Technical Review of AI-Driven Predictive Modelling, Signal Generation, and Risk Control in Capital Markets

Shourya Gupta

University of Bath, United Kingdom

Abstract

Artificial intelligence (AI) has moved from experimental signal research to a core capability across modern capital markets, influencing return forecasting, portfolio construction, execution, risk, and surveillance. The change is driven by three forces: (1) richer and faster data (tick/limit-order-book, news, filings, alternative data), (2) scalable compute, and (3) algorithmic advances in machine learning (ML), deep learning (DL), and reinforcement learning (RL). Yet capital markets remain a uniquely difficult domain: signal-to-noise is low, non-stationarity is high, transaction costs matter, and small backtesting mistakes can create convincing but false "alpha." This paper introduces major AI approaches used in capital markets, maps them to key tasks, and provides a comparative analysis of methods, data requirements, interpretability, and deployment risks. We synthesize evidence from influential empirical finance and market microstructure research (e.g., ML in asset pricing and limit order book forecasting), multi-modal transformers combining text and time series, RL for portfolio selection under transaction costs, and practical evaluation methods to reduce overfitting risk.

1. Background and Motivation

Capital markets are information-processing systems: prices aggregate beliefs and constraints under uncertainty. Traditional quantitative finance has long used statistical models (linear factor models, ARIMA/GARCH families, Kalman filters, etc.). AI differs primarily in (a) capacity to learn non-linear interactions among predictors, (b) ability to exploit high-dimensional data, and (c) emphasis on out-of-sample predictive performance and scalable pipelines.

A key reason AI gained traction is that many canonical problems in finance can be reframed as prediction or decision tasks: expected return estimation, volatility/liquidity forecasting, cross-sectional ranking, and dynamic portfolio rebalancing. A widely cited milestone is **Empirical Asset Pricing via Machine Learning**, which compares a broad suite of ML methods (regularized linear models, trees/ensembles, neural networks) on return prediction and shows that non-linear ML can deliver economic gains in portfolio strategies in their setting.

In market microstructure, growth of electronic trading produced limit order book (LOB) data and ultra-high-frequency time series. Deep learning models have been proposed to learn patterns from the LOB directly, for example "Deep learning for limit order books" (Quantitative Finance) and more recent benchmarking that highlights reproducibility/generalization gaps when models are moved across datasets and regimes.

Table 1. Why AI methods became attractive in capital markets

Driver	**What changed**	**Practical consequence**
Data richness	Tick/LOB, alt data, text streams	More features; need representation learning
Non-linear effects	Interactions, thresholds, regime shifts	Trees/NNs can outperform linear baselines in some tasks
Compute + tooling	GPUs, distributed training, MLOps	Faster iteration, larger models
Automation	Straight-through research-to-trade pipelines	Faster deployment, higher model risk if evaluation is weak

2. Data Landscape in Capital Markets

AI performance is often dominated by data choices and labeling. Capital market datasets fall into four broad categories:

1. **Price/volume time series** (daily to millisecond), including OHLCV and derived technical indicators.
2. **Market microstructure data** such as LOB snapshots, order flow, and "market-by-order" (MBO) messages that capture the full order stream.
3. **Fundamental and macro data** (accounting variables, rates, inflation, earnings).
4. **Textual and alternative data**: news, filings, social media, transcripts, web traffic, satellite imagery, supply-chain signals, etc.

The trend is toward **multi-modal modeling** that fuses numeric time series with text features. Recent work proposes transformer-based architectures that explicitly integrate modalities (numerical time series + categorical/text inputs) for forecasting.

On the microstructure side, studies show that modeling choices (representation, normalization, labeling horizon) can dominate results. For example, deep learning on MBO data has been explored as an information source complementary to LOB snapshots.

Table 2. Common data types, typical tasks, and pitfalls

Data type	**Typical AI task**	**Label examples**	**Common pitfalls**
Daily OHLCV	return/volatility forecast; ranking	next-day return, drawdown	look-ahead via corporate actions; survivorship bias
LOB snapshots	mid-price direction classification	up/down/flat in next k seconds	leakage in label construction; brittle across venues
MBO order messages	order-flow prediction; short-horizon price move	next-tick direction	data volume, synchronization, exchange-specific microstructure
News / text	sentiment → price impact	abnormal return after news	timestamp alignment; stale news; duplication
Multi-modal (text + numeric)	joint forecasting	return, volatility, risk regime	modality imbalance; missingness; overfitting

3. Core AI Approaches Used in Capital Markets

3.1 Supervised learning (tabular + time series)

Most production quant stacks still rely heavily on supervised learning: predict a target (y) (return, volatility, direction, spread, default probability) from features (x). Methods include:

- Regularized linear models (ridge/lasso/elastic net)
- Tree ensembles (random forests, gradient boosting)
- Shallow neural networks

Evidence from empirical asset pricing suggests tree-based methods and neural networks can capture non-linear interactions and provide improvements over linear baselines in certain return-prediction settings.

3.2 Deep learning for sequences and microstructure

Deep learning is used when representation learning matters:

- CNNs for "image-like" encodings of time series or LOB states
- RNN/LSTM/GRU for sequential dependencies
- Attention/transformers for long-range dependencies

An example in high-frequency forecasting converts LOB states into images and applies CNNs; the published study reports competitive performance and provides a concrete DOI-linked implementation reference point in the *International Journal of Forecasting*.

The broader LOB-based literature also warns that strong benchmark scores can collapse out-of-sample across new periods or venues, motivating careful generalization testing and standardized evaluation frameworks.

3.3 NLP for financial text and sentiment signals

NLP is widely used to extract structured signals from:

- news headlines and articles,
- earnings-call transcripts,
- regulatory filings.

A representative study examines whether news sentiment can be traded and investigates deep learning as the modeling tool for sentiment extraction and trading relevance.

More recently, surveys focus specifically on NLP in finance and highlight typical tasks (sentiment classification, event extraction, forecasting) and open challenges such as domain shift and label quality.

3.4 Reinforcement learning for trading and portfolio decisions

RL reframes trading as sequential decision-making under uncertainty and costs. In practice, RL must handle:

- partial observability,
- non-stationarity,
- transaction costs and market impact,

- risk constraints.

Recent open-access work proposes a deep RL framework for portfolio selection with explicit transaction cost and risk awareness in the reward structure (Global Finance Journal).

3.5 Explainable AI (XAI) and model risk controls

Interpretability is not optional in many capital markets contexts (risk, compliance, regulated institutions). Explainability helps with:

- diagnosing spurious signals,
- validating stability under regime changes,
- communicating model behavior to risk committees.

A recent ACM Computing Surveys article reviews XAI in financial time-series forecasting. A separate systematic review focuses on model-agnostic XAI methods in finance and discusses limitations and challenges.

Table 3. Method families and "where they fit best"

Method family	Strength	Weakness	Best-fit market problems
Regularized linear	stable, interpretable	limited non-linearity	factor-like signals, risk models
Tree ensembles	strong on tabular; non-linear interactions	can overfit; limited sequence modeling	cross-sectional ranking, feature-rich forecasting
CNN/RNN	learns representations for sequences	tuning/instability; drift	short-horizon patterns, LOB encodings
Transformers	long-range dependencies; multi-modal	data hunger; heavy compute	fused text+price forecasts
RL (deep RL)	handles sequential actions + costs	evaluation is hard; fragile	dynamic portfolio rebalancing, execution
XAI layers	improves trust/diagnostics	may be misleading if misused	regulated deployment, model governance

4. Application Areas Across the Trade Lifecycle

4.1 Alpha research and return prediction

Return prediction is a classic "hard" problem: low predictability, noisy labels, and shifting regimes. Still, ML can be useful for:

- cross-sectional ranking (long-short deciles),
- regime-conditional signals,
- combining many weak predictors.

The empirical asset pricing literature shows that ML can exploit a large predictor set and non-linearities to construct economically meaningful strategies in their tested setting.

4.2 Portfolio construction and rebalancing

Portfolio problems introduce constraints (turnover, sector exposure, risk budgets) and multi-objective trade-offs. RL approaches explicitly model sequential rebalancing. A recent TD3-based framework embeds transaction costs and risk aversion into the reward and compares against benchmarks.

4.3 Execution and microstructure-aware trading

Execution quality depends on spread, depth, volatility, and market impact. LOB-based forecasting has become a major research area, but real-world robustness is challenging. Recent benchmarking work emphasizes performance drops on unseen data and highlights overfitting risk in popular LOB datasets. An example of LOB representation engineering is transforming high-frequency LOB data into images and applying CNNs for short-term trend prediction, with reported improvements over some baselines in that study.

4.4 Text-driven trading and event response

News and sentiment-based strategies are attractive because text can reveal information not immediately embedded in prices. However, alignment and causality are difficult. A representative study examines trading on news sentiment and deep learning-based NLP approaches.

A 2025 survey of NLP in finance synthesizes tasks and methods, including forecasting and risk-related text analytics.

4.5 Compliance, surveillance, and operational risk

AI is widely used for:

- anomaly detection (spoofing-like patterns, wash trading signatures),
- entity resolution and network analytics (beneficial ownership, collusion),
- alert triage to reduce false positives.

Even when these systems are not “alpha,” they are mission-critical and must be auditable, which increases the importance of XAI and evaluation rigor.

Table 4. Capital markets use-cases mapped to AI approach choices

Use-case	Typical horizon	Preferred methods	Notes
Cross-sectional stock ranking	days–months	boosting, NN, regularized linear	many weak predictors; stability matters
LOB mid-price direction	milliseconds–seconds	CNN/RNN/attention	generalization across venues is hard
News-driven signals	minutes–days	NLP + time-series fusion	timestamp alignment is a first-class problem
Dynamic portfolio allocation	days–weeks	RL + constraints	transaction costs must be explicit
Forecast explainability	any	SHAP/LIME, counterfactuals, attribution	avoid “comfort explainability”

5. Evaluation, Backtesting, and Comparative Analysis

5.1 Why evaluation is unusually tricky in markets

Two models can have identical prediction accuracy but very different trading performance after costs. Common failure modes include:

- **Look-ahead bias** (using future information in features/labels),
- **Survivorship bias** (ignoring delisted names),
- **Data snooping** (selecting models after many trials),
- **Improper cross-validation** for time series (leakage across time).

A well-cited framework introduces the **Probability of Backtest Overfitting (PBO)** and combinatorially symmetric cross-validation (CSCV) to quantify overfitting risk in strategy selection.

5.2 Comparative analysis: which methods win under which constraints?

Comparisons should be done on three layers:

1. **Predictive layer**: log-loss, AUC, MSE, calibration, directional accuracy.
2. **Trading layer**: Sharpe/Sortino, drawdowns, turnover, capacity, cost-adjusted returns.
3. **Operational layer**: stability, latency, interpretability, monitoring burden.

Below is a practical comparative matrix based on the strengths/limitations emphasized in the empirical asset pricing ML literature, LOB deep learning studies, multi-modal transformer work, and RL portfolio research.

Table 5. Comparative analysis of AI approaches in capital markets

Approach	Data requirement	Strength in markets	Typical weak spot	Cost sensitivity	Interpretability	Best "first use"
Regularized linear	Low–Medium	robust baselines, easy governance	misses non-linear interactions	Medium	High	factor/risk models, forecasting with few signals
Gradient boosting	Medium	strong tabular performance; handles non-linearities	regime drift; feature leakage	Medium–High	Medium	cross-sectional ranking, medium-horizon signals
Deep seq (CNN/RNN)	Medium–High	representation learning on sequences	brittle OOS; tuning complexity	High (latency/cost)	Low–Medium	microstructure forecasting prototypes
Transformers (multi-modal)	High	fuses text + numeric; long context	data hunger; heavy compute	High	Low–Medium	text+price fusion forecasts
Deep RL	High	sequential decisions with costs/constraints	evaluation fragility; instability	Very High	Low	portfolio rebalancing research sandbox
XAI overlays	N/A	governance, debugging, accountability	can be misleading if used blindly	N/A	Medium–High	risk/compliance-facing models

5.3 Practical "minimum bar" evaluation checklist

Table 6. Backtest and model validation checklist (practical)

Category	Minimum checks
Data	point-in-time data, corporate action adjustments, delisting returns, timestamp alignment
Splits	walk-forward or purged time-series CV; no leakage across horizons
Costs	explicit transaction costs + slippage; turnover constraints
Robustness	stress by regime (high vol/low vol), feature perturbations, alternative universes
Selection bias	quantify trial count; consider PBO/CSCV-style thinking
Monitoring	drift detection, performance attribution, kill-switch rules

6. Governance, Ethics, and Market Integrity

Even when a model "works," capital markets impose constraints beyond typical ML deployments:

- **Fairness and market impact**: strategies can amplify volatility or degrade liquidity if widely replicated.
- **Model risk management**: institutions require documentation, validation, and change control.
- **Explainability**: for auditability and accountability, especially for risk/surveillance.
- **Robustness and generalization**: evidence from LOB benchmarking highlights performance drops on new data, raising questions about "paper alpha" versus deployable alpha.

Two recent review streams are helpful for governance: (1) XAI surveys tailored to financial time series forecasting, and (2) systematic reviews of model-agnostic XAI methods in finance.

Table 7. Governance questions to ask before deployment

Domain	Question	Why it matters
Explainability	Can we explain top drivers and failure modes?	reduces black-box operational risk
Stability	Does performance persist across regimes/universes?	prevents "lucky backtests"
Controls	Are there limits, alerts, and human overrides?	mitigates tail risks
Compliance	Are inputs licensed/allowed and decisions auditable?	regulatory & contractual constraints

7. Future Directions

Several trajectories are likely to shape the next phase of AI in capital markets:

1. **Multi-modal foundation modeling**: richer fusion of numeric + text + alternative data (transformer-centric).
2. **More honest robustness standards**: broader benchmarking and reporting of performance under distribution shift (reinforced by LOB benchmark-style studies).
3. **Decision-focused learning**: optimizing for trading objectives directly (cost-aware losses, RL with constraints).
4. **Explainability as a requirement, not an add-on**: improved XAI taxonomies and evaluation of explanations in finance-specific settings.
5. **Backtest discipline at scale**: formal quantification of selection bias and overfitting risk as experimentation speeds up (PBO/CSCV lens).

Table 8. Research opportunities (actionable and finance-specific)

Opportunity	What to build	Main challenge
Cross-venue generalization	models that transfer across exchanges	microstructure differences
Causal event modeling	separate correlation from event-driven causality	confounding + timing
Cost-aware training	train on net-of-cost objectives	stable cost estimation
Explanation quality metrics	score explanations, not just predictions	defining “faithful” explanations
Safer model selection	integrate PBO-like controls into MLOps	operationalizing rigorous CV

Conclusion

AI in capital markets is best understood as a toolkit for (1) predicting market-relevant quantities and (2) making sequential decisions under uncertainty and costs. Supervised ML remains the workhorse, deep learning expands representational capacity for microstructure and multi-modal settings, and RL provides a principled framework for dynamic allocation and execution when carefully constrained. The main differentiator between “research alpha” and “deployable alpha” is not model sophistication, but evaluation discipline: point-in-time data hygiene, cost-aware backtesting, robustness checks, and explicit control of selection bias and overfitting risk. Surveys and benchmarks increasingly emphasize interpretability, generalization, and governance as first-order requirements rather than optional polish, suggesting that the next generation of capital markets AI will be judged as much by reliability and auditability as by headline backtest returns.

References

Arsenault, P.-D., Wang, S., & Patenaude, J.-M. (2025). A survey of explainable artificial intelligence (XAI) in financial time series forecasting. *ACM Computing Surveys, 57*(6), 1–38. https://doi.org/10.1145/3729531

Bailey, D. H., Borwein, J. M., López de Prado, M., & Zhu, Q. J. (2016). The probability of backtest overfitting. *Journal of Computational Finance, 22*(4), 1–31. https://doi.org/10.21314/JCF.2016.322

David, Z. (2019). Information leakage in financial machine learning research. *Algorithmic Finance, 8*(3-4), 61–78. https://doi.org/10.3233/AF-190900

Emami Gohari, H., Fauvel, K., Ghorbani, M., & Yu, S. (2024). Modality-aware Transformer for financial time series forecasting. *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 4413–4424. https://doi.org/10.1145/3677052.3698654

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hoang, D., Wiegratz, J., & Pukthuanthong, K. (2023). Machine learning methods in finance: Recent applications and prospects. *European Financial Management, 29*(5), 1457–1501. https://doi.org/10.1111/eufm.12408

Jiang, Y., Olmo, J., & Atwi, M. (2024). Deep reinforcement learning for portfolio selection. *Global Finance Journal, 61*, 101016. https://doi.org/10.1016/j.gfj.2024.101016

Khan, F. S., Kim, R., Lee, J., & Park, S. (2025). Model-agnostic explainable artificial intelligence methods in finance: A systematic review and research agenda. *Artificial Intelligence Review, 58*(1), 55. https://doi.org/10.1007/s10462-025-11215-9

Li, R., Chen, X., & Wang, J. (2025). NLP in finance: A comprehensive survey. *Operations Research Perspectives, 13*, 100402. https://doi.org/10.1016/j.orp.2025.100402

Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. *Applied Soft Computing, 97*(Part B), 106384. https://doi.org/10.1016/j.asoc.2020.106384

Prata, M., Masi, G., Berti, L., Donati, M., & Fantacci, C. (2024). Lob-based deep learning models for stock price trend prediction: A benchmark study. *Artificial Intelligence Review, 57*(7), 212. https://doi.org/10.1007/s10462-024-10715-4

Sirignano, J. A. (2019). Deep learning for limit order books. *Quantitative Finance, 19*(4), 549–570. https://doi.org/10.1080/14697688.2018.1546053

Vicari, M., & Gaspari, M. (2021). Analysis of news sentiments using natural language processing and deep learning. *AI & SOCIETY, 37*(4), 1437–1455. https://doi.org/10.1007/s00146-020-01111-x

Ye, W., Yang, J., & Chen, P. (2024). Short-term stock price trend prediction with imaging high frequency limit order book data. *International Journal of Forecasting, 40*(2), 539–554. https://doi.org/10.1016/j.ijforecast.2023.10.008

Zhang, Z., Zhao, Q., & Cui, J. (2021). Deep learning for market by order data. *Applied Mathematical Finance, 28*(2), 99–125. https://doi.org/10.1080/1350486X.2021.1967767