WCRG Global Research WORLD COMET RESEARCH GROUP

International Journal of Artificial Intelligence & Digital Transformation
Vol. 2, No. 2, 2019
Doi: 10.67228/30713315/IJAIDT-2019PII2P24
PP. 01-14

Original Article

Engineering Data for Explainable AI in Financial Forecasting Systems

Dr. Fatima Zahra El Idrissi[1], Marta Silva[2]

[1,2]*Department of Business Studies, Atlas Valley University, Fez, Morocco.*

Received: 11-04-2019 Revised: 11-24-2019 Accepted: 12-02-2019 Published: 12-05-2019

ABSTRACT

In high-stakes financial forecasting systems, the demand for both predictive accuracy and interpretability has elevated the role of Explainable AI (XAI). However, achieving meaningful explainability requires more than just applying model-agnostic techniques it demands a robust data engineering foundation that ensures transparency, traceability, and consistency across the data lifecycle. This paper presents a novel data engineering framework tailored for explainable AI applications in financial forecasting. We explore how pipeline design, feature provenance, and metadata management influence the quality and reliability of XAI outputs. Through a case study on stock market prediction using SHAP-enabled XGBoost models, we demonstrate how engineered data pipelines contribute to regulatory compliance, model transparency, and trustworthiness. Our findings suggest that data engineering is not just a support mechanism but a critical enabler of explainability in AI-driven financial decision-making systems.

KEYWORDS

Data Engineering, Explainable Ai (Xai), Financial Forecasting, Feature Lineage, Shap, Stock Market Prediction, Transparent Ai, Metadata Management, Regulatory Compliance, Data Pipeline.

ISSN: 3071-3315
PP: 01-14.

1. INTRODUCTION

1.1. Motivation for Financial Forecasting Accuracy and Transparency

In modern financial systems, forecasting models are pivotal for making informed decisions across diverse applications such as asset pricing, credit risk assessment, portfolio optimization, and algorithmic trading. Given the rapid influx of structured and unstructured financial data—from real-time market feeds to economic indicators and news sentiment—organizations are increasingly relying on AI and machine learning models to enhance forecasting precision. However, as these models become more sophisticated, their decision-making processes often become less interpretable, leading to a "black box" problem. In finance, where decisions carry substantial monetary and regulatory implications, opaque predictions can erode stakeholder trust, obscure model accountability, and even introduce systemic risk. Therefore, ensuring both high predictive accuracy and interpretability is not just desirable—it is essential for responsible and auditable AI deployment in the financial domain.

1.2. Limitations of Black-Box Models in Finance

While deep learning models, ensemble techniques, and other complex AI algorithms have demonstrated superior forecasting capabilities, their inner workings remain largely inscrutable to end users and auditors. This lack of transparency is particularly problematic in finance, where regulatory bodies like the SEC, FINRA, and European Banking Authority increasingly demand explainable decision-making processes. Black-box models fail to provide insight into the rationale behind a prediction, making it difficult to identify biases, validate the consistency of outputs, or align outcomes with fiduciary responsibilities. Furthermore, in high-stakes scenarios such as loan approvals, investment recommendations, or fraud detection, decisions must be traceable and justifiable. The absence of interpretability in such models not only poses compliance risks but also undermines institutional credibility and customer confidence.

1.3. Importance of Data Engineering in Enabling Explainability

The ability to explain an AI model's prediction does not begin with the model itself it begins with the data. Data engineering lays the groundwork for explainability by ensuring that the datasets fed into models are clean, well-documented, traceable, and contextualized. Techniques such as feature lineage tracking, standardized metadata schemas, schema versioning, and real-time data validation play critical roles in making both the data and subsequent explanations transparent. In financial forecasting, where data comes from heterogeneous sources such as trading platforms, economic bulletins, and textual news feeds, the complexity of the data pipeline can significantly influence the reliability of XAI methods. Without a robust data engineering infrastructure, explanations provided by tools like SHAP or LIME may be misleading, inconsistent, or technically infeasible. Thus, engineering data with explainability in mind is foundational to building transparent, trustworthy AI systems in finance.

ISSN: 3071-3315
PP: 01-14.

1.4. Paper Contributions and Structure

This paper contributes to the growing discourse on trustworthy AI by presenting a data engineering framework designed to support explainable AI in financial forecasting applications. Specifically, we propose architectural components and pipeline design principles that enhance feature transparency, enable real-time feature attribution, and support regulatory auditability. The paper is structured as follows: Section 2 reviews existing literature on financial forecasting, explainable AI techniques, and data engineering practices. Section 3 discusses the unique challenges of engineering data for explainability in financial contexts. Section 4 presents our proposed framework and its integration with SHAP-based explanations. Section 5 provides a case study on stock market prediction using a transparent data pipeline. Section 6 evaluates the system on accuracy, interpretability, and performance metrics. Section 7 discusses implications, trade-offs, and generalizability, followed by concluding remarks and directions for future research in Section 8.

2. BACKGROUND AND LITERATURE REVIEW

2.1. Overview of Financial Forecasting Systems

Financial forecasting refers to the use of historical and real-time financial data to predict future economic conditions, asset prices, or corporate performance metrics. Traditional forecasting systems often relied on econometric models such as ARIMA, GARCH, and linear regressions, which assume statistical stationarity and limited feature complexity. In recent years, however, these models have increasingly been supplanted or augmented by machine learning (ML) and deep learning (DL) algorithms, including random forests, gradient boosting machines, and recurrent neural networks (RNNs), due to their ability to capture non-linear relationships and temporal dependencies in large datasets. These systems ingest multi-modal data ranging from time-series price movements to textual news feeds and macroeconomic indicators. Despite improvements in accuracy, these models often lack interpretability, especially in production environments, leading to challenges in trust, validation, and regulatory oversight. This evolution underscores the need to revisit how data is curated and presented for these models especially in the context of explainability.

2.2. Explainable AI Methods in Financial Domains

Explainable AI (XAI) encompasses a suite of techniques designed to make black-box models more transparent and their outputs more understandable to human stakeholders. In finance, popular model-agnostic approaches include SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-Agnostic Explanations), and Integrated Gradients. These methods provide insights into which features influenced a given prediction and to what extent. SHAP, in particular, has become widely adopted in finance due to its theoretical grounding in cooperative game theory and its ability to provide consistent, global and local feature attributions. However, the efficacy of XAI tools is often constrained by the quality and structure of the underlying data. Financial institutions have started embedding these tools in applications such as credit scoring, portfolio management, and fraud detection. Yet, challenges persist around interpretability across time

ISSN: 3071-3315
PP: 01-14.

horizons, consistency in dynamic data environments, and explaining model updates in real-time systems all of which require data engineering interventions.

2.3. Role of Data Engineering in AI/ML Pipelines

Data engineering forms the backbone of any AI/ML system, encompassing the processes of data ingestion, transformation, validation, enrichment, storage, and provisioning. In financial AI systems, this entails curating massive volumes of transactional data, integrating third-party feeds, and ensuring data compliance with regulatory frameworks. Crucially, for XAI to function correctly, these pipelines must preserve information about feature derivation, transformations, and temporal validity. For instance, if a model explanation reveals that "feature X" significantly impacted a forecast, it is essential to know how "feature X" was derived, whether it was lagged, normalized, or aggregated, and from which source it originated. Engineering this transparency requires pipeline components that support metadata tracking, feature lineage, and audit logging. Moreover, reproducibility—a key requirement in regulated industries—depends on the version control of both data and pipeline configurations. Thus, data engineering does not merely support ML; it governs the fidelity and usability of explainability in production systems.

2.4. Gaps in Current Research

Despite the proliferation of XAI tools and the growing focus on responsible AI, existing research often treats explainability as a post-modeling step, isolated from the data lifecycle. Few studies investigate how upstream data engineering decisions—such as pipeline architecture, feature encoding strategies, and data storage formats—affect the interpretability of downstream models. This lack of integration between data engineering and XAI introduces risks of explanation inconsistency, information loss, and audit failure in real-world deployments. Moreover, much of the current literature focuses on benchmark datasets or simplified environments, overlooking the complexities of live financial systems, which involve high-frequency data, evolving schema, and dynamic feature sets. There is a clear research gap in proposing and validating end-to-end frameworks that embed explainability considerations directly into the data pipeline. This paper addresses this gap by foregrounding the role of data engineering in shaping the transparency and accountability of AI-driven financial forecasting systems.

3. CHALLENGES IN DATA ENGINEERING FOR EXPLAINABLE FINANCIAL AI

3.1. Data Heterogeneity in Financial Systems (Market Data, Transactions, News, etc.)

Financial systems are inherently complex and data-rich, generating vast volumes of heterogeneous data from diverse sources such as stock market feeds, economic indicators, customer transaction logs, social media sentiment, and financial news articles. These data streams differ significantly in format, velocity, structure, and semantics. For instance, while market tick data arrives in high-frequency numerical form, regulatory filings are often semi-structured documents, and news sentiment is derived from unstructured text. This heterogeneity poses a major challenge for data

ISSN: 3071-3315
PP: 01-14.

engineering pipelines designed to support explainable AI. Standardizing disparate data formats into a unified schema suitable for machine learning models often requires intricate transformation logic and feature synthesis. More critically, when integrating such varied data sources, ensuring that the relationships and provenance of engineered features are preserved becomes essential for valid interpretation. Without a consistent representation and alignment of heterogeneous data, XAI outputs may become misleading, making it difficult to attribute predictions to meaningful real-world events or signals.

3.2. Data Quality and Labeling Issues

The effectiveness of both machine learning models and explainability tools is fundamentally tied to the quality of the data used in training and inference. In financial forecasting, data quality issues such as missing values, duplicate records, outliers, or incorrect timestamps can severely compromise model performance and explanation fidelity. Furthermore, financial data often contains noisy or delayed updates, particularly in transaction logs and third-party feeds. Labeling is another significant challenge especially in supervised learning contexts where ground truth is either delayed, ambiguous, or unavailable. For example, predicting future stock trends involves labels that are affected by future events not observable at the time of prediction. Poorly labeled data not only misguides model learning but also undermines the trustworthiness of post hoc explanations, as the interpretability tools may justify patterns that are artifacts of data noise rather than true signals. Addressing these quality and labeling challenges requires robust data engineering practices such as anomaly detection during ingestion, schema validation, temporal alignment, and label reconciliation mechanisms across distributed data systems.

3.3. Feature Lineage and Traceability

Explainable AI methods rely on the ability to attribute model predictions to specific input features; however, in most real-world financial AI systems, those input features are not raw they are the result of complex transformations across multiple pipeline stages. These transformations may include aggregation (e.g., moving averages), normalization, categorical encoding, temporal lags, or even feature extraction from textual or graphical content. Without proper lineage tracking, it becomes virtually impossible to trace a model's output back to its original data source or to understand how a particular feature was constructed. This lack of feature traceability undermines both interpretability and reproducibility, making it difficult to conduct model audits, root-cause analyses, or compliance reviews. Feature lineage systems must record not only the feature's derivation logic but also its dependencies, version history, and contextual metadata (e.g., which stock exchange a price feed originated from or how sentiment scores were derived). Integrating this into the data engineering pipeline is essential for building an explanation framework that is defensible, transparent, and audit-ready.

ISSN: 3071-3315
PP: 01-14.

3.4. Real-Time Data Constraints and Pipeline Latency

In financial forecasting, especially in domains like algorithmic trading or risk management, decisions must often be made within milliseconds of data availability. Supporting such real-time or near-real-time performance demands low-latency data pipelines that can ingest, transform, and feed data to models continuously and without delays. However, the integration of explainability mechanisms into these pipelines introduces additional complexity. XAI methods such as SHAP or LIME are computationally expensive and not optimized for high-throughput, low-latency environments. Moreover, real-time data often lacks the temporal stability required for consistent explanations features may fluctuate rapidly, and model behavior can change in response to evolving data patterns. Balancing speed with interpretability thus becomes a significant engineering trade-off. To mitigate this, data engineers must implement pipeline architectures that support asynchronous explanation computation, real-time feature monitoring, and caching of intermediate representations. The challenge lies in preserving explanation accuracy without degrading system performance, especially in latency-sensitive financial applications.

3.5. Regulatory Compliance (e.g., GDPR, Basel III)

Regulatory frameworks in the financial sector increasingly emphasize the need for transparency, accountability, and explainability in AI-driven decision-making. Regulations such as the European Union's General Data Protection Regulation (GDPR) and financial industry-specific mandates like Basel III or MiFID II introduce legal obligations for institutions to provide meaningful explanations for automated decisions. GDPR's "right to explanation" implies that individuals affected by algorithmic decisions such as loan rejections or investment allocations must be able to understand how and why those decisions were made. This regulatory pressure necessitates the implementation of robust data governance and lineage systems throughout the AI lifecycle. Data engineering pipelines must therefore embed compliance checks, maintain audit trails, and facilitate traceable mappings from input data to model output. Furthermore, sensitive financial data must be handled with strict privacy controls, including data minimization, anonymization, and access monitoring. Achieving explainability within this regulatory landscape requires that data engineering practices not only enable technical interpretability but also align with legal and ethical standards.

4. PROPOSED FRAMEWORK

4.1. Architecture of the Proposed Data Engineering Pipeline

The proposed framework is designed as a modular and scalable data engineering pipeline that prioritizes explainability in financial forecasting systems. At its core, the architecture consists of four primary layers: data acquisition, processing and transformation, feature engineering and metadata management, and model serving with integrated explainability. These layers are orchestrated using a microservices-based architecture that facilitates decoupled development and monitoring of each component. The data flows through various services via event-driven or streaming paradigms (e.g., Kafka or Apache Pulsar), which ensures real-time responsiveness and

ISSN: 3071-3315
PP: 01-14.

resilience. A dedicated orchestration layer, implemented using tools like Apache Airflow or Prefect, governs the scheduling, lineage tracking, and inter-service dependencies. Unlike conventional pipelines focused solely on performance or throughput, this architecture embeds observability and traceability into every component, ensuring that data transformations are transparent and auditable. The pipeline supports bidirectional integration with model repositories and XAI services, thereby creating a closed-loop system where data and explanations co-evolve in real time.

4.2. Data Ingestion (Structured/Unstructured)

The ingestion layer is designed to handle both structured and unstructured data, which is essential for financial forecasting systems that rely on diverse sources. Structured data, such as market feeds, economic indicators, and transactional logs, are ingested using message brokers or batch-based connectors (e.g., JDBC, Apache Nifi, AWS Glue). Unstructured data, including financial news articles, analyst reports, and social media feeds, are collected via web scraping, APIs, or streaming text sources and processed using natural language processing (NLP) pipelines. The ingestion layer normalizes these sources into a unified data lake or lakehouse architecture, allowing downstream components to query and transform data consistently. Importantly, ingestion modules apply tagging and initial metadata annotation to ensure that the origin, frequency, and format of the data are preserved. This early tagging is critical for downstream traceability and ultimately supports the interpretability of the features extracted from these raw inputs.

4.3. Feature Engineering with Explainability Constraints

Feature engineering is the most pivotal stage in the pipeline when aiming to support explainability. In this framework, features are not only optimized for predictive performance but are also designed with constraints that facilitate human interpretability and causal analysis. For instance, complex polynomial features or latent embeddings are only introduced if they are paired with interpretable summaries or explanations. Time-lagged variables, ratios, rolling aggregates, and domain-driven indicators (e.g., moving averages, Bollinger bands, sentiment polarity scores) are preferred because of their intuitive meaning and relevance to financial experts. Feature generation modules log every transformation step, including aggregation windows, filtering criteria, and encoding strategies. This structured logging allows XAI methods to relate output explanations back to the specific transformations applied to raw data. Moreover, the pipeline enforces version control over feature sets so that model retraining and explanation re-generation can be performed reproducibly, even after the underlying data changes over time.

4.4. Metadata Management and Feature Provenance

Metadata management is integrated across all layers of the pipeline and serves as the backbone of explainability. Each dataset and feature generated within the pipeline is annotated with rich metadata including source information, schema version, update frequency, and transformation lineage. A centralized metadata repository, built using tools like Apache Atlas or Amundsen, enables

ISSN: 3071-3315
PP: 01-14.

efficient querying and visualization of data and feature provenance. This system makes it possible to answer critical questions such as: where did this feature originate, what transformations were applied, and which model version used it? Additionally, the metadata repository interfaces with data cataloging and access control systems to ensure data governance. Feature provenance is further reinforced through lineage graphs that show the entire journey from raw data to model input, which can be programmatically accessed by both developers and auditors. This capability not only supports transparency but also enables regulatory audit trails, facilitates debugging, and allows explanation tools to generate context-aware interpretations.

4.5. Integration with XAI Toolkits

To bridge the gap between model predictions and human understanding, the proposed framework includes native support for integrating explainable AI toolkits such as SHAP, LIME, and Integrated Gradients. These tools are embedded into the model-serving layer and can be triggered automatically with each inference request or in an asynchronous post-processing mode, depending on latency requirements. The integration is bi-directional: models send prediction contexts to the XAI module, and the explanations generated are fed back into downstream visualization dashboards or compliance reporting systems. To ensure that explanations remain meaningful, the pipeline enforces consistency between the features used in the model and those visible to the explanation tool—this includes maintaining feature ordering, normalization parameters, and data type consistency. Additionally, the explanation outputs themselves are stored with metadata tags, allowing them to be queried historically and compared across model versions. This integrated approach ensures that explainability is not an afterthought but a continuous, versioned, and reproducible component of the AI lifecycle.

4.6. Design Principles: Traceability, Reproducibility, Transparency

The entire pipeline is governed by three foundational design principles: traceability, reproducibility, and transparency. Traceability ensures that every data point, feature, and prediction can be traced back to its source and transformation history, which is essential for debugging, compliance, and ethical accountability. This is achieved through automated lineage tracking and immutable logs of transformations and metadata. Reproducibility guarantees that any experiment, training process, or model inference can be repeated under the same conditions and produce consistent outputs. This involves strict version control of datasets, feature sets, model parameters, and explanation configurations. Transparency refers to the accessibility and intelligibility of both the data and the model behavior to stakeholders, including non-technical users. This is supported through human-readable feature documentation, visualization of explanation results, and clear documentation of pipeline stages. Together, these principles form the philosophical and technical foundation of a data engineering pipeline that is not only performant but also explainable, ethical, and trustworthy for financial decision-making.

ISSN: 3071-3315
PP: 01-14.

5. CASE STUDY: FORECASTING STOCK MARKET MOVEMENTS

5.1. Dataset: Stock Prices, Sentiment Data, Financial Indicators

For the case study, we utilize a composite dataset comprising historical stock prices, market sentiment data, and fundamental financial indicators to simulate a real-world financial forecasting environment. The stock price data includes daily open, high, low, close (OHLC) prices, and trading volume from publicly traded companies listed on major exchanges such as NASDAQ and NYSE over a 5-year window. To complement the price signals, we integrate sentiment scores derived from financial news articles, analyst reports, and social media (e.g., Twitter, Reddit), using a custom NLP pipeline that assigns polarity and subjectivity measures to each text input. Financial indicators such as earnings per share (EPS), debt-to-equity ratios, and price-to-earnings (P/E) ratios are sourced from quarterly filings and structured into time-aligned series with the price data. This combination of structured numerical data and semi-structured sentiment data provides a rich feature space for both model training and interpretation. The heterogeneity of the dataset also reflects the complexities that real-world financial systems must handle and makes it a suitable testbed for the proposed data engineering and explainability framework.

5.2. Model: XGBoost or LSTM + SHAP Explanations

The predictive modeling component of this case study employs two types of models: a gradient-boosted decision tree (XGBoost) and a Long Short-Term Memory (LSTM) neural network, both widely used in financial time-series forecasting. XGBoost is chosen for its ability to handle non-linear relationships and sparse features, as well as its compatibility with SHAP (SHapley Additive exPlanations), which offers theoretically grounded, consistent feature attribution. The LSTM, on the other hand, is used to evaluate the framework's performance on sequential data where temporal dependencies are critical. While the LSTM architecture is more complex and less inherently interpretable, SHAP's deep explainer and temporal aggregation techniques allow us to generate meaningful interpretations from the hidden state representations. Both models are trained on rolling windows of historical data to predict short-term stock price movements (e.g., 1-day or 5-day forward returns). Importantly, SHAP is integrated as a native module within the pipeline, enabling both local (per-instance) and global (aggregated) explanations to be generated in real time as new data is processed.

5.3. Data Pipeline: Steps from Raw Data to Explainable Predictions

The data pipeline implemented in this case study begins with real-time ingestion of stock price feeds, financial news streams, and company filings, which are harmonized and transformed into a unified schema. Textual sentiment is extracted using a fine-tuned BERT-based model, and all data sources are timestamp-aligned to ensure temporal consistency. The feature engineering module creates interpretable variables such as rolling averages, volatility bands, sentiment momentum, and earnings surprise metrics, all of which are logged and versioned through the metadata management system. These features are fed into the selected model (XGBoost or LSTM), and predictions are

ISSN: 3071-3315
PP: 01-14.

generated for future stock price direction. The SHAP engine is then triggered to compute feature attributions, which are stored alongside the predictions. Explanations are pushed to a dashboard interface where analysts can view not just the model's forecast but also the key contributing features and their magnitudes, making the prediction process transparent and auditable. The pipeline ensures end-to-end traceability by maintaining logs of data sources, transformation scripts, model parameters, and explanation configurations for each prediction instance.

5.4. Visualization of Feature Contributions

To make the model's decision-making process interpretable to human analysts, SHAP-based visualizations are integrated into the output layer of the forecasting system. These include local force plots, which illustrate the contribution of each feature to a specific prediction, and global bar plots that rank features by average absolute SHAP values across the dataset. For time-series models like LSTM, we use temporal SHAP summaries that show how a feature's importance evolves over time, offering insights into dynamic dependencies. Additionally, we generate clustered heatmaps to reveal correlations between features and explanations, enabling analysts to detect potential redundancies or misleading signals. These visualizations are accessible through a web-based dashboard that supports filtering by time period, asset class, and model version. This graphical interface transforms the raw numerical output of SHAP into actionable financial intelligence, helping portfolio managers, auditors, and regulators understand the rationale behind each forecast and how data inputs drive model behavior.

6. EVALUATION AND RESULTS

6.1. Explainability Metrics (e.g., Feature Importance Stability, Sparsity, Fidelity)

To rigorously assess the quality of model interpretability, we evaluate the system using several explainability-specific metrics. Feature importance stability is measured by assessing how consistently SHAP identifies top contributing features across different temporal slices and training runs. High stability implies that the model's behavior is consistent and reliable a critical property for financial applications. Sparsity is quantified by calculating the average number of features that account for a given threshold of total SHAP value (e.g., 90%), which reflects how focused or diffuse the model's decision logic is. Models with sparse explanations are generally easier to interpret and more robust to noise. Fidelity is evaluated by comparing the predictive performance of surrogate models trained on SHAP explanations to the original model's outputs, with high fidelity indicating that explanations accurately represent model behavior. These metrics are computed for both XGBoost and LSTM variants to benchmark how explainability behaves across different modeling paradigms.

6.2. Forecasting Performance (RMSE, MAE, R^2)

Beyond explainability, the predictive accuracy of the forecasting models is evaluated using standard regression metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics are computed on a rolling forecast basis using

ISSN: 3071-3315
PP: 01-14.

a test set comprised of out-of-sample market data. The XGBoost model achieves relatively lower RMSE and MAE scores due to its capability to handle non-linearities and interactions among engineered features. The LSTM model, while slightly less performant in some static tests, exhibits stronger adaptability during periods of high market volatility due to its temporal memory structure. Importantly, the integration of interpretable feature engineering improves overall model performance compared to raw data inputs, reinforcing the hypothesis that explainability-oriented data engineering can benefit both transparency and predictive capability. Performance results are also segmented by sector (e.g., tech, healthcare, finance) to assess generalizability across asset classes.

6.3. System Efficiency (Latency, Throughput)

System efficiency is a critical consideration for real-time financial forecasting systems. Latency is measured as the time elapsed between data ingestion and the generation of a complete, explained prediction, while throughput is defined as the number of predictions processed per second under typical load. The system achieves sub-second latency for XGBoost-based predictions, including SHAP computation, owing to efficient parallel processing and pre-computed model states. For LSTM models, latency increases modestly due to the complexity of deep learning inference and explanation generation, but remains within acceptable bounds for use in tactical investment scenarios. Throughput benchmarks show that the pipeline can handle thousands of predictions per minute when deployed on a distributed cloud environment with GPU support. These performance characteristics validate the feasibility of integrating explainability into real-time systems without compromising operational efficiency, a major advancement over conventional batch-mode XAI frameworks.

6.4. Regulatory Auditability and Interpretability Scores

To measure compliance readiness, we assess the system's auditability and interpretability using qualitative and quantitative criteria aligned with financial regulations such as GDPR's "right to explanation" and Basel III's model risk management guidelines. Auditability is validated through simulated audits where external reviewers trace model predictions back through the data pipeline, verifying inputs, transformations, and SHAP-generated explanations. All stages pass reproducibility checks, and version histories are preserved in immutable logs. Interpretability is quantified using structured expert evaluations wherein financial analysts rate the clarity, completeness, and actionability of the explanations on a Likert scale. The system scores consistently above industry benchmarks, particularly in clarity and completeness, confirming that the integration of engineered features, metadata management, and SHAP explanations leads to interpretable and justifiable forecasts. These evaluations affirm that the framework not only enhances model transparency but also meets the interpretability standards required for responsible AI adoption in regulated financial domains.

ISSN: 3071-3315
PP: 01-14.

7. DISCUSSION

7.1. Trade-offs: Performance vs Explainability vs Engineering Cost

A central theme in this study is the careful balance between model performance, interpretability, and the engineering resources required to support both. High-performing models, such as deep neural networks, often come at the cost of reduced transparency, which can hinder stakeholder trust and regulatory compliance. While simpler models are inherently more interpretable, they may not capture the intricate patterns present in high-dimensional financial datasets. Our integration of explainability into complex models like XGBoost and LSTM illustrates that it is possible to achieve a middle ground—where sufficient model performance can be retained without sacrificing interpretability. However, enabling this trade-off requires substantial investment in data engineering: from sophisticated metadata tracking and lineage mapping to maintaining reproducible feature sets and scalable pipeline orchestration. These engineering efforts incur not only financial costs but also operational complexity and the need for interdisciplinary collaboration among data engineers, data scientists, domain experts, and compliance officers. Therefore, the decision to implement explainability features must be guided by a clear understanding of business priorities, regulatory obligations, and resource availability. In mission-critical applications such as algorithmic trading or investment risk management, the marginal cost of explainability is often justified by the increase in trust, reliability, and auditability of model decisions.

7.2. Adaptability to Other Financial Domains (Risk Scoring, Fraud Detection)

Although the case study focuses on stock market forecasting, the proposed framework exhibits strong adaptability to other domains within financial services, such as credit risk scoring, fraud detection, insurance claim evaluation, and anti-money laundering (AML) systems. These applications also rely heavily on predictive analytics but have distinct requirements in terms of input data, model explainability, and compliance. For instance, in credit risk scoring, explainability is not optional—it is a regulatory requirement under frameworks like the Fair Credit Reporting Act (FCRA) and Basel II/III. The data pipeline described in this paper can be extended to include credit bureau records, income statements, and behavioral data from mobile apps, all engineered with traceability and transparency in mind. Similarly, in fraud detection, where anomaly detection models must justify rare and high-impact decisions, integrating real-time SHAP explanations can help operational teams and compliance officers quickly understand suspicious transactions and take appropriate action. Because the pipeline architecture is modular, with standardized ingestion, transformation, metadata, and explanation components, it can be readily adapted to fit different data modalities and regulatory contexts. This flexibility underscores the framework's value as a foundational infrastructure for explainable AI across a broad spectrum of financial applications.

7.3. Alignment with Regulatory and Ethical AI Guidelines

As financial institutions increasingly integrate AI into decision-making processes, the alignment of these technologies with regulatory and ethical frameworks becomes imperative.

ISSN: 3071-3315
PP: 01-14.

Regulatory bodies such as the European Banking Authority (EBA), U.S. Securities and Exchange Commission (SEC), and global financial watchdogs now emphasize model transparency, accountability, and governance as foundational principles for AI adoption. This paper's proposed framework directly supports these goals by embedding explainability throughout the data engineering lifecycle. By preserving feature provenance, maintaining immutable logs, and integrating post hoc explanation tools, the system provides the structural support necessary to fulfill audit requests, conduct internal risk assessments, and meet consumer protection mandates such as the GDPR's "right to explanation." Beyond compliance, the framework also aligns with ethical AI principles, including fairness, accountability, and transparency (FAT). For example, by documenting feature transformations and ensuring that model behavior is interpretable, institutions can proactively detect and correct discriminatory patterns or biases that may emerge in training data. Moreover, the use of transparent and well-documented models enhances user trust and promotes responsible innovation. As regulatory and ethical requirements continue to evolve, frameworks like the one presented in this paper offer a proactive and technically sound path toward sustainable and trustworthy AI systems in the financial sector.

8. CONCLUSION AND FUTURE WORK

This paper presents a comprehensive framework for engineering data pipelines that support explainable artificial intelligence (XAI) in financial forecasting systems, addressing the pressing demand for transparency, traceability, and accountability in high-stakes decision-making environments. Through a detailed architectural proposal and a case study on stock market prediction using SHAP explanations, we demonstrate how data engineering practices can be structured to maintain metadata lineage, enable real-time feature transparency, and seamlessly integrate XAI toolkits into model workflows. Our study contributes to the growing body of literature by emphasizing the foundational role of data quality, feature provenance, and pipeline observability in explainable machine learning (ML) deployment, especially in regulated financial contexts. The proposed system is designed with scalability and modularity in mind, making it adaptable to other domains such as credit risk assessment and fraud detection. However, the implementation is not without limitations—particularly in terms of computational overhead, the complexity of pipeline orchestration, and the trade-offs between model complexity and interpretability. Additionally, while tools like SHAP and LIME provide useful local explanations, they often require careful tuning and may not generalize well across temporal or multimodal datasets. Future work should explore the integration of automated XAI (AutoXAI) pipelines that dynamically select and calibrate explanation methods based on the data type and model behavior. Further, the emergence of multimodal financial data including social media sentiment, satellite imagery, and IoT sensor data necessitates the development of unified data representations that retain interpretability across modalities. Another promising direction involves extending the framework to support edge deployment in scenarios requiring low-latency decision-making, such as high-frequency trading or mobile-based financial advisory. These advancements, combined with continual alignment to evolving regulatory and

ISSN: 3071-3315
PP: 01-14.

ethical AI guidelines, will be essential to scale explainable AI in finance. Ultimately, we posit that investing in explainability-driven data engineering is not just a technical endeavor but a strategic imperative for financial institutions seeking to build AI systems that are not only accurate and efficient, but also auditable, fair, and trustworthy.

REFERENCES

[1] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.

[2] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD*, 1135–1144.

[3] Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.

[4] Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Interpretable machine learning: Definitions, methods, and applications. *PNAS*, 116(44), 22071–22080.

[5] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2019). A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR)*, 51(5), 1–42.

[6] Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138–52160.

[7] Amershi, S., Chickering, M., Drucker, S. M., Lee, B., Simard, P., & Suh, J. (2015). ModelTracker: Redesigning performance analysis tools for machine learning. *CHI*, 337–346.

[8] Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2), 2053951717743530.

[9] Breck, E., Polyzotis, N., Roy, S., Whang, S. E., & Zinkevich, M. (2019). Data validation for machine learning. *Proceedings of SysML Conference*.