawaki, M., & Tokuoka, S. (2007). Adaptive use of technical indicators for the prediction of intra-day stock prices. *Physica A: Statistical Mechanics and its Applications*, *383(1)*, 125-133.

Thomsett, M. C. (1998). Mastering Fundamental Analysis. Chicago: Dearborn Publishing.

Wang, W., Li, W., Zhang, N., & Liu, K. (2019). Portfolio formation with preselection using deep learning from long-term financial data. *Expert Systems with Applications*, 143, 113042.

Wanjawa, B. W. (2014). A Neural Network Model for Predicting Stock Market Prices at the Nairobi Securities Exchange (Doctoral dissertation, University of Nairobi).

Wen, M., Li, P., Zhang, L., & Chen, Y. (2019). Stock market trend prediction using high-order information of time series. *Ieee Access*, 7, 28299-28308.

Wu, C. H., Lu, C. C., Ma, Y. F., & Lu, R. S. (2018, November). A new forecasting framework for bitcoin price with LSTM. *IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 168-175). IEEE.

Zelenka, J. (2021). Does LSTM neural network improve factor models' predictions of the European stock market?.

Zhuge, Q., Xu, L., & Zhang, G. (2017). LSTM Neural Network with Emotional Analysis for prediction of stock price. *Engineering letters*, 25(2).