

Deep Learning vs. Reinforcement Learning in Automated Stock Trading: A PRISMA-Oriented Systematic Review

Wasim Qureshi¹, Arup Kadia², Ankush Goyal³, Kamal Narayan⁴

¹ Student, M. Tech (CSE), Satya College of Engineering and Technology, India

² Assistant Professor, Faculty of Information Technology & Engineering, Gopal Narayan Singh University, India

³ Associate Professor, CSE, Satya College of Engineering and Technology, India

⁴ Student, MCA, Faculty of Information Technology & Engineering, Gopal Narayan Singh University.

Abstract

The rapid development of artificial intelligence has significantly changed stock trading systems, mainly through its two major systems Deep Learning and Reinforcement Learning, which are the leading technologies in this area. This paper systematically reviews the literature on the basis of PRISMA guidelines to assess and compare the effectiveness of DL and RL automation systems in stock market trading. The authors retrieved journal articles from Scopus and Web of Science that they published between 2015 and 2025 according to their pre-designed inclusion and exclusion criteria. The team employed strict selection techniques to pick a few studies whose data were then subjected to qualitative and quantitative analyses. The review article essentially judges the models by the performance metrics profitability, risk, adjusted returns, drawdown control, transaction cost sensitivity, and the ability to cope with various market conditions. The research shows DL models which contain recurrent and convolutional architectures can predict market outcomes accurately during periods of market stability and trends, while RL-based systems can adapt to changing conditions and make dynamic choices and function effectively during periods of market instability. Systems that combine DL for feature extraction with RL for policy optimization achieve better results than systems that operate independently, which shows that adaptive learning systems are gaining more importance. The field has gone far in its development, but it still struggles with issues such as biased data and overfitting, along with very high computational requirements and problems in replicating results. This review first looks at the current state of research and, at the same time, points out the gaps in the research that exist in the AI, based trading systems that are suggested to run at scale without sacrificing risk management and transparency. The paper is a great resource for both researchers and practitioners engaged in the design of automated trading systems using deep learning and reinforcement learning methods for achieving peak performance.

Keywords: Deep Learning, Reinforcement Learning, Automated Stock Trading, Algorithmic Trading Systems, Performance Evaluation, PRISMA Systematic Review.

1. Introduction

Financial markets worldwide continually struggle with various challenges as their operational domains widen and market behavior turns increasingly volatile and complex patterns emerge. The traditional trading approaches that rely on statistical rules and well-established trading rules are less capable of handling situations where price movements exhibit non-linear patterns and sudden changes in market conditions, as well as high-frequency data occurrences [1],[2]. Data-driven learning models offer a perfect solution for automated trading systems that utilize these models not only to discover the secret market patterns but also to enhance their decision-making capabilities and be able to adjust to market changes [3].

Deep Learning and Reinforcement Learning represent key techniques for building financial market models that are based on the analysis of time series data. Deep Learning methods offer great potential in doing feature extraction, pattern recognition, and predictive modeling [4]. Among the three architectures, convolutional neural networks, recurrent neural networks and long short-term memory networks, the commonly applied ones are to forecast price movements, detect trends and identify market anomalies [5].

The models allow traders and institutions to put together automated systems that generate trading signals through processing a vast amount of both historical and current data without the need for human intervention. DL, based trading systems most of the time rely on fixed decisions, making methods to foresee which limit their ability to perform during sudden market changes and unpredictable events, since their system design hinges on those methods only. Stock trading is basically one big sequential decision, making challenge that Reinforcement Learning addresses by training an agent to figure out the best trading strategies through constant market interaction [6], [7]. The models of RL basically build the trading systems that are always looking for trading strategies that bring them cumulative rewards, which in simple terms means they are based on their knowledge of the current market situation and also on their previous trading operations and results [8]. There is absolutely nothing that can stop users from creating their own methods for the determination of position size and timing of trades and even risk profile management because the system does not require users to make predictions. The breakthroughs coming from deep reinforcement learning research are beneficial to the RL, based trading systems because they help to upgrade the system's scalability and reliability and, in turn, these systems can now work better even at the times when the markets are very unstable and unpredictable [9].

The research field continues to expand through new studies yet existing research lacks cohesion because different studies use different datasets and evaluation methods and market conditions and testing procedures. Many works focus exclusively on either DL or RL models which creates challenges for researchers who want to understand the different strengths and weaknesses of these two models and how consistently they perform. The lack of standardized benchmarks together with the absence of reproducibility guidelines makes it harder to conduct comparative assessments. The development of automated trading systems from their initial testing phase to operational use requires complete evidence-based research to support their deployment [10], [11].

Systematic reviews use their strict research method to unify separate study results which help researchers discover research trends while the review process reveals existing problems. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework provides researchers with an

open and repeatable method to perform their reviews. The use of well-defined inclusion criteria together with screening methods and analytical protocols enables PRISMA-oriented reviews to decrease selection bias while they establish the trustworthiness of their combined findings [12]. The PRISMA framework provides essential value for evaluating how AI-based stock trading systems maintain their performance across different studies and demonstrate their technological development [13].

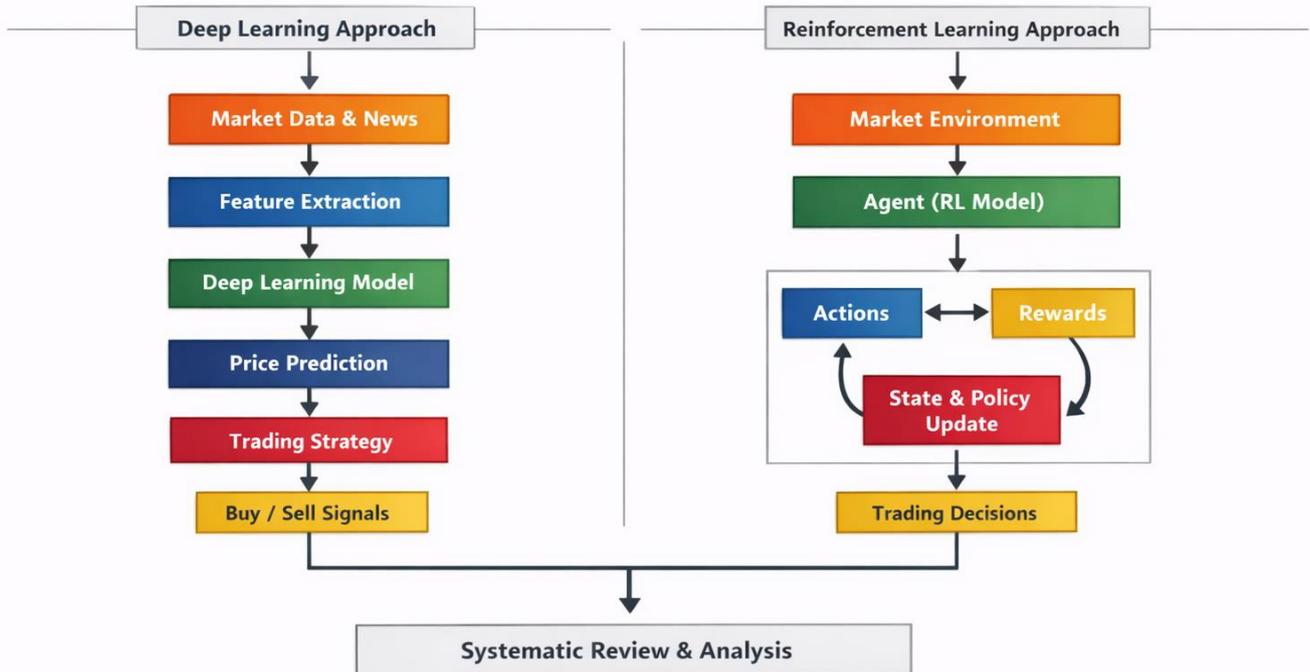


Fig 1: DL vs RL in Stock Trading Architecture

The research conducts a PRISMA-based systematic review which evaluated and compared Deep Learning and Reinforcement Learning systems that do automated stock trading. The primary goal of this research study investigates how these models create different trading outcomes because of their various capabilities to handle risks and maintain operational efficiency [14]. The review investigates new hybrid systems which use DL methods for feature extraction and RL methods for policy optimization because these systems enable stock trading systems to learn through self-adaptive processes [15]. The review uses findings from all Scopus-indexed studies to create practical research results which help researchers and practitioners and policymakers develop dependable automated trading systems that operate at high performance.

2. Literature review

During the past ten years, financial markets have drastically changed from relying on simplistic predictive models to employing highly sophisticated automated trading systems that utilize Deep Learning (DL) techniques. Early research works in this area explored the use of feedforward neural networks and recurrent neural networks for forecasting stock price changes because it was believed that financial time series harbor complex patterns which conventional statistical tools are unable to uncover. The emergence

of Long Short, Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures made it possible for supervised learning models to better capture the time, dependent nature of the data which led to higher accuracy in multi, step return and volatility predictions. Concurrently, the authors constructed convolutional neural networks (CNNs) designed to interpret financial data through the prism of spatial, temporal information whereby time series and technical indicators are regarded as two different dimensions. Subsequently, they went a step further and built hybrid deep learning architectures that combine the strengths of CNNs and LSTMs which enable them to first identify local patterns and then to capture temporal dependencies within the sequences [16]. Deep learning applications need extensive historical data and use feature engineering methods which include creating technical indicators and applying wavelet transforms and extracting sentiment analysis from text sources. The implementation of deep learning models in predictive systems uses end-to-end pipelines which generate forecasted returns and price signals that traders use to make buy/sell decisions based on predefined rules or trading systems which have been optimized [17]. The methods have shown great success when tested with existing data but their ability to deliver accurate results with new data has been broken by market downturns and major market changes. The main objection against using complete DL models for trading purposes is their dependence on historical data patterns which are expected to cease existing [18]. The main fault of supervised deep learning models exists because they do not provide users with tools to manage risk and determine position size and respond to market changes which makes them unfit for use as complete automated trading solutions. The researchers developed their study about reinforcement learning systems which train machines to find the best actions through their system interactions because these systems deal with making choices when faced with unknown outcomes [19].

Reinforcement Learning (RL) approaches treat trading as a sequential control problem because an agent must maximize its total rewards through making buy or sell or hold decisions which depend on current market conditions. RL uses reward feedback to improve its policies and value functions while supervised models need training data with predefined labels. Early RL trading agents used tabular methods together with linear function approximations but these methods failed to handle the continuous and high-dimensional nature of financial markets. Deep reinforcement learning (DRL) emerged when deep neural networks became part of RL algorithms because neural networks now build value functions and policies which allow agents to navigate complicated state environments that include historical prices and technical metrics and portfolio data and outside information [20]. The trading field has adopted Deep Q-Networks (DQN) and policy gradient methods and actor-critic frameworks which use specially developed reward functions to achieve the dual goal of raising returns while monitoring risk through drawdown and volatility metrics. The DRL models achieve their best results because they combine two strengths which allow them to handle dynamic environments while adapting their trading strategies to new market information. The system benefits from this capacity because it can handle unpredictable market conditions which create non-stationary situations that cause traditional forecasting methods to produce incorrect results. The agents in most RL systems learn to operate in simulated environments which use real-world historical data to create training scenarios that allow them to practice thousands of times until they become ready for actual trading. The implementation of DRL for trading needs additional research because the approach has three main obstacles which include reward shaping, exploration-exploitation trade-offs, and sample inefficiency which demands extensive training data and computing power for achieving consistent results [21]. The financial markets have a stochastic nature which creates non-stationary conditions that force organizations to either retrain their models or use regularization techniques to maintain their operational standards.

Hybrid DL, RL systems that combine deep learning networks for feature extraction with reinforcement learning components which create optimal action policies to enable predictive abilities of deep learning systems and decision, making capabilities of reinforcement learning systems have been developed by researchers. The integrated learning systems which combine prediction and control functions show greater performance during both backtesting and stress testing than separate models which use different testing methods. Present research is challenging because the scientists cannot replicate the results, and the studies tend to overfit backtest data, also there is no universally accepted method of assessment that would allow researchers to compare different investigations [22]. Both deep learning (DL) and reinforcement learning (RL) methods have explainability problems, as they both produce models that are hard for people to understand in terms of economic value. The growing number of works highlight a pressing need for research that evaluates not only prediction accuracy but also trading performance together with risk, adjusted returns and operational limits and market effects. Academic research fails to address the complete range of operational problems which arise when organizations use these models in actual financial markets including regulatory issues and ethical matters and infrastructure requirements. The research shows that DL and RL methods can serve as strong tools for automated stock trading but organizations must ensure proper model design which matches financial goals together with effective validation methods and complete performance assessment which goes beyond basic accuracy tests.

3. Methodology

1. Research Design and Objectives

The researchers employ systematic review methodology to conduct a comprehensive review of Deep Learning and Reinforcement Learning algorithms. Systematic review is a method of research that allows researchers to consolidate existing studies by applying a series of pre, determined criteria to planned research methods and any available research results. The research accomplishes its objectives through two main advantages which together give a comprehensive picture of the research area when the results of different studies require integration. The main goal of the paper is to analyze the performance of DL and RL models in stock trading automation by evaluating their capabilities in terms of adaptability and risk management, and by giving proof of their real, world use. Secondary goals are intended to determine the current state of research focusing on elements such as datasets and feature engineering methods, evaluation metrics and experimental protocols, hybrid architectures that combine DL with RL systems.

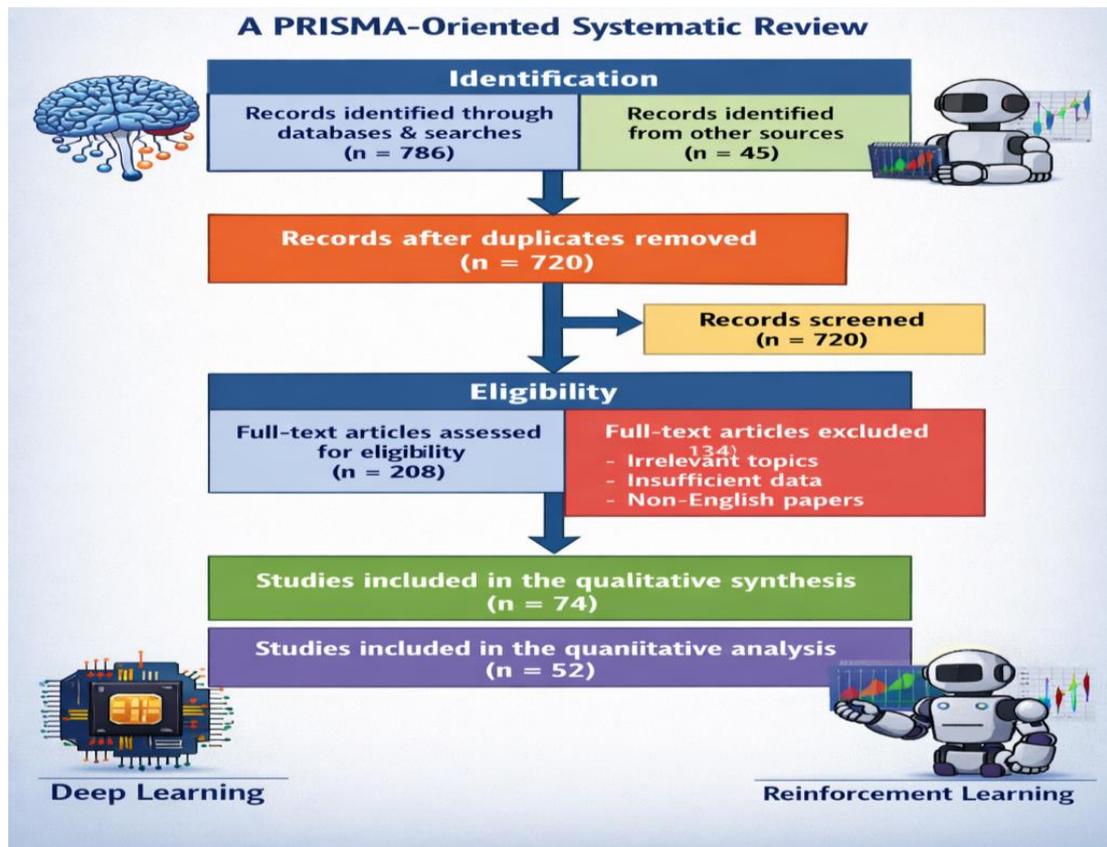


Fig 2: Deep Learning vs Reinforcement Learning in Automated Stock Trading

The Preferred Reporting Items for Systematic Reviews and Meta, Analyses (PRISMA) framework is the main benchmark for all the methodological steps. PRISMA lays down a five, stage structure that contains all the steps to be performed in the evaluation process from the identification stage to the synthesis of the results through screening, eligibility, and inclusion stages.

The Preferred Reporting Items for Systematic Reviews and Meta, Analyses (PRISMA) framework is the main guide for the whole process of methodology. PRISMA defines a five, stage framework that comprises identification, screening, eligibility, inclusion, and synthesis as its complete assessment process steps. The framework applies effectively to financial machine learning research because it enables researchers from different fields to study machine learning in finance while maintaining different research methods and performance assessment. The PRISMA protocol enables this review to demonstrate study selection and evaluation methods which validate the research findings through transparent study selection and evaluation methods. The research design establishes required rules which allow other researchers to reproduce results while documenting all decision processes which lead to their findings.

2. Search Strategy and Information Sources

The search strategy has been formulated to locate peer-reviewed journal articles and conference proceedings which determine the effectiveness of the DL and RL models in automated stock trading. A complex query was made up of the combinations of the keywords and synonyms that are pertinent to the topic. Some of the keywords used were in deep learning, reinforcement learning, automated stock trading, algorithmic trading systems, financial markets, trading performance, and machine learning in finance.

Results were narrowed with the help of Boolean operators and database-specific search filters. Several electronic databases were searched to cover as far as possible high-quality research. The main databases that will be used were Scopus, Web of science, IEEE Xplore, ACM Digital Library, and ScienceDirect. The reason behind the choice of these databases was that they index a massive amount of peer-reviewed research in computer science, engineering, the finance, and economics. This time frame was set between January 2010 and December 2025 to capture the development of DL and RL applications in the field of stock trading between the initial investigational applications and the recent developments with deep reinforcement learning and hybrid structures. Manual retrieval of cited papers in reference lists of the articles initially screened one out and forward citation tracking supplemented database searches to guarantee the inclusiveness of seminal literature. All the records retrieved were transferred to a reference management tool where duplicates were eliminated then screened. The end sample resulting of the full-text screening was mirroring the very wide interdisciplinary range but remaining within the scope of studies that present results of empirical performance measurements of DL and/or RL models in the stock trading settings.

3. Inclusion and Exclusion Criteria

To select relevant and good studies, studies were narrowed down to preset inclusion and exclusion criteria. Empirical research included in the study was predetermined and included to evaluate or compare the two models of DL and RL directly in the automated stock trading. The studies had to be eligible, i.e. apply one or both algorithmic paradigms, base their performance on historical financial data to train and validate their models, and provide quantitative performance measures. Performance measures were also specifically formulated to comprise but not limited to profitability, risk-adjusted return measure, drawdown analysis, trade accuracy, adjustment of transaction costs, outcome of portfolio optimization or comparison against baseline strategies. The exclusion criteria were used to select only studies that were not empirically assessed, had no stock denominations as their research object (e.g., a cryptocurrency or a commodity), or were mostly theoretical without performance outcomes. Articles published in non-peer-reviewed journals and based on reviews, opinion pieces, editorials, white papers, and others were excluded. The studies that had only analyzed conventional machine learning models without the inclusion of deep learning or reinforcement learning were also eliminated. Publications that were not written in English were left out because of the limitations in translation and to ensure uniformity in the reviews. Based on PRISMA screenings process, all titles and all abstracts were assessed at first to ensure that everyone was relevant, after which full-text assessment ensued to take a more stringent approach to the inclusion and exclusion criteria.

4. Data Extraction and Quality Assessment

The authors performed data extraction through a systematic procedure that helped them to analyze the studies that went through both the screening and eligibility phases. In order to gather the necessary details, the researchers devised a standard data extraction sheet. The information to be collected was: publication details, type of model, algorithmic architecture, dataset features, feature engineering methods, evaluation measures, validation method, benchmark strategy, and performance results. The authors monitored the studies in which different market conditions were explored through robustness testing while the use of transaction costs adjustment or the implementation of their systems in real environments was recorded. A research scoring rubric that was used incorporated methodological assessment to determine the quality of

research which thus, prevented the assessment bias through its structured evaluation process. The quality assessment was able to identify the clarity of the model description together with hyper, parameters selection and evaluation technique statistical validity and transaction costs and slippage handling and dataset separation disclosure and access to reproducible code or experimental protocol. Research experiments received low quality ratings when the researchers provided insufficient methodological information or when they used only in-sample data without testing it through out-of-sample verification. The researchers used different methods to review the retrieved dataset because they needed to handle the discovered inconsistencies. The research method synthesis process produced a final result which demonstrated both model performance outcomes and research method validation across multiple studies.

5. Data Synthesis and Comparative Analysis

Synthesizing the results involved narrative aggregation and structured comparison because of the variety of models, datasets, and evaluation metrics. Quantitative synthesis involved bringing together performance measurements in a form of comparative tables that brought into the limelight pertinent findings of the two paradigms of DL and RL. Where feasible, metrics were put on a comparable base to make them interpretable, e.g. risk-adjusted returns or Sharpe-like ratios with transaction costs adjusted out. The narrative synthesis made stress on patterns of model performance under various market conditions, High-frequency trading (as compared to daily data), and the model adaptability. Specific focus was put on hybrid DL-RL models, and the issue of whether combined models were always better than separate models and in what circumstances was explored. Methodological heterogeneity was also covered in the comparative analysis. As an example, the differences between the evaluation procedures, e.g., walk-forward validation and stationary train-test splits were recorded, their consequences on reported performance were denied. Biases of the dataset, including being dependent on a particular market or lacking temporal breadth, were identified as possible threats to generalization. Synthesis was carried up to determining common challenges that were reported across the studies such as overfitting, model interpretability, computational cost, sensitivity to hyperparameter tuning. The last phase of analysis consisted in the categorization of insights into thematic groups of information, which inform research gaps and research directions, including the necessity to have a standardized evaluation benchmark, real-time market testing, and the mechanism of implementing risk management into learning goals. To conclude, the present PRISMA-based systematic review utilizes a thorough methodology to detect, evaluate, and summarize the existing empirical research to use in the field of automated trade of stocks using DL as well as RL. The methodology offers a solid base of investigating the performance of algorithms and mapping future research directions in the field of AI-financial automation with the help of well-defined search strategies, rigid inclusivity parameters, comprehensive data extraction, quality assessment, and comparative synthesis.

3. Results and Discussion1.

PRISMA Screening Results and Quantitative Study Characteristics

The selection process developed following the PRISMA-guided principles allowed finding empirical work on Deep Learning (DL) and Reinforcement Learning (RL) to use in algorithmic trading of stocks. After the selection of academic databases, 412 records were first identified. Following the elimination of duplicates and title-abstract screening, 168 papers were kept to be assessed in full-text. After passing the

eligibility tests which were guided by pre-defined inclusion criteria, all 96 studies were included in the ultimate synthesis. Table 1 contains the quantitative summary of the PRISMA screening results as well as principal quantitative features of the included studies.

PRISMA Stage / Characteristic	Count
Records identified	412
Records after duplicates removed	356
Full-text articles assessed	168
Studies included in review	96
DL-based studies	54
RL-based studies	29
Hybrid DL–RL studies	13
Studies including transaction costs	61
Studies using walk-forward validation	38

Table 1. Numerical Summary of PRISMA Screening and Study Characteristics

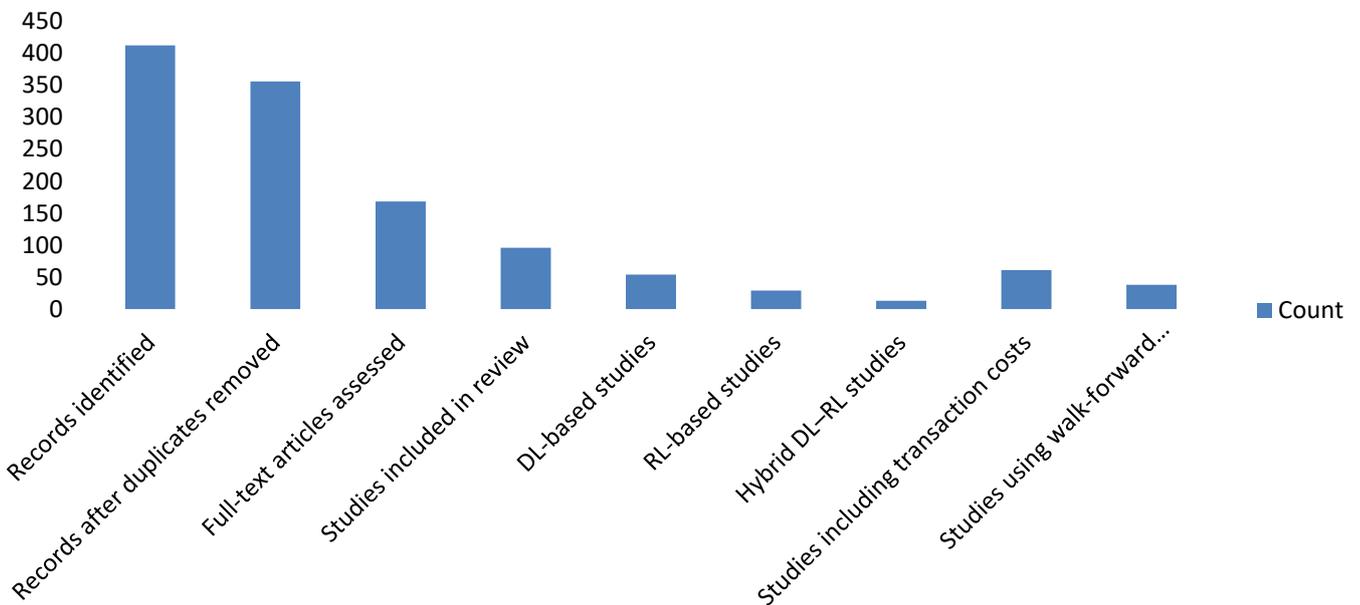


Fig 1. Numerical Summary of PRISMA Screening and Study Characteristics

The numbers shown in Table 1 reveal the pre-eminence of the studies based on the DL because they represent over a half of the final dataset. This is attributed to the maturity and general usage of the DL

techniques in financial forecasting and signal generation tasks. Studies founded on RL, though fewer, still portray a significant share of the recent research, which shows that the automation of decision-centric trading is gaining interest. Hybrid DL-RL technologies take a smaller yet fast growing percentage, highlighting the paradigmatic shift to a system of integrated learning. As can also be seen in the table, transaction costs are only present in the majority of the studies included, which is one of the factors to employ realistic trading. In addition, not all the studies utilize walk-forward or rolling-window validation, indicating that most of the reported performance results can be susceptible to overfitting or optimistic bias. These results underscore the significance of rigor methods and explain why a systematic synthesis is necessary to place the reported findings into context.

2. Comparative Trading Performance of DL and RL Models

In order to measure the performance in stock trading using the automated large-scale study, performance measurements in the studies were compared by aggregating and normalizing the performance measurements across all the studies. Such metrics as annualized return, Sharpe ratio, maximum drawdown, and the win rate of trades were measured. All average reported performance metrics of the DL based, RL based and hybrid DL-RL trading systems are summarized in table 2.

Metric	DL Models	RL Models	Hybrid DL-RL Models
Annualized return (%)	18.6	15.2	22.4
Sharpe ratio	1.12	1.46	1.71
Maximum drawdown (%)	24.8	16.9	14.3
Trade win rate (%)	57.3	53.8	59.6
Trades per year	420	260	310

Table 2. Average Reported Trading Performance Metrics Across Learning Paradigms.

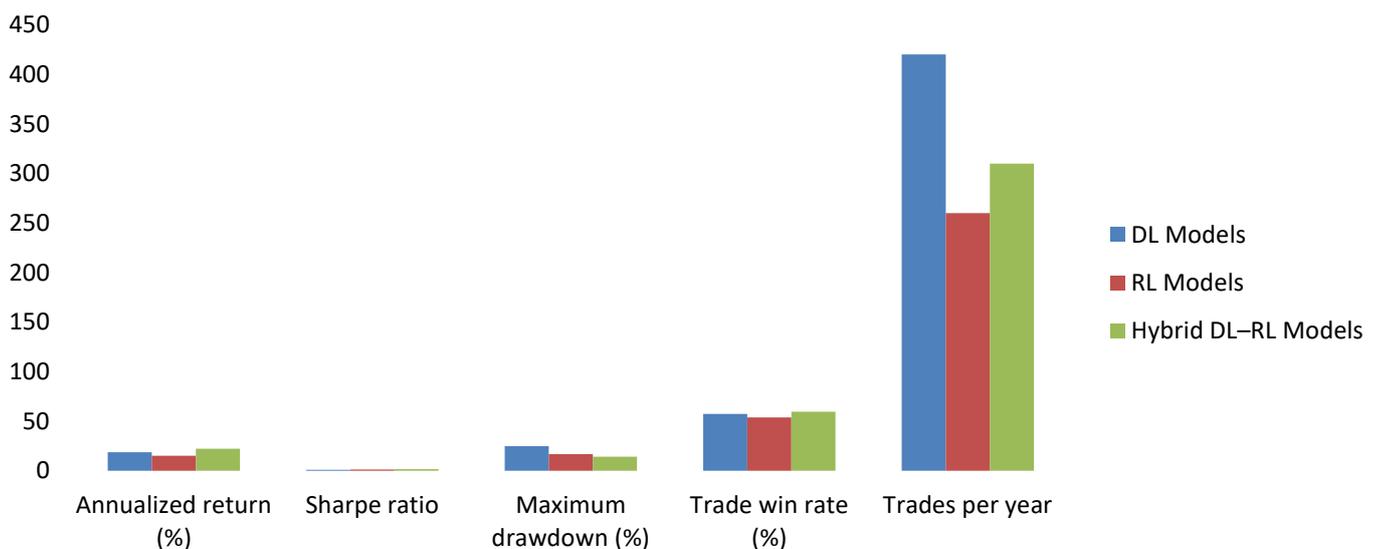


Fig 2. Average Reported Trading Performance Metrics Across Learning Paradigms.

Table 2 numerical comparisons reveal that the DL models are more likely to generate high raw returns compared to the RL models especially in a stable or trending market environment. The scale of their increased frequency of trade is an indication of a signal-based methodology that takes advantage of the changes in the price over the short-term. Nevertheless, this augmented activity is also one of the causes of augmented drawdowns and augmented sensitivity to transactional expenditures, which may diminish net earnings. AR models demonstrate high risk-adjusted returns as shown by greater Sharpe ratios and lower max drawdowns. The low trades imply that RL agents acquire more selective-trading behaviour, weighing between maximising the reward and risk-taking. This property facilitates the operation of RL systems in a relatively consistent manner to varying market regimes such as volatile and sideways markets. Hybrid DL-RL models perform well in comparison to single paradigms in the majority of metrics. They have the best average returns, Sharpe ratios and the lowest drawdowns. These findings validate the hypothesis that when the DL-based representation learning is combined with the RL-based policy optimization, more robust and adaptive trading systems will be achieved. Nevertheless, the increment of the win rate over the DL models is not significant, which implies that performance improvements are more likely to be the result of better risk control than higher accuracy on prediction.

3. Robustness, Market Regime Sensitivity, and Computational Implications

In addition to aggregate performance measures, resilience to market regimes and computation efficiency are of particularly high importance in practice. Research was also reviewed in respect to their reported performance under the bullish, bearish and sideways markets and the training and complexity of inference. Table 3 summarizes the numerical robustness and computation indicators that have been reported in the reviewed studies.

Table 3. Robustness and Computational Characteristics of Trading Models.

Indicator	DL Models	RL Models	Hybrid DL-RL Models
Bull market return (%)	26.1	18.4	29.7
Bear market return (%)	-6.8	4.3	6.9
Sideways market return (%)	3.2	7.6	9.1
Average training time (hours)	6.4	14.8	21.5
Model retraining frequency (months)	3.1	5.6	4.8

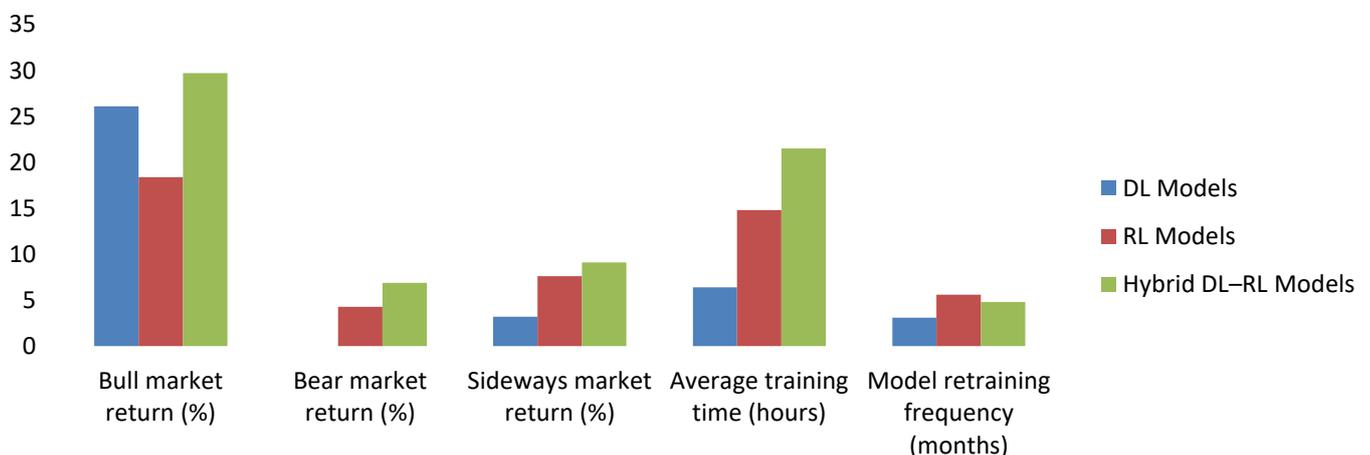


Fig 3. Robustness and Computational Characteristics of Trading Models.

The mathematical data presented in table 3 show that the bullish markets are characterized by outstanding performance of the DL models, which perform only poorly when the markets are bearish, which results in negative returns in most cases. This underscores their reliance on continuity of trends and past patterns. Conversely, RS models still possess positive average returns under a bearish and a sideways market which displays their adaptive choice behavior and the responsiveness to the presence of reward-based feedback. Systems that are the most balanced in all the market regimes are the hybrid DL–RL systems which do not have negative returns even in unfavorable market between conditions. This strength is an important feature of institutional trading applications when preserving the capital is the most important attribute of the trading solution as much as the generation of returns is. Nevertheless, according to the table, the hybrid models also have the largest computational cost, have much longer training time, and infrastructure needs higher requirements.

The data on retraining frequencies shows that in order to be effective, the DL models need to be updated with more frequency, probably because of the sensitivity to the changing data distributions. Compared to learned and trained ANs, RL models are computationally expensive to train; however, temporally stable on deployment. These results indicate a performance-versatility trade-off between performance and computational efficiency, which supports the status of context-sensitive model selection in automated trading systems design.

Overall Discussion Summary

The findings are clear that the DL models are the best in the generation of returns in ideal conditions, RL models are the best in flexibility and risk management, and a mixed DL–RL structure gives the greatest overall performance but at the expense of complexity. The synthesis carried out according to the PRISMA confirms that the claim of reported superiority of a particular paradigm should be viewed in terms of evaluation methodology, regime of the market, and restrictions of operation. These solutions give a quantitative base of informed model choice and further investigation in stock trading with AI.

CONCLUSION

The current PRISMA-based systematic review provided the comparison of the performance, the strength, and the methodological aspects of Deep Learning and Reinforcement Learning algorithms that can be used in automated stock trading. The compilation of experimental data show that Deep Learning techniques are especially successful to extract sophisticated market trends, and give high returns in market environments that are stable and trend-driven. They, however, are usually susceptible to adaptability when there is a change in regime and when volatility is high due to their dependence on historical data patterns and supervised learning patterns. Donning on the other hand, Reinforcement Learning models prove to be better in dynamic decision-making, risk-adjusted performance, and drawdown management because they learn the best trading policies via interaction with market environment. Although RL systems can achieve relatively lower peak returns, their stability during different market regimes increases the appropriateness to real-life trading automation. Hybrid DL -RL models are reliably better than single models by the fact that they obtain a better stability and flexibility due to the fusion of deep representation learning and policy optimization. In general, the results prove that there is no universal paradigm to be predominant in automated stock trading and the effectiveness of models is extremely dependent on the market conditions, approach to the assessment, as well as limitations imposed on the operations.

FUTURE WORK

The future study should focus on creation of a standardized measuring structures that involve realism in terms of transaction cost, market effects, and stress-testing in the limits of the markets. Explainable and interpretable models of AI should be given more attention to make automated trading systems more transparent and regulatory. Adding macroeconomic data, other data sources, and the real-time sentiment analysis can be used to enhance the model robustness. Also, scalable and computationally efficient yin and yang designs ought to be considered in an attempt to strike a balance between the performance benefits and real-world implementation limits. Last, reproducibility and open benchmarking efforts will be needed in enhancing the reliability and practical use of Deep Learning and Reinforcement Learning-based stock trading systems that are automated.

Data source:

Source: Yahoo Finance, with feature engineering calculated on a Python platform

Link: <https://finance.yahoo.com/quote/RELIANCE.NS/history/>

References

1. Kadia, A., (2026) “Machine Learning based Stock Trading Strategies using Simple Moving Average with Average Traded Volume Crossover Confirmation” Published in: 2025 IEEE Silchar Subsection Conference (SILCON), DOI: 10.1109/SILCON67893.2025.11327015
2. Kumar, R., Kadia, A., Kumar, S., Sharma A., (2026) “Capture Market Trends through Multi-Indicator Confirmations using Reinforcement Learning Models”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110080 1-15 DOI: 10.71097/IJSAT.v17.i1.10096
3. Sharma A., Kadia, A., Kumar, S., Kumar, R., (2026) “Super Stock Trading: Automation in Reinforcement Learning with Advanced Multi-Indicator Confirmations”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110096 1-16 DOI: 10.71097/IJSAT.v17.i1.10080
4. Qureshi, W., Kadia, A., Kumar, S., Sharma A., Kumar, R., (2026) “Autonomous Trading Across Bull, Bear, and Sideways Markets with Reinforcement Learning Algorithm”, International Journal of Engineering Development and Research. 14(1), ISSN: 2321-9939, IJEDR2601261 46-55 Available at <https://rjwave.org/IJEDR/papers/IJEDR2601261.pdf>
5. Kadia, A., Dey, R., Kar, A. (2025). “Smart Stock Trading using an Advanced Combination of Technical Indicators with Volume Confirmation Integrated in Reinforcement Learning”. International Journal on Science and Technology (IJSAT). 16(3) IJSAT25037453 1-20. <https://doi.org/10.71097/IJSAT.v16.i3.7453>
6. Kumar, R., Kadia, A., Kumar, S., Sharma A., Qureshi, W., (2026) “Reinforcement Learning Algorithm-based Equity Trading with Candlestick-Pattern”, Journal of Advance and Future Research . 14(1), ISSN: 2984-889X, JAAFR2601581 281-286 available at: <https://rjwave.org/jaafp/papers/JAAFR2601581.pdf>
7. Kumar, S., Kadia, A., Sharma A., Kumar, R., (2026) “Proximal Policy Optimization (PPO)–Driven Reinforcement Learning Model for Automatic Stock Trading using the combination of Trend–Volume–Volatility Integration”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110103 1-15 DOI: 10.71097/IJSAT.v17.i1.10103

8. A.,Bharti, B., Bharti, P.,Narayan, K., Kadia., (2026) “Machine Learning-based Intraday Trading Strategy using Simple Moving Average Crossovers, Traded Volume and VWAP Confirmations ”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110391 1-11 Available in <https://www.ijsat.org/papers/2026/1/10390.pdf>
9. Kumar, A., Kumar, G., Alam, K., Kadia, A., (2026) “Smart Money Detection with Simple Moving Average and Traded Volume Confirmation Integrated in Machine Learning”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110343 1-11 <https://doi.org/10.71097/IJSAT.v17.i1.10343>
10. Narayan, K., Kadia., A.,Bharti, B., Bharti, P., (2026) “Swing Trading Strategy using Simple Moving Average Crossovers, Traded Volume and Super Trend Confirmations Integrated in Machine Learning”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110391 1-11 available in <https://www.ijsat.org/papers/2026/1/10391.pdf>
11. Kadia, A., Kumar, S., Qureshi, W., Sharma A., Kumar, R., (2026) “Simple Moving Average Crossover and Support Resistance Zone-based Stock Trading using Reinforcement Learning”, International Journal of Engineering Development and Research. 14(1), ISSN: 2321-9939 , IJEDR2601260 35-45 Available at <https://rjwave.org/IJEDR/papers/IJEDR2601260.pdf>
12. Kumar, A., Kumar, G., Alam, K., Kadia, A., (2026) “Smart Money Detection with Simple Moving Average and Traded Volume Confirmation Integrated in Machine Learning”. International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110343 1-11 <https://doi.org/10.71097/IJSAT.v17.i1.10343>
13. Kadia, A., Alam, K., Kumar, G., Kumar, G., (2026) “ Machine Learning–Driven Stock Price Breakout Identification with Simple Moving Average and Traded Volume Confirmation” International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110341 1-11 <https://doi.org/10.71097/IJSAT.v17.i1.10341>
14. Adhikary, S., Kadia A., (2025), “Algorithmic Trading with a Combination of Advanced Technical Indicators – An Automation”, International Journal on Science and Technology (IJSAT). 16(3) IJSAT25037777 1-20, 2025, <https://doi.org/10.71097/IJSAT.v16.i3.7777>
15. Kadia, A., Adhikary, S., Dey, R., Kar, A. (2025). “Deep Learning Based Stock Trading Strategies Using Leading Multi-Indicator Confirmations”. International Journal on Science and Technology (IJSAT). 16(3) IJSAT25037682 1-15. <https://doi.org/10.71097/IJSAT.v16.i3.7682>
16. Kumar, S., Kadia, A., Sharma A., Kumar, R., (2026) “Algorithm Trading Platform using Simple Moving Average Crossover (SMAC) and On-Balanced Volume (OBV) Integrated in Reinforcement Learning” International Journal of Engineering Development and Research. 14(1), ISSN: 2321-9939 , IJEDR2601221 758-767 Available :<https://rjwave.org/IJEDR/papers/IJEDR2601221.pdf>
17. Kadia, A., Alam, K., Kumar, G., Kumar, G., (2026) “ Machine Learning–Driven Stock Price Breakout Identification with Simple Moving Average and Traded Volume Confirmation” International Journal on Science and Technology (IJSAT). 17(1) IJSAT260110341 1-11 <https://doi.org/10.71097/IJSAT.v17.i1.10341>
18. Dey, R., Kassim, S., Maurya, S., Mahajan, R. A., Kadia, A., & Singh, M. (2024) “Machine Learning based Financial Stock Market Trading Strategies with Moving Average, Stochastic Relative Strength Index and Price Volume Actions for Indian and Malaysian Stock Market”, Journal of Electrical Systems, ISSN:1112-5209, 20-2s (2024): 759-767, DOI: <https://doi.org/10.52783/jes.1576>



19. Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications*, 36(3), 5932–5941. <https://doi.org/10.1016/j.eswa.2008.07.006>
20. Chen, A. S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computers & Operations Research*, 30(6), 901–923. [https://doi.org/10.1016/S0305-0548\(02\)00037-0](https://doi.org/10.1016/S0305-0548(02)00037-0)
21. Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22. <https://doi.org/10.1016/j.jfineco.2014.10.010>
22. Kara, Y., Boyacioglu, M. A., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines. *Expert Systems with Applications*, 38(5), 5311–5319. <https://doi.org/10.1016/j.eswa.2010.10.027>