Utilizing Genetic Algorithms in Conjunction with ANN-Based Stock Valuation Models to Enhance the Optimization of Stock Investment Decisions

Ying-Hua Chang * and Chen-Wei Huang

Department of Information Management, TamKang University, New Taipei 251301, Taiwan
* Correspondence: yhchang@mail.tku.edu.tw; Tel.: +886-2-26215656 (ext. 2976)

Abstract: Navigating the stock market’s unpredictability and reducing vulnerability to its volatility requires well-informed decisions on stock selection, capital allocation, and transaction timing. While stock selection can be accomplished through fundamental analysis, the extensive data involved often pose challenges in discerning pertinent information. Timing, typically managed through technical analysis, may experience delays, leading to missed opportunities for stock transactions. Capital allocation, a quintessential resource optimization dilemma, necessitates meticulous planning for resolution. Consequently, this thesis leverages the optimization attributes of genetic algorithms, in conjunction with fundamental analysis and the concept of combination with repetition optimization, to identify appropriate stock selection and capital allocation strategies. Regarding timing, it employs deep learning coupled with the Ohlson model for stock valuation to ascertain the intrinsic worth of stocks. This lays the groundwork for transactions to yield favorable returns. In terms of experimentation, this study juxtaposes the integrated analytical approach of this thesis with the equal capital allocation strategy, TAIEX, and the Taiwan 50 index. The findings affirm that irrespective of the Taiwan stock market’s bullish or bearish tendencies, the method proposed in this study indeed facilitates investors in making astute investment decisions and attaining substantial profits.

Keywords: combination with repetitions; generic algorithms; Ohlson model; deep learning; optimization techniques; investment strategies

1. Introduction

Considering the increasing variety of investment options, numerous investment methods and information are available. In this era of low interest rates, the rewards from keeping funds in the bank have diminished. Therefore, investors are compelled to seek higher returns on their investments. There are many investment tools available, such as stocks, insurance, funds, and futures. Many investors prefer equities due to the transparency and convenience of stock market information, coupled with high liquidity and potential for high returns. However, higher returns often come with higher risks. In the rapidly fluctuating stock market, investors can become victims of high risk if they lack sufficient knowledge or receive incorrect information. Therefore, it is crucial for investors to have a sound financial and investment strategy to make informed decisions and achieve high returns.

Nevertheless, numerous factors influence stock prices, leading to changes in the investment environment and increased complexity. Even seasoned investors cannot accurately predict stock price movements. With a vast array of stocks available, selecting the right ones to invest in poses a challenge. Investment decisions vary based on individual preferences, the information received, and risk tolerance. However, the ultimate goal remains the same for all investors—to achieve the highest return at the lowest risk.

Many studies focus solely on addressing one or two aspects, such as stock selection, the timing of buying and selling, or capital allocation, lacking a comprehensive approach that encompasses all three key areas currently. The absence of an integrated approach in
research may result in investors struggling to obtain a comprehensive investment strategy. Furthermore, while fundamental analysis and technical analysis are widely used in investment decision-making, effectively integrating them remains a challenge. Simultaneously, with the increase in market data volume, the ability to process and analyze these data becomes crucial, necessitating exploration into enhancing the speed and efficiency of data processing to support effective investment decisions. Filling these research gaps will help enhance the overall effectiveness of stock investment and provide investors with more robust tools to tackle market challenges.

Recently, research in artificial intelligence has seen vigorous development, with applications in the financial sector gradually maturing and relevant studies continually emerging. Artificial intelligence has proven to be an effective tool for investment decision-making, especially in solving optimization problems with large search spaces. Studies by Bao [1], Bermúdez et al. [2], Chang Chien et al. [3], Chang et al. [4], Chen et al. [5], Dastkhan et al. [6], Fu et al. [7], Gorgulho et al. [8], Kamali [9], Leu et al. [10], NG et al. [11], Orito et al. [12], Papadamou et al. [13] and Rafaely et al. [14] have used genetic algorithms to solve portfolio optimization problems, achieving high returns at low risk.

Similarly, Chavarnakul et al. [15], Du et al. [16], Kara et al. [17], Lazo et al. [18], Liu et al. [19], Nazemi et al. [20] and Varnami et al. [21] have utilized neural networks to construct various algorithms and prediction methods, discussing stock trading profitability and achieving excess returns. These studies demonstrate that artificial intelligence can indeed provide optimal solutions. Notably, genetic algorithms do not require additional assumptions or limitations to obtain the best overall satisfactory solution. This study, therefore, uses genetic algorithms in artificial intelligence as a foundation for research.

The Ohlson model, proposed in 1995, is a price evaluation model that uses accounting information from financial statements. This model can determine the true value of a company based on defined variables. A nonlinear relationship exists between these variables and the company’s value. Therefore, this study leverages the learning ability of neural networks to simulate the relationships between variables and the company’s value, aiding in determining the company’s correct value. When making timing decisions, investors can compare the true value of a company’s stock with its market value, enabling them to make informed buy or sell decisions.

Through effectively making these three key decisions, investors can attain substantial risk premiums. Therefore, this study utilizes the optimization properties of genetic algorithms, combined with the principles of fundamental analysis and repetitive portfolio optimization, to seek suitable timing, stock selection, and capital allocation strategies. In terms of timing, while previous research mostly employs technical analysis to determine the timing for buying and selling, this paper integrates the stock valuation Ohlson model of fundamental analysis with deep learning to effectively provide timing judgment information, thus addressing the lagging issue of technical analysis in timing resolution. To augment the speed and efficiency of processing stock data, this paper introduces an adept genetic algorithm encoding method founded on the principle of optimizing resource allocation through combination with repetitions. The objective is to effectively tackle the optimization challenges associated with stock selection and capital allocation in stock investment. In terms of experimentation, this study compares the proposed comprehensive analysis method with equal capital allocation strategies, TAIEX, and the Taiwan 50 index. The results demonstrate that regardless of whether the Taiwanese stock market is in a bull or bear market, the method proposed in this study can indeed assist investors in making outstanding investment decisions and achieving substantial profits.

2. Literature Review
2.1. Stock Investment Research

Investing in the stock market necessitates a careful evaluation of risk and potential returns. To mitigate investment risk, investors can distribute their capital across various
investment options. An effective investment strategy should encompass the selection of stocks, the timing of investments, and the allocation of capital.

The objective of timing is to pinpoint the most opportune moment to buy or sell a stock. Various studies, including those by Allen et al. [22], Badawy et al. [23], Bao [1], Chang Chien et al. [3], Chang et al. [24], Huang et al. [25], Jiang et al. [26] and Korczak et al. [27] have integrated genetic algorithms with other methodologies like technical analysis, regression, and support vector machine to identify optimal stock market trading opportunities. These studies aimed to showcase impressive returns through experimental demonstrations. The use of genetic algorithms to optimize the stock index prediction model has proven effective, as demonstrated by Du et al. [16], Fu et al. [28], Khan et al. [29], Sahin et al. [30], Shen et al. [31], and Versace et al. [32]. Furthermore, Jang et al. [33], Kim et al. [34] and Yang et al. [35] developed an intelligent stock trading system that uses neural networks to predict stock prices and market trends. Some researchers have also employed genetic algorithms to address the slow convergence and low learning efficiency issues associated with neural networks. Mahmoodi et al. [36] used three different models to predict stock market signals. Two models were better at predicting when to sell, and one was better at predicting when to buy. The model using the Imperialist Competitive Algorithm performed the best. Solares et al. [37] used artificial neural networks and fundamental analysis to predict future stock prices. They also used a method called differential evolution and fundamental analysis to choose the most sensible stocks. Lastly, they used genetic algorithms and statistical analysis to create the best investment portfolio. The tests showed that most of the time, their methods did better than the standard methods with statistical importance.

The objective of a selection strategy is to discern superior investment targets from a multitude of potential stocks. This process necessitates an evaluation of the return and risk associated with stock investment, leading to the identification of an optimal portfolio. To enhance the precision of stock price predictions, Cheng et al. [38] proposed a hybrid forecasting model that integrates genetic algorithms with rough set theory. This model has the potential to outperform both the buy-and-hold strategy and traditional genetic algorithms in terms of returns. Chiu et al. [39] fused genetic algorithm, regression, and technical analysis to formulate a stock selection model, thereby identifying high-yield stocks as potential investment targets. Chun et al. [40] and Oh et al. [41] employed genetic algorithms, with the incorporation of a novel fitness function design, to select investment targets predicated on stock prices. This methodology can yield consistent returns irrespective of the fluctuations in the stock market. Gold et al. [42] amalgamated artificial intelligence with fundamental analysis to optimize portfolios, thereby validating the applicability of artificial intelligence within the realm of investment. Ozcalici et al. [43] used artificial neural networks to pick stocks. The stocks they picked did better than just buying and holding onto stocks. They also used special rules and artificial neural networks to help investors make decisions and potentially earn more money. Yodmun et al. [44] suggested a methodology for stock selection that initially categorizes stocks and then determines investment weights via fuzzy quantitative analysis. Subsequently, stocks are ranked according to the fuzzy hierarchy analysis method. Investors can then select stocks based on these final investment weights. Yun et al. [45] employed a genetic algorithm coupled with learning regression models to pinpoint the ideal feature subset. Their objective was to augment the interpretability of deep learning in the context of stock price prediction.

Capital allocation is designed to aid investors in maximizing portfolio return by distributing funds among various investment targets. Adebiyi et al. [46] used genetic algorithms to make investment portfolios better. They suggested ways to distribute assets to obtain the best returns. This method helps investors make smart decisions about risk when choosing and investing in portfolios, and it suggests how to distribute assets to obtain the best returns. Chen et al. [47] suggested a relational genetic algorithm guided by new mutation operations to boost evolution efficiency and solve a comprehensive portfolio optimization problem. Kocak [48] employed the Shapely value in cooperative
game theory to calculate the weight of each stock in the portfolio, thereby determining the capital allocation ratio. Shoaf et al. [49] introduced a novel method of encoding the genetic algorithm for capital allocation. Soleimani et al. [50] combined several heuristic and non-heuristic algorithms to propose a portfolio optimization model that addresses the nonlinear portfolio optimization problem and enhances the rate of return.

The literature review above indicates that stock portfolio optimization primarily concentrates on the selection of investment targets, the timing of buying or selling stocks, and the amounts invested in each investment target. However, a review of the past literature reveals that few studies have been able to address these three issues concurrently, with most only able to tackle one or two. Furthermore, strategies on how to enhance returns while mitigating risk also hold significant importance.

2.2. Deep Learning

Deep learning is a type of machine learning that mimics the way the human brain operates, using neural networks to learn from and extract valuable information from large amounts of data. This learning approach has wide applications in many fields, including speech recognition, image recognition, and natural language processing. Long Short-Term Memory (LSTM) is a special type of Recurrent Neural Network (RNN) proposed by Hochreiter and Schmidhuber in 1997, designed specifically to solve the long-term dependency problem. LSTM is a special deep learning model designed to handle time series data. What makes LSTM unique is its “memory cell”, which can store and access past information for long periods. This enables LSTM to capture long-term dependencies in time series data, something that many other models struggle to do. The reason for choosing LSTM to predict stock prices is that stock prices are typical time series data, and their variations depend not only on the current market conditions but also on past price trends. The memory cell in LSTM allows it to remember past price information and use it to predict future prices.

In recent years, numerous studies have proven the effectiveness of Long Short-Term Memory (LSTM) in stock price prediction. LSTM can be utilized not only for predicting stock prices but also for making stock investment decisions. Akita et al. [51] utilized LSTM neural networks to predict the closing prices of publicly listed companies. They incorporated both textual and numerical information as features and compared LSTM with various other neural networks. The results demonstrated that LSTM neural networks exhibited better profitability and were more capable of capturing the impact of time-series data. Furthermore, the inclusion of textual information as feature data led to more accurate prediction results compared to considering only numerical information. Bao et al. [1] employed LSTM to forecast the prices of the Shanghai Stock Exchange Composite Index and compared its predictive performance with other models, showing that LSTM outperformed the other models. Chen et al. [52] employed genetic algorithms combined with LSTM to build a model for stock price prediction. The experiments confirmed that their proposed predictive model effectively enhanced the accuracy of stock predictions. Chung et al. [53] employed genetic algorithms combined with LSTM to establish a predictive model for the Korean stock price index (KOSPI). The experimental results validated the lower Mean Squared Error (MSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) of the proposed GA-LSTM model. Ding et al. [54] employed LSTM to forecast stock price movements and made buy or sell decisions based on the prediction results. Their findings revealed that LSTM’s predictive performance surpassed that of traditional technical analysis methods and could generate stable positive returns. Fischer et al. [55] used LSTM to predict the daily returns of the German DAX index. Their findings demonstrated that LSTM effectively captured the market’s nonlinear dynamics and yielded predictive performance surpassing that of the random walk model. Krauss et al. [56] used LSTM to predict the daily returns of the German DAX index and made buy or sell decisions based on the prediction results. Their findings indicated
that LSTM effectively captured the market’s nonlinear dynamics and yielded predictive performance surpassing that of the random walk model.

Nelson et al. [57] developed an LSTM classification model to predict the future trends of stock prices in 2014 based on historical prices and technical analysis indicators. The research results also demonstrated that their proposed model, with few exceptions, outperformed other methods, and the risk associated with the LSTM model was relatively lower. Selvin et al. [58] also utilized LSTM to predict stock market prices and obtained similar results. They compared the performance of three different deep learning architectures—RNN, LSTM-RNN, and CNN—with the traditional ARIMA method. The results revealed that the deep learning models consistently outperformed the traditional ARIMA model. Siami-Namini et al. [59] compared the predictions of Dow Jones, Hang Seng, NASDAQ, Nikkei, S&P 500 indices, and several companies’ stocks using ARIMA and LSTM. The results showed that LSTM outperformed ARIMA.

These studies indicate that LSTM is a powerful tool that can help us better understand and predict changes in stock prices. However, there are still many unresolved issues and challenges in this field, such as how to choose appropriate model parameters and how to handle the uncertainty of data. These issues require further research and exploration. Due to its unique structure and capabilities, LSTM can effectively handle time-series data, such as stock price data, and has been proven to have superior performance in many studies. However, there are also some challenges. First, data quality is a critical issue. Stock price data may be influenced by many factors, such as market sentiment and policy changes, which may increase the noise in the data and affect the predictive performance of the model. Second, the interpretability of the model is another challenge. Although LSTM can generate highly accurate predictions, its internal workings are often difficult to understand, which may limit its use in practical applications.

The advantages of LSTM-based stock price prediction over traditional econometric models have been confirmed in multiple studies ([2,7,8,35,58,60]). These research papers provide valuable insights into and empirical results regarding the superiority of LSTM-based stock price prediction over traditional econometric models. By comparing the performance of different models in predicting stock prices, these studies can provide a deeper understanding, helping investors make wiser decisions in practice. LSTM can capture complex nonlinear patterns and long-term dependencies in the data, which may be challenging for traditional econometric models. Furthermore, LSTM can handle and analyze large amounts of historical stock price sequence data, extracting meaningful patterns from it, thus achieving more accurate predictions. Traditional econometric models may struggle to effectively capture the complex dynamics and nonlinear relationships present in stock price fluctuations. However, it is worth noting that comparisons between LSTM and traditional econometric models may vary depending on the specific dataset, market conditions, and analysis time range. While LSTM has shown promising predictive performance in many cases, in some situations, traditional econometric models may still provide competitive results. Overall, the advantage of LSTM lies in its ability to handle sequence data and capture complex patterns, which can provide a better edge in some stock price prediction tasks than traditional econometric models. Therefore, future research can focus on more effective data preprocessing and feature selection methods to improve data quality and also integrate other models to enhance the interpretability of LSTM.

2.3. Ohlson Equity Evaluation Model

The equity evaluation model proposed by Ohlson in 1995 leverages accounting information found in financial reports, specifically earnings and book value, to assess stock value. This model is based on three hypotheses: the present value of expected dividend (PVED), clean surplus relation (CSR), and the linear information dynamic model (LIM). These hypotheses lead to a linear evaluation model where the firm value is equal to the
book value plus abnormal earnings and non-accounting information. This relationship can be represented by the following Formula (1).

\[ P_t = bv_t + a_1 X_t^a + a_2 V_t \]  

And \[ a_1 = \frac{\omega}{(R_f - \omega)} \geq 0; a_2 = \frac{R_f}{(R_f - \omega)(R_f - \gamma)} \geq 0 \]

\( P \) = market price; \( bv \) = book value; \( X \) = abnormal earnings; 
\( V \) = non-accounting information

In the model, \( \omega \) and \( \gamma \) symbolize the parameters for abnormal earnings and other information related to sustainability. The larger these values, the more sensitive they are to the company’s value. Among these variables, the book value signifies the actual asset value of the stock, while abnormal earnings gauge current profitability. Other non-accounting information adjusts future profitability. However, the model does not specify what this other non-accounting information comprises.

White et al. [61] demonstrated that the Ohlson model has a strong explanatory power, reaching between 60% and 70%. Subsequent research by various scholars also affirmed the effectiveness of the Ohlson model ([62,63]). However, because Ohlson did not explicitly define the non-accounting information, many researchers made relevant adjustments, such as incorporating the corporate governance variable into the non-accounting information. This study found that the inclusion of corporate governance-related variables enhanced the explanatory power of the original Ohlson equity evaluation model. The research also confirmed a significant correlation between the characteristics of corporate governance and equity value and showed that the Ohlson equity evaluation model becomes more comprehensive with the addition of this variable.

This study employs the accounting-based Ohlson model to evaluate the true value of various types of stocks in the stock market. Factors influencing the value of the stock are divided into non-accounting information variables and accounting information variables. The accounting information, which serves as the basic variables for equity valuation, includes book value and earnings per share. The non-accounting information variable consists of eight corporate governance variables. These variables, proposed by Lee et al. [64], have been shown to effectively explain the equity value of a company. They include the supervisor and director shareholding ratio, large shareholder ratios, managers’ shareholding ratio, the group shareholding ratio, the external holdings ratio, shares surplus deviation variables, seat surplus deviation variables, and seat deviation variables.

### 3. Research Framework

This study leverages the search optimization capabilities of the genetic algorithm to facilitate appropriate investment decisions for stock selection and capital allocation. It also employs neural networks in conjunction with the Ohlson model to predict various stock prices, serving as a basis for buy or sell decisions in stock transactions. By integrating genetic algorithms, LSTM neural networks, and the Ohlson model, it is possible to unify three investment strategies: selection, timing, and capital allocation. The validity of this approach is verified in this paper by comparing it with several benchmarks.

Figure 1 illustrates the research framework. This study considers stocks listed on the Taiwan Stock Exchange as investment targets. Stock selection and capital allocation decisions are made based on the monthly rate of return of these stocks. For market timing, a correlation is identified between the input and output variables of the Ohlson model through the combination of a neural network with the Ohlson model to calculate the real value of stocks.
Hybrid Genetic Algorithms

Investment Decision Optimization

Figure 1. Research framework.

The real value of stocks can be used to determine whether the market value of the stock is overvalued or undervalued in the stock market. If the stock price is undervalued in the market, the proportion of funds obtained from genetic algorithms is used to buy. Conversely, if the market price is excessively high, the short selling method can be used to sell stocks and achieve a high rate of return on investment.

The pseudocode for the hybrid genetic algorithm proposed in this study is as follows:

Hybrid GA ( )

| Randomly set the initial population (Applied Combination with Repetition Encoding) |
| Calculate the fitness of the chromosomes |
| While (Unchanged fitness for 3000 generations ends evolution) |
| { |
| Selection (Roulette wheel selection) |
| Crossover (Adaptive crossover operation) |
| Mutation (Adaptive mutation operation) |
| Fitness evaluation of the chromosomes |
| (LSTM-Ohlson fusion model (timing) + Evaluation Criteria) |
| } |

The system initially generates a population of parent chromosomes based on the combination with a repetition encoding method. After generating these initial chromosomes, the system calculates the fitness value of each chromosome, representing its effectiveness in solving the problem at hand. Following this, the system enters the evolutionary process of the hybrid genetic algorithm. During the evolutionary process, selection operations are first performed based on the roulette wheel method. This method ensures that chromosomes with higher fitness values have a greater chance of being selected as parent chromosomes for the next generation. The selected chromosomes then undergo crossover operations using the adaptive crossover method proposed in this study, ensuring that the crossover process produces offspring chromosomes with enhanced advantages.

After crossover, the chromosomes undergo adaptive mutation operations according to a specified mutation rate. This step is intended to increase the diversity of the chromosomes and prevent the algorithm from becoming trapped in a local optimum. Subsequently, the system employs the LSTM–Ohlson fusion model to determine specific trading strategies. This model, which combines Long Short-Term Memory (LSTM) neural networks and the Ohlson model, aims to improve the timing accuracy of investment decisions. Upon completing these steps, the system evaluates each chromosome based on evaluation criteria and recalculates its fitness values. This process is repeated continuously, with chromosomes being optimized in each generation until the fitness values show no significant change for 3000 consecutive generations. At this point, the system concludes that the evolutionary process has converged and identifies the optimal investment strategy.

This process, by integrating genetic algorithms and deep learning models, achieves intelligent optimization of investment strategies. The initial population of parent chromo-
somes provides diverse solutions, while the selection operation using the roulette wheel method and the adaptive crossover and mutation mechanisms ensure that the offspring chromosomes progressively improve their fitness values. The application of the LSTM–Ohlson fusion model in determining trading strategies further enhances the model’s timing prediction accuracy and practical application effectiveness. Ultimately, through continuous iteration and optimization, the system is capable of automatically generating an optimal investment strategy, achieving the goal of intelligent investment decision-making.

The proposed method introduces an efficient combination with a repetition encoding scheme for chromosomes. In this scheme, the selection operation employs the roulette wheel method. To accommodate the combination with the repetition encoding scheme, special mating patterns are utilized for the crossover and mutation operations, as detailed in Section 3.1. The evaluation method for the fitness of chromosomes is also explained in Section 3.1.

3.1. Stock Selection and Capital Allocation

The investment targets of this study include eight categories of stocks and cash reserves. Chromosomes are evaluated using multiple criteria. Through the mechanisms of selection, crossover, and mutation, the genetic algorithm constantly evolves to solve optimization problems effectively, identifying the optimal capital allocation for the eight categories of stocks and cash reserves. This approach evaluates the return on the investment strategy, providing investors with a reference for decision-making that encompasses both stock prediction accuracy and practical application effectiveness. Ultimately, through continuous source allocation optimization. The capital allocation in investment represents an optimization problem of resource allocation, akin to solving the equation \(X_1 + X_2 + \ldots + X_n = r\) for non-negative integer solutions. This refers to arranging \(n-1\) 1’s and \(r\) 0’s as permutations of multi-sets. “+” is represented by 1, and 0 represents the number of resources; the equation \(X_1 + X_2 + \ldots + X_n = r\) can be interpreted as a problem of distributing \(r\) units of capital among \(n\) different stocks, where \(X_i\) represents the amount of capital allocated to the \(i\)-th stock. This study proposes a new encoding scheme for genetic algorithms based on the concept of combination with repetition to solve the optimization problem of capital allocation. To ensure that the chromosomes retain their repetitive combination characteristics during the evolutionary processes of mating and mutation in genetic algorithms, which is crucial for solving the capital allocation challenges in stock investment. This

3.1.1. Encoding and Decoding for Chromosome

Capital allocation is indeed a form of resource allocation problem. This study proposes a new encoding scheme for genetic algorithms to address the stock selection and capital allocation problems in resource optimization. In this study, the genetic algorithm used binary encoding with the concept of a homogeneous product, which randomly generated chromosomes by a combination of 0 and 1, as shown in Figure 2. The present study employs the concept of combinations with repetitions for encoding. In combinations with repetitions, from a set of \(n\) distinct items, any \(r\) items were selected as a group, where the number of each type of item was no less than \(r\), and repetitions were allowed. This type of combination is referred to as “\(n\) choose \(r\)” with repetitions.

\[
\begin{align*}
\text{Cash reserve} & \quad \text{the 1st category stocks} & \quad \text{the 2nd category stocks} & \quad \text{the 4th category stocks} & \quad \text{the 5th category stocks} & \quad \text{the 6th category stocks} & \quad \text{the 7th category stocks} & \quad \text{the 8th category stocks} \\
0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1
\end{align*}
\]

Figure 2. Chromosome encoding.

The capital allocation in stock investment can indeed be viewed as a problem of resource allocation optimization. The capital allocation in investment represents an optimization problem of resource allocation, akin to solving the equation \(X_1 + X_2 + \ldots + X_n = r\) for non-negative integer solutions. This refers to arranging \(n-1\) 1’s and \(r\) 0’s as permutations of multi-sets. “+” is represented by 1, and 0 represents the number of resources; the equation \(X_1 + X_2 + \ldots + X_n = r\) can be interpreted as a problem of distributing \(r\) units of capital among \(n\) different stocks, where \(X_i\) represents the amount of capital allocated to the \(i\)-th stock. This study proposes a new encoding scheme for genetic algorithms based on the concept of combination with repetition to solve the optimization problem of capital allocation. To ensure that the chromosomes retain their repetitive combination characteristics during the evolutionary processes of mating and mutation in genetic algorithms, which is crucial for solving the capital allocation challenges in stock investment. This
study introduces specialized crossover and mutation operators. The explanation is further supported by practical examples, as illustrated in Figures 3 and 4.

The utilization of the concept of combinations with repetitions, in conjunction with genetic algorithms, aims to optimize resource allocation problems. Chromosome encoding represents how much capital is allocated to investment targets. 0 represents the ratio of the funds invested in the investment target. 100% represents the total amount of investment. If there is m 0s, each 0 represents invest \( \frac{100}{m} \% \) proportion of funds. Moreover, 1 represents segmentation between investment targets. The number of 0 between two 1 in a chromosome represents the funds of corresponding investment target. For example, in eight categories of shares and cash reserves as investment targets, it must have eight 1 in a chromosome. If each 0 indicates that 10% of funds were invested, then there will be ten 0 in chromosomes.

Figure 2 shows that cash reserves, the first category stocks, the fourth category stocks, the fifth category stocks, and the seventh category stocks were invested into with 10% of funds; the sixth category stocks have 20% capital invested into them; and the second category stocks have 30% capital invested into them. No investments were made into the third and eighth category stocks.

### 3.1.2. Crossover and Mutation

During the evolutionary process, the chromosomes of the parent population undergo crossover and mutation operations. The combination with a repetition encoding scheme can reduce the search space for optimal solutions in genetic algorithms. Using conventional crossover and mutation operations would enlarge the search space for optimal solutions, thus reducing the efficiency of finding the best solution. Therefore, to enhance the efficiency of genetic algorithms, this study proposes a special crossover and mutation operation that complements the combination with a repetition encoding scheme. In order to ensure that chromosomes can still accurately represent the proportion of funds invested in various stocks after crossover and mutation operations, the number of 0s and 1s in the offspring chromosomes must be the same as those in their parent chromosomes.

The steps for performing crossover operations are as follows:

1. Randomly select two parent chromosomes to prepare for the crossover operation.

2. Randomly select two crossover points, as shown in Figure 3, and record the number and positions of 1s between the two crossover points. For example, the first parent chromosome has a total of three 1s, appearing at positions 0, 2, and 5, while the second parent chromosome has a total of two 1s, appearing at positions 1 and 4.
Step 3. Execute the crossover operation. The number of 1s in the offspring chromosomes must match that of the parent chromosomes. Thus, for the first offspring chromosome, randomly choose three positions from positions 0, 1, 2, 4, and 5, as illustrated in Figure 3. For instance, randomly select positions 0, 1, and 4. For the second offspring chromosome, select the remaining positions, such as positions 2 and 5. Generate new offspring chromosomes accordingly.

Step 4. Continue executing Steps 1 to 3 iteratively until all chromosomes in the parent population have undergone the crossover operation.

After completing the crossover operation, proceed with the mutation operation based on the mutation rate. Following the mutation operation, the count of 0s and 1s in the chromosomes must remain the same as before the mutation. The mutation operation involves randomly selecting one 1 and one 0 within a chromosome and swapping them to complete the mutation. For instance, randomly select the 1 at position 1 and the 0 at position 6 for swapping to accomplish the mutation, as illustrated in Figure 4.

3.1.3. Evaluation

For experimental analysis, this study designed three criteria as investment guidelines for investors, which serve as the evaluation function for the genetic algorithm, as shown in Table 1. De Souza et al. [65] explored the application of the Sharpe ratio in assessing global financial stability, emphasizing its effectiveness and ease of use in measuring risk-adjusted returns. This paper argued that the Sharpe ratio, by combining risk and return, offers a comprehensive, user-friendly, and meaningful tool for portfolio performance measurement. Hence, the first criterion adopts the Sharpe ratio.

<table>
<thead>
<tr>
<th>Number</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sharpe ratio for eight industry category stocks</td>
</tr>
<tr>
<td>2</td>
<td>The investment capital of Electronic category stocks is between 10% and 30%.</td>
</tr>
<tr>
<td>3</td>
<td>Investors hold 5% to 10% of funds in hand.</td>
</tr>
</tbody>
</table>

According to Wong et al. [66], Taiwanese investors’ preference for electronic stocks is related to international market trends because a significant portion of Taiwan’s exports comprises electronic products. Investors have high confidence in electronic stocks, believing they can provide stable returns. Consequently, Taiwanese investors tend to favor purchasing electronic stocks. Therefore, the second criterion sets the capital allocation for investing in electronic industry stocks to be between 10% and 30%.

Moreover, based on Jahan-Parvar et al.’s [67] study on the impact of liquidity on investment returns and risk, it is emphasized that investors should maintain a certain level of liquidity in their funds. Hence, the third criterion assumes that investors would want to hold between 5% and 10% of their funds in cash reserves to mitigate risk.

Fitness value of chromosome =

\[
\text{Criterion 1 (Sharpe ratio for eight industry category stocks)} + \\
\text{Criterion 2 (capital ratio for electronic category stocks)} + \\
\text{Criterion 3 (capital retention rate in hands)}
\]

Furthermore, in determining whether to buy or sell stocks, the LSTM–Ohlson fusion model must be invoked to make decisions and calculate the investment return of buying or selling the stock. Since each criterion is measured in different units, the values for each criterion must be normalized to ensure that scores fall between 5 and 100. The worst-performing criterion is assigned a score of 5 to give chromosomes a chance to survive to the next generation, while the best-performing criterion is assigned a score of 100 to increase the likelihood of chromosomes surviving to the next generation. The chromosome’s fitness
value is used as a reference for selection operation execution. Evolution is carried out based on the principle of survival of the fittest to obtain the optimal solution.

3.2. Timing

In terms of market timing, this study diverges from previous research that used technical analysis to identify trading opportunities for investment targets. Instead, it employs fundamental analysis to uncover the true value of a company. The Ohlson model’s book value, earnings per share, and corporate governance variables of non-accounting information are used to determine the true value of all stocks. However, these variables and the true value of stocks have a nonlinear relationship, making it challenging to discern a pattern using statistical methods. To address this, the study utilizes the widely used LSTM neural network to emulate the Ohlson model. The variables of the Ohlson model serve as nodes in the input layer, and the actual true value of stocks acts as a node in the output layer, establishing the architecture of the artificial neural network.

If the value of shares derived from the Ohlson model is higher than the actual market price, it indicates that the market undervalues these stocks. This suggests that the stock price will rise, making it a good opportunity for investors to buy these shares. Conversely, if the value of shares obtained from the Ohlson model is lower than the market price, it signifies that the market overestimates the value of these shares. The stock price is likely to fall in the future, so investors should consider selling these shares. If investors do not currently hold these shares, they can engage in short selling, which involves borrowing stocks from an investment institution to sell.

4. Experimental Design and Analysis

4.1. Hybrid Genetic Algorithms

The chromosome of integrated genetic algorithms employs binary encoding, and its length is 58 bits. There are eight 1s to distinguish the eight categories of stocks and cash reserves. Invested funds are put into 50 equal parts, so there are 50 0s, and for each, 0 represents 2% of all funds. The population size is 100. When evolution has the same fitness value, 3000 continuous generations that stand achieve convergence status and stop evolution. Roulette selection mechanism and elite law adopted for evolution. The crossover rate is set to 1. And mutation rate is set at 0.01 according to a rule of thumb.

4.2. LSTM–Ohlson Fusion Model

This paper employs LSTM to simulate the Ohlson model and to learn the relationships between financial variables and the true values of the company. The neural network architecture, which integrates the Ohlson model, is depicted in Figure 5. LSTM is particularly suited for this task as it is designed to handle time series data, making it ideal for financial forecasting and understanding the temporal dynamics of stock prices.

The input layer of the neural network has ten nodes, which consist of two basic variables and eight corporate governance variables related to equity valuation, as shown in Figure 5. The output layer node represents the true value of stocks. Through training, testing, and validation, the neural network can learn the relationship between these relevant variables and the real value of stocks. When new input variable values are introduced to the neural network prediction model, it can predict the actual market value of the shares. This approach allows for a dynamic and adaptable model that can adjust to new information and provide accurate predictions.

In the LSTM neural network’s stock assessment, a sliding window mechanism is used to improve forecasting performance. Specifically, data from every three months is used to predict the data for the fourth month. The training architecture of the LSTM neural networks can span three years. The data for these experiments are sourced from the Taiwan Economic Journal (TEJ), covering the period from 2008 to the end of 2015.
4.4. Experimental Analysis

Crossover operator adopts two-point crossovers. The crossover rate is 1. Each chromosome is 63 bits long, as shown in Figure 6. The section between the eighth bit and the seventh bit represents the proportion of capital reserved. Moreover, 68 represents the capital reserved being 68% of the total funds. The section between the eighth bit and the seventh bit represents the proportion of capital reserved. Moreover, 68 represents the proportion of investing in the first class of shares. Moreover, 76 says that investors must employ 76% of the total funds to invest in the first class of shares. The proportion of each class of stocks and capital reserves may be re-arranged according to the mutation rate, one gene of the chromosome is randomly selected; then switching the gene between the two crossover points and retained as offspring.

Figure 7 reveals the crossover of chromosomes. Two randomly selected crossover points, then switching the gene between the two crossover points and retained as offspring.

Chromosome mutation is shown in Figure 8. When a chromosome mutation must execute according to the mutation rate, one gene of the chromosome is randomly selected; the gene changes from 1 to 0 and vice versa. Evolutions achieve the mutation operator by this method.
In addition, the proportion of each class of shares and capital reserves may be retained between 0% to 100%. The sum of the proportion of each class of shares and capital reserves must equal 100%. The evaluation method of evolution employs the penalty function. When the proportion of the total funds is more or less than 100%, the number of penalty points will be determined according to the proportion of exceeds. The assessment of evolution also uses other criteria. When the electronic stocks investment ratio is less than 10% or greater than 30%, and cash reserves in the hands of less than 5% or greater than 10%, then the mechanism will take the penalty function. The population size is 100. The selection mechanism for SGA uses the roulette method, employing elite law. The Crossover operator adopts two-point crossovers. The crossover rate is 1. Each chromosome’s mutation rate is set to 0.01.

4.4. Experimental Analysis

The experimental analyses in this study were conducted according to the proposed methodology and framework. Eight categories of stocks were targeted for investment. An artificial neural network, combined with the Ohlson model, was used to assess the actual value of these stocks. Each month served as a trading point to determine the appropriate timing for entry and exit and to settle the rate of return. The return on investment for the Taiwan top 50 Index (TWN50) was used as a benchmark. This was compared to the rate of return on the average capital invested in the targeted stocks and the proportion of capital allocation obtained through a simple genetic algorithm for these stocks. The data used for the experiment were sourced from the Taiwanese Economic Journal (TEJ), covering the period from January 2008 to December 2015. To ensure higher accuracy, the experimental results were based on an average of ten experiments.

This study integrates the Ohlson stock assessment model and genetic algorithms to propose an investment strategy model. Experimental data from 2008 to 2015 are encoded based on quarter periods, resulting in thirty-two periods over eight years, as shown in Table 2. During these thirty-two periods, Taiwan’s stock market experienced five bull markets, five bear markets, and one market correction. Regardless of the market state, the method proposed in this study outperformed the market and the TWN50 91% of the time (29 out of 32 periods). During the bull market, this proposed method outperformed simple genetic algorithms, the average allocated capital strategy, TAIEX, and the TWN50
85% of the time (11 out of 13 periods). In the bear market, this proposed method defeated competitors 89% of the time (17 out of 19 periods). During the market correction, this proposed method outperformed all other benchmarks 90% of the time (9 out of 10 periods).

Table 2. Experiment results for January 2008–December 2015.

<table>
<thead>
<tr>
<th>Year</th>
<th>Average Return</th>
<th>1st Quarter</th>
<th>2nd Quarter</th>
<th>3rd Quarter</th>
<th>4th Quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>TWN50</td>
<td>1.07</td>
<td>−0.08</td>
<td>−4.39</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>TAIEX</td>
<td>1.00</td>
<td>−0.89</td>
<td>−4.21</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Return on average allocated capital</td>
<td>0.93</td>
<td>−0.69</td>
<td>−3.36</td>
<td>−0.34</td>
</tr>
<tr>
<td></td>
<td>SGA</td>
<td>1.62</td>
<td>2.50</td>
<td>−2.76</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Hybrid GA</td>
<td>4.23</td>
<td>4.76</td>
<td>1.18</td>
<td>4.65</td>
</tr>
<tr>
<td>2014</td>
<td>TWN50</td>
<td>0.67</td>
<td>0.71</td>
<td>−0.34</td>
<td>1.89</td>
</tr>
<tr>
<td></td>
<td>TAIEX</td>
<td>0.94</td>
<td>0.62</td>
<td>0.47</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>Return on average allocated capital</td>
<td>1.09</td>
<td>0.79</td>
<td>1.38</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>SGA</td>
<td>0.61</td>
<td>3.30</td>
<td>1.63</td>
<td>4.57</td>
</tr>
<tr>
<td></td>
<td>Hybrid GA</td>
<td>0.58</td>
<td>8.11</td>
<td>3.97</td>
<td>10.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Average Return</th>
<th>2011</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>TWN50</td>
<td>−0.67</td>
<td>−0.58</td>
</tr>
<tr>
<td></td>
<td>TAIEX</td>
<td>−1.02</td>
<td>−0.74</td>
</tr>
<tr>
<td></td>
<td>Return on average allocated capital</td>
<td>−1.49</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>SGA</td>
<td>1.76</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Hybrid GA</td>
<td>1.69</td>
<td>3.25</td>
</tr>
<tr>
<td>2014</td>
<td>TWN50</td>
<td>3.95</td>
<td>6.95</td>
</tr>
<tr>
<td></td>
<td>TAIEX</td>
<td>4.71</td>
<td>7.78</td>
</tr>
<tr>
<td></td>
<td>Return on average allocated capital</td>
<td>4.25</td>
<td>11.70</td>
</tr>
<tr>
<td></td>
<td>SGA</td>
<td>2.88</td>
<td>10.52</td>
</tr>
<tr>
<td></td>
<td>Hybrid GA</td>
<td>11.91</td>
<td>13.82</td>
</tr>
</tbody>
</table>

TAIEX = Market index returns (TAIEX), SGA = simple genetic algorithm (SGA). Hybrid GA = Hybrid genetic algorithm.

Figure 9 is divided into 11 sections, among which the method proposed in this study outperforms in 9 sections. Furthermore, in the 11 quarters, even when the TWN50, TAIEX, and average allocated capital strategy yield a negative rate of return, the proposed method still manages to achieve a 4.20% return on investment. With a total average rate of return of up to 6.00%, the proposed method exhibits the best performance. This confirms that the integration of genetic algorithms and the Ohlson value assessment model can significantly enhance the return on investment.

In 2010, the European debt crisis began to spread through the Eurozone, and Taiwan was also affected during this period. Table 3 presents the monthly returns obtained from various comparison methods from January 2010 to December 2010. According to the average rate of return from January 2010 to June 2010, the framework proposed in this study yielded a 9.51% rate of return. This performance is superior to the TWN50 market index returns, the return on average allocated capital, and the return produced by the simple genetic algorithm (SGA) evolution. The returns obtained from this study outperformed the TWN50’s average return in five out of six months. From July 2010 to December 2010, the average rate of return using this study’s framework also outperformed the TWN50 average return, although it was higher than the TWN50’s average return in only three out of six months. Nevertheless, the rate of return from other methods used in the experiments indicates that this study’s framework consistently yields better returns.
Figure 9 is divided into 11 sections, among which the method proposed in this study outperforms in 9 sections. Furthermore, in the 11 quarters, even when the TWN50, TAIEX, and average allocated capital strategy yield a negative rate of return, the proposed method still manages to achieve a 4.20% return on investment. With a total average rate of return of up to 6.00%, the proposed method exhibits the best performance. This confirms that the integration of genetic algorithms and the Ohlson value assessment model can significantly enhance the return on investment.

In September 2008, the bankruptcy of Lehman Brothers triggered a global financial crisis that severely impacted stock markets worldwide. The Taiwan stock market was not spared, suffering significant losses, with its value shrinking from 22.37 trillion to 12.65 trillion. The stock market index also fell below the 5000-point mark. Despite this challenging environment, the integrated genetic algorithm used in this study outperformed all benchmark methods in 2008. The architecture of this study achieved a 12.05% rate of return, surpassing all average rates of return from competing methods in the first half of the year. Although the integrated genetic algorithms experienced negative returns in July,
September, and October, the method proposed in this paper still achieved a six-month average return rate in the second half of the year that was 4.91% higher than the simple genetic algorithm (SGA), 10.54% higher than TWN50, 10.33% higher than the market index returns, and 10.28% higher than the return on average allocated capital. The proposed method still managed to yield a 2.75% rate of return. A comparison of the rates of return for each method is shown in Table 4.

Table 4. Experiment results for 2008.

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWN50</td>
<td>−10.64</td>
<td>9.78</td>
<td>0.58</td>
<td>4.54</td>
<td>−2.89</td>
<td>−10.50</td>
<td>−1.52</td>
</tr>
<tr>
<td>TAIEX</td>
<td>−11.58</td>
<td>11.86</td>
<td>1.90</td>
<td>4.05</td>
<td>−3.37</td>
<td>−12.71</td>
<td>−1.64</td>
</tr>
<tr>
<td>Return on average allocated capital</td>
<td>−3.76</td>
<td>17.20</td>
<td>4.16</td>
<td>0.57</td>
<td>−2.94</td>
<td>−15.95</td>
<td>−0.12</td>
</tr>
<tr>
<td>SGA</td>
<td>−1.56</td>
<td>16.32</td>
<td>4.22</td>
<td>4.25</td>
<td>3.09</td>
<td>16.24</td>
<td>7.09</td>
</tr>
<tr>
<td>Integrated GA</td>
<td>5.36</td>
<td>23.56</td>
<td>9.57</td>
<td>8.91</td>
<td>5.52</td>
<td>19.35</td>
<td>12.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWN50</td>
<td>−8.83</td>
<td>1.40</td>
<td>−15.77</td>
<td>−16.40</td>
<td>−10.23</td>
<td>3.10</td>
<td>−7.79</td>
</tr>
<tr>
<td>TAIEX</td>
<td>−6.64</td>
<td>0.31</td>
<td>−18.83</td>
<td>−14.84</td>
<td>−8.42</td>
<td>2.93</td>
<td>−7.58</td>
</tr>
<tr>
<td>Return on average allocated capital</td>
<td>−6.01</td>
<td>−7.45</td>
<td>−23.91</td>
<td>−12.64</td>
<td>3.40</td>
<td>1.43</td>
<td>−7.53</td>
</tr>
<tr>
<td>SGA</td>
<td>−10.10</td>
<td>5.43</td>
<td>−7.36</td>
<td>−11.39</td>
<td>7.92</td>
<td>−1.30</td>
<td>−2.80</td>
</tr>
<tr>
<td>Integrated GA</td>
<td>−5.99</td>
<td>11.17</td>
<td>−5.57</td>
<td>−2.01</td>
<td>13.18</td>
<td>5.69</td>
<td>2.75</td>
</tr>
</tbody>
</table>

The experimental results show that regardless of whether the market index rises or falls, the integrated genetic algorithms proposed in this study are capable of generating higher excess returns or minimizing losses. Even when compared with the top 50 professional investment funds, the integrated genetic algorithm consistently outperforms the returns of the TWN50. In any investment environment, the proposed method yields positive results. Apart from the genetic algorithm’s ability to make sound stock investment decisions (both in terms of selection and capital allocation), the integration of LSTM neural networks and the Ohlson model can accurately identify the optimal trading points, thereby addressing the issue of market timing decisions. By properly integrating selection, timing, and capital allocation decisions, investors can truly achieve superior excess returns. This study demonstrates the potential of using advanced AI techniques, such as genetic algorithms and LSTM neural networks, in enhancing investment strategies and outcomes.

5. Conclusions

Based on the past literature, few studies have simultaneously integrated the three strategies of stock market investment—selection, timing, and capital allocation. Most market timing strategies differ from previous ones that used technical indicators to determine the point of sale. Instead, this paper uses the Ohlson model from fundamental analysis to ascertain the true value of stocks. By combining LSTM neural networks and the Ohlson model, this study learns the nonlinear relationship between variables and the true value of the stocks. The calculated true value of stocks is then compared with the stock market’s price to serve as the basis for buy or sell decisions. This approach offers a more comprehensive and potentially more effective strategy for stock market investment.

To enhance the speed and efficiency of processing stock data, this paper also introduces an efficient genetic algorithm encoding method based on the concept of optimizing resource allocation through combination with repetitions. The goal is to effectively address the optimization of stock selection and capital allocation in stock investment. The experimental findings underscore that the utilization of genetic algorithms with optimization characteristics can refine capital allocation effectively. Employing this capital allocation configuration for stock investment yields returns surpassing those attained through diversi-
fication across targeted stocks, average returns of targeted stocks, market index returns, and returns yielded by the simple genetic algorithm. In a six-month average return competition, the method proposed in this study surpasses the average return of TWN50. These results firmly establish that the method presented herein indeed offers investors a superior basis for making investment decisions, yielding higher returns.

Author Contributions: Conceptualization, Y.-H.C.; methodology, Y.-H.C.; software, Y.-H.C. and C.-W.H.; validation, Y.-H.C.; resources, C.-W.H.; data curation, Y.-H.C. and C.-W.H.; writing—original draft preparation, Y.-H.C.; writing—review and editing, Y.-H.C.; project administration, Y.-H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: The study did not involve humans.

Data Availability Statement: The experimental data has been recorded in this research paper.

Conflicts of Interest: The authors declare no conflict of interest.

References


37. Solares, E.; De-León-Gómez, V.; Salas, F.G.; Díaz, R. A comprehensive decision support system for stock investment decisions. *Expert Syst. Appl.* 2022, 210, 118485. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.