Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v17n1p89

Predicting Warrant Prices Using an Artificial Neural Network Model: Experimental Comparison with Black Scholes Metron Model

By Tam et al.

15 March 2024

Electronic Journal of Applied Statistical Analysis
Vol. 17, Issue 01, March 2024, 89-114
DOI: 10.1285/i20705948v17n1p89

Predicting Warrant Prices Using an Artificial Neural Network Model: Experimental Comparison with Black Scholes Metron Model

Phan Huy Tam *[a], Nguyen Thi Hai Yen[a], Nguyen Minh Tu[a], Do Thi Lan Phuong[a], Nguyen Thuy Hong Thanh[a], Le Nguyen Phu Loc[a], and Pham Thi Kim Anh[b]

[a] *University of Economics and Law, Ho Chi Minh City, Vietnam and Vietnam National University, Ho Chi Minh City, Vietnam.*
[b] *Faculty of Finance and Accounting - University of Industry and Trade, Ho Chi Minh City, Vietam.*

15 March 2024

The main objective of this study is to build an artificial neural network (ANN) model to predict warrant prices in Vietnam with data collected from 2019 to 2021 from nearly 300 different warrants. The ANN model is applied on a case-by-case basis depending on the status of the ITM or OTM warrants to examine further the model's pricing performance of the proposed model's price relative to the actual warrant's price. In addition, to compare with the ANN model, the Black Scholes Merton (BS) model is also used for warrant pricing. The ANN model is built with structure of 3 hidden layers using ReLU activation and 1 hidden layer using Softplus activation. The research results show that the ANN model has a more significant error performance in the case of more significant data than in the other two cases. BS model, there is no specific conclusion that applying the model, in any case, will be more effective. Regarding performance comparison between the two models, the ANN model outperforms both the BS model.

keywords: ANN, Black Scholes, deep learning, machine learning, warrant prediction.

*Corresponding authors: tamph@uel.edu.vn

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

1.1 The urgency of the topic

Options or covered warrants have only been widely traded as an investment or hedging instrument in Vietnam from 2018 to the present. Usually, securities companies often issue warrants to promote securities trading between investors and securities companies. The form of warrants is similar to stock options with similar parameters such as exercise price, conversion rate, and time to maturity. In Vietnam, warrants have a reasonably short maturity, usually less than a year, and the price of warrants can vary from a few thousand to a few tens or even a few hundred thousand dongs. Moreover, like options, the leverage ratio of warrants is substantial. With just a tiny amount of capital, we can get a similar return as investing in stocks.

The problem is how to quickly and accurately value the warrant's value to the observed price in the market for investors. Here, valuation is understood as the model will try to learn the patterns of the market in giving the value of that warrant at a particular time and try to give the closest approximation to the market. The valuation of a home, for example, will include the home's intrinsic value and market fluctuations at that time. For traditional valuation models, although many different formulas exist for the type of financial asset that is warranted, the author has reviewed several previous research papers and decided to use Black Scholes as a reference for the ANN model. According to Abd Aziz et al. (2020), Black Scholes is a well-known and commonly used model in valuing assets such as warrants or options. This model has existed for a long time, around 1960-1970, and is still widely used and developed today in terms of research and actual trading in options pricing. Other assets are similar to warrants. In addition, the study of Aziz et al. (2015) shows the model's effectiveness for both European and American style options, so this is also one of the reasons why Black Scholes is currently the model. Taking the lead in valuing warrants in Vietnam when this asset class being traded in our country is of the European type.

At the same time, the Black Scholes model is said to be easier to use and has fewer computational resources than other models (Wu et al., 2012). In a recent paper, Frino et al. (2019) have shown insufficient evidence that the Black Scholes model is mispricing or that there are market biases in the pricing of options in Australia. Despite its success, the model still has some flaws in being constrained by assumptions such as constant volatility, which leads to bias when applying Black Scholes to different markets where volatility Implied volatility usually has a skew or smile-shape distribution (Funahashi, 2021). Therefore many experimentalists have used other models related to local volatility (LVM) and stochastic volatility (SVM) to overcome this, such as Cox and Ross (1976) for LVM or Heston's (1993) model for LVM. SVM. However, much analysis and computation are often required under these models because the computational cost is explosively significant (Funahashi, 2021).

From some stated assumptions of the Black Scholes formula, the author finds that this limitation can reduce the performance of the traditional model when valuing warrants in a highly volatile trading market. As well as, there always exist other external factors such as transaction fees,... Therefore, along with the development of technology and computers with higher computing capabilities and the increasingly available financial data sets and machine learning algorithms, the author is confident about building a guaranteed warrant pricing model with higher efficiency than the Black Scholes pricing model, specifically. An ANN machine learning model aims to provide another frame of reference for warrant investors in the market to refer to. Among the valuation models, the neural network model is considered a highly efficient model. Liu et al. (2019) have developed an artificial neural network model solution - ANN, to reduce the computation time of options pricing, especially for multi-dimensional financial models. The authors test the ANN method in three different solvers, including the closed-form solution for the Black-Scholes equation, the COS (Fourier-cosine series) method for the Heston model, and the Brent root method for degrees rough seas imply. Their numerical results show that ANN can efficiently and powerfully calculate options prices and implied movements. Culkin and Das (2017) show that a simple deep learning model can value options with low error. The grid architecture of this deep learning model can easily be extended to real-world pricing options without knowing the option pricing theory. New technology has popularised deep learning, so it is effortless for an investment manager or trader to implement these models.

Therefore, the ANN machine learning model and the traditional Black Scholes formula are the two objects studied and used in this study by the author. After going through the training process in the ANN model and conducting the warrant pricing in both models, the author will compare the results of the predicted warrant prices from ANNs and Black Scholes. Furthermore, this study provides a discussion of the differences between factors such as the average performance, error, or how well each model performs on each case of disaggregated warrant data.

1.2 The purpose of the topic

1.2.1 General objective

The main objective of this study is to build a warrant pricing model based on the ANN machine learning algorithm. Also, compare the performance between the two traditional and machine learning models to discuss further each model's effectiveness and appropriateness in each warrant state case. Investors can make accurate and faster reference prices and keep pace with the market.

1.2.2 Detailed objectives

From the general objective, the authors have determined the detailed objectives that need to be done as follows:

- Built an ANN machine learning model to quickly value warrants and give approximate results to market prices
- Compare the average performance of two ANN models and BS models based on error indexes. Then, continue to compare the performance of two ANN models

and BS models in three cases (complete data, ITM and OTM state)

- Evaluate the ANN model's effectiveness from the results compared with the BS model. Discuss the advantages and limitations of the ANN model. From there, propose new research directions in the future.

1.2.3 Research question

After considering the possible aspects and issues of the topic of the study, the authors posed the following questions:

- Question 1: Why choose the ANN model and the BS model for warrant pricing for the Vietnamese market?
- Question 2: Can the results of the ANN model show better performance than that of the BS model?
- Question 3: In each case of different data, is the performance of the ANN model and the BS model different?

1.3 Research subjects

The subjects of this study include: (i) ANN model and BS model when applying warrant pricing in Vietnam. (ii) The students of the University of Economics - Law, investors need to value warrants. (iii) The effectiveness of the warrant pricing model by ANN helps students and investors who want to price warrants in Vietnam.

1.4 Research scope

In terms of spatial scope, the research carried out is limited to the territory of Vietnam. Regarding the time range, the research data is collected over two years from 2019 - the time warrants are launched in Vietnam's stock market - to 2021. In this study, the data set for analysis, evaluation and model running will be collected from reliable and specialized information sites in the financial sector in general and securities in particular. At the same time, there are references from relevant documents and research articles of precise and scientific origin.

1.5 Research methods

The research is carried out using the ANN model and BS model and random parameters of the ANN model for warrant pricing. Methods of information collecting: including reading documents and collecting data from financial data sources. To evaluate the model's performance on the training set, the author uses indexes to measure the accuracy, such as MSE (loss function), RMSE, MAE, and MAPE, between the difference between the given model results and the actual historical price of warrants. To evaluate the model's performance on the test set, the author uses RMSE, MAE, MAPE, and R2.

Regarding the performance comparison on the test set, there will be two parts, the first part will compare the average performance of the two models, and the second part will compare the performance of each model in each case. Finally, the group will compare and conclude. The data includes price movement data for more than 294 covered warrants in Vietnam and warrant pricing parameters for 2019-2021. Data is collected from finance.vietstock.vn.

1.6 Research structure

This study is organized into five sections: Chapter 1 serves as the Introduction. Chapter 2 lays down the theoretical groundwork, encompassing the theory of option pricing, the principles behind the Black Scholes Merton model and the artificial neural network model, and an overview of past research. Chapter 3 discusses the data used in the study and the research methodology. Chapter 4 presents the findings and facilitates a discussion on them. Finally, Chapter 5 provides the conclusion and outlines potential avenues for future research.

2 Literature Review

2.1 Theorical framework

2.1.1 Option pricing theory

Options pricing theory estimates the fair value of an option that traders combine with their strategies to maximise profits or hedge risk. The main goal of option pricing theory is to calculate the probability that an option will be exercised when the option expires. Louis Bachelier first developed the theory in 1900. A payoff chart depicts the profit and loss position of a call option. Accordingly, an alternative will have two primary states: out-of-the-money (OTM) and in-the-money (ITM) status. An OTM position where the underlying asset price (S) is below the strike price (K) and the call option holder will not exercise their right. ITM state or a state where S is greater than (K) and the option is exercised. If S equals K plus the cost of buying the call (B), the investor will be in a breakeven position. And finally, if S is greater than B, then the investor will start to make a profit.

1.jpg

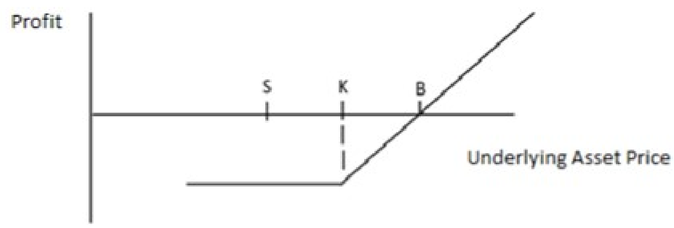


Figure 1: Bachelier's payoff chart for call options.

Bachelier needs to describe the distribution of the underlying asset's price to describe how options are priced. By modelling successive price changes in a particular way, he used the central limit theorem to determine the normal distribution of price movements. About sixty years later, Osborne (1959) produced similar results. Although subsequent studies by Mandelbrot (1963), Fama (1965), Brealey (1970), and others began to gradually reject the assumption of a normal distribution of stock prices, Bachelier's basic premise was conserved by Black and Scholes (1973) in their model with the belief that stock returns are log-normally distributed. Bachelier's description of the underlying asset's price assumed those price movements were similar and independent over time. This is an essential assumption in the basic formulation of the efficient market theory – if the prices of assets are priced efficiently. When prices change, it will also fully reflect new and independent information through time. This implies that stale information is useless in predicting the asset's price. From the above, we conclude that the factors that can affect the options price are the current price of the underlying asset and the exercise price, also known as the option's intrinsic value.

In the years that followed, many more studies on option pricing and the theory and model of Black and Scholes introduced in their research are still valid and famous today. According to Black and Scholes, in addition to the price of the underlying asset and the strike price, other factors can influence the price of an option, which we now call the option's time value. An option's time value includes time to maturity, annual fluctuations in the underlying asset's price, and the risk-free rate. Time to expiration is the number of days remaining for the option to expire, and it is proportional to the value of the option. Because the longer the option's expiration time, the longer the option will stay out of the market and the more likely it is that the underlying asset's price will exceed the strike price. The annual volatility of the underlying asset will be directly proportional to the value of the option because the higher the volatility of the underlying option, the more likely the underlying asset's price will exceed the strike price level is more extensive. Regarding the risk-free rate, it usually affects an option if the time to maturity is long otherwise the risk-free rate often acts as the discount rate on the option's strike price to the present to subtract the portion the investor receives from the underlying asset's price.

To sum up, option pricing theory has existed for a long time, researched and developed. The value of an option can be divided into two parts: intrinsic value and time value. The factors that make up the two parts of an option's value can affect the value of the option, but not the value of the option. The intrinsic value consists of the underlying asset's price and the option's exercise price. Time value includes factors such as the time to expiration of the option, the volatility of the underlying asset, and the risk-free rate.

2.1.2 Black Scholes Merton model

Different options pricing models were developed from Louis Bachelier's option pricing theory. Standard valuation models include the Black-Scholes-Merton model, the Monte Carlo simulation, and the binomial model. In particular, the current Black-Scholes-Merton model is still popularly and widely used.

Black-Scholes or Black-Scholes-Merton (BS) is a mathematical model applied to the pricing of several financial products, typically European-style options. The model was developed by Fischer Black and Myron Scholes and published in a 1973 paper (Black and Scholes, 1973). The model is considered an essential milestone in derivatives trading and the development of hedging strategies. It provides investors with safety with properly structured financial asset portfolios. Initially, the BS model was widely used to price European-style options and was widely accepted as a mathematical formula for option pricing. Before that, options traders used different and inconsistent formulas. The BS pricing model has been empirically analysed that the formula gives a value close to the observed values of the option price. Black and Scholes provided the first form of the model as individual differential equations. Robert Merton published a more mathematical method for easier understanding and deeper understanding of the model. And the formula after his research has the following form:

$$C(S_0, t) = S_0 N(d_1) - K e^{-rT} N(d_2) \tag{1}$$

Where:

- S_0 is the current price of the underlying asset.
- $C(S_0, t)$ is the price of the call option at S_0 and time t.
- K is the strike price of the option.
- T is the expiration time of the option.
- r is the annual risk-free rate.
- $N(d_1)$ and $N(d_2)$ are cumulative distribution functions for a random variable of a normal distribution.

The parameters d_1 and d_2 are calculated according to the following formulas:

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r + \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}} \tag{2}$$

$$d_2 = d_1 - \sigma\sqrt{T} = \frac{\ln\left(\frac{S_0}{K}\right) + \left(r - \frac{\sigma^2}{2}\right) T}{\sigma\sqrt{T}} \tag{3}$$

Where σ is the annual volatility of the stock.

The usefulness of the Black-Scholes formula is based on several assumptions about ideal market conditions as follows: Short-term interest rates are known in advance and do not change over time; The stock price follows a log-normal distribution; Because it is a European-style option, it can only be exercised at expiration; There are no transaction costs and taxes when buying or selling stocks or options; Short-selling is allowed to use the proceeds; No dividends are paid during the derivatives implementation; The process of securities trading is taking place continuously.

An option is theoretically priced according to the parameters of the stock price, the exercise price, the volatility of the underlying asset, the interest rate, and the time remaining to the option's expiration date. According to research by Black and Scholes (1973), in terms of stock prices and exercise prices, the fair value of a call option at a critical point in time is equal to the stock price minus the discounted value of the strike price at that time. Theoretically, the lower the current stock's price, which is lower than the strike price, the more likely that the call will not be exercised at expiration, and thus the value of the call will be zero. When assuming other factors such as stock prices remain unchanged, the longer the expiration date, the higher the actual value of the call option. The Black-Scholes model evaluates the likely price range of an option at expiration based on the current price of the underlying asset, the strike price, and other factors. It calculates two distinct probabilities - $S_0N(d_1)$ and $Ke^{-rT}N(d_2)$. $S_0N(d_1)$ represents the current price of the stock multiplied by the probability distribution function. The function returns zero if the stock price is below the strike price. If the stock price exceeds the strike price, the function returns a value representing the stock's potential value at the time of expiration. $Ke^{-rT}N(d_2)$ represents the strike price of the option adjusted for the time value of money (discounted to present value), multiplied by the probability distribution function for scenarios where the option's price exceeds the strike price at the time of expiration. In essence, the Black-Scholes model is designed to calculate the potential profit from an option, factoring in the cost of the option itself.

However, things will become relatively more complicated when valuing warrants for various reasons such as price adjustment from dividends, from company mergers, the company issuing more shares, or volatility is not uniform over the life of the warrant (Black and Scholes, 1973). The above problems have now been solved through adjusted formulas from the Black Scholes formula by subtracting the discounted dividend stream at the time i from the stock price or the diluted Black Scholes formula from stocks (Galai and Schneller, 1978), constant elasticity of variance (CEV) (Cox, 1975), Longstaff Extendible-Warrant formula (Lauterbach and Schultz, 1990). In addition to options, the Black Scholes formula is also widely used in pricing warrants. Among them, there are research papers by Lauterbach and Schultz (1990); Kremer and Roenfeldt (1992).

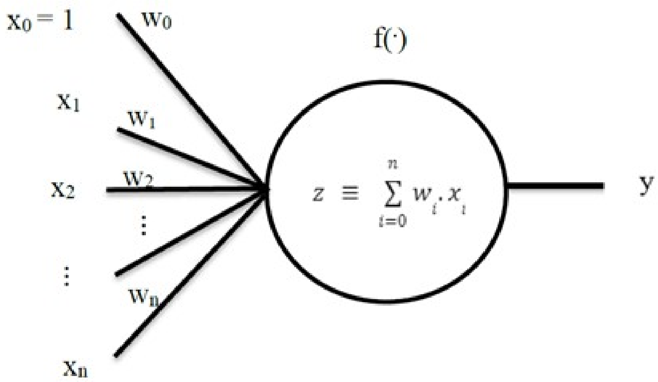
2.1.3 Artificial neural network model- ANN

Deep Learning is becoming more and more popular and vital in the financial sector. Deep Learning tools are supported in many programming languages, notably Python and R. As a result, the tools have enabled financial applications to achieve high levels of responsiveness thanks to viable trading algorithms. And it is often used to train models to predict outputs, classify securities or build credit risk models. One of the most widely applied models is its feasibility. and precisely the artificial neural network (ANN) model.

Over the historical periods, ANN has had many development steps based on the influences of the human brain structure. A human brain contains billions of neurons, and they are linked together to form a giant network known as a neural network. ANN is a mathematical model that simulates structure. In the 1940s, scientists initially mimicked the most straightforward neuronal modes to establish primitive ANNs. Neurophysicist

Warren McCulloch and mathematician Walter Pitts (1943) proposed a simple mathematical model of neurons to simulate the computing power of the human brain. Later, Rosenblatt (1958) presented a perceptron. This is the simplest model of an artificial neuron, but his approach can only solve the linear problem because the expression of the two layers is relatively poor. To overcome this limitation, Rumelhart et al. (1986) extended it to many classes, which can be learned by back-propagation to solve problems considered indivisible linear.

The model of an artificial neural network with more than two layers is called a multi-layer perceptron or multi-layer perceptron (MLP) model. In particular, MLP was developed based on one of the simplest and first machine learning models, the perceptron - a model used to solve binary classification problems. A general perceptron has the basic structure shown in figure 2.



2.jpg

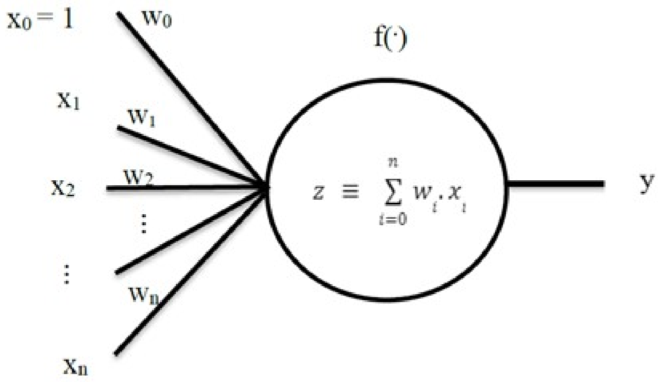
Figure 2: Components of an artificial neuron.

The above model is represented by the formula $y = f(z)$. In which, $z = \sum_{i=0}^{n} w_i x_i$ is the composite function, f is the activation function, n represents the number of input variables, x_i are the input values, y is the unique output value, w_i are the association weights, usually initialized randomly according to some distribution and are continuously updated during the learning process. At the first link weight w_0 the bias is given with $x_0 = 1$.

A feedforward neural network is an ANN model in which information moves only in the forward direction, i.e. from input nodes through hidden nodes to output nodes. The neurons in each layer will have corresponding weights fully connected to the neurons in the next layer. Each neuron in any network layer is connected to all neurons in the preceding layer to refine and optimise the prediction or classification. This type of network consists of many layers and is shown in Figure 2.

Typically, an ANN model will consist of three layers:

- The input layer will provide and receive input data to the network in vector form;



2.jpg

Figure 3: The basic structure of the ANN model.

- The hidden layer is in charge of processing the input data by performing multiplications with weights and summing the products, and then sending the results to the activation function;
- The output layer will return the output data and pass the information to the user.

In it, the activation function determines whether a particular neuron is active or not to decide whether to continue transmitting data and, if active, how much to transfer. In addition, the activation function limits the output of a neuron to a specific range of values. There are two types of activation functions f: linear and nonlinear functions. Because when MLP uses a linear process, there is always a single-layer network that is entirely equivalent. It is necessary to use a nonlinear function to take advantage of the multilayer network so that the model can be adapted to many types of data and generalizations. This is the difference between single-layer perceptron and multilayer perceptron.

The data is divided into two parts: training data and test data. The training set is the data on which the algorithms are applied to train the machine learning model. The test set is the data used to evaluate the model's accuracy and has a smaller amount of data than the training set. If the output of the ANN does not fit the test set, the weights will be adjusted using gradient descent. This method changes the weights iteratively through each training data to minimize the loss function. To apply the above method, it is necessary to calculate the derivative of the loss function according to each weight matrix w and deviation vector b. The process of adjusting the weights in each layer is repeated through the formula:

$$x_{t+1} = x_t - \eta f'(x_t) \tag{4}$$

where x_t is the point found after loop t, x_{t+1} is the point found after loop $t+1$, η is the learning rate and $f'(x_t)$ is the derivative at x_t. The sign $(-)$ here represents the opposite direction of x_{t+1} compared to the sign of $f'(x_t)$. It is very important to determine the input data x_t and the convergence rate η.

Besides, estimating the derivative value directly requires a lot of computational resources. This harms the algorithm's performance as it increases the time needed to process the data, especially with large data sets and high complexity. Therefore, most algorithms with artificial neural networks use backpropagation to overcome this problem. Suppose the feedforward process goes in one direction from left to right. In that case, the backpropagation helps calculate the derivative in the opposite direction from the last layer to the first layer. The previous layer gets the first derivative because it's closest to the output prediction and the loss function. The derivative of the weight matrices between the aforementioned classes and the deviation vector of the last layer will be calculated according to the normal series rule for the derivative of the composite function.

2.2 Research model review

2.2.1 BSM model

The valuation of warrants is a relatively widely studied topic, but the approaches of the studies are not uniform. The authors have used various models such as GARCH, Black–Scholes Merton, Brownian, Binomial, *etc.* Wu et al. (2012) proposed a method for valuing warrants when the underlying asset follows the GARCH dispersion model. However, because the GARCH dispersion model is non-linear, it consumes many computational resources when all must use simulation methods or numerical methods on differential equations. These processes are computationally intensive, and it is not feasible for large transaction books to be evaluated quickly and frequently. Therefore, the study had to use the perturbation method for the partial differential equation (PDE) of the characteristic function for the underlying asset price, transforming the non-linear PDE into an approximate linear PDE. An analytical approximation to the European option price is then derived using the Fourier transform. To estimate the parameters for the model, the study used the maximum likelihood estimation (ML) method, in which the likelihood function is determined by using a range-efficient importance sampling (EIS) Richard and Zhang (2007). Experimental results show that the GARCH dispersion model has higher initial accuracy than the Black–Scholes model. However, the GARCH dispersion model consumes too much time and resources in the computation process compared to the Black–Scholes model.

Aziz et al. (2015) discussed warrant pricing in Malaysia using the Black–Scholes model. Then, a new study on warrant pricing by binomial model and implied volatility was further developed by Abd Aziz et al. (2020). Both studies used warrant data listed on the Bursa Malaysia Exchange and randomly selected from the Mara University of Technology (UiTM) data line. The parameters related to warrant pricing covered in both papers are exercise price, interest rate, expiration date and volatility. After the experiment, the warrant price obtained from the model is compared with the actual price to check the accuracy and consistency of the model. The mean square values (MSE) obtained from the valuation summary table of 4 American-style warrants randomly selected to test the model are relatively low (0.074% - 0.98%). Abd Aziz et al. (2020) have shown that the warrant price from the Binomial model is almost the same as the actual price. Therefore, the Binomial model is one of the methods used for default warrant pricing, although it was developed for option pricing. In addition, Segara and Sagara (2007) argue that the basic features of warrants are similar to the options and obligations of the issuer for each share. Therefore, options pricing models can also calculate warrant prices because of their similar characteristics. However, the authors do not use the binomial tree model in this study because the model has some limitations in valuation. A binomial tree model is only helpful when valuing American-style warrants, while the author's research object is European-style warrants (Vietnamese warrants). Moreover, in the binomial tree model, the underlying property is only likely to be one of two outcomes, but in reality, the value of that property at any point in time can be any number.

Inheriting the results from previous studies, the authors decided to continue using the traditional Black–Scholes model to perform the warrant valuation in the Vietnamese market because this is a popular and variable model for option pricing. Moreover, the model can be used to price European or American options Aziz et al. (2015). Additionally, it can be seen that the Black–Scholes model is easier to use and requires fewer computational resources than other methods.

2.2.2 Machine learning model

Given the sensitive and volatile nature of financial market externalities, using non-parametric models that do not depend on pre-existing assumptions, such as machine learning, is expected to increase accuracy when determining the prices of financial products. Since the late 20th century, there have been many studies choosing machine learning models to evaluate or compare the performance from theoretical point of view (Poojary et al., 2023)and when predicting the prices of different asset classes from crude oil (Gabralla et al., 2013), real estate (Li et al., 2009), stocks (Kohara et al., 1997; Gururaj et al., 2019) or closest to the existing research are options prices (Hutchinson et al., 1994; Garcia and Gençay, 2000).

Choosing a suitable machine learning model is essential to improve the use-value of the property valuation formula that this thesis is building. It also depends on the efficient database, the ability to collect relevant information and the auxiliary algorithm used for the model (Alpaydin, 2020). In this study, in parallel with the traditional Black–Scholes valuation formula, the authors decided to use the artificial neural network (ANN) model to build and evaluate the performance when warrant pricing through the data of the Vietnamese stock market. Some studies show the potential computational efficiency of ANN compared to other types of machine learning algorithms. The research results of Kara et al. (2011) show that both models can be considered valuable predictors in predicting stock price movement, but ANN is superior with an average performance higher than 75.74% of the Support Vector Machine - SVM algorithm. Madhu et al. (2021) also used two models, SVM and ANN, to compare the performance of these algorithms when predicting SPY option prices from the 2015 training dataset. The results still favor the ANN model as more optimal than SVM, with minor RMSE error and predictions closer to the observed market price. In the same year, Ivașcu (2021) published a study on call option pricing performance among machine learning models including ANN, Support Vector Regression, Genetic Algorithms, Decision Tree variant models Random Forest, XGBoost, LightGBM with two classical methods Blackscholes and Corrado-su according to historical volatility and implied volatility. The results show that the machine learning models are better than the traditional models. ANN is only behind XGBoost and LightGBM overall, with not much difference.

As one of the most commonly used algorithms in machine learning, ANN has been chosen for use in many valuation performance evaluation studies. Hutchinson et al. (1994) researched to realize the ability of non-parametric network model training for option pricing. The same author has compared 3 ANNs algorithms, including Multilayer Perceptron (MLP), radial-basis function network (RBF) and projection pursuit (PPR) regression, with the traditional Black–Scholes formula. Using two input data variables, the underlying asset's price normalized to the option's exercise price and the expiration time, the ANN artificial neural network model gives a prediction result closer to the option's price. The observed market and absolute error are lower than the Black–Scholes formula with historical volatility. According to Yao et al. (2000), using a backpropagation neural network to price options on the Nikkei 225 index is more optimal than Black–Scholes because financial markets are always volatile and changing and not the same as the specific assumptions made. Estimating a general option pricing formula with a functional shape similar to the conventional Black–Scholes formula using a feedforward neural network model is the work of (Garcia and Gençay, 2000). They used two input variables similar to Hutchinson et al. (1994) for ANN but divided into two variants: the original ANN and the other is a modified model with a calculation formula called ANN With Hint.

Besides, instead of choosing another machine learning model as a comparison reference system, some studies choose to exploit the ANN algorithm more deeply and then make post-improvement conclusions. Amilon (2003) relied on the MLP model from Hutchinson et al. (1994) and further developed it compared with the Black–Scholes formula with historical and implied volatility. Also, they used both the bid and ask price of the call option on the Swedish stock exchange as input variables instead of the closing price or the average price. The original MLP model of Guresen et al. (2011) was compared with two mixed versions called Dynamic Artificial Neural Networks (DAN2) and Hybrid Neural Networks, which further incorporated the GARCH formula. The conclusion shows that MLP has more reliable predictive results when predicting stock prices with minor absolute deviation MAD and mean error MSE than the two models.

Overview of the reviewed studies, the author found that the ANN artificial neural network model is a widely used machine learning algorithm in the topics of asset valuation. The model's price prediction results are better than the traditional parameter formulas. Thanks to its ability to make predictions based on training from real historical data sets, ANNs do not have to depend on a definite number of variables. They can better adapt to changes in the market. Moreover, the ability to learn, process and compute data of ANN has shown superiority compared to other machine learning algorithms. This network formula is constantly being improved in training methods model training, so this is also a potential model used by the author to evaluate warrant pricing performance.

2.2.3 Comparison of ANN and BS models

For option pricing, it has become quite common to use the ANN model to compare with a benchmark such as the Black Scholes formula. Theoretically, artificial neural networks are believed to overcome the characteristic depending on certain variables or specific assumptions of the parameter formula (Hutchinson et al., 1994). The author finds that the assumptions made make the Black Scholes formula less applicable in fundamental financial markets with high volatility. The research results of Yao et al. (2000) also highlight the possibility of Black Scholes option pricing in a stable and relatively perfect market, while ANN has more potential for the market and lots of volatility as well as options with an extensive range of price movements. Therefore, ANNs are increasingly used to evaluate when the predictions are the result of training the model from the actual data in the past to find the non-parametric relationship between the models' input and output variables so that errors are significantly reduced, and a more accurate pricing formula is obtained for a volatile market (Obthong et al., 2020).

Recent studies have also shown the advantages of the ANN machine learning model compared to the traditional Black Scholes model. According to Hutchinson et al. (1994), although the difference is not too clear, the results of ANN networks, especially MLP, are closer to the actual observed price. The error when predicting the risk performance is also smaller than Black Scholes. The author argues that network formulas are a more practical alternative when the price dynamics of the underlying asset are unknown, or it is not known exactly what factors will lead to the price difference. Network models will have several significant advantages over traditional parametric models because they do not rely on restrictive parametric assumptions such as predictability or sample-path continuity. Second, they adapt and react to structural changes during data generation in a way that parametric models do not. Finally, they are flexible enough to include various derivatives and underlying asset price dynamics.

Amilon (2003) show that the ANN model, specifically MLP, has higher performance in valuation and hedging. Compared with Black Scholes with historical volatility, MLP's RMSE is consistently lower across price segments. For OTM options, Black Scholes often overvalues the actual price, deviating wildly from the observed data in the market. Compared with Black Scholes with implied volatility, neural networks typically price ITM options slightly higher, while Black Scholes price them lower. In general, except for ITM bids, the mean squared error of ANNs is lower than that of Black Scholes.

The overall results are similar to the study by Gençay and Salih (2003). They demonstrate that Black Scholes is not suitable for pricing in highly volatile markets, especially for predicting prices for options, both buy and sell options in the OTM group. According to Stark (2017), although the ANN is not over-predictive for short-term options, the longer the expiration time, the more stable the prediction results. The prediction results of ANN in this study are the best in all long-term options. When divided by moneyness, the group with the slightest error is OTM, then ITM. Whereas Black Scholes is only more effective in valuing short-term options and is very likely to overvalue OTM options, this is the same result of some of the studies mentioned, such as Gençay and Salih (2003); Amilon (2003).

From the results of the above studies, the artificial neural network model has shown a superior predictive potential compared to the Black Scholes model. The author can build a model to predict the warrant price with a minor error, thanks to ANN. Compared to Black Scholes, ANN ensures a more stable long-term accuracy for prediction results, avoiding maximally over-priced prediction cases, especially for OTM warrants. Through judgments based on finding out the relationship between variables, ANN can be developed into a formula for valuing contracts with a more accurate and faster reference price. To improve the objectivity of the comparison, the research team uses the input data of the ANN model similar to the data variables in the formula of the Black Scholes model, including stock price (S), stock price exercise (K), volatility of the underlying (σ), the risk-free rate (r), and the time remaining to the expiration date of the warrant (T). After going through the training process from the given data, the model will provide outputs, including the forecasted warrant prices, and then evaluate the error of the ANN model and the Black Scholes formula when comparing the same actual price on the data set.

3 Methodology

3.1 Data

This study uses data on Vietnam warrants collected from finance.vietstock.vn for valuation by machine learning model and compares with the traditional Black Scholes method. The dataset includes 31,497 observations. Table 1 provides information on descriptive statistics of 294 warrant codes collected from October 6, 2019, to June 7, 2021, with different maturities from 3 months to 11 months. And The number of trading days (TradingDays) is at least 39 days, at most 214 days. Data used for the warrant pricing machine learning model and the Black Scholes model

Table 1: Description of warrant data

	Maturity (Months)	TradingDays
No of obs	294	294
Mean	5.629	102.031
Standard Deviation	2.039	43.161
Min	3	39
Median	6	103
Max	11	214

3.2 Model building

Many empirical studies have shown the effectiveness of random search in optimising parameters in artificial neural network models (Bergstra and Bengio, 2012; Greff et al., 2017; Wang et al., 2019). The parameters in a neural network model include the number of times the data is passed through the network to optimise the loss function (epoch), about 10 to 1000; neuron activation functions; the number of hidden layers; the number of neurons in each layer from 2 to 1000. In the backpropagation process to optimise the loss function, Adam's algorithm - an optimisation algorithm that combines RMSprop (friction) and Momentum (momentum), was used (Kingma and Ba, 2014). The Adam algorithm is widely used and proven to be quite effective with deep learning network models (Kingma and Ba, 2014; Reimers and Gurevych, 2017; Wang et al., 2019). With the support of TensorFlow and Keras - popularly used open-source Machine Learning libraries, the model is built by the team according to the following structure.

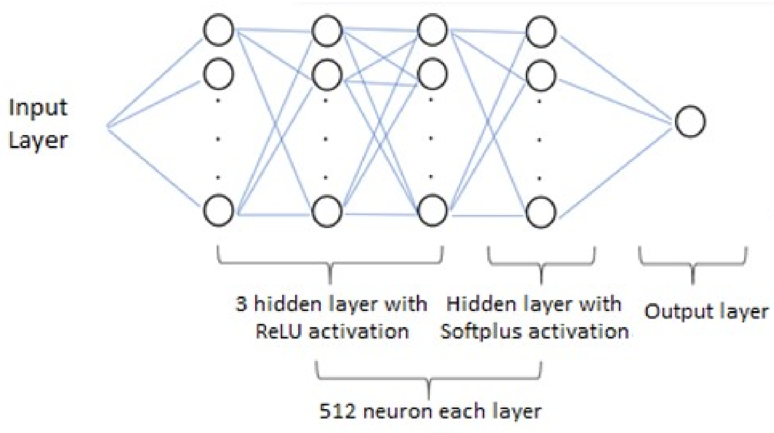
The model consists of 1 input data layer (input layer) with some data features of 6, including features taken from the BS model and adding a conversion rate; 4 hidden layers; and an output layer. Each hidden layer contains 512 neurons, and the team uses the Rectified Linear Unit - ReLU activation function to transform the calculated result data. The ReLU function is very effective in training neural network models because it helps the model train and optimise the loss function faster (Krizhevsky et al., 2012). Mainly, the output data of the hidden layer is finally used with the Softplus activation function referenced from Itkin (2019) to produce results in the form of real numbers greater than 0, ensuring compliance with the warrant price conditions is not negative. The ReLU and softplus functions have the following formula:

$$\text{ReLU:} \quad y = \max(0, z) \tag{5}$$

where $z = \mathbf{w}^T.\mathbf{X}$ and $\mathbf{w}$ is the weight matrix and $\mathbf{X}$ is the feature matrix.

$$\text{Softplus:} \quad y = \log(1 + e^z) \tag{6}$$

The data will be divided into two parts, the data part for training and the data part for testing after training is complete (test). During the training process, to avoid overfitting the model, the group also divides the training set into two more small sets, the leading training set and the validation set at the rate of 10% training set and the Early Stopping technique. Early Stopping technique is used to monitor the loss value on the reinforcement set to avoid the model learning too deeply on the training set (the loss value is minimal) while on the reinforcement set, what has been known is applied false again (loss value increases). Suppose it finds that the loss value on the reinforcement set does not tend to decrease after a few epochs. The model will automatically stop training and save the weighting value between the classes, making the loss function on the reinforcement set the minimum to a file for later use on the test set. In addition, in each iteration (epoch), putting too much data in at the same time will lead to a significant slowdown in the computational model, so the team uses a smaller number of



4.jpg

Figure 4: Artificial neural network model .

random samples taken from the training set (batch size) in each epoch to include in the model is 64.

To construct the loss function between the model output versus the actual value of the warrant price, the team decided to use Mean square error (MSE) in the back-propagation process. This loss function is commonly used with many problems, especially with valuation problems, value prediction, or regression problems. The MSE formula is given below.

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(\mathrm{Predicted}_i - \mathrm{Actual}_i)^2}{n} \tag{7}$$

where n is number of observations.

3.3 Research process

Split the dataset: Then divide the above data into two parts, ITM and OTM. The data section at ITM has 13,953 observations. The data section at OTM has 17,433 observations. Each data set will then be divided into two random training and testing sets (train and test) with a training rate of 70% and a test rate of 30%. A reinforcement data set will be extracted a small part from the training set to help the model check again to avoid overfitting. The training set will include the input variables x of the machine learning model: the risk-free rate, the closing price of the underlying asset, the option strike price, the number of days remaining, and the annual stock volatility. Daily stock price volatility, warrant conversion rate, HOSE matching volume. And an output variable y is the closing price of the warrant. The test set also has the same input and output variables.

Standardize the data: During training, the model will use algorithms in the Gradient branch to optimise the weights between classes in the model, thereby minimising the loss function. The general Gradient algorithm is to calculate the derivative of the loss function, then let the variable x go in the opposite direction to the derivative sign to move towards 0. However, in machine learning models, activation functions are often used. Machine learning models have high derivative sensitivity for data with too large numbers. When the data exceeds a specific limit, the derivative will be zero, the Vanishing Gradient phenomenon will occur, and the model will not learn well. So before putting data into the model, we need to structure and normalise the data using different methods. In this article, my team used Standard Scaler. Before entering the model, the team used a standardisation technique called Standard Scaler. Standard Scaler is a method of normalising data. It will recalculate that observation by subtracting it from the data set's mean and dividing it by the standard deviation. The formula is given by:

$$Z = \frac{\chi - \mu}{\sigma} \tag{8}$$

Where,

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{9}$$

and,

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2} \tag{10}$$

Train the model: Next, the research team will use the Early Stopping technique to prevent the model from overfitting, causing the model to learn by rote and make incorrect predictions on the validation data set. After training the machine learning model, the model's weight will be saved for evaluation on the test set. After the training is complete, the training set results will be evaluated using the functions MSE, RMSE, MAE, and MAPE. Use the functions RMSE, MAE, MAPE, and R2 to evaluate the results for the test set.

Valuation of warrants by machine learning model and BS. Performance evaluation: The saved model will evaluate the CW on the test set and compare the results for y of the test set and y of the Black Scholes model using indicators like RMSE, MAPE, MAE, and R2. In which RMSE and MAE have units of Vietnam dong (VND), MAPE and R2 have units of percentage. Accordingly, the model with lower RMSE, MAPE, MAE and higher R2, that model is better. The formula are as followed:

$$\text{Root Mean Squared Error (RMSE):} \quad RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\text{Actual}_i - \text{Predicted}_i}{\text{Actual}_i}\right)^2}$$

$$\text{Mean Absolute Percentage Error (MAPE):} \quad MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\text{Actual}_i - \text{Predicted}_i}{\text{Actual}_i}\right|$$

$$\text{Mean Absolute Error (MAE):} \quad MAE = \frac{1}{N}\sum_{i=1}^{N}|\text{Actual}_i - \text{Predicted}_i|$$

$$\text{R-squared (R2):} \quad R^2 = 1 - \frac{\sum_{i=1}^{N}(\text{Actual}_i - \text{Predicted}_i)^2}{\sum_{i=1}^{N}(\text{Actual}_i - \text{Mean}_i)^2}$$

4 Results & Discussion

4.1 Descriptive statistics

Table 2 provides descriptive statistics about the independent variables, dependent variables and control variables for the valuation machine learning model and the Black Scholes model, including risk-free interest rate (r), the closing price of the asset underlying asset (S_0), the option strike price ($C(S_0, t)$), days remaining (T), annual stock volatility (σ), warrant conversion rate (CR - conversion rate), the closing price of the warrant (C).

Table 2: Summary of variables used in the model

	Mean	Standard deviation	Min	Median	Max
r (%)	0.833	0.756	0.259	0.413	3.217
S_0 (VND)	57,739.292	37,555.921	2,090	48,150	176,600
K (VND)	56,373.955	39,277.572	7,227	42,000	173,137
T (day)	85.228	62.429	2	72	308
σ	0.322	0.070	0.147	0.326	0.610
CR	4.721	4.475	0.820	3	20
C (VND)	3,752.569	5,251.150	10	1,950	51,400

In addition to the variables available in the Black Scholes model, such as the risk-free rate, the closing price of the underlying asset, the option strike price, the option's expiration date, and the annual stock volatility, the group's study uses the conversion rate variable of warrants to apply to the machine learning model. Since the warrant's price is after multiplying the conversion rate, if this variable is missing, the model will not recognise it and lead to errors in the prediction process.

The dataset consists of 31,197 observations. The risk-free rate ranges from 0.259% to 3.217%. The closing price of the underlying asset with the minimum price from 2,090 (VND) to 176,600 (VND). The exercise price of the call option ranges from 7,227 (VND) to 173,137 (VND). The number of days waiting to maturity ranges from 2 to 308 days. The stock fluctuates in a range of 14.7% to 61% annually. The warrant conversion rate is from 0.82 (times) to 20 (times). The closing price of warrants ranges from 10 (VND) to 51,400 (VND).

Table 3: Correlation matrix between variables

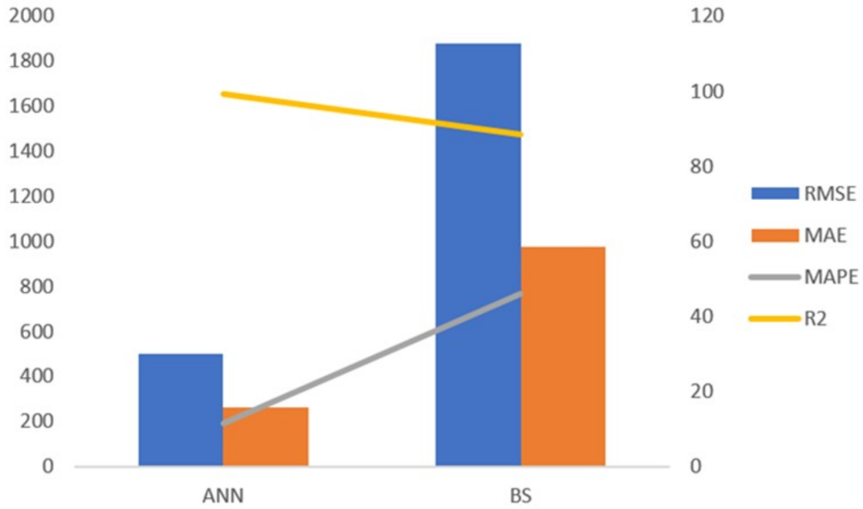
	r	$S0$	K	T	σ	CR	C
r	1						
S_0	-0.046	1					
K	0.057	0.920	1				
T	0.064	-0.061	0.017	1			
σ	-0.567	-0.309	-0.377	-0.071	1		
CR	-0.176	0.649	0.666	0.027	-0.130	1	
C	-0.081	0.133	-0.066	-0.077	0.060	-0.296	1

Table 3 reflects the correlation of variables through the correlation coefficient. The analysis results show that the correlation between the warrant closing price variable ranges from -0.081 to 0.133. The positive correlation between C and $S0$ shows that the warrants will have a higher price when the price of the underlying asset increases. It is worth noting that $S0$ is highly and positively correlated with K, indicating that the underlying asset price will increase the exercise price.

4.2 Model results

Between the ANN model and the BS model, we see that the ANN model gives a lower error level than the BS model on all four indexes. Specifically, we see that ANNs with RMSE are about 4-5 times lower than BS (400-500 compared to 1800-1900). ANN also outperforms in MAE and MAPE compared to BS. When ANN's MAE is only in the range of 200 and BS is in the field of 900, ANN's MAPE is less than 20%, while BS is higher than 40%. ANN's R^2 also gives better results, explaining the volatility of warrant prices at about 99% compared to only 86 to 89% of the BS model.

Table 5 on the next page shows the detailed results of the effectiveness of the ANN model and the BS model in each case. Overall, we can see that the ANN model gives



5.jpg

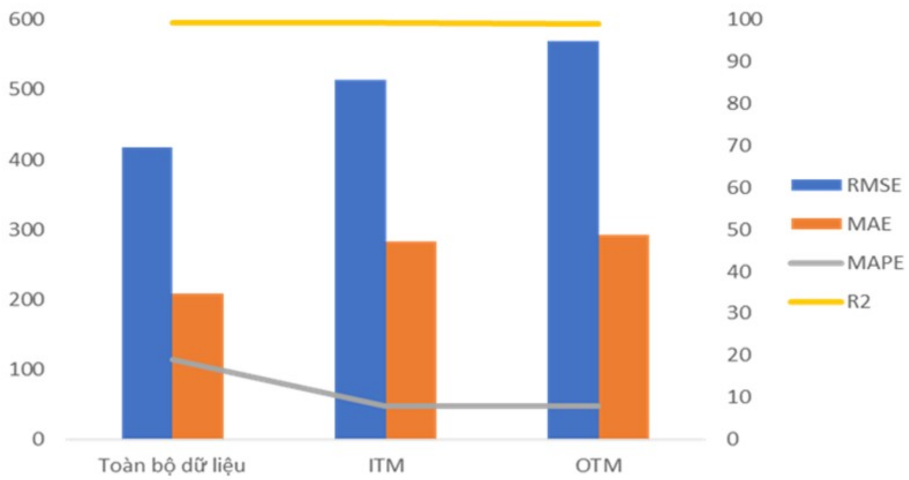
Figure 5: The average efficiency of the ANN model and BS model

better error results than the BS model. Specifically, in the case of all data, the ANN model provides RMSE and MAE indexes of VND 417.24 and VND 208.23 compared to VND 1,913.64 and 980.99 of the BS model. The other two indexes are MAPE and R^2 of the ANN model in case the whole data also gives better results when the MAPE of the ANN model is only 18.92% while that of the BS model is 46.33%. And the coefficient of determination R^2 of the ANN model is 99.37% compared to only 86.73% of the BS model. The remaining two cases of the ANN model also give similar results when the error indexes at absolute values are 3-4 times lower, and the error indexes in percentage form are also better than those of the other two cases with the BS model.

	RMSE (VNĐ)		MAE (VNĐ)		MAPE (%)		R2 (%)	
	ANN	BS	ANN	BS	ANN	BS	ANN	BS
Complete data	417.24	1,913.64	208.23	980.99	18.92	46.33	99.37	86.73
ITM	513.69	1,870.61	283.36	965.30	7.87	45.59	99.32	89.15
OTM	569.89	1,855.55	292.65	973.53	7.91	45.71	99.17	89.32

Table 4: Evaluation results on the training dataset after training

We can see that the results are consistent with the pre-comparisons between the ANN model and the BS model. Particularly in the pricing of warrants, the ANN model is superior to the BS model in that it does not follow the assumption that the underlying asset's price movement is constant. Moreover, the ANN model can also learn those volatility characteristics to adjust the model accordingly and provide a more accurate valuation. In addition, for artificial neural network models in general, before applying the model to the test data to verify its effectiveness, the model must undergo training on the training data trained to be able to adjust the model so that it fits the data the best. The error is the smallest on the test data set. So that difference also helps the ANN model give better results than the BS model.

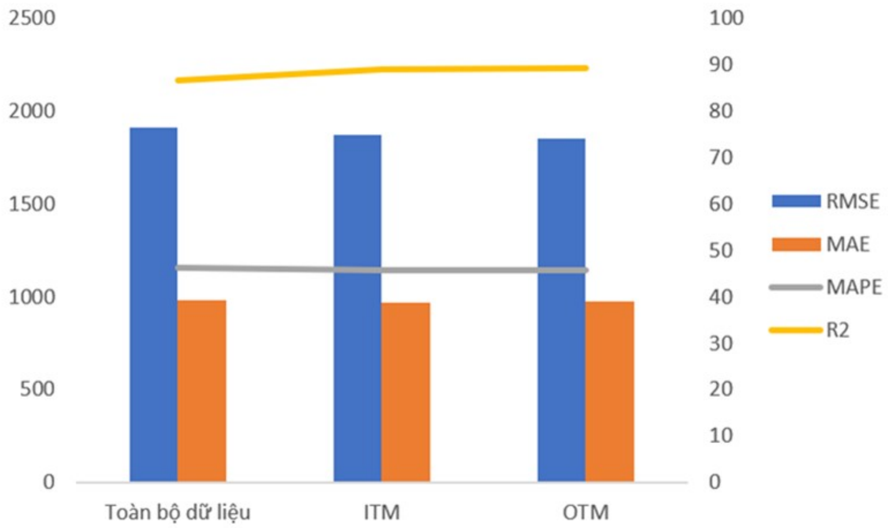


6.jpg

Figure 6: Evaluation of the ANN model in 3 cases

After the training is complete, the model with the best results will be saved and used on the test set of each warrant case that, according to the author, has divided all data (case 1), the warrant has ITM status (case 2), and warrants have OTM status (case 3). When using the ANN model in case 1, we see that the error indexes are superior and better than the other two cases except for MAPE and the reason explained in the training results section. Specifically, the RMSE of the ANN model in case 1 is worth about 400 VND, while case 2 is worth more than 500 VND and case 3 is nearly 600 VND. The MAE of the ANN model in the training case on the whole data is lower than in the other two cases. The coefficient of determination R^2 in all 3 cases is about 99% equal.

Regarding the BS model on the test set divided from the data set according to 3 warrant cases, the error criteria of the BS when used in case 1 are lower than in the other 2 cases. Specifically, the RMSE of BS case 1 has an error of 1913.64 VND compared to 1855.55 VND and 1870.61 VND of OTM and ITM. The MAE of BS case 1 has an error of about 20-30 VND more than case 2 and case 3 (980.99 compared to 965.30 and 973.53). The MAPE and R^2 of case 2 and case 3 are not much different at around 45%.



7.jpg

Figure 7: Evaluation of BS model in 3 cases

5 Conclusions & Recommendations

This study embarked on this project to scrutinize the effectiveness of the Artificial Neural Network (ANN) machine learning model and the traditional Black-Scholes (BS) model in pricing warrants. Through our experiments, we observed that the ANN model outperforms the BS model in terms of the error metrics such as RMSE, MAE, MAPE, and R^2, thereby suggesting that the ANN model could serve as a more reliable tool for warrant pricing.

A key attribute of our study was the decomposition of data into three distinct cases: complete data, in-the-money (ITM), and out-of-the-money (OTM). This approach allowed us to discern the performance of each model under these different scenarios, providing valuable insights for investors to make informed investment decisions.

The ANN model was chosen for its proven efficacy in asset valuation projects, superior computational abilities, and the ability to handle highly variable data, particularly when normalised. The BS model, on the other hand, is a commonly used method for warrant pricing due to its simplicity and minimal computational demands.

While the ANN model showed promising results, it is essential to note its limitations. The model's performance was less consistent across different cases, which suggests the need for caution when applying the ANN model to specific situations such as ITM or OTM warrants. On the contrary, the BS model maintained steady results across all cases, though its performance was generally inferior to the ANN model.

Despite the promising results, our study has its limitations. The data used was exclusively from Vietnam, which could potentially bias the model's performance when applied to other countries. Practical aspects such as transaction costs, hedging of financial investments, and additional variables were recognised but not resolved. Furthermore, the focus was solely on the ANN algorithm, thereby neglecting other potential machine learning models with superior performance.

Recommendations for future research include the exploration of other machine learning algorithms, expansion of the research data to encompass other countries, and the incorporation of real-world factors such as transaction costs into the model. We believe that these directions can yield further valuable insights and advances in the field of finance and warrant pricing. Ultimately, our hope is that this study forms a foundation for more comprehensive and reliable models for warrant pricing in the future, assisting investors in making informed decisions and mitigating financial risk.

References

Abd Aziz, K. A., Mohd Idris, M. F. I., Wan Daud, W. S., and Mazlan, M. N. (2020). Pricing warrant by using binomial model: comparison between historical and implied volatility. *Borneo Akademika*, 4(2):13–21.

Alpaydin, E. (2020). *Introduction to machine learning*. MIT press.

Amilon, H. (2003). A neural network versus black–scholes: a comparison of pricing and hedging performances. *Journal of Forecasting*, 22(4):317–335.

Aziz, K. A. A., Idris, M. F. I. M., Saian, R., and Daud, W. S. W. (2015). Adaptation of warrant price with black scholes model and historical volatility. In *AIP Conference Proceedings*, volume 1660. AIP Publishing.

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

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654.

Brealey, R. A. (1970). The distribution and independence of successive rates of return from the british equity market. *Journal of Business Finance*, 2(2):29–40.

Cox, J. (1975). Notes on option pricing. *Unpublished draft.*

Cox, J. C. and Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of financial economics*, 3(1-2):145–166.

Culkin, R. and Das, S. R. (2017). Machine learning in finance: the case of deep learning for option pricing. *Journal of Investment Management*, 15(4):92–100.

Fama, E. F. (1965). The behavior of stock market prices. *Journal of Business*, 38:34–105.

Frino, A., Lodh, S., and Khan, E. (2019). The black scholes call option pricing model and the australian options market: Where are we after 15 years. *The International Journal of Accounting and Business Society*, 1(1):40–57.

Funahashi, H. (2021). Artificial neural network for option pricing with and without asymptotic correction. *Quantitative Finance*, 21(4):575–592.

Gabralla, L. A., Jammazi, R., and Abraham, A. (2013). Oil price prediction using ensemble machine learning. In *2013 International Conference on Computing, Electrical and Electronic Engineering (ICCEEE)*, pages 674–679. IEEE.

Galai, D. and Schneller, M. I. (1978). Pricing warrants and the value of the firm. *Journal of Finance*, 33(5).

Garcia, R. and Gençay, R. (2000). Pricing and hedging derivative securities with neural networks and a homogeneity hint. *Journal of Econometrics*, 94(1-2):93–115.

Gençay, R. and Salih, A. (2003). Degree of mispricing with the black-scholes model and nonparametric cures. *Economics and Finance. Annals*, 4:73–101.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., and Schmidhuber, J. (2017). Lstm: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2222–22232.

Guresen, E., Kayakutlu, G., and Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 38(8):10389–10397.

Gururaj, V., Shriya, V., and Ashwini, K. (2019). Stock market prediction using linear regression and support vector machines. *Int. J. Appl. Eng. Res*, 14(8):1931–1934.

Hutchinson, J. M., Lo, A. W., and Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The journal of Finance*, 49(3):851–889.

Ivașcu, C. F. (2021). Option pricing using machine learning. *Expert Systems with Applications*, 163:113799.

Kara, Y., Boyacioglu, M. A., and Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. *Expert systems with Applications*, 38(5):5311–5319.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *CoRR*, arXiv:1412.6980.

Kohara, K., Ishikawa, T., Fukuhara, Y., and Nakamura, Y. (1997). Stock price prediction using prior knowledge and neural networks. *Intelligent Systems in Accounting, Finance & Management*, 6(1):11–22.

Kremer, J. and Roenfeldt, R. (1992). Warrant pricing: Jump-diffusion vs.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

Lauterbach, B. and Schultz, P. (1990). Pricing warrants: An empirical study of the black–scholes model and its alternatives. *Journal of Finance*, 45(4).

Li, D.-Y., Xu, W., Zhao, H., and Chen, R.-Q. (2009). A svr based forecasting approach for real estate price prediction. In *2009 International conference on machine learning and cybernetics*, volume 2, pages 970–974. IEEE.

Liu, S., Oosterlee, C. W., and Bohte, S. M. (2019). Pricing options and computing implied volatilities using neural networks. *Risks*, 7(1):16.

Madhu, B., Rahman, M., Mukherjee, A., Islam, M., Roy, R., and Ali, L. (2021). A comparative study of support vector machine and artificial neural network for option price prediction. *Journal of Computer and Communications*, 9(5):78–91.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36:394–419.

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133.

Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., and Wills, G. (2020). A survey on machine learning for stock price prediction: Algorithms and techniques. pages 63–71.

Osborne, M. F. (1959). Brownian motion in the stock market. *Operations Research*, 7:145–73.

Poojary, S. et al. (2023). A hybrid regression model for improving prediction accuracy. *Electronic Journal of Applied Statistical Analysis*, 16(3):784–801.

Reimers, N. and Gurevych, I. (2017). Optimal hyperparameters for deep lstm-networks for sequence labeling tasks. *CoRR*, arXiv preprint arXiv:1707.06799.

Richard, J.-F. and Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141(2):1385–1411.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.

Segara, L. and Sagara, R. (2007). Intraday trading patterns in the equity warrants and equity options markets: Australian evidence. *Australasian Accounting, Business and Finance Journal*, 1(2):42–60.

Stark, L. (2017). Machine learning and options pricing: A comparison of black-scholes and a deep neural network in pricing and hedging dax 30 index options.

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

Wu, X., Ma, C., and Wang, S. (2012). Warrant pricing under garch diffusion model. *Economic Modelling*, 29(6):2237–2244.

Yao, J., Li, Y., and Tan, C. L. (2000). Option price forecasting using neural networks. *Omega*, 28(4):455–466.