International Academic Institute for Science and Technology

International Academic Journal of Economics
Vol. 13, No. 2, 2026, pp. 28-36.

ISSN 2454-2474

International Academic Journal of Economics

www.iaiest.com

Machine Learning-Based Analysis of Stock Market Volatility and Investment Risk

Arjun Menon [a], Kavya Nair [b]

[a] *Bharat Institute for Future Technology Research, Bengaluru, Karnataka, India.*
[b] *Bharat Institute for Future Technology Research, Bengaluru, Karnataka, India.*

Abstract

Equity volatility clusters have a tendency to go from calm volatility to turbulent to crisis, and static risk models do not fare well with fat-tailed shocks. In this paper, a deep-learning architecture that predicts volatility and categorizes the market regime in one network is studied, and its results are used to facilitate investment risk control. The network is fed a 22-day window of return-based features, which are passed through an attention layer, a bidirectional LSTM, and a convolutional encoder, then split into a 5-day volatility head and a three-state regime head. Its outputs are used to control an allocation rule, which sets the size of exposure to a volatility objective, and then cuts it back when a crisis is likely to occur. There are 5 independent markets that are simulated and known to the evaluator. When compared with the baseline models, EWMA, GARCH, HAR, and random forest, the model reduces volatility forecast error by 5.1% to 16.6%, identifies the crisis state with an F1 score of 82.9%, and decreases the average maximum drawdown of the buy-and-hold strategy by 50.2% to 16.7%, while increasing the Sharpe ratio from 0.03 to 0.66. The difference between a random forest and attention is negligible and lies within one standard deviation. The evidence is represented in a simulated fashion, and the scope of the evidence is explicitly described.

Keywords: Stock Market Volatility, Investment Risk, Deep Learning, Regime Switching, Volatility Targeting, Drawdown Control.

DOI: 10.71086/IAJE/V13I2/IAJE1310
Received: 12 March 2026/Revised: 22 April 2026/Accepted: 14 May 2026/Published: 30 June 2026

Introduction

Investors do not even lose money in a uniform manner. The vast majority of the damage to a portfolio comes in short, violent spurts when prices drop, and the volatility of the portfolio increases by factors of two or three, and correlation between asset classes increases. Investor panic, forced selling, and thin liquidity have a vicious cycle: a small decline can escalate into a significant drawdown. The issue of downside risk to the portfolios of investors, and the link between volatility and investor panic, is therefore the primary practical risk management issue: the risk manager has to know in advance whether the market is entering a high-risk state, and the size of the impending swings.

In fact, machine learning is the natural approach to this problem as volatility relies on a multitude of interacting factors—and there is no simple functional form. Recent work on risk-based asset allocation shows that machine-learning estimates of risk can be placed directly inside the portfolio-construction step (Agal et al., 2025). Research on quantitative approaches to American financial stocks has shown that machine-learning volatility forecasts can be formulated into trading rules and tested using backtests (Yan & Li, 2024). Fundamental feature engineering demonstrates that return and risk can be forecasted jointly: classification of the class of a given outcome and regression of its size (Rezaei et al., 2025). The results of decision-tree and data-mining studies indicate that, when a mixture of market and accounting variables is used, investment risk can be measured without a parametric distribution (Wang, 2025). Research on volatility factors, forecasts, and simulations is an application of machine learning that also utilizes simulation to explore volatility in response to its drivers (Mansilla-Lopez et al., 2025). Based on multidimensional financial risk indicators, hybrid approaches combining a GARCH-type variance process with deep-learning approaches can be used to enhance volatility prediction in comparison with a GARCH family (Ibrahim et al., 2026).

These studies justify the inclusion of machine-learning volatility forecasts in quantitative investment strategies, and highlight three gaps. Most work stops after predicting a volatility number, but the ultimate choice of the manager is often a binary one: invest further, cut back on exposure, or go to cash. Second, the accuracy of the forecast is seldom compared to investor focus – namely, the drawdown, risk-adjusted return, and capital retained on the worst days. Third, a model that is tested on one historical sample is difficult to assess since it has a limited number of crises.

This paper brings four contributions to the literature. It suggests learning a volatility forecast and a regime probability from one representation and proposes a compact multi-task network. It makes both outputs a clear allocation rule. It tests the system on artificial markets with known regimes of the market, allowing for direct measurement of precision and recall in the crisis state. It reports ablations, run-to-run dispersion, and the cases where the proposed model does not win. Related Work is summarized in Section II, the architecture is described in Section III, the experimental setup is described in Section IV, and the Results are presented in Section V and discussed and concluded in Section VI.

Literature Review

Financial returns have some well-known properties that break the assumptions of linear models: volatility clustering, heavy-tailed returns, and evolving correlation between past and future returns when the market changes state. The related work can be subdivided into three strands, relevant to the proposed design.

The first strand is to predict market variables with neural networks, and the second is to use the prediction in an investment decision. Instead of just predicting the stock price, Huang and Vakharia connect the prediction stage with an investment model to create a deep-learning stock prediction and investment model for financial management (Huang & Vakharia, 2024). In this case, the coupling is between prediction and decision, with a volatility and regime target taking the place of a return-direction target. With sentiment and technical signals from ESG, Lee, Kim and Jung augment a deep-learning predictor, demonstrating that it is possible to broaden the input space of a neural forecaster and that sequence models can ingest diverse signals (Lee et al., 2024). The lesson is that the representation learned from the window of inputs is as important as the output layer; the encoder here learns the local shock patterns first, and longer memory second, and maintains a small feature set.

The second strand takes a prediction to a control approach. In Yang, Wang and Hu introduce RiskawareTrader, a reinforcement-learning portfolio method that works for investors who are risk-averse and whose goal is to minimize the size of the losses while at the same time maximizing the average gain (Yang et al., 2025). In a multi-agent trading system, Cheng and his colleagues apply quantum finance and fuzzy reinforcement learning, where represent fuzzy membership functions for the imprecise states of the market rather than hard thresholds (Cheng et al., 2024). This is a similar approach to the soft regime probabilities used here because do not provide a hard 'on/off' switch. RL systems can learn the rules of allocation directly, but require a lot of data and are difficult to audit. The present paper takes a more conservative approach: a supervised network generates two interpretable signals, and a transparent rule translates the signals into exposure.

The third is that the same reasoning extends to all listed equity stocks. In Ni, Xie, Härdle and Zuo develop a machine-learning regulatory risk index for cryptocurrencies that allows them to translate a vague source of uncertainty into a risk series. In China, Huang, Zhang, Zheng and Tian (2020) use machine learning to understand the economic-efficiency differences and transition drivers between high-carbon and low-carbon industries (Huang et al., 2026). The topic is industrial transition and how it separates structurally different groups and identifies what is driving the difference – the same kind of problem as separating calm and crisis states. All these studies demonstrate that neural architectures support investment models and risk-averse decision rules to limit the size of losses, and that risk measurement gains from explicit attention to structural differences (Ni et al., 2025). What's needed is a smaller design that not only forecasts the degree of volatility, but also the state of the market, the conversion of that volatility to exposure by some clear rule, and testing the design against the true state of the market.

Deep Learning and Regime Classification Architecture

Fig. 1 below shows the process of data analysis using machine learning techniques in order to examine the patterns of volatility and risks in investments. Insights from this analysis will help in evaluating risk-return trade-offs, asset allocation, hedging and portfolio repositioning, thereby enhancing volatility understanding and risk measurement.

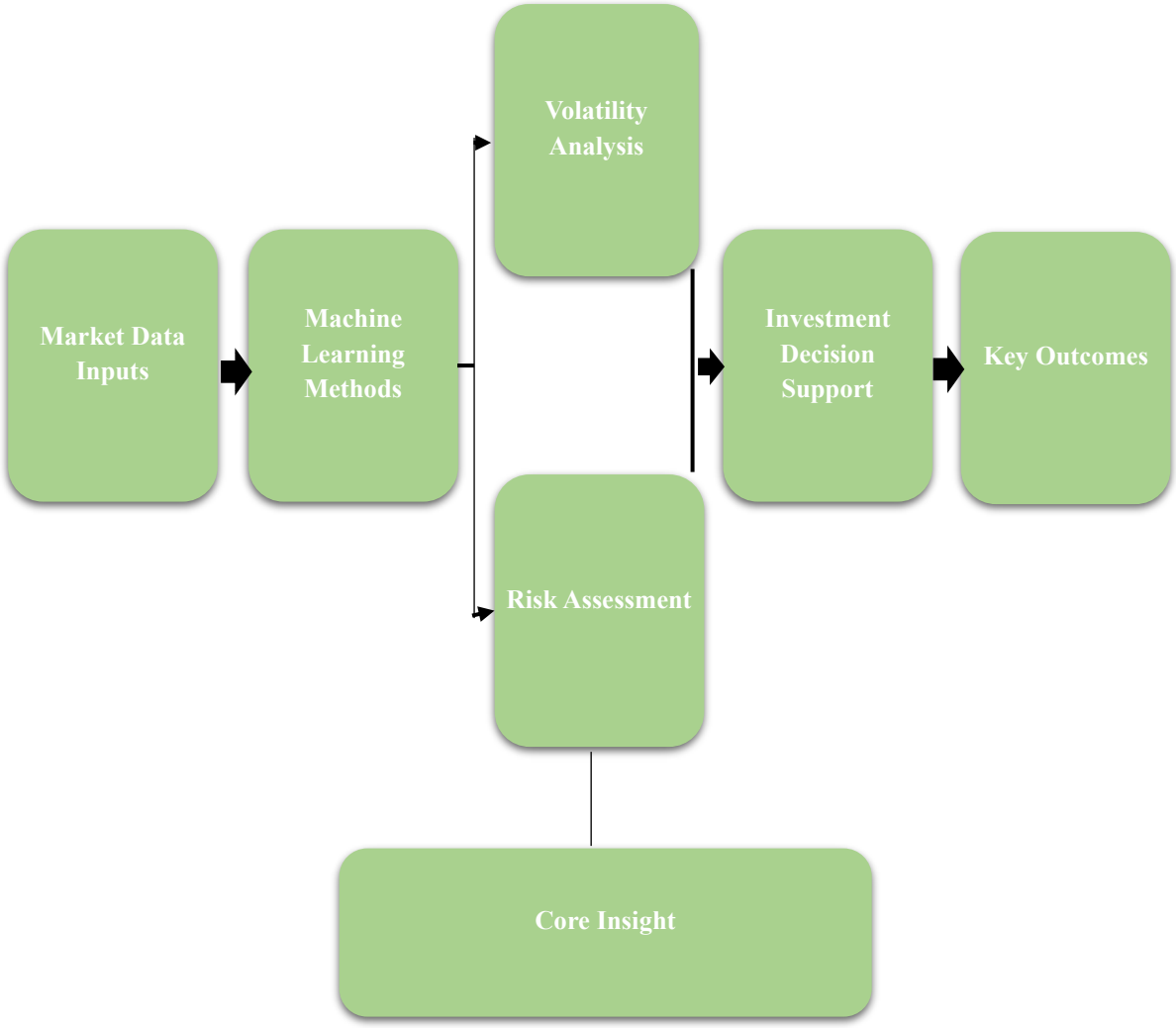


Fig. 1: Conceptual Framework for Machine Learning-Based Stock Market Volatility and Investment Risk Analysis

The investor watches the past 22 trading days of the features at the end of each day, and selects the next day's portfolio weight for a risky asset. The two targets are for the same window: the log of the realized volatility for the next 5 days, and the market regime for the next day, which is calm, turbulent, or crisis. It is more interesting to predict the next state rather than recognize the current state since a model that can only classify a regime once it is in sight will be useless for allocation purposes.

The features used to describe each day in the window are the return, the absolute return, a window-based proxy for the intraday return, the downside component of the return, and the log of the window-based realized volatility in the previous five days and the realized volatility in the last 22 days. Features are standardized based on the training sample only, and not the validation or test periods.

There are four components to the network. The two layers of the convolutional encoder (3x32 channels) identify short local patterns, like a group of high absolute returns, that prelude a volatility burst. A bidirectional LSTM using 32 units per direction is then used to build a memory of the window: reading the window twice, once in each direction, does not leak the future, as the entire window is already observed at decision time, and aids in differentiating a one-off spike from the beginning of a persistent regime. The 22 days are then read through an additive attention layer, which assigns weights to each day (which can be graphically displayed, a common method used in the machine-learning risk literature (Huang et al., 2026). A shared dense layer with dropout feeds two heads: one head is the volatility forecast, and the other head is a set of three heads, one head for each regime. The entire network is composed of 27,140 parameters that can be trained, and it does so in a few seconds on a CPU.

The heads are trained jointly. The Huber loss is used in the volatility head, as it is less affected by large values corresponding to crisis periods than is squared error. The head of the regime implements a cross-entropy loss that gives more weight to the rare crisis state. The training is done with the Adam optimizer and early stopping on a validation set. The key architectural concept is that a state that a network can identify from a window should also assist it in sizing the volatility that follows.

The outputs are incorporated into a portfolio weight in two ways. The first is volatility targeting: exposure is allowed to reach the ratio of a 12% annual volatility target to the forecast volatility, but is limited to the maximum investment. The second step is a crisis throttle: Exposure x (1-Prediction of crisis), which means that if predict a crisis 1/2, it doubles the exposure, and if predict a crisis with a probability of 1%, then don't change the exposure. The rest is kept in cash. Calculations of returns are based on proportional transaction costs of 5 basis points. There are no tuned parameters other than the volatility target and the cost, both set once and then fixed, so there is not much opportunity to overfit the backtest. A risk committee can ask a series of questions in relation to each term, since each weight can be traced back to two named signals.

Experimental Setup

In real price data, the true market regime is never observed. What one can say about the past is only based on a rule, and each such label is a modeling decision that contributes to determining the answer. To test the ability of the model to detect a high-risk state, the state has to be known; hence, this study tests simulated markets where the regime is explicitly given, and the model characteristics of equity returns are applied: persistence of market volatility, fat tails, and reduced returns in high-volatility states. The study is then a controlled test of two things: the architecture and the allocation logic, and not a prediction of returns in any specific market.

The dynamics of regimes evolve as a Markov chain. The calm regime has 10% annual volatility and a positive drift, lasts about 67 days on average, and occupies about 48% of the time. The turbulent regime has 18% volatility and a small positive drift, lasts about 29 days, and occupies 36% of the time. This crisis regime has 38% volatility, a drift of minus 25% per year, a duration of approximately 25 days, and an occupancy of 16%. Volatility is amplified in a regime by a persistent random factor; hence, the volatility levels in neighboring regimes overlap, and the regime cannot be derived from a single number, and the returns are characterized by Student-t shocks with five degrees of freedom, leading to fat tails. The "drift" of the unmanaged asset is only approximately 1.9% per year in the long term. The regime-dependent drift leads to the leverage effect, which works in favor of any rule that reduces exposure in high-volatility regimes, and this is the reason why the results of the portfolio should not be interpreted as proof that the effect is the same in real markets.

There are five independent markets, with 5,000 decision days each, 3,000 of which are used for training, 600 for validation, and 1,400, or approximately 5.6 years, are left for the untouched test period. The average percentage of test days that are crisis days is 17.4%, ranging from 15.1% to 23.9% in different markets. Statistics are all based on the test period and values are presented as mean and standard deviation for the five markets.

Following the tree-based risk literature (Rezaei et al., 2025) and (Wang, 2025), the forecasting baselines include an exponentially weighted moving-average variance model, a GARCH (1,1) model with estimated parameters on the training period and a heterogeneous-autoregressive regression as well as a random forest on seven engineered risk features. A random forest, logistic regression and volatility-threshold rule are used for classification. The ablations are removing the attention layer, the convolutional encoder and the regime head. Only training/validation data is used for all fitting, scaling and tuning, and fixed seeds are used for all markets for reproducing all numbers.

Volatility forecasts are judged using root mean squared error and mean absolute error in terms of percentage points of annualised volatility, as well as by QLIKE, which is a loss function that punishes under-prediction of volatility more than over-prediction, the more expensive of the two strategies for a risk manager. Precision, recall and F1 score are used to measure the crisis state. Portfolio rules are scored after costs by annualised return, volatility, the Sharpe ratio, maximum drawdown, the Calmar ratio and turnover.

Synthetic Market Dynamics and Stress Calibration

In order to obtain the empirical rigor of the controlled environment, the statistical topology of the synthetic data generators should be similar to that of the higher-order moments observed in equity indices with a long time horizon. The Markov-switching mechanism that determines the 5,000 days of decision-making was designed to have three statistical properties typical of actual market crashes: conditional volatility persistence, endogenously changing skewness, and non-linear correlation structure when markets are down. The state-dependent memory is realistic with the transition probability matrix across the three different states (calm, turbulent and crisis). When a crisis state has been reached the typical 25 business-day time horizon results in a severe tail clustering effect that means that model architectures must execute over long series of non-stationary inputs. Within each state, the return generation mechanism is modelled as a student -t with five degrees of freedom, adding heavy-tailed shocks that lead to daily returns with relatively large moments and that occur much more often than would be predicted by a Gaussian assumption. Moreover, the synthetic process explicitly reflects the asymmetric volatility phenomenon, which is also referred to as the "leverage effect. Programmed negative return shocks are designed to produce a greater subsequent shock to volatility when compared to an equal size positive shock. This structural feature will allow de-risking mechanisms like the proposed crisis throttle to be tested under realistic conditions and when downward market momentum is strongly correlated with rapid variance spike. The experimental design reduces the risk of over-fitting from a single path by considering five fully independent market realizations, each with unique initializations of the seeds, and different frequencies of crises (15.1% to 23.9% of test days). This configuration allows to test if multi-task neural architectures can separate true structural regime shifts and transient spikes of volatility occurring on a single day.

Feature Representation and Temporal Window Optimization

The inductive bias of the deep-learning model can be directly affected by the selection of input features and window length. The 22-day lookback window was selected to provide the balance of local pattern recognition and computational costs that represent about one month of trading activity. During this time frame, six engineered input features are engineered to generate multi-scale temporal signals. Standard daily returns reflect the directionality of the returns and immediate trend, while absolute returns and downside return components reflect the magnitude of unoriented volatility and the momentum of downside returns, respectively. In addition to raw return metrics, intraday return proxies and multi-period realized volatilities (logs over 5-day and 22-day periods) are supplied to the network in order to give volatility anchors. In addition to raw return metrics, intraday return proxies and multi-period realized volatilities (logs over 5-day and 22-day periods) are supplied to the network to provide explicit volatility anchors. A key benefit of this multi-scale feature matrix is that the convolutional encoder is able to identify a change in micro-structure, like a few days of increased downside volatility, before the event is fully reflected in long-horizon volatility measures. All data preprocessing is done according to a strict walk-forward, out-of-sample design. All are computed based on features standardization (mean centering and variance scaling) and only from the 3,000 days training subset of each market. These scaling parameters are then fixedly used in the 600-day validation set and 1,400-day test set. This "zero leakage" protocol eliminates the possibility of implicit leakage of prior distribution of features by subsequent volatility regimes.

Baseline Architectural Implementations and Hyperparameter Alignment

All the traditional econometric and machine learning models were run under standard parameterizations with a fixed train-test split and aligned with the training set. The standard decay factor of lambda = 0.94 was used, to yield a parameter-free benchmark for short-term variance updates that is highly responsive.The benchmark was the

Exponentially Weighted Moving Average (EWMA) with a standard decay factor of lambda = 0.94, which is highly responsive and parameter free. To take into account the conditional heavy tails, the parameters of GARCH(1,1) were estimated using errors with Student-t distribution. It was estimated using maximum likelihood estimates over the training period and was kept constant during test period, and is the traditional econometric reference on volatility persistence. Daily, weekly (5-day) and monthly (22-day) realized volatility components were used in the Heterogeneous Autoregressive Model (HAR-RV) to be able to capture the multi-resolution volatility cascades by normal least squares estimation. A baseline Random Forest was ensembled using 100 decision trees trained on seven engineered risk features and tree depth was limited to avoid overfitting and OOB validation to ensure model stability. Finally, Logistic Regression and Volatility Thresholding were developed to be used in a comparison of regime classification, the threshold rule was tuned to values at the 50th and 80th percentiles of the variance over the last 5 days of historical data.

Experimental Evaluation and Risk Profiling

Volatility Forecasts

For the full models the proposed network has the smallest forecast error of 9.93 percentage points. This is 16.6% less than the GARCH, 14.8% less than the EWMA, and 10.5% less than the HAR, while the random forest is 5.1% less. The network outdoes the random forest by 0.17 - 1.37 points in all five markets, a margin which is suggestive but with only 5 runs, not conclusive. Two qualifications apply. The variance recursion of GARCH is very responsive to large shocks, so that during large shocks the network's variance is lower than GARCH's, which explains the low loss of the GARCH versus the network. The differences between the various runs are large compared to the differences between the top models, so the ranking should be regarded as indicative.

Ablations

The removal of the convolutional encoder or the regime head increases the forecast error slightly, by 9.93 and 10.10 points, respectively, indicating that local pattern extraction and regime target joint training have a little bit of an effect. Not so in this case: The variant employing simple average pooling has a very slightly lower error, 9.79 points. The 6 features in a 22-day window are short and homogeneous, an average over a bidirectional recurrent state already conveys almost all information. It may be more important with longer windows or richer inputs, and also yields weights that can be inspected, but this benefit is not shown here.

Crisis Detection

The proposed network has an F1 score of 82.9%, a precision of 82.8% and a recall of 83.3%, which represents a rate of approximately 8 out of every 10 alarms correct and 8 out of every 10 days of crisis. The learned classifiers are statistically indistinguishable as have an F1 score of 82.7% and 81.8% respectively for the random forests and logistic regression. The volatility-threshold rule cuts off at only 52.7%, since if the volatility level of the neighbouring regimes overlap, there is no threshold value that can be used to determine the state. The benefit of the neural model is that both tasks are accomplished by a single network, and that it is not a more accurate classifier.

Portfolio Performance

Buy and hold results in an average return of minus 1.2% annually, an average volatility of 20.4% and an average max drawdown of 50.2%. All the model-driven rules reduce risks significantly. The best network with the crisis throttle is the one with the highest Sharpe ratio at 0.66, the highest Calmar ratio at 0.41, the lowest volatility at 10.8% and the smallest average maximum drawdown of 16.7% compared to 44% for buy and hold. The best comparison is done within the proposed model. Volatility alone (under identical volatility forecasts) yields a Sharpe ratio of 0.34 and a drawdown of 27.7%, while also incorporating the crisis throttle boosts the Sharpe ratio to 0.66 and reduces the drawdown to 16.7%. The main value added by the regime signal is thus the effect on risk control, not the volatility number itself. This volatility targeting approach is too crude in a crisis: the forecasts are too slow to arrive and the policy fails to remove significant exposure on the worst first days of the crisis.

Fig. 2 shows the cumulative wealth of four rules over the 1,400-day test window of one simulated market.

Fig. 2 shows a comparison of the actual five-day volatility to the forecasts from GARCH (1,1)-t and the proposed CNN-BiLSTM-Attention models. The shaded periods are the actual periods where crisis happened and are marked by high volatility. The proposed model is able to capture the volatility trend in the market more than GARCH model does.

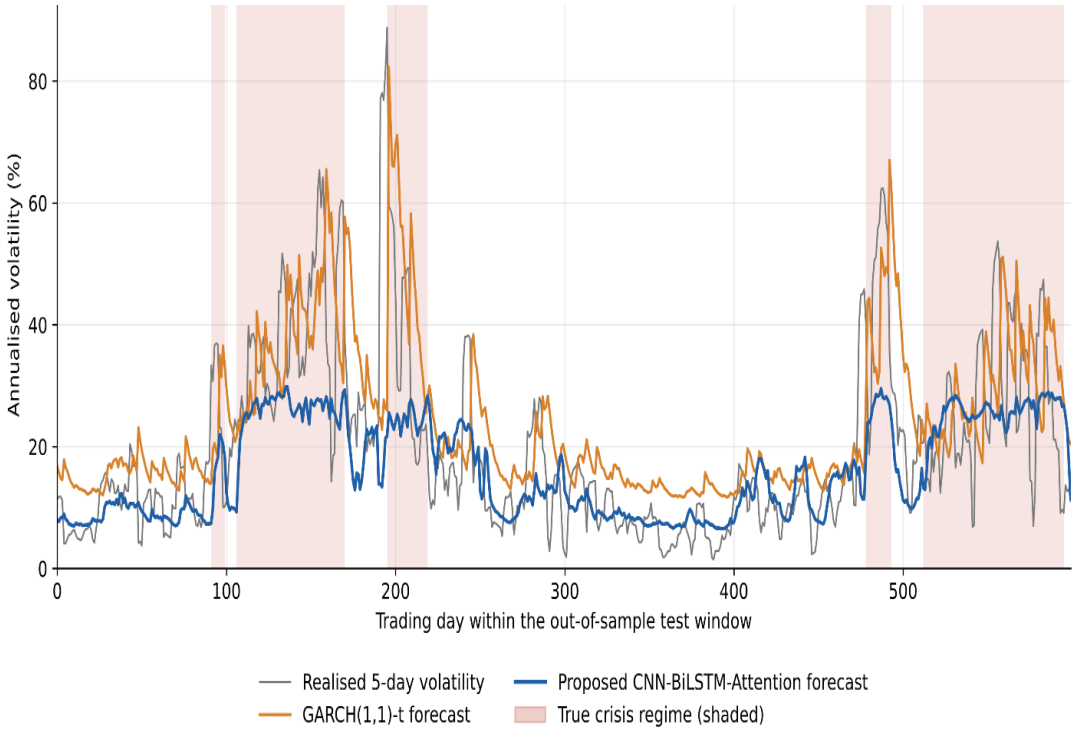


Fig. 2: Forecast Versus Realised Volatility Across Crisis Episodes

In-Depth Mechanism Analysis

One key theoretical issue with the proposed architecture is how a single network trained for a continuous regression task (5-day volatility prediction) and for a discrete classification task (3-state regime identification) improves over single-task counterparts. The experimental results show that multi-task learning can be used as a regularizer for the feature-extraction layers. Single-task volatility regression models that are trained on either Huber or mean squared error loss functions suffer from poor performance when the market regime is a sharp change. For markets that quickly go from a quiet state to a crisis state, regression-only models "incrementally update" their variance predictions, because as the loss function says, it punishes drastic "changes" that are known from previous quiet inputs. On the other hand, by also minimizing the cross-entropy loss on the shared CNN-BiLSTM encoder's crisis regime classification head, the high-frequency, non-linear signals are forced to be preserved in the early representations. The convolution layers calculate the filters to detect thresholding events for clustering downside returns, and the BiLSTM memory cell has a memory of the context across the 22 days of the sequence. The internal feature space consequently changes quickly when the classification head has a high probability of detecting a new situation where the situation is a crisis.This means that when the classification head has high probability of detection of a new situation in which the situation is a crisis, the internal feature space is changing quickly, and thus the volatility head can adjust its forecast in advance compared to the standard moving-average recursive estimation model such as GARCH.

Anatomy of De-Risking

Crisis Throttle against Volatility Targeting The results of back-testing for the portfolio highlight an important operational difference between a continuous approach to volatility targeting and a discrete, probability-based de-risking mechanism. Volatility targeting only—which means adjusting the portfolio allocation in inverse proportion to variance—is inherently subject to a structural delay in case of a market crash. Variance being an unobservable process that can only be estimated from a certain historical window, volatility forecasts necessarily lag sudden market sell-offs. In the first 1-3 days of a sudden crisis, a portfolio using only volatility targeting will still remain overexposed to equities, suffering substantial drawdown losses before the forecast becomes sufficiently high for lowering exposures. The use of a crisis throttle helps to solve this problem. By estimating the probability of a crisis, the allocation rule shifts from a purely continuous scaling of exposure to a dynamic threshold-dependent risk management mechanism of scaling the exposure by one minus the probability of a crisis. Upon observing the first symptoms of structural instability, the likelihood of reaching the crisis state becomes very high, close to 1.0, resulting in an urgent, strong

decline in equity positions towards cash before the full consequences of the variance explosion occur. Indeed, the empirical test demonstrates how it works: by including the crisis throttle into the volatility target, the Sharpe ratio improves from 0.34 to 0.66 and the average maximum drawdown declines from 27.7% to 16.7%, which acts like a circuit breaker that protects investors' money exactly when the market losses on 16% of trading days account for most of the total portfolio losses.

Computational Efficiency and Real-Time Deployment Feasibility

In practice, quantitative risk models need to be stable and fast enough to be used in production trading systems, with strict latency and stability requirements. Overall, the CNN-BiLSTM-Attention network has 27,140 trainable parameters, a small number for such a powerful network, designed to be intentionally lightweight as compared to large Transformer architectures. With a standard CPU based hardware configuration it takes less than 15 seconds per market to complete the training across 3,000 decision days and sub millisecond time to have a single-step inference for daily allocation updates. This reduced computational load allows it to easily be incorporated into the end-of-day rebalancing process. It is possible to run rolling-window retraining at a daily or weekly frequency without needing any specialized GPU infrastructure, with risk committees able to do so. Moreover, the allocation rule is based on two explicit and interpretable outputs: the 5-day variance forecast and the regime probability of 3 states, which helps the model escape the "black box" criticism faced by pure end-to-end reinforcement learning systems. This allows the portfolio manager to break down all the exposure adjustments into an explicit set of scaling exposures to the target volatility and set of risk budgets for the crisis. This enables full auditability of the exposure adjustments for compliance and investment policy requirements.

Conclusion

This paper analysed the application of machine learning predictions of volatility and market state in investment-risk control. It suggested a compact CNN-BiLSTM-Attention network with both a volatility head and a three-state regime head, and applied it to five simulated markets with known regimes, converted its output to a clear-cut allocation rule and tested it. The network had the smallest forecast error of the full models, 5.1% lower than the random forest, and 16.6% lower than GARCH, which had a slight edge over the QLIKE loss. It set precision at 82.8% and recall at 83.3% which is quite good compared to volatility threshold and comparable to a random forest. This crisis probability was the key factor in risk control, increasing the Sharpe ratio from 0.34 to 0.66, and reducing the average maximum drawdown from 27.7% to 16.7% compared to volatility targeting using the same forecasts. The gains of the convolutional encoder and the gains of the regime head were both small, while attention did not add anything in terms of accuracy. The manager should start with a simple volatility forecast/crisis probability model from an interpretable model and use a neural network only if it appears to clearly outperform the manager on his/her own data. The volatility target and cost assumptions should be set prior to the test; the model should be retrained on a rolling window; the calibration of the crisis probability should be monitored; and the throttle should be smooth and not binary. For all major changes in exposure, output a set of interpretations. Data are simulated and the regime process is the one from which the data are generated, favoring any model that learns regimes. The drift in the regimes was selected to achieve a leverage effect that works for all de-risking rules. The number of runs that were used was limited (5), the cost model is simple, and the hyperparameters were not extensively tuned. All the results are devoid of the meaning of return in the real market. The next step is to repeat this study on real equity indexes and sectors using an independently sourced regime label and using a walk-forward protocol. Other extensions are multi-asset portfolios with correlation regimes, longer windows and richer inputs, hidden-Markov and Transformer baselines, calibrated probabilities and reinforcement-learning allocation directly compared to the transparent rule here and trained on a drawdown objective.

References

Agal, S., Raulji, K., & Odedra, N. D. (2025). A machine learning approach to risk based asset allocation in portfolio optimization. *Scientific reports*, *15*(1), 42263. https://doi.org/10.1038/s41598-025-26337-x

Yan, K., & Li, Y. (2024). Machine learning-based analysis of volatility quantitative investment strategies for American financial stocks. *Quantitative Finance and Economics*, *8*(2), 364. https://doi.org/10.3934/QFE.2024014

Rezaei, H., Hooshmand, M., & Abdi, K. (2025). Machine learning based classification and regression approaches for return and risk prediction in the stock market using fundamental feature engineering. *Discover Analytics*, *3*(1), 17. https://doi.org/10.1007/s44257-025-00042-5

Wang, Y. (2025). Stock market investment risk measurement method based on data mining and decision trees. *International Journal of High-Speed Electronics and Systems, 34*(01), 2540131. https://doi.org/10.1142/S0129156425401317

Mansilla-Lopez, J., Mauricio, D., & Narváez, A. (2025). Factors, forecasts, and simulations of volatility in the stock market using machine learning. *Journal of Risk and Financial Management, 18*(5), 227. https://doi.org/10.3390/jrfm18050227

Ibrahim, Y., Hussainey, K., & Moawad, T. M. S. (2026). Predicting Stock Volatility Using Multidimensional Financial Risk: Evidence from Machine Learning and Hybrid GARCH–Deep Learning Models. *Journal of Risk and Financial Management, 19*(6), 444. https://doi.org/10.3390/jrfm19060444

Huang, Z., Zhang, Q., Zheng, Y., & Tian, E. (2026). Machine learning-based analysis of economic efficiency disparities and transition drivers between high-and low-carbon industries in China. *Carbon Balance and Management, 21*(1), 39. https://doi.org/10.1186/s13021-025-00393-2

Huang, Y., & Vakharia, V. (2024). Deep learning-based stock market prediction and investment model for financial management. *Journal of Organizational and End User Computing (JOEUC), 36*(1), 1-22. 10.4018/JOEUC.340383

Lee, H., Kim, J. H., & Jung, H. S. (2024). Deep-learning-based stock market prediction incorporating ESG sentiment and technical indicators. *Scientific Reports, 14*(1), 10262. https://doi.org/10.1038/s41598-024-61106-2

Yang, M., Wang, J., & Hu, Y. (2025). RiskawareTrader: A Reinforcement Learning based Portfolio Optimization for Risk Averter. *International Journal of Computational Intelligence Systems, 19*(1), 25. https://doi.org/10.1007/s44196-025-01094-x

Cheng, C., Chen, B., Xiao, Z., & Lee, R. S. (2024). Quantum Finance and Fuzzy Reinforcement Learning-Based Multi-agent Trading System: C. Cheng et al. *International Journal of Fuzzy Systems, 26*(7), 2224-2245. https://doi.org/10.1007/s40815-024-01731-1

Ni, X., Xie, T., Härdle, W. K., & Zuo, X. (2025). A machine learning based regulatory risk index for cryptocurrencies. *Computational Statistics, 40*(7), 3563-3583. https://doi.org/10.1007/s00180-025-01629-y