Pricing Options Using Machine Learning Algorithms

Jonathan Kinlay,
Intelligent Technologies

Aug 3, 2023

Abstract

In the intricate world of quantitative finance, the potential of machine learning to predict financial instruments' prices is increasingly evident. This research evaluates various machine learning models in forecasting vanilla put option prices. By generating a dataset rooted in the Black-Scholes model, the Neural Network model displayed outstanding out-of-sample performance, closely trailed by tree-based models.

Introduction

Options, pivotal in financial markets, serve in risk mitigation, hedging, and speculative strategies. Traditional models, such as Black-Scholes, have been foundational in option pricing. Yet, market evolution challenges the assumptions of these conventional models, ushering in potential pricing discrepancies. The advent of machine learning offers a potential avenue to apply these advanced algorithms in option pricing, potentially surpassing classical model limitations.

Methodology

Data Generation:

Emulating real-world scenarios, 100,000 option samples were synthesized, embedding parameters to mirror genuine market characteristics. A pronounced volatility skew, commonly observed for out-of-the-money options, was seamlessly incorporated.

Machine Learning Models:

We selected a diverse palette of five models, traversing from linear models to ensemble techniques and deep learning. The objective was a comprehensive evaluation, spanning models from the simplistic to the intricate, examining both their predictive acumen and elucidation capabilities.

Evaluation Metrics:

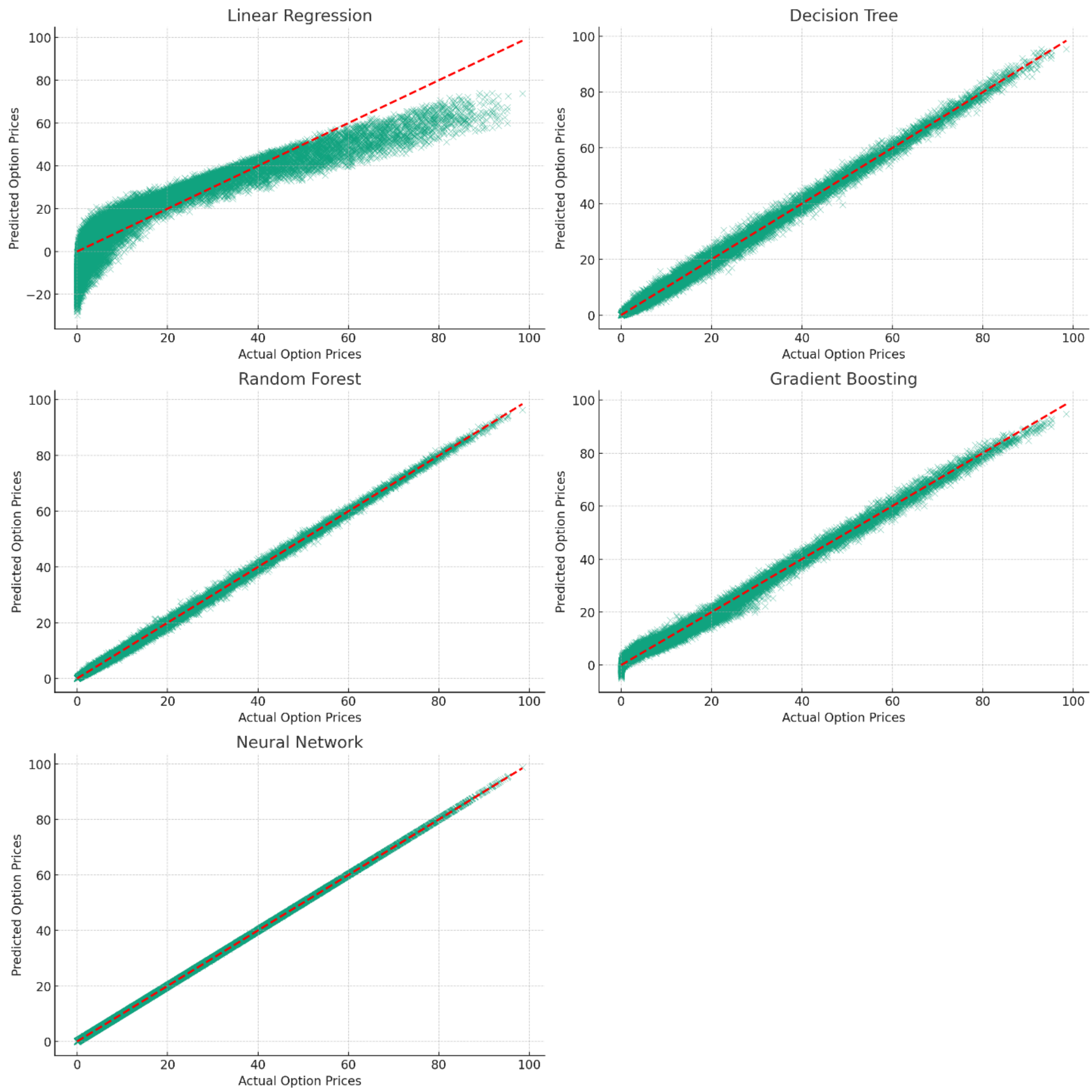
Model evaluations employed metrics illuminating both error magnitude (MSE, MAE) and the proportion of variance elucidated R^2.

Results

For a rigorous assessment, all models were tested on out-of-sample data. The summary of results is as follows:

Model	MSE	MAE	R2
Linear Regression	67.16	6.66	0.8384
Decision Tree	2.11	1.05	0.9949
Random Forest	0.5333	0.5131	0.9987
Gradient Boosting	3.14	1.39	0.9925
Neural Network	0.0102	0.0756	0.99998

Visual scatterplots, representing each model's actual vs. predicted prices for out-of-sample data, are shown below:



Discussion

A few observations emerge:

1. Neural Network's Superiority: The Neural Network's remarkable performance highlights the potential of deep learning techniques in the realm of financial modeling. Its near-perfect predictions underscore its capacity to capture complex nonlinear relationships inherent in option pricing.

2. Tree-based Models: While the Neural Network stood out, tree-based models, particularly the Random Forest, showcased impressive predictive abilities. These models offer the advantage of better interpretability, making them invaluable in scenarios where understanding model decisions is pivotal.

3. Linear Regression's Limitations: Linear Regression, though foundational, struggled to capture the complexities of option pricing, evident from its lower R^2 value compared to other models. This suggests that the relationships in the data might be inherently nonlinear, necessitating more intricate models.

Conclusion

This research underscores the potential of machine learning models, especially deep learning, in the realm of option pricing. Their ability to model intricate relationships and adapt to complex data structures makes them compelling candidates for financial modeling. However, while these models show promise, practitioners must exercise caution. Thorough validation on real-world data and a comprehensive understanding of model intricacies and limitations are imperative before deploying them in live financial scenarios.

References

Black, F., Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy.

Hull, J. (2017). Options, Futures, and Other Derivatives. Pearson.

Rubinstein, M. (1994). Implied Binomial Trees. The Journal of Finance.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. Springer.

Hyndman, R.J., Koehler, A.B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Breiman, L. (2001). Random Forests. Machine Learning.