Networks and Heterogeneous Media

https://www.aimspress.com/journal/nhm

NHM, 20(3): 987–1009.
DOI: 10.3934/nhm.2025043
Received: 08 June 2025
Revised: 01 September 2025
Accepted: 10 September 2025
Published: 18 September 2025

Research article

Reliable option pricing through deep learning: An anomaly score-based approach

Jihong Park[1], **Jeonggyu Huh**[2] **and Jaegi Jeon**[3,*]

[1] Department of Mathematics and Statistics, Chonnam National University, Gwangju 61186, Korea
[2] Department of Mathematics, Sungkyunkwan University, Suwon 16419, Korea
[3] Graduate School of Data Science, Chonnam National University, Gwangju 61186, Korea

* **Correspondence:** Email: jaegijeon@jnu.ac.kr; Tel: +82-625305795.

Abstract: We propose a neural-network variant integrating the Isolation Forest anomaly detection algorithm into its loss function. By incorporating anomaly scores as weights—effectively treating them as inverse measures of data reliability—the model suppresses outlier impact, yielding modest but consistent accuracy gains. Using KOSPI 200 option price data from 2019 to 2023, our experiments show that this anomaly-based approach enhances predictive accuracy by an average of 4.77% on the test set compared to a baseline neural network. Moreover, performance gains are generally observed across various market conditions, including different moneyness states, trading volumes, and time to maturity. Analysis of the identified anomalies reveals that trading volume and time to maturity are key factors strongly associated with irregularities in option data. Option moneyness also contributes to these irregularity patterns, particularly with other market conditions or at extreme levels. In contrast, interest rates show a less direct impact on anomaly scores in our dataset. These findings are broadly consistent with established market regularities, suggesting the anomaly detector's effectiveness in capturing characteristics of market inefficiencies or challenging pricing conditions. Overall, the proposed methodology contributes to the development of a more robust option pricing framework by better reflecting actual market dynamics. It shows potential during periods of heightened volatility, offering useful insights for further academic and practical applications.

Keywords: anomaly detection; data reliability; deep learning; Isolation Forest; option pricing
Mathematics subject classification: 91G20, 68T07

1. Introduction

Options are fundamental instruments in modern finance, widely used for hedging and speculation. Accurate option valuation therefore underpins market efficiency and robust risk management [1, 2].

Closed-form models, such as the Black–Scholes (BS) framework, rely on simplifying assumptions—including log-normal returns and constant volatility—that are frequently violated in real-world markets. These violations often lead to systematic pricing biases and phenomena like volatility smiles [3, 4], thereby motivating the exploration of more flexible, data-driven alternatives.

Since the pioneering work of Malliaris and Salchenberger [5] and Hutchinson et al. [6], neural networks have been increasingly explored as nonparametric option-pricing models. Subsequent research adopted more complex architectures, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), to capture spatial and temporal dependencies within option data (e.g., [7, 8]). For a broader survey of machine learning applications in option pricing, see, for example, Ruf and Wang [9]. Other studies have focused on infusing financial domain knowledge directly into the network design to improve consistency and predictive performance [10].

A persistent challenge in applying deep learning models to real-world financial data, however, is the prevalence of noisy, high-volatility, or stale quotes, often generated by market microstructure effects or abrupt volatility spikes. Short-maturity or thinly traded option contracts are particularly prone to such anomalous observations [11, 12]. If not appropriately handled, these outliers can disproportionately influence the model training process, skew gradient descent, and ultimately erode predictive accuracy and model reliability [13].

To address these concerns, researchers have developed various techniques to detect and mitigate the impact of outliers. Unsupervised anomaly detection methods—most notably Isolation Forest [14]—have proven effective in identifying atypical observations by, for instance, leveraging random partitioning. Such techniques have already found applications in finance, for example, in uncovering stock-market manipulation [15]. More recently, a strand of machine-learning research has focused on using anomaly scores not merely to discard data points but to quantify data reliability. This involves reweighting each sample's contribution during the training process, an approach that has demonstrated efficacy in diverse fields, including sensor prognostics and weakly supervised defect detection [16, 17]. Although promising in these domains (often distinct from finance), the specific application of such reliability-weighted learning to refine deep learning models for option pricing has seen limited exploration.

This study aims to bridge this gap. We propose and evaluate a novel deep learning-based option pricing model that incorporates unsupervised anomaly detection directly into its training mechanism. Specifically, our approach utilizes Isolation Forest-derived anomaly scores to dynamically weight each data sample within the neural network's loss function. This effectively reduces the influence of unreliable or anomalous option quotes without resorting to wholesale data deletion, enabling the model to learn more robustly from trustworthy information. We apply our methodology to a substantial dataset of KOSPI 200 index options from the Korea Exchange, analyzing model performance across various moneyness states, liquidity levels, and time-to-maturity periods.

Our contributions are threefold:

- we adapt reliability-weighted learning to the option-pricing setting by incorporating Isolation Forest anomaly scores directly into the loss function;
- we provide a comprehensive empirical validation of this approach using an extensive set of real-world market quotes;
- we demonstrate that the learned anomaly-based weights align with financial domain knowledge, thereby enhancing model interpretability and trustworthiness, alongside improving predictive accuracy.

The remainder of this paper is organized as follows: Section 2 reviews the fundamental concepts of European option pricing and the Isolation Forest anomaly detection method. Section 3 describes the KOSPI 200 options dataset used in this study, along with our data preprocessing pipeline and the architecture of the proposed anomaly-weighted model. In Section 4, we present a comprehensive evaluation of the model's performance, including benchmark comparisons and robustness checks. Section 5 provides an in-depth discussion of the results, interpreting the significance of the anomaly scores and performance implications, while also identifying avenues for future research. Finally, Section 6 concludes the paper with a summary of our findings and contributions.

2. Theoretical background

This section first reviews key concepts of European option valuation, including the put–call parity identity. Then, it summarizes the classical BS framework and outlines unsupervised anomaly detection, specifically Isolation Forest, on which our weighting strategy is built.

2.1. European option-pricing basics

An option grants its holder the right, but not the obligation, to buy (a call option) or sell (a put option) an underlying asset at a predetermined price, known as the strike price, on or before a specific future date (the expiration or maturity date). This study focuses on European-style options, which can only be exercised at the maturity date. Options serve as versatile instruments in financial markets, widely used for purposes such as hedging against risk and speculating on price movements.

The fair value or price of an option is influenced by several key factors. These primarily include the current market price of the underlying asset S, the strike price K, the time to maturity τ, the expected future volatility of the underlying asset's price σ, and the prevailing risk-free interest rate r.

The seminal BS model, introduced by Black, Scholes, and Merton et al. [1, 2], provided a theoretical framework for pricing European options by mathematically relating these five key factors. While the BS model offers a closed-form benchmark and identifies crucial pricing determinants, it relies on several restrictive assumptions, such as constant volatility and constant interest rates, as well as log-normal returns of the underlying asset. These assumptions are often violated in real-world markets [3, 4]. This discrepancy motivates data-driven approaches, like the one proposed in this study, that aim to learn pricing dynamics directly from observed market data without relying on these strict theoretical assumptions.

A crucial concept for characterizing options is moneyness. This term describes the relationship between an option's strike price K and the current price of its underlying asset S, which in turn indicates whether the option currently holds intrinsic value. For modeling purposes, we often use the ratio of the underlying asset price to the strike price, S/K. Based on this ratio, an option's state is typically categorized as in-the-money (ITM; e.g., $S > K$ for a call or $S < K$ for a put, indicating positive intrinsic value), at-the-money (ATM; $S \approx K$), or out-of-the-money (OTM; e.g., $S < K$ for a call or $S > K$ for a put, indicating no intrinsic value).

Generally, higher expected volatility tends to increase option prices, as it increases the likelihood of favorable price movements for the option holder. Conversely, as an option approaches its expiration date ($\tau \to 0$), its time value diminishes, typically leading to a decrease in its price, all else being equal.

Finally, a fundamental no-arbitrage relationship known as put–call parity connects the prices of European call options C and European put options P that share the same underlying asset S, strike price

K, and time to expiration τ, and are subject to the same risk-free interest rate r. The relationship is expressed as:

$$C - P = S - Ke^{-r\tau}. \tag{2.1}$$

This principle ensures consistency across option prices in an arbitrage-free market. As detailed in Section 3.1, this identity was utilized in our data preprocessing steps.

2.2. Isolation Forest anomaly detection

Real-world financial quote data often deviate from idealized theoretical assumptions due to factors such as market microstructure noise (e.g., bid–ask spreads, price discreteness), abrupt volatility spikes, or behavioral patterns like herding during periods of market stress [18]. Identifying and appropriately handling such anomalous observations is therefore a crucial prerequisite for developing robust predictive models in finance.

In this study, we employ the Isolation Forest algorithm [14], an efficient unsupervised method well-suited for detecting anomalies. It does not rely on distance or density measures, making it particularly effective in high-dimensional spaces. The core idea behind Isolation Forest is elegantly simple and founded on the principle that anomalies, being "few and different", are inherently more susceptible to isolation than normal data points. The algorithm operates as follows:

(i) It randomly partitions subsets of the data by constructing an ensemble of isolation trees (iTrees). For each tree, points are recursively partitioned by randomly selecting a feature and then a split value between the minimum and maximum values of that feature for the data points in the current node, until individual points are isolated or a predefined tree depth is reached.
(ii) It counts the number of splits required to isolate a specific data point x, known as its path length $h(x)$ in a given iTree. Anomalous points are expected to have significantly shorter path lengths on average.
(iii) To obtain a robust measure, the expected path length $E[h(x)]$ is calculated by averaging $h(x)$ for point x across all iTrees in the ensemble.

Because atypical points are few and different, they tend to be separated by fewer splits and thus exhibit shorter expected path lengths $E[h(x)]$.

To produce a normalized and interpretable anomaly score, Isolation Forest compares $E[h(x)]$ to an average path length estimate for a sample of size n (where n is the size of the subsamples used for tree construction). This estimate is given by $c(n)$, defined as:

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}, \tag{2.2}$$

where $H(\cdot)$ is the harmonic number (approximated by $\ln(\cdot) + 0.5772156649...$ (Euler's constant)), and n is the number of instances in the subsample. The final anomaly score $s(x)$ for a data point x is then defined as:

$$s(x) = 2^{-E[h(x)]/c(n)}. \tag{2.3}$$

This score, $s(x) \in (0, 1]$, provides an intuitive measure of anomaly: A score close to 1 indicates a high likelihood of being an anomaly (implying $E[h(x)] \approx 0$), whereas values near 0 suggest normality (implying $E[h(x)]$ is close to the expected average $c(n)$). While the threshold can be tuned, often

a score $s(x) > 0.5$ is considered indicative of a potential anomaly, a convention we note following Liu et al. [14]. The Isolation Forest algorithm has found applications in various domains, including finance—for instance, to detect stock-price manipulation (see also [15])—further underscoring its suitability for analyzing market data.

In our proposed framework, this anomaly score $s(x)$ serves as a quantifiable proxy for the reliability or trustworthiness of each data point. As detailed in Section 3.2, we integrate this score directly into the neural network's loss function to modulate the influence of each sample during training, thereby aiming to enhance model robustness against anomalous data points prevalent in financial markets.

3. Methodology

3.1. Data description

This study utilizes KOSPI 200 European-style option quote data obtained from the Korea Exchange (KRX)*. The dataset covers the period from January 2, 2019, to December 30, 2023, and includes daily closing prices along with relevant market variables required for modeling.

The primary input features selected for our neural network model are based on theoretical option pricing factors and market information:

- Moneyness (S/K): The ratio of the underlying KOSPI 200 index price S to the option's strike price K, as defined in Section 2.1. This normalization facilitates consistent comparisons across options with different strikes [6].
- Time to maturity (τ): The remaining life of the option, precisely computed in years based on the actual expiration dates and accounting for Korean holidays and specific non-trading days.
- Volatility (VKOSPI, annualized %): The Korean volatility index, analogous to the U.S. VIX, which reflects the market's expectation of future 30-day volatility derived from current KOSPI 200 option prices.
- Risk-free interest rate (KOFR, annualized %): Analogous to the U.S. Secured Overnight Financing Rate (SOFR), the Korea Overnight Financing Repo rate (KOFR) is the officially designated benchmark risk-free rate in Korea (since February 2021), providing a stable and credible measure.
- Trading volume: The number of contracts traded for the specific option during the day, indicating market activity and liquidity.
- Open interest: The total number of outstanding contracts for the option at the end of the day, reflecting market participation.

The target variable (P) for our prediction model is the option's daily closing price.

Several preprocessing steps were applied to ensure data quality and suitability for modeling:

1. **Zero-volume removal:** First, observations with zero daily trading volume were excluded to avoid using potentially stale or unreliable price quotes. Of the initial 1,172,586 records, 816,762 rows (approximately 69.7%) were removed. This initial filtering left 355,824 observations.
2. **Put-call parity adjustment:** Second, to standardize the dataset and effectively treat all options as puts (thereby eliminating option type as an explicit input feature), call option prices were converted to their corresponding implied put prices using the put-call parity relationship, Eq 2.1, as discussed

*Data obtained from the KRX Information Data System (http://data.krx.co.kr/)

in Section 2.1. This approach allows the model to learn directly from a unified set of prices and enforces theoretical consistency. Any resulting implied put prices that were negative were floored at zero. This practice of flooring or adjusting prices that violate no-arbitrage bounds is common in literature to maintain non-negativity while preserving sample size (see, e.g., [19]).

3. **Missing value removal:** Finally, any rows containing missing values in the selected input features or the standardized target variable (P) were removed to ensure complete records for modeling, yielding the final dataset of 346,675 observations used for training and evaluation.

Remark 1 (parity preprocessing)**.** *Let* (C^*, P^*) *satisfy Eq 2.1, and let observed quotes be* $C = C^* + \varepsilon_C$ *and* $P = P^* + \varepsilon_P$*. Converting calls via* $\widetilde{P} = C - S + Ke^{-r\tau}$ *gives* $\widetilde{P} - P^* = \varepsilon_C$ *under exact parity; thus the target is* P^* *plus a single-side microstructure error, regardless of whether the source quote was a put or a call. For mild parity violations, define the empirical gap*

$$\delta := C - P - (S - Ke^{-r\tau}) = \varepsilon_C - \varepsilon_P.$$

Then $\widetilde{P} = P + \delta$ *and hence* $\widetilde{P} - P^* = \varepsilon_P + \delta = \varepsilon_C$*. Empirically, both* ε_P *and* δ *tend to be larger in illiquid or extreme-maturity regimes, and the same regions are flagged by high anomaly scores, supporting the heteroskedastic modeling in Assumption 1. For a formal likelihood-based interpretation of the weighted loss, see Section 3.4.*

For the base model development and evaluation, this final cleaned dataset was randomly shuffled to prevent any ordering bias and then split into three subsets: 64% for training, 16% for validation, and 20% for testing. The random shuffling ensures that each subset generally reflects the overall distribution of the data. The use of time-series splits for further robustness checks is detailed in Section 4.3.

Table 2 summarizes the input features and the target variable used in the neural network. Descriptive statistics for these variables in the final processed dataset (2019–2023) are presented in Table 1.

Table 1. Descriptive statistics of the final dataset (N = 346,675, data from 2019 to 2023).

Variable	Mean	Std	Median	Min	Max
Moneyness	1.10	0.29	1.03	0.50	3.15
VKOSPI	19.56	6.88	17.75	1.00	69.24
Volume	3 998.09	16 321.83	42	1	480 413
Open interest	4 377.71	7 701.25	1 176	12	159 624
Interest rate	1.75	1.12	1.48	0.31	4.02
Time to maturity	0.24	0.20	0.17	0.02	1.08
Target price	20.53	31.53	4.65	0.01	237.15

3.2. Model architecture and loss

The proposed deep learning architecture combines a standard Multi-Layer Perceptron (MLP) for regression with a custom loss function weighted by anomaly scores from Isolation Forest. The key components are detailed below. Figure 1 visually summarizes this architecture.

Table 2. Summary of input and output variables for the neural network.

Type	Variable	Description
Input	Moneyness	Moneyness ratio (S/K)
	VKOSPI	Korean volatility index (annualized %)
	Volume	Daily trading volume (contracts)
	Open interest	Number of open option contracts
	Interest rate	Risk-free interest rate (KOFR, annualized % p.a.)
	Time to maturity	Remaining time to maturity (years)
Output	Target price	Option closing price (P)

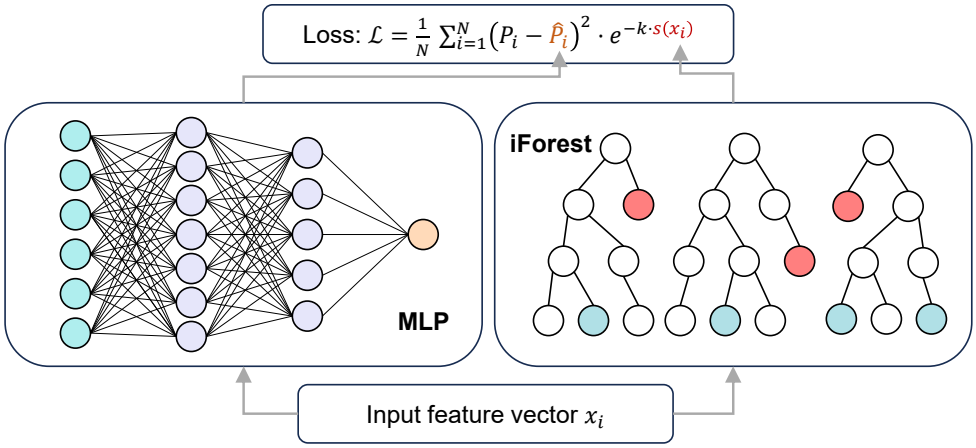


Figure 1. Overview of the anomaly-weighted MLP option-pricing model, illustrating the integration of Isolation Forest scores (from the same input features) into the custom loss function to guide robust training.

- **MLP network structure:** The core model is an MLP designed for the regression task of predicting option prices. It takes the six input features detailed in Section 3.1 (moneyness, VKOSPI, etc.) as input. Prior to being fed into the network, these features were selectively scaled. Specifically, interest rate and VKOSPI values were converted from percentages to decimals by dividing by 100, while trading volume and open interest were normalized using min-max scaling. The network consists of two fully connected hidden layers with 100 neurons each, utilizing the ReLU activation function [20] for nonlinearity, followed by a linear output layer with a single neuron predicting the put option price ($\hat{P}$). This architecture contains a total of $(6 \times 100 + 100) + (100 \times 100 + 100) + (100 \times 1 + 1) = 700 + 10100 + 101 = 10,901$ trainable parameters.
- **Anomaly-weighted loss function:** The distinctive feature of our approach lies in the loss function used for training. We modify the standard mean squared error (MSE) by incorporating the anomaly

score $s(x_i) \in [0, 1]$ obtained from the Isolation Forest algorithm (detailed in Section 2.2) for each data point x_i. The loss, denoted as $\mathcal{L}$, is calculated as a weighted MSE:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^{N} (P_i - \hat{P}_i)^2 \cdot w_i, \tag{3.1}$$

where P_i is the actual closing price (implied put price), $\hat{P}_i$ is the model's predicted price, N is the number of samples in the batch, and w_i is the anomaly-based weight for the i-th sample. This weight is defined as:

$$w_i = \exp(-k \cdot s(x_i)). \tag{3.2}$$

The exponential form of the weight (Eq 3.2) was chosen to provide a smooth, non-negative decay in influence as the anomaly score $s(x_i)$ increases. The hyperparameter k (determined experimentally as $k = 0.7$, see Section 4.1) controls the rate of this decay. This formulation assigns lower weights to samples deemed highly anomalous by Isolation Forest, thereby reducing their influence on the gradient updates and encouraging the model to learn primarily from more reliable data points.

- **Training details:** Network weights were initialized using He normal initialization [21], which is appropriate for ReLU activations. Optimization was performed using the Adam optimizer [22] with a learning rate of 5×10^{-3} (0.005) and a batch size of 500. An early stopping mechanism based on the validation set performance (monitored loss) was employed during training to prevent overfitting, with a patience of 20 epochs.

Code availability. Implementation details and scripts/notebooks used in this study are publicly available in our GitHub repository (see Code and data availability).

3.3. Choice of anomaly-detection method

To select the most suitable anomaly detector for our framework, we conducted a benchmark comparison between Isolation Forest, Local Outlier Factor (LOF), and One-Class SVM (OC-SVM). To establish a standardized benchmark across the different anomaly detectors, the raw scores from each method were mapped to a common range of [0, 1] via min-max scaling. This ensures a fair comparison and a consistent interpretation of the weights derived from our scheme Eq 3.2.

After min-max normalization on the training fold, we obtain scores $s \in [0, 1]$. Our loss uses the exponential weights in Eq 3.2; for a generic k, the support is $[e^{-k}, 1]$. Figure 2 displays the resulting distribution for Isolation Forest when $k = 0.7$, the value later selected by validation (see Section 4.1). For Isolation Forest, the normalized scores place relatively little mass near $s = 1$, so the induced weights are skewed but well spread rather than collapsing at the lower bound; this preserves a large effective sample size and yields steadier optimization while still down-weighting high-variance quotes. In contrast, LOF and OC-SVM exhibit heavier upper tails in s, which translates into more samples near the weight floor (Appendix A).

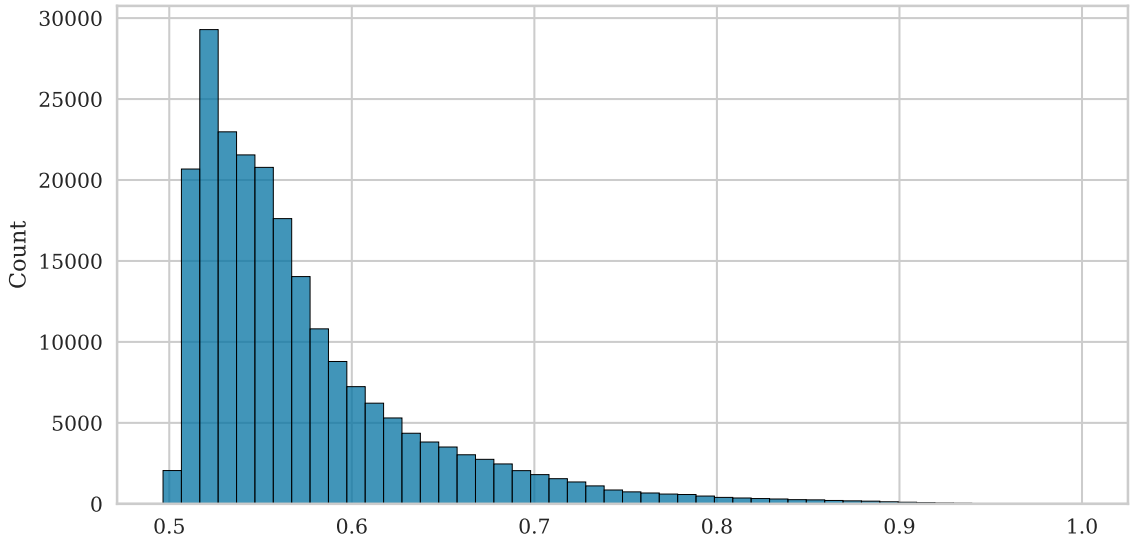


Figure 2. Distribution of Isolation Forest–induced weights $w = \exp(-0.7\, s)$ after training-fold min–max normalization of scores $s \in [0, 1]$. The weight support is $[e^{-0.7}, 1] \approx [0.496, 1]$.

In a 10-run benchmark under this primary protocol, Isolation Forest demonstrated a slight advantage in validation performance (best MSE of 4.742) and higher stability across runs compared to LOF (4.762) and OC-SVM (4.766). More critically, its computational efficiency was greater, with a fit time of just 2.3 s, compared to 12.6 s for LOF and 640.7 s for OC-SVM.

We further performed a sensitivity analysis using z-score standardization instead of min-max scaling. While the relative performance ranking remained, this change highlighted the instability of LOF, whose unbounded scores led to a significant degradation in downstream performance. Given its superior combination of predictive accuracy, stability, and computational speed, we selected Isolation Forest as the anomaly detection component for our model. Full results for both scaling protocols are detailed in Appendix A.

3.4. Formal probabilistic justification of the weighted loss

We formalize the proposed weighting via a heteroskedastic Gaussian regression in which error variance increases with the anomaly score through a log–linear link (Assumption 1). The aim of this subsection is to provide a probabilistic rationale for the exponential weights (Eq 3.2). We view $k \geq 0$ as a tuning parameter chosen by validation (Section 4.1), rather than estimated by (quasi-)likelihood; all results below are thus conditional on a fixed k. Under Assumption 1, Proposition 1 shows that, for any fixed k, minimizing our weighted squared loss over θ is equivalent to maximum likelihood under the log–linear heteroskedastic model.

Assumption 1 (Heteroskedastic Gaussian noise with log-linear variance link)**.** *Let $P_i \in \mathbb{R}$ be the (implied) put price observed at features x_i, and let $f_\theta : \mathcal{X} \to \mathbb{R}$ be a predictor with parameters θ. Conditional on x_i, assume*

$$P_i \mid x_i \;=\; f_\theta(x_i) + \varepsilon_i, \qquad \varepsilon_i \sim \mathcal{N}\big(0,\, \sigma_i^2\big), \quad \sigma_i^2 \;=\; \sigma_0^2 \exp(k\, s_i),$$

where $s_i = s(x_i) \in [0, 1]$ is the normalized anomaly score (computed on the training fold and held fixed on validation/test), $\sigma_0^2 > 0$ is a baseline variance, and $k \geq 0$ controls the rate at which the variance increases with anomalies. Assume moreover that, conditional on $\{x_i\}$, the errors $\{\varepsilon_i\}$ are independent across i.

Under Assumption 1, the negative log-likelihood is

$$\mathcal{J}(\theta, \sigma_0^2, k) = \frac{1}{2} \sum_{i=1}^{N} \left\{ \log(2\pi) + \log(\sigma_0^2 e^{k s_i}) + \frac{(P_i - f_\theta(x_i))^2}{\sigma_0^2 e^{k s_i}} \right\}. \tag{3.3}$$

Proposition 1 (MLE equivalence (for fixed k) to the exponential weighting)**.** *For fixed (σ_0^2, k), minimizing Eq 3.3 over θ is equivalent to minimizing the exponentially weighted squared loss Eq 3.1 with Eq 3.2*

$$\mathcal{L}_k(\theta) = \frac{1}{N} \sum_{i=1}^{N} (P_i - f_\theta(x_i))^2 w_i, \qquad w_i \equiv \exp(-k\, s_i).$$

Proof. Expanding Eq 3.3 and dropping terms independent of θ yields $\sum_i (P_i - f_\theta(x_i))^2 / (\sigma_0^2 e^{k s_i})$, which is proportional to $\sum_i (P_i - f_\theta(x_i))^2 e^{-k s_i}$.

Remark 2 (Choice of k in practice)**.** *Proposition 1 only requires k to be fixed and does not prescribe how it is chosen. In our experiments, we treat k as a hyperparameter and select it by validation (see Section 4.1); we do not estimate k from the likelihood in this paper.*

Because $w(x)$ depends only on x, the population target remains the Bayes predictor $f^*(x) = \mathbb{E}[P \mid X = x]$ (target invariance). Under mild smoothness and identifiability conditions, giving smaller weights to higher-variance observations reduces the estimator's large-sample (asymptotic) variance; a canonical choice is inverse-variance weighting $w(x) \propto \sigma^{-2}(x)$. With the log–linear variance link in Assumption 1, the exponential weights $w_i = e^{-k s_i}$ implement this idea: when k matches the variance-link slope, they coincide with σ_i^{-2} up to a constant multiplier; otherwise, they serve as a monotone surrogate. A simple mean-model illustration of variance reduction appears in Appendix B.

4. Results

This section presents the quantitative evaluation of the proposed anomaly-weighted option pricing model. We first detail the hyperparameter tuning process for the anomaly weighting, followed by a comprehensive comparison of the model's predictive performance against a baseline neural network on a randomly selected hold-out test set. Finally, we assess the model's robustness under varying market conditions using a time-series rolling split methodology. All experiments were implemented in PyTorch, primarily utilizing an NVIDIA GeForce RTX 3090 GPU, supported by an Intel Xeon CPU, for efficient model training and evaluation.

4.1. Hyperparameter tuning for anomaly weight (k)

The effectiveness of the anomaly weighting depends on the hyperparameter k in the weight function Eq 3.2. We determined the optimal value for k by evaluating the model's performance on the validation

set across a range of k values from 0.0 (equivalent to the Basic Model) to 1.0. Table 3 summarizes the validation loss results.

Table 3. Comparison of validation MSE (median of 10 runs) for different k-values.

k-value	0.0	0.1	0.2	0.3	0.4
Median validation MSE	0.7092	0.7092	0.6851	0.6776	0.7102
Standard deviation	0.1061	0.0513	0.0280	0.0763	0.0982
k-value	0.5	**0.7**	0.8	0.9	1.0
Median validation MSE	0.6788	**0.6722**	0.7285	0.6923	0.6885
Standard deviation	0.1106	**0.0666**	0.0649	0.0384	0.0756

As shown in Table 3, the validation loss was minimized when $k = 0.7$. Therefore, this value was selected for the Anomaly Model in all subsequent performance evaluations on the test set.

4.2. Benchmark comparison on random split

This subsection presents the predictive performance of the proposed Anomaly Model (incorporating anomaly scores with $k = 0.7$) against the Basic Model (a standard MLP without anomaly weighting) on the entire test dataset and across various challenging option characteristic subsets.

Overall performance comparison The Anomaly Model demonstrated improved predictive accuracy over the Basic Model on the entire test dataset. The performance metrics reported for this overall comparison are averages from five independent experimental runs, presented as mean ± standard deviation (SD). As indicated by the overall performance row in Table 4 (where results for specific categories are also detailed), the Anomaly Model achieved an overall test MSE of 5.37 ± 0.18, while the Basic Model recorded an MSE of 5.64 ± 0.38. This corresponds to a 4.77% reduction in MSE for the Anomaly Model.

Performance analysis across challenging scenarios The benefits of the Anomaly Model were further investigated across different option characteristics:

- **Moneyness:** Reflecting the conversion of all call options to their equivalent put prices, options were categorized based on moneyness S/K for these (implied) puts as follows: ITM ($S/K < 0.95$), ATM ($0.95 \leq S/K \leq 1.05$), and OTM ($S/K > 1.05$).
- **Time to maturity:** Options were grouped by their remaining time to maturity. These categories are referred to for brevity as "Two-week" and "One-month" (short-term); "Quarterly" and "Semi-annual" (medium-term); and "Annual" (long-term), with their precise definitions also provided in Table 4.
- **Trading volume:** Options were classified by daily trading volume (Very Low, Low, High, Very High), with definitions provided in Table 4.

The detailed MSE results for these analyses across moneyness, time to maturity, and trading volume are presented in Table 4.

Table 4. Test MSE comparison for Basic and Anomaly Models across different option characteristics (moneyness, time to maturity, and trading volume).

Characteristic	Category (definition)	Basic model (MSE ± SD)	Anomaly model (MSE ± SD)	Relative imp. (%)
Overall	–	5.6387 ± 0.38	5.3700 ± 0.18	4.77
Moneyness	ITM ($S/K < 0.95$)	15.4516 ± 1.13	14.7323 ± 0.54	4.66
	ATM ($0.95 \leq S/K \leq 1.05$)	2.2857 ± 0.12	2.1556 ± 0.07	5.69
	OTM ($S/K > 1.05$)	0.3329 ± 0.02	0.3165 ± 0.01	4.93
Time to maturity	Two-week ($\tau < 1/24$)	4.5915 ± 0.27	4.4851 ± 0.32	2.32
	One-month ($1/24 \leq \tau < 1/12$)	3.5399 ± 0.26	3.4055 ± 0.14	3.80
	Quarterly ($1/12 \leq \tau < 1/4$)	4.2575 ± 0.28	4.0904 ± 0.11	3.92
	Semi-annual ($1/4 \leq \tau < 1/2$)	6.4319 ± 0.53	6.0283 ± 0.23	6.27
	Annual ($\tau \geq 1/2$)	13.8190 ± 1.03	13.1170 ± 0.86	5.08
Trading volume	Very low (0–10)	2.6272 ± 0.21	2.4867 ± 0.10	5.35
	Low (10–100)	4.0540 ± 0.30	3.8703 ± 0.14	4.53
	High (100–1000)	5.4440 ± 0.32	5.2722 ± 0.25	3.16
	Very high (> 1000)	8.5818 ± 0.66	8.1110 ± 0.24	5.49

Referring to the moneyness categories in Table 4, the Anomaly Model showed lower MSE values than the Basic Model in all categories: ITM (Anomaly: 14.73 ± 0.54; Basic: 15.45 ± 1.13), ATM (Anomaly: 2.16 ± 0.07; Basic: 2.29 ± 0.12), and OTM (Anomaly: 0.32 ± 0.01; Basic: 0.33 ± 0.02). The relative MSE reductions were 4.66% for ITM, 5.69% for ATM, and 4.93% for OTM options, respectively.

For time to maturity, Table 4 shows that the Anomaly Model yielded consistent performance gains across all maturity categories, with improvements ranging from 2.32% for two-week options to 6.27% for semi-annual options.

Regarding trading volume, Table 4 indicates that the Anomaly Model achieved substantial improvements in both the Very Low volume (5.35% improvement) and Very High volume (5.49% improvement) categories, with consistent gains across all volume levels.

4.3. Robustness across rolling time-series splits

To assess performance under changing market conditions over time, a rolling train-test split methodology was employed using KOSPI 200 option data from 2019 to 2023. As illustrated in Figure 3, models were trained on one year of historical data and evaluated on the ensuing six months, resulting in eight distinct time splits. Each evaluation was repeated five times for statistical robustness.

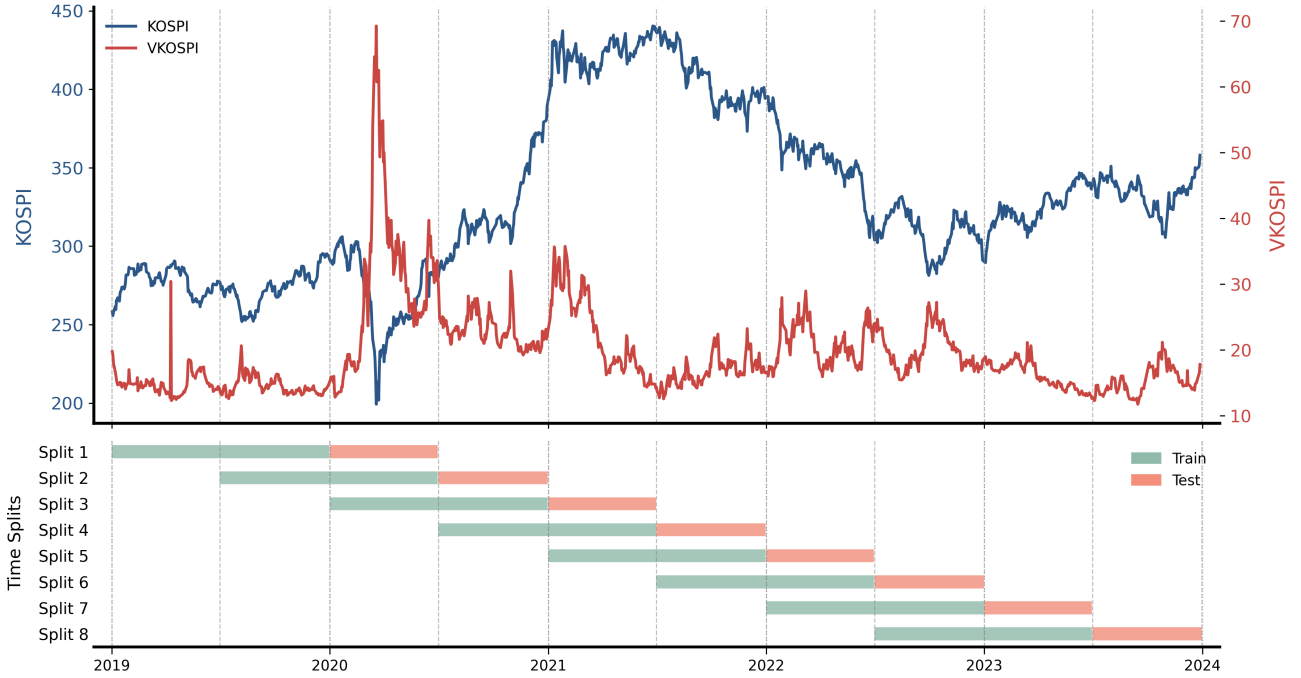


Figure 3. Market context and the rolling train-test split methodology. The top panel displays the daily KOSPI and VKOSPI indices for the 2019–2023 period, highlighting the varying market regimes. The bottom panel illustrates the structure of the eight time-series splits, where each split comprises a one-year training period (green bars) followed by a six-month testing period (red bars).

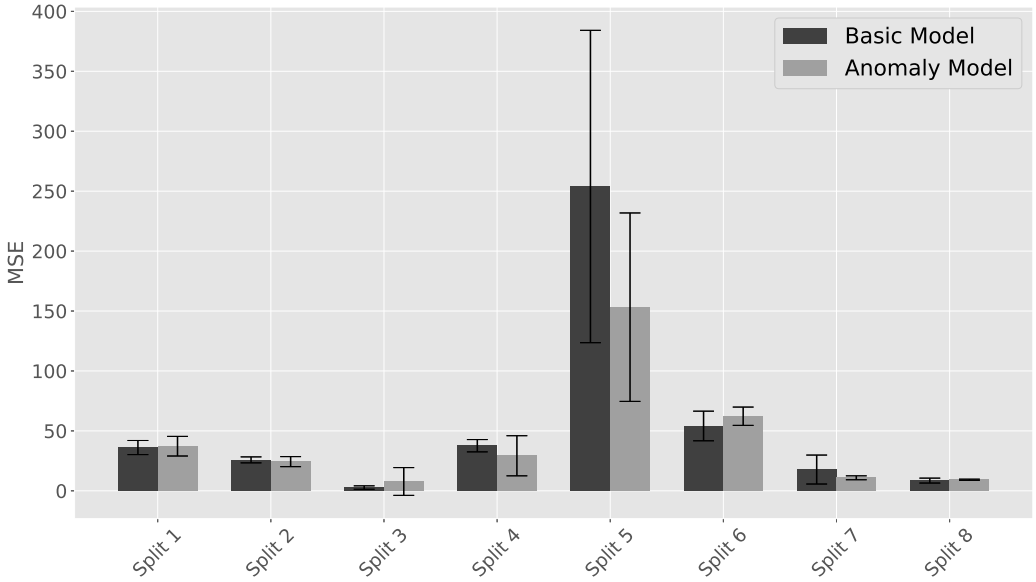


Figure 4. MSE comparison for the rolling time-series splits. Each bar indicates the mean MSE over five runs for a given split, with error bars representing one standard deviation. Lower values indicate better performance.

The aggregated MSE results are summarized in Figure 4. Overall, the Anomaly Model achieved comparable or, in several instances, improved predictive performance relative to the Basic Model across these time-series splits. For instance, during the first split, which tested the first half of 2020 coinciding with the COVID-19 onset, both models exhibited high prediction errors (Basic Model MSE = 36.11 ± 5.90; Anomaly Model MSE = 37.23 ± 8.17). In the second split, the Anomaly Model (MSE = 24.35 ± 4.18) slightly outperformed the Basic Model (MSE = 25.82 ± 2.50).

A particularly notable performance difference was observed in the fifth split (testing the first half of 2022). In this period, the Anomaly Model (MSE = 153.2 ± 78.8) substantially outperformed the Basic Model (MSE = 253.9 ± 132.2). However, the absolute MSE values for both models were considerably higher than in most other splits, as shown in Figure 4. Conversely, in the sixth split, the Basic Model (MSE = 54.10 ± 12.36) outperformed the Anomaly Model (MSE = 62.24 ± 7.64).

5. Discussion

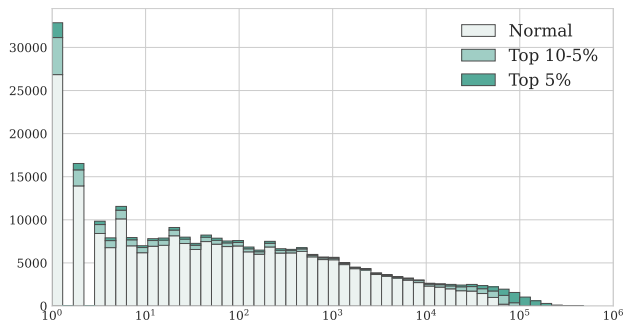
The quantitative results presented in Section 4 indicate that the proposed anomaly-weighted neural network offers improvements in option pricing accuracy and robustness. This section delves into the interpretation of these findings, focusing on why the anomaly weighting scheme is effective and discussing the implications of the model's performance. It then outlines avenues for future research that build upon this work.

5.1. Interpretation of anomaly scores and weighting effectiveness

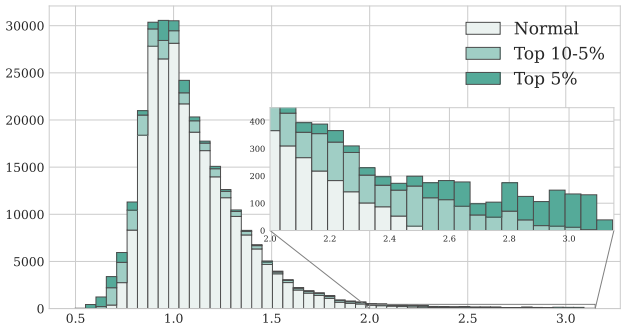
The effectiveness of our model hinges on its ability to leverage anomaly scores from Isolation Forest to differentiate data reliability. The underlying premise that these scores identify observations from challenging or inefficient market conditions is well supported by their distribution across key financial features (Figure 5).

Notably, higher anomaly scores were concentrated in scenarios widely recognized as problematic for reliable pricing. For instance, options with very low trading volume (Figure 5a) received high scores. Similarly, the distribution for time to maturity (Figure 5d) reveals that options with both very short and very long remaining maturities were frequently flagged as anomalous. This is consistent with financial intuition. Options with a short time to maturity are prone to pricing instabilities due to heightened sensitivity (e.g., high gamma), particularly for at-the-money options. Conversely, options with a long time to maturity, much like those with generally low trading volume, often suffer from illiquidity, leading to wider bid-ask spreads, price discreteness, or stale quotes that do not reflect current market information.

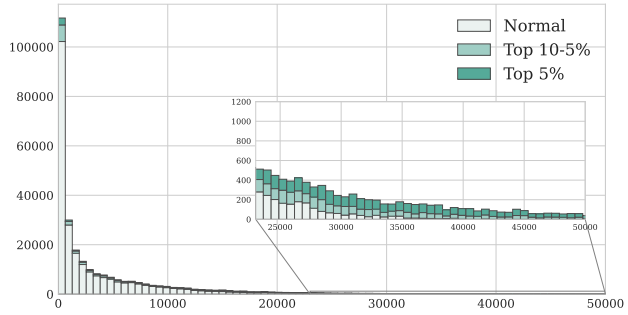
Periods of high market volatility (VKOSPI exceeding 50, Figure 5f) also correlated with increased anomaly scores, reflecting inherent uncertainty and price fluctuations. Interestingly, a concentration of anomalies was also observed at very high trading volumes, potentially indicating unusual market events or noise from large trades. A similar pattern was observed for options with very high open interest (Figure 5c). High open interest often signifies a "crowded trade" or concentrates around key strike prices near major events like expiration days (e.g., the quarterly simultaneous expiration day in the Korean market, similar to a Triple Witching Day). On such days, large-scale hedging or unwinding pressures can lead to temporary price dislocations that deviate from typical model behavior. The model's ability to flag these instances as anomalous—despite high liquidity—demonstrates its sensitivity not just to illiquidity but also to unusual market dynamics driven by concentrated positions.



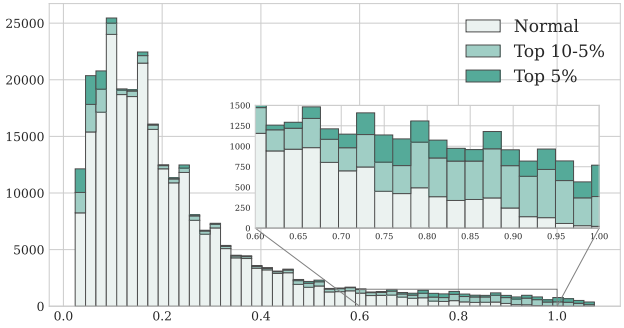
(a) Trading Volume.



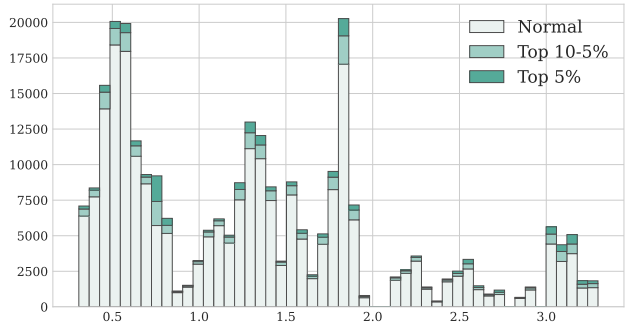
(b) Moneyness.



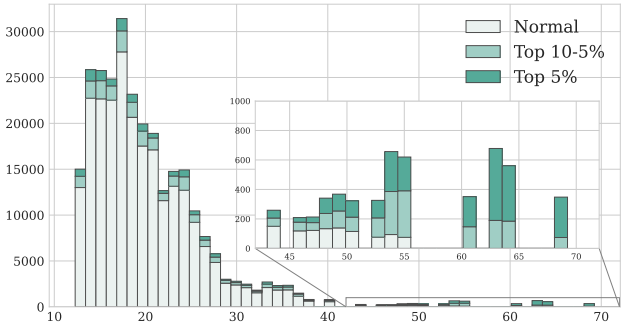
(c) Open Interest.



(d) Time to Maturity.



(e) Interest Rates.



(f) Volatility Index (VKOSPI).

Figure 5. Composition of anomaly score levels across the distributions of key input features. In each subplot, data points are categorized based on their Isolation Forest anomaly score: "Normal" (scores below the 90th percentile), "Top 10%-5%" (scores between the 90th and 95th percentiles), and "Top 5%" (scores above the 95th percentile). Higher color intensity (darker shades) indicates a greater concentration of highly anomalous data within a given range of the input feature. The subplots show the distributions for (a) trading volume, (b) moneyness, (c) open interest, (d) time to maturity, (e) interest rates, and (f) VKOSPI.

The impact of moneyness (Figure 5b) also reveals a pattern consistent with market liquidity. While ATM options tended to have lower anomaly scores, the scores generally increased as options moved further away from the money, either deep ITM or deep OTM. This phenomenon is well-aligned with trading behavior, as these deep ITM/OTM options are typically less liquid and less actively traded than their at-the-money counterparts. Such illiquidity often leads to wider bid-ask spreads or stale price quotes, which our anomaly detection model correctly identifies as less reliable. In contrast, interest rates (Figure 5e), on their own, did not exhibit a strong, independent pattern, suggesting that their impact on data reliability is likely more intertwined with other market conditions.

The alignment of these Isolation Forest-identified anomalies with established financial intuition lends economic credibility to the weighting scheme. By systematically down-weighting these less reliable observations, the Anomaly Model focuses on more stable and representative data. This refined learning process underpins the observed enhancements in predictive accuracy, particularly in illiquid or highly sensitive market segments, as detailed in Section 4. The model's capacity to mitigate noise, even in very high volume scenarios, further underscores its utility.

5.2. Implications of model performance

The Anomaly Model showed consistent, albeit modest, overall improvements in predictive accuracy on the random split test set (see Section 4.2 for details). The anomaly-weighting strategy appeared beneficial in certain scenarios, particularly for critical ATM options and across various liquidity conditions. Modest MSE reductions were observed in most liquidity buckets, including in thinly and extremely active markets (refer to Table 4).

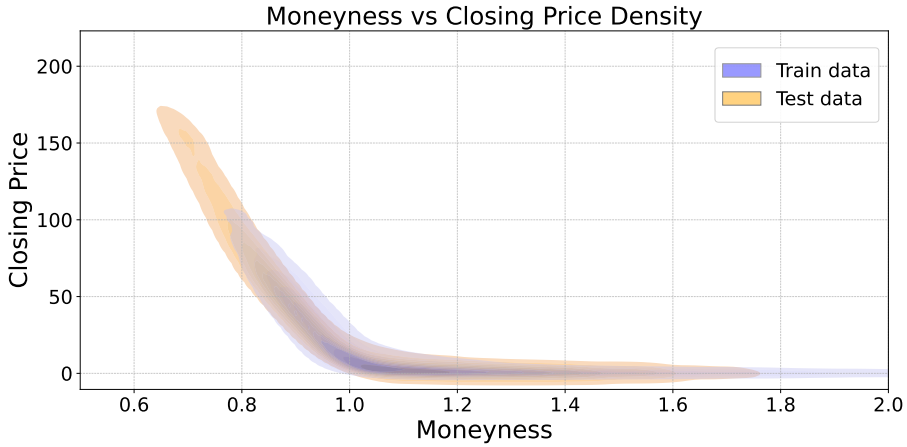


Figure 6. Density plot comparing distributions of training data (2021) and test data (first half of 2022) for Time Split 5 (moneyness vs. closing price), highlighting a distributional shift.

The rolling time-series analysis provided further insights into the model's robustness. Contrary to expectations of high prediction errors during extreme market stress, both models performed with relative stability during the COVID-19 onset (first half of 2020), with the Anomaly Model showing only a marginal advantage. A more pronounced performance difference emerged during the continuous bear market of the first half of 2022 (Split 5). While the absolute prediction errors for both models increased significantly in this period, the Anomaly Model's MSE was considerably lower than that of the Basic Model. This general performance decline in Split 5 appears to be linked to the distributional shift

illustrated in Figure 6. Specifically, the training data (2021) contained very few ITM options, whereas the test set (first half of 2022) included a large number of ITM, and particularly Deep ITM, options. This suggests that both models struggled when predicting on data types that were underrepresented during their training phase. The relative outperformance of the Anomaly Model in this challenging split indicates that its weighting mechanism may help in adapting to such regime shifts by moderating the influence of unfamiliar or outlier data points.

However, the Anomaly Model's superiority was not universal across all splits, with the Basic Model performing better in some instances (e.g., Split 6). This highlights that any relative advantage is contingent on the specific nature of market dynamics and anomalies within each period. The occasionally higher MSE variance for the Anomaly Model in certain volatile splits, combined with its apparent capacity to better adapt to certain distributional shifts as seen in Split 5, remains a noteworthy characteristic that warrants further investigation.

5.3. Limitations and future work

While this study provides a solid foundation for anomaly-aware option pricing, several avenues remain for future research to build upon these results:

- **Robustness to data scarcity and distributional shifts:** Our rolling time-series analysis revealed that the model's performance was highly sensitive to the specific training window. While our anomaly-based weighting provided benefits in specific regime shifts, further exploration using varied training window lengths or expanded datasets could better clarify these advantages. This contrasts with the more stable results from the random split benchmark, which benefited from a much larger and more diverse training set. This suggests that while our anomaly-weighting is beneficial, its effectiveness is maximized when the model is trained on a rich, representative dataset. Therefore, future work should not only implement a more granular walk-forward validation framework—exploring schemes like expanding windows in addition to the rolling window approach—to better simulate live trading but also explore methods for robust training on datasets that may be limited or subject to regime shifts, such as transfer learning or data augmentation techniques.
- **Model scalability and expressiveness:** A compact MLP was employed to clearly isolate the impact of the anomaly weighting mechanism. A significant next step involves integrating this anomaly-aware loss function into more expressive and scalable architectures, such as Transformer-based models for capturing long-range dependencies or graph neural networks (GNNs) designed to exploit the interconnected structure of option chains.
- **Anomaly characterization and weighting calibration:** The current approach relies on a general-purpose anomaly detector (Isolation Forest) and a statically tuned weighting hyperparameter (k). Future refinements could explore the use of domain-specific anomaly detection techniques tailored to financial market idiosyncrasies. Additionally, investigating more sophisticated, potentially dynamic or learnable calibration methods for the weighting function could enhance its adaptability to varying market states and data characteristics.
- **External validity and broader benchmarking:** The empirical validation of this study focused on KOSPI 200 European options. Key extensions will involve assessing the methodology's generalizability across diverse asset classes (e.g., individual equities, commodities), different option

types (e.g., American, exotic options), and varied market conditions. Furthermore, conducting comprehensive benchmarking against well-calibrated traditional financial models (e.g., Heston, Bates) will be crucial for a fuller understanding of the practical advantages and relative positioning of this data-driven approach.

Finally, although our experiments focus on KOSPI 200, the pipeline is structurally market-agnostic. A detector produces raw reliability signals that are mapped on the training fold to normalized scores $s \in [0, 1]$ (the same map applied to validation/test), and the network is trained with exponential weights $w = \exp(-ks)$ (see Section 3.2). Adapting the model to a new market therefore amounts to exchanging market inputs and performing light preprocessing while keeping this core structure unchanged.

For equity index and single-stock options, moneyness and time-to-maturity remain universal features; the risk-free rate input is mapped to the local benchmark curve (e.g., SOFR in the U.S. market), and dividends can be handled by introducing a yield q or by working with forward prices $F = S\, e^{(r-q)\tau}$. This principle extends to foreign exchange and commodity options, where the forward calculation incorporates the foreign risk-free rate or a convenience yield, respectively. For European-style options, our put–call parity unification and non-negativity treatment carry over directly. For American-style options, one may subtract a tractable early-exercise premium to create a European-equivalent target. Because the scores are normalized ($s \in [0, 1]$), the weights are bounded within $w \in [e^{-k}, 1]$, which preserves a large effective sample size and stabilizes optimization across regimes (see Appendix B).

6. Conclusion

In this study, we explored a method to enhance option price prediction in financial markets by combining deep learning with an anomaly detection algorithm. Traditionally, option pricing has relied on mathematical models such as the BS framework. However, given the high nonlinearity, volatility, and noise inherent in financial data, deep learning's potential advantages have increasingly drawn attention. Against this backdrop, we used KOSPI 200 option data from the KRX to propose an MLP model that incorporates anomaly scores—derived from the Isolation Forest method—into the loss function as weights. This approach aims to mitigate distortions and noise arising under extreme market conditions, thereby more accurately reflecting actual trading data and improving predictive performance.

Experimental results show that the model utilizing anomaly scores generally achieves lower prediction errors than the baseline model across various conditions, including option types, trading volume levels, and maturities. The benefits were most evident in scenarios of extremely low or extremely high trading volume and for short-dated options exposed to sharp price fluctuations. During certain periods of market stress, the anomaly-based model showed resilience, suggesting its potential for robust performance. These findings suggest that integrating anomaly detection with deep learning models may improve their capability to handle complex financial data patterns, and that anomaly detection effectively minimizes the impact of extreme values, allowing for more robust predictions.

This proposed approach enhances the stability and accuracy of prediction models using real-world financial market data. While this study focused on KOSPI 200 options, the underlying rationale for our method has strong potential for generalization. The data irregularities that our model addresses, such as price distortions in illiquid deep ITM or OTM options and noise from short-maturity contracts, are not unique to the Korean market but are fundamental features of option markets globally. Therefore, our anomaly-based framework is expected to be effective across different assets and market structures.

Future research will proceed in several key directions to validate this generalizability. A primary goal is to apply and evaluate the model on more liquid and diverse markets, such as the U.S. S&P 500 index options. Furthermore, we plan to test its performance on different asset types, such as single-stock options, and more complex contract structures, like American-style options, to assess the model's ability to price the early exercise feature. Ultimately, these extensions could pave the way for developing real-time analysis and prediction systems, thereby empowering financial institutions to respond more rapidly to market changes and strengthen risk management.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Code and data availability

All analysis code (training/evaluation scripts, bucketed analyses, rolling time-series splits, and EDA) is publicly available at this GitHub repository https://github.com/Pjihong/Reliable-option-pricing_anomaly-score. For long-term reproducibility, an immutable, citable snapshot of the exact version used in this paper is archived on Zenodo: 10.5281/zenodo.17018138. Please cite the software using this DOI.

The KRX raw data used in this study are subject to licensing and therefore cannot be redistributed. Readers who independently obtain the KRX data can run the code by supplying the variables described in Section 3.1.

Acknowledgments

This research was supported by the BK21 FOUR (Fostering Outstanding Universities for Research, NO.5120200913674) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea (NRF). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00562904). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00355646).

Conflict of interest

The authors declare there is no conflict of interest.

Author contributions

Jihong Park: Data curation, Software implementation, Writing–original draft; Jeonggyu Huh: Methodology, Conceptualization, Resources, Writing–review & editing; Jaegi Jeon: Investigation, Validation, Formal analysis, Theoretical analysis, Funding acquisition, Writing–review & editing. All authors have read and approved the final version of the manuscript.

References

1. F. Black, M. Scholes, The pricing of options and corporate liabilities, *J. Polit. Econ.*, **81** (1973), 637–654. https://doi.org/10.1086/260062
2. R. C. Merton, Theory of rational option pricing, *Bell J. Econ. Manage. Sci.*, **4** (1973), 141–183. https://doi.org/10.2307/3003143
3. R. Cont, Empirical properties of asset returns: Stylized facts and statistical issues, *Quant. Finance*, **1** (2001), 223–236. https://doi.org/10.1080/713665670
4. J. C. Hull, *Options, futures, and other derivatives*, 5th ed., Pearson Education, 2003.
5. M. Malliaris, L. A. Salchenberger, Neural network model for estimating option prices, *Appl. Intell.*, **3** (1993), 193–206. https://doi.org/10.1007/BF00871937
6. J. M. Hutchinson, A. W. Lo, T. A. Poggio, Nonparametric approach to pricing and hedging derivative securities via learning networks, *J. Finance*, **49** (1994), 851–889. https://doi.org/10.1111/j.1540-6261.1994.tb00081.x
7. A. Sharma, C. K. Verma, P. Singh, Enhancing option pricing accuracy in the Indian market: A CNN-BiLSTM approach, *Comput. Econ.*, **65** (2025), 3751–3778. https://doi.org/10.1007/s10614-024-10689-z
8. C. Quek, M. Pasquier, N. Kumar, A novel recurrent neural network-based prediction system for option trading and hedging, *Appl. Intell.*, **29** (2008), 138–151. https://doi.org/10.1007/s10489-007-0052-4
9. J. Ruf, W. Wang, Neural networks for option pricing and hedging: A literature review, *J. Comput. Finance*, **24** (2020), 1–46. https://doi.org/10.21314/JCF.2020.390
10. J. Huh, Pricing options with exponential Lévy neural network, *Expert Syst. Appl.*, **127** (2019), 128–140. https://doi.org/10.1016/j.eswa.2019.03.008
11. Y. Aït-Sahalia, J. Yu, High frequency market microstructure noise estimates and liquidity measures, *Ann. Appl. Stat.*, **3** (2009), 422–457. https://doi.org/10.1214/08-AOAS200
12. A. Madhavan, Market microstructure: A survey, *J. Financ. Mark.*, **3** (2000), 205–258. https://doi.org/10.1016/S1386-4181(00)00007-0
13. B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu, D. Cho, et al. Deep autoencoding Gaussian mixture model for unsupervised anomaly detection, In: *Proc. Int. Conf. Learn. Represent. (ICLR)*, (2018). Available from: https://openreview.net/forum?id=BJJLHbb0- .
14. F. T. Liu, K. M. Ting, Z. H. Zhou, Isolation forest, In: *2008 Eighth IEEE International Conference on Data Mining*, Pisa, Italy, (2008), 413–422. https://doi.org/10.1109/ICDM.2008.17
15. K. Golmohammadi, O. R. Zaiane, Time series contextual anomaly detection for detecting market manipulation in stock market, *Expert Syst. Appl.*, **42** (2015), 3635–3644. https://doi.org/10.1109/DSAA.2015.7344856
16. M. Ren, W. Zeng, B. Yang, R. Urtasun, Learning to reweight examples for robust deep learning, *arXiv preprint*, arXiv:1803.09050 (2018). Available from: https://arxiv.org/abs/1803.09050 .

17. B. Han, Q. Yao, X. Yu, G. Niu, M. Xu, W. H. Hu, et al. Co-teaching: Robust training of deep neural networks with extremely noisy labels, *Adv. Neural Inf. Process. Syst.*, **31** (2018), 8527–8537. Available from: https://proceedings.neurips.cc/paper/2018/hash/a19744e268754fb0148b017647355b7b-Abstract.html .
18. E. F. Fama, Efficient capital markets: A review of theory and empirical work, *J. Finance*, **25** (1970), 383–417. https://doi.org/10.2307/2325486
19. G. Bakshi, N. Kapadia, D. Madan, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Rev. Financ. Stud.*, **16** (2003), 101–143. https://doi.org/10.1093/rfs/16.1.101
20. V. Nair, G. E. Hinton, Rectified linear units improve restricted Boltzmann machines, In: *ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning*, (2010), 807–814. Available from: https://dl.acm.org/doi/10.5555/3104322.3104425 .
21. K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification, In: *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, (2015), 1026–1034. https://doi.org/10.1109/ICCV.2015.123
22. D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, *arXiv preprint*, arXiv:1412.6980 (2014). Available from: https://arxiv.org/abs/1412.6980 .

Appendix A. Supplementary comparison of anomaly detection methods

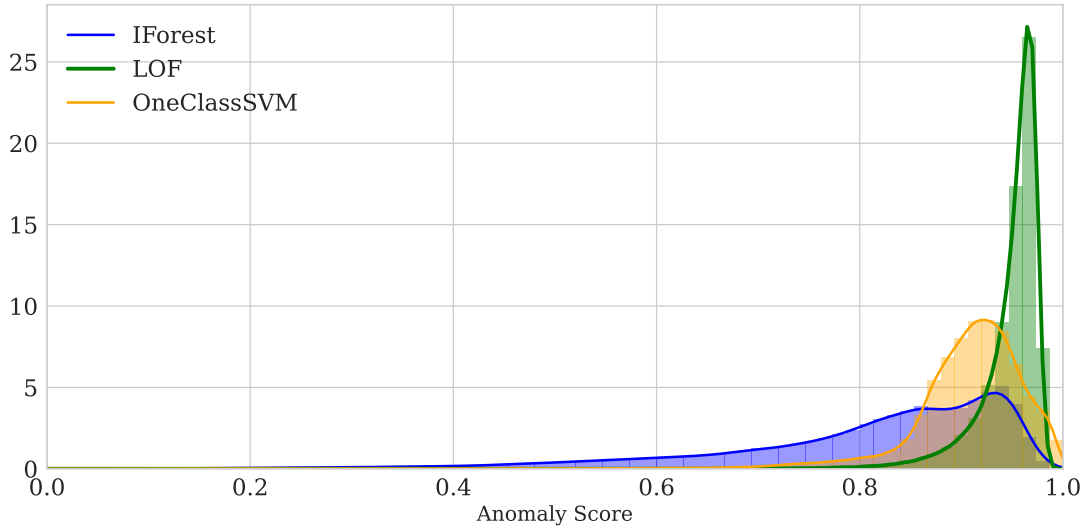


Figure A1. Training-fold densities of normalized anomaly scores $s \in [0, 1]$ for Isolation Forest, LOF, and OC-SVM.

We compare three detectors (Isolation Forest, LOF, OC-SVM) over 10 runs with distinct seeds and early stopping (mean ~350 epochs). Scores are min–max scaled to $[0, 1]$ in the main protocol; a z-score sensitivity check is also reported. The downstream MLP and data splits are identical across detectors, and the weighting hyperparameter was fixed to $k = 0.7$, which was optimized for the Isolation Forest-based model.

The training-fold densities in Figure A1 show that Isolation Forest is more diffuse with relatively little mass near $s = 1$, whereas LOF exhibits a pronounced spike around $s \approx 1$ and OC-SVM is strongly right-skewed. Under the exponential weighting $w = \exp(-ks)$, this translates into more balanced weights and therefore a larger effective sample size for Isolation Forest relative to LOF and OC-SVM.

Table A1. Validation performance under two scalers (10 runs) and anomaly-detector fit time (detector only; downstream MLP excluded). Lower is better for MSE.

	Isolation forest	LOF	OC-SVM
Min-Max scaling			
Best Val MSE (mean ± sd)	**4.742 ± 0.101**	4.762 ± 0.147	4.766 ± 0.202
Across-run spread (min-max Δ)	**0.343**	0.485	0.701
Standard (z-score) scaling			
Best Val MSE (mean ± sd)	**5.066 ± 0.157**	10.232 ± 0.127	5.159 ± 0.187
Across-run spread (min-max Δ)	0.448	4.159	**0.403**
Fit time (detector only)	**2.3 s**	12.6 s	640.7 s

Appendix B. Minimal error analysis

We demonstrate variance reduction under a simplified heteroskedastic mean model. This setting can be viewed as a local approximation to the general model by either considering a small neighbourhood in which $f_\theta(x)$ is approximately constant, or equivalently applying the analysis to residuals $P_i - f_{\theta^*}(x_i)$ around a reference θ^*. We also assume that $\{\varepsilon_i\}$ are independent.

Proposition B1 (Variance reduction with exponential weights)**.** *Consider the heteroskedastic mean model $P_i = \mu + \varepsilon_i$, $\mathbb{E}[\varepsilon_i] = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma_i^2 = \sigma_0^2 e^{k^* s_i}$, where $s_i \in [0,1]$ are fixed anomaly scores and $k^* \geq 0$. For any nonnegative weights w_i, the weighted least-squares estimator $\hat{\mu}(w) = \frac{\sum_i w_i P_i}{\sum_i w_i}$ satisfies*

$$\mathrm{Var}\left(\hat{\mu}(w)\right) = \frac{\sum_i w_i^2 \sigma_i^2}{\left(\sum_i w_i\right)^2}.$$

In particular, for the exponential weights $w_i = e^{-ks_i}$,

$$\mathrm{Var}\left(\hat{\mu}(k)\right) = \sigma_0^2 \frac{\sum_i e^{-(2k-k^*)s_i}}{\left(\sum_i e^{-ks_i}\right)^2}, \tag{B.1}$$

which is minimized at $k = k^*$ (the inverse-variance choice), yielding $\mathrm{Var}\left(\hat{\mu}(k^*)\right) = \sigma_0^2 / \sum_i e^{-k^* s_i}$.

Moreover, compared to the unweighted estimator ($k = 0$),

$$\mathrm{Var}\left(\hat{\mu}(k^*)\right) \;\leq\; \mathrm{Var}\left(\hat{\mu}(0)\right) = \sigma_0^2 \frac{\sum_i e^{k^* s_i}}{N^2},$$

with equality if and only if all s_i are equal.

Proof. The variance formula for $\hat{\mu}(w)$ is immediate from independence: $\mathrm{Var}\left(\sum_i w_i P_i\right) = \sum_i w_i^2 \sigma_i^2$. Substituting $w_i = e^{-ks_i}$ gives Eq B.1; minimizing in k clearly occurs at $k = k^*$ since $w_i \propto \sigma_i^{-2}$ at that value. For the comparison with the unweighted mean, write

$$\mathrm{Var}\left(\hat{\mu}(k^*)\right) = \frac{\sigma_0^2}{\sum_i e^{-k^* s_i}}, \qquad \mathrm{Var}\left(\hat{\mu}(0)\right) = \sigma_0^2 \frac{\sum_i e^{k^* s_i}}{N^2}.$$

By Cauchy–Schwarz with $a_i = e^{k^* s_i/2}$ and $b_i = e^{-k^* s_i/2}$,

$$N^2 = \Big(\sum_i a_i b_i\Big)^2 \le \Big(\sum_i a_i^2\Big)\Big(\sum_i b_i^2\Big) = \Big(\sum_i e^{k^* s_i}\Big)\Big(\sum_i e^{-k^* s_i}\Big),$$

which rearranges to $\frac{1}{\sum_i e^{-k^* s_i}} \le \frac{\sum_i e^{k^* s_i}}{N^2}$. Equality holds iff all s_i are identical.

Around θ^*, the first-order expansion $f_\theta(x_i) \approx f_{\theta^*}(x_i) + J_i(\theta - \theta^*)$ with $J_i = \nabla_\theta f_{\theta^*}(x_i)^\top$ turns the weighted objective into a local weighted least-squares problem. Consequently, giving smaller weights to high-variance observations reduces the covariance of the local estimator; Proposition B1 captures this effect in its simplest form. Intuitively, the exponential weights shrink the contribution of high-variance examples to both the objective and its gradients. Under the heteroskedastic model, this is equivalent to (approximate) inverse-variance weighting, which yields the MLE in the Gaussian case and reduces the variance of the estimator via the local weighted least-squares normal equations. Hence weighted training primarily lowers the variance component of the generalization error.

AIMS Press