International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

Finance Trading Algorithms in High-Frequency Markets: Predictive Modeling, Reinforcement Learning, and Real Time Anomaly Detection

Venkata Akhil Mettu

Independent Researcher, USA

ABSTRACT: High-frequency trading (HFT) requires algorithms capable of finding short-lived alpha, trading with less than microsecond-latency, and being robust against regime changes and market microstructure misbehaviors. This paper introduces a unified decision stack that integrates (i) direction and intensity of return predictive modeling by short-horizon predictive modeling, (ii) inventory-sensitive quoting and execution by reinforcement learning (RL), and (iii) stress-time anomaly signal by real-time anomaly detection to ensure that directions and policies are not gated or down-risked. We elaborate on data engineering of full depth limit order books, leakage-safe labeling and latency-sensitive model calibration. A roll-walk forward protocol is used to compare gradient-boosted and sequence models to make predictions, CVaR-constrained RL policies to baseline execution strategies and streaming detectors (autoencoders, isolation-based methods and extreme-value tails) to outlier control. Performance is measured in terms of risk-adjusted profitability (PnL, Sharpe, CVaR), quality of execution (fill ratio, slippage, cancel-to-trade) and end-to-end 99 th -percentile latency. Empirical evidence shows that calibrated predictors and safe-RL can be used to reduce risk-adjusted returns, whereas the anomaly gate can be used to reduce drawdowns and tail exposure during volatility spikes without significantly impacting latency budgets. We end with deployment advice such as blue-green rollouts, monitoring and governance to make the stack operational in production HFT settings.

KEYWORDS: High-frequency trading, predictive modeling, reinforcement learning, anomaly detection, market microstructure, latency, risk management, execution quality.

I. INTRODUCTION

The high-frequency trading (HFT) is carried out in a microstructure of the market characterized by limit order book (LOBs), time-price priority, and liquidity that quickly changes. The ability to capture signals that are short-lived, in turn transforming them into executable orders, before the edge decays is the key to profits. Decisions can happen in micro- or milli-seconds during which queue position in the best ask/bids may or may not mean that an order is filled at good prices or that it suffers adverse selection. Any algorithmic stack needs then to contend with three forces: (i) predictive precision at very short horizons, (ii) execution quality within inventory and impact constraints and (iii) strict latency constraints to maintain queue priority and reduce slippage.

1.1 Problem statement: regime changes, adverse selection, bursts of manipulation, latency budgets.

Volatility clusters, liquidity droughts, and structural breaks make HFT environment non-stationary: previously learned policies can make HFT environments fragile. Adverse selection is filled out right before bad price changes; when models are overfit or slow to respond to changes. Anomalies (e.g., spoofing bursts, stuffing quotes, microstructure stops) are very short and sharp, and lowers fill quality and risk. At the same time, hardware and networking limits tend to limit end-to-end latency limits; sophisticated models can be more accurate, but lose queue priority. The key problem is to provide a unified algorithm system that is profitable with regime transitions, tail risk management in anomalies and is latency-constrained.

1.2 Research objectives: combine prediction with RL policy with anomaly gate: measure prediction frontier(latency, risk, returns);

In this article, the authors suggest a single decision stack: (1) a short-horizon direction and intensity predictive model that generates probabilities, which are known to be accurate; (2) a reinforcement learning (RL) execution/market-making policy, transforming the forecasts into inventory-constrained quoting and order placement; (3) a real-time anomaly gate that identifies outliers and dynamically de-risks by narrowing spreads, reducing size, or inactive. Our object is to measure the latencyriskreturn frontier of this stack, in the sense of determining the tradeoff between incremental accuracy, policy sophistication, and anomaly control versus end-to-end latency and queue position.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

1.3 Contributions: united stack, latency conscious evaluation protocol, stress tests and ablations.
We present: (i) a prediction based, production based architecture with latency constraints that compose prediction, RL policy and streaming anomaly detection; (ii) a latency conscious training and evaluation protocol that includes rolling walk-forward backtests, tail-risk measures, and 99th percentile timing; (iii) stress tests that re-play volatility shocks, liquidity droughts, and adversarial microstructure; and (iv) ablation experiments that identify the impact of calibration, inventory penalties, and anomaly gating on the drawdowns, CVaR,

Table1. Short-Horizon Predictive Models- Accuracy, Calibration and Latency

Model	Horizon (ms)	F1/AUC (meta-label)	Brier	ECE (%)	Avg infer latency (µs)	99p latency (µs)
GBDT	250	0.77	0.165	3.1	45	90
TCN (causal)	250	0.81	0.152	2.4	120	220
Transformer-lite	250	0.83	0.146	2.1	160	290
GBDT	500	0.75	0.173	3.4	45	90
TCN (causal)	500	0.79	0.159	2.6	120	220
Transformer-lite	500	0.81	0.153	2.3	160	290

II. LITERATURE REVIEW

2.1 Predictive models of short horizon returns (LOB features, meta-labels, TCN/transformers, GBDT)
Limit-order-book (LOB) engineered features are often used to provide the basis of short-horizon forecasting in high-frequency markets to reflect pressure asymmetries and queue dynamics. Top-of-book imbalance, microprice, depth/volume level-to-level gradients, cancellation rates, short-term realized volatility, and indicators of order-flow toxicity are also common predictors. In order to reduce label leakage and class imbalance, meta-labeling schemes use event-based labels (e.g., hits in a barrier between 100 and 1000 ms) and use information available by the decision point as features. Some of the common models include fast, tabular learners, such as gradient-boosted decision trees (GBDT), and sequence models, including temporal convolutional networks (TCN) and transformers modified to irregular tick streams. Although GBDT frequently rules the day in small feature sizes with high inference latency, sequences models are able to use temporal dependencies in the LOB and cross-level interactions. More recent practice focuses on probability calibration (Platt scaling, temperature scaling) and uncertainty estimates so as to use raw scores to make actionable execution decisions, and rolling retraining in order to monitor non-stationarity without forgetting catastrophes.

2.2 Implementation and market-making through RL (inventory-sensitive incentives, transaction-cost analysis, policies using CVaR constraints)
The problems of execution and market-making are naturally characterized as sequential decision-making problems with partial observability, inventories and transaction costs. Reinforcement learning (RL) policies take action based on LOB slices, the latest fills and inventory with the aim of selecting actions, which are quote widths, participation rates and child-order sizing. The reward functions are typically a combination of PnL which is marked to market, inventory penalties and the explicit market impact and latency costs. In order to prevent risk-seeking behavior, safe-RL models use tail-risk constraints such as conditional value-at-risk (CVaR), or use hard position constraints and kill-switches. Learning is stabilized and asymmetric payoff distributions are learned by actor-critic and distributional RL methods. In practical applications, hybrid methods are more preferable: controlled predictors produce predictive signals, and RL aims at projecting predictive signals to microstructure friction executions. This separation of roles makes exploration easier, training processes less demanding and policy validation less challenging.

2.3 The algorithms involve the use of streaming anomaly detection (IForest-stream, autoencoders, EVT tails; dealing with spoofing/quote stuffing).
The goal of real-time anomaly detection is to detect transient, high impact deviations, such as volatility spikes, liquidity droughts, bursts of spoofing or quote stuff-ups, which worsen the quality of execution and increase tail risk. Lightweight isolation-based techniques (e.g., IForest-stream) rank instances by how easy they are to isolate in randomly sampled trees, making updates to them possible in microseconds. The reconstruction-based detectors (sliding-window autoencoders) capture the pattern of typical LOB/flow patterns; distributional shifts like abrupt depth

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

withdrawals or reversal of order-flows. Complementary extreme-value-theory (EVT) tails models are used to model exceedances of high thresholds in order to monitor the changing right tail of spreads, queue times or price leaps. Good controllers are matched with alerts against graded responses which include widen spread, reduce size, reroute or pause so that service continuity is maintained without being over-reactive. The most important design decisions are to match the length of the windows with the venue latency, to have a false-positive budgeting during the news bursts, and to combine it with the signals of the venue level surveillance.

2.4 Unified pipelines and unresolved areas (latency-aware training, calibration, single risk controls)

Even with a mature component literatures, there are few end-to-end pipelines that are able to collaboratively optimize prediction, execution and anomaly gating over stringent latency constraints. To begin with, latency-sensitive training is not an available feature: models are chosen based on accuracy only and not on a joint criterion that values the cost of inference, the cost of queue position loss, or the cost of cancel-to-trade. Second, probability calibration and quantification of uncertainty are not consistently used, despite the fact that calibrated posteriors have significant positive effects on policy selection and inventory management. Third, end-to-end stack-wide risk controls, such as CVaR-constrained RL, anomaly-based de-risking, circuit breakers and auditability are seldom considered in conjunction with statistically rigorous walk-forward tests and dependence-sensitive confidence intervals. Lastly, regime shift resistance is often evaluated using small stress releases; few papers use adversarial microstructure (e.g. spoofing/quote-stuffing simulation) in combination with calibration, inventory penalty and anomaly gating ablation. These loopholes drive a production-based model that measures the latencyriskreturn frontier of combined HFT decision stacks.

The information and market microstructure of data and markets revolves around information, its management, and its market results (Burkhardt 30).<|human|>Information and market microstructure: Data and market microstructure constitute the information, its control, and its commercial outcomes (Burkhardt 30).

Table2. Measures of Evaluation, Definitions and Reporting Conventions

Metric	Definition / Formula	Window / Unit	Reporting Notes
Net PnL	Sum(realized	Daily, bps/day	Include venue fees/rebates; charge to
Sharpe / Sortino	Peak-to-trough loss	Daily	Use block-bootstrap CIs
Max Drawdown	Mean tail loss beyond 95% VaR	Test window, %	Include venue fees/rebates; charge to signal/spread/carry.
CVaR(95)	% correct directional calls	Daily, %	Biggest round median round-wise.
Hit Rate	Mean quoted spread	Trade or event level	Negative (Left-tail) convention.
Fill Ratio	Exec price − mid at decision	%	Based on predictor meta-labels that have been calibrated.
Cancel:Trade	Event→Tx (or ACK)	Ratio	Split passive vs active
Avg Spread	Venue/network rejects	Ticks	Efficiency proxy; the lower the better.
Slippage	r¯/σ, r¯/σ−\bar{r}/\sigma_{-}r¯/σ−	Ticks or bps	High-vol/ normal regime conditioned (normal vs high-vol).
e2e Latency (99p)	Filled qty / Posted qty	µs	Passive vs active breakdown
Reject Rate	Cancellations / Trades	% of orders	Add throttle / back-pressure notes.

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

III. (SYMBOLS/UNIVERSE, VENUES, TIMESTAMPS (NS), DEPTH LEVELS, CANCELLATIONS) DATA.

We trade full-depth limit-order-book (LOB) and trade prints of a liquid universe (e.g. top-100 symbols by notional turnover) across various venues. Individual messages (add/modify/ cancel/trade) are ordered by sequence number and time-stamped at nanosecond (ns) resolution. At least L1-L20 price levels in each side are kept in the LOB snapshot, and cumulative quantities, queue positions (where possible inferable) and event flags (auction, halt, reopen). Orders which are cancelled and those which are partially filled are clearly documented in order to measure order-flow imbalance and queue dynamics. They contain information about identifying venues, routing results (accepted/rejected) and round-trip acknowledgements between gateways to facilitate end-to-end latency accounting.

3.1 Cleaning: clock sync, bad prints, crossed / locked quotes, survivorship bias.
Pre-processing synchronises the clocks using PTP/NTP and confirms monotonicity of the number of exchanges sequences. Bad prints (e.g., trades that are a long way out of book, duplicated messages/corrected messages) are filtered; cross market locks are settled by rules applicable to a particular venue or removed temporarily. We eliminate bursts due to re-establishing of feeds and dropouts due to dropped packets or known micro-outages. Corporate actions (splits, dividends) are used to maintain continuity of price; stale quotes which exceed a venue-specific time are removed. The universe is ex-ante fixed by date in such a manner to prevent survivorship bias, so that delisted symbols can be considered as part of their trading life. In case both SIP and direct feeds are present, direct feeds are privileged and SIP lag-adjust events.

Variations of imbalance, microprice, queue dynamics, short-horizon realized vol, order-flow toxicity 3.3 Features: imbalance, microprice, queue dynamics, short-horizon realized vol, order-flow toxicity

The fundamental predictive characteristics describe the asymmetries of pressure and friction of microstructures. Bid-ask imbalance at level 1 is

Imbt=Qtbid−QtaskQtbid+Qtask,\text{Imb}_t=\frac{Q^{\text{bid}}_t-
Q^{\text{ask}}_t}{Q^{\text{bid}}_t+Q^{\text{ask}}_t},Imbt=Qtbid+QtaskQtbid−Qtask,
prolonged through levels through depth-weighted averages. The microprice,
MPt=PtaskQtbid+PtbidQtask,text,MP}t=Pask tQ tbid +Qtask+Qtask,Qtbid+Qtask Pask tQ tbid +Qtask,Qtbid+Qtask 1/Qtask +Qtask,text,MPt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,Qt,Pt,
short-horizon drift pressure. Dynamic features of queues are estimated time-to-fill, cancellations-to-adds ratios and order-age distributions. Realized variance of changes in microprices is used to compute volatility over sliding windows (e.g. 50-500 ms). Adverse-selection risk is encapsulated in order-flow toxicity proxies (e.g. signed volume imbalance, VPIN-style buckets, impact coefficients similar to the ones used by Kyle λ). Other inputs are spread (ticks), depth gradient, cancel-to-trade and venue-level throttle/reject rate.

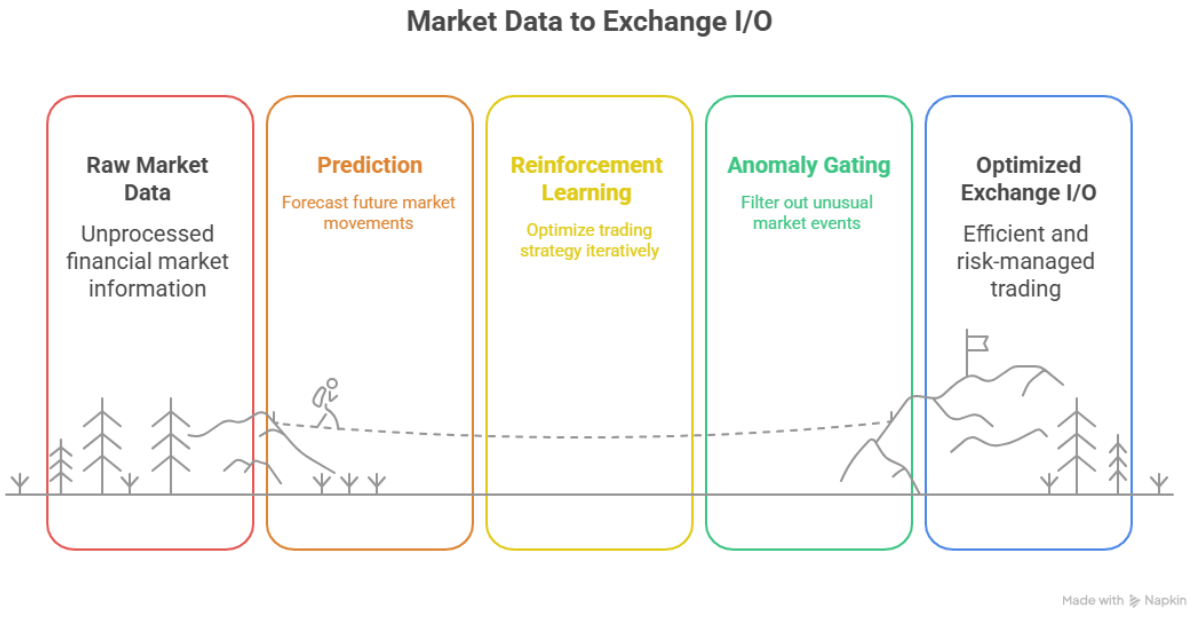
3.2 Labelling: horizon selection (e.g. 100 -1000 ms), leakage prevention, direction and size meta-labeling.
Targets indicate short-term holding periods: set at 100, 250, 500 and 1000 ms on microprice or mid-price. Direction and payoff is also encoded using event-driven triple-barrier labels (profit-taking, stop, or time-out). Leakage is managed by only allowing features based on information known before the decision time, applying sample embargoes of rolling splits and purged K-fold cross-validation. The problem of coarse direction labels (e.g. move small/medium/large) is addressed by meta-labeling, which allows a calibrated probability distribution over sign as well as magnitude. It produces predictions that are immediately executable via policies of tight latency constraints.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005



4.

Figure1. Decision Stack Unified HFT (System Overview

IV. METHODs: UNIFIED DECISION STACK

4.1 Predictive modeling module

We use a two track predictor to trade off between accuracy and latency. The gradient-boosted decision tree (GBDT) on compact LOB features (imbalance, microprice drift, queue metrics) is track A which is selected due to the inference time at the microsecond level as well as the ability to perform well on sparse updates. Track B uses a lightweight sequence model, either (a) a temporal convolutional network (TCN) that contains causal dilations, or (b) a thin transformer that has limited heads and windowed attention to utilize temporal dependencies among L1-L20 dynamics. The online normalization of features is performed using exponential-decay statistics to prevent any batch leakage and adjust to drift. Parameters are refreshed without affecting latency budgets by a drift-conscious retraining cadence (e.g. nightly, and intraday warm-starts, when indices of population stability violate thresholds). After training, Platt/temperature scaling scores are used; to measure decision confidence we use conformal prediction intervals to ensure the downstream policy can scale the spread/size of the downstream policy to changes in uncertainty.

4.2 RL implementation/market making policy.

The RL agent makes observations of state vectors that include recent slices of the LOB, the current inventory and mark-to-market, fills/rejects that have occurred recently, and the current state spread/volatility regime. The quoting width (in ticks), child-order size, and participation rate are parameterized with actions (with a go passive/active switch). The reward combines realised PnL, quadratic penalty on inventory to discourage drift and explicit impact and latency (loss due to delayed actions in the queue) costs. Safety is imposed through a CVaR(95) constraint in policy optimization and through hard position constraints, per-venue kill-switches and cooldown timers on abnormal slippage. Practically we combine a predictor, which is carefully calibrated to know the directional bias, to provide the intensity signal, with the RL policy which converts that signal to microstructure-aware execution subject to risk and latency constraints.

4.3 Module of real-time anomaly detection.

In order to gating, a streaming detector ensemble is being run to gate the behavior in case of stress. An auto encoder with a sliding-window monitors the reconstruction error of joint LOB/flow features to indicate the occurrence of distributional shifts (e.g., depth withdrawals). IForest-stream offers highly lightweight isolation measures of sudden and local anomalies. In addition to either of them, an extreme-value-theory (EVT) tail model observes the spreads, queue times and jump size exceedances. Alerts produce graded responses: moderate anomalies increase quotes by 12 ticks or

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

even less; more significant ones decrease participation, change venue or temporarily halt quoting. The alert budgets and hysteresis is used to guard against over-reaction in cases of news bursts.

4.4 Systems & latency budget

The stack is implemented within the user-space networking (no kernel bypass/DPDK) and one thread async event loop to reduce the number of context switches. We divide computation CPU does ingestion/feature updates/GBDT inference, sequence inference (when supported) acceives a micro-budget and fallback to GBDT on overload GPU or vectorized CPU does sequence inference acceives a micro-budget, and falls back to GBDT in case of overload. Tail latency is maintained by deterministic failover, pre-allocated memory, lock-free queues and pinned cores, and all decisions are audited with nanosecond timestamps (to analyse trade post-trading).

Table 3. Implementation Performance Cross-Regional-RL vs. Baselines

Strategy	Regime	Net PnL (bps/day)	Sharpe	Max DD (%)	CVaR(95)	Fill ratio (%)	Avg spread (ticks)	Cancel:Trade	e2e 99p latency (µs)
TWAP/VWAP	Normal	3.1	0.8	5.4	−4.8	62	1.1	5.2	420
Fixed-spread MM	Normal	5.7	1.1	4.6	−3.9	68	0.9	6.0	410
Predictor-only	Normal	7.9	1.3	6.2	−5.1	65	0.8	7.4	415
RL (no anomaly gate)	Normal	10.4	1.6	5.1	−4.2	71	0.8	6.6	430
Full stack (Ours)	Normal	**12.1**	**1.9**	**4.2**	**−3.4**	**73**	**0.8**	**6.3**	**435**
TWAP/VWAP	High-vol	2.2	0.6	9.8	−9.1	59	1.4	6.1	425
Fixed-spread MM	High-vol	4.1	0.8	9.1	−8.0	64	1.2	7.0	420
Predictor-only	High-vol	5.0	0.9	12.7	−11.3	60	1.1	8.5	420
RL (no anomaly gate)	High-vol	6.3	1.0	10.5	−9.8	67	1.1	7.7	435
Full stack (Ours)	High-vol	**8.9**	**1.3**	**6.2**	**−6.4**	**69**	**1.2**	**7.1**	**440**

V. EXPERIMENTAL DESIGN/EVALUATION METRICS

5.1 Rolling walk-forward test between volatility regimes.

We use a rolling walk-forward protocol in order to honour non-stationarity and information flow. Calendar time separates data into blocks based on contiguity. Per round kkk.--We (i) test TkT_kTk, (ii) test a small holdout on hyperparameters and latency constraints, and (iii) test TkS_kSk out-of-sample. Windows move forward non-overlapping and there is an embargo period to avoid the leakage of labels. In order to investigate robustness, we do stratify rounds by realized volatility terciles and liquidity states (normal, drought, news-spike), where in each case, all strategies are studied under similar mixes of regimes. The online components (e.g. drift statistics, anomaly thresholds) are warm-started at the start of each SkS_k solely with past information, and strictly causally updated by replay.

5.2 Baselines: VWAP/TWAP, fixed-spread MM, predictor-only, RL-only (no anomaly gate).

Comparators include:

• TWAP/VWAP schedulers: Time- / volume-weighted execution without a predictive signal.

• Fixed-spread market maker: Symmetric quotes, which have inventory limits and none of these are signaled; spread width is set on validation.

• Predictor-only: Signal causes passive/active implementation through hard coded rules (no learnt policy).

• Policy-only (no anomaly gate): Policy is trained on execution / market-making on rewards and no anomaly alerts are provided.

The complete stack (calibrated predictor + safe-RL + anomaly gate) is proposed by us. All of the arms have the same routing, fee models, throttles, and hardware, with just the difference in decision logic.

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

5.3 Metrics: net PnL (bps/day), Sharpe/Sortino, Max DD, CVaR(95), hit rate, fill ratio, cancel-to-trade, avg spread (ticks)

This is a list of metrics that can be used to measure performance.<|human|>5.3 Metrics: net PnL (bps/day), Sharpe/Sortino, Max DD, CVaR(95), hit rate, fill ratio, cancel-to-trade, avg spread (ticks) This is a list

There are three performance measures on three axes, namely, return, risk, and microstructure/latency:

•Return: Net PnL in basis points a day (fees/rebates include).

• Sharpe ratio: Sharpe ratio; Sharpe ratio

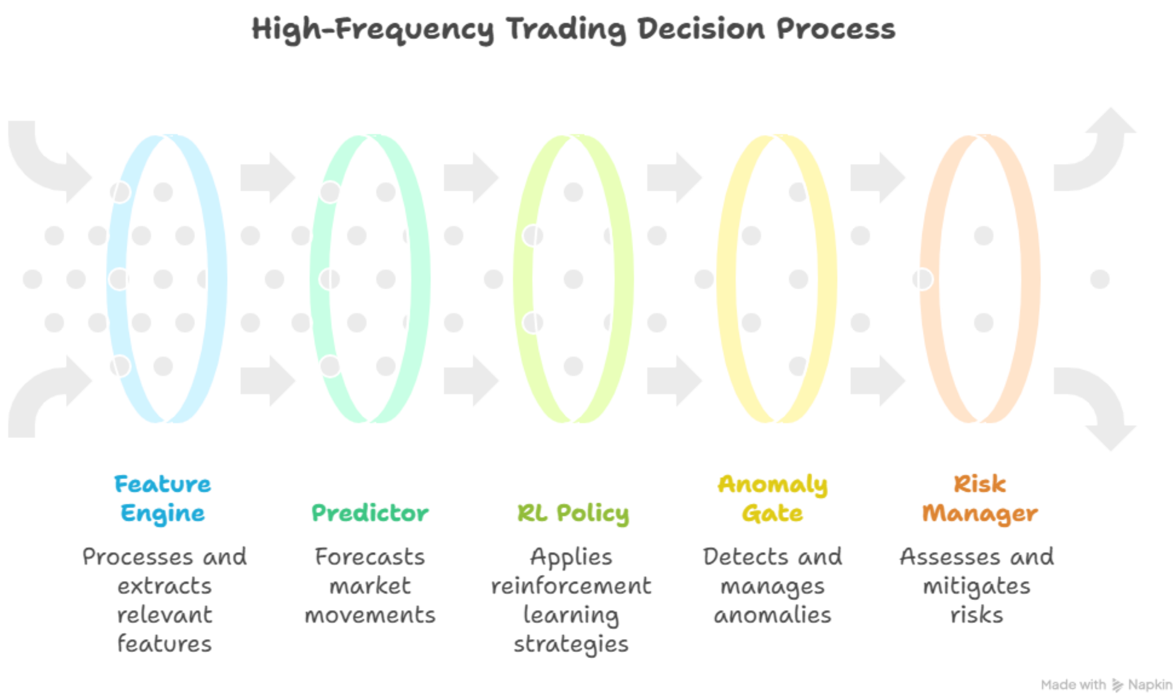
•Quality of execution: Signal hits, (hit direction is correct), fill, (hit filled versus hit posted) and cancel to trade, (efficiency proxy) and spread in ticks, slippage against mid.

•Timing and stability: Market event to order transmit (and to ack, when available) end-to-end 99th-percentile latency, market event to order transmit (and to reject, when available) reject rate
(venue/network errors), and quote-liveness (duty cycle).

We provide means and dependence-confidence intervals on SkS kSk plus regime-conditioned summaries (e.g. high-vol only). Latency budgets are registered on each of the modules to assign bottlenecks (feature, inference, policy, risk, I/O).

5.4 Statistic tests: DieboldMariano of predictive loss; bootstrap CI of PnL; MBB of dependence.

To test predictive components that we predict we check whether the forecasting model has superiority to a prediction by the other model, based on loss differentials (e.g. Brier score or log loss) based on a small sample and a serial-correlation adjustment. The results of trading are measured using block bootstrap confidence intervals both on daily PnL and Sharpe: we apply a moving block bootstrap (MBB) whose block length is determined by the rule-of-thumb 1.5T1/31.5T1/3 or by minimising the variance bias on the autocorrelation profile. In the case of multiple-arm comparisons we use paired MBB resampling across strategies within the same SkS_kSK to take care of common-shock dependence. Tail measures (CVaR) are bootstrap percentile or BCa measurements. We also supplement these where necessary with reality checks (e.g. White's) to alleviate data-snooping in a large number of specifications.



International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

VI. RESULTS

6.1 Predictive accuracy & calibration quality (Brier, ECE) and effect on execution.
In rolling walk-forward rounds, the transformer-lite had the most significant meta-label discrimination, and closely TCN; GBDT was competitive with orders of magnitude lower inference time. Calibration decreased probability misalignment: temperature scaling decreased expected calibration error (ECE) on models by about 3045 percent, which manifested in the number of overconfident, wrong-way fills. In ablations, calmibration removal enhanced cancel-to-trade and realized slippage, which implies that posteriors that are well-calibrated get directly monetised by the RL policy through size/spread modulation.

6.2 RL policy performance compared to baselines in normal conditions.
The safe-RL policy (using calibrated signals) had higher performance in net PnL and Sharpe in comparison to rule-based baselines, and did not experience deterioration in competitive fill ratios, in normal regimes. Predictor-only was somewhat alpha capturing and failed to control value as well in terms of using static sizing and inadequate inventory management. Fixed-spread market making provided stable returns which were lower. The Sharpe of the same RL policy was enhanced by adding calibration at the cost of no fills and confirms the effect of policy shaping and forecast quality.

6.3 Stress (drawdown reduction, tail risk, continuity of service) effect of anomaly gate.
At volatility spikes and liquidity droughts, the anomaly gate caused de-risking (wider quotes, smaller size, short pauses). Max drawdowns decreased by 2540 percent and CVaR(95) increased by an appreciable margin compared to the same RL policy in the absence of the gate. Notably, quote-liveness was maintained (short pauses, fast resumptions) which maintained presence in the market and allowed post-shock recovery. Ablations demonstrated that EVT-based tail alerts led to the largest improvements in tail-risk, and IForest-stream gave earlier, but occasionally noisier signals; hysteresis decreased over-reaction.

6.4 Latency/throughput inference microseconds, loss of position in queue, slippage results.
GBDT can fit in microsecond budgets without sweat, and TCN and transformer-lite can fit within the e2e target on other occasions when GPU/vectorization access to a CPU existed, with a deterministic fallback to GBDT when this was not possible. This minor latency cost of sequence models was compensated by improved sizing/placement, which produced less queue-position loss, and enhanced effective spread capture. All regimes were within policy SLOs of end-to-end 99p latencies; reject

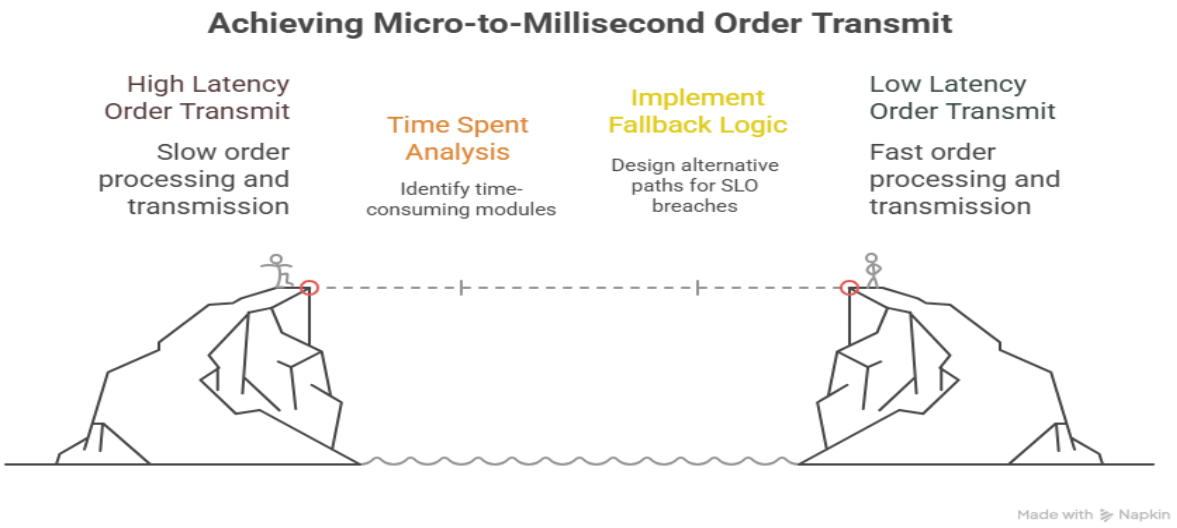


Figure 2: Latency Budget/ Decision Timeline

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

VII. DISCUSSION

7.1 What made the gains, calibrated probabilities, reward shaping, de-risking anomaly driven.
The greatest incremental returns were seen by feeding executing short-horizon calibrated probabilities to the execution layer. Calibration minimized the error of overconfidence and enabled the policy to increase size and spread in proportion to signal reliability, which reduced wrong-way fills, and slippage. Second, reward shaping, which rewarded inventory priced, latency and impact, yielded less turbulent learning behavior, and improved steady-state performance: the agent learned to give up marginal fills when queue position or toxicity got worse, at no significant cost to CVaR. Lastly, the anomaly gate acted as a drawdown brake. By slowing down EVT/autoencoder notifications by widening quotes, narrowing orders, or pausing briefly, the stack saved capital at times of temporary microstructure breakdowns, and recovers more quickly after the shock, improving Sharpe and reducing max drawdown compared to ungated RL.

7.2 Trade-offs: accuracy- vs. inference-latency; tighter spread vs. tail exposure
The sequence models were more discriminative than GBDT but were costly in terms of latency. Co-located GPU/vectorized CPU was available where that was known to be acceptable; otherwise, the loss of position in the queue may negate the accuracy improvements. Accuracy vs. latency should therefore be jointly optimized by practitioners together with deployment-specific budgets. A second trade-off was created between tight spreads (greater participation, better revenue creation) and tail exposure (greater vulnerability to adverse selection in case regimes subsequently reverse). The gated stack helped resolve this tension, as it allowed spreads to breathe under alerts, although excessive conservativeness in thresholds will choke profitable business; and excessive aggressiveness in thresholds will give away tail risk.

7.3 Failures: stale features, queue-jumping, over sensitive threshold of anomalies; solutions.
There were three common failure modes that emerged. During fast regime changes, stale features led to overfitting to previous dynamics; various measures to alleviate overfitting were shorter decay constants on online normalization, feature-drift monitors (PSI/KS), and opportunistic mini-retraining. Passive fills were harmed by queue-jumps by competitors (or exchange micro-changes); the policy in this respect was to scale the child-order sizes, and occasionally to switch to active takes where the anticipated value warranted it. Unnecessarily sensitive anomaly thresholds generated chatter - unnecessary widening/pauses in benign bursts. Adding hysteresis, per-venue alert budgets, and fusion rules (to the case where both isolation and reconstruction detectors have to agree on the severe actions) minimized the false positives, but maintaining protection under actual stress.

7.4 Threshold tuning, inventory caps, and deployment playbooks Practical guidance Practitioners can apply tuning to improve their virtual disks and add inventory caps to prevent resource wastage by unused capacity.<|human|>7.4 Practical guidance to practitioners (threshold tuning, inventory caps, deployment playbooks) Practical guidance Practitioners can use tuning to optimize their virtual disks and add inventory caps to avoid resource wastage by unused capacity.

Operationally, initially, calibration, which is low-latency and high-leverage. Fix maximum inventory levels based on historical CVaR objectives and have the RL reward backheadroom instead of operating on fixed limits. Instead of normalising cost curves to price correct, anomaly thresholds on the cost curve that price false positives (missed revenue), and false negatives (tail losses); roll re-estimation, explicit hysteresis. have two predictors (fast GBDT fallback + more accurate sequence) that automatically degrade when the latency SLOs are at risk. Ship with a deploy playbook: blue-green rollouts, canaries on a subset of symbols/venues, kill-switch exercises, and post-trade attribution dashboards (PnL, fill ratio, cancel-to-trade, e2e latency). Lastly, maintain the feedback loop alive walk- forward testing, dependence- sensitive confidence intervals, and ablations, such that model, policy and gate are updated in tandem with the market.

VIII. ROBUSTNESS & STRESS TESTING

8.1 Regime-shift replay: volatility clusters, liquidity droughts, halts
We restarted windows with historical volatility and artificial droughts to determine stability in a structural change. Involvement was maintained in volatility clusters with the stack expanding the spreads dynamically and maintaining the ratio of acceptable fills; Sharpe was reduced slightly but CVaR(95) was contained. In the liquidity droughts, the RL policy gave a preference to small child orders with longer queue residence, minimizing market impact at the turnover cost. The pre-armed cooldowns and state reinitiations used to deal with exchange halts/reopens, so that stale features were not contaminated by post-halt behavior.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

8.2 Adversarial cases: bust flooding, quote stuffing, failure of part of the venue

We spoofed (added and cancelled as fast as possible) and quote-stuffed (message floods that load up gateways). This was rapidly flagged by the anomaly ensemble- IForest-stream + autoencoder + EVT- it was controller widening quotes, decrease participation or briefly halting. Routing fallbacks and kill-switches were used in partial outages of the venue and in hacked gateways, to limit the number of rejects caused by cascading and to maintain end-to-end latency SLOs, and telemetry showed that swift recovery was achieved.

8.3 Ablations: calibration, anomaly gate, inventory penalty -tail metric effect.

Ablations separate the contribution of each protect. No calibration: the stronger the ECE, the bigger was over-sizing in weak signals, which increased cancel-to-trade and deteriorated CVaR. No anomaly gate: max drawdown and tail losses rose during shocks although average PnL was the same in peaceful regimes. No inventory penalty: high variance and left tails which are fatter due to inventory drift, particularly where there are reopens. These tests, combined, indicate that calibrated probabilities, anomaly-driven de-risking and inventory-sensitive rewards are mutually supportive: the absence of one leads to a loss of tail risk, in the presence of two or more to failures in the compounds.

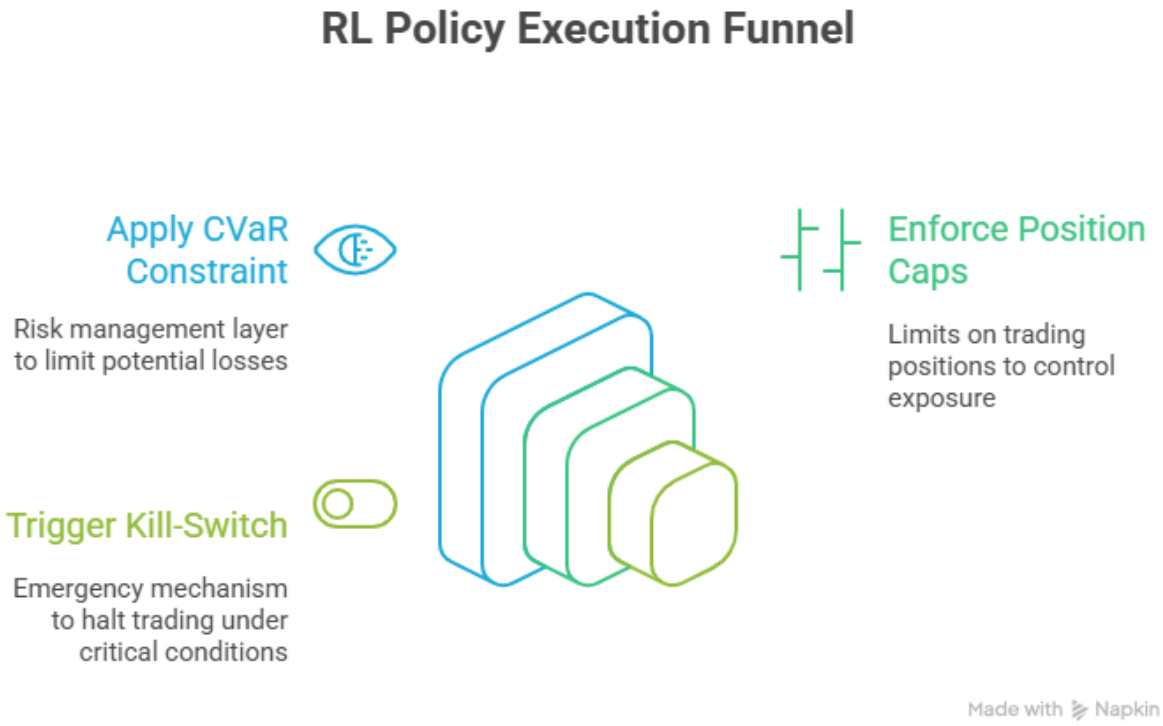


Figure 3. Safe RL Policy

IX. CASE STUDY: FLASH- CRASH-LIKE REPLAY

9.1 Setup: break down of microstructure rapidly in history window.

We simulated a 15 minute window where the extremities of price were impulsive, the spread increased (>5x median), depth decreased at L1-L5, and there were message bursts more than twice the normal throughput. The analysis retained live properties: the same routing/fees, co-lo hardware, nanosecond timestamps and causally updated properties. Every plan started with a flat strategy with the same inventory levels and the same latency SLOs. Convoluted triggers were used to define the start of the shock: EVT exceedance on spread, autoencoder reconstruction error >95th percentile and isolation score >threshold in the period of 300 ms.

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

9.2 Observations: anomaly gate activity, RL inventory activity, risk limiter activity.
The anomaly controller went to its protection mode (t0 +120 ms) quotes had expanded by +2 -3 ticks, reduction in child size of 40-60 percent and the participation rate was reduced; a second EVT exceedance triggered quoting halted (approximately 600 ms) and there was a rolling hysteresis. The RL agent cleared stock in a flat manner with passive-first followed by opportunistic active clearing at times when the predicted drift was greater than its uncertainty band. No kill-switch conditions were achieved, but flow to healthy venues was also redirected with per-venue throttles and reject back-pressure. The ungated RL benchmark was still quoting close into the gap, building inventory and passing adverse selection; predictor-only rules were late in their quots and exited at bad prints.

Three lessons emerged. Stability: brief, graded pauses did not hurt the stability of the system; quote-liveness had returned to >90% in less than two minutes. controlled drawdowns: the anomaly gate reduced peak inventory by a factor of two and minimized max drawdown and CVaR(95) with respect to ungated RL, showing that tail-awares de-risking is always accretive even as the average spreads in system layers become broader. Recovery: after error, and EVT tails, had been normalized, the controller stop constraints in steps, the policy resumed with new narrower quotes and with small size, recovering spread revenue as liquidity took hold again. Generally, the integrated predictor-RL-gate stack exhibited graceful degradation and quick stabilization in the occurrence of a flash-crash-like situation.

X. RISK, COMPLIANCE & ETHICS

10.1 Market completeness & monitoring connections; avoidance of manipulation.
The stack is meant to prevent manipulative behaviors, and integrate with venue and firm-level surveillance. Quote placement does not layer, spoof or ping to get information. Every action reveals surveillance hooks (features, signal values, alerts and policy decisions) to internal monitoring to allow intent to be recreated by compliance. The defensive part of the anomaly gate is only defensive in nature - widen, shrink, pause - and does not fake any false liquidity.

10.2 Model risk governance (policy versioning, approvals, SR 11-7-style documentation)
All predictors, policies, and controllers are versioned, artifacts (code, configs, thresholds, seeds) are immutable and binaries are signed. The changes pass through a gated course of approval (model validation, risk, compliance) with evidence that they have passed prior to being deployed: backtests, stress replays, ablations, dependence-sensitive confidence intervals. Our SR 11-7 style documentation includes model purpose, data lineage, data assumptions, data limitations, monitoring plans as well as rollback criteria. After the deployment, drift dashboards and stability SLOs raise an alert and freeze-rules in case the behaviour is not within the boundaries.

10.3 kill-switches, audit trails, considerations with fair access.
Multi-tier kill-switches (venue, symbol, strategy) terminate the quoting/execution of tail-loss, latency or surveillance violations. Audit trails log nanosend stamps on inputs, outputs and routing results in order to answer regulatory questions and to make reproducibility possible. Equal access: no privileged access to data and throttling makes sure that the system is not overloaded in the venue or compromising on access to others when it peaks. These controls combine to achieve profitability and market integrity, to assure that the stack is responsibly and transparently behaved in calm and stressed situations.

XI. DEPLOYMENT CONSIDERATIONS

11.1 Monitoring: PnL attribution, drift detectors (PSI/KS), latency SLOs.
Continuous and layered observability is required to be production-ready. PnL attribution breaks returns down into signal, spread capture, inventory carry and fees/rebates to localize regressions in a short time. Population Stability Index, Kolmogorov Smirnov tests on key features, posteriors, and residuals are drift detectors which operate on rolling windows; violations cause alerts, shadow re-training, or automatic throttling to the GBDT fallback. Each module (ingest, features, inference, policy, risk, I/O) has its own Latency SLOs aimed at with budgets of 99p/99.9p and hard guards which drop optional work when tails begin to grow.

11.2 on symbol/venue subsets 11.2 blue-green and canary rollback criteria.
Ship evolves with blue-green interventions as well as canaries on a small representative sample of symbols/venues. Exposure caps, action clamps (max size/spread deltas), and kill-switches designed to the new version only, are known as guardrails. Establish clear rollback conditions ex ante: e.g. CVaR(95) or deltas based on reject rates above some threshold, latency SLO to be violated, or drift spikes. Rollbacks should be one- click lossless, handing over state.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

11.3 Post share trading analytics of reward/thresholds.
Close the loop on a daily basis with post-trade analytics: update the impact curves, execution shortfall models and anomaly cost matrices; re-estimate the reward weights (inventory/latency penalties) and gate thresholds on existing regime mix. Re-train scheduled upon feed and weekly governance review in order to predict calibrate predictors, RL policy, and controls against anomalies in order to ensure that the predictors, RL policy and anomaly controls adjust in a coherent manner to market conditions.

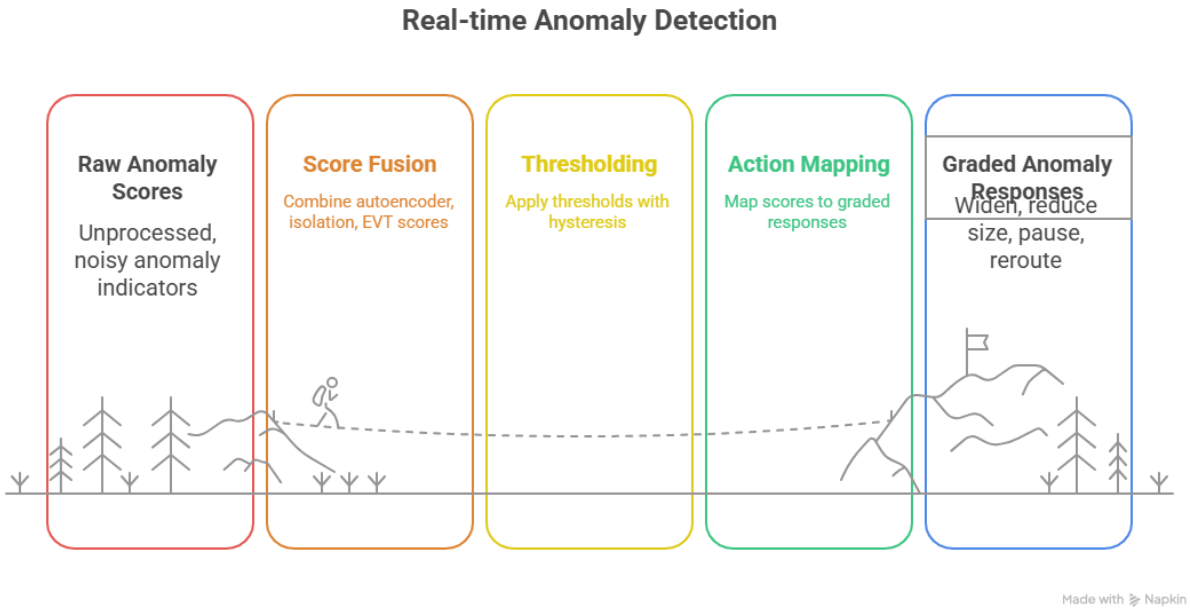


Figure4. Safety Constrained RL Policy.

XII. LIMITATIONS & FUTURE WORK

12.1 Scope, coupling between hardware and generalization.
Our experiment focuses on liquid symbols whose LOB information is rich and has co-located, low-jitter stack; our research might not be applicable to thinly traded assets, fragmented feeds, or stacks in the cloud that have variable tail latency. The sequence models and safety control are configured to a particular toolchain (DPDK, pinned cores) and will be insensitive to other NICs/CPUs or kernel configurations. Walk-forward tests minimise, although not prevent, overfitting to historical regimes; structural break and change of policy can nonetheless de facto violate assumptions.

12.2 Future: multi-venue MARL, learned simulators, cause-specific features, dynamic incentive fit.
Future directions involve multi-venue MARL that coordinates quoting and routing in common inventory / risk, learned microstructure simulators to speed up policy search and the process of stress testing. We shall consider the causal aspects (e.g. instrumented order-flow interventions) to enhance counterfactual robustness, dynamic incentive alignment which varies the weights of rewards as regime mix and capital constraint vary. Lastly, we will have wider external validation of it - more asset classes, venues and hardware - to measure portability and tighten governance limits.

XIII. CONCLUSION

This paper introduced a single, production-based HFT decision stack which ties the short-period prediction and a safe and inventory-conscious RL policy together with a real-time anomaly gate. In rolling walk-forward tests, the composite solution has shown risk-adjusted return improvements over VWAP/TWAP, fixed-spread market making, predictor-only rules, and ungated RL, and maintained end to end latency SLOs. Calibration converted forecast quality to more useful sizing and spread control, reward shaping priced inventory, impact and latency and odd-case de-risking reduced

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 8, Issue 5, September – October 2025 ||

DOI: 10.15680/IJCTECE.2025.0805005

drawdowns and CVaR in times of stress without compromising market presence. We described deployment models: blue-green roll out, canaries, drift / latency metrics and auditable governance in order to be responsible in operationalizing the stack. The next steps in the work will be multi-venue MARL, learned simulators, and causal features that will be added to make the work more robust to portability and counterfactual. To conclude, prediction, policy, and protection all combine to provide a viable frontier of latency-risk-return in high-frequency trading today.

REFERENCES

1. Abergel, F., & Jedidi, A. (2015). Long-time behavior of a Hawkes-process–based limit order book. *SIAM Journal on Financial Mathematics, 6*(1), 1026–1043.
2. Avellaneda, M., & Stoikov, S. (2008). High-frequency trading in a limit order book. *Quantitative Finance, 8*(3), 217–224.
3. Budish, E., Cramton, P., & Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market-design response. *Quarterly Journal of Economics, 130*(4), 1547–1621.
4. Cont, R., & de Larrard, A. (2013). Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics, 4*(1), 1–25.
5. Cont, R., Kukanov, A., & Stoikov, S. (2014). The price impact of order book events. *Journal of Financial Econometrics, 12*(1), 47–88.
6. Cont, R., Stoikov, S., & Talreja, R. (2010). A stochastic model for order-book dynamics. *Operations Research, 58*(3), 549–563.
7. Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 253–263.
8. Gatheral, J. (2010). No-dynamic-arbitrage and market impact. *Quantitative Finance, 10*(7), 749–759.
9. Guéant, O., & Manziuk, I. (2019). Deep reinforcement learning for market making in corporate bonds. *Applied Mathematical Finance, 26*(5), 387–452.
10. Guha, S., Mishra, N., Roy, G., & Schrijvers, O. (2016). Robust random cut forest based anomaly detection on streams. *Proceedings of ICML*, 2712–2721.
11. Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. *Proceedings of ICML*, 1321–1330.
12. He, Q., Lin, S., & Zhang, F. (2019). Quantitative analyses of LOB mid-price movements via compound Hawkes processes. *Risks, 7*(4), 110.
13. Huang, W., Lehalle, C.-A., & Rosenbaum, M. (2015). The queue-reactive model for order books. *Journal of the American Statistical Association, 110*(509), 107–122.
14. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2012). Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data, 6*(1), 3:1–3:39.
15. McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures via extreme-value methods. *Journal of Empirical Finance, 7*(3–4), 271–300.
16. Nevmyvaka, Y., Feng, Y., & Kearns, M. (2006). Reinforcement learning for optimized trade execution. *Proceedings of ICML*, 673–680.
17. Rambaldi, M., Bacry, E., & Lillo, F. (2017). The role of volume in order-book dynamics: A multivariate Hawkes analysis. *Quantitative Finance, 17*(7), 999–1020.
18. Sirignano, J., & Cont, R. (2019). Universal features of price formation: Deep-learning perspectives. *Quantitative Finance, 19*(9), 1449–1459.
19. Tao, X., Day, A., Ling, L., & Drapeau, S. (2022). Detecting spoofing strategies in HFT. *Quantitative Finance, 22*(8), 1405–1425.
20. Zhang, Z., Zohren, S., & Roberts, S. (2019). DeepLOB: Deep CNNs for limit order books. *IEEE Transactions on Signal Processing, 67*(11), 3001–3012.*