International Journal of Business and Management Study – IJBMS
Volume 2 : Issue 2 [ISSN : 2372-3955]
Publication Date: 19 October, 2015

Exploring Nonparametric Strategies for Pricing American Index Options

Hyunwoong Ji, Youngdoo Son, Hyeongmin Byun, and Jaewook Lee

Abstract— **Nonparametric models such as machine learning methods that are counterparts of parametric financial models were applied to pricing American index options. 10 year S&P 100 Index American options were adopted as experimental dataset and the both training (in-sample) and test (out-of-sample) errors of machine learning and ad-hoc pricing which is a conventional financial pricing model were calculated and compared to each other. We found out that Bayesian neural network outperforms the other pricing methods. Furthermore, we suggested an ensemble method which takes advantage of both machine learning method and ad-hoc pricing method and as a consequence, it shows the best performance.**

***Keywords*—American Option, Option Pricing, Machine Learning, Ad-hoc Pricing.**

I. Introduction

Option, the right to buy or sell an asset at a pre-specified price by a pre-determined date, is one of the popularly traded financial derivatives in extensive financial markets. Among the various features, the type of option, European or American, is the most fundamental feature that determines the characteristics of an option.

For European options, several pricing methods has been developed for the cases from the most traditional geometric Brownian [1] to the recent Levy models [2]. For American-type options, however, there are no explicit pricing formulae for most cases, thus some approximations or heuristics are required.

Tree models [3, 4] and finite difference methods [5, 6, 7] are popular ones for American-type option pricing, but both of them do not well-operate when the number of state variables becomes large. Other popular methods are Monte Carlo-based methods [8, 9, 10, 11]. Their computational complexity does not depend on the number of state variables but it requires a lot of computations to achieve accurate results.

The machine learning methods can be alternatives for those pricing methods. They have been successfully applied to the several nonlinear regression tasks including European option pricing [12, 13].

In this research, we explore the several machine learning methods to find American options pricing and compared the results with the ad-hoc approach which showed good performance for American option pricing [14, 15] as well as with one another. Furthermore, we construct the ensemble model combining the models explained above.

Hyunwoong Ji, Youngdoo Son, Hyeongmin Byun, Jaewook Lee
Seoul National University
Republic of Korea
{meditate, hand02, ragnarok, jaewook}@snu.ac.kr

The remainder of the paper is as follows. In the next section, the data set used for the research and the screening methods are described. In section III, we briefly review the exploited machine learning methods. In section IV, we explain the experimental results. Finally, we conclude the paper with some future directions in section V.

II. Data Description

A. *Data Description*

Table 1 contains the summary of the option data that we've used. Option from 2003/01/02 to 2013/08/30 were considered. Call option was not of our interest, because estimating put option is financially more complex than call option. The underlying of the option is S&P 100 index. Obviously, the exercise type is American. The total number of instances is 584,755 and the data is from Optionmetrics of IVY database. In the Optionmetrics data, there are 23 attributes for each option.

The attributes are listed in Table 1. Among them only time to maturity and moneyness which is underlying over strike price were selected as most relevant attributes to predict option price. Therefore, we used only the two attributes to predict option price.

TABLE 1 SUMMARY OF ATTRIBUTES FROM OPTIONMETRICS DATABASES

Attribute	Format	Example
Date	YYYY/MM/dd	20030102
Time to maturity	Days	16
Strike Price times 1000	-	405000
Highest bid	Dollar	0.4500
Lowest offer	Dollar	0.6000
Volume	-	126
Imp. Vol.	-	0.3582
Dividend	Continuously compounded rate (%)	1.9797
RV(10 days)	-	0.2684
RV(14 days)	-	0.2379
RV(30 days)	-	0.2234
RV(60 days)	-	0.2285
RV(91 days)	-	0.2695

RV(122 days)	-	0.2935
RV(152 days)	-	0.3016
RV(182 days)	-	0.3323
RV(273 days)	-	0.3016
RV(365 days)	-	0.2752
RV(547 days)	-	0.2566
RV(730 days)	-	0.2546
RV(1825 days)	-	0.2307
Interest rate	Continuously compounded rate (%)	1.3730
Underlying	-	459.5

B. *Data Screening*

Before making a model for option price, some portion of data should be removed because such data is not valuable being predicted for some financial reason. Totally, 4 filters were used to reduce irrelevant data.

The filters are as follows. Only traded option were considered because non traded options are usually overpriced. Time to maturity of an option should be between 7days and 90days because options with short time to maturity tend to be expensive due to time premium and options with long time to maturity are traded not frequently. The data of the first three year were removed because it shows unstable pattern. At last, only options that are more than 1($) were considered to alleviate the effect of price discreteness. After applying these exclusionary criteria, 83,700 instances remained which was originally 584,755.

III. Nonparametric Methods

There are a lot of researches to predict option price. In [16], GARCH volatility model and Black-Scholes model were combined to price American option. In [17] and [18], three machine learning methods, committee of neural networks, multi-layer perceptron and support vector machine, were applied to price American option. However, it used strike price and time to maturity as attributes so information of underlying asset was excluded, and this lead to poor performance. In

In this paper, we used 4 parametric or nonparametric methods to price American option price. The first one is a parametric model. It is an ad-hoc black-Scholes model which is popular method to price European option. To apply an ad-hoc scheme to American option case, a closed form valuation of American option is needed. So closed form approximation of American option which was suggested by [19] were used. Ad-hoc Black Scholes model is consist of 3 stages. At the first stage, volatility surface is generated from training points. At the next stage, the volatility surface substitute volatility part of closed form valuation. At the last stage, price of new point is predicted by using the volatility surface acquired at the second stage and the closed approximation [19].

The last three are nonparametric models that are state-of-arts machine learning methods; support vector regression [20], Neural network [21], and Bayesian neural network [22].

In support vector regression case, there are also three stages. At the first stage, data of window is separated into 2 parts. The first part is validation set which is the most latest options and the remained data is used as training data. The next step is that parameters which best fit to the validation data are acquired by solving the corresponding optimization problem. At the second stage all the data of window become new training data and the corresponding model is calculated using the best parameters of previous stage. Finally option price is estimated at new points using the model of the previous stage.

In Neural network case, option price is estimated in same way as support vector regression case. The only difference is that the machine learning model had been changed from SVR to Neural network.

In Bayesian neural network case, the previous first stage and second stage are combined into one stage because it does not need validation set for early stopping. At the first stage, a model is calculated and at the second stage option price is estimated at new points using the model of the previous stage.

IV. Ensemble Method

By analyzing both errors of Bayesian neural network and ad-hoc pricing model, we have found out two things: (i) Variance of error of Bayesian neural network is higher than that of ad-hoc model. (ii) Large portion of error of Bayesian neural network is mainly due to over-fitting problem (However, ratio between the over-fitted data and the counterpart of it is significantly small). Therefore, we suggested the following ensemble method:

$$P_{emsemble} = \begin{cases} P_{adhoc} & if \left|\frac{P_{BNN}}{P_{adhoc}} - 1\right| > \alpha \\ \frac{P_{adhoc} + P_{BNN}}{2} & otherwise \end{cases}$$

Where α is a predefined threshold level, P_{BNN} is an estimated price of Bayesian neural network and P_{adhoc} is that of ad-hoc model. The pricing scheme of ensemble method is straight forward. As we could realize, the estimated price of ad-hoc model gives not best but consistently quite good approximation of options. Therefore, if the discrepancy between the two model prices is larger than \alpha, we could take it as a signal that the Bayesian model is suffering from over-fitting problem and use the price of ad-hoc model as the estimated price of ensemble model. If it is not the case, we use arithmetic mean of the two models as the estimated price of ensemble model, which is called bagging and generally improves performance of prediction.

V. Experimental Results

Table 2 contains mean average percentage error (MAPE) for each model and each window size. In sample error, 1 day ahead forecast error and 7 day ahead forecast error were considered and for all cases Bayesian neural network was the best model for pricing option values. This is because the option data is sparse so Bayesian approach is more suit for option pricing than other machine learning methods.

TABLE 2 MEAN AVERAGE PERCENTAGE ERROR (MAPE) FOR EACH MODEL AND EACH WINDOW. THE DARKER SHADED REGION AND LIGHTER SHADED REGION ARE THE BEST FITTED MODEL AMONG MACHINE LEARNING AND AD-HOC MODELS, RESPECTIVELY. ALSO, THE BEST FITTED MODEL AMONG WHOLE MODELS ARE IN BOLD TYPE.

MAPE (Mean Average Percentage Error)				
Window Length (day)	In-sample			
	SVR	NN	BNN	Ad-hoc
1	0.700	0.296	**0.008**	0.070
6	0.498	0.148	0.068	0.123
11	0.510	0.164	0.096	0.146
16	0.528	0.161	0.116	0.163
21	0.543	0.167	0.136	0.178
Window Length (day)	1 day ahead forecast			
	SVR	NN	BNN	Ad-hoc
1	0.765	0.436	**0.119**	0.136
6	0.533	0.187	0.131	0.148
11	0.542	0.192	0.136	0.164
16	0.555	0.185	0.146	0.177
21	0.567	0.191	0.159	0.190
Window Length (day)	7 day ahead forecast			
	SVR	NN	BNN	Ad-hoc
1	0.839	0.659	0.297	0.225
6	0.639	0.328	0.229	0.212
11	0.634	0.259	**0.202**	0.218
16	0.641	0.256	0.205	0.224
21	0.650	0.253	0.211	0.230

Table 3 shows the best models in terms of in-sample error, for each time to maturity and moneyness. One can see that BNN performs consistently better than Ad-hoc and our proposed methods, for every value of time to maturity and moneyness. This result is obvious and expectable because objective of machine learning method is minimizing error between estimated and true price of option over in-sample data.

Table 4 shows the best models in terms in 1-day ahead forecasting error, for each time to maturity and each moneyness. Unlike previous situation where BNN outperforms Ad-hoc and proposed method for every pair of time to maturity and moneyness, our proposed method has smaller error than BNN and Ad-hoc method when time to maturity is short or long. This phenomenon can be described as the effect of bagging, which reduces the effect of extrapolation with respect to time to maturity. On the other hand, for medium range of time to maturity BNN performed well relative to proposed and Ad-hoc method, since the model already interpolates well without ensemble.

Table 5 shows the same for 7-day ahead forecasting error. One can see similar pattern with previous table, which used 1-day ahead forecasting error. Here, unlike previous 1-day result, our proposed method works relatively well for short time to maturity only. This implies, for long-term prediction, our method only reduces the effect of extrapolation for relative short levels of time to maturity.

TABLE 3 BEST FITTED MODEL IN TERMS OF IN-SAMPLE ERROR FOR EACH TIME TO MATURITY AND MONEYNESS. THE ERRORS OF BEST FITTED MODEL AMONG THREE MODELS (AD-HOC, BNN, AND PROPOSED ONE) ARE IN BOLD-TYPE.

Best Model (in-sample)							
Time to maturity, τ,(day)	Model	Moneyness, S/K=κ					
		ITM		ATM		OTM	
		κ<0.94	0.94< κ<0.97	0.97< κ<1.00	1.00< κ<1.03	1.03< κ<1.06	1.06< κ
τ <30	Ad-hoc	0.019	0.026	0.049	0.078	0.070	0.107
	BNN	**0.002**	**0.002**	**0.005**	**0.013**	**0.016**	**0.022**
	Proposed	0.009	0.013	0.026	0.041	0.038	0.059
60< τ<90	Ad-hoc	0.023	0.032	0.041	0.040	0.052	0.105
	BNN	**0.001**	**0.001**	**0.003**	**0.005**	**0.009**	**0.010**
	Proposed	0.011	0.016	0.021	0.021	0.028	0.056
90< τ	Ad-hoc	0.022	0.031	0.032	0.036	0.054	0.146
	BNN	**0.001**	**0.001**	**0.001**	**0.002**	**0.003**	**0.005**
	Proposed	0.011	0.015	0.016	0.018	0.027	0.075

TABLE 4 BEST FITTED MODEL IN TERMS OF 1 DAY AHEAD ERROR FOR EACH TIME TO MATURITY AND MONEYNESS. THE ERRORS OF BEST FITTED MODEL THREE MODELS (AD-HOC, BNN, AND PROPOSED ONE) ARE IN BOLD-TYPE.

Best Model (1 day ahead)							
Time to maturity, τ,(day)	Model	Moneyness, S/K=κ					
		ITM		ATM		OTM	
		κ<0.94	0.94< κ<0.97	0.97< κ<1.00	1.00< κ<1.03	1.03< κ<1.06	1.06< κ
τ <30	Ad-hoc	**0.025**	0.038	0.076	0.144	0.166	0.225
	BNN	0.050	0.034	0.068	0.183	0.180	0.240
	Proposed	0.032	**0.031**	**0.066**	**0.129**	**0.155**	**0.213**
60< τ<90	Ad-hoc	0.034	0.049	0.063	0.076	0.102	0.184
	BNN	0.039	**0.032**	**0.041**	**0.059**	**0.086**	0.152
	Proposed	**0.031**	0.036	0.048	0.064	0.088	**0.149**
90< τ	Ad-hoc	0.040	0.053	0.057	0.065	0.092	0.365
	BNN	0.053	0.049	0.051	0.064	0.083	**0.240**
	Proposed	**0.038**	**0.043**	**0.045**	**0.054**	**0.075**	0.266

TABLE 5 BEST FITTED MODEL IN TERMS 7 DAY AHEAD ERROR FOR EACH TIME TO MATURITY AND MONEYNESS. THE ERRORS OF BEST FITTED MODEL AMONG THREE MODELS (AD-HOC, BNN, AND PROPOSED ONE) ARE IN BOLD-TYPE.

Best Model (7 day ahead)							
Time to maturity, τ,(day)	Model	Moneyness, S/K=κ					
		ITM		ATM		OTM	
		κ<0.94	0.94< κ<0.97	0.97< κ<1.00	1.00< κ<1.03	1.03< κ<1.06	1.06< κ
τ <30	Ad-hoc	**0.031**	0.052	0.110	0.265	0.345	0.453
	BNN	0.062	0.044	**0.092**	0.259	0.399	0.539
	Proposed	0.035	**0.041**	**0.092**	**0.240**	**0.326**	**0.436**
60< τ<90	Ad-hoc	0.039	0.059	0.095	0.149	0.200	0.299
	BNN	0.053	**0.037**	**0.065**	**0.120**	0.187	0.304
	Proposed	**0.037**	0.043	0.074	0.127	**0.184**	**0.283**
90< τ	Ad-hoc	0.046	0.069	0.090	0.123	0.161	0.389
	BNN	0.058	**0.042**	**0.063**	**0.095**	**0.133**	**0.232**
	Proposed	**0.044**	0.050	0.072	0.104	0.140	0.298

VI. Conclusion and Discussion

Throughout this paper, we have shown that machine learning approach out-performs ad-hoc pricing model, which means it gives academic insight: there could be a better financial stochastic process which explains option market better than current financial model.

Also we have suggested an ensemble method which exploits only advantages of both financial models and machine learning models and the result of it improved accuracy of prediction.

For future work, we are planning to try other machine learning methods including multi-layer perceptron, neural network using 1-norm regularization, and find a model which does not violate arbitrage condition.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2011-0017657).

References

[1] Black, Fischer, and Myron Scholes. "The pricing of options and corporate liabilities." *Journal of Political Economy* (1973): 637-654.

[2] Carr, Peter, and Dilip Madan. "Option valuation using the fast Fourier transform." *Journal of Computational Finance* 2.4 (1999): 61-73.

[3] Cox, John C., Stephen A. Ross, and Mark Rubinstein. "Option pricing: A simplified approach." *Journal of Financial Economics* 7.3 (1979): 229-263.

[4] Boyle, Phelim P., Jeremy Evnine, and Stephen Gibbs. "Numerical evaluation of multivariate contingent claims." *Review of Financial Studies* 2.2 (1989): 241-250.

[5] Schwartz, Eduardo S. "The valuation of warrants: Implementing a new approach." *Journal of Financial Economics* 4.1 (1977): 79-93.

[6] Hirsa, Ali, and Dilip B. Madan. "Pricing American options under variance gamma." *Journal of Computational Finance* 7.2 (2004): 63-80.

[7] Cont, Rama, and Ekaterina Voltchkova. "A finite difference scheme for option pricing in jump diffusion and exponential Lévy models." *SIAM Journal on Numerical Analysis* 43.4 (2005): 1596-1626.

[8] Tilley, James A. "Valuing American options in a path simulation model." *Transactions of the Society of Actuaries* 45.83 (1993): 104.

[9] Barraquand, Jérôme, and Didier Martineau. "Numerical valuation of high dimensional multivariate American securities." *Journal of Financial and Quantitative Analysis* 30.03 (1995): 383-405.

[10] Broadie, Mark, and Paul Glasserman. "Pricing American-Style Securities using Simulation." *Journal of Economic Dynamics and Control* 21.8 (1997): 1323-1352.

[11] Longstaff, Francis A., and Eduardo S. Schwartz. "Valuing American Options by Simulation: A Simple Least-Squares Approach." *Review of Financial studies* 14.1 (2001): 113-147.

[12] Han, Gyu-Sik, Bo-Hyun Kim, and Jaewook Lee. "Kernel-Based Monte Carlo Simulation for American Option Pricing." *Expert Systems with Applications* 36.3 (2009): 4431-4436.

[13] Park, Hyejin, and Jaewook Lee. "Forecasting nonnegative option price distributions using Bayesian kernel methods." *Expert Systems with Applications* 39.18 (2012): 13243-13252.

[14] Dumas, Bernard, Jeff Fleming, and Robert E. Whaley. "Implied volatility functions: Empirical tests." *The Journal of Finance* 53.6 (1998): 2059-2106.

[15] Brandt, Michael W., and Tao Wu. "Cross-sectional tests of deterministic volatility functions." *Journal of Empirical Finance* 9.5 (2002): 525-550.

[16] Weber, Michael, and Marcel Prokopczuk. "American option valuation: Implied calibration of GARCH pricing models." *Journal of Futures Markets* 31.10 (2011): 971-994.

[17] Dindar, Zaheer A., and Tshilidzi Marwala. "Option pricing using a committee of neural networks and optimized networks." *2004 IEEE International Conference on Systems, Man and Cybernetics. Vol. 1*. IEEE, 2004.

[18] Pires, Michael Maio, and Tshilidzi Marwala. "American option pricing using multi-layer perceptron and support vector machine." *2004 IEEE International Conference on Systems, Man and Cybernetics. Vol. 2*. IEEE, 2004.

[19] Bjerksund, Petter, and Gunnar Stensland. "Closed form valuation of American options." Working Paper, 2002.

[20] Smola, Alex J., and Bernhard Schölkopf. "A tutorial on support vector regression." *Statistics and Computing* 14.3 (2004): 199-222.

[21] Hagan, Martin T., Howard B. Demuth, and Mark H. Beale. *Neural network design*. Boston: Pws Pub., 1996.

[22] Demuth, Howard, and Mark Beale. "Neural network toolbox for use with MATLAB." Mathworks. 1993.

About Author (s):

Hyunwoong Ji is a Ph.D Candidate in the Department of Industrial Engineering, Seoul National University.

Youngdoo Son is a Ph.D Candidate in the Department of Industrial Engineering, Seoul National University.

Hyeongmin Byun is a M.S Student in the Department of Industrial Engineering, Seoul National University.

Jaewook Lee, Ph. D is a Professor in the Department of Industrial Engineering, Seoul National University.