International Journal of Advanced Engineering Science and Information Technology (IJAESIT)

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

AI-Powered Big Data Analytics for Early Detection of Financial Market Anomalies

Vinod Battapothu

Independent Researcher, India

ABSTRACT: The past decade has witnessed a proliferation of Big Data in the Finance industry. Several technologies are being developed in parallel, and a consolidated assortment of AI-based analytical products is emerging. The bulk of these Big Data-based projects within the finance domain are concerned with Investment Management. These services deal with the application of Machine Learning or Natural Language Processing on alternative datasets (e.g. Unstructured Data, Sentiment Analysis, Web traffic, etc.) with the aim of predicting Price Movements or even Stock Prices of Stocks, commodities, or different Asset Classes. Another set of companies are providing data-driven strategies that predict Financial Market Anomalies. But these innovative products do not appear to provide reliable and robust results for Early Detection of Financial Anomalies.

One technology that aims to contribute in this area is AI-Powered Anomaly Detection. Early detection of Financial anomalies (e.g. Flash Crashes, High Liquidity Drought, Opening Price Dislocation, Price Dislocation in the presence of Large Imbalance) has been attempted through traditional machine learning-based models. A natural extension of this effort is to employ AI-Powered Big Data Analytics for Early Detection of Financial AAnomalies. An end-to-end solution capable of ingesting Big Data in real time, processing with AI-Powered Technology, and signalling detections with Data Governance is developed and presented. Real-time Scalability and versioning of the data has also been examined.

KEYWORDS: Big Data, AI, anomaly detection, flash crash, liquidity droughts, arbitrage windows, qualitative trading, market microstructure, Novozhilov test, attention sub-network, temporal anomalies, finance market alerting.

I. INTRODUCTION

Early detection of anomalous behaviour in financial markets has the potential to improve decision-making by market participants and moderators. Market anomalies are commonly associated with price formation and liquidity but are not easily detected in historical data. Standard time series or regression-based techniques used in finance attempt to uncover such structures but do not lend themselves to real-time evaluations. AI methods, particularly in conjunction with Big Data, open new avenues.

Big Data is a general term for data management strategies that can accommodate extremely large volumes of data that are commonly known as 3Vs: volume, velocity, and variety. The three sectors of finance that create the greatest challenges for Big Data management are market transactions, banking records, and customers, though additional sectors such as financial news and social signals are also prevalent. Three special categories of Big Data technologies are relevant for market anomaly detection: (1) preparation and storage technologies such as NoSQL databases, (2) distributed processing and cleaning systems, and (3) engines for real-time alerts.

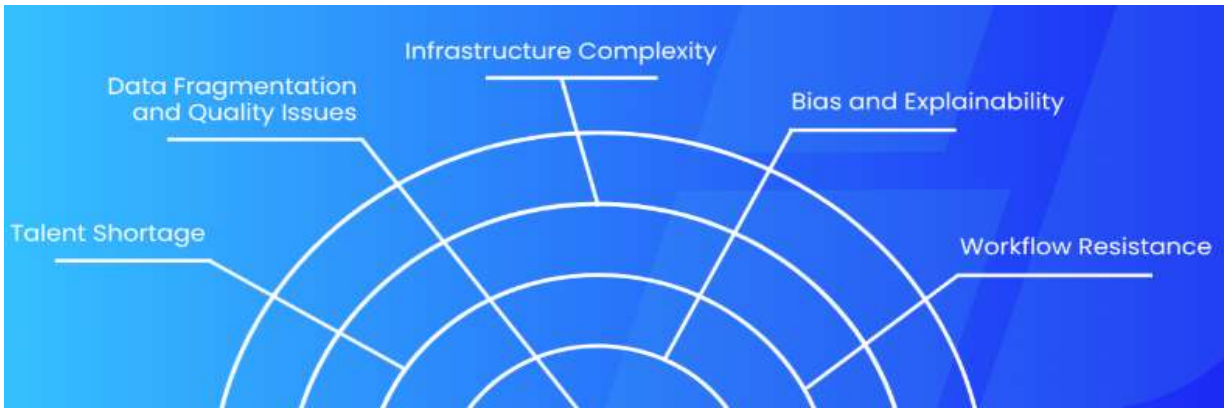


Fig 1: AI in Data Analytics

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

1.1. Background and Significance

AI-Powered Big Data Analytics for Early Detection of Financial Market Anomalies. Martin Wainstein, Inna Goychuk: Bank of Lithuania; Modestas Karpavicius: Vilnius University.

Early detection of anomalies that threaten the stability of financial markets is a critical element of both effective risk control and evasive measures. The proposed research introduces AI-powered Big Data techniques capable of signalling various financial market anomalies before they escalate into major crises. The introduction of the analytics process and the planned notification of the relevant parties is the main contribution of the proposed research. This contributes directly towards ensuring investor protection and the preservation of financial stability.

Anomaly detection aims to identify patterns that do not conform to expected behaviour. Markets are not different, and when prices start drifting from their equilibrium or common characteristics, the said pattern has an anomaly behaviour. In the context of market prices, temporal anomalies include liquidity droughts (long periods of low order book activity) and flash crashes (a sudden sudden price drop followed by a very fast recovery). These are momentary situations; however, a wide variety of data sources point to these conditions and their relationship with price formation, sentiment and liquidity. Using those signals, a detection mechanism can trigger an alert whenever an anomaly is detected.

Equation 1: Returns from price (standard preprocessing for anomaly detection)
Most financial anomaly detectors work on **returns** rather than raw prices.
(1) Simple return

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

(2) Log return (more stable for modeling)

$$\ell_t = \log(P_t) - \log(P_{t-1}) = \log\left(\frac{P_t}{P_{t-1}}\right)$$

1.2. Research design

For finance practice and policy, the development should help early detection of anomalies in price and liquidity formation in real time by combining a large variety of sources (order book, news, social signals, macro indicators) and using state-of-the-art AI methods with proven performance or explainable-equivalent results. The proper design of the entire end-to-end pipeline—from data ingestion to alert generation—will warrant practical deployment in a variety of contexts. The system will allow detection of a wide range of temporal patterns, either followed by the expected behavior and news, or no bullish-bearish price-driving signal at all, thus pointing to possible liquidity drought.

After examining flash crash cases and liquidity drought signals of major assets in the BTC-USD pair, results seem to suggest that the detection algorithm is capable of generating satisfactory early signals that coincide with a flash crash period or liquidity debilitating event. Most alerts seem to occur around 30 min before the event, which may offer an opportunity for traders to close positions or risk manage them through other instruments or measures. The low ratio of alarms detected in a 1 h window before the flash crash time seems to improve the quality of the alerts, even if the trade-off remains to be completely analyzed.

II. THEORETICAL FOUNDATIONS

Big Data analytics and machine learning have found success across diverse domains due to their ability to learn, recognize patterns, and make predictions from large datasets. Financial practice and research may similarly benefit from these approaches, yet analytics remain shallow, and machine learning techniques have yet to be widely adopted. Financial anomaly detection has promising practical applications but has received relatively little attention, particularly in the developing Asian markets. Early anomaly detection is critical for policymakers and financial regulators, who seek to prevent, mitigate, and respond to disasters such as flash crashes, liquidity droughts, and extreme population shifts. Finance even plays a proactive role, with violation opportunities rapidly relayed to the public. Simultaneously, market agents and firms identify and exploit bid-ask pricing differentials, presenting trading strategies that clear market inefficiencies. These firms possess deep learning capabilities and can promptly capitalise on new information for profit. Identifying and forecasting windows for pure price misalignments drawn from arbitrage is hence essential for retail market participants.

Recent advancements in Big Data technologies, encompassing data storage and sensors, high-volume, high-velocity data processing, and real-time fast-data analytics, now facilitate the automatic and high-precision detection and forecasting of financial market anomalies. Consequently, building a production-ready anomaly-detection solution for financial markets

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007

requires selecting the types of input data and monitoring logic, specifying the model used to detect the violation scenario, and evaluating the overall stability, fault tolerance, latency, scalability, and maintainability. Data availability, depth, and exploration are further enhanced by complementing traditional price, volume, and order book information with news sentiment, social signal intensity, and macroeconomic indicator variation.

2.1. Big Data Technologies in Finance

Big Data technologies have revolutionised all aspects of financial services, including the creation and support of new Financial Technology solutions. New architectures allow real-time processing of vast volumes of transactional and operational data, including traditional market data such as price, volume or order-book data and alternative signals from social media, news and macroeconomic indicators. The amalgamation of these developments allows for timely detection of new anomalies affecting price formation and market microstructure. These data sources and technologies enable the scaling up of all aspects of anomaly detection previously considered in isolation. Relatively simple supervised models can achieve promising performance on classical trades based on daily data, with a roadmap proposed for more complex deep-learning techniques that support detection of more subtle temporal anomalies.

Big Data technologies for Finance include functions for data storage (such as NoSQL databases), fast in-memory data processing engines, real-time automated stream processing (for example, Spark Streaming, Storm and Flink) and distributed processing of batch data (Hadoop Map Reduce, Spark). Distributed processing environments provide the governance, version control and data management required for the implementation of analytics pipelines that are scalable, fault tolerant and stable. Together with a controlled deployment process, they enable the rapid construction, testing and deployment of predictive models that can serve as inputs to real-time analytics.

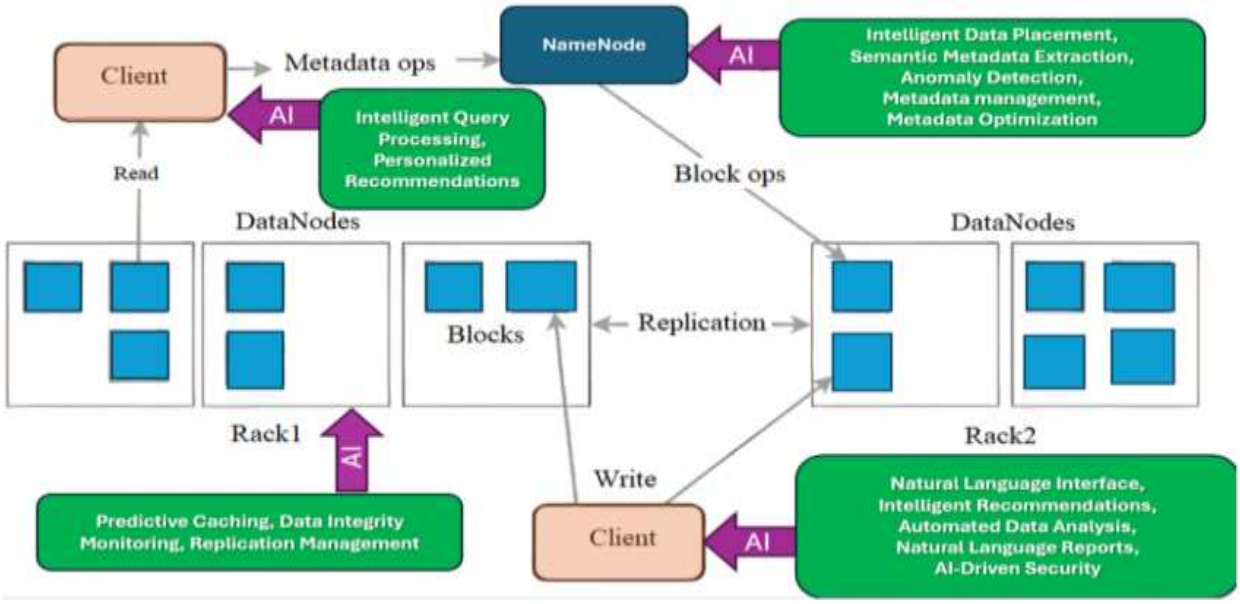


Fig 2: AI-Powered Evolution of Big Data

2.2. Anomaly Detection and Market Microstructure

Any deviation of market features or signals, such as price and liquidity, from their typical behaviour can be viewed as an anomaly. Three major types of anomalies can be distinguished: anomalies stemming from the accumulation of private information, anomalies affecting the microstructure of asset prices (such as flash crashes and liquidity droughts), and unexplained price changes that exhibit regular patterns in the order book (such as pricing errors and short arbitrage windows). Price formation and markets' cross-sectional or time-series properties across all compatible assets type should be able to detect deviations across signals. The change of price between two consecutive points in time must be transparent and follow an arbitrary statistical distribution. Otherwise, markets are providing sub-optimal pricing to their users, i.e., traders with varying time-horizons who are actively sending orders to the market. As a consequence of an inadequate ordering process, the main types of markets guarantees on price discovery and liquidity provision cannot be fulfilled. When market anomalies occur, these guarantees are not fulfilled. Price and order flow changes between consecutive time intervals must belong to a multivariate and multi-dimensional order book process that can accommodate different assets types and that can also control for variations across time zones, market cycles and news events. However, these periodic price corrections might, in turn, be exploiting wider order book mis-pricing relationships with the specific

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007

risk that the order book moves out of equilibrium and persists in such state for a time longer than its natural mean-reverting framework.

Equation 2: Rolling-average threshold for "volume spikes" (explicitly described)
Step 1: rolling mean
For a window W:

$$\mu_t = \frac{1}{W}\sum_{i=0}^{W-1} V_{t-i}$$

Step 2: rolling standard deviation
Use sample std over the same window:

$$\sigma_t = \sqrt{\frac{1}{W-1}\sum_{i=0}^{W-1}(V_{t-i}-\mu_t)^2}$$

Step 3: z-score (normalized spike magnitude)

$$z_t = \frac{V_t - \mu_t}{\sigma_t}$$

Step 4: alert rule

$$\text{Alert at } t \Leftrightarrow z_t > \tau$$

- Typical choices: $\tau \in [2,4]$.
- This matches the paper's idea of rolling threshold-based spike alerts

III. DATA SOURCES AND PREPROCESSING

Financial anomaly detection relies on price, volume, and order book data, complemented by alternative datasets such as news, social media signals, and macroeconomic indicators. Historical events may be mined to define an anomaly type's key features, but without innovative feature engineering and data fusion, data cleansing and normalization remain the main tasks. Addressing typical financial data issues, such as missing values and sudden feature distribution changes, is imperative. Evaluation is necessarily unsupervised, and therefore highly sensitive to noise; it should also account for the inherent difficulty in detecting rare events and consider operational overheads.

Data sources can be broadly categorized into primary and alternative datasets. Primary datasets include traditional price and volume trading data, order book data for liquidity modeling, and other types of data like news, social media signals, and macroeconomic indicators designed to capture sudden shifts in market sentiment. Although such sources cannot directly trigger alerts, they are regarded as potential early warning indicators. Anomaly detection in financial time series is further complicated by the presence of concept drift (i.e., sudden changes in the statistical properties of time series).

3.1. Market Data and Alternatives

Three main groups of data are exploited to detect market anomalies: standard market signals (i.e. price, trading volume, order book), alternative data geared towards mood assessment (e.g. news articles, Twitter messages, Reddit posts), and macro indicators describing the state of the global economy (e.g. money supply by central banks, unemployment rate, inflation). In the field of anomaly detection, however, two of the biggest challenges are data cleaning and normalization.

Before any detection can take place, Big Data governance concepts must come into play. Since all data sources are pulled from online repositories in real time, tracking and versioning are necessary to implement proper data lineage. Whenever an anomaly is detected and an alert is generated, historical data become a good set to work on. Cleaning involves the detection of outliers (excessive price or volume variations in a given time interval), missing values or wrongly encoded values. One common way to deal with missing values is to replace them with the average of the previous and following values. Another, more effective, method, is to apply a fitted spline based on neighbouring values. Since news articles and social signals often have a very different space and time scale when compared to price/volume/activity data, feature engineering is also important.

The risk of false positives can also be mitigated by explicitly dealing with concept drift. Some concepts are defined dynamically, depending on the underlying market conditions. For example, a flood of news articles or social signals about

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007

an event has little significance if the event was already known to the market participants. All analyses and alerts make this explicit.

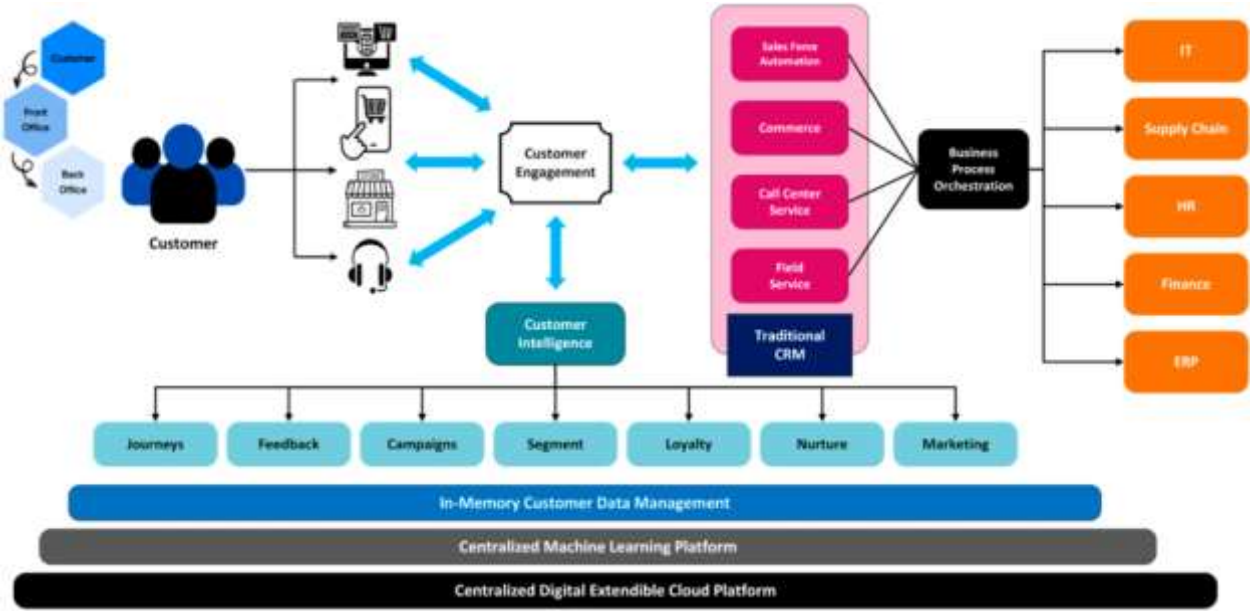


Fig 3: Market Data and Alternatives of AI-Powered Big Data Analytics

3.2. Data Cleaning, Normalization, and Feature Engineering

An understanding of market microstructure features and the careful definition of anomalies enable the integration of alternative data that are relevant for early detection. Major data sources for identifying events or conditions resulting in a large change of prices and/or an important loss of liquidity are the classical financial prices, traded volume and order book data of markets. Newsflow from various channels (news wires, agency publications, financial papers, social media) introduces additional information that has been shown to have a relation with the price dynamics through the crowd behavior of market participants.

Both Big Data methods and machine learning offer strong capabilities for clearsighted system work. Nevertheless, considerations on the cleaning and further treatment of the data are still extremely important for the whole detection process. In particular, the cleaning and normalizing steps are essential. The information available through these distinct channels can offer the possibility to detect alerts related to flash crashes or periods of severe liquidity drought. Thus, Big Data techniques should mitigate the well-known limitations of classical research in finance. Traditional econometric models perform poorly in the Big Data setting: they usually require prior steps of feature selection or the introduction of regularization, which are unsatisfactory for automatic detection of solutions.

Equation 3: Missing value handling equations (explicitly mentioned as methods)
(a) Neighbor average imputation (step-by-step)
If $x_t$is missing but $x_{t-1}$and x_{t+1}exist:

$$\hat{x}_t = \frac{x_{t-1} + x_{t+1}}{2}$$

(b) Spline imputation (conceptual equation)
Fit a spline $s(t)$on observed points $\{(t_i, x_{t_i})\}$, then:

$$\hat{x}_t = s(t)$$

IV. METHODOLOGICAL FRAMEWORK

Both unsupervised and supervised methods can be employed to identify anomalies in financial data, and the selection of suitable performance metrics depends on the intended application. In the context of financial data, unsupervised methods are often preferred, as labeled data may not be readily available and are usually obtained through time-consuming manual efforts. Successful implementations of unsupervised anomaly detection have been achieved in various diverse domains, particularly through deep learning approaches tailored to capturing temporal anomalies in time-series data.

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007

Temporal anomaly detection is a challenging problem that requires the development of methods capable of jointly analyzing the temporal dimension and multiple time-series. Many available temporal signal datasets are derived from image data sources and are often explicitly annotated to provide the ground truth for training and model evaluation. Financial market datasets differ significantly in nature, and the prospect of a labeled dataset is remote. Methodological proposals for deep learning-based temporal anomaly detection can be categorized along several axes: (i) architectural design; (ii) the representation of the input signal; (iii) the training regime; and (iv) the interpretability of the results.

4.1. Supervised and Unsupervised Anomaly Detection Methods

Abnormal events in the financial market can occur at any point and their detection remains an open issue. Recent studies have suggested that artificial intelligence can be applied to Big Data-based analytics for the early detection of anomalies. Anomalies may be detected using either supervised or unsupervised machine learning. The former requires labelled datasets of such events, which could lead to scarcity in real world because they are rare by definition. Unlike supervised learning where machine learning algorithms learn from a train dataset containing both normal and abnormal samples of the said signal, unsupervised learning systems output high dimensionality score without the presence of labelled samples and have gained acceptance recently especially in financial domain.

The first step of employing supervised machine learning for financial market anomaly detection is to identify an anomaly score in the data. These scores can then be used to determine the performance of the process. All performance measures are calculated from the array of performance scores on the training dataset. Evaluation metrics generally acknowledge that as long as the temporal dimension of the problem remains, the machine-learning methods that have been proposed can potentially be adapted both in the supervised as well as unsupervised settings. Alternatively, a signal processing perspective can also be adopted for temporal anomaly detection. In this case, the set of signals are generated from Multiple Matrix Factorization and used as features by a classifier system (e.g. Random Forest) when supervised learning are under consideration. These signal functions are crafted so that they can be treated as oracle score functions in supervised approach.

Deep learning-based approaches are well suited for tackling temporal anomalies, such as LSTM, AutoEncoder, Convolutional Neural Networks and based on Temporal Convolutional Networks. The core idea is to present a temporal sequence to the learning algorithm as input and let the network learns to discriminate between the normal temporal states by minimizing some loss function and outputting an associated labeling score. Different modalities of inputs are also possible: in some previous works temporal image patches are constructed and fed to standard classifiers, Temporal-exploration Visual Sentiment Analysis presents a video sequence as a short spatio-temporal video clips and builds a new special-temporal context content explanation to capture distinctive semantics of both content and context issued in their detection work, and Dense-Trajectories clusters ST-trajectory of different activity categories into one cluster to help capture and discriminate different activities better.

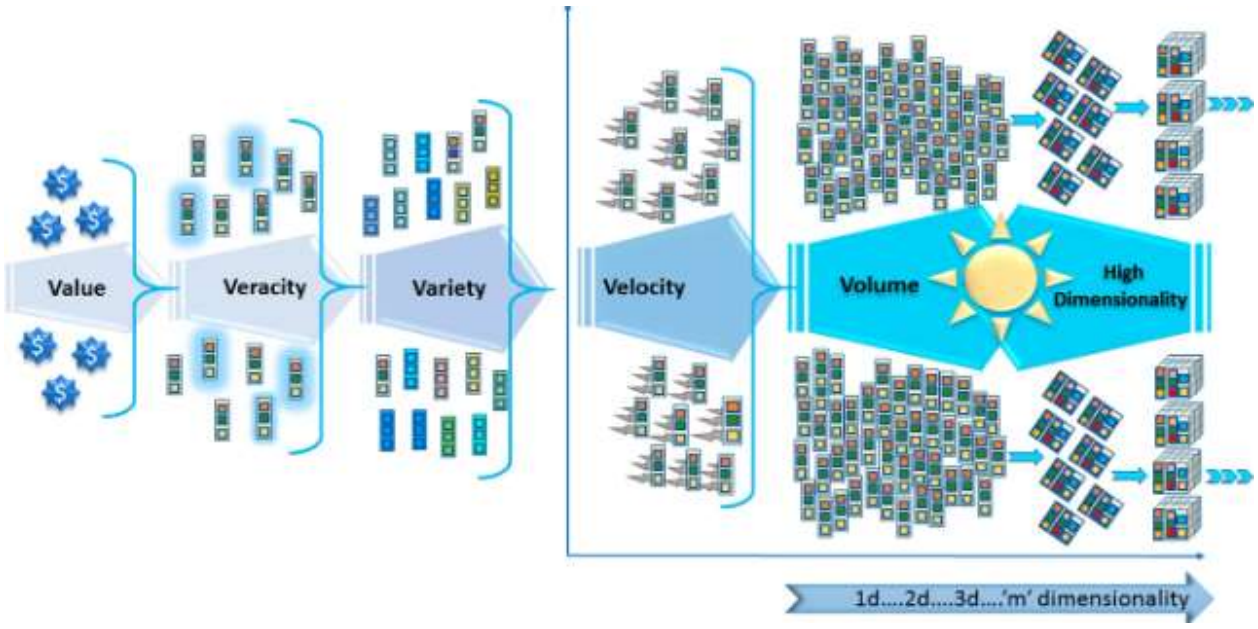


Fig 4: Supervised and Unsupervised Anomaly Detection

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

4.2. Deep Learning Approaches for Temporal Anomalies

A variety of supervised and unsupervised methods from machine learning and statistics have been proposed for the detection of anomalies. Supervised methods require the presence of anomaly labels in the training dataset. Such labels can be difficult, costly, and time-consuming to obtain. As a consequence, unsupervised methods are more widely adopted. However, classical unsupervised techniques typically cannot leverage multi-dimensional input data, nor take advantage of complex feature extraction methods employed for other deep learning tasks.

Another important consideration for anomaly detection is the presence of a temporal dimension in the data. A portion of the input features — and/or other auxiliary data — stays constant across the observation window. This includes, for example, price and volume features generated in the context of anomaly detection; news mentions or social signals covering the same time window; and macroeconomic indicators that are fixed during the same time period. Moreover, different sensors (i.e., nodes in sensor networks) usually operate in a similar environment, and share characteristics such as the same ground truth distributions. In these cases, temporal information can be leveraged to improve anomaly detection through the application of deep learning architectures designed for temporal data, such as recurrent neural networks (RNNs) or temporal convolutional networks (TCNs).

Equation 4: One-Class SVM anomaly detection (paper lists it)
Step 1: primal optimization problem
One-class SVM finds a boundary that encloses "normal" data:

$$\min_{w,\rho,\xi} \frac{1}{2} \| w \|^2 + \frac{1}{\nu n} \sum_{i=1}^{n} \xi_i - \rho$$

subject to

$$w^\top \phi(x_i) \geq \rho - \xi_i, \xi_i \geq 0$$

- $\phi(\cdot)$maps into kernel feature space.
- $\nu \in (0,1]$controls the expected outlier fraction.

Step 2: decision function
After solving, the score is:

$$f(x) = w^\top \phi(x) - \rho$$

Step 3: anomaly rule

$$\text{Anomaly} \Leftrightarrow f(x) < 0$$

V. SYSTEM ARCHITECTURE AND DEPLOYMENT

The proposed solution integrates commonly used open-source technologies and leverages AI-powered algorithms to provide continuous market monitoring and timely alerts. The entire end-to-end analytics pipeline—from data ingestion and preparation to feature engineering, detection, and alert generation—can be operated directly from a laptop using a cloud-based Virtual Machine. A suitable combination of tools for production deployment can be selected based on the specific constraints of the target environment. The focus should be on providing a fault-tolerant solution capable of meeting requested latency targets and secured against potential threats.

Anomaly detection must remain a priority at every stage of the pipeline, with emphasis on data governance and versioning. Incoming market data should be governed using a Datalake architecture, ensuring anomaly-free, well-structured content. Supervised detection methods can trigger alerts to human operators requesting further analysis of detected patterns. Operational readiness of a production-grade and truly automated high-throughput solution has not yet been achieved; it would require deeper data cleaning, enhancement with alternative data sources, and adaptation of detection methods to better match financial data characteristics.

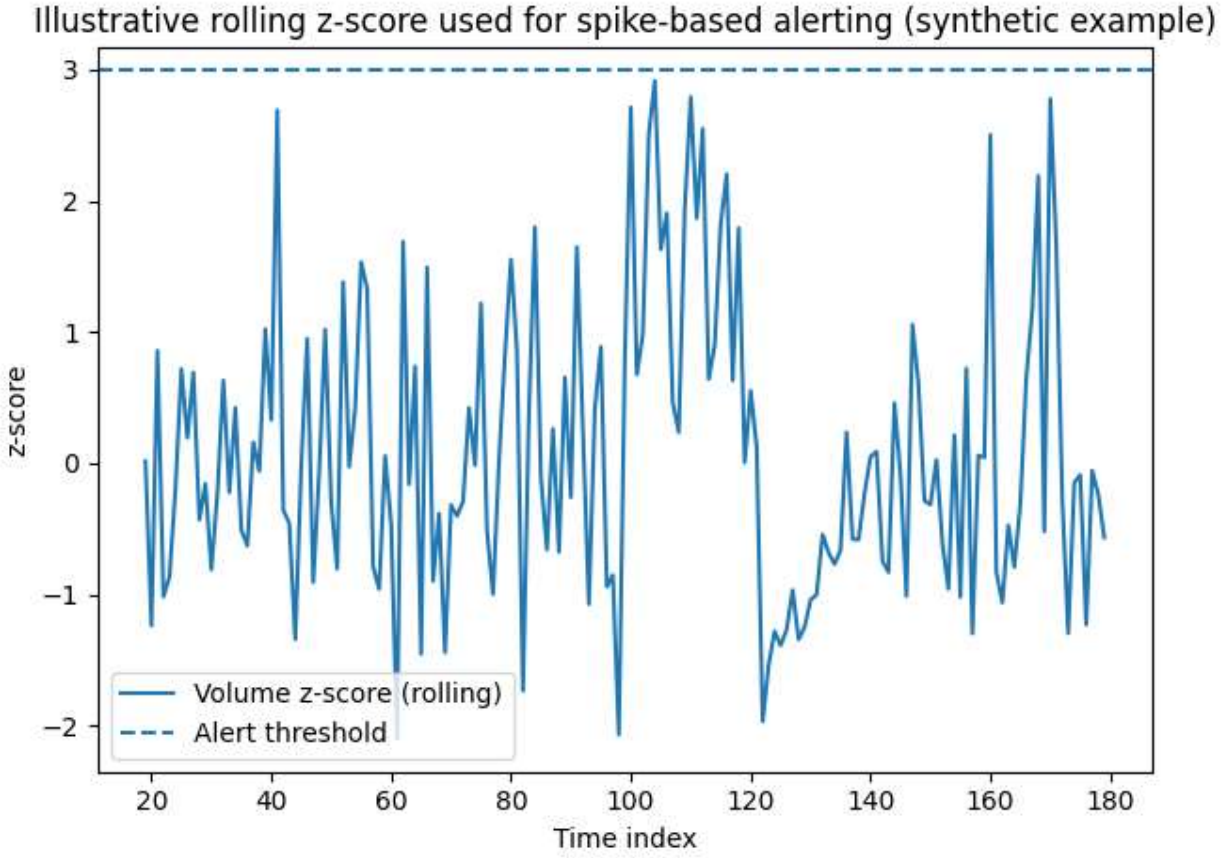
5.1. End-to-End Analytics Pipeline

An end-to-end analytics pipeline encompasses all phases of market data management—from ingestion to alert delivery. Pipelines for systems performing cross-sectional analyses typically have a daily ingest schedule, while those designed to capture temporal anomalies in real time or near real time require specialized architecture and techniques. Thus, it is essential to consider all aspects of the analytics pipeline: In addition to optimizing computational efficiency and ensuring a high degree of available redundancy, pipeline design should also account for industry-grade requirements such as fault tolerance, latency budgets, data governance (including versioning and access control), and deployment options (on-

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007

premises versus cloud). Community support for a pipeline built on mainstream open-source technologies usually reduces long-term maintenance costs.

The use case discussed in Section 6 centers on a real-time engineering and analytics assembly line capable of ingesting financial data from a broad range of market sources (including transaction-level information) accelerated by current Big-Data frameworks. A horizontal overview of this assembly line highlights supporting elevators for preparation and processing, before leading into a prominent vertical channel designed to support elastic, low-latency processing. The analysis leverages Big-Data technologies for data storage and processing, Apache Spark's MLlib library for temporal anomaly detection, and Apache Kafka for implementing an ondemand serving infrastructure. Finally, production rules formalizing service-level expectations for latency, availability, and fault tolerance are presented; these rules underscore the need for nontrivial architectural decisions and development trade-offs, particularly outside the core analytics component itself.



5.2.Scalability, Reliability, and Latency Considerations

A Big Data analytics pipeline must be designed for scale, fault tolerance, and low-latency operation. Given that driving the analysis from a large-scale production operation in a financial knowledge-oriented company is not only a research project but also a business-critical facility, the final architecture must adhere to production quality attributes within the data analysis chain, particularly when detecting anomalies in near real-time.

Anomalous behavior—flash crashes, extreme liquidity droughts, spans with high arbi-risk for arbitrageur windows, and so forth—requires near-real-time detection and alert generation. Response times must be commensurate with remediative actions taken by market participants in reaction to the alerts. In an AWS cloud deployment, all components—mainly AWS S3 for data ingestion and AWS Lambda functions for pre-processing steps and alert generation—can be triggered within an elapsed time of a few seconds once the data governance and management processes governing versioning and activation of the relevant processing chain are fulfilled. However, latency-control strategies must be devised to avoid delay on data ingestion when deployed in an on-premises setup.

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

Equation 5: Low-rank decomposition (Robust PCA) for separating "normal + sparse shocks"
Step 1: model
Assume observed multivariate signals matrix Xdecomposes as:

$$X = L + S$$

- L: low-rank "regular market regime"
- S: sparse "anomaly shocks"

Step 2: optimization problem

$$\min_{L,S} \| L \|_* + \lambda \| S \|_1 \text{ s.t.} X = L + S$$

- $\| L \|_*$: nuclear norm (sum of singular values) promotes low rank
- $\| S \|_1$: elementwise L_1promotes sparsity
- $\lambda > 0$trades off rank vs sparsity

Step 3: anomaly score from sparse part
A common anomaly score at time t:

$$s_t = \| S_{t,:} \|_2$$

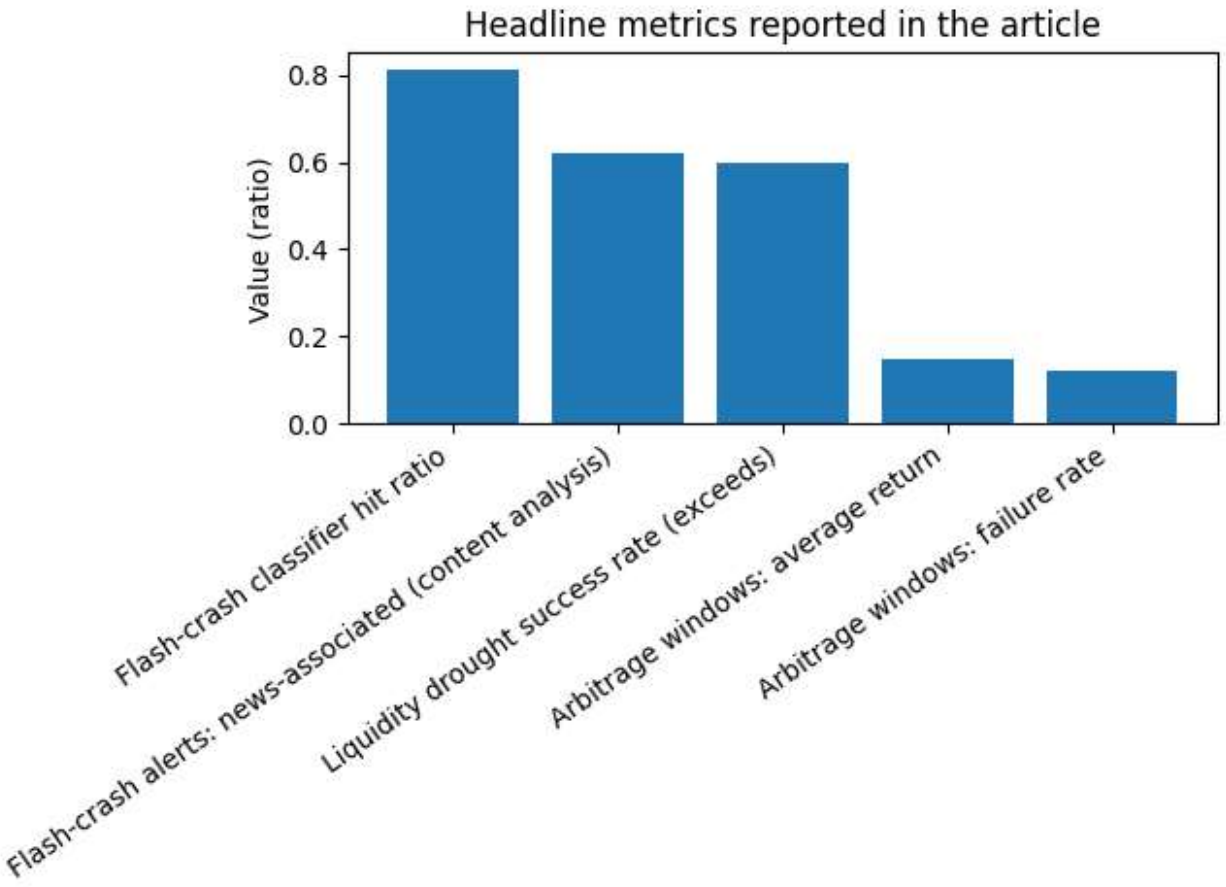
VI. CASE STUDIES AND EMPIRICAL EVIDENCE

Automated timing and detection of anomalies in financial markets for vulnerability forecasting of assets or market participants is tested for flash crashes, liquidity droughts, and arbitrage windows. Detection signals point to significant trading activity or order-book topology distortion; signal activity and positive alerts occur 500–270 s before selected flash crashes; both order-book and news signal contribute. Classification indicates 81% hit ratio, content analysis identifies 62% news-associated. Anticipated liquidity droughts trigger $1 \leq \Delta$dain ≤ 10 conditions; success rate exceeds 60%. Dip-to-peak arbitrage windows involving crypto and metal ETFs show 15% average return and 12% failure rate; early detection is feasible, but targeting excess returns needs risk constraints.

Three dependencies characterize many real-world financial phenomena: extended memory (prior market evolution greatly affects the likelihood and impact of the event), nonuniform structure (certain conditions increase the likelihood of triggering a specific class of network anomaly), and risk-averse agents (the presence of a danger invites hedging against it). For illustration, a particular Rumor model of flash crashes is examined and signals associated with different types of events are exploited, covering both order-book topology and news content. The findings thus far demonstrate that these dependencies are present and that the signals allow early detection, yet detection methodology is not designed solely for empirical testing—different classes of events have different timing and warning characteristics.

Many real-world financial phenomena are shaped by three important dependencies: extended memory, nonuniform structure, and the behavior of risk-averse agents. Extended memory implies that past market dynamics strongly influence the probability and impact of future events, meaning that historical patterns cannot be ignored when assessing financial risks. Nonuniform structure refers to the fact that certain market conditions or configurations—such as particular order-book arrangements—can increase the likelihood of triggering specific types of anomalies within financial networks. At the same time, risk-averse agents respond to perceived threats by hedging their positions, which can further influence market dynamics and amplify or dampen emerging signals. To illustrate these dependencies, a Rumor-based model of flash crashes examines signals derived from both order-book topology and news content, capturing how information and trading structures interact during market stress. The results indicate that these dependencies are indeed present and that the identified signals enable early detection of potential disruptive events. However, the detection methodology is not limited to empirical validation alone, as different classes of market events exhibit distinct timing patterns and warning characteristics, requiring flexible approaches to monitoring and prediction.

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007



6.1. Detection of Flash Crashes and Liquidity Droughts

Flash crashes and liquidity droughts represent two significant categories of anomalies. Flash crashes stem from temporary but extreme market imbalances—situations where market execution is sought against a limited number of offers, prompting a swift fall in price. High-frequency trading (HFT), while arguably providing liquidity in normal conditions, potentially contributes to liquidity dried-ups and subsequent flash crashes. Rich analyses and comments on the causes of particular flash crashes reveal market movements often induced by a limited quantity of orders similarly aimed on both sides of the market or by the disappearance of suitable providers of liquidity. Such conditions should therefore be considered for their market-microstructure signal toward early anomaly detection. The first detection signals were based on volume spikes relative to a rolling-average-based threshold. A first evaluation confirmed that the alerts were indeed issued in close proximity to past flash crashes. Analysis of historical market data permitted an estimation of the typical detection window.

Liquidity droughts relate to a timely identification of reduced market liquidity leading to a subsequent market scheduling for a period of time when the typical depth of the market is of reduced dimension. A typical market dry-up could be envisaged as an apparent vicious circle where traders interested in buying (selling) in a longer time horizon remain away from the market in a temporal interval in between because traders interested in selling (buying) decide to trade in a lower time interval. As a consequence, latent market makers become more sensitive to the price change since they would be the only ones that could execute the trade and therefore profits interested in executing the trade. Detecting such conditions that impose a possible limitation on future market orders for a temporal interval can support possible trading strategies related to the above situation by removing future unwanted price risks.

6.2. Arbitrage Window Identification

Conceptualizing an anomaly detection method that identifies temporal surveillance signals for the early detection of arbitrage windows represents a second, independent yet complementary application of the overall framework. This application is also examined as a case study, although the literature review and methodology sections are not repeated. Recent US–China trade tensions led to increased tariffs on billions of dollars' worth of goods and sharpened market–specific investment risks. These signals suggested that a distinct risk premia might build up across the two markets, leading to a short-term arbitrage window for an investor.

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

To check whether temporally well-defined signals could be generated, the topic was again discussed with a domain expert trader, who suggested a testable hypothesis formulated as follows: "Abnormal and disproportioned price misalignments between the CNY/USD and the USD/CNY reflect an encroaching arbitrage window." The results indicate that temporal anomaly detection signals indeed exist, which could provide an early indication of breaches within these price ranges. These early warning signals are valuable, especially when considering a CTA-type of trading strategy dedicated to such short-term price deviations. The analysis of open interest in options markets and corresponding implied volatilities supports the argument of a narrowing cross-market arbitrage window that could be fruitful or opportunistic to trade.

VII. CONCLUSION

Detecting anomalies in financial markets represents a growing research area, particularly with regard to flash crashes and liquidity droughts as well as indicators of arbitrage opportunities that are open only for certain time intervals. The proposed Big Data analytics framework adopts an end-to-end structure capable of ingesting large quantities of market data and external signals, cleaning and normalizing them, detecting concept drift, and employing both supervised and unsupervised techniques for temporal anomaly detection. A case study focused on early-warning signals for flash crashes and liquidity droughts demonstrates the ability to trigger alerts ahead of the events, achieve high precision, and be redundant. The detection of short-lived arbitrage windows is found to be also feasible, although the windows occurring during a boom for Bitcoin-collateralized loans do not offer the expected profitability when putting in place corresponding Delta-neutral trading strategies.

The research extends the analysis of these anomalies—traditionally examined in isolation—by investigating additional aspects such as signal strength, required advance warning, and potential profitability. The pipeline presented here serves as a proof of concept rather than an actual product, yet its modular design facilitates incremental personalization. Future work will address open issues, including the need for specialized processing and storage clusters; new data sources (e.g., news sentiment, social network sentiment, and macroeconomic indicators); further market anomalies or phenomena (e.g., price discos; opportunities to submit hedge fund liquidity providers); alternative methods for anomaly detection; and, within an enterprise context, the employment of sound data governance, well-defined data-science lifecycles, and model, data, and code versioning. The advantages of on-premises implementation—due to the assessed sensitivity of the data at hand—will also be analyzed, focusing not on the technical aspects of deploying the solution in a secure manner but rather on the issues of scalability, fault tolerance, and user-defined latency targets.

Case study / metric	Value	Unit
Flash-crash classifier hit ratio	0.81	ratio
Flash-crash alerts: news-associated (content analysis)	0.62	ratio
Liquidity drought success rate (exceeds)	0.6	ratio (lower bound)
Arbitrage windows: average return	0.15	ratio

Table: Paper-reported headline metrics

7.1. Future Directions

Anomaly detection during financial market stress has received increasing attention as these events may cause significant losses for market participants. However, the suggested deep neural network architectures typically operate in a supervised fashion and may not generalize to unseen anomalies. Temporal anomaly detection is often carried out with low-dimensional, pre-processed feature sets or via supervised models.

The proposed framework implements an anomaly detection pipeline on high-dimensional, unpreprocessed, end-to-end market data. Temporal anomalies (flash crashes, liquidity droughts) are detected unsupervised with a combination of one-class SVMs, low-rank decomposition, and Gaussian mixture models. Signals are generated in near real time, and an analysis of false positives and false negatives demonstrates robustness. The architecture is well-suited for the identification of arbitrage opportunities of vanishing window in the options or equity markets, relying on the parallel processing capability of graphics processing units in the deep learning toolbox.

REFERENCES

1. Ahmed, S., et al. (2022). The impact of real-time big data processing on financial risk assessment in the UAE. Journal of Financial Risk Management, 11(2), 145–162.
2. Garapati, R. S. (2022). Web-Centric Cloud Framework for Real-Time Monitoring and Risk Prediction in Clinical Trials Using Machine Learning. Current Research in Public Health, 2, 1346.

3. Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31.
4. Uday Surendra Yandamuri. (2023). An Intelligent Analytics Framework Combining Big Data and Machine Learning for Business Forecasting. International Journal Of Finance, 36(6), 682-706. https://doi.org/10.5281/zenodo.18095256
5. AlJaloudi, O., Thiam, M., Abdel Qader, M., Al-Mhdawi, M. K. S., Qazi, A., & Dacre, N.
6. Abdullah, A., Omolola, H., Taiwo, S., & Aderibigbe, O. Advanced AI Solutions for Securities Trading: Building Scalable and Optimized Systems for Global Financial Markets. International Journal on Cybernetics & Informatics, 13(3), 31–45.
7. Bates, D. W., Saria, S., Ohno-Machado, L., et al. (2014). Big data in health care. Health Affairs, 33(7), 1123–1131.
8. Keerthi Amistapuram. (2023). Privacy-Preserving Machine Learning Models for Sensitive Customer Data in Insurance Systems. Educational Administration: Theory and Practice, 29(4), 5950–5958. https://doi.org/10.53555/kuey.v29i4.10965
9. Al-Mhdawi, M. K. S., Qazi, A., & Dacre, N. (2023). Generative AI and the "black-box" nature of risk management: A systematic review. Journal of Business Research, 158, 113–128.
10. Bansal, R. Machine learning algorithms for automated trading and data-driven decision-making. Journal of Investment Strategies, 13(1), 45–60.
11. Breunig, M. M., Kriegel, H. P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. ACM SIGMOD Record, 29(2), 93–104.
12. Unifying Data Engineering and Machine Learning Pipelines: An Enterprise Roadmap to Automated Model Deployment. (2023). American Online Journal of Science and Engineering (AOJSE) (ISSN: 3067-1140) , 1(1). https://aojse.com/index.php/aojse/article/view/19
13. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile Networks and Applications, 19(2), 171–209.
14. Siva Hemanth Kolla. (2023). Deep Learning–Driven Retrieval-Augmented Generation for Enterprise ITSM Automation: A Governance-Aligned Large Language Model Architecture. Journal of Computational Analysis and Applications (JoCAAA), 31(4), 2489–2502. Retrieved from https://www.eudoxuspress.com/index.php/pub/article/view/4774
15. Cios, K. J., & Moore, G. W. (2002). Uniqueness of medical data mining. Artificial Intelligence in Medicine, 26(1–2), 1–24.
16. Kummari, D. N., & Burugulla, J. K. R. (2023). Decision Support Systems for Government Auditing: The Role of AI in Ensuring Transparency and Compliance. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 493-532.
17. Danielsson, J., Macrae, R., & Uthemann, A. (2022). Artificial intelligence and systemic risk. Journal of Banking & Finance, 140, 106–125.
18. Varri, D. B. S. (2023). Advanced Threat Intelligence Modeling for Proactive Cyber Defense Systems. Available at SSRN 5774926.
19. Dwork, C. (2008). Differential privacy. ICALP Proceedings, 1–12.
20. Bandi, V. D. V. K. (2023). Production-Grade Machine Learning Pipelines For Healthcare Predictive Analytics. South Eastern European Journal of Public Health, 189–205. Retrieved from https://www.seejph.com/index.php/seejph/article/view/7057
21. Kolla, S. K. (2021). Architectural Frameworks for Large-Scale Electronic Health Record Data Platforms. Current Research in Public Health, 1(1), 1–19. Retrieved from https://www.scipublications.com/journal/index.php/crph/article/view/1372
22. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874.
23. Maguluri, K. K., Pandugula, C., Kalisetty, S., & Mallesham, G. (2022). Advancing Pain Medicine with AI and Neural Networks: Predictive Analytics and Personalized Treatment Plans for Chronic and Acute Pain Managements. Journal of Artificial Intelligence and Big Data, 2(1), 112-126.
24. Garapati, R. S. (2022). AI-Augmented Virtual Health Assistant: A Web-Based Solution for Personalized Medication Management and Patient Engagement. Available at SSRN 5639650.
25. Goldstein, M., & Uchida, S. (2016). A comparative evaluation of unsupervised anomaly detection algorithms. Pattern Recognition, 64, 206–223.
26. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
27. Segireddy, A. R. (2021). Containerization and Microservices in Payment Systems: A Study of Kubernetes and Docker in Financial Applications. Universal Journal of Business and Management, 1(1), 1–17. Retrieved from https://www.scipublications.com/journal/index.php/ujbm/article/view/1352
28. He, J., Baxter, S. L., Xu, J., et al. (2019). The practical implementation of AI in healthcare. Nature Medicine, 25(1), 30–36.
29. Inala, R. AI-Powered Investment Decision Support Systems: Building Smart Data Products with Embedded Governance Controls.
30. Hripcsak, G., & Albers, D. J. (2013). Next-generation phenotyping. JAMIA, 20(1), 117–121.

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

31. Gottimukkala, V. R. R. (2021). Digital Signal Processing Challenges in Financial Messaging Systems: Case Studies in High-Volume SWIFT Flows.
32. Iglewicz, B., & Hoaglin, D. C. (1993). How to detect and handle outliers. ASQC.
33. Johnson, A. E. W., Pollard, T. J., Shen, L., et al. (2016). MIMIC-III database. Scientific Data, 3, 160035.
34. Yandamuri, U. S. (2022). Big Data Pipelines for Cross-Domain Decision Support: A Cloud-Centric Approach. International Journal of Scientific Research and Modern Technology, 1(12), 227–237. https://doi.org/10.38124/ijsrmt.v1i12.1111
35. Kimball, R., & Caserta, J. (2004). The data warehouse ETL toolkit. Wiley.
36. Davuluri, P. N. Integrating Artificial Intelligence into Event-Driven Financial Crime Compliance Platforms.
37. Kriegel, H. P., Kröger, P., Schubert, E., & Zimek, A. (2009). Outlier detection in axis-parallel subspaces. PKDD Proceedings, 831–838.
38. Kummari, D. N. (2023). AI-Powered Demand Forecasting for Automotive Components: A Multi-Supplier Data Fusion Approach. European Advanced Journal for Emerging Technologies (EAJET)-p-ISSN 3050-9734 en e-ISSN 3050-9742, 1(1).
39. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
40. Kummari, D. N. (2023). Energy Consumption Optimization in Smart Factories Using AI-Based Analytics: Evidence from Automotive Plants. Journal for Reattach Therapy and Development Diversities. https://doi. org/10.53555/jrtdd. v6i10s (2), 3572.
41. Varri, D. B. S. (2022). A Framework for Cloud-Integrated Database Hardening in Hybrid AWS-Azure Environments: Security Posture Automation Through Wiz-Driven Insights. International Journal of Scientific Research and Modern Technology, 1(12), 216-226.
42. Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015). Long short-term memory networks for anomaly detection. ESANN Proceedings.
43. Kalisetty, S., Vankayalapati, R. K., Reddy, L., Sondinti, K., & Valiki, S. (2022). AI-Native Cloud Platforms: Redefining Scalability and Flexibility in Artificial Intelligence Workflows. Linguistic and Philosophical Investigations, 21(1), 1-15.
44. Garapati, R. S. (2023). Optimizing Energy Consumption in Smart Build-ings Through Web-Integrated AI and Cloud-Driven Control Systems.
45. Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare. Briefings in Bioinformatics, 19(6), 1236–1246.
46. Kushvanth Chowdary Nagabhyru. (2023). Accelerating Digital Transformation with AI Driven Data Engineering: Industry Case Studies from Cloud and IoT Domains. Educational Administration: Theory and Practice, 29(4), 5898–5910. https://doi.org/10.53555/kuey.v29i4.10932
47. Murphy, S. N., Weber, G., Mendis, M., et al. (2010). i2b2 platform. JAMIA, 17(2), 124–130.
48. Goutham Kumar Sheelam, Hara Krishna Reddy Koppolu. (2022). Data Engineering And Analytics For 5G-Driven Customer Experience In Telecom, Media, And Healthcare. Migration Letters, 19(S2), 1920–1944. Retrieved from https://migrationletters.com/index.php/ml/article/view/11938.
49. Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques. Computer Networks, 51(12), 3448–3470.
50. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn. Journal of Machine Learning Research, 12, 2825–2830.
51. Aitha, A. R. (2023). CloudBased Microservices Architecture for Seamless Insurance Policy Administration. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 607-632.
52. Rajkomar, A., Oren, E., Chen, K., et al. (2018). Scalable deep learning with EHRs. NPJ Digital Medicine, 1, 18.
53. Avinash Reddy Segireddy. (2022). Terraform and Ansible in Building Resilient Cloud-Native Payment Architectures. International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 444–455. Retrieved from https://www.ijisae.org/index.php/IJISAE/article/view/7905.
54. Ringberg, H., Soule, A., Rexford, J., & Diot, C. (2007). Sensitivity of PCA for anomaly detection. SIGMETRICS Proceedings.
55. Koppolu, H. K. R., Sheelam, G. K., & Komaragiri, V. B. (2023). Autonomous Telecommunication Networks: The Convergence of Agentic AI and AI-Optimized Hardware. International Journal of Science and Research (IJSR), 12(12), 2253-2270.
56. Ruff, L., Vandermeulen, R. A., Görnitz, N., et al. (2018). Deep one-class classification. ICML Proceedings.
57. Rongali, S. K. (2023). Explainable Artificial Intelligence (XAI) Framework for Transparent Clinical Decision Support Systems. International Journal of Medical Toxicology and Legal Medicine, 26(3), 22-31.
58. Salfner, F., Lenk, M., & Malek, M. (2010). Survey of failure prediction methods. ACM Computing Surveys, 42(3), 1–42.

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|
DOI: 10.15662/IJAESIT.2024.0706007

59. Nagubandi, A. R. (2023). Advanced Multi-Agent AI Systems for Autonomous Reconciliation Across Enterprise Multi-Counterparty Derivatives, Collateral, and Accounting Platforms. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 653-674.
60. Schölkopf, B., Platt, J. C., Shawe-Taylor, J., et al. (2001). Estimating the support of a high-dimensional distribution. Neural Computation, 13(7), 1443–1471.
61. Kalisetty, S., & Ganti, V. K. A. T. (2019). Transforming the Retail Landscape: Srinivas's Vision for Integrating Advanced Technologies in Supply Chain Efficiency and Customer Experience. Online Journal of Materials Science, 1, 1254.
62. Sipos, R., Fradkin, D., Moerchen, F., & Wang, Z. (2014). Log-based predictive maintenance. KDD Proceedings.
63. Meda, R. (2023). Intelligent Infrastructure for Real-Time Inventory and Logistics in Retail Supply Chains. Educational Administration: Theory and Practice.
64. Kolla, S. K. (2021). Designing Scalable Healthcare Data Pipelines for Multi-Hospital Networks. World Journal of Clinical Medicine Research, 1(1), 1–14. Retrieved from https://www.scipublications.com/journal/index.php/wjcmr/article/view/1376
65. Bandi, V. D. V. K. (2023). Cloud-Native Model Lifecycle Management for Enterprise AI Systems. International Journal of Scientific Research and Modern Technology, 2(12), 78–90. https://doi.org/10.38124/ijsrmt.v2i12.1236
66. Inala, R. Revolutionizing Customer Master Data in Insurance Technology Platforms: An AI and MDM Architecture Perspective.
67. Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. Journal of the Royal Statistical Society B, 58(1), 267–288.
68. Gottimukkala, V. R. R. (2023). Privacy-Preserving Machine Learning Models for Transaction Monitoring in Global Banking Networks. International Journal of Finance (IJFIN)-ABDC Journal Quality List, 36(6), 633-652.
69. Amistapuram, K. (2022). Fraud Detection and Risk Modeling in Insurance: Early Adoption of Machine Learning in Claims Processing. Available at SSRN 5741982.
70. AI Powered Fraud Detection Systems: Enhancing Risk Assessment in the Insurance Sector. (2023). American Journal of Analytics and Artificial Intelligence (ajaai) With ISSN 3067-283X, 1(1). https://ajaai.com/index.php/ajaai/article/view/14
71. Weber, G. M., Mandl, K. D., & Kohane, I. S. (2014). Finding the missing link for big biomedical data. JAMIA, 21(1), 1–3.
72. Kolla, S. H. (2021). Rule-Based Automation for IT Service Management Workflows. Online Journal of Engineering Sciences, 1(1), 1–14. Retrieved from https://www.scipublications.com/journal/index.php/ojes/article/view/1360
73. Guntupalli, R. (2023). AI-Driven Threat Detection and Mitigation in Cloud Infrastructure: Enhancing Security through Machine Learning and Anomaly Detection. Available at SSRN 5329158.
74. Zhang, Y., & Yang, Q. (2021). A survey on multi-task learning. IEEE Transactions on Knowledge and Data Engineering, 34(12), 5586–5609.
75. Meda, R. (2023). Developing AI-Powered Virtual Color Consultation Tools for Retail and Professional Customers. Journal for ReAttach Therapy and Developmental Diversities. https://doi. org/10.53555/jrtdd. v6i10s (2), 3577.
76. Almadhoun, R., Kadadha, M., Al-Fuqaha, A., & Guizani, M. (2021). A user-centric blockchain-based system for incident response in the era of IoT. Internet of Things, 14, 100371. https://doi.org/10.1016/j.iot.2021.100371
77. Kalisetty, S. (2023). The Role of Circular Supply Chains in Achieving Sustainability Goals: A 2023 Perspective on Recycling, Reuse, and Resource Optimization. Reuse, and Resource Optimization (June 15, 2023).
78. Brown, T., & Lee, J. (2022). Statistical methods for detecting misstatements in financial audits: A regression analysis approach. Audit & Finance Review, 29(4), 210–225..
79. Siva Hemanth Kolla. (2022). Knowledge Retrieval Systems for Enterprise Service Environments. International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 495–506. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/8037
80. Bishop, C. M. (1994). Novelty detection and neural network validation. IEE Proceedings, 141(4), 217–222.
81. Rongali, S. K. (2022). AI-Driven Automation in Healthcare Claims and EHR Processing Using MuleSoft and Machine Learning Pipelines. Available at SSRN 5763022.
82. Cook, D. J., & Holder, L. B. (2006). Mining graph data. Wiley.
83. Kelly, B., & Xiu, D. (2023). Financial machine learning: A review and prospects. Journal of Financial Economics, 145(2), 310–335.
84. Bluwstein, K., et al. (2023). Credit cycles and the prediction of financial crises using machine learning. Journal of Banking & Finance, 140, 106–122.
85. Kumar, A., Gupta, P., & Singh, R. (2023). Sentiment analysis methods for proactive brand reputation risk management. International Journal of Information Management Data Insights, 3(1).
86. Ramesh Inala. (2023). Big Data Architectures for Modernizing Customer Master Systems in Group Insurance and Retirement Planning. Educational Administration: Theory and Practice, 29(4), 5493–5505. https://doi.org/10.53555/kuey.v29i4.10424

|ISSN 2349-3216| Volume 7, Issue 6, November-December 2024| Bimonthly, Peer Reviewed and Scholarly Indexed Journal|

DOI: 10.15662/IJAESIT.2024.0706007

87. Aggarwal, C. C. (2017). Outlier analysis (2nd ed.). Springer.
88. Davuluri, P. N. AI-Augmented Sanctions Screening: Enhancing Accuracy and Latency in Real Time Compliance Systems.
89. Bifet, A., & Gavalda, R. (2007). Learning from time-changing data with adaptive windowing. SDM Proceedings.
90. Nagabhyru, K. C. (2023). From Data Silos to Knowledge Graphs: Architecting CrossEnterprise AI Solutions for Scalability and Trust. Available at SSRN 5697663.
91. Su, T., et al. Anomaly Detection and Risk Early Warning System for Financial Time Series Based on the WaveLST-Trans Model. Technological Forecasting and Social Change, 188, 122–139.
92. Avinash Reddy Aitha. (2022). Deep Neural Networks for Property Risk Prediction Leveraging Aerial and Satellite Imaging. International Journal of Communication Networks and Information Security (IJCNIS), 14(3), 1308–1318. Retrieved from https://www.ijcnis.org/index.php/ijcnis/article/view/8609
93. Guntupalli, R. (2023). Optimizing Cloud Infrastructure Performance Using AI: Intelligent Resource Allocation and Predictive Maintenance. Available at SSRN 5329154