

A Machine Learning Approach to Forecasting Option Prices Beyond the Black Scholes Model

Tahmid Tamrin Suki^{1*}, Md Motinur Rahman², Subrata Saha², Sayma Alam Suha¹, Nazneen Akhter³

Abstract

The Black-Scholes (BS) model has been used as the foundation for pricing options in the stock market for years. However, this mathematical model is very simple, which limits its capability to predict the present market dynamics for stock prices. In this work, stock price data of twenty years of a leading energy company is used to predict option prices directly from the BS closed-form solutions. Afterwards, four different machine learning (ML) models, namely linear regression, XGBoost regression, support vector machine, and random forest regression, are used to forecast the stock prices 3 days and 10 days ahead of option maturity. The performances of all these ML models are compared using the coefficient of determination R^2 , root mean square error (RMSE) and mean absolute error (MAE). Among all, the XGBoost algorithm achieved the highest R^2 value of 0.85. Since this ML model showed the closest fit to the market stock prices, the residuals of option pricing using XGBoost have been calculated and hence compared with BS as a justification of the estimation. The result of the residue is around ± 0.11 . It indicates superior predictive capability compared to other approaches. The lower RMSE and MAE values further confirm that XGBoost provides the most accurate fit between predicted and actual option prices by analyzing the stock prices of the company. It is anticipated that this study will be helpful in preventing sudden stock market crashes and improving the stability of the market.

Keywords: Option Pricing, Black-Scholes, XGBR, Machine Learning, Forecasting

1. Introduction

Option pricing is an important player in the stock market and is rising up every year in terms of number of contracts and value. In 2023, Cboe Global Markets

*Corresponding Author email: tahmid.suki@bup.edu.bd

¹ Department of CSE, Bangladesh University of Professionals, Dhaka, Bangladesh

² IE, Bangladesh Atomic Energy Commission, Dhaka, Bangladesh

³ Department of ICT, Bangladesh University of Professionals, Dhaka, Bangladesh

DOI: <https://doi.org/10.64494/JFST/v4i1/ts/2026/02/23-40>

reported a record average daily options trading volume of 44.2 million contracts, an 8% increase from 2022 (Cboe Global Markets, 2023). In 2025, Bloomberg, using Cboe Global Markets data, reported that roughly \$4 trillion worth of U.S. equity options trade per day (Goyder, n.d.). In reference (Smith, 1976), option pricing is defined as the process of figuring out the fair value of these contracts by modeling the price behavior, volatility, and time to expiration of the underlying asset to guess what their expected payoffs will be when there is uncertainty. Options are financial contracts that give the holder no legal obligation to buy or sell the underlying asset but represent a right to exercise it by paying a cost for the option. This characteristic distinguishes options from other financial derivatives like futures and forwards. In general, there are two types of options: options that give the holder right to buy the underlying asset by a certain date (expiration date or maturity) for a certain price (exercise price or strike price) are calls, whereas options giving the holder right to sell it are puts. European-style options can only be practiced at maturity. On the other hand, American options can be exercised at any time till maturity (Hull, 2003).

There are many well-known methods of option pricing, such as the Binomial Option Pricing Model (BOPM), Monte Carlo Simulation Model, Black Model (Black–76), Black Scholes (BS) Model, etc. However, all mathematical models have their own advantages and limitations. For example, BS hardly reflects the real market observations due to its straightforward assumptions, such as constant variance. Thus, it cannot explain the volatility surface (Ke & Yang, 2019).

The BS Model, also known as the Black-Scholes-Merton model, is one of the most preliminary yet popular option pricing models. It holds the assumption that price of heavily traded financial instruments follows a Geometric Brownian motion with constant mean and variance. The focus of the model is to hedge options by buying and selling the underlying assets and, hence, reducing risk. Fischer Black and Myron Scholes introduced the primary idea of option valuation through this model. Robert C. Merton first ever published a paper explaining the mathematical understanding of this model by Itô process and stated the term “Black-Scholes option pricing model.” In 1997, Merton and Scholes received “The Nobel Memorial Prize in Economic Science” for this novel work. The BS model has a revolutionary impact on the world of finance. This model provides a simple closed-form solution for European options, making it a straightforward approach unlike other iterative methods, such as the binomial model or Monte Carlo simulations. Due to the simplicity of its assumptions, the Black-Scholes-Merton option pricing model makes itself the benchmark among all option pricing models in practice. Almost half a century after it was

introduced, a massive amount of literature has been and is still being written as extensions or modifications of the BS model by addressing the limitations of its assumptions. The Heston Stochastic Volatility Model and jump-diffusion models are good examples of such modified versions (Karagozoglu, 2022; Pesantez-Narvaez et al., 2019). As the Heston model takes random volatility into consideration, it simulates some sort of real market scenario such as volatility smile or volatility skew. It is true that it performs better in creating scenarios of real markets and also performs better than BS, but it fails to model extreme events like market crashes or sudden shock (Suki & Hossain, 2019).

Monte Carlo method is another popular technique in option pricing where it can handle complex features, and it works even where other model fails, for example in case of Asian options valuation. However, it is computationally very expensive, as the cost increases with the number of iterations.

Though there are many studies available on option pricing based on BS model, only a few approaches have been taken in the field of option pricing to predict the optimum option price corresponding to a stock price process using ML. It is true that some of the ML approaches have been taken to forecast option premium in some recent studies, but none of them compare their effectiveness with the BS model and combine the two to have better prediction.

In this study, an ML approach has been taken to forecast the stock option price 3 days ahead and 10 days ahead of the actual day. MATLAB uses 20 years of data and trains the ML models to predict the pattern of the stock values. It also predicted the option price of the stock using the BS model as well as the highest-performing ML model. In this way, a real effort has been made to combine and compare the performances of the renowned BS model with the most common ML models to estimate the accurate option price.

2. Methodology

In this study, the strike price is initially assumed from the stock data, and hence the cost of call and put options is estimated by comparing it with the current stock price using the BS analytic solution. Here, the strike price is the fixed value at which someone might buy/sell the item to be exchanged under the option contracts at the expiry date. If the estimated strike price is cheaper than the market price, then the holder of call option can make a profit by selling the stock. Thus, a call option gets more valuable as the stock price increases. This positive relationship between stock price and call price can be further analyzed using sensitivity analysis through option Greeks, namely 'delta,' as an extension of this study. Similarly, when the stock price is lower than the strike price, the call price becomes cheaper, and the holder might not exercise the option to limit their loss to the option price only. On the other hand, put options become cheaper with the

right to sell at a lower price than the current price, as it involves less risk for the contractor. This case is more likely to be a loss for the stockholders, but it still provides some extent of protection for them due to a sudden change in the stock market. In figure 1, for the strike price of p_2 , the option price is m_5 (both call and put), and this is the most balanced value for buying or selling.

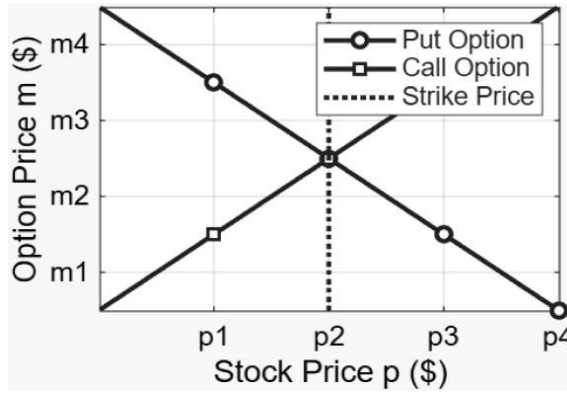


Figure 1: Call and put option price change with stock price.

This strategy is the basis of option pricing in the stock market. In the case of the ML approach, the stock prices are forecasted through different regression models and hence compared with the real market data. The ML technique that showed best compliance with real market price is further used to predict option prices and hence compared with the results from BS model with a view to justify approximations made by it. Figure 2 provides a step-by-step illustration of the procedures used in this study.

2.1 Data Compilation and Preprocessing

NextEra Energy, Inc. (NEE) company's historical stock market data spanning 20 years is collected from Yahoo Finance as CSV dataset containing Date, Open, High, Low, Close, Adjusted Close, and Volume (Yahoo Finance, n.d.). The historical volatility data has been collected from the S&P 500 over 20 years with the same data interval that has been added to the column of Yahoo Finance (*VIX Volatility Index | Historical Chart | Data | 1990-2026*, n.d.). The date column is converted to MATLAB datetime format for temporal consistency. Any non-numeric entries are converted to numeric form, and missing values are handled using forward-fill interpolation to preserve time-series continuity.

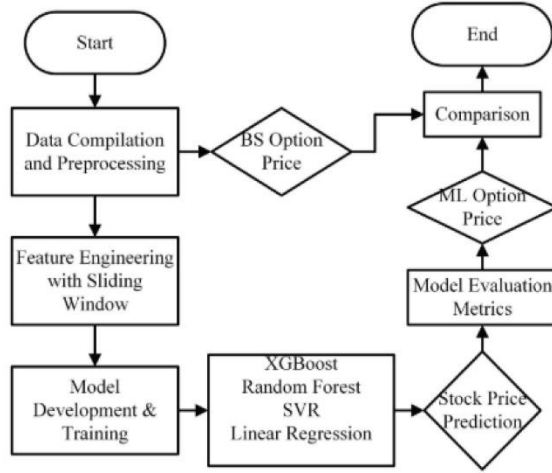


Figure 2: Structure of Methodology

2.2 Option Price Computation Using BS Model

BS model used the following parameters in equations (1), (2), and (3):

Risk-free interest rate, $r = 3\%$; Volatility, σ is computed from historical log returns; Time to maturity, T is 30 trading days (expressed in years); Stock price, S ; The strike price, K is equal to the net spot price (ATM options).

For each trading day, the parameters d_1 and d_2 were computed as:

$$d_1 = \frac{\ln(S/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, d_2 = d_1 - \sigma\sqrt{T} \quad (1)$$

Next, the call and put prices were computed as follows:

$$C = SN(d_1) - Ke^{-rT}N(d_2) \quad (2)$$

$$P = Ke^{-rT}N(-d_2) - SN(-d_1) \quad (3)$$

The resulting call and put prices are added to the dataset and subsequently utilized as extra variables.

2.3 Feature Engineering with Sliding Window Representation

Since financial time-series data has strong temporal dependence, for each prediction, the model looks back at the previous 60 trading days. For each set of training samples, the windows move forward one day. OHLC prices (Open, High, Low, Close), Adjusted Close price, Trading Volume, Computed Call Prices,

Computed Put Prices features are used to find the prediction framework. The target variable is set as the adjusted closing price three and ten days ahead, which enables short-term forecasting. Feature normalization was applied to ensure scale invariance across different magnitudes.

2.4 Model Development and Training

The dataset was divided into subsets of 20% for testing and 80% for training. The following machine learning models are trained and assessed using multiple regression:

2.4.1 Trees with Gradient Boost

XGBoost (eXtreme Gradient Boosting) constructs an ensemble of D Classification and Regression Trees (CARTs) sequentially, where each successive tree is trained to correct the residual errors of the earlier fitted model (Nova et al., 2025). The final prediction for the i -th observation is obtained by summing the contributions of all trees in the ensemble, as shown in equation (4).

$$\hat{Y}_i = F(X_i) = \sum_{d=1}^D f_d(X_i), f_d \in \mathcal{F}, i = 1, \dots, n \quad (4)$$

Where F denotes the function space of CART models. Each function f_d represents an independent tree structure q that maps the input features to leaf nodes, $q: \mathbb{R}^p \rightarrow \mathcal{T}$, with each leaf assigned a score $w \in \mathbb{R}^T$. The prediction of a tree is thus given by $f(X) = w_{q(X)}$, and the ensemble prediction is the sum of leaf scores across all trees (Pesantez-Narvaez et al., 2019).

2.4.2 Random Forest

Random Forest predicts by averaging the outputs of many decision trees, following equation (5).

$$\hat{Y}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (5)$$

Random Forest introduces randomness in feature selection by considering only a subset of predictors at each split, which decorrelates individual trees and enhances generalization (Olaniran et al., 2025).

2.4.3 Regression with Support Vector (SVR)

SVR is a supervised machine learning method developed for regression problems involving the prediction of continuous valued outputs. Consider a training dataset, $\{(x_i, y_i)\}_{i=1}^I$, where x_i denotes the input feature vector and y_i is the corresponding target value. In the linear ϵ -SVR setting, the regression function is defined as (6).

$$f(x) = \mathbf{w} \cdot \mathbf{x} + \mathbf{b} = \mathbf{w}^T \mathbf{x} + \mathbf{b} \quad (6)$$

where $\mathbf{w}$ is the weight vector and $\mathbf{b}$ is the bias term (Zhang & O'Donnell, 2020).

2.4.4. Linear Regression

To evaluate linear dependency assumptions, a baseline ridge-regularized linear regression model was incorporated. Simple linear regression estimates a quantitative dependent variable (DV), denoted as Y , based on a single independent variable (IV), denoted as X . It assumes an approximately linear relationship between X and Y . Mathematically; it is defined as equation (7).

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (7)$$

Where β_0 is the intercept, β_1 is the slope, and ϵ is the error term, which accounts for the variation in Y .

Multiple linear regression (MLR) means incorporating two or more predictor variables to model a continuous response variable. MLR can capture a greater proportion of the variability in the dependent variable compared to a univariate linear model. However, there are some assumptions that the multiple linear regression relies on. Like, the assumption of independence requires that residuals be statistically independent of one another; a linear relationship between the dependent variable and each independent variable (Starbuck, 2023).

2.5 Model Evaluation Metrics

Model performance was evaluated using the following metrics.

2.5.1 Mean Absolute Error (MAE)

The MAE measures the average magnitude of prediction errors without considering their direction. For a set of n observed values y_i ($i=1, 2, \dots, n$) and corresponding model predictions $\hat{y}_i$, MAE is defined as equation (8) (Hodson, 2022).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

2.5.2 Root Mean Squared Error (RMSE)

The RMSE quantifies the average magnitude of prediction errors while assigning greater weight to larger deviations. For n number of observations, it is defined as equation (9).

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (y_i - \hat{y}_i)^2} \quad (9)$$

RMSE is derived from the MSE and retains the same units as the response variable, facilitating interpretation. Due to the squaring of errors, RMSE is more sensitive to large prediction errors and is particularly informative when large deviations are undesirable (Hodson, 2022).

2.5.3 Coefficient of Determination (R^2)

The coefficient of determination R^2 measures the proportion of variance in the observed data that the predictive model explains. For a set of n observed values y_i ($i=1, 2, \dots, n$), $\bar{y}$ denotes the mean of the observed values, and the corresponding predicted values $\hat{y}_i$, R^2 is defined as equation (10).

$$R^2 = 1 - \frac{\sum_1^n (y_i - \hat{y}_i)^2}{\sum_1^n (y_i - \bar{y})^2} \quad (10)$$

An R^2 value of 1 corresponds to perfect agreement between observed and predicted values (Chicco et al., 2021). These metrics provided insights into prediction accuracy.

2.6 Forecasting Short-Term Prices

The trained ML model is utilized to forecast the stock price for the following day using the most recent sliding window in order to show practical applicability. This step uses data derived from past prices and options to simulate real-time forecasting. Predicted and actual prices were plotted for each model to visually assess forecasting performance.

Considering risk-free interest of 3% BS model has generated call and put option prices from (2) and (3) corresponding to the stock price, as shown in figure 3, using 20 years of stock data of NEE company through MATLAB with columns like high, low, open, close, volume, etc. The graph shown in figure 3 depicts how the stock prices for options (x-axis) are related to their corresponding prices (y-axis) according to the BS formula. The curve for call options (denoted in red) is higher at stock prices greater than the assumed strike price. On the other hand, as stock prices increase, the price for put options decreases, as represented by the blue line. The strike price is marked by the red dashed line close to \$78.

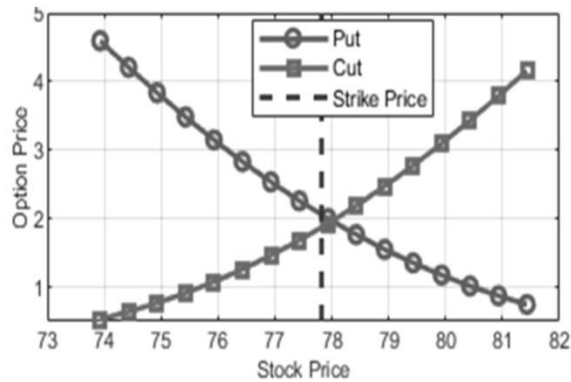


Figure 3: Call and put option vs stock price determined using the BS Model

At this point, the options are at-the-money. Neither call nor put options have any intrinsic value at this point. They possess considerable extrinsic values as a result of potential market changes.

Now the prediction of stock prices from ML is shown in figures 4 and 5. It is easily observed that in the case of 3 days ahead prediction depicted by figure 4, the anticipated stock prices from both XGBoost and Random Forest show a close match with actual price for about 800 samples. Except for some small dissimilarities of these two models around 100 and 200 samples, and for all the other points, it is almost impossible to distinguish their respective plots from real market prices at all other points. However, in the case of the SVR and Linear regression models, the match between the actual and predicted stock prices is not as close as the two methods mentioned above. The SVR approximations always end up with higher values, as observed in figure 4. Similarly, it is quite evident for linear regression models that the price difference between the prediction and actual price tends to \$15 mostly.

In the case of ten days ahead of stock price prediction, the gap between the estimated and actual price is shown in figure 5. It is quite logical that as the time gap increases, the number of factors that may impact market dynamics multiply over time. For example, market factors like natural disasters or changes in demand, etc., might change the value of the stock long-term ahead of the real timeline. This phenomenon fuels the chance of error in prediction. Here, the gap between the real stock price and the forecasted price in the case of XGBoost is about \$4 at most, observed around sample 550. The gap extends in the case of Random Forest, SVR, and linear regression. However, the XGBoost forecasting still carries a pattern analogous to the real data for 10 days ahead of real time.

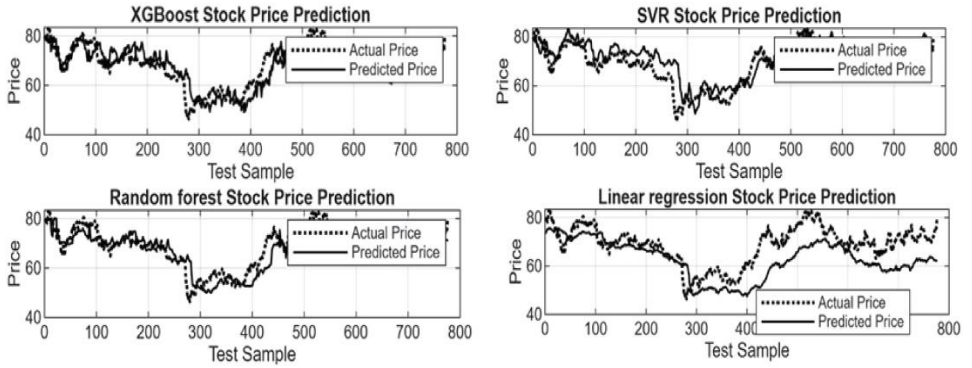


Figure 4: Comparison of 3 days ahead of the predicted stock price to the actual price using ML models

To obtain better results than the other models taken into consideration in the study, grid search CV, a technique commonly used for hyperparameter tuning (Suzana & Kabir, 2025), was applied. In the optimization process of the hyperparameters, a Min Objective vs Function Evaluations has been done, and it is observed that in the beginning, there is a large decrease in the objective value, which indicates that the optimization process quickly identified a significant performance improvement. After several evaluations, both the observed and estimated objectives become flat, which suggests the optimizer found the optimal or nearly optimal solution shown in figure 6.

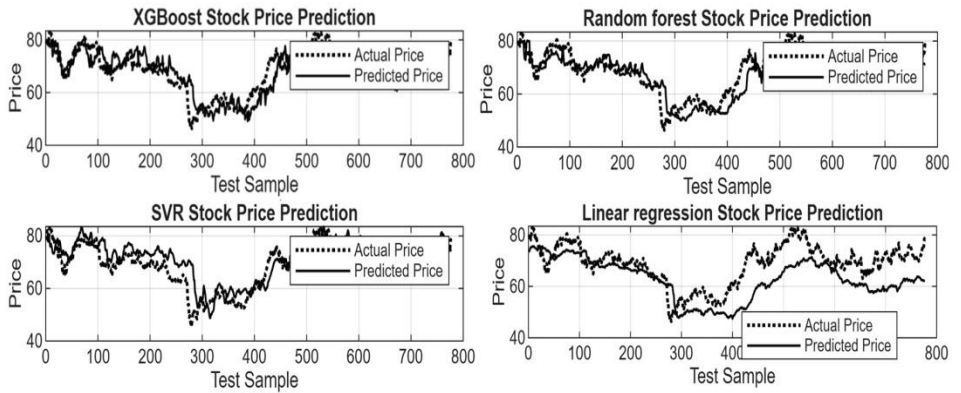


Figure 5: Comparison of 10 days ahead of the predicted stock price to the actual price using ML models

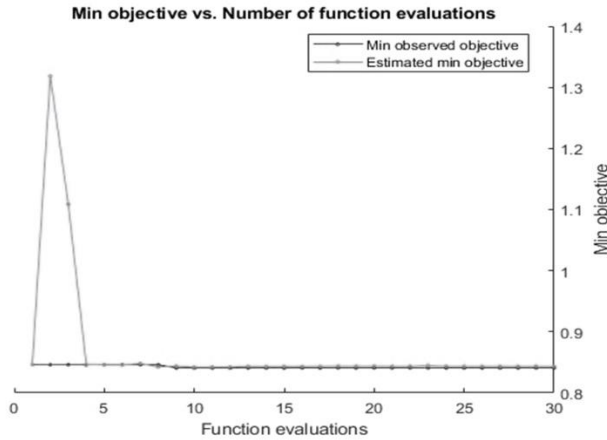


Figure 6: Optimization of the hyperparameters of the XGBoost

Table 1: Stock Price Prediction Model Performance Comparison

Model	Metrics					
	RMSE (10-Day)	MAE (10-Day)	R ² (10-Day)	RMSE (3-Day)	MAE (3-Day)	R ² (3-Day)
XGBoost	4.587	3.639	0.645	2.978	2.361	0.850
SVR	7.196	5.872	0.128	3.368	2.647	0.809
Linear	9.110	7.759	-0.396	12.581	11.889	-1.658
Random Forest	4.287	3.338	0.690	2.979	2.373	0.850

To understand the accuracy of the prediction as well as the validation of different models in terms of RMSE, MAE, and R^2 values has been used. It helps to select the best model to predict the stock and the best fit to the data. Figure 7(a) depicts the model's performance in predicting 3 days ahead of actual options and describes how well XGBoost, SVR, Linear, and Random Forest regression performed. It is observed that both XGBoost and Random Forest regression both have low error in terms of RMSE and MAE. They also fit well with the data, as the R^2 is close to one for both models. On the other hand, SVR has slightly higher error rates, and correlation with the data is decent. However, the linear regression model R^2 value is negative, which indicates the model does not fit the data well and performs worse than a simple mean-based model. In addition, the errors are also pretty much higher than all other models. The pattern is similar for 10 days ahead of prediction of the option price shown in figure 7(b), where both XGBoost and Random Forest regression have lower error and higher R^2 value

compared to linear regression and the SVR model. Though the SVR model results in further errors, R^2 value is still positive. But in case of linear regression, no correlation has been established between data, and it is obvious that the error is around 8 from the actual option price. The negative R^2 value, which is possible for poor models like the linear regression here, indicates that the model is worse than just a mean horizontal line and worse at predicting.

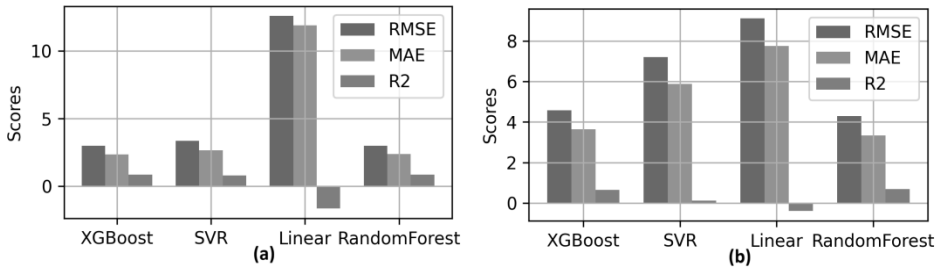


Figure 7: (a) Three days ahead prediction performance of different models in predicting stock price; (b) ten days ahead prediction performance.

Though XGBoost and Random Forest showcase similar R^2 values, other error metrics (RMSE, MAE) are lower for XGBoost in the case of predicting stock prices 3 days ahead of maturity. On the other hand, XGBoost outperforms Random Forest in case of 10 days ahead stock price prediction, as its R^2 value is lower than Random Forest, with other error metrics being close enough for these models. Nevertheless, XGBoost is more widely used for high-performance and competition settings, while Random Forest remains mostly used for its simplicity and reliability.

After finding the best prediction model, which is XGBoost in the study, we used it to predict the option prices 10 days ahead of time. Since both option price and stock price depend on similar underlying features like asset price and historical data, models that predict the best stock price can also predict a reliable option price. The prediction of both call and option prices is shown in figure 8(a). The predicted option price is then compared to the BS model-based option price, and the residual of these two models is depicted in figure 8(b). It is observed that the price difference of the BS model and XGBoost model is almost between ± 0.2 , which can be seen in the figure 8(b), except for a few samples when the difference of the ML predicted option price differs from BS model predicted option price. For example, around sample 280, the residue of the two predictions is at a peak, which is around 0.5. The RMS of the residue has also been calculated, which is 0.129 for the call option and 0.110 for the put option. In

comparison to the Black-Scholes benchmark, the ML-predicted option prices show small residuals (RMSE ≈ 0.11 – 0.12). It suggests that the model captures additional nonlinear patterns not included in the standard BS model.

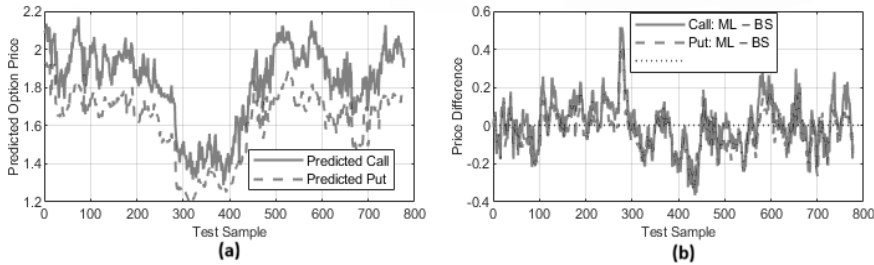


Figure 8: (a) Prediction of option price using XGBoost; (b) Comparison of price differences of ML model and BS model

In the following discussion, we have run a stress test to compare the performance of all the ML models discussed so far to analyze and compare their performance in case of high volatility, such as the 2020 COVID-19 market shock. To do this, we exported as mentioned in 2.1. Figure 9 shows how the ML models (SVR, XGboost, Linear & Random Forest) perform in the normal market vs the shock market. Additionally, to prevent models from seeing the future in this training time series, a cross-validation procedure has been followed rather than simple 80/20 percent splitting.

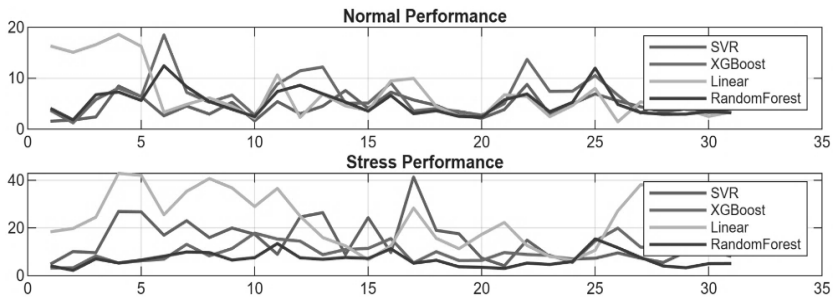


Figure 9: Normal & Stress performance of ML models after adding volatility index

Table 2 shows the RMSE error of different ML models after adding share market shock, particularly the shock of COVID-19 that occurred in 2021. Evidently, SVR & Linear Regression display remarkable error in shock. In contrast, Random Forest (RMSE 30.07%) and XGBoost (RMSE 51.02%) demonstrate more stable behavior.

Table 2: RMSE error of different ML models

Metrics Models	Normal RMSE	Stress RMSE	Degradation RMSE (%)
XGBoost	6.5519	9.8947	51.02
SVR	4.4047	14.7251	234.30
Linear	7.5351	23.2488	208.54
Random Forest	5.1862	6.7619	30.38

This prediction using the mathematical models, like BS and ML, helps millions of shareholders, especially the small investors, to decide the purchase of options, as it can predict a pattern of the stock price well ahead of time. Although option pricing is practiced all over the world, it has not yet been introduced in the share market of Bangladesh. Due to this reason, the stock exchange is more volatile and less controlled here. There are some notorious collapses of the share market in 2011 (Abu Rayhan et al., n.d.; Azad et al., 2014; Habib, n.d; Siddiquee & Begum, 2016) that made thousands of shareholders penniless overnight. However, the South Asian countries like Pakistan, Bangladesh, and India face a similar kind of manipulation of the market that makes the stock market more volatile (Azad et al., 2014; Gul et al., 2013). Right now, there are some options like international brokers like Interactive Brokers (IBKR), which offer extensive option products, pricing, and tools (*Best Options Trading Platforms in Bangladesh in 2026 – Fee Comparison Included*, n.d.). This creates a huge gap with the international stock market, and the stock market becomes less accessible. The option prices also restrict the manipulation of the stock price by miscreant people like arbitrators, both from inside the stock exchanges and outside of the system. The step towards option pricing is a positive move towards the transparency and stability of the future stock market of Bangladesh. This study would surely help in guiding the stock exchange market to the path of option pricing.

Finally, this work compares different ML models' performance against a baseline BS model rather than a realistic model like the Heston model that can incorporate volatility dynamics into the model. Hence, it restricts the scope for comparative evaluation. In addition, this study depends on historical data that may not perform well during market shocks or regime changes. The results of the ML models can be evaluated against the Heston model in future work, though it will add complexity because of the volatility dynamics.

4. Conclusion

The BS model, a widely accepted option pricing model, is aggregated along with four different ML models to predict the option prices well ahead of expiry date. The errors of ML models are also calculated, and hyperparameter tuning is done to optimize the XGBoost ML model (the high-performing model). Based on the results and error metrics, it can be concluded that XGBoost is the best model in predicting both call and put option prices compared to the other models considered in the study. The performance of the prediction can also be improved by accumulating more recent data and more features. Feature selection was very crucial in the study, which can impact the prediction results.

5. Limitations and Future Works

Lastly, this work depends on the historical dataset that does not work during the market crash because every time there is a market crash for very different reasons that might seem like a very little factor in a normal scenario, its impact on the market is huge during the sudden shock. Hence, this model might not predict well during sudden shock in the market. In the future, this study can also be extended towards the Monte Carlo method, which is one of the most used models in option price prediction.

Acknowledgement

The authors gratefully acknowledge Arkansas State University for its support of this research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data used in this study is publicly available from Yahoo Finance.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

AI Declaration

AI-assisted tools were solely used for proofreading, grammatical correction, and limited code debugging. No AI contributed to the data analysis decisions or interpretation of results.

References

- Abu Rayhan, M., Md Al-Emran Sarker, S., & Mohammad Sayem, S.-(2011). The Volatility of Dhaka Stock Exchange (DSE) Returns: Evidence and Implications. *ASA University Review*, 5(2), 87-99.
- Azad, A. S. M. S., Azmat, S., Fang, V., & Edirisuriya, P. (2014). Unchecked manipulations, price–volume relationship, and market efficiency: Evidence from emerging markets. *Research in International Business and Finance*, 30, 51–71. <https://doi.org/10.1016/j.ribaf.2013.05.003>
- Best Options Trading Platforms in Bangladesh in 2026 – Fee Comparison Included (2026). <https://brokerchooser.com/best-brokers/best-options-platforms-and-brokers-in-bangladesh>. [Access Date: 20 December 20 25]
- Cboe Global Markets, I. (2023). *Annual Report 2023*.
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, <https://doi.org/10.7717/peerj-cs.623>
- Goyder, B. (2025). Options Clearing Corp. Slams \$4 Trillion Options Market's Loophole Shut - Bloomberg. *Bloomberg* from <https://www.bloomberg.com/news/articles/2025-08-29/options-clearing-corp-slams-4-trillion-options-market-s-0dte-loophole-shut>. [Access Date: 10 November 2025]
- Gul, S., Khan, M. T., Saif, N., Rehman, S. U., & . R. (2013). Stock Market Reaction to Political Events (Evidence from Pakistan). *Journal of Economics and Sustainable Development*, 4(1), 165–174. <https://iiste.org/Journals/index.php/JEDS/article/view/4115>
- Habib, A. (2024). Why stock market was overrun by price manipulations | The Daily Star. *The Daily Star* from <https://www.thedailystar.net/business/economy/news/why-stock-market-was-overrun-price-manipulations-3717441>. [Access Date: 4 May 2026]
- Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, 15(14), 5481–5487. <https://doi.org/10.5194/gmd-15-5481-2022>
- Hull, J. C. (2003). *Options, Futures, And Other Derivatives* (D. Battista (Ed.); 8th ed.). Pearson Education Ltd.

- F.Sarkar, M.M. Rahman;Moni (2024). An Empirical Analysis of Capital Market Manipulation in Bangladesh and Its Impact on Investors' Confidence, *ICAB publication*, pp. 12–24.
- Karagozoglu, A. K. (2022). Option Pricing Models: From Black-Scholes-Merton to Present. *The Journal of Derivatives*, 29(4), 61–80. <https://doi.org/10.3905/jod.2022.1.158>
- Ke, A., & Yang, A. (2019). Option pricing with deep learning. *Department of Computer Science, Stanford University, In CS230: Deep learning*, 8, 1-8.
- Nova, R. M., Fahim, J. A., & Seema, S. J. (2025). A Study of Machine Learning Models with Population and Swarm Based Optimization for Cardiovascular Disease Prediction. *Journal of FST*, 3(1), 01–24. <https://doi.org/10.64494/JFST/v3i1/ss/2025/01/1-24>
- Olaniran, O. R., Olaniran, S. F., Alzahrani, A. R. R., Alharbi, N. M., & Alzahrani, A. A. (2025). Random Forest Adaptation for High-Dimensional Count Regression. *Mathematics*, 13(18), 3041. <https://doi.org/10.3390/math13183041>
- Pesantez-Narvaez, J., Guillen, M., & Alcañiz, M. (2019). Predicting Motor Insurance Claims Using Telematics Data—XGBoost versus Logistic Regression. *Risks*, 7(2), Article 70. <https://doi.org/10.3390/risks7020070>
- Siddiquee, M. N., & Begum, N. N. (2016). Volatility of Dhaka Stock Exchange. *International Journal of Economics and Finance*, 8(5), Article 220. <https://doi.org/10.5539/IJEF.V8N5P220>
- Smith, C. W. (1976). Option pricing. *Journal of Financial Economics*, 3(1–2), 3–51. [https://doi.org/10.1016/0304-405X\(76\)90019-2](https://doi.org/10.1016/0304-405X(76)90019-2)
- Starbuck, C. (2023). Linear Regression. In *The Fundamentals of People Analytics* (pp. 181–206). Springer International Publishing. https://doi.org/10.1007/978-3-031-28674-2_10
- Suki, T. T., & Hossain, A. S. (2019). A Comparative Analysis of the Black-Scholes- Merton Model and the Heston Stochastic Volatility Model. *GANIT: Journal of Bangladesh Mathematical Society*, 39, 127–140. <https://doi.org/10.3329/ganit.v39i0.44168>
- Suzana, N. N., & Kabir, K. H. (2025). HDP: Simulated IoT Architecture Design Applying GridSearchCV. *Journal of FST*, 3(1), 87–110. <https://doi.org/10.64494/JFST/ns/2025/04/87-110>

- VIX Volatility Index | Historical Chart | Data* | (2026). <https://www.macrotrends.net/datasets/2603/vix-volatility-index-historical-chart>. [Access date: 4 May 2026]
- Yahoo Finance. (2026). *NextEra Energy, Inc. (NEE) Options Chain* from <https://finance.yahoo.com/quote/NEE/options>. [Access date: 18 December 2025]
- Zhang, F., & O'Donnell, L. J. (2020). Support vector regression. In *Machine Learning* (pp. 123–140). Elsevier. <https://doi.org/10.1016/B978-0-12-815739-8.00007-9>