Deep Risk Management in Finance

Final Report

Indian Institute of Technology Madras

Chennai, India

Submitted by

Saptarshi Pramanik

Dual Degree Programme

Department of Civil Engineering / Quantitative Finance

Project Guide

Prof. Neelesh Shankar Upadhye

1 Introduction

The problem of hedging financial derivatives is central to risk management in quantitative finance. Classical hedging approaches, most notably delta hedging derived from the Black–Scholes framework, aim to replicate the payoff of a contingent claim by dynamically trading in the underlying asset. Under idealized assumptions such as continuous trading, frictionless markets, and known constant volatility, these methods provide theoretically optimal hedging strategies.

In practice, however, financial markets deviate significantly from these assumptions. Trading occurs at discrete times, transaction costs are non-negligible, and asset dynamics often exhibit features such as stochastic volatility that render markets incomplete. In such settings, perfect replication is no longer feasible, and classical hedging strategies may perform suboptimally when evaluated from a risk management perspective.

The Deep Hedging framework, introduced by Bühler et al. (2019), addresses these limitations by formulating hedging as a risk minimization problem rather than a replication problem. Instead of seeking to eliminate hedging error pathwise, the objective is to directly optimize a risk measure defined on the terminal profit-and-loss (P&L) of the hedged portfolio. This perspective naturally accommodates market frictions, model uncertainty, and risk preferences.

A key feature of the Deep Hedging approach is the use of neural networks to represent dynamic trading strategies. By parameterizing the hedging policy as a function of observable market states and training it using simulated data, the framework leverages modern machine learning techniques to solve high-dimensional and nonlinear hedging problems that are analytically intractable.

The objective of this report is to develop and evaluate a working implementation of the Deep Hedging framework as described in the original paper. The focus is on reproducing key experimental results that demonstrate risk-adjusted performance under both complete and incomplete market settings, and on comparing the learned strategies with classical delta hedging benchmarks. This work forms the initial phase of a broader project aimed at studying deep learning–based approaches to financial risk management.

2 Problem Setup and Mathematical Formulation

We consider the problem of dynamically hedging a European-style contingent claim written on a single underlying asset over a finite trading horizon. The hedging problem is formulated in discrete time, reflecting the practical constraints of real-world trading.

2.1 Trading Horizon and Market Information

Let $T > 0$ denote the maturity of the derivative contract. Trading is allowed at a finite set of dates

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

where $N \in N$ denotes the number of rebalancing periods. All trading decisions at time t_k are based on the information available up to that time.

2.2 Underlying Asset Dynamics

Let S_{t_k} denote the price of the underlying asset at time t_k. The asset price evolution is modeled under a given stochastic process, and Monte Carlo simulation is used to generate sample paths for training and evaluation.

In this work, both complete and incomplete market settings are considered:

- the Black–Scholes model with constant volatility, representing a complete market, and
- the Heston stochastic volatility model, representing an incomplete market due to unhedgeable volatility risk.

The specific model dynamics are used solely for data generation and are not assumed to be known analytically by the hedging strategy.

2.3 Contingent Claim

Let $Z(S_T)$ denote the terminal payoff of the derivative at maturity. Throughout this work, we focus on a European call option with strike price K, whose payoff is given by

$$Z(S_T) = \max(S_T - K, 0).$$

The analysis is conducted from the perspective of the option seller, who receives the option premium at inception and seeks to manage the residual risk through dynamic trading in the underlying asset.

2.4 Hedging Strategy

A hedging strategy is defined as a predictable sequence

$$\pi = (\pi_{t_0}, \pi_{t_1}, \dots, \pi_{t_{N-1}}),$$

where π_{t_k} denotes the number of units of the underlying asset held over the interval $[t_k, t_{k+1})$.

The strategy is assumed to be self-financing, meaning that portfolio value changes arise solely from gains and losses due to trading in the underlying asset, subject to market frictions when transaction costs are present.

2.5 Trading Gains and Terminal P&L

The cumulative trading gains generated by a hedging strategy π are given by

$$G_T(\pi) = \sum_{k=0}^{N-1} \pi_{t_k} \left(S_{t_{k+1}} - S_{t_k} \right),$$

with appropriate modifications when proportional transaction costs are included.

The terminal profit-and-loss (P&L) of the hedged position is defined as

$$P\&L(\pi) = G_T(\pi) - Z(S_T).$$

The objective of the hedging problem is to construct a strategy π that controls the risk of the terminal P&L distribution under the chosen market model.

3 Risk Measures and Objective Function

Classical hedging approaches often evaluate performance using variance or mean-squared hedging error. While such criteria are analytically convenient, they do not explicitly encode risk preferences and may inadequately capture tail risk, particularly in incomplete or frictional markets.

The Deep Hedging framework instead formulates hedging as a risk minimization problem by optimizing a convex risk measure defined on the terminal profit-and-loss (P&L) of the hedged portfolio.

3.1 Convex Risk Measures

A convex risk measure $\rho(\cdot)$ assigns a real number to a random payoff and satisfies properties such as monotonicity, convexity, and translation invariance. Convexity reflects the principle that diversification should not increase risk, while translation invariance ensures that adding a deterministic amount of cash reduces risk by the same amount.

Convex risk measures provide a flexible and economically meaningful way to quantify risk beyond variance-based criteria, particularly in settings where payoff distributions are asymmetric or heavy-tailed.

3.2 Entropic Risk Measure

In this work, the entropic risk measure is employed, defined for a random variable X as

$$\rho_\lambda(X) = \frac{1}{\lambda} \log E\left[e^{-\lambda X}\right],$$

where $\lambda > 0$ denotes a risk aversion parameter.

The entropic risk measure penalizes adverse outcomes exponentially and is closely related to exponential utility maximization. Larger values of λ correspond to higher sensitivity to downside risk.

3.3 Risk-Based Hedging Objective

Let $P\&L(\pi)$ denote the terminal profit-and-loss associated with a hedging strategy π. The Deep Hedging objective is to find a strategy that minimizes the entropic risk of the terminal P&L:

$$\min_{\pi} \ \rho_\lambda(P\&L(\pi)).$$

Equivalently, this problem can be written as

$$\min_{\pi} \ E\left[e^{-\lambda\, P\&L(\pi)}\right],$$

since the logarithm and scaling by λ do not affect the minimizer.

This formulation directly links the hedging strategy to the chosen risk preference and avoids reliance on replication-based arguments that may fail in incomplete or frictional markets.

3.4 Certainty Equivalent

Associated with the entropic risk measure is the certainty equivalent (CE), defined as

$$CE(\pi) = -\frac{1}{\lambda} \log E\left[e^{-\lambda\, P\&L(\pi)}\right].$$

The certainty equivalent represents the guaranteed amount of terminal P&L that yields the same utility as the risky P&L distribution under exponential utility. It provides a scalar performance metric that is directly aligned with the optimization objective.

In this work, the certainty equivalent is used as a diagnostic and evaluation metric to compare hedging strategies trained under the same risk aversion parameter.

4 Deep Hedging Framework and Policy Representation

The Deep Hedging framework formulates dynamic hedging as a sequential decision-making problem, where trading decisions are made at discrete times based on observable market information. Rather than deriving an explicit analytical hedging rule, the strategy is learned directly by optimizing a risk-based objective through simulation.

4.1 Policy-Based Hedging

At each trading time t_k, the hedger selects a position in the underlying asset based on the current market state. This decision is modeled as a policy function

$$\pi_{t_k} = \pi_\theta(X_{t_k}),$$

where π_θ is a parametric function with parameters θ, and X_{t_k} denotes the observable state at time t_k.

In this work, the state is defined as

$$X_{t_k} = (S_{t_k}, t_k),$$

consisting of the current underlying asset price and the current time. The inclusion of time as an input allows the policy to adapt its behavior across the trading horizon while maintaining a time-consistent structure.

4.2 Neural Network Representation

The policy function π_θ is represented by a feedforward neural network that maps the current state to a hedge ratio. The same network is reused across all trading times, with time explicitly provided as an input variable.

This parameter sharing enforces temporal consistency and avoids the need to train separate models for each time step. Neural networks are employed due to their ability to approximate complex, nonlinear functions and to generalize across a wide range of market states.

4.3 Sequential Decision Structure

The cumulative trading gains and terminal P&L depend on the entire sequence of hedging decisions made along a price path. As a result, the effect of each action is only fully realized at maturity. This structure is analogous to policy-based reinforcement learning, where:

- the hedger acts as the agent,
- the financial market acts as the environment,
- the hedge ratio represents the action, and
- the terminal P&L plays the role of a delayed reward.

However, unlike classical reinforcement learning algorithms, the Deep Hedging framework does not rely on Bellman recursion or value function approximation. Instead, it directly optimizes the expected risk-adjusted objective over simulated trajectories using stochastic gradient-based methods.

4.4 Learning Objective

By combining the policy representation with the entropic risk objective defined in the previous section, the hedging problem is reduced to the optimization of neural network parameters θ:

$$\min_\theta \; E\left[e^{-\lambda\, P\&L(\pi_\theta)}\right].$$

Gradients of this objective are computed via automatic differentiation through the simulated trading process, enabling end-to-end optimization of the hedging strategy.

5 Implementation Details

This section summarizes the key implementation choices made in developing the Deep Hedging framework. The emphasis is on describing the overall structure and numerical considerations rather than presenting exhaustive algorithmic or code-level details.

5.1 Simulation of Asset Price Paths

Monte Carlo simulation is used to generate sample paths of the underlying asset price over the discrete trading grid. Two classes of market models are considered:

- **Black–Scholes model:** The underlying asset follows a geometric Brownian motion with constant volatility. This setting corresponds to a complete market and serves as a baseline for validating the implementation.
- **Heston model:** The underlying asset price evolves jointly with a stochastic variance process. The presence of unhedgeable volatility risk renders the market incomplete and provides a more realistic testbed for the hedging strategy.

All simulations are performed under the physical probability measure and are used solely for training and evaluation of hedging strategies.

5.2 Transaction Costs

To incorporate market frictions, proportional transaction costs are introduced in selected experiments. Trading costs are incurred whenever the hedge position is adjusted and are proportional to the absolute change in the hedge ratio and the underlying asset price.

The inclusion of transaction costs discourages excessive rebalancing and alters the optimal hedging behavior, making perfect replication infeasible even in otherwise simple market settings.

5.3 Policy Architecture

The hedging policy is parameterized by a feedforward neural network that takes the current state (S_{t_k}, t_k) as input and outputs a scalar hedge ratio. Input normalization is applied to improve numerical stability during training.

A single network is shared across all trading dates, with time explicitly included as an input variable. This design enforces time consistency while allowing the policy to adapt dynamically across the trading horizon.

5.4 Training Procedure

Model parameters are optimized using stochastic gradient descent with the Adam optimizer. At each training iteration, a mini-batch of simulated price paths is generated, and the corresponding trading gains and terminal P&L are computed by sequentially applying the policy along each path.

The expectation in the entropic risk objective is approximated using Monte Carlo averaging over the mini-batch. Gradients are computed via automatic differentiation through the simulated trading process, enabling end-to-end optimization of the hedging strategy.

5.5 Numerical Stability Considerations

Optimizing exponential risk objectives can lead to numerical instabilities due to the rapid growth of the exponential function for large positive or negative values of the terminal P&L. To address this issue, the training objective is reformulated in a numerically stable manner.

In particular, the terminal P&L is scaled by its empirical standard deviation within each mini-batch before being used in the entropic risk objective. This normalization stabilizes the magnitude of the loss without altering the relative ordering of outcomes within the batch.

Additionally, the entropic objective is implemented using a log-sum-exp formulation, which improves numerical robustness when aggregating exponential terms. Gradient clipping is applied during optimization to prevent excessively large parameter updates.

These stabilization techniques are employed solely to facilitate reliable training and do not modify the underlying definition of the hedging problem or the evaluation of terminal P&L distributions.

6 Experimental Design

The experimental evaluation is designed to assess the effectiveness of the Deep Hedging framework in controlling risk under progressively more realistic market conditions. Rather than exhaustively reproducing all numerical experiments from the original paper, the focus is on a carefully selected set of experiments that highlight the key conceptual advantages of risk-based hedging.

6.1 Benchmark Strategy

Classical delta hedging is used as the primary benchmark throughout the experiments. The delta is computed under the same model assumptions used for data generation, ensuring a fair comparison between learned and analytical strategies.

Delta hedging serves as a natural reference point, as it is widely used in practice and is theoretically optimal under idealized conditions such as complete markets and frictionless trading.

6.2 Market Scenarios

Experiments are conducted under three representative market scenarios:

- **Complete market:** The Black–Scholes model with constant volatility, used as a sanity check for the implementation.
- **Incomplete market:** The Heston stochastic volatility model, introducing unhedgeable sources of risk.
- **Market frictions:** The Heston model augmented with proportional transaction costs, capturing the effect of trading frictions.

This progression allows isolation of the impact of market incompleteness and transaction costs on hedging performance.

6.3 Risk Preference Specification

Hedging strategies are trained under a fixed entropic risk aversion parameter λ. To study the effect of risk preferences on hedging behavior, additional experiments are conducted for different values of λ, while keeping all other model parameters unchanged.

All comparisons between Deep Hedging and classical delta hedging are performed at the same value of λ, ensuring consistency of risk preferences across strategies.

6.4 Evaluation Metrics

Hedging performance is evaluated using the distribution of terminal profit-and-loss (P&L) obtained from an independent set of simulated price paths. In addition to visual inspection of P&L distributions, risk-adjusted performance is summarized using the certainty equivalent associated with the entropic risk measure.

The certainty equivalent provides a scalar performance metric that is directly aligned with the optimization objective and enables comparison of hedging strategies trained under a fixed risk aversion parameter.

7 Results and Discussion

This section presents experimental results obtained using the Deep Hedging framework under progressively more realistic market settings. The focus is on qualitative behavior of terminal profit-and-loss (P&L) distributions and risk-adjusted performance, rather than on exact numerical dominance over classical benchmarks. All results are evaluated out-of-sample on independently simulated price paths.

7.1 Heston Model without Transaction Costs

We first consider the Heston stochastic volatility model without transaction costs. In this incomplete market setting, volatility evolves as an unhedgeable risk factor, making perfect replication infeasible.

Figure 1 compares the terminal P&L distributions of Deep Hedging and classical Black–Scholes delta hedging under identical Heston price paths. Both strategies exhibit negatively skewed distributions with heavy left tails, reflecting volatility risk.

The Deep Hedging strategy produces a broader P&L distribution, with greater dispersion than delta hedging. This behavior reflects optimization of a risk-based objective rather than pathwise replication. The results demonstrate that Deep Hedging remains stable and well-defined in incomplete markets where classical hedging assumptions fail.

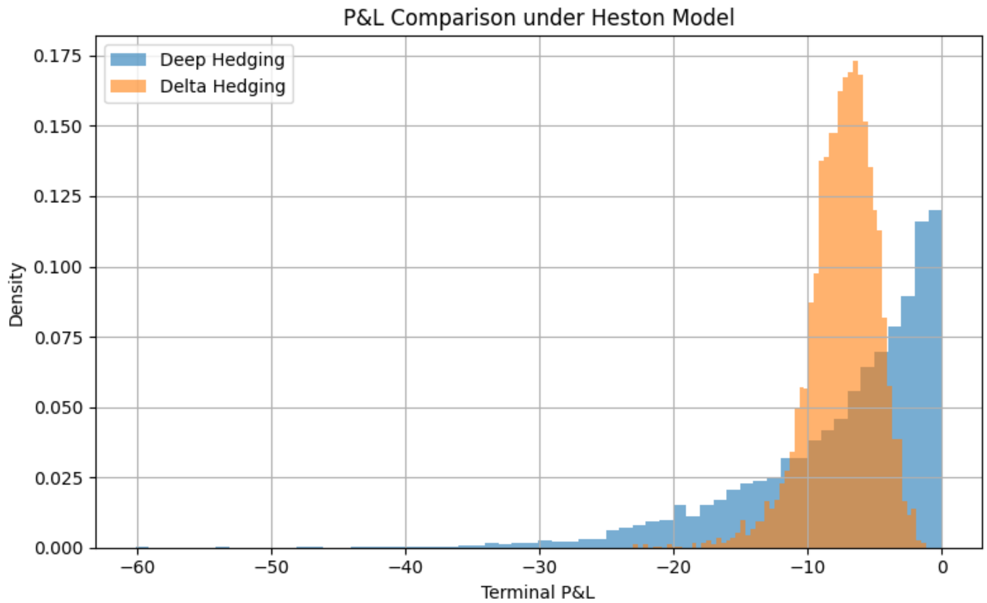

Figure 1: Terminal P&L distributions under the Heston model for Deep Hedging and classical delta hedging.

7.2 Heston Model with Transaction Costs

We next introduce proportional transaction costs into the Heston model, further increasing market frictions. Trading costs penalize frequent rebalancing and alter the optimal hedging behavior.

Figure 2 shows the terminal P&L distributions for Deep Hedging and delta hedging with transaction costs. Both strategies experience increased dispersion and downside risk compared to the frictionless case.

Deep Hedging adapts by learning smoother hedge adjustments that reduce excessive trading, while delta hedging continues to rebalance aggressively under frictionless assumptions. Although Deep Hedging does not uniformly reduce tail losses, it better balances hedging accuracy against trading costs.

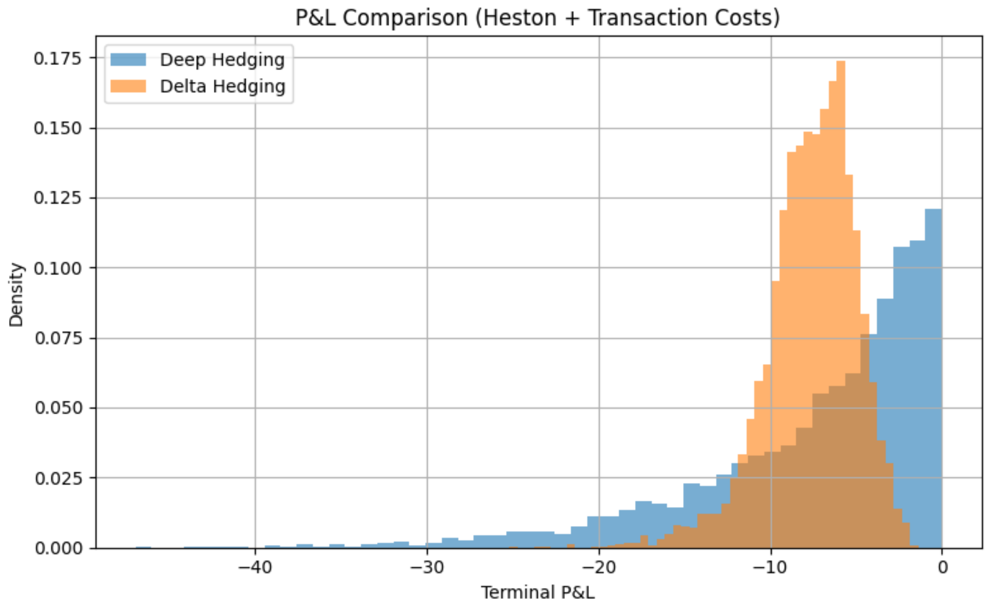

Figure 2: Terminal P&L distributions under the Heston model with transaction costs for Deep Hedging and classical delta hedging.

7.3 Effect of Risk Aversion

To examine the role of risk preferences, additional experiments are conducted for different values of the entropic risk aversion parameter λ, while keeping all market and model parameters fixed. Figure **??** illustrates the terminal P&L distributions obtained for different values of λ under the Heston model with transaction costs.

As λ increases, the learned hedging strategy places greater emphasis on penalizing adverse outcomes in the entropic risk objective. This leads to observable changes in the shape of the P&L distribution, reflecting a shift in the trade-off between hedging aggressiveness and exposure to unfavorable price movements.

Rather than uniformly reducing tail losses, higher risk aversion results in more conservative trading behavior, which can increase the frequency of moderate losses while limiting excessive rebalancing under transaction costs. These results highlight that the effect of risk aversion is mediated by market frictions and the chosen risk measure, and should be interpreted in the context of the underlying optimization objective.

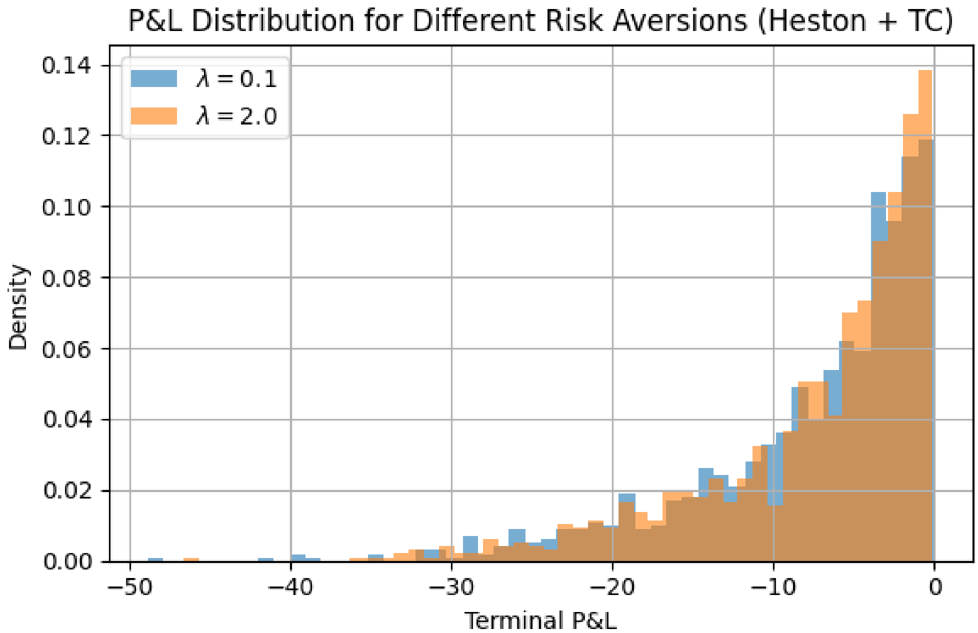

Figure 3: Terminal P&L distributions for different values of the risk aversion parameter λ under the Heston model with transaction costs.

7.4 Risk-Adjusted Performance

In addition to visual inspection of terminal P&L distributions, hedging strategies are evaluated using the certainty equivalent (CE) associated with the entropic risk measure. The certainty equivalent provides a scalar summary of risk-adjusted performance that is directly aligned with the optimization objective.

All comparisons of certainty equivalents are performed at a fixed value of the risk aversion parameter λ, ensuring consistency of risk preferences across strategies. Under this criterion, Deep Hedging achieves certainty equivalent values that are comparable to those obtained using classical delta hedging in incomplete and frictional market settings.

These results emphasize the importance of evaluating hedging strategies using risk-adjusted metrics that are consistent with the underlying optimization objective, rather than relying solely on variance-based performance measures.

7.5 Interpretation of Tail Risk

The observed P&L distributions highlight the distinction between minimizing a convex risk measure and explicitly minimizing tail risk. The entropic risk measure penalizes losses exponentially but does not directly target extreme quantiles such as Value-at-Risk or Conditional Value-at-Risk.

As a result, Deep Hedging does not necessarily produce thinner left tails than delta hedging, particularly in incomplete or frictional markets. Instead, it learns strategies that reflect the chosen risk preference embedded in the optimization objective.

These results emphasize the importance of aligning the selected risk measure with the desired notion of risk control.

8 Limitations and Future Work

The results presented in this report correspond to an initial implementation of the Deep Hedging framework and should be interpreted in light of several limitations.

8.1 Model Scope and Market Assumptions

The current implementation is restricted to single-asset European-style options. While this setting is sufficient to validate the core Deep Hedging methodology, it does not capture the full range of complexities encountered in practical risk management, such as multi-asset portfolios, path-dependent payoffs, or early exercise features. Extending the framework to these settings constitutes a natural direction for future work.

In addition, the experimental analysis focuses primarily on the Black–Scholes and Heston models. Although these models capture key features of financial markets, real-world asset dynamics may involve jumps, regime shifts, or other forms of non-stationarity that are not considered here.

8.2 Choice of Risk Measure

All experiments in this work are conducted using the entropic (exponential) risk measure, following the primary setup in the Deep Hedging paper. While this choice provides a smooth and convex objective that is well suited for gradient-based optimization, it does not explicitly target extreme tail outcomes.

Future work could investigate alternative convex risk measures, such as Conditional Value-at-Risk (CVaR), to study how the choice of risk criterion influences the structure of learned hedging strategies and their tail-risk characteristics.

8.3 Optimization and Training Dynamics

The training procedure relies on stochastic gradient descent with Monte Carlo estimates of the objective function. In the presence of transaction costs and stochastic volatility, the optimization landscape can become noisy, leading to slow convergence or sensitivity to hyperparameters.

More advanced optimization techniques, variance reduction methods, or alternative policy parameterizations may improve training stability and efficiency in these settings.

8.4 Out-of-Sample Evaluation

The experimental evaluation in this report is primarily qualitative and illustrative. A more systematic assessment involving multiple random seeds, extensive parameter sweeps, and formal statistical testing would provide stronger empirical validation of the observed results.

Overall, these limitations do not detract from the validity of the current implementation but rather highlight opportunities for extending and refining the Deep Hedging framework in future stages of the project.

9 Conclusion

This report presented an implementation and empirical study of the Deep Hedging framework for financial risk management. By formulating dynamic hedging as a risk-based optimization problem and leveraging neural network policies, the framework provides a flexible alternative to classical replication-based approaches, particularly in settings where perfect hedging is infeasible.

The implementation closely followed the original Deep Hedging formulation, employing an entropic risk objective and evaluating hedging performance under both complete and incomplete market models. Experiments conducted under the Black–Scholes and Heston stochastic volatility models, with and without transaction costs, demonstrated that the framework can be trained stably and applied across a range of market conditions without structural modification.

The empirical results highlight that Deep Hedging does not universally outperform classical delta hedging across all regimes, especially for simple payoff structures and relatively mild forms of market incompleteness. Instead, the behavior of the learned strategies reflects the choice of risk measure and the presence of market frictions. In particular, optimization under the entropic risk criterion emphasizes downside-sensitive risk control rather than explicit minimization of extreme tail losses, underscoring the importance of aligning optimization objectives with the desired notion of risk.

This phase of the project focused on understanding and implementing the core Deep Hedging framework as described in the original paper. Subsequent phases may explore alternative risk measures, richer payoff structures, and more comprehensive empirical evaluations, thereby extending the applicability of deep learning–based approaches to realistic financial risk management problems.

References

[1] H. Bühler, L. Gonon, J. Teichmann, and B. Wood, *Deep Hedging*, Quantitative Finance, Vol. 19, No. 8, pp. 1271–1291, 2019.

[2] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, Vol. 81, No. 3, pp. 637–654, 1973.

[3] S. L. Heston, *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, Review of Financial Studies, Vol. 6, No. 2, pp. 327–343, 1993.