Deep Risk Management in Finance

Initial Implementation and First 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 Problem Setup

We consider the problem of dynamically hedging a European-style contingent claim written on a single underlying asset. Trading takes place over a finite time horizon $[0, T]$, which is discretized into N trading dates

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

1.1 Underlying Asset Dynamics

Let S_t denote the price of the underlying asset at time t. The asset price is assumed to evolve according to a specified stochastic process. In this initial phase of the project, both complete and incomplete market settings are considered, including:

- the Black–Scholes model with constant volatility, and
- the Heston stochastic volatility model.

The precise dynamics used for simulation are specified in the implementation section. The market is assumed to be frictionless unless otherwise stated; extensions incorporating proportional transaction costs are introduced later.

1.2 Hedging Strategy

A hedging strategy is defined as a predictable sequence of positions

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

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 changes in portfolio value arise solely from gains and losses due to trading in the underlying asset.

1.3 Trading Gains and Terminal P&L

The cumulative trading gains generated by a hedging strategy π over the trading horizon 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 transaction costs are included.

Let $Z(S_T)$ denote the terminal payoff of the contingent claim at maturity. For a European call option with strike K, the payoff is

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

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

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

This formulation corresponds to the perspective of an option seller who receives the option premium at inception and seeks to manage the residual risk through dynamic trading in the underlying asset.

2 Deep Hedging Framework

Classical hedging approaches typically aim to replicate the payoff of a contingent claim by dynamically adjusting positions in the underlying asset. In contrast, the Deep Hedging framework formulates hedging as a sequential decision-making problem, where trading strategies are optimized directly with respect to a chosen risk criterion.

In this initial phase of the project, the objective is to study and implement the core components of the Deep Hedging methodology, focusing on policy representation and risk-based optimization, rather than on exhaustive architectural or algorithmic refinements.

2.1 Policy Representation

At each trading time t_k, the hedging decision is modeled as a function of the current observable state. The state is defined as

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

where S_{t_k} denotes the underlying asset price and t_k denotes the current time.

The hedging strategy is represented by a parametric policy

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

where π_θ is a neural-network-based function with parameters θ. The same policy function is applied at all trading times, with time explicitly included as an input variable. This formulation enforces time consistency while allowing the strategy to adapt dynamically to the current state.

2.2 Risk-Based Objective

Rather than minimizing replication error or variance, the Deep Hedging framework optimizes a risk-adjusted objective defined on the terminal P&L. In this work, a convex risk measure of the entropic (exponential) form is employed:

$$\min_\theta \; E\left[\exp\left(-\lambda \, P\&L(\pi_\theta)\right)\right],$$

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

This objective places asymmetric penalties on adverse outcomes and provides a flexible mechanism for incorporating risk preferences directly into the learning problem. The focus of this initial implementation is on validating the use of such a risk-based objective within the Deep Hedging framework.

3 Implementation Details

This section summarizes the key implementation choices made in developing the initial Deep Hedging framework. The emphasis is on describing the overall structure and design decisions rather than presenting exhaustive algorithmic or code-level details. The complete implementation is provided separately as supplementary material.

3.1 Simulation of Asset Price Paths

Monte Carlo simulation is used to generate sample paths of the underlying asset price over the discrete trading grid. The framework supports multiple market models, including the Black–Scholes model with constant volatility and the Heston stochastic volatility model. These models allow evaluation of the hedging strategy under both complete and incomplete market settings.

For each training iteration, a batch of simulated price paths is generated and used to compute the corresponding trading gains and terminal P&L. This enables estimation of expectations appearing in the risk-based objective.

3.2 Policy Parameterization

The hedging policy is parameterized by a feedforward neural network that maps the current state, consisting of the underlying price and time, to a hedge ratio. The same network is applied across all trading dates, ensuring time-consistent decision-making. Input normalization is applied to improve numerical stability during training.

The choice of a feedforward architecture reflects the Markovian structure of the problem and avoids unnecessary model complexity in this initial implementation.

3.3 Training Procedure

Model parameters are optimized using stochastic gradient descent, with gradients computed via automatic differentiation through the simulated trading process. The expectation in the objective function is approximated using mini-batches of simulated price paths.

To ensure numerical stability when optimizing the exponential risk objective, appropriate scaling and clipping of the terminal P&L are applied during training. These transformations are used solely for optimization purposes and do not alter the underlying definition of the hedging problem.

3.4 Evaluation and Baseline Comparison

After training, the learned hedging policy is evaluated on independently simulated price paths. Performance is assessed using distributions of terminal P&L and risk-adjusted summary statistics.

For benchmarking purposes, the Deep Hedging strategy is compared against classical delta hedging under identical market conditions. This comparison serves as a reference point for interpreting the behavior of the learned strategy, rather than as a claim of universal dominance.

The implementation is designed to be modular and extensible, allowing additional market models, risk measures, and constraints to be incorporated in subsequent phases of the project.

4 Experimental Results

This section presents initial experimental results obtained using the Deep Hedging framework described earlier. The purpose of these experiments is to validate the correctness of the implementation and to examine the qualitative behavior of the learned hedging strategies under different market settings. The results should be interpreted as exploratory rather than definitive.

4.1 Experimental Setup

All experiments are conducted in a discrete-time setting with a fixed number of rebalancing dates over the trading horizon. For each experiment, the hedging policy is trained using Monte Carlo simulation of asset price paths and subsequently evaluated on an independent set of simulated paths.

Results are reported for multiple market configurations, including complete and incomplete market models. Performance is assessed using the distribution of terminal profit-and-loss (P&L) as well as selected risk-adjusted summary statistics.

4.2 Black–Scholes Model

As a baseline experiment, the Deep Hedging framework is first evaluated under the Black–Scholes model with constant volatility. This setting corresponds to a complete market, where classical delta hedging is known to perform well under idealized assumptions.

The learned hedging strategy produces stable P&L distributions, and the training procedure converges reliably. Comparisons with classical delta hedging indicate that both approaches exhibit similar qualitative behavior in this regime, as expected given the favorable market assumptions.

This experiment serves primarily as a validation check for the implementation rather than as a setting in which substantial performance differences are anticipated.

4.3 Heston Stochastic Volatility Model

To examine the framework under market incompleteness arising from stochastic volatility, experiments are conducted using the Heston model. In this setting, volatility evolves as an additional stochastic factor that cannot be perfectly hedged using the underlying asset alone.

Under the Heston dynamics, the Deep Hedging framework continues to train stably and produces well-defined hedging strategies. The resulting P&L distributions exhibit greater dispersion compared to the Black–Scholes case, reflecting the increased uncertainty introduced by stochastic volatility.

Classical delta hedging is used as a benchmark under the same market conditions. While delta hedging remains competitive in this regime, the experiment demonstrates that the Deep Hedging framework can be applied without modification in incomplete market settings.

4.4 Transaction Costs

Further experiments incorporate proportional transaction costs into the trading process, introducing additional market frictions. In this setting, frequent rebalancing incurs explicit costs, and perfect replication becomes infeasible even in otherwise simple market models.

The inclusion of transaction costs leads the learned hedging policy to adopt smoother trading behavior, with reduced turnover relative to frictionless settings. Terminal P&L distributions reflect the trade-off between hedging accuracy and transaction cost minimization.

Comparisons with classical delta hedging under transaction costs indicate that delta hedging may still perform well when transaction costs are low and rebalancing remains frequent. These results highlight that the benefits of learned hedging strategies depend on the severity and nature of market frictions.

4.5 Risk-Adjusted Metrics

In addition to summary statistics such as mean and standard deviation of terminal P&L, risk-adjusted performance is evaluated using the certainty equivalent corresponding to the entropic risk measure employed during training. This metric provides a scalar summary aligned with the optimization objective.

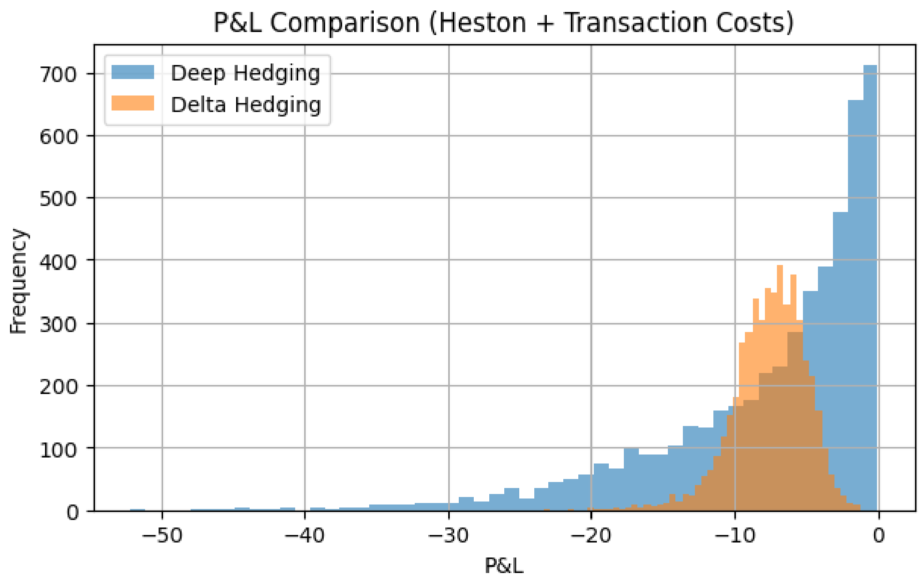

Figure 1: Terminal P&L distributions under the Heston model with transaction costs for Deep Hedging and classical delta hedging.

Across the considered experimental settings, the certainty equivalent reflects the interaction between hedging accuracy, transaction costs, and risk preferences. Differences between Deep Hedging and classical hedging strategies are observed to depend sensitively on the chosen market configuration and parameter values.

These observations underscore the importance of aligning evaluation metrics with the underlying optimization objective.

5 Current Limitations and Next Steps

The results presented in this report correspond to an initial implementation of the Deep Hedging framework and are subject to several limitations. First, the experiments are restricted to single-asset European-style derivatives, and the analysis does not yet consider multi-asset portfolios or path-dependent payoffs.

Second, the choice of risk measure is limited to the entropic risk measure. While this choice is theoretically well motivated and computationally convenient, alternative risk measures such as Conditional Value-at-Risk (CVaR) may offer different perspectives on tail risk and warrant further investigation.

Third, the experimental evaluation is primarily qualitative and exploratory. A more systematic analysis involving extensive parameter sweeps, stress testing, and statistical validation is deferred to subsequent phases of the project.

Future work will focus on extending the framework along these dimensions, refining the empirical evaluation, and improving the modularity and robustness of the implementation. These extensions will form the basis of the final implementation and analysis.

References

[1] H. Bühler, L. Gonon, J. Teichmann, and B. Wood, *Deep Hedging*, Quantitative Finance, 19(8), 1271–1291, 2019.

[2] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, 81(3), 637–654, 1973.

[3] S. L. Heston, *A Closed-Form Solution for Options with Stochastic Volatility*, Review of Financial Studies, 6(2), 327–343, 1993.