Deep Learning in Model Risk Neutral Distribution for Option Pricing

Chin Chou
Institute of Information Management
National Chiao Tung University
Hsinchu, Taiwan, R.O.C
chinchou.iim06g@nctu.edu.tw

Jhih-Chen Liu
Institute of Information Management
National Chiao Tung University
Hsinchu, Taiwan, R.O.C
j0753418.iim07g@ nctu.edu.tw

Chiao-Ting Chen
Institute of Information Management
National Chiao Tung University
Hsinchu, Taiwan, R.O.C
rsps971130@gmail.com

Szu-Hao Huang
Department of Information Management and Finance
National Chiao Tung University
Hsinchu, Taiwan, R.O.C
szuhaohuang@nctu.edu.tw

***Abstract*—Option pricing has been studied extensively in recent years. An important issue in option pricing is the estimation of the risk neutral distribution of an underlying asset. Better estimation of this distribution can lead to a more rational investment, enabling one to earn an equal return with lower risk. To price options precisely and correctly, traditional financial engineering methods make some assumptions for the risk neutral distribution. However, some assumptions of traditional methods have proved inappropriate and insufficient in empirical option pricing analysis. To address these problems in option pricing, this study adopts a data-driven approach. Owing to advances in hardware and software, studies have been using deep learning methods to price options; however, these have not adequately considered the risk neutral distribution. This may cause an uncontrollable risk, thereby preventing the real-world application of the model. To overcome these problems, this study proposes a deep learning method with a mixture distribution model. Further, it generates a rational risk neutral distribution with accurate empirical pricing analysis.**

Keywords—option pricing, deep learning, mixture distribution model, risk neutral distribution.

I. Introduction

In recent years, option trading markets have become increasingly competitive. Many different types of options and underlying assets can be traded on the market. Irrespective of the type of option, all options aim to avoid possible risk on future uncertainty. Some investors engage in the option trading market to construct a risk-free portfolio, whereas others use options for speculation [1][2]. The most important issue in option trading is the option pricing method. A well-structured pricing method can perform risk measurement fairly and result in a profitable investment [3]. Certain experts, called market makers play an important role in option markets. The market makers' main role is to provide market liquidity [4][5]. They make profits from the spread of the bid and ask prices. However, if market makers misprice an option and leave too much position in the inventory, they will be exposed to unnecessary risk. Therefore, they urgently require a well-structured option pricing model to offer appropriate quotations [6]. Traditionally, market makers used the Black-Scholes model and adjusted its result based on their own experience. However, the classical Black–Scholes (BS) model has proved insufficient in empirical pricing analysis for European options owing to economic assumptions such as the constant variance and normal distribution of the log return of underlying assets [7]. These unreasonable assumptions have restricted the achievement of good pricing results. Therefore, various studies have tried to deal with different restrictions [8][9][10][11].

This paper proposes a deep neural network structured with a Gaussian mixture model to handle European option pricing. To prevent the model from making irrational assumptions and to enable it to price European options correctly and accurately, a large amount of historical data of Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX) options are collected from the Taiwan Futures Exchange (TAIFEX). Apart from pricing accuracy, risk problems are also considered. The proposed deep structure neural network can identify temporal changes in the risk neutral distribution. Therefore, the proposed model is more practical and rational for real-world applications.

II. Related Work

Option pricing has been an important concern in computational finance. The well-known BS model, proposed in 1973, was the first to address this issue [12]. This model is used by some traders even today. However, the assumptions of this model have been proven to be inappropriate through considerable empirical evidence, such as a large jump in underlying assets and the volatility smile in real pricing. To address the large jump in underlying assets, Merton [8] proposed a model in which the jump process was mixed with the original Weiner process. To address the discontinuous return of underlying assets, Kou [13] and Carr and Geman [14] constructed the jump process differently to obtain a more theoretically appropriate model. This type of model is also called the Levy model, and it can capture unusual events that impact the underlying asset [8][13][14]. To relax constant volatility assumptions, Heston [9] established a model to describe the volatility of the underlying asset by using another Wiener process. All of the abovementioned models can handle problems of varying degrees. However, these methods are not suitable for real-world applications. Therefore, this paper proposes a model that adopts a mixture distribution model with a hybrid structure to overcome the inadequacies of the original assumptions of the BS model.

In recent decades, many studies have applied deep learning to option pricing. Initially, Malliaris [15] proposed a neural network for options of Standard and Poor's (S&P) 100. Hutchinson [16] proposed nonparametric learning networks for options of S&P 500. Both these studies were considered innovative. Subsequently, Yao [11] proposed another simple neural network model for options of Nikkei 225 index futures that differed from the conventional BS model. This model partitions the option data into three groups according to the moneyness. Similarly, Bennell and Sutcliffe [10] reported a model for FTSE 100 Options and experimented with various inputs with different moneyness. Lajbcygier [17] tried to constraint the neural network model with economic principles to realize improved pricing performance. Gradojevic [18] used a modular neural network to fit different conditions by grouping data with different moneyness and maturity. However, this model relies on manual design. Yang [19] proposed a gated neural network (GNN) model to generate the weights of different modules by automatically fitting different conditions.

This paper proposes a novel deep structure neural network. In the BS model, we consider that the future price of an underlying asset on the settlement date will form a probability distribution, called a risk neutral distribution, and that the current price of an option can be derived by calculating the expected value with the condition of the strike price. By contrast, the proposed model captures the risk neutral distribution by a mixture distribution model. This approach breaks the inappropriate normality. Further, this study assumes that the volatility over the market cannot be extracted using a single value of the current underlying asset but requires a series of historical trading data of the underlying asset. Therefore, the model contains a deep structure dealing with the series data.

III. Proposed Method

This section describes the proposed model in detail and uses TAIEX options to calculate the experiment result. TAIEX options are European options. Option pricing models are constructed to deal with the problem of how much a contract is worth. The European-style call option price can be calculated as follows:

$$C(S_t, K, T) = e^{-r(T-t)} \int \max(0, S_T - K) f(S_T|S_t) dS_T \quad (1)$$

$C(S_t, K, T)$ is the call option price; t, the current date; S_t, the price of the underlying asset on the current date; K, the strike price on the settlement date; T, the settlement date; annualized $(T - t)$, the time to maturity; S_T, the price of the underlying asset on the settlement date; and r, the risk-free rate. In this formula, the risk neutral distribution term $f(S_T|S_t)$ is the most important factor, and it has been studied for decades.

A. Black–Scholes model

Most researchers usually estimate the risk neutral distribution term by constructing a stochastic process model on the underlying asset using some parameters from meaningful economic principles. In the BS model, the dynamics of the underlying asset are constructed using geometric Brownian motion with the following equation:

$$dS_t = rS_t dt + \sigma S_t dW_t \quad (2)$$

where S_t is the underlying asset at time t; r, the risk-free rate (same as (1)); σ, the volatility involved in the returns of the underlying asset; and W_t, the Wiener process. By solving the stochastic calculus, (3) is obtained from (2):

$$S_T = S_t e^{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{(T-t)}Z} \quad (3)$$

S_T is the price of the underlying asset at time T, and it is a necessary term to estimate the expected return of an option. Through this process, the pricing result can be derived using a computing method such as the Monte Carlo method, fast Fourier transform, or finite difference method. However, these methods entail high computational costs. Fortunately, the following solution has long been available:

$$C = S_t N(d_1) - K e^{-r(T-t)} N(d_2)$$
$$d_1 = \frac{\ln(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t} \quad (4)$$

With the closed-form solution (4), the call pricing result can be calculated directly without simulation. Here, N is the normal cumulative distribution function; $N(d_1)$, the delta, that is, the hedge against the option position; and $N(d_2)$, the probability that the option is in-the-money, that is, the price of the underlying asset on the settlement date is higher than the strike price. term to the S_T of (3) which is represented as τ. This term is calculated from a historical price series, and it can release the constraint that reduces the peak of the risk neutral distribution. The proposed structure can learn this pattern by itself. The correction term is added to (3) and shown below.

B. Proposed model

To solve the problem of the constraints in the stochastic process model, the proposed model first makes a correction:

$$S_T = S_t e^{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{(T-t)}Z} + \tau$$
$$\tau = f(S_t, \dots, S_{t-n}, V_t, \dots, V_{t-n}; \theta) \quad (5)$$

To estimate σ and τ in (5), the most important factor that can generate a more complex risk neutral distribution is the adoption of the concept of a Gaussian mixture model, which ensembles several Gaussian distribution models to fit the complex probability distribution more efficiently. The new formula with the combination of i distributions is

$$S_T = \sum_i w_i \left(S_t e^{\left(r - \frac{1}{2}\sigma_i^2\right)(T-t) + \sigma_i\sqrt{(T-t)}Z} + \tau_i\right) \quad (6)$$

i sets of parameters are estimated using the proposed structure in (5). Each set of estimated parameters is given a weight w_i obtained from the proposed neural network. Typically, irrespective of how large the volatility is, the most probable S_T, which is also the peak of the risk neutral distribution, remains around the current price. However, when the volatility increases, a large jump or drop should be expected relative to the current price. If the risk neutral distribution can reflect this situation, a bimodal distribution should be observed with one peak each toward the high and the low price regions. For this purpose, the weight parameter is estimated using the series data of historical price and trading volume because the proposed method assumes that the needed implied information is contained in both series data. The whole model is shown in Fig. 1. With the design of the Gaussian mixture model, the constraints of the original model are released. The risk neutral distribution can now have two peaks or more. As i in (6) increases, distributions with even some odd shapes can be observed; this better reflects the real-world situation.

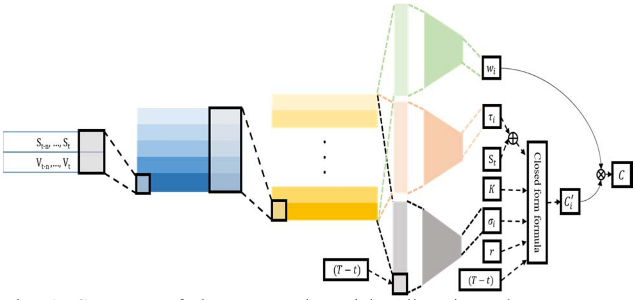

Fig. 1. Structure of the proposed model. All estimated parameters are independently generated by fully connected layers that take the time series features extracted by a one-dimensional convolutional neural network.

C. Data Description

In the experiments, TAIEX options collected from Taiwan Futures Exchange are used for European call option data. As the option has already been traded since 2001 in the Taiwanese market, the option data from Taiwan Futures Exchange should be suitable for testing the proposed method. Data from January 2016 to March 2018 is collected, and the data frequency is per minute data. Only near month contracts are used and weekly options are excluded.

IV. Experimental Results

The previous section introduces the proposed method with the sequences of historical data and concept of mixed distribution. Therefore, the model can be flexibly tuned for different option data. For example, high-frequency-trading option data may need a larger distribution to cover the variety of data. Therefore, the model settings are chosen by first constructing a setting experiment. After choosing the model settings, the temporal variation of the risk neutral distribution is analyzed. As different conditions exist in the market, the risk neutral distribution is skewed in different ways. Next, comparisons with other option pricing models, including some traditional financial engineering models and some deep learning models, are performed. The results for TAIEX options from 2017 are compared. Finally, to check the practicality in real-world applications, a trading strategy comparison is performed.

A. Setting experiment

To verify an appropriate experimental setting and to compare the results of different models, different sequence lengths and different numbers of distributions are applied. Because TAIEX options data is used for the baseline judgement, TAIEX option data for 2016 are used in the experiment. To make this experiment fair, a rolling test is conducted between May 2016 and December 2016. t is the current date, and training data for 60 trading days from $(t - 60)$ to t and $(t + 1)$ is used for testing. The Chinese New Year is celebrated in February, and the resulting long period with no trading may cause a large bias. Therefore, the experimental setting starts from option contracts due in May 2016. After trying different settings, the setting of 150 sequence length with five weighted distributions is finally used.

B. Pricing results comparison

This section discusses the empirical pricing result of the proposed model compared with that of other pricing methods including BS, Heston, Kou Jump (Kou), Variance Gamma (VG), and GNN. The BS, Heston, Kou, and VG econometric models are calibrated with the last five days of training data and the GNN deep learning method is trained with seven single models. Both GNN and the proposed model are trained with 60 trading days and tested on the day after the last training day. The results are shown with both mean absolute percentage error (MAPE) and mean absolute error (MAE) in Fig. 2.

Each model is tested for all trading days in 2017 (Fig. 2). However, for quick and easy comparisons, the daily pricing error is transformed to the near-term trading period error. There are 12 near-term trading periods in 2017. The results clearly indicate that in most periods, the proposed model outperforms other methods in terms of both MAPE and MAE.

Fig. 2. Pricing error of each model in terms of both MAPE (top) and MAE (bottom). The pricing error is calculated on every trading day, and the results shown in both figures are the mean of each near-term contract trading period in 2017. Therefore, there are 12 points in both figures. The proposed method indicated by the blue line clearly outperforms other methods.

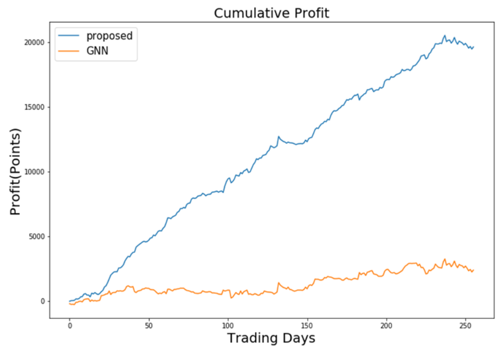

Fig. 3. Cumulative profit by simulated trading result.

C. Trading strategy

Irrespective of the pricing model applied, the model should be useful in real-world applications. To determine the practicality of the proposed model, a simple trading strategy is simulated. Fig. 3 shows a trading sample. When the option price of the market is understated, implying that the model price is higher than the market price, a call contract is longed. By contrast, when the option price is overstated, implying that the market price is higher than the model price, a call contract is shorted. In the whole trading period, the holding position is constrained to one for each strike. All positions are closed at the end of the trading day, implying that only intraday trading is performed. Fig. 4 shows the profit curve. Both GNN and the proposed method perform well in the trading simulation. When the proposed model prices the option more precisely, the profit is much higher than that obtained with the GNN. However, this result is obtained with simple simulated conditions that omit many real-world factors such as the trading fee and trading tax. Another important factor to be considered is whether the trading volume is sufficient for simulated trading.

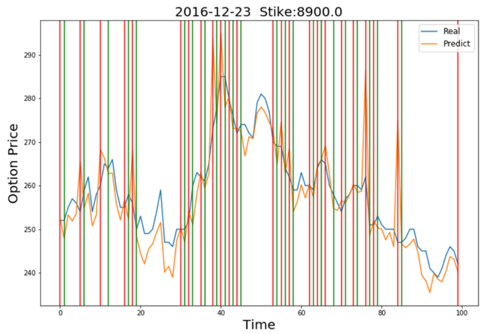

Fig. 4. Trading sample for 2016-12-23. The blue line indicates the real price of this option with 8900 strike. The orange line indicates the model prediction. The red and green vertical lines respectively indicate the buy and the sell point.

V. CONCLUSION

This study proposed an option pricing model with a deep learning structure to model the risk neutral distribution. This model can outperform existing option pricing models in terms of both MAPE and MAE. Despite the pricing results, the distribution diversity is captured by the proposed model through the application of historical sequence data, which is important to closely simulate real-world situations and fix simple normality assumptions.

Future work could focus on eliminating the Brownian motion and obtaining the risk neutral distribution only from the data. The trading strategy is another factor that could be considered. With reinforcement learning becoming increasingly powerful, an option pricing model constructed through deep learning could be combined with reinforcement learning to construct an option portfolio such as Long Straddle or Butterfly Spread. Therefore, the option pricing model can be trained from a trading perspective. Further, AI could be used to train the option pricing model to make it interpretable. This will enable users to understand why they should long or short a given option.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Science and Technology, Taiwan for partially financially supporting this research under contract number MOST 108-2218-E-009 -049 and MOST 108-2221-E-009-081 through Pervasive Artificial Intelligence Research (PAIR) Labs and Taiwan Information Security Center at NTHU. This work was also supported by the FinTech Innovation Research Center, National Chiao Tung University.

REFERENCES

[1] Lakonishok, J., Lee, I., Pearson, N. D., & Poteshman, A. M. (2006). Option market activity. The Review of Financial Studies, 20(3), 813-857.

[2] Aragon, G. O., & Martin, J. S. (2012). A unique view of hedge fund derivatives usage: Safeguard or speculation?. Journal of Financial Economics, 105(2), 436-456.

[3] Rebonato, R., & Jäckel, P. (2011). The most general methodology to create a valid correlation matrix for risk management and option pricing purposes.

[4] Comerton□Forde, C., Hendershott, T., Jones, C. M., Moulton, P. C., & Seasholes, M. S. (2010). Time variation in liquidity: The role of market□maker inventories and revenues. The Journal of Finance, 65(1), 295-331.

[5] Hendershott, T., & Seasholes, M. S. (2007). Market maker inventories and stock prices. American Economic Review, 97(2), 210-214.

[6] Garleanu, N., Pedersen, L. H., & Poteshman, A. M. (2008). Demand-based option pricing. The Review of Financial Studies, 22(10), 4259-4299.

[7] Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. The Journal of finance, 52(5), 2003-2049.

[8] Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. Journal of financial economics, 3(1-2), 125-144.

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

[10] Bennell, J., & Sutcliffe, C. (2004). Black–Scholes versus artificial neural networks in pricing FTSE 100 options. Intelligent Systems in Accounting, Finance & Management: International Journal, 12(4), 243-260.

[11] Yao, J., Li, Y., & Tan, C. L. (2000). Option price forecasting using neural networks. Omega, 28(4), 455-466.

[12] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of political economy, 81(3), 637-654.

[13] Kou, S. G. (2002). A jump-diffusion model for option pricing. Management science, 48(8), 1086-1101.

[14] Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for Lévy processes. Mathematical finance, 13(3), 345-382.

[15] Malliaris, M., & Salchenberger, L. (1993). A neural network model for estimating option prices. Applied Intelligence, 3(3), 193-206.

[16] Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. The Journal of Finance, 49(3), 851-889.

[17] Lajbcygier, P. (2004). Improving option pricing with the product constrained hybrid neural network. IEEE Transactions on Neural Networks, 15(2), 465-476.

[18] Gradojevic, N., Gençay, R., & Kukolj, D. (2009). Option pricing with modular neural networks. IEEE transactions on neural networks, 20(4), 626-637.

[19] Yang, Y., Zheng, Y., & Hospedales, T. M. (2017, February). Gated Neural Networks for Option Pricing: Rationality by Design. In AAAI (pp. 52-58).