

---

# A Comprehensive Study on Integration of Big Data and AI in Financial Decision-Making

---

Karim Elalkaoui\*, Safaa Moqqaddem  
and Mounir Ait Kerroum

*Laboratory of Research in Informatics (LaRI), Faculty of Sciences, Ibn Tofail  
University, Kenitra, Morocco*

*E-mail: karim.elalkaoui@uit.ac.ma*

*\*Corresponding Author*

Received 17 November 2025; Accepted 16 February 2026

## **Abstract**

The rapid proliferation of large-scale financial data, coupled with advancements in Artificial Intelligence (AI), has significantly transformed modern financial decision-making. This paper presents a comprehensive state-of-the-art review of AI-driven approaches supported by Big Data infrastructure in the financial domain. We analyse recent academic contributions (2023–2025) across machine learning (ML), deep learning and hybrid ensemble techniques applied to forecasting, portfolio optimisation, risk assessment and fraud detection. Emerging data architectures such as streaming frameworks and lakehouse platforms are assessed in terms of their ability to support real-time analytics and large-scale model deployment. We highlight the transition towards multimodal and attention-based models that integrate structured and unstructured data sources, and identify key challenges including concept drift, explainability, privacy and robustness. A detailed case study involving

*Journal of ICT Standardization, Vol. 14.2, 199–222.*

doi: 10.13052/jicts2245-800X.1423

© 2026 River Publishers

a GPU-accelerated hybrid deep learning and ensemble model for BTC–USD price prediction demonstrates practical benefits and current limitations: the hybrid model achieved an RMSE of 2656.69, a MAPE of 2.14%, and an  $R^2$  of 0.9626 on the test set. Although the absolute RMSE reflects the inherent volatility of the asset class, the low MAPE (2.14%) and high  $R^2$  confirm the model’s predictive efficacy during regime shifts, highlighting the necessity for future integration of macro-scenarios.

**Keywords:** Artificial intelligence (AI), big data, financial decision-making, deep learning, machine learning (ML), hybrid models, ensemble learning, real-time analytics, financial forecasting, risk assessment, fraud detection, concept drift, explainability, privacy, robustness, multimodal models, streaming frameworks, lakehouse architecture, BTC-USD prediction.

## 1 Introduction

### 1.1 Background and Motivation

The financial sector has witnessed a profound transformation in the past decade, driven by two key technological forces: Big Data and Artificial Intelligence (AI). This convergence has accelerated Financial Technology (FinTech), where modelling and forecasting financial instruments (e.g., equities, foreign exchange, derivatives, and cryptocurrencies) has become central to decision-making. A *lakehouse* architecture combines the flexibility of data lakes (raw, heterogeneous data at scale) with the governance and performance of data warehouses, providing a unified foundation for analytics and machine-learning workflows while reducing data duplication. Despite the abundance of market data, forecasting financial time series remains difficult because prices are shaped by interacting drivers such as macroeconomic conditions, policy changes, fundamentals, geopolitical shocks, and investor sentiment, often producing sudden event-driven discontinuities.

Financial time series prediction therefore represents a high-stakes, data-intensive problem at the intersection of economics, statistics, and AI. Traditional statistical models such as Linear Regression (LR), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA) have historically dominated this space, providing interpretable but often limited forecasting capability under non-stationarity and noise [1]. Classical finance foundations – including Modern Portfolio Theory [51], the Efficient Market Hypothesis [52], and the Black–Scholes option pricing

framework [53] provide important context for why forecasting is challenging and why rigorous evaluation is necessary. ML approaches such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Random Forests (RF) improved predictive power in non-linear settings [2, 4]. However, many classical ML pipelines still rely on substantial feature engineering and can struggle to capture long-range temporal dependencies.

More recently, deep learning has revolutionized time series modelling, offering architectures such as Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Transformer models that can capture complex temporal dynamics [5, 8]. The shift from hand-crafted features and fixed lag structures to automatic representation learning and long-range dependency modelling has improved predictive accuracy and robustness in many financial forecasting tasks.

## 1.2 Challenges in Financial Time Series Prediction

Financial time series exhibit statistical properties that distinguish them from generic time series. These include non-linearity, non-stationarity, volatility clustering, heavy tails, and regime dependence, which weaken the assumptions behind many traditional forecasting models [9, 12]. Behavioral and event-driven influences – such as investor psychology, government policy shifts, or earnings announcements – add further uncertainty and abrupt discontinuities [13]. Another crucial challenge is temporal heterogeneity: financial data contain both high-frequency microstructure signals and long-term macroeconomic trends, and ignoring either dimension can degrade performance. As a result, robust models must capture short-term volatility while also modelling long-range dependencies. Moreover, financial datasets are often noisy, incomplete, and prone to sudden regime shifts, requiring models to generalize under changing data distributions. This phenomenon is commonly referred to as *concept drift*: the relationship between input features and prediction targets evolves over time, causing models trained on historical patterns to lose accuracy in new market conditions.

### 1.2.1 Unified taxonomy of challenges and research directions

Financial AI challenges can be organized into four linked layers: (1) data and regimes (noise, non-stationarity, drift), (2) model reliability (overfitting, robustness, uncertainty), (3) explainability and governance (interpretability, auditability, compliance), and (4) deployment constraints (latency, throughput, monitoring, retraining). These layers directly motivate the research

agenda: drift-aware learning for regime changes, uncertainty-aware and robust validation for reliability, explainable AI for transparent decisions, and scalable Big Data pipelines for low-latency deployment with continuous monitoring.

### **1.3 The Need for Hybrid and Interpretable Models**

While deep learning can achieve state-of-the-art performance, its black-box nature raises concerns in high-risk domains such as finance. Domain experts and regulators increasingly demand transparency, explainability, and risk-awareness. There is growing interest in hybrid frameworks that combine the pattern-recognition capabilities of deep networks with the interpretability and stability of traditional models [14, 16]. This includes integrating statistical priors, economic knowledge, or ensemble learning to enhance robustness and reduce overfitting. In practice, explainability also supports auditability and model-risk governance in regulated settings.

### **1.4 Objective and Scope of This Study**

The objective of this study is to provide a comprehensive and practical review of modern AI approaches for financial time series forecasting. Unlike prior surveys, we jointly synthesise recent advances in attention-based, multimodal, and hybrid/ensemble models with deployment-oriented Big Data considerations (e.g., latency, scalability, and monitoring). We focus on studies published between 2023 and 2025 because this period reflects a major acceleration in attention-based and multimodal models, as well as production-grade Big Data architectures (e.g., streaming and lakehouse systems), which reshaped both research directions and real-world deployment requirements in finance. Our focus is threefold:

- First, we survey the evolution from statistical to ML and deep learning approaches, highlighting their comparative strengths and limitations.
- Second, we identify the major challenges that arise from the statistical nature of financial data and the operational requirements of real-world financial systems.
- Third, we propose a hybrid deep learning architecture combining CNN, GRU, and Transformer modules, followed by ensemble meta-learners, and we evaluate its performance using a real-world financial dataset.

This study is designed for a dual audience: financial practitioners seeking to deploy robust AI architectures and academic researchers aiming to identify critical gaps in the current literature regarding scalability and infrastructure.

To position our contribution with respect to existing review studies, Table 1 summarizes representative surveys published between 2020 and 2024, highlighting their primary focus, key limitations, and how this work extends the literature.

**Table 1** Comparison with prior review studies (2020–2024)

Ref/Year	Focus	Key Limitations	Our Contribution
[1]/2020	DL forecasting review (2005–2019).	Older scope; limited attention/multimodal and modern deployment.	Extends to 2023–2025; links AI to streaming/lakehouse and deployment constraints.
[3]/2021	DL applications in finance.	Less emphasis on latency/cost, governance, and drift-aware deployment.	Adds a system-level view with drift, privacy, and explainability for production.
[2]/2023	DL for financial time-series forecasting.	Model-centric; limited Big Data engineering integration.	Provides an end-to-end AI + Big Data view; maps methods to real-world constraints.
[4]/2023	Fraud detection review.	Task-specific; not forecasting-centric.	Broader coverage across tasks and a unified scalable architecture view.
[5]/2024	DRL portfolio review/benchmark.	Narrow scope (portfolio/DRL); not end-to-end.	Broader taxonomy and BTC–USD hybrid case study for validation.

## 1.5 Paper Structure

The remainder of this paper is structured as follows: Section 2 outlines related literature on machine learning and deep learning for financial prediction. Section 3 presents the proposed hybrid deep architecture and discusses its methodological components. Section 4 details the experimental setup, use case, and evaluation metrics. Section 5 discusses the results, compares them to benchmark models, and analyzes limitations. Section 6 concludes the paper with recommendations for future research and practical deployment.

## 1.6 Methodology

Relevant literature was identified through structured keyword-based searches across major academic databases, including IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect, and Google Scholar. The search strategy combined terms related to AI methods and financial applications, such as: “financial time series forecasting”, “deep learning finance”, “hybrid model stock prediction”, “transformer time series”, “multimodal finance”, “concept drift detection”, and “big data architecture finance”.

To ensure consistency and relevance, we restricted our selection to peer-reviewed journal and conference publications written in English and published between January 2023 and July 2025. Preprints (e.g., arXiv) were included only when they demonstrated clear novelty and had been cited by at least one peer-reviewed publication. We excluded studies that lacked empirical evaluation, focused on non-financial domains, or provided insufficient methodological detail for meaningful comparison.

The selection process followed two screening stages: first, titles and abstracts were reviewed to remove irrelevant or duplicate records; second, full-text screening was conducted to confirm eligibility and extract evidence. Each retained study was then analysed using a unified template capturing the financial domain (forecasting, risk, portfolio, fraud), modelling approach (ML/DL/hybrid/ensemble), datasets used, evaluation metrics, and reported limitations. Overall, an initial pool of **312** records was retrieved, of which **87** studies were retained for detailed analysis and form the evidence base for the subsequent sections.

## 2 Foundations

### 2.1 Overview of Financial Decision-Making

Financial decision-making refers to processes in which individuals or institutions allocate capital and manage risk under uncertainty, including investment, financing and operational decisions. While classic economic and statistical models—such as regression-based forecasting and the Markowitz portfolio framework—remain prevalent, they often fail to capture the complex temporal dependencies and nonlinear patterns observed in modern financial markets [6]. Recent studies demonstrate that purely model-driven methods struggle when faced with volatile conditions and high-dimensional data structures [7].

## **2.2 Role of Big Data and AI in Finance**

Big Data technologies provide the infrastructure needed to manage the enormous volume, velocity, and variety of financial data. Distributed computing platforms such as Hadoop and Spark enable scalable storage and processing of heterogeneous data streams. AI techniques, including ML and deep neural networks, make it possible to extract complex patterns and transform raw data into actionable insights, thereby improving decision-making processes [8].

From an operational perspective, the suitability of streaming and lakehouse platforms for financial AI is determined by three constraints: latency (time from data arrival to model inference), throughput (ability to ingest and process high-volume streams), and cost (compute, storage, and governance overhead). Message queues such as Apache Kafka support scalable ingestion of real-time feeds, while stream processors such as Apache Flink enable low-latency event-time computations for continuous analytics [34, 35]. On the storage layer, lakehouse implementations such as Delta Lake provide ACID-like reliability over cloud object storage, supporting consistent batch and streaming pipelines [36], and engines such as Apache Spark enable large-scale feature engineering and training workloads [37]. In practice, finance use cases must balance strict service-level requirements (e.g., fraud detection) against the cost of continuous processing and model retraining, especially under bursty market conditions.

Recent research has shown that combining Big Data infrastructures with AI-based analytics leads to significant improvements in real-time forecasting and fraud detection performance [9, 10].

## **3 AI Techniques Used in Financial Decision-Making**

### **3.1 Classical ML Models**

Classical ML algorithms remain widely used because of their interpretability and empirical robustness. Support Vector Machines (SVMs) have been employed for credit risk assessment and market state classification [17]. Ensemble methods such as Random Forests and Gradient Boosted Trees (e.g., XGBoost, LightGBM) have been shown to outperform traditional econometric models for early-warning financial systems [14, 16]. Their main limitation is the need for extensive feature engineering and their limited ability to capture long-term temporal dependencies.

## **3.2 Deep Learning Models**

### **3.2.1 Recurrent networks (LSTM/GRU)**

LSTM and GRU networks are particularly effective for sequential forecasting of stock prices and volatility because of their ability to capture long-range temporal dependencies [11, 12].

### **3.2.2 Convolutional neural networks (CNNs)**

Convolutional Neural Networks have been applied to limit order book data and candlestick-chart representations, where they extract spatial and temporal patterns that are not visible to traditional models [26].

### **3.2.3 Attention-based (transformer) models**

Transformers replace recurrence with self-attention, allowing each time step to learn weighted dependencies over all other time steps. This enables parallel training and better modeling of long-range temporal relationships compared to purely recurrent architectures. In financial forecasting, attention is particularly useful for capturing delayed effects (e.g., macro signals influencing price trends) and for learning cross-feature interactions when many engineered indicators and contextual variables are included. However, Transformers are typically data- and compute-intensive, can overfit noisy markets if regularization and leakage control are weak, and may degrade under regime shifts unless combined with drift-aware retraining or robust validation. Their interpretability is also partial: attention weights offer insight into influential time segments, but should be complemented with explanation methods (e.g., SHAP) for feature-level attribution [19]. Transformers were introduced in [13] and later adapted for forecasting through architectures such as TFT and Informer [6, 49].

### **3.2.4 Multimodal models (text–price fusion)**

Multimodal learning extends financial forecasting beyond numerical price and volume series by integrating unstructured signals such as financial news, social-media sentiment, earnings-call text, or even chart-based visual representations. A common design is to use separate encoders for each modality (e.g., a language model for text and a temporal encoder for numerical features), followed by a fusion module that combines the learned representations. Fusion can be performed via early fusion (concatenating inputs), late fusion (merging latent embeddings), or cross-attention, where one modality (e.g., news) guides the model to focus on relevant temporal patterns in the

price sequence. Pre-trained financial language models such as FinBERT have been widely used to extract sentiment and event representations from text, which can improve robustness during news-driven volatility [27]. More recent approaches adopt transformer-based multimodal fusion mechanisms to handle unaligned sequences and complex interactions between modalities [30], and financial-specific multimodal architectures have demonstrated competitive performance by jointly learning from news and chart information for stock movement prediction [46]. However, multimodal models introduce additional challenges, including timestamp alignment, noisy or conflicting signals, higher computational cost, and potential information leakage if text is not carefully synchronized with market time. As a result, rigorous data alignment, validation, and ablation studies are essential to ensure that performance gains reflect genuine predictive value rather than artefacts of data availability.

Overall, The progression from recurrent to attention-based and multimodal architectures reflects a trade-off between temporal expressiveness, computational efficiency, and robustness to market noise. While LSTM and GRU networks capture sequential dependencies with moderate computational cost, they may struggle with very long-horizon interactions and abrupt regime shifts in non-stationary markets [11, 12]. CNN-based models complement this by extracting local spatial-temporal patterns from order-book or chart-style representations, but remain limited in modelling distant dependencies [26]. Transformer-based approaches address these limitations through self-attention, enabling effective learning of long-range relationships and cross-feature interactions, albeit at higher computational cost and a higher risk of overfitting under volatile conditions [13]. Multimodal models further extend this paradigm by integrating textual, visual, and numerical signals, improving context-awareness during event-driven volatility, while introducing challenges in alignment, fusion design, and signal noise [27, 30, 46]. Overall, no single architecture dominates all forecasting scenarios; model choice should balance predictive performance, interpretability requirements, latency constraints, and the available data modalities.

### **3.3 Hybrid and Ensemble Approaches**

Hybrid and ensemble methods aim to combine the pattern-recognition strength of deep networks with the stability of classical models. A common

strategy is feature-stacking, where recurrent encoders (e.g., LSTM/GRU) extract latent time-series representations that are then used by a tree-based learner such as XGBoost for final prediction [11, 12, 15]. Another approach is model-stacking, where outputs from heterogeneous learners (e.g., CNN/LSTM and Random Forest) are aggregated using a meta-learner following the stacked generalization principle [14, 18]. These approaches can improve robustness to noise and non-stationarity, but they introduce additional computational overhead and may reduce transparency unless complemented with explanation tools such as SHAP or LIME [19, 20], creating a trade-off between performance gains and operational interpretability.

## **4 Big Data Frameworks and Technologies in Finance**

### **4.1 Data Sources and Infrastructures**

Financial analytics increasingly depends on a wide range of heterogeneous data sources, including structured market data (e.g., stock prices, trading volumes), semi-structured transaction logs and unstructured data such as news articles, reports, and social media feeds. The integration of such diverse data requires scalable storage and processing infrastructures. Distributed computing platforms such as Apache Hadoop and Apache Spark are commonly adopted in financial firms to handle batch and iterative processing of large-scale datasets [21, 22]. In addition, lakehouse architectures have recently emerged as a unified paradigm combining the flexibility of data lakes with the transactional consistency of data warehouses, thereby enabling more efficient analytics and easier integration with ML pipelines [23].

### **4.2 Real-Time Data Processing and Streaming**

Real-time decision-making in areas such as algorithmic trading and fraud detection requires fast ingestion and low-latency processing of streaming data. Frameworks such as Apache Kafka and Apache Flink facilitate high-throughput data ingestion, while stream-processing engines allow continuous model execution over incoming data streams [24]. Recent studies demonstrate that combining stream-processing architectures with AI models can significantly improve real-time anomaly detection and market trend prediction [25, 26]. In addition, edge computing strategies are increasingly being explored to enable faster local inference while reducing latency in time-critical financial applications [27].

## **5 Recent Advances(2023–2025)**

### **5.1 Risk Assessment and Fraud Detection**

Recent research has focused heavily on developing AI models that enhance the accuracy and speed of fraud detection and credit risk assessment. Graph-based neural networks have been proposed to model transactional relationships and detect anomalous behaviour in financial networks [28]. In addition, hybrid deep learning–ensemble schemes have demonstrated improved generalisation in real-time credit default prediction scenarios [29]. Several studies report that incorporating alternative data (e.g., customer browsing activity) into AI models significantly increases early detection rates of fraudulent events compared to traditional rule-based systems [30].

### **5.2 Portfolio Management and Trading**

In the area of portfolio optimisation and quantitative trading, Transformer-based time-series models have shown promising results, outperforming LSTM baselines in predicting multi-asset returns over long horizons [31]. Other recent studies have explored reinforcement learning approaches that dynamically adjust portfolio allocations based on market conditions and risk constraints [32]. Combining reinforcement learning with attention-based architectures has been reported to improve financial performance metrics such as Sharpe and Sortino ratios in simulated trading environments [33].

### **5.3 LLMs and Multimodal AI in Finance**

Large Language Models (LLMs) have recently been applied to extract actionable information from unstructured text such as company earnings calls and macroeconomic announcements. For instance, transformer-based language models fine-tuned on financial text have been shown to improve earnings surprise prediction and sentiment-based trading strategies [34]. Hybrid multimodal systems that integrate numeric data with textual and visual features (e.g., combining price series with news sentiment and candlestick images) have also been proposed to enhance forecasting performance and increase robustness under volatile conditions [35, 36].

## **6 Challenges and Open Research Problems**

### **6.1 Non-Stationarity and Concept Drift**

One of the major challenges in financial decision-making is the non-stationary nature of financial data. Market conditions evolve over time due

to regulatory changes, macroeconomic shifts and unexpected geopolitical events, which can cause previously learned patterns to become invalid. Standard machine learning and deep learning models often assume that the training and deployment data are drawn from the same distribution; this assumption rarely holds in real-world financial environments.

Recent studies have highlighted the importance of concept drift detection and adaptive learning strategies in order to maintain model performance under changing conditions. For example, online learning and incremental training techniques have been applied to LSTM-based forecasting models to adapt to drifting distributions in stock price time series [37]. Other work proposes the use of drift-robust architectures that integrate uncertainty quantification and Bayesian model updating to cope with non-stationarity [38]. Despite these advances, developing AI systems that remain reliable under severe regime shifts (e.g., financial crises) remains an open research problem.

## **6.2 Model Explainability and Regulation**

The increasing complexity of AI models used in financial applications—particularly deep and hybrid architectures—raises significant concerns about transparency and interpretability. Financial institutions must operate under strict regulatory frameworks that require not only high predictive performance but also clear explanations of how decisions are made. Traditional explainability techniques, such as feature importance scores or local surrogate models, are often insufficient for capturing the full behaviour of deep neural networks [39].

Recent research efforts have attempted to combine explainable AI (XAI) techniques with financial models in order to meet regulatory standards. For example, attention-weight visualisation has been used to identify which time-series segments contribute most to a prediction, while SHAP (Shapley Additive Explanations) values have been applied to interpret hybrid LSTM–ensemble models in credit scoring [40]. Nevertheless, developing explainable models that satisfy both regulatory requirements and performance criteria is still an open challenge, especially in real-time decision-making environments.

## **6.3 Data Privacy and Governance**

Financial datasets often contain highly sensitive personal and transactional information, which makes data privacy and governance a critical concern

in AI-driven decision-making. Strict regulations such as the General Data Protection Regulation (GDPR) impose legal constraints on how financial institutions collect, process and store data. These regulations challenge the development of data-hungry models and complicate the integration of heterogeneous data sources across different jurisdictions.

To address these issues, recent studies have explored privacy-preserving machine learning techniques, including federated learning and differential privacy, which enable collaborative training without exposing raw data [41]. Such approaches have been applied to credit scoring and fraud detection systems and have demonstrated a balance between model performance and privacy preservation [42]. However, implementing these techniques in production remains difficult due to high system complexity and trade-offs between privacy guarantees and predictive accuracy.

## **6.4 Robustness and Generalisation**

AI models deployed in financial systems must be robust to noise, adversarial manipulation and unexpected market events. However, many existing approaches tend to overfit to historical data and exhibit poor out-of-sample generalisation. This issue is particularly critical in high-stakes applications such as automated trading or systemic risk assessment. Recent research has highlighted the importance of incorporating regularisation techniques, adversarial training and uncertainty estimation methods into financial models in order to improve robustness [43]. For example, ensemble-based Bayesian deep learning models have been shown to provide better uncertainty estimates and more reliable predictions during periods of high volatility [44]. Nevertheless, ensuring reliable model performance in unseen or extreme conditions (e.g., black-swan events) continues to be a major unresolved challenge in the field.

## **7 Discussion and Future Directions**

### **7.1 Implications for Practice**

The integration of AI and Big Data into financial decision-making has significant implications for practitioners. Financial institutions can benefit from enhanced forecasting accuracy, improved risk management, and faster detection of fraudulent activities. However, successful deployment requires not only advanced models but also robust data infrastructure and strong governance frameworks. Organisations must therefore invest in scalable data

pipelines, cross-disciplinary teams, and risk-management procedures that account for model uncertainty and regulatory requirements. In particular, explainable AI capabilities should be integrated at the design stage to enable better auditability and stakeholder trust.

## 7.2 Future Research Directions

Several promising research directions emerge from the current literature. First, there is a need for drift-aware and adaptive models that can dynamically update in response to changing market conditions. Second, multimodal learning approaches that combine numerical, textual, and visual data streams should be further explored to capture complementary information in complex financial environments. Third, the development of privacy-preserving and explainable architectures remains a high priority, especially in regulated domains. Finally, further work is required to evaluate AI models under stress-testing and extreme-event scenarios, in order to better understand their robustness and decision-making reliability.

## 8 Use Case: Hybrid AI Experiment on BTC-USD Forecasting

We validate the practical impact of the proposed approach through a reproducible BTC–USD forecasting experiment. Bitcoin is a demanding test case due to its high volatility and frequent regime shifts. The goal is to demonstrate an end-to-end workflow that combines feature-rich data preparation with a hybrid deep-learning and ensemble architecture.

### 8.1 Data Workflow and Experimental Setup

The dataset was built from daily BTC–USD OHLCV data collected through Yahoo Finance using the `yfinance` API. To add broader market context, we also included SPY, QQQ, VIX, and the U.S. Dollar Index (DXY), aligned by date and merged into a single table. We then engineered features covering multi-horizon returns and log-returns, common technical indicators (moving averages, RSI, MACD, Bollinger Bands), volatility and momentum measures, market-relative signals, and calendar/cyclical variables. Highly correlated features were removed using a correlation threshold to reduce redundancy.

For preprocessing, input features were scaled with `RobustScaler` to limit the impact of outliers, while the prediction target was defined as next-step log

return and standardized with StandardScaler. The deep-learning branch was trained on rolling sequences of 60 time steps, whereas the ensemble learners used non-sequential feature vectors. Data were split chronologically to avoid leakage: the final 20% was reserved for testing, and the remaining 80% was split into training and validation (roughly 64% train, 16% validation, 20% test overall). Performance was reported using RMSE, MAE, MAPE, and  $R^2$  on reconstructed prices.

## 8.2 Hybrid Model Architecture

The proposed model follows a hybrid design that combines three deep-learning branches trained in parallel: a CNN→LSTM branch to capture local patterns and short-term dynamics, a bidirectional GRU branch as an alternative sequential encoder, and a Transformer branch to model longer-range dependencies through self-attention. Each branch outputs a compact latent representation, and these representations are concatenated and passed through fusion layers with a residual connection to form a final shared embedding.

**Table 2** Hybrid deep-learning branches used in the architecture

Branch	Purpose
<b>CNN → LSTM</b>	Local temporal smoothing and sequential learning.
<b>GRU</b>	Alternative sequential representation.
<b>Transformer</b>	Long-range attention and context modelling.

The branch outputs are concatenated into a 128-dimensional latent vector and refined through residual fusion layers. In parallel, a stacking ensemble (Random Forest, Gradient Boosting, ElasticNet) is trained, and a Ridge meta-learner combines their predictions to produce the final forecast.

## 8.3 Results and Interpretation

After GPU-accelerated training, the model achieved the following test-set results:  $RMSE = 2656.69$ ,  $MAPE = 2.14\%$ , and  $R^2 = 0.9626$  (Table 3).

**Table 3** Test-set performance metrics

Metric	Value
RMSE	2656.69
MAPE	2.14%
$R^2$	0.9626

The hybrid pipeline improved performance by 39% compared to the baseline deep-learning model, highlighting the benefit of combining complementary architectures. While the overall fit is strong (high  $R^2$ ) and the relative error remains moderate (MAPE), the absolute RMSE is relatively large in dollar terms because BTC–USD operates at a high price level and exhibits heavy volatility, abrupt jumps, and regime shifts; under these conditions, even small percentage deviations can translate into large absolute errors. Moreover, purely offline training on historical patterns may under-represent event-driven movements and macro-regime transitions. These results suggest that future work should prioritize drift-aware updating and macro-scenario integration to improve robustness under changing market conditions.

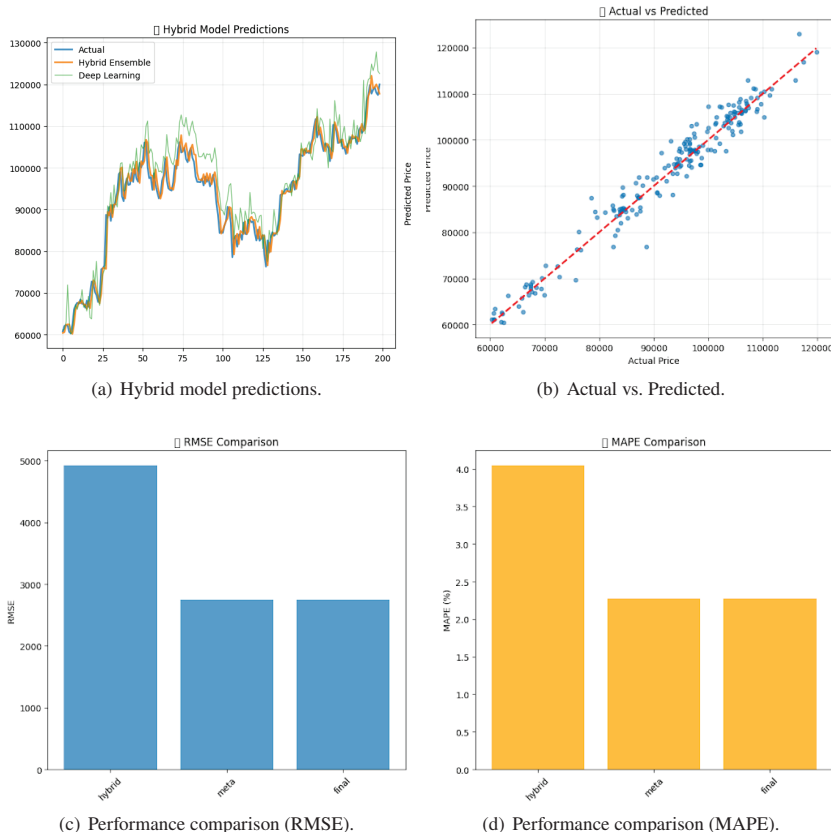
**Ablation (stacking effect):** To quantify the contribution of the stacked classical learners, we compared the deep-only predictor (hybrid network output) against the full deep+stacking pipeline (meta-learner output). Table 4 reports the corresponding test-set metrics.

**Table 4** Ablation study: deep-only vs. deep+stacking (test set)

Metric	Deep-only	Deep+Stacking
RMSE	3262.53	2572.36
MAE	2610.71	1928.61
MAPE	2.58%	1.93%
$R^2$	0.9252	0.9535

Compared to the deep-only predictor, adding the stacking ensemble reduced RMSE from 3262.53 to 2572.36 (21.15% improvement) and MAPE from 2.58% to 1.93% (25.33% improvement), while  $R^2$  increased from 0.9252 to 0.9535 ( $\Delta R^2 = +0.0283$ ), confirming that the ensemble layer contributes materially to robustness under BTC–USD’s noisy, non-stationary dynamics.

Figure 1 shows the prediction and performance results of the proposed hybrid model. Subfigures (a) and (b) illustrate that the hybrid ensemble closely follows the BTC–USD trajectory and maintains a strong overall fit across the test window. Subfigures (c) and (d) compare RMSE and MAPE against baseline variants, highlighting the relative advantage of the hybrid design.



**Figure 1** Comparison of prediction and performance results for the proposed hybrid model.

### 8.4 Future Improvements

In future work, we will enhance robustness to non-stationarity by adding drift-detection and walk-forward retraining so the model can adapt when market regimes change. We also plan to incorporate macro-scenario signals—such as monetary-policy proxies, inflation/interest-rate indicators, and macro sentiment—to better capture event-driven movements that are not fully represented in historical OHLCV patterns. Finally, we will refine the fusion stage by exploring more flexible attention-based fusion mechanisms and will validate the approach on additional asset classes (e.g., equities and commodities) to assess generalizability beyond crypto markets.

## 9 Conclusion

The integration of Artificial Intelligence (AI) and Big Data is reshaping financial decision-making by enabling scalable analytics, faster risk-aware responses, and more automated forecasting pipelines. This study showed that meaningful progress in financial prediction requires not only stronger models, but also production-oriented data architectures capable of handling high volume, velocity, and heterogeneous sources under real operational constraints.

Across recent work (2023–2025), a clear trend emerges from classical machine learning toward deep, hybrid, and multimodal systems. Recurrent models (LSTM/GRU) remain effective for sequential dependence, CNNs capture local structure in market representations, and attention-based Transformers better model long-range interactions, while multimodal approaches enrich forecasting by combining numerical series with unstructured signals such as news and sentiment. In parallel, modern lakehouse and stream-processing platforms provide a practical foundation for governance, reproducibility, and near real-time deployment at scale.

Our BTC–USD case study provides practical evidence of this integration: a GPU-accelerated hybrid architecture combining CNN→LSTM, GRU, and Transformer branches with a stacked ensemble improved performance over a single deep-learning baseline and achieved  $R^2 = 0.9626$  with  $\text{MAPE} = 2.14\%$  on the test set. However, the remaining gap in absolute error highlights persistent challenges in highly volatile markets, where regime shifts and event-driven jumps reduce the reliability of offline-trained models.

Overall, the findings emphasize that future progress depends on drift-aware updating, macro-scenario integration, and stronger interpretability mechanisms that support trustworthy decision support. Continued development of adaptive, explainable, and multimodal AI—paired with robust Big Data infrastructures—will be essential for reliable deployment in real financial environments.

## References

- [1] O.B. Sezer, M.U. Gudelek, and A.M. Ozbayoglu. Financial time series forecasting with deep learning: A systematic literature review (2005–2019). *Applied Soft Computing*, 90, 106181, 2020.
- [2] Xu. Zhang, W. Hussain, and W. Chen. Deep learning for financial time series forecasting: A review. *Computer Science Review*, 49, 100567, 2023.

- [3] E. Mienye, N. Jere, G. Obaido, I. D. Mienye and K. Aruleba. Deep Learning in Finance: A Survey of Applications and Techniques *IEEE Access*, 9, 107222–107248, 2021.
- [4] A. Ali, S. A. Razak, S. H. Othman, T. A. E. Eisa, A. Al-Dhaqm, M. Nasser, T. Elhassan, H. Elshafie and A. Saif. Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Journal of Financial Crime*, 30(2), 506–526, 2023.
- [5] Y. Jiang, J. Olmo, and M. Atwi. Deep reinforcement learning for portfolio selection. *Global Finance Journal*, 43(2), 345–372, 2024.
- [6] B. Lim, S.O. Arik, N. Loeff, and T. Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748–1764, 2021.
- [7] B.N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *International Journal of Forecasting*, 36(3), 1091–1102, 2020.
- [8] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 36(3), 1181–1191, 2020.
- [9] S. Makridakis, I. Spiliotis, and V. Assimakopoulos. M5 accuracy competition: Results, findings and conclusions. *International Journal of Forecasting*, 38(4), 1346–1364, 2022.
- [10] G. Woo, C. Liu, D. Sahoo, A. Kumar, and S. Hoi. ETSformer: Exponential smoothing transformers for time-series forecasting. *arXiv preprint arXiv:2202.01381*, 2022.
- [11] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8), 1735–1780, 1997.
- [12] K. Cho, B. v. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.
- [13] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, 30, 2017.
- [14] L. Breiman. Random forests. *Machine Learning*, 45, 5–32, 2001.
- [15] T. Chen and C. Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794, 2016.
- [16] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T. Liu. LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems*, 30, 2017.

- [17] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20, 273–297, 1995.
- [18] D.H. Wolpert. Stacked generalization. *Neural Networks*, 5(2), 241–259, 1992.
- [19] S.M. Lundberg and S.-I. Lee. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems*, 30, 2017.
- [20] M.T. Ribeiro, S. Singh, and C. Guestrin. Why should I trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144, 2016.
- [21] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, D. Pedreschi, F. Giannotti A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 1–42, 2019.
- [22] Y. Chen , R. Calabrese and B. Martin-Barragan Interpretable machine learning for imbalanced credit scoring datasets *European Journal of Operational Research*, 312(1), 357–372, 2024.
- [23] Y. Gal and Z. Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning*, 1050–1059, 2016.
- [24] B. Lakshminarayanan, A. Pritzel, and C. Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems*, 30, 2017.
- [25] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, M. Shahzad, W. Yang, R. Bamler and X. X. Zhu A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56, 1513–1589, 2023.
- [26] Z. Zhang, S. Zohren, and S. Roberts. DeepLOB: Deep convolutional neural networks for limit order book data. *IEEE Transactions on Signal Processing*, 67(11), 3001–3012, 2019.
- [27] D. Araci. FinBERT: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*, 2019.
- [28] D. Cheng, Y. Zou, S. Xiang, C. Jiang Graph neural networks for financial fraud detection: A review. *Frontiers of Computer Science*, 239, 122300, 2025.
- [29] D. Shah, V. Shah, and H. Patel. A comprehensive survey on applications of graph neural networks in fraud detection. *Knowledge and Information Systems*, 65, 4027–4064, 2023.

- [30] Y-H. H. Tsai, S. Bai, P. P. Liang, J. Zico Kolter, L-P. Morency, R. Salakhutdinov Multimodal transformer for unaligned multimodal language sequences. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 6558–6569, 2019.
- [31] J. Moody and M. Saffell. Reinforcement learning for trading. In *Advances in Neural Information Processing Systems*, 11, 1998.
- [32] Y. Deng, F. Bao, Y. Kong, Z. Ren; Q. Dai Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3), 653–664, 2016.
- [33] H. Yang, X. Zhang, A. Walid, and X-Y. Liu. Deep reinforcement learning for automated stock trading: An ensemble strategy. *SSRN Electronic Journal*, 2020.
- [34] J. Kreps, N. Narkhede, and J. Rao. Kafka: A distributed messaging system for log processing. In *Proceedings of the 6th International Workshop on Networking Meets Databases*, 2011.
- [35] P. Carbone ,A. Katsifodimos ,S. Ewen ,V. Markl,S. Haridi,K. Tzoumas Apache Flink: Stream and batch processing in a single engine. *IEEE Data Engineering Bulletin*, 38(4), 2015.
- [36] M. Armbrust, T. Das, L. Sun, B. Yavuz, S. Zhu, M. Murthy, J. Torres, H. v. Hovell, A. Ionescu, A. Łuszczak, M. Świtakowski, M. Szafranski, X. Li, T. Ueshin, M. Mokhtar, P. Boncz, A. Ghodsi, S. Paranjpye, P. Senster, R. Xin and M. Zaharia Delta Lake: High-performance ACID table storage over cloud object stores. *Proceedings of the VLDB Endowment*, 13(12), 3411–3424, 2020.
- [37] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M. J. Franklin, A. Ghodsi, J. Gonzalez, S. Shenker and I. Stoica Apache Spark: A unified engine for big data processing. *Communications of the ACM*, 59(11), 56–65, 2016.
- [38] C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4), 211–407, 2014.
- [39] H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson and B. Agüera y Arcas Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 1273–1282, 2017.
- [40] P. Kairouz, H.B. McMahan, B. Avent, A. Bellet, M. Bennis, A.N. Bhalaji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R.G.L. D’Oliveira, H. Eichner, S. El Rouayheb, D. Evans, J. Gardner,

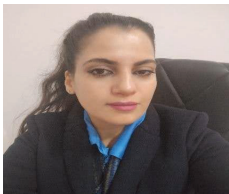
- Z. Garrett, A. Gascón, B. Ghazi, P.B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konečný, A. Korolova, F. Koushanfar, S. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Özgür, R. Pagh, M. Raykova, H. Qi, D. Ramage, R. Raskar, D. Song, W. Song, S.U. Stich, Z. Sun, A.T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F.X. Yu, H. Yu, and S. Zhao, Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1–2), 1–210, 2021.
- [41] Y. Wu, Y. Wang, and X. Yuan. A novel deep reinforcement learning framework for portfolio allocation with state representation learning. *Expert Systems with Applications*, 238, 121828, 2024.
- [42] A. Sadighi, M.S. Nosrati, and S.H. Hasheminejad. A deep attention-based reinforcement learning framework for algorithmic trading under regime shifts. *Engineering Applications of Artificial Intelligence*, 127, 107258, 2024.
- [43] Y. Wang, J. Zhang, and L. Wu. Market regime adaptive deep reinforcement learning for quantitative trading. *Neurocomputing*, 566, 127026, 2024.
- [44] R. Chandra, Y. He Bayesian neural networks for stock price forecasting before and during COVID-19 pandemic *Information Sciences*, 657, 119920, 2024.
- [45] R. C. Cavalcante and A. Oliveira An approach to handle concept drift in financial time series based on Extreme Learning Machines and explicit Drift Detection *IEEE International Joint Conference on Neural Network*, 153, 111298, 2024.
- [46] P. Chen, Z. Boukouvalas and R. Corizzo A deep fusion model for stock market prediction with news headlines and time series data *Knowledge-Based Systems*, 21229–21271, (2024)
- [47] J. Park and H. Kim. Integrating technical indicators, news sentiment, and stock correlation for enhanced portfolio optimization. *Neural Computing and Applications*, 54, 1057–1073, 2024.
- [48] R. Maulik, R. Egele, K. Raghavan and P. Balaprakash. Quantifying uncertainty for deep learning based forecasting and flow-reconstruction using neural architecture search ensembles *Physica D: Nonlinear Phenomena*, 158, 107059, 2024.
- [49] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong and W. Zhang. Informer: Beyond efficient transformer for long sequence time-series

- forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(12), 11106–11115, 2021.
- [50] Y. Liu, T. Hu, H. Zhang, H. Wu, S. Wang, L. Ma and M. Long. Transformer: Inverted transformers are effective for time series forecasting. *arXiv preprint arXiv:2310.06625*, 2023.
- [51] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1), 77–91, 1952.
- [52] E.F. Fama. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417, 1970.
- [53] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654, 1973.

## Biographies



**Karim Elalkaoui** is a PhD candidate at Ibn Tofail University in Kenitra, Morocco, at the Laboratory of Research in Informatics (LaRI), Faculty of Sciences. He obtained a master's degree in Big Data and Cloud Computing from Ibn Tofail University. His research interests include artificial intelligence, big data analytics, and financial decision-making systems.



**Safaa Moqqaddem** is an Assistant Professor at the National School of Commerce and Management (ENCG), Ibn Tofail University, Kenitra, Morocco.

She received a master's degree in computer science and telecommunications and a PhD in computer science and telecommunications from the Faculty of Sciences, Ibn Tofail University. Her research interests include artificial intelligence, big data, computer vision, machine perception (particularly 3D object detection and tracking), image processing, and intelligent systems. She is also a member of the LaRI Laboratory at the Faculty of Sciences, Kenitra.



**Mounir Ait Kerroum** received his Master's Degree (DESA) in Computer Science and Telecommunications from the Faculty of Sciences at Mohammed V University, Rabat, Morocco, in 2003, followed by a PhD from the same institution in 2010. In March 2010, he joined the National School of Business and Management (ENCG) of Kénitra, Ibn Tofail University, as an Assistant Professor, before being promoted to Associate Professor in 2014. He is a founding and permanent member of the Laboratory for Research in Computer Science and Telecommunications (LaRIT), established in 2010. From 2017 to 2020, he headed the 'Networks, Telecommunications, and Artificial Intelligence' research team within the lab. He is also an Associate Researcher at the Laboratory for Research in Computing and Telecommunications (LRIT) at the Faculty of Sciences, Rabat. Since December 2020, he has held the rank of Full Professor. His current research interests include Artificial Intelligence, Machine Learning, and Deep Learning applied to pattern recognition in remote sensing and medical imaging. His expertise also extends to the application of Explainable AI (XAI), NLP, and Deep Learning for financial market and cryptocurrency forecasting, as well as the use of sentiment analysis to enhance model transparency and explainability.