Artificial Intelligence-Assisted Decision Systems for Brokerage Operations and Capital Market Risk Management in Saudi Arabia.

Abstract:

Artificial intelligence is moving from experimental analytics to operational decision support across brokerage, trading, surveillance, and market-risk functions. In Saudi Arabia, this transition intersects with a rapidly modernizing capital market, expanding digital participation, and a policy environment that centers market development and technological capability in economic diversification. This review critically evaluates how artificial intelligence-assisted decision systems can be designed, governed, and applied to brokerage operations and capital market risk management in the Saudi context. An integrative review approach synthesizes peer-reviewed research on machine learning, deep learning, reinforcement learning, explainable artificial intelligence, asset pricing, forecasting, algorithmic trading, portfolio management, and financial risk. The evidence indicates that artificial intelligence can improve signal extraction, execution support, client segmentation, portfolio construction, volatility forecasting, and surveillance, but predictive gains do not automatically translate into robust economic or regulatory outcomes. Model instability, regime shifts, transaction costs, data leakage, limited explainability, and weak human oversight can reverse apparent benefits. Saudi-specific studies demonstrate promising applications to Tadawul forecasting and automated trading, yet the literature remains fragmented and concentrated on prediction rather than end-to-end brokerage governance. The review develops a decision architecture that combines predictive models with risk constraints, explainability, validation, human escalation, and auditability. It concludes that the strongest path for Saudi brokerage firms is not autonomous artificial intelligence, but controlled augmentation in which machine intelligence expands analytical capacity. Accountable professionals retain authority over material decisions.

Keywords: artificial intelligence; brokerage operations; capital market risk; Saudi Arabia; Tadawul; explainable AI; algorithmic trading; model risk

1. Introduction

Artificial intelligence (AI) has become a prominent component of financial decision-making because brokerage and capital market activities combine large data volumes, fast information arrival, non-linear relationships, and repeated choices under uncertainty. Reviews of AI in finance show a shift from rule-based expert systems toward machine learning, deep learning, natural-language processing, and reinforcement learning applied to forecasting, trading, portfolio selection, fraud detection, and risk analysis (Goodell et al., 2021; Cao, 2023; Bahoo et al., 2024). This evolution is especially relevant to brokerage operations, where firms must simultaneously improve execution quality, personalize client service, comply with conduct requirements, manage operational risk, and control exposures created by increasingly automated workflows.

The central problem is not whether AI can generate accurate predictions, but whether an AI-assisted decision system can convert predictions into decisions that remain economically sound, explainable, and governable. Forecasting studies repeatedly show that neural networks and tree-based methods can detect interactions that traditional linear models miss, while asset-pricing research documents improvements in out-of-sample return prediction from flexible machine-learning specifications (Gu et al., 2020; Chen et al., 2024). Yet market prediction is unusually vulnerable to non-stationarity, low signal-to-noise ratios, data snooping, and changing market microstructure. Kumbure et al. (2022), Jiang (2021), Thakkar and Chaudhari (2021), and Olorunnimbe and Viktor (2023) therefore emphasize reproducibility, backtesting design, variable selection, and realistic evaluation rather than headline accuracy alone.

Saudi Arabia provides a strategically important setting for this debate. Research using Tadawul data has explored machine-learning trading systems, deep-learning index prediction, transformer-based reinforcement learning, and the influence of oil and international markets on Saudi equity predictability (Alsulmi and Al-Shahrani, 2022; Malibari et al., 2022; Jarrah and Derbali, 2023; Abdou et al., 2024). More recent work has compared recurrent neural networks with GARCH models for Value-at-Risk and Expected Shortfall in Tadawul, finding that neural models can capture nonlinear dynamics. In contrast, econometric models can remain stronger for extreme-tail validation (Dawod et al., 2025). These results suggest complementarity rather than technological substitution.

For brokerage firms, the operational implications extend beyond directional forecasting. AI systems may support order routing, portfolio rebalancing, client suitability analysis, anomaly detection, margin monitoring, market surveillance, liquidity assessment, and scenario generation. Each use case shifts where judgment sits within the organization. A prediction engine that merely informs an analyst creates a different risk profile from a system that triggers an order, changes a client limit, or automatically escalates suspicious behavior. The relevant unit of analysis is therefore the decision system: data, model, constraints, human roles, interfaces, controls, and post-decision monitoring considered together.

This review develops that system-level perspective. It draws on recent finance and explainable-AI literature, with particular attention to evidence from the Saudi stock market. It treats AI performance as multidimensional: statistical accuracy, economic value, execution feasibility, risk sensitivity, transparency, robustness, and governance quality. This framing responds to a recurring weakness in the literature, where technically strong models are evaluated without sufficient attention to deployment conditions. The paper consequently shifts the question from "Which model predicts best?" to "Under what organizational and market conditions can AI improve brokerage decisions without weakening risk accountability?"

2. Aim of the Study

This review aims to develop a critical, evidence-based framework for understanding how AI-assisted decision systems can improve brokerage operations and capital market risk management in Saudi Arabia while preserving explainability, human accountability, operational resilience, and risk discipline. The study integrates technical evidence on predictive and optimization models with organizational considerations concerning model validation, workflow design, governance, and Saudi-market characteristics.

3. Research Objectives

The review pursues five objectives. First, it identifies the principal AI techniques and decision-support applications relevant to brokerage operations and capital market risk. Second, it evaluates the evidence on forecasting, algorithmic trading, portfolio management, and risk measurement, distinguishing predictive performance from deployable economic value. Third, it examines Saudi-specific empirical research to identify what is known about Tadawul-focused AI applications and what remains untested. Fourth, it synthesizes governance requirements concerning explainability, validation, human oversight, data controls, and model risk. Fifth, it proposes a practical architecture and research agenda for AI-assisted brokerage decision systems in Saudi Arabia.

4. Research Questions

Four questions guide the review: How can AI-assisted systems support brokerage decisions across research, execution, client service, surveillance, and risk management? Which technical and organizational factors determine whether predictive improvements produce reliable economic outcomes? What evidence from the Saudi capital market supports or limits the transfer of international AI methods to Tadawul brokerage operations? Finally, what governance design is required to make AI-enabled decisions explainable, auditable, and resilient under changing market conditions?

5. Review Methodology

A structured integrative review methodology was adopted because the topic spans finance, information systems, computer science, decision science, and risk management. The approach combines systematic search logic with critical conceptual synthesis rather than statistical meta-analysis. The attached MDPI article was used only as a structural and academic-writing reference for section sequencing and presentation; no wording, data, hypotheses, or substantive claims were transferred from it.

Searches were designed for Scopus and Web of Science coverage using combinations of terms such as “artificial intelligence,” “machine learning,” “deep learning,” “reinforcement learning,” “explainable AI,” “algorithmic trading,” “brokerage,” “portfolio management,” “stock market prediction,” “market risk,” “Value-at-Risk,” “Expected Shortfall,” “Saudi stock market,” and “Tadawul.” Peer-reviewed journal research published from 2020 through 2025 was prioritized, supplemented where necessary by influential articles that define widely used concepts or methods. Relevant studies were cross-checked through publisher pages and DOI records. The final reference set contains exactly thirty peer-reviewed articles, each with a verified DOI.

Inclusion criteria required a direct contribution to at least one of four domains: AI-enabled financial decision-making; trading, asset pricing, or portfolio decisions; risk prediction, explainability, or model governance; or empirical analysis relevant to Saudi equity markets. We excluded studies that were non-peer-reviewed, primarily promotional, focused on unrelated financial technologies, lacked sufficient methodological detail, or duplicated the contribution of a stronger source. Saudi-specific studies were retained even with narrow samples because geographic evidence remains limited and is essential for assessing contextual transferability.

Screening was qualitative and did not assign fabricated record counts. We assessed titles and abstracts for topical fit, then examined methods, evaluation design, data setting, findings, and limitations. Quality assessment emphasized peer-review status, methodological transparency, out-of-sample evaluation where applicable, attention to benchmark models, treatment of overfitting and transaction costs, interpretability, and relevance to actual decision processes. Review articles mapped broad methodological patterns, while empirical studies evaluated claims about performance and market applicability.

Synthesis proceeded through thematic coding. Evidence was organized into prediction and signal generation; execution and automated trading; portfolio and client decision support; market-risk measurement; explainability and model governance; and Saudi-market evidence. We retained contradictory results rather than averaging them away. For example, deep models may improve point forecasts while conventional GARCH variants remain more dependable for tail-risk backtesting. This distinction is important because brokerage decisions depend on loss control as well as expected return.

Methodological limitations arise from heterogeneity in datasets, horizons, performance metrics, market regimes, and implementation assumptions. Much of the AI-finance literature evaluates historical data under conditions that may not represent live execution. Publication bias can also favor successful models. Accordingly, this review does not infer a universal ranking of algorithms. It instead identifies conditions under which model classes are more or less suitable and emphasizes decision-system governance as a necessary complement to predictive performance.

Table 1. Analytical domains and review interpretation

Domain	**Typical AI methods**	**Decision contribution**	**Critical review criterion**
Forecasting and signals	Trees, neural networks, LSTM, transformers	Return, volatility, regime and factor signals	Out-of-sample stability; leakage; economic value
Trading and execution	Supervised learning, reinforcement learning	Entry/exit support, routing, rebalancing	Costs, slippage, reward design, limits
Portfolio and advice	Optimization, ranking, multi-criteria models	Allocation and client recommendation support	Suitability, diversification, explainability
Market risk	RNNs, ensembles, GARCH hybrids	VaR/ES, stress and volatility monitoring	Tail calibration, backtesting, benchmark resilience
Governance	XAI, drift monitoring, validation tools	Challenge, escalation and auditability	Traceability, accountability, human control

6. Thematic Literature Review and Critical Analysis

6.1 Prediction, signal generation, and market intelligence

The forecasting literature consistently finds that AI is valuable when relationships are nonlinear, interactions are numerous, and data include both market and non-market signals. Gu et al. (2020) show that trees and neural networks can improve risk-premium estimation by capturing nonlinear interactions, while Chen et al. (2024) demonstrate that deep architectures can embed economic restrictions within flexible asset-pricing models. Reviews by Jiang (2021), Thakkar and Chaudhari (2021), Kumbure et al. (2022), and Barboza et al. (2023) nevertheless show substantial variation in preprocessing, feature selection, training windows, evaluation metrics, and reproducibility. These differences matter operationally because brokerage decisions are made after costs, latency, market impact, and risk constraints, not at the prediction stage.

Saudi evidence broadly supports the usefulness of nonlinear methods but also demonstrates contextual dependence. Jarrah and Derbali (2023) apply multivariate deep learning to the Saudi stock index, while Abdou et al. (2024) show that the predictive influence of oil and major global equity markets on Saudi equities changes across historical regimes. Khurayshi and Alharbi (2025) compare several machine-learning algorithms for selected Saudi stocks, reinforcing the value of comparing models rather than relying on a single architecture. These studies imply that a Saudi brokerage system should treat macro-financial linkages, sector structure, and regime shifts as dynamic inputs rather than fixed relationships.

6.2 Algorithmic trading and execution support.

AI-assisted trading research ranges from supervised prediction to reinforcement learning, where the model learns actions based on reward functions. Wang and Yan (2021) document the rapid expansion of deep learning in algorithmic trading, while Wu et al. (2021) show how reinforcement learning can support portfolio management under equity-market-neutral objectives. In Saudi Arabia, Alsulmi and Al-Shahrani (2022) combine machine learning with risk-control principles for automated trading on Tadawul, and Malibari et al. (2022) integrate transformer and reinforcement-learning methods with risk-adjusted reward functions.

The critical issue is reward specification. A trading agent optimized for raw return can rationally take exposures that are unacceptable to a brokerage risk function. Malibari et al. (2022) are therefore important because they compare reward formulations linked to downside and risk-adjusted performance. Yet even risk-aware rewards do not substitute for external controls. Live systems require pre-trade limits, restricted-instrument rules, concentration thresholds, kill switches, market-impact safeguards, and independent monitoring. Reinforcement learning should consequently sit inside a constrained decision envelope rather than define the envelope itself.

6.3 Portfolio, advisory, and client decisions.

AI can assist security selection, portfolio construction, rebalancing, and individualized recommendations. Reviews of AI applications in finance identify forecasting and decision support as dominant use cases, but they also warn that outputs become more consequential when they affect customers directly (Li et al., 2023; Pattnaik et al., 2024). Multi-criteria decision research offers a useful bridge because brokerage advice rarely depends on return forecasts alone. Suitability, liquidity, investment horizon, drawdown tolerance, diversification, and product constraints must be considered jointly (Černevičienė and Kabašinskas, 2022).

This is where human-AI complementarity matters most. An AI engine can rank securities or suggest allocation changes, but a brokerage workflow should separate prediction from recommendation and recommendation from authorization. The same model output may be acceptable for an experienced institutional client but unsuitable for a retail investor with different objectives. Decision systems should therefore encode client-level constraints before model suggestions reach an execution interface. The benefit of AI is expanded analytical coverage and consistency, not removal of fiduciary or suitability judgment.

6.4 Capital market risk measurement.

Financial risk prediction is an established application of deep learning, especially where data are heterogeneous, multi-source, or imbalanced (Peng and Yan, 2021; Shi et al., 2022). For market risk, recurrent architectures can capture volatility clustering and nonlinear temporal dependence, but tail estimation remains difficult because extreme observations are sparse. Dawod et al. (2025) provide directly relevant Saudi evidence: recurrent networks improved some forecasting dimensions, whereas heavy-tailed GARCH specifications remained more reliable for extreme VaR and Expected Shortfall validation. The implication is methodological pluralism. Brokerages should combine AI-based forecasts with econometric benchmarks, stress scenarios, and conservative overrides rather than replacing established risk models wholesale.

Operational risk must also be included. AI-assisted brokerage expands model risk, data-pipeline risk, cyber exposure, vendor dependence, and the possibility of correlated automated responses. Alonso Robisco and Carbó Martínez (2022) show in credit modeling that predictive improvement should be evaluated alongside model-risk costs, a principle transferable to market-risk systems. A model that produces slightly higher accuracy but is unstable, opaque, or difficult to validate may be inferior for a regulated production process.

6.5 Explainability, validation, and responsible deployment.

Explainability is central because brokerage decisions can affect client wealth, market exposure, and regulatory obligations. Arrieta et al. (2020) frame explainability as part of responsible AI, while Bussmann et al. (2020) demonstrate the financial usefulness of Shapley-based explanations. Weber et al. (2024) find that explainable-AI research in finance is concentrated in risk, portfolios, and stock-market applications, and Chen et al. (2023) document growing attention to transparency and risk assessment. Explainability should, however, be treated as a validation tool rather than a cosmetic report. Feature attributions can help reveal unstable dependencies, leakage, or economically implausible drivers, but they do not prove causality.

For Saudi brokerages, an explainable system should provide different views to different users. Traders need concise drivers and confidence indicators; risk managers need exposure decomposition, stress behavior, and threshold breaches; model validators need lineage, training data, hyperparameters, and challenge results; senior management needs aggregated decision and exception reporting. A single explanation format cannot satisfy all these purposes. Governance quality therefore depends on role-specific transparency linked to explicit decision rights.

6.6 Data architecture, operating resilience, and Saudi localization.

A brokerage AI system is only as dependable as the data process that surrounds it. Market data, client information, corporate disclosures, macroeconomic variables, and alternative signals arrive at different frequencies and with different error structures. In live brokerage environments, stale timestamps, corporate-action adjustments, survivorship bias, missing records, or inconsistent identifiers can create operationally false and apparently sophisticated signals. The literature on stock forecasting repeatedly notes that preprocessing and feature construction influence results as much as model selection (Kumbure et al., 2022; Barboza et al., 2023). Consequently, data controls should be considered part of the model rather than a separate information-technology concern. Every production feature should have documented provenance, transformation logic, latency expectations, and a defined response when its source becomes unavailable.

Saudi localization further requires caution with alternative variables. AL-Najjar et al. (2022), for example, examine relationships between air-quality indices and the Tadawul exchange index using machine-learning techniques. The value of such work is not that any particular alternative signal should automatically enter trading models, but that AI makes it technically easy to test variables outside conventional price and accounting datasets. This capability raises a governance question: whether a detected association is economically interpretable, stable across periods, and available when a decision would actually have been made. A disciplined brokerage research process should therefore require an economic rationale, timestamp audit, stability test, and incremental-value test before approving an alternative feature for production.

Operational resilience is equally important because AI concentrates dependency on shared pipelines and services. A brokerage may deploy different models for recommendations, order routing, surveillance, and risk, yet all may depend on the same market-data feed, feature store, cloud service, or identity layer. A single failure can therefore propagate across supposedly independent applications. Model inventories should link to technology-dependency maps so risk managers can identify common points of failure. Business-continuity plans should specify degraded operating modes, such as reverting to simpler benchmark models, tightening limits, or requiring manual authorization when confidence in automated analytics falls below a defined threshold.

The distinction between confidence and authority should also be explicit in user interfaces. A model may be highly confident because a current observation resembles its training data, but confidence estimates can themselves be miscalibrated during structural breaks. Conversely, a low-confidence output can still contain useful information for an analyst. Interfaces should therefore display prediction, uncertainty, relevant drivers, constraint status, and data-quality warnings separately. Combining these elements into a single traffic-light score can conceal important trade-offs. Traders and risk managers should be able to see why an action is suggested, which limits bind, and what evidence would cause the system to abstain.

A further design principle concerns separation of duties. Teams that develop revenue-generating trading models should not be solely responsible for validating their risk behavior. Independent validation should reproduce key results, challenge feature logic, test sensitivity to regime changes, inspect performance by sector and liquidity bucket, and evaluate failure under stressed inputs. For reinforcement-learning applications, validation should examine the reward function and the learned behavior under scenarios not represented in ordinary training. An agent can satisfy a mathematical objective while exploiting unrealistic simulation assumptions. An independent challenge is therefore needed at the environment level, not only at the algorithm level.

Finally, deployment should be staged. Shadow-mode testing, in which a model produces recommendations without controlling real orders, can reveal latency, data, and workflow problems before financial exposure. Limited pilots can then impose conservative position sizes, narrow instrument universes, and stronger human approvals. Expansion should depend on realized evidence, including execution quality, risk-adjusted outcomes, override patterns, incidents, and model drift. This staged approach aligns technological experimentation with the brokerage obligation to maintain stable controls while learning from actual market conditions.

6.7 Cybersecurity, vendor, and concentration risk.

AI-enabled brokerage often relies on external model libraries, cloud infrastructure, data vendors, and specialized analytics providers. These dependencies can accelerate deployment but complicate accountability. Third-party models should meet the same validation standards as internally developed models, including documentation of intended use, performance boundaries, update procedures, and incident responsibilities. Brokerages should also assess whether several critical services depend on one vendor or technology stack, because concentration can transform a local outage into an enterprise-wide decision failure. Security testing should cover model endpoints, data interfaces, access privileges, and input manipulation. Adversarial or corrupted data need not cause a dramatic system breach to be damaging; small distortions can alter rankings, risk estimates, or order timing. Vendor governance, cyber controls, and model governance should therefore be connected through common ownership and escalation procedures rather than managed as separate compliance exercises.

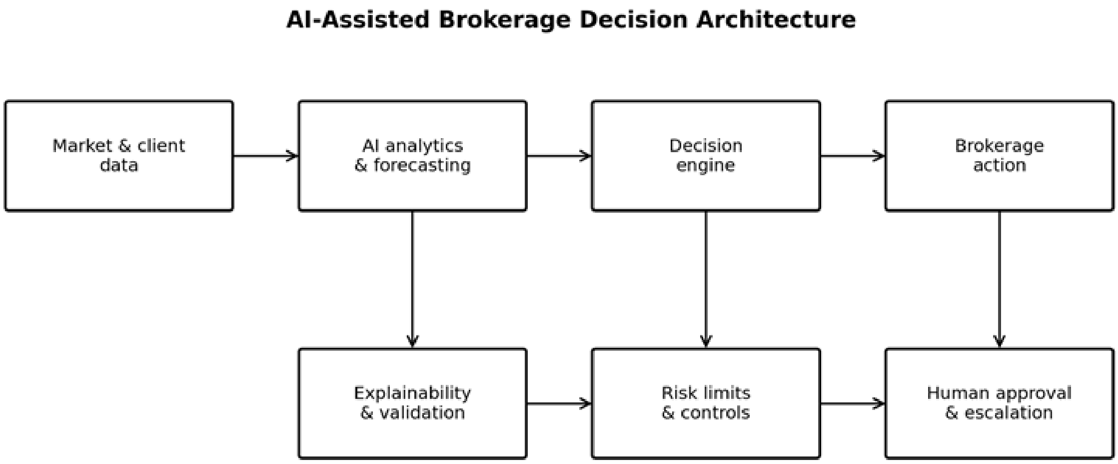

Figure 1. Proposed AI-assisted brokerage decision architecture.

7. Discussion

The literature supports a layered interpretation of AI value. At the first layer, models transform data into forecasts, classifications, scores, or recommended actions. At the second, optimization and business rules translate model outputs into candidate decisions. At the third, risk controls constrain those decisions. At the fourth, human governance determines whether an action is authorized, escalated, rejected, or reviewed after execution. This architecture resolves an important tension in the literature: technically sophisticated models can be useful without being allowed to act autonomously.

The strongest evidence concerns analytical augmentation. Broad reviews show increasing AI adoption across forecasting, trading, risk, and financial operations (Milana, 2021; Goodell et al., 2021; El Hajj and Hammoud, 2023; Pattnaik et al., 2024). Saudi studies show that Tadawul contains exploitable nonlinear structure and that models can be designed around risk-adjusted objectives (Alsulmi and Al-Shahrani, 2022; Malibari et al., 2022). However, the same evidence does not establish that fully autonomous brokerage is desirable. Most studies are retrospective, use limited instrument sets, or abstract away live market impact and organizational controls.

A second conclusion is that hybrid systems are likely to dominate. Forecast combinations, benchmark comparison, and layered validation reduce dependence on any one model class. The Saudi VaR evidence is especially instructive: recurrent networks and econometric volatility models have different strengths (Dawod et al., 2025). A brokerage risk platform can exploit this diversity by using machine learning for nonlinear pattern detection, GARCH-type models for benchmark tail estimation, scenario analysis for structural breaks, and human judgment for events outside the training distribution.

Third, treat local market structure as a first-class modeling variable. Abdou et al. (2024) show that the drivers of Saudi market predictability can change materially across regimes. A production system should therefore monitor not only prediction errors but also data drift, feature importance drift, liquidity changes, and shifts in cross-market dependence. Retraining schedules should be evidence-driven, not merely calendar-based.

Finally, governance shapes model performance. Cao (2023) emphasizes the distinctive challenges of financial data and business processes, while Weber et al. (2024) and Černevičienė and Kabašinskas (2022) show why explanation and multi-criteria reasoning matter in financial decisions. An AI system that cannot be challenged, reproduced, or audited should not be considered high-performing, even if its backtest is statistically impressive.

8. Research Gaps and Future Research

The most important gap is the scarcity of end-to-end studies on Saudi brokerage. Existing Tadawul research focuses mainly on prediction, selected-stock trading, or index-level forecasting. Future work should evaluate complete workflows linking signals to order decisions, transaction costs, market impact, compliance checks, portfolio constraints, and realized risk. Such studies should compare AI-assisted human decisions with both purely human and highly automated alternatives.

A second gap concerns high-frequency and microstructure data. Much Saudi research uses daily or aggregated market series. Brokerage operations require evidence on intraday liquidity, order-book imbalance, execution probability, slippage, and queue dynamics. Research should test whether deep learning materially improves execution after realistic latency and fees.

Third, model governance needs empirical evaluation. Explainability studies show a growing toolbox, but little evidence identifies which explanations actually improve trader judgment, validator effectiveness, or risk escalation. Saudi research could use controlled experiments to compare explanation formats across professional roles. Related work should measure model drift and define statistically defensible thresholds for retraining, fallback, or suspension.

Fourth, research should investigate Arabic-language and bilingual financial information. Natural-language processing could incorporate Arabic corporate disclosures, news, analyst commentary, and social signals, but such models require careful treatment of dialect, terminology, misinformation, and timestamp integrity. Multimodal systems combining text, prices, fundamentals, and macroeconomic variables may be particularly valuable in a market influenced by domestic reforms and global energy conditions.

Fifth, systemic effects deserve attention. If multiple brokers adopt similar AI signals or execution policies, individually rational systems may amplify crowding, liquidity withdrawal, or feedback loops. Future agent-based and network studies should examine correlated automation, stress periods, and the market-wide consequences of common model architectures. This research would connect firm-level model risk with capital-market resilience.

9. Practical, Managerial, and Policy Implications

For brokerage management, AI deployment should begin with decision inventory rather than technology procurement. Classify each use case by decision materiality, speed, reversibility, client impact, and regulatory sensitivity. Low-risk applications such as research summarization can tolerate broader experimentation; order generation, margin actions, suitability assessments, and risk-limit changes require stronger validation and authorization controls.

Managers should establish a three-line model governance structure: model owners in the business, independent risk and validation functions, and periodic audit assurance. Production approval should require documented training data, benchmark comparisons, out-of-sample testing, stress testing, explainability review, cybersecurity assessment, and rollback procedures. Table 2 summarizes controls that map to each deployment stage. Log human overrides as data because repeated overrides may reveal model weakness or poor workflow design.

For Saudi policymakers and market institutions, a principles-based approach should encourage innovation while requiring traceability and accountability for material automated decisions. Regulatory expectations can focus on model inventories, validation independence, incident reporting, data governance, and demonstrable human accountability. Sandboxed testing environments may help execution and surveillance applications where live-market externalities matter. Shared technical guidance on explainability and model-risk documentation could reduce compliance uncertainty without prescribing specific algorithms.

Brokerages should also invest in workforce design. AI changes the skills required of traders, advisers, risk managers, and compliance teams. Professionals need enough statistical and technical literacy to challenge models, while data-science teams need market-microstructure and regulatory knowledge. The organizational objective is not to turn every employee into a programmer, but to build cross-functional teams that can identify when a model is outside its competence.

Table 2. Governance controls across the AI decision lifecycle

Lifecycle stage	Primary control	Evidence retained	Escalation trigger
Use-case approval	Materiality and client-impact classification	Business rationale; decision owner	High client or market impact
Development	Data lineage and benchmark testing	Datasets; code version; benchmark results	Unstable or economically implausible drivers
Validation	Independent challenge, stress and XAI	Validation report; limitations; approvals	Failed robustness or explanation review
Deployment	Pre-trade/risk limits and human authorization	Limits; access rights; release record	Limit breach or abnormal conditions
Monitoring	Drift, incidents, overrides and outcome tracking	Monitoring dashboard; override log	Performance decay, drift or repeated overrides

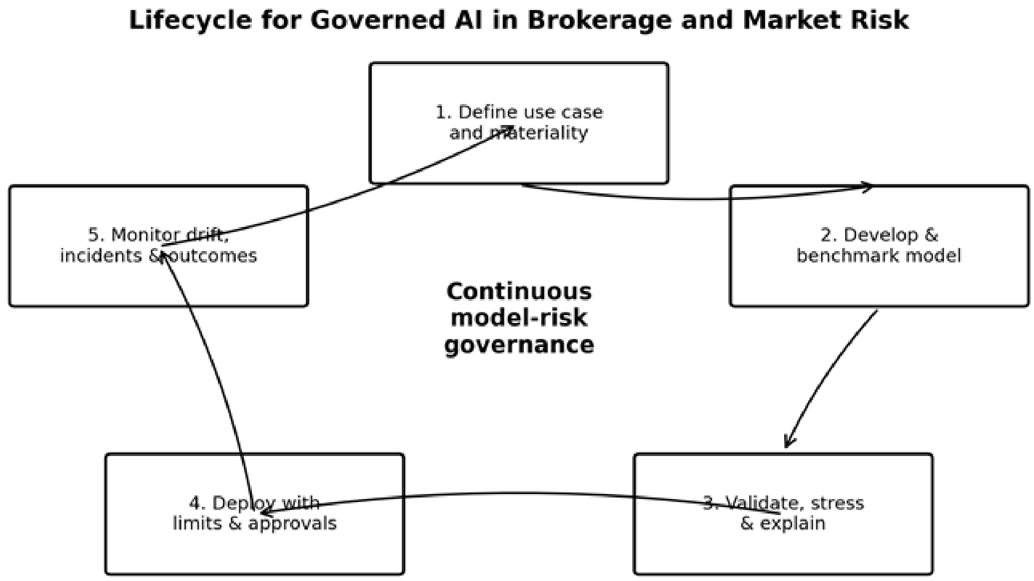

Figure 2. Continuous lifecycle for governed AI deployment in brokerage and capital market risk.

10. Limitations

This review has four limitations. First, the evidence base is heterogeneous, preventing meaningful quantitative aggregation of performance. Second, the Saudi-specific literature is still relatively small and often uses restricted samples, so conclusions about production-scale brokerage should be treated with caution. Third, rapid methodological change means that new architectures can emerge faster than peer-reviewed evaluation. Fourth, the review prioritizes journal research with verified DOIs and therefore excludes some useful regulatory papers, industry reports, technical standards, and practitioner evidence. These limitations strengthen rather than weaken the case for controlled deployment, because governance design should reflect uncertainty about external validity.

11. Conclusion

AI-assisted decision systems can materially strengthen brokerage operations and capital market risk management in Saudi Arabia when they are designed as governed decision architectures rather than stand-alone prediction engines. The literature shows credible benefits in nonlinear forecasting, signal generation, portfolio support, algorithmic trading, and risk monitoring, including promising results from Tadawul-focused studies. It also shows that performance is conditional on data quality, market regime, evaluation design, transaction costs, tail behavior, and organizational controls.

The central synthesis is that predictive sophistication and decision reliability are different objectives. A model can improve forecast accuracy yet remain unsuitable for automated action if it is fragile, opaque, poorly calibrated, or hard to constrain. The most defensible model for Saudi brokerages is therefore controlled augmentation: AI expands analytical capacity; formal rules define permissible actions; independent risk functions validate models and limits; and accountable professionals retain authority over material decisions.

This approach also fits the empirical evidence favoring hybrid methods. Machine learning can identify nonlinear structure, conventional models can provide stable benchmarks, explainability can expose questionable dependencies, and stress testing can address scenarios absent from historical training data. When these elements are integrated, AI becomes part of an auditable risk process rather than an autonomous source of risk. Future Saudi research should move beyond isolated forecasting contests toward end-to-end evaluation of execution, client outcomes, model drift, operational resilience, and market-wide effects. The strategic opportunity is substantial, but durable value will depend on disciplined integration of technology, finance, governance, and human judgment, with measurable safeguards.

References

1. Alsulmi, M., & Al-Shahrani, N. (2022). Machine Learning-Based Decision-Making for Stock Trading: Case Study for Automated Trading in Saudi Stock Exchange. Scientific Programming, 2022, 6542862. https://doi.org/10.1155/2022/6542862

2. Abdou, H. A., Elamer, A. A., Abedin, M. Z., & Ibrahim, B. A. (2024). The impact of oil and global markets on Saudi stock market predictability: A machine learning approach. Energy Economics, 132, 107416. https://doi.org/10.1016/j.eneco.2024.107416

3. Jarrah, M., &Derbali, M. (2023). Predicting Saudi Stock Market Index by Using Multivariate Time Series Based on Deep Learning. Applied Sciences, 13(14), 8356. https://doi.org/10.3390/app13148356

4. Malibari, N., Katib, I., & Mehmood, R. (2022). Smart Robotic Strategies and Advice for Stock Trading Using Deep Transformer Reinforcement Learning. Applied Sciences, 12(24), 12526. https://doi.org/10.3390/app122412526

5. Dawod, A. B. A., Nsaif, M., Honi, D. G., Abdi, I. H., &Neamah, H. A. (2025). Value-at-risk estimation in the Tadawul exchange: A comparison of recurrent neural networks and GARCH models. Systems and Soft Computing, 7, 200411. https://doi.org/10.1016/j.sasc.2025.200411

6. Khurayshi, A., & Alharbi, A. (2025). Saudi Arabia stock prices prediction using machine learning. Advances and Applications in Statistics, 92(6), 813-833. https://doi.org/10.17654/0972361725034

7. AL-Najjar, D., AL-Najjar, H., Al-Rousan, N., &Assous, H. F. (2022). Developing Machine Learning Techniques to Investigate the Impact of Air Quality Indices on Tadawul Exchange Index. Complexity, 2022, 4079524. https://doi.org/10.1155/2022/4079524

8. Goodell, J. W., Kumar, S., Lim, W. M., &Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. Journal of Behavioral and Experimental Finance, 32, 100577. https://doi.org/10.1016/j.jbef.2021.100577

9. Cao, L. (2023). AI in finance: Challenges, techniques, and opportunities. ACM Computing Surveys, 55(3), Article 64, 1-38. https://doi.org/10.1145/3502289

10. Bussmann, N., Giudici, P., Marinelli, D., &Papenbrock, J. (2020). Explainable AI in Fintech Risk Management. Frontiers in Artificial Intelligence, 3, 26. https://doi.org/10.3389/frai.2020.00026

11. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

12. Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. Expert Systems with Applications, 197, 116659. https://doi.org/10.1016/j.eswa.2022.116659

13. Jiang, W. (2021). Applications of deep learning in stock market prediction: Recent progress. Expert Systems with Applications, 184, 115537. https://doi.org/10.1016/j.eswa.2021.115537

14. Thakkar, A., & Chaudhari, K. (2021). A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions. Expert Systems with Applications, 177, 114800. https://doi.org/10.1016/j.eswa.2021.114800

15. Olorunnimbe, K., & Viktor, H. (2023). Deep learning in the stock market-a systematic survey of practice, backtesting, and applications. Artificial Intelligence Review, 56, 2057-2109. https://doi.org/10.1007/s10462-022-10226-0

16. Wang, Y., & Yan, G. (2021). Survey on the application of deep learning in algorithmic trading. Data Science in Finance and Economics, 1(4), 345-361. https://doi.org/10.3934/DSFE.2021019

17. Peng, K., & Yan, G. (2021). A survey on deep learning for financial risk prediction. Quantitative Finance and Economics, 5(4), 716-737. https://doi.org/10.3934/QFE.2021032

18. 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

19. Chen, L., Pelger, M., & Zhu, J. (2024). Deep Learning in Asset Pricing. Management Science, 70(2), 714-750. https://doi.org/10.1287/mnsc.2023.4695

20. Wu, M.-E., Syu, J.-H., Lin, J. C.-W., & Ho, J.-M. (2021). Portfolio management system in equity market neutral using reinforcement learning. Applied Intelligence, 51, 8119-8131. https://doi.org/10.1007/s10489-021-02262-0

21. Pattnaik, D., Ray, S., & Raman, R. (2024). Applications of artificial intelligence and machine learning in the financial services industry: A bibliometric review. Heliyon, 10(1), e23492. https://doi.org/10.1016/j.heliyon.2023.e23492

22. Weber, P., Carl, K. V., & Hinz, O. (2024). Applications of Explainable Artificial Intelligence in Finance-a systematic review of Finance, Information Systems, and Computer Science literature. Management Review Quarterly, 74, 867-907. https://doi.org/10.1007/s11301-023-00320-0

23. Bahoo, S., Cucculelli, M., Goga, X., &Mondolo, J. (2024). Artificial intelligence in Finance: a comprehensive review through bibliometric and content analysis. SN Business & Economics, 4, 23. https://doi.org/10.1007/s43546-023-00618-x

24. Li, X., Sigov, A., Ratkin, L., Ivanov, L. A., & Li, L. (2023). Artificial intelligence applications in finance: a survey. Journal of Management Analytics, 10(4), 676-692. https://doi.org/10.1080/23270012.2023.2244503

25. Barboza, F., Nunes Silva, G., & Augusto Fiorucci, J. (2023). A review of artificial intelligence quality in forecasting asset prices. Journal of Forecasting, 42(7), 1708-1728. https://doi.org/10.1002/for.2979

26. El Hajj, M., & Hammoud, J. (2023). Unveiling the Influence of Artificial Intelligence and Machine Learning on Financial Markets: A Comprehensive Analysis of AI Applications in Trading, Risk Management, and Financial Operations. Journal of Risk and Financial Management, 16(10), 434. https://doi.org/10.3390/jrfm16100434

27. Černevičienė, J., &Kabašinskas, A. (2022). Review of Multi-Criteria Decision-Making Methods in Finance Using Explainable Artificial Intelligence. Frontiers in Artificial Intelligence, 5, 827584. https://doi.org/10.3389/frai.2022.827584

28. Milana, C. (2021). Artificial intelligence techniques in finance and financial markets: A survey of the literature. Strategic Change, 30(3), 189-209. https://doi.org/10.1002/jsc.2403

29. Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: a systemic review. Neural Computing and Applications, 34(17), 14327-14339. https://doi.org/10.1007/s00521-022-07472-2

30. Chen, X.-Q., Ma, C.-Q., Ren, Y.-S., Lei, Y.-T., Huynh, N. Q. A., & Narayan, S. (2023). Explainable artificial intelligence in finance: a bibliometric review. Finance Research Letters, 56, 104145. https://doi.org/10.1016/j.frl.2023.104145