Advances in Economics, Business and Management Research, volume 215
Proceedings of the 2022 7th International Conference on Social Sciences and Economic Development (ICSSED 2022)

Application of Machine Learning in Option Pricing: A Review

Wenda Li[1,*]

[1]Graduated from University of Manchester, Manchester, UK
[*]Corresponding author. Email: ace_wenda_li@hotmail.com; 2015011477@student.cup.edu.cn

ABSTRACT

Options occupy a certain position in the derivatives market. Researchers, speculators, and other traders all hope to get a reasonable price for each option. There are only limited options that we can get an accurate solution to the price, and most options we need to get the price numerically. The classical method is poor in processing large data sets and high-dimensional data, and the calculation is slow. With the development of artificial intelligence in recent years, such as machine learning methods, optimization of target values has become easier and easier. So, some scholars, investors and traders began to care the application of artificial intelligence to different kinds of option pricing. This article is a review of the use of different methods in the pricing of different options in the past years and compare the pros and cons of different methods on accuracy and robust.

Keywords: *Machine Learning, Option Pricing, Deep Learning, Neural Network.*

1. INTRODUCTION

With the rapid development of economy, the scale of the financial market is also rapidly expanding, more funds have begun to flow into the financial market, and various financial derivatives have also been produced. *Financial derivatives* are contracts that derive their value from the performance, mainly price, of underlying assets, which play a very important role in the market. In the financial market, *options* are a type of financial derivatives with very high frequency transactions. Their trading volume is huge, exceeding 50% of the total global derivatives trading. The holders of options, by paying a specified amount of option *premiums*, has the right to choose whether to buy or sell an agreed number of certain assets at or before a specified date T, *maturity date* or *expire date*, in the future, at a predetermined price K, namely *exercise price* or *strike price*. Unlike futures trading, the rights and obligations of options are separated from each other. For both to an option transaction, only the seller needs to pay a certain margin in advance, and the amount will change with the fluctuation of the price of the underlying assets. For sellers, the risk is uncertain, and there is no fixed upper limit; For buyers, the biggest risk is the loss of all option premiums.

According to the rights of buying or selling assets, options can be divided into call options and put options. And according to the payoff and admitted exercise time, options can also be divided into such as European options, American options, and Asian options and other types. As a highly leveraged and flexible financial derivative, options play an important role in many fields and could be used for risk measurement, asset allocation and price discovery, etc.

To price the options fairly researchers have made many efforts. Black and Scholes (1973) constructed the Black-Scholes model, which became the most classical one. [1] They assumed such seven ideal conditions for their model:

a) The short-time interest rate r is known and constant before the option expires.

b) One could save or borrow money with the short-term interest rate r.

c) The stock price S_t is an Ito process, i.e., $dS_t = \mu S_t dt + \sigma_t S_t dW_t$, where the constant μ is called *drift*, constant σ is called *diffusion* or *volatility*, the standard deviation of the stock price, and W_t is the Wiener process.

d) There are no dividends from the stock.

e) The holder could only exercise the option at maturity, if he/she want and we call this type of option *European*.

f) The market is frictionless, that is, when people buy or sell, there are no transaction costs.

g) A seller could sell an asset even if he/she does not own one, namely “short selling”.

Note that the volatility is not constant in the real financial market, Merton (1976) introduced the jump diffusion term into the volatility. [2] Then got the Black-Scholes-Merton (BSM) model and gave the BSM option pricing formula. Although in the following years Cox and Ross (1976), Engle (1982), Hull and White (1987) and Heston (1993) took many amendments to the BSM model, it is still the most widely used until now. [3-6] For one option and share, we could use excel, calculator, table of Gaussian distribution or even by hand to work out the option price, since there is only one option and share. However, traders would face millions of options and shares every day and the market changes all the time. Although there is non-trading day, the market is still working. For solving this problem, Hutchinson et al. (1994) first tried to explore the artificial neural networks (ANN) method to pricing options. [7] And from then on, machine learning method began to be applied into option pricing. The discussions included but were not limited to the comparison of parametric model, semi-parametric model and non-parametric model, the difference among different machine learning methods etc. What’s more, some methods could abandon the limited unreal assumptions above.

The rest of this paper is organized as follows. The applications on different types of options are discussed in the second part. And conclude the research significant, future work and limitations in the third part.

2. MACHINE LEARNING METHODS ON OPTION PRICING

According to the exercise time and payoff function, there is many types of options. The most popular are European and American options. We also call then vanilla options due to the popularity.

2.1. Methods for European Options

European options could only be exercised at the maturity date. So, we need not to worry about when to exercise, nor to care the path of the price of the underlying assets. There are different ways to get the option price. For example, we can use machine learning method to get the volatility σ and prediction of stock price S_t in the future in order to use B-S formula and get the price, which is still parametric model. Also, using machine to learn the relationship between some basic feature of the option and its price directly is feasible.

Many researchers started from ANN, which simulates neuron activity with a mathematical model. [7-11] ANN is an information processing system established based on imitating the structure and function of the neural network of the brain. ANN has self-learning, self-organization, self-adaptation and strong nonlinear function approximation ability, and has strong fault tolerance. It can realize functions such as simulation, binary image recognition, prediction and fuzzy control, and is a powerful tool for dealing with nonlinear systems.

Hutchinson et al. (1994) [7] used the ANN method to price the option on the S&P 500 futures. They chose the statistical volatility on a sliding window of 60 days to get the B-S price. A good result was shown in their paper. There still some topic to discuss like how to construct the network well, that is how to choose the hyper-parameter, and how to measure the performance of the new method. Now we know to choose some loss functions to measure how well the models perform. In recent year, Culkin and Das (2017) also used ANN method with the Google’s TensorFlow. [8] They picked six variables as input: stock price, exercise price, maturity time, dividends, interest rate and volatility. After trained, they have obtained the low-error model. The disadvantage is that they only mentioned four activation function without comparison. And they missed to discuss which one is better for some certain cases.

Salvador et al. (2020) used ANN to pricing both European and American options and gave some theoretical derivation of PDE solution to construct the loss function. [9] Their model needed six parameters: interest rate, maturity date, exercise price, dividends, constant known volatility and the stock price at infinity time. There paper showed the error is small but regrettably they only gave the unique and lacked comparison with other Feature selection or other learning method. Moreover, the last two are ideal values so that there is a gap between there research and the real investment scenario. Shao (2021) constructed a two-step model, using full join neural network with three feature volatility of last 30 days, moneyness (S/K) and time to maturity to get the implied volatility surface, then Gaussian process regression to predict the option price. [10] What’s more, a priori value from B-S model and day-by-day training made more information about market useful. Unfortunately, inverse matrix occurs in Gaussian process and costs much. We could consider the numerical method to get the inverse cheaper and faster.

Lu (2021) found B-S model could not reflect the random jump, explain volatility smile nor describe the phenomenon of stock spikes and thick tails, then she used Carr-Geman-Madan-Yor (CGMY) model to price Taiwan Index Options with ANN to verify the model. [11] As discussion above, ANN could give the option price easier and faster than the original B-S-M model. We find in the recent years, researchers used ANN but discussed less about itself. We only need to choose some useful feature with high information value and then train it without any financial interpretation. Some feature such as maturity data has explicit financial meaning and moneyness describes the intrinsic value of an option's premium in the market. Interestingly, the former is in the B-S formula but the latter is not. Maybe someone could think in the formula there is S and K, so the function of those two also makes sense. But if we choose the function instead of the original variable, we would lose some information. If moneyness do be more useful in the option pricing, we need to find the reason rather than only depending on the train outputs. If it is not useful, we also need to interpret why it perform well in the ANN. Those are the important task in the future.

And is there any other method better? With the birth and generalization of other machine learning methods, other methods have also been applied to the pricing of options.

Chen and Lee (1997) used the Genetic Algorithms (GA) to price the European call options and compared with the ones from Black-Scholes formula. [12] There were 100 generations for each GA run. Although the distribution covers the true price, the result is not good because when $S/K > 0.8$, the maxima of the distribution is off the true price. That is the most likely scenario is not the true price. Hence, in their opinion GA could work in limited effect and need to be revised.

Some other method like Support Vector Machine (SVM). [13,14] And some researchers especially focused on the same options on the Shanghai 50ETF. [13,16-18] Luo (2021) considered the generalized autoregressive conditional heteroskedastic (GARCH) model and used improved B-S model and GARCH-SVM method to price option on 50ETF. [13] He chose five parameters: stock price, exercise price, historical volatility, maturity date and interest rate. Comparing to the traditional parameter model, SVM only needed to learn the relationship between those economic parameters and the option price. While the model got a good result, he only picked the sigmoid as kernel function. We could consider adding ant colony algorithm and particle swarm algorithm, first proposed respectively by M. Dorigo et al. and J. Kennedy et al., to the model for comparison. [14,15]

Also, on 50ETF, Wu (2021) chose other tools instead of revising the model. [16] She compared multiple linear regression, SVM and random forest and she chose 20 features to train the models. The random forest could reflect the function between inputs and output best among those three methods. Then based on the training result, she gave the trading strategy to arbitrage and hedge and totally got 171.16% profit. However, the high volatility of strategy and how to find the optimal hyper-parameter have not been resolved yet. So, we cannot assert that the strategy is good because the high proceed may came from the high risk.

Still on 50ETF, Xie and You (2018) showed another method, long-short term memory (LSTM) method. [17] The real data are non-linear with high dimension. What's more, some financial data are time series, so they consider Recurrent Neural Networks (RNN) first. While they found that gradient disappearance and gradient explosion make it learn only short time before. Then, they consider the LSTM to control the speed of accumulating the information and forget some useless memory in order that it could learn long term information. They picked 5 features and compared it with Monte Carlo method. Finally, the LSTM performed better and faster.

Again, on 50ETF, Ke (2020) took investor sentiment into account. [18] Considering that LSTM neural network has strong nonlinear approximation ability by virtue of multi-layer hidden layer, and the introduction of memory unit can learn the timing law contained in data, Ke used BP neural network, SVM, and XGBoost prediction model as the control group. The 50ETF call options are samples, and MAE, MSE, and MRE are used as the evaluation criteria for model prediction accuracy. He did not consider the implied volatility into the LSTM and the choice of investor sentiment is with subjective arbitrariness.

Ivascu (2020) chose 5 features, stock price, exercise price, maturity date, interest rate and historical volatility (60 days standard deviation) after the preliminary model. [19] Then compared SVM, random forest, XGBoost, Light GBM, ANN, and GA with 1465 options in 121,488 records, then he found boosting models are best among them not only in short term but also in long term. Unfortunately, he did not consider the LSTM. So, in the future we could mainly focus on the methods such as LSTM and XGBoost that have proven their advantages. And for the not very good ones like GA, we could propose improved methods based on the basic model and take numerical experiments.

2.2. Methods for American Options

Although it is important, the research on machine learning methods on American options started later than European, especially for the high-dimension cases. For American options, the problem goes harder. The maturity is fixed, but the option holder could exercise it at any time before maturity. Hence, we hope to know which time is the best time or the optimal stopping time to exercise and give a fair price. Scientists have done some research on different aspects with many methods. [20-24] For example they used Markov decision process (MDP) and quantum mechanics mixed with the machine learning methods. With those novel ideas, research on pricing American options has taken steps forward.

Jang and Lee (2019) directly introduced Generative Bayesian Neural network (GBNN) to calibrate and predict on American options, using the S&P 100 American put options from 2003 to 2012. [20] Then they had the conclusion of Bayesian neural network is superior to the CGMY model. One of the shortcomings is in their model, they picked Laguerre polynomials as the basis functions without interpreting the reason. And, stock index option is a one-dimensional option they did not discuss the case of high-dimensional. In order to solve such a problem Chen and Wan (2021) and Becker et al. (2021) focused on the high-dimensional American options. [21,22] Chen and Wan constructed a neutral network framework for this case. The novel point is learning the differences between the price functions of the adjacent timesteps and then minimizing the residual of the backward SDE that couples both prices and deltas. They got not only the price and delta at the initial time but also the entire spacetime which is of good importance to hedging in the real transacting scenarios. The only thing that is not satisfactory is the final method is quadratic of the dimension. We could study whether there are other time-saving methods in the future. Becker et al. (2021) derive the optimal stopping time theoretically and extended their methods into Bermudian options.

To solve the problem of finding optimal exercise time, Zhao (2019) considered MDP. [23] He thought when the market is in a specific state, the Agent can change the state of the environment by performing specific actions. After the state changes, it will return an observation value to the Agent and also get feedback, so that the Agent can take a new one based on the returned information. The action is repeated. How an agent chooses an action is called a strategy. The task of MDP is to find a strategy to maximize this feedback. Hence, we could use ANN, SVM and reinforcement learning method to learn how the MDP should do to make trading strategy. The traditional pricing methods are relatively time-consuming and dependent on the model. Using MDP could stimulate the path of the stock price. Although the origin paper only check efficiency on European options, due to the simulation path, American and Asian options are also accessible. What's more, MDP is easy to calibrate and could adapt to more types of data. And the existence of market transaction friction making the traditional parameter model give the bias result, pricing can also be promoted with machine learning methods. The limitation of this article is that he only focused on 12 American companies, and in the future, more data should be used to verify this method.

Not like Zhao as above, Chowdhury et al. (2020) used software called RapidMiner with three learning algorithms, decision tree, ANN, and ensemble learning method, which are used to predict the stock price of eleven companies and then to price the options. [24] They got the result that the ensemble learning method was the best result to predict the stock price among those three. While they did not show some details such as just setting but not explaining why to choose 0.3 as the learning rate and 0.2 as the momentum. Maybe the software recommended or something else. Interestingly, this paper introduced quantum mechanics to explain the B-S option pricing equation. Under Wick rotation, the time turns to imaginary time and they would do further research on it in the future. And that could be used on American options.

2.3. Methods for Other Options

Some other exotic options are also important such as Asian options, which can only be exercised at maturity date but the payoff function is dependent on the average price of the underlying asset. Unlike the European option depending only on S_T, the average could avoid some extreme cases such as the S_t changing drastically near the maturity. According to the choice between arithmetic average and geometric average, there are two types of Asian options.

Gan et al. (2019) used backward propagation neural network (BPNN) to train both arithmetic and geometric Asian options and got accurate results. [25] Again, the deep learning method is a model-free approach so that avoid the ideal non-realistic assumptions. There is some claw in this paper. For example, they only use one BPNN in this paper. We do not know whether there is another better method.

Li (2021) compared ANN and Mont Carlo method for the options. Both European and Asian, both 1-dimension and high-dimension cases, both arithmetic and geometric. [26] But in this paper, he just compared the results between the numerical solutions, not to mention any analytical solutions, even if the trivial European options.

As for Bermudian options, unlike the European and Asian options with fixed exercise date, Bermudian options, more like the American ones, could be exercised on several but only predetermined dates before the expire date. Becker et al. did some numerical experiment on it and got good results. [22]

Under the machine learning methods, we could ignore whether one option is European or Asian, because they both could exercise only the one day. And the logic inside would be less important. If an ANN learn some type of options, then it would output the same one accurately and fast, but nobody knows which one it is.

3. CONCLUSION

Nowadays, machine learning method such as neural networks in financial market has been a hot issue. Among them, derivatives pricing plays an important role in both academia and actual transactions. Deep learning algorithms that keep pace with the times also have good model generalization capabilities. And its prediction accuracy has surpassed the traditional financial model. For example, compared with the more traditional Monte Carlo method, LSTM achieves higher accuracy with fewer times. What's more, to the non-linear, high-dimension data, machine learning methods would do better than traditional methods. Because the former does not need to construct a complicated model while parts of the latter are linear. In addition, we can select different features and enter different variables to improve the accuracy of prediction. While for traditional methods, we might need to construct another PDE or SDE to picture the dynamic.

By far, LSTM, derived from RNN, has been one of the best methods to learn financial time series data. We could re-examine the original model and make corrections on this basis, and obtain a learning model that is also applicable to financial data. While for American options there is an additional question, finding the optimal stopping time and provide a reasonable explanation because the optimal exercise time could not be learned from the market information directly. Moreover, the feature selection in deep learning is of good importance, but not each picked feature would have the financial meaning. So, we need to find and interpret the relation between the learning theory and the real financial markets. What's more, we could use different measure to evaluate a model. Such as Hou (2021) in her paper, root mean square error (RMSE) to reflect the accuracy, mean absolute error (MAE) and mean absolute percent error (MAPE) to represent the robust. [27]

ANN and its derivatives are all black boxes. We could get a satisfactory value but probably we can give neither a rigorous derivation proof nor a financial explanation. Unlike the traditional methods, the neural network does not care about the specific meaning of the input parameters, but the degree of its influence on the model. we cannot rely solely on new methods and stop discussing traditional models. The two directions should develop together.

REFERENCES

[1] Black F, Scholes M. The pricing of options and corp orate liabilities[J]. Journal of Political Economy, 19 73, 81(3): 637-654.

[2] Merton R C. Option pricing when underlying stock r eturns are discontinuous[J]. Journal of financial eco nomics, 1976, 3(1-2): 125-144.

[3] Cox J C, Ross S A. The valuation of options for alter native stochastic processes[J]. Journal of financial e conomics, 1976, 3(1-2): 145-166.

[4] Engle R F. Autoregressive conditional heteroscedasti city with estimates of the variance of United Kingd om inflation[J]. Econometrica: Journal of the econo metric society, 1982: 987-1007.

[5] Hull J, White A. The pricing of options on assets wit h stochastic volatilities[J]. The journal of finance, 1 987, 42(2): 281-300.

[6] Heston S L. A closed-form solution for options with stochastic volatility with applications to bond and c urrency options[J]. The review of financial studies, 1993, 6(2): 327-343.

[7] Hutchinson J M, Lo A W, Poggio T. A nonparametri c approach to pricing and hedging derivative securi ties via learning networks[J]. The journal of Financ e, 1994, 49(3): 851-889.

[8] Culkin R, Das S R. Machine learning in finance: the case of deep learning for option pricing[J]. Journal of Investment Management, 2017, 15(4): 92-100.

[9] Salvador B, Oosterlee C W, Meer R. European and A merican Options Valuation by Unsupervised Learni ng with Artificial Neural Networks[C]//Multidiscip linary Digital Publishing Institute Proceedings. 202 0, 54(1): 14.

[10] Shuai Shao. The research of European option pricin g based on Gaussian process regression (in Chinese) [D]. Shandong University. 2021

[11] Yuxin Lu. Research on option pricing based on CG MY model and FCOS method [D]. Shandong Univ ersity,2021.

[12] Chen S H, Lee W C. Option pricing with genetic al gorithms: a second report[C]// Proceedings of Inter national Conference on Neural Networks (ICNN'9 7). IEEE, 1997, 1: 21-25.

[13] Jian Luo. Research on option pricing based on GAR CH-SVM volatility modelling for improved B-S m odel (in Chinese) [D]. Shandong University. 2021

[14] Dorigo M, Maniezzo V, Colorni A. Ant system: opt imization by a colony of cooperating agents[J]. IEE E Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 1996, 26(1): 29-41.

[15] Kennedy J, Eberhart R. Particle swarm optimization [C] Proceedings of ICNN'95-international conferen ce on neural networks. IEEE, 1995, 4: 1942-1948.

[16] Sirong Wu. SSE 50ETF trading strategies in the opt ions market (in Chinese) [D]. Shanghai Normal Un iversity, 2021. DOI:10.27312/d.cnki.gshsu.2021.00 1716.

[17] Heliang Xie, Tao You. Research on European Stoc k Index Option Pricing Based on Deep Learning Al gorithms——Evidence from 50ETF Option Market (in Chinese) [J]. Statistics & Information Forum,20 18,33(06):99-106.

[18] Chuqiu Ke, A study of option price prediction base d on LSTM neural network (in Chinese) [D]. South east University, 2021,

[19] Ivașcu C F. Option pricing using machine learning [J]. Expert Systems with Applications, 2021, 163: 1 13799.

[20] Jang H, Lee J. Generative Bayesian neural network model for risk-neutral pricing of American index options[J]. Quantitative Finance, 2019, 19(4): 587-603.

[21] Chen Y, Wan J W L. Deep neural network framework based on backward stochastic differential equations for pricing and hedging American options in high dimensions[J]. Quantitative Finance, 2021, 21(1): 45-67.

[22] Becker S, Cheridito P, Jentzen A, et al. Solving high-dimensional optimal stopping problems using deep learning[J]. European Journal of Applied Mathematics, 2021, 32(3): 470-514.

[23] Xinrui Zhao, the application of machine learning in option pricing (in Chinese) [D]. University of Science and Technology of China. 2019.

[24] Chowdhury R, Mahdy M R C, Alam T N, et al. Predicting the stock price of frontier markets using machine learning and modified Black–Scholes Option pricing model[J]. Physica A: Statistical Mechanics and its Applications, 2020, 555: 124444.

[25] Gan L, Wang H, Yang Z. Machine learning solutions to challenges in finance: An application to the pricing of financial products[J]. Technological Forecasting and Social Change, 2020, 153: 119928.

[26] Fangqi Li. Numerical Algorithm of option pricing based on deep learning (in Chinese) [D]. Beijing University of Posts and Telecommunications, 2021. DOI: 10.26969/d.cnki.gbydu.2021.001531.

[27] Ning Hou. Study on option pricing based on parametric model and nonparametric machine learning model (in Chinese) [D]. Northwest University. 2021.