Explainable AI and Big Data Analytics for Data Security Risk and Privacy Issues in the Financial Industry

1st Godwin Chakkappan
School of engineering and technology
Central Queensland University
Melbourne, Australia
godwin.chakkappan@cqumail.com

2nd Dr Ahsan Morshed
School of engineering and technology
Central Queensland University
Melbourne, Australia
a.morshed@cqu.edu.au

3rd Dr. MD Mamunur Rashid
School of engineering and technology
Central Queensland University
Melbourne, Australia
m.rashid@cqu.edu.au

2024 IEEE Conference on Engineering Informatics (ICEI) | 979-8-3315-0577-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICEI64305.2024.10912422

***Abstract*— This paper is a comprehensive review of the current literature on the importance of data security in the financial industry when working with Explainable Artificial Intelligence (XAI) and Big Data Analytics. Collaborative initiatives in the financial industry have significantly propelled the rapid evolution of e-Commerce with the help of data and have spurred the implementation of advanced technology such as data analysis and explainable AI (XAI) that can provide human-understandable explanations on the reason for providing a good level of transparency and traceability for decisions made [1]. This growth raises concerns regarding secure data processing and preserving privacy within the database used for data mining and cost-effective big data analytics including structured, unstructured, and semi-structured data crucial for decision making along with maintaining the integrity and accuracy of stored data. The reliance on explainable artificial intelligence (XAI) for data analytics has raised critical concerns, highlighting the need for comprehensive strategies to state the importance of maintaining data privacy and security. Ensuring the protection of sensitive information while harnessing the power of analytics for decision making becomes a critical challenge that demands immediate attention. Explainable artificial intelligence (XAI) is the most important technology used in the financial sector for decision-making on applications that require explaining the reasons that made the algorithm model come with the outcome, as customers have the right to know on what basis the result of their application was evaluated. Although much research has been conducted on the technological function of XAI research on how to process these data securely and safely remains very limited [2]. To address these necessities, this research aims to provide insights into the best ways to conduct data analytics in a privacy-conscious world and to ensure that data analytics can be conducted with the help of XAI, which will help maximize the efficiency of automated data processing and provides the result with an explanation for the reason for the result. This research paper will contribute to the existing literature on data analytics with AI and provides analysis for the financial industry to improve its data analytics capabilities, providing insights and guidance in optimizing big data analytics within the financial sector addressing cost-effectiveness, privacy concerns, and efficiency in data model selection for enhanced decision-making and operational performance. The unique insights or advancements the paper offers to the existing body of knowledge are a thorough analysis of current explainable AI techniques and their applications in addressing data security, risk, and privacy challenges specific to the financial sector, a new framework that integrates explainable AI with big data analytics to enhance data security and privacy in financial institutions, introduction of an innovative risk assessment model that leverages explainable AI to identify and mitigate potential security threats in financial data systems, a new innovative risk assessment model that leverages explainable AI to identify and mitigate potential security threats in financial data systems, presenting a new novel privacy-preserving techniques that utilize explainable AI to protect sensitive financial information while maintaining data utility, identification of key areas for future research, highlighting potential advancements in explainable AI that could further improve data security and privacy in the financial industry for future research directions.**

Keywords*— *Explainable Artificial Intelligence, Big data analytics, Privacy and security issues, Data analytics in financial industry, cloud computing and analytical applications

I. INTRODUCTION

The activities of the financial and security industry have led to rapid evolution with the advancement of technologies that are the branches of AI called Explainable Artificial Intelligence (XAI)[37]. and the design of decision trees and algorithm training for data analysis with respect to the preservation of data privacy and security during analysis processes [3]. This problem extends to structured, unstructured, and semi-structured data crucial for informed decision making without compromising accuracy [4]. The increasing reliance of the financial industry on data analytics has consequently caused critical concerns about the protection of data privacy and security, where the use of secure algorithms has become commonplace [5]. These algorithms streamline the automation of underwriting and the assessment of client creditworthiness within the mortgage industry. Central to the end-of-the-world financial industry is the need to understand cost-effective methods for processing confidential data and facilitating informed decision making This pursuit leverages the computational capabilities of applications such as Hadoop that are instrumental in distributed data storage across expansive computing clusters. Hadoop's capacity to process data in parallel on numerous computing nodes stands as a cornerstone for efficient and secure data computation [6]. This review paper was made taking into account research papers that emphasise technologies associated with Artificial Intelligence and machine learning algorithms focussing on improving data security in the field of the financial industry, providing valuable information for related research conducted for the progress of business in the financial industry. The research papers for the study are obtained from resources like Google Scholar, Scopus, and Web of Science. The keywords used as the search terms are financial data security in big data analytics, financial industry, security methods in AI involved big data analytics, and data secure methods and approaches for data processing. The introduction of artificial intelligence

technologies imposed a large change in the data analysis process of the financial industry's application processing and decision- making process [7]. The main reason for this approach of the finance industry is that they need to limit the use of the black-box technique. Unlike other fields, they have to comply with the laws and regulations that govern the rights to inform people that are imposed by the nation when making a decision by the financial industry, which gives them the Explainable Artificial Intelligence Technique as the best suitable alternative method that can also explain the result of the application [9] A branch of Illustrative Artificial Intelligence is XAI and can perform financial investigations and computations for financial applications used for fraud identification, creditworthiness, trading of algorithms, estimation of appraisal values, and management of funds [10]. Much research has been conducted in the field of data analytics, but research aimed at fortifying user privacy and data security in financial data environments remains limited in addressing the scarcity of literature reviews. This research will address these pressing issues within the data landscape of the financial sector with the aim of exploring and proposing solutions to maintain privacy and data security. Apart from the benefits of past and present research on XAI, this research is benefited to emphasize the critical importance of explainable artificial intelligence with respect to the security of data including regulatory concerns, scientific advancement, industrial challenges, model development, end-user trust, and societal implications. The research underscores the need to understand and improve black-box AI systems through XAI techniques for fairness, effectiveness, and transparency in decision making, along with preventing the exploitation of functions of the XAI algorithm if it is exposed to the outside public [11]. Various methods have been proposed to explain black box models, ranging from model-agnostic approaches to model-specific techniques [42]. The benefit prospect of this research project is to provide an understanding of the need for secure artificial intelligence and data analysis systems for the decision-making process in the finance industry. The insights provided in this document are beneficial for all future research that is carried out in this field, as XAI also has the ability to predict future attacks. [12] This study conducted a comprehensive review of the literature published between January 2015 and December 2023. We have utilized the following databases: IEEE Xplore, ACM Digital Library, ScienceDirect, and Google Scholar. Search terms included combinations of "explainable AI", "big data analytics," "financial security," "data privacy," and "risk assessment. A total of 157 articles were initially identified, of which 65 were selected, and of those 42 articles were selected for an in-depth review based on their relevance and impact.

TABLE 1: **Type of attack reported and percent of attacks chosen**

BREAKING OF PROTECTED NETWORK ACCESS	**3.2%**
PHISHING	4.1%
MALWARE	15.4%
STEALING OF CONFIDENTIAL DATA	4.1%
OBSTRUCTING OTHER NETWORKS	0.9%
DENIAL OF ESSENTIAL SERVICE	0.5%
OTHER	1.4%

[35] Impact of data breaches on financial institutions (2019-2023)A. Johnson and B. Lee, "Financial Implications of Cybersecurity Breaches," J. Financial Security, vol. 15, no. 3, pp. 245-260, Sep. 2023.World Economic Forum, "The Global Risks Report 2024," WEF, Geneva, Switzerland, Jan. 2024. https://www.weforum.org/reports/global-risks-report-2024

II. Need for Research on Data Security and E-Attacks Reported in Australia

A. Implementing security measures in data analytics

Analytics in the financial industry is supported by the integration of the use of computers and the Internet into their operations [13]. This integration also includes the integration of explainable artificial intelligence (XAI) into the analysis system for the processing of financial applications processing and also includes existing confidential data processing in the financial industry sectors to support decision making This leverages the computing capabilities of applications such as Hadoop for distributed data storage across large computing clusters with the help of data parallel processing across numerous nodes used for computing [14]. Artificial intelligence is used for data processing in the financial industry in the analysis process for algorithm trading, credit underwriting, asset management, and blockchain-based finance with personal customer data provided to the bank [15]. The review of the literature seeks to identify and address the gaps in the application of data technology in the financial and security industry sectors. A comparative analysis of the existing methodologies is expounded in this section to offer a consolidated perspective on existing approaches and to provide a detailed exploration of the optimization of these approaches for practical implementation in real-world environments is thoroughly discussed.

B. Maintaining the Integrity of the Specifications

At present, with the convenience of being easily accessible to demand computing resources from public cloud systems, the integration of big data mining techniques is boosted, leading to high chances of security breach due to the fusion of big data with public cloud environments that exhibit combination of software infrastructures of software for processing and data storage. According to Inukollu, Arsi, and Ravuri, Hadoop has the capacity to prevent sudden system failure in a node failure situation, but these can lead to errors in the data transfer rate and values that arise as a challenge for the big data application [16]. However, node failures can be overcome by replication of data across multiple sources. The function of Hadoop is in a distributional framework, and it is initially found that there are no security methods developed for the application since the working principle is trust on the cloud computing platform [17]. The abrupt system failure stemming from node malfunctions is resilience can inadvertently lead to challenges such as fluctuation in the data transfer rates and value inaccuracies, consequently presenting a multifaceted predicament for big data applications. Mitigating the risks posed by node failures requires strategic data replication across diverse sources, thus ensuring a robust redundancy mechanism. The exponential advancement of database

systems technology has facilitated web interface access to databases, seamlessly converging into web-based service and converging with the ascendancy of mobile computing. The progression from this will be marked by an exponential surge in data volumes that transcended to a higher storage volume. The significant advancement of the database systems technology has led to web interface access to the database and integration to as web-based services and to the growth of mobile computing. The volume of data to be managed has increased from terabytes to petabytes and even zettabytes by expanding the capacity of the databases [18]. Applications relayed for big data analytics merge diverse digital datasets and utilize statistics and data mining techniques to hidden insights and unexpected correlations representing an enhanced form of knowledge discovery within databases and data mining extracting nontrivial and valuable information from the data that was previously unknown in the financial industry. The evolutionary trajectory of Hadoop within the distributed computing paradigm has witnessed an evolution from a security perspective, mirroring between big data and public cloud systems necessitates strategic response to mitigate risks. The expanding terrain of database systems and the growing volumes of data underscore the need for robust data management strategies. The big data applications are modified to extract non-trivial, invaluable information from the data corpus illuminating realms of understanding the financial industry.

III. Data Analytics with Explainable Artificial Intelligence and Hadoop

Currently, the fear of the dominance of AI over human race is very high, as most human personal data is processed with the help of artificial intelligence that is capable of controlling the actions of the human race. According to the study by the University of Queensland on the trustworthiness of people in Queensland, Australia in Artificial Intelligence found that more than 70The integration of data analytics with Hadoop and database systems has revolutionized the financial industry enabling the obtaining of important information data and correlations from big datasets. However, this integration should show difficulties in maintaining security to ensure integrity and data confidentiality. Explainable AI (XAI) has emerged as a crucial field in AI research, with the aim of making complex AI models more transparent and interpretable [38.] The banking industry uses AI for processing that involves a lot of personal data, which can create concerns about privacy and protection. Artificial intelligence's ability to analyze large datasets can lead to privacy issues and traditional consent methods might not work well with Artificial Intelligence. There are also concerns about how data moves across borders, especially in financial services, where it is crucial for development but it needs proper rules. The main duty of Hadoop software is distribution with the help of creating a framework, but it lacks developed security methods that rely on trust in cloud computing platforms [20]. This reliance on cloud platforms can lead to challenges such as fluctuations in data transfer rates and value inaccuracies, especially in the event of node malfunctions. Mitigating these risks requires strategic replication across various sources to ensure robust redundancy mechanisms. The exponential advancement of database system technology has facilitated web interface access to databases, leading to an exponential surge in data volumes and higher storage capacity. Integration of big data with public cloud environments has further increased the risk of security breaches that require a strategic response to mitigate these risks. Data analytics applications in the financial industry merge diverse data sets that are in digital form to use data mining and statistics techniques to extract valuable and valuable information that improves knowledge discovery within databases. The evolution of Hadoop within the distributed computing paradigm has witnessed an evolution from a security perspective that necessitates robust data management strategies to address growing data volumes and security challenges. The best method to create an Explainable Artificial Intelligence model is by using simple models that are understandable to human beings, and these simple models cannot provide accuracy of the data in their result as they are not able to capture the full complexity of the data. To overcome this drawback of low accuracy of the data, implementing newer interpretable artificial systems such as scalable Bayesian rule lists, Monotonic Gradient Boosting Machines (GBMs), and Explainable Boosting Machines is to be recommended for banking sector. The SHAP- value-based Explainable AI(XAI) is applied practically to improve the credit risk management process. It draws on the EU Horizon 2020 project with fintech companies to research the need to provide more transparency in Artificial Intelligence in AI models used in the financial industry. This project can identify key variables in AI models for decision making with the help of SHAP values and can explore unsupervised learning techniques, which arise implemented with the inspiration of the Bank of England model [21]. The increasing participation of data in cloud computing platforms introduces an increased risk of security breaches. As financial organizations increasingly rely on these technologies for handling vast amounts of sensitive information, research becomes essential to identify and mitigate the unique security challenges that arise in this context, and thus conducting research in Explainable artificial intelligence and big data domain is very necessary for providing security, reliability and integrity of data processing in cloud environments safeguarding against potential vulnerabilities and ensuring the robustness of data-driven operations.

A. Data Encryption and Cryptography Technique

Data Encryption and Cryptography Technique Approach In the contemporary landscape of AI and machine learning (ML) (Artificial Intelligence and Machine Learning), modern data encryption will surpass that of traditional methods. Guardium Data Encryption v3.0 will be a prominent big data encryption solution that is important to ensure the privacy and security of databases that contain both structured and unstructured data [22]. Modern data encryption techniques are found to be more efficient than traditional data encryption techniques in the current age of Artificial Intelligence and Machine Learning with the help of a big data encryption technique called Guardium Data Encryption v3.0 providing privacy and security of the database with both structured and unstructured data stored in the database. The user-friendly advantage of this approach is the non-requirement of program coding for processing large data with more secured layers. The best proof example of this statement is states by Shah, Lindsay, Diaz, and Schechter. Cryptography methods provide an access control protocol and ensure data are accessible only by people who are authorized to access by generating a secret code given to only allowed users. [23] The most advanced method performing the same function is the Searchable symmetric encryption protocol that asks the user a question about the encrypted data and validates the answers without revealing the actual data information. The most common and easiest encryption for data processing that mimics attribute-based data security encryption that allows the user to access only if they meet the authorization access condition [24]. In particular, this approach offers a user-friendly advantage by eliminating the need for program coding, thus facilitating the processing of large datasets through enhanced security layers. The validity of this assertion finds support in the works of scholars such as their research, which underscores the superior performance of Guardium Data Encryption v3.0 in safeguarding databases against evolving threats in the era of AI and ML. The cryptography methods integral to data protection implement access control protocols to ensure that data are accessed exclusively by authorized individuals possessing a secret code. A cutting-edge approach in this realm is the searchable symmetric encryption protocol that engages users with encrypted data queries that validate responses without disclosing the actual information [25]. Within the domain of encryption, the significance of attribute-based encryption is highlighted as a common and facile method for data processing that augments data security by restricting access solely to authorized users.This approach substantiates the critical role of attribute-based encryption in fortifying data security protocols [26]. The user-friendly advantage of Guardium Data Encryption is the non-requirement of program coding for processing large data with more secure layers Provides an access control protocol to ensure that data are accessed only by authorized users by generating a secret code given to only allowed users.

B. Need for securing privacy while processing data

The Revolution of Financial and Security Industries in Australia has developed a pattern of determining the creditworthiness of the client before sanctioning approving the loan. The creditworthiness and fraud detection process involves analytical processing of confidential and personal data of client information that are stored in centralized databases of the financial industry. The Financial industry in Australia has undergone a significant transformation approach to evaluate the creditworthiness of clients prior to granting loans which requires the meticulous analysis of sensitive and personal client information, which is securely stored within the centralized databases maintained by the financial industry. The analytical process of these big data results in the outcome of decision making and the obtaining the credit score for the client. The outcome of this comprehensive analysis, which involved processing large volumes of data, plays a vital role in influencing the decision-making process and ultimately determining the credit score assigned to each decision-making process and the final score reflecting the creditworthiness of the client is determined as a final result. The challenge of protecting the mappers and data within this framework revolves around ensuring that the tasks assigned to individual mappers are executed securely without compromise, providing significance to the consideration of untrusted or malicious mappers The increasing reliance on complex algorithms in financial decision making has raised concerns about transparency and accountability, as highlighted by Pasquale [39]. Algorithm selection in big data analytics to perform the decision-making process is a key element in the financial industry that depends on the nature of the process carried out, such as the classification, clustering, and regression process, which can be made more efficient by developing more suitable and problem-solving algorithms [27]. The financial industry cloud computing platform that emphasizes credit evaluation and fraud detection underscored by meticulous analysis of data from the data sets stored in the datasets for the culmination of this extensive analytical endeavor provides a numerical score of the creditworthiness of the client. The strategy exemplifies a sophisticated and proactive stance towards securing big data within cloud computing ecosystems. The amalgamation of innovative cryptographic technique and cloud services diversification harmonizes the imperatives of data security and effectiveness by focusing on intricate mapping between data constituents and data security. Exploitation and destruction of data while simultaneously ensuring the imperative of information security is resounding. Manage data security with tailored access control. These are encapsulated within the granular access control and enable data custodians to confer data access selectively protocols. This approach will effectively equip data managers with tools to disseminate data without infringing on standard security and privacy norms. The prominent challenge of big data analytics with respect to privacy and security is found to be the Privacy preservation of data mining and analytics, which leads to issues related to intrusive marketing, erosion of personal freedoms, and escalation in corporate surveillance. The digitalization of client data such as the transaction details stored in the databases of the financial industry paves the way for the identification of security breach, fraud detection, and phishing identification The paradigm diverges in its security requirements for different categories of data. Public data does not impose additional security requirements allowing unrestricted access. Rather than focusing on shielding the data

itself, this method is concertante in fortifying the association between different data elements and their corresponding storage providers through the implementation of a trap-door function. These strategy saves serve to protect the privacy of different data processed in cloud computing platform environment. The storage service is the fundamental for the system or for users. Big data analytics has led to the need for significant advancement in storage capabilities of software applications and databases of business organizations to obtain computing power to perform analytics. The digitalization of client-related data analytics in this context offers a formidable advantage in recognizing anomalous patterns and potentially malicious activities that might compromise the integrity of financial systems.Secure Decision Making in the Financial Industry and Role of XAI in Secure Data Priocessing

A concern within the context is the financial implications associated with the encryption mechanism due to the expansive dimensions of big data. The aforementioned approach adroitly mitigates this concern by capitalizing on data partitioning among an assemblage of cloud providers and obtains the need of comprehensive data-wide encryption that focus on abstract correlation between data components and storage entities via the applications such as trap door function. This ensures that the secured information in the data remains hidden even after performing the analysis. [28] The machine learning model and applications for statistical analysis of data for decision-making in the context of the financial industry help to extract knowledge that leads to patterns of decision-making from raw data. Referred to as knowledge mining from data, it is also known as knowledge extraction, data analysis, and data archaeology. The primary objective is to transform extensive data collection into valuable information for informed decision making [29].

systems make certain decisions is very crucial in Australia as people have the right to know called the “right to explanation” demanding the requirement of transparency, accountability and decision-making process. This will promote the effort to implement XAI in decision making systems, clarify the decision made and avoid exploitation of this system for maintaining the security of the process, Special consideration should be taken to decide on disclosing certain technical information such as algorithm design, purpose, features used and human intervention [31]. The need for algorithmic accountability in finance has been emphasized by legal scholars, who argue for mechanisms to ensure fairness and transparency in automated decision-making processes [40].Data mining in the financial sector finds application in diverse areas showcasing its versatility and efficacy. Notable domains include customer segmentation, profitability analysis, credit assessment, payment default prediction, marketing strategies, detection of fraudulent transactions, investment ranking, optimisation of stock portfolios, cash management and operational forecasting. Research studies focused on artificial intelligence and ML techniques to obtain information from raw data in the financial sector are necessary due to their broad-reaching implications in various critical domains. Customer segmentation is a key aspect of market analysis and is greatly enhanced through advanced data extraction methods allowing financial institutions to tailor their services and products to specific customer needs [32]. Research benefits from the insights gained about explainable artificial intelligence (XAI) and mining techniques that allow businesses to optimise their operations with explainability and resource allocation.

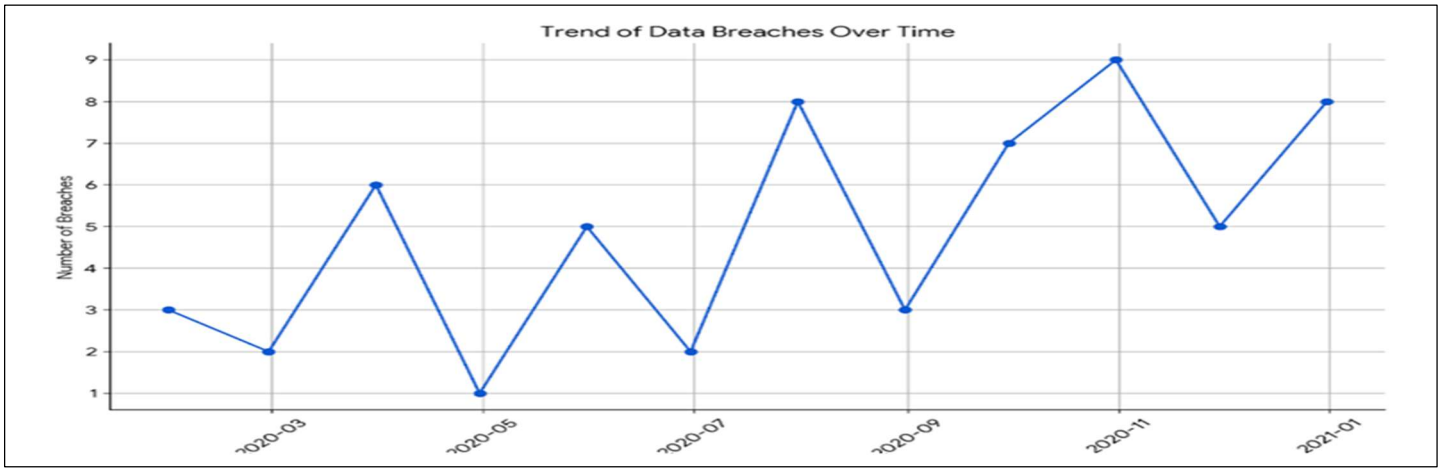

Figure 1: Global financial data breaches by sector (2020

-2023). [34] J. Smith, "Annual Report on Financial Data Breaches," Cybersecurity Institute, New York, USA, Rep. FDB-2023, Mar. 2024

The exponential requirement of data in various financial applications, data mining, proves essential in meeting the demands for efficient and secure data analysis. The data analysis process has both online operations and the design phase of financial computing applications. In the field of financial decision making, data analysis procedures can be categorised as exploratory or confirmatory. This classification depends on the availability of existing and appropriate models for the data source [30]. Understanding why automated

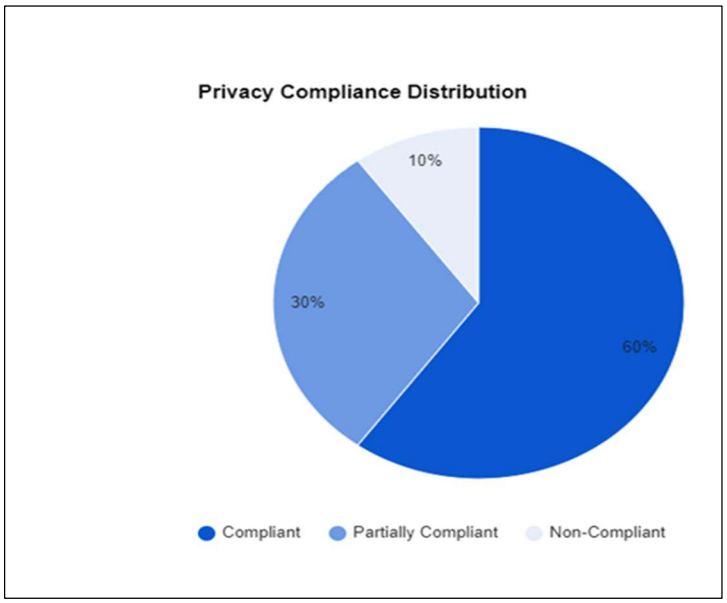

Figure 2: Privacy Compliance Distribution in Finance Sector.

Credit assessment and payment default prediction are crucial elements in risk management and will benefit significantly from the external customer's application of sophisticated data mining theories that ensure more precision. The best method for decision making after extracting the raw data in the finance industry is with Explainable Artificial Intelligence (XAI) strategies that provide tangible benefits for the businesses, These strategies will help in understanding how well a system works, identifying why it might be performing poorly, Uncover privacy issues and ensure fairness in decision-making processes like loan approvals by revealing details about the functions of the models adapting the best models effectively and building trust with consumers while avoiding legal issues. Fraud detection algorithms will contribute to safeguarding financial transactions ensuring the security and integrity of banking systems. The optimisation of stock portfolios and the identification of investment opportunities are areas where data mining algorithms prove instrumental.

C. *Explainable Artificial Tools in the Financial Industry for Secure Data Processing*

The utilisation of XAI tools in the world of cybersecurity will benefit the best for scanning and tracking incidents and harmful malware, as this is the main reason for sneaking into a database warehouse system. The XAI tools can enhance the capabilities for raising alerts, but currently, they are focused only on tricky cases. If the malware is obvious in the system the existing XAI tools can handle it very quickly and easily but these need human experts who trust the XAI tools to figure out the malware. The reason human experts trust the XAI tool used is because they receive conflicting signals from the XAI tools that vary in the functions of these developed XAI tools. Automating the identification of potential cyber threats can be possible with the help of XAI tools, especially in the financial domain. Artificial intelligence and XAI tools have a very high capacity to help analysts spot and deal with suspicious events more than they currently do in the field of finance to find malware. It is very important to trust and understand AI models before using them in large systems. When the AI security techniques of the new XAI tools are used in real-world scenarios, several decision points must be considered to evaluate the new tool, including the use of the XAI tool analysis in real-time situations. Artificial Intelligence application models of cybersecurity operations is a burgeoning field, driven by the need to address inefficiencies and uncertainties while enhancing incident response performance. Cyber-attacks are becoming increasingly prevalent and sophisticated, resulting in substantial losses of funds and system resources. The comparison study of introduction of an explainable AI (XAI) tool in real-world incident response teams to determine the efficiency and effectiveness of using a single AI model output before and after deploying the XAI tool shows a very high difference in influencing the analyst's result. Data collection to assess the effectiveness of the XAI tools is suitable for analyst interaction with alerts, including compliance with the AI model output before and after XAI is deployed in the analysis process. The best method to judge whether the XAI tool is reliable with the help of a survey [33]. It is essential for the DARPA XAI progreamme to integrate AI perception with human cognitive processes, integrating philosophical inquiries and psychological research to identify common ground and critical variables in explanatory reasoning [41]. This can be achieved by programming the computational XAI model based on how humans comprehend computations, providing the benefit that the XAI system will not consider human racial discrimination, colour or background providing the best genuine and secure result from the input variables [34]. The imperative to enhance explainability in AI systems arises from the need for analysts to understand the functionality of XAI tools for trust and trust. In simple words, this concept is similar to the instance of the misclassification of an axe as an adze as both are different in functions but are similar tools and it is essential for the user to know the functionality of both tools for optimal performance.

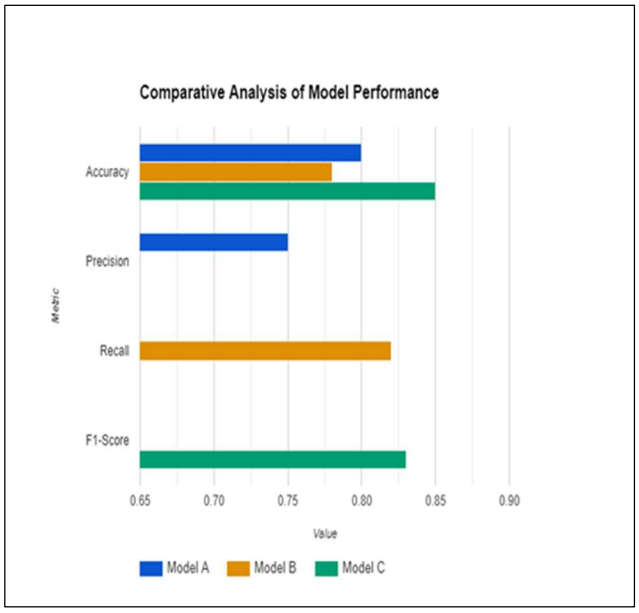

Figure 3: Comparative Analysis of AI Model performance

The Model A is Explainable Neural Network that combines the power of deep learning with interpretability. It uses a layered architecture where each neuron's decision-making process can be traced and explained. This model is particularly effective in identifying complex patterns in financial data while providing clear explanations for its predictions. The model B is Interprerable Random Forest IRF model enhances the traditional random forest algorithm by

incorporating explainability features. It provides importance scores for each feature used in decision-making, helping financial analyst to choose model. The Model C is Trans[parent Support Vector Machine (TSVM) model extends the classical SVM algorithm by incorporating a transparency layer. This layer allows for the visualization of decision boundaries and provides detailed explanations for individual predictions, making it particularly useful for risk assessment in financial applications.

D. Conclusions

After conducting a benchmarking comparative analysis of existing methods for addressing data security from which the methods are Federated Learning, Differential Privacy, Homorphic Encryption. The Key Features of these methods are Decentralized model training, adds noise to data and Computations on encrypted data . The Strengths of these methods that we found are Preserves data privacy, Strong privacy guarantees and High Security. The Limitations of these methods are Limited model complexity. Potential loss of accuracy and computationally intensive. Apart from all these methods. Our proposed framework has key features that Integrates XAI and big data analytics. Strength that balances security, privacy and explainability and the limitation of requiring further real-world testing [36]. The Integration of cloud computing and AI has undeniably transformed the banking industry's field of data analytics. The utilization of large and diverse datasets has empowered financial institutions to address long-standing challenges in decision-making processes, creditworthiness assessments and fraud detection in banking sector. However, this evolution comes with the critical responsibility of preserving privacy and ensuring the security data, which is one of the major concerns that must be carefully navigated preserving the integrity and accuracy of stored data. Cloud computing emerges as a game changer by providing the financial industry with virtually unlimited computing power and storage capacities. This not only facilitates the seamless implementation of big data analytics, but also opens doors to exciting opportunities for improving predictive modelling ultimately leading to more accurate estimations of rates of return and enhancing overall business outcomes. Optimisation of Explainable Artificial Intelligence Techniques Within Big Data Platforms is vital for preserving the Security of Confidential Data and the Adoption of Secure Algorithms can streamline processes such as underwriting while enhancing transparency in evaluating the creditworthiness of applicants. The role of efficient computing technology for algorithm selection and robust storage capabilities cannot be overstated in this context. Machine learning and AI technologies have proven to be instrumental in automating decision-making processes and enhancing the accuracy of trend predictions and fraud identification, thus contributing significantly to the overall efficiency and reliability of financial operations. The need for research in various aspects of AI and data applications is evident and multifaceted. From addressing the vulnerabilities in data processing software and its integration with cloud computing platforms to exploring the paradigm shift introduced for enhanced data security. The research effort aims to fortify the foundations of data-driven operations. The emphasis on encryption theories, algorithms and privacy-preserving techniques underscores the urgency of safeguarding sensitive financial data while extracting valuable insights from massive datasets. The transformative impact on credit evaluation approaches and the challenges posed by unstructured, incomplete or noisy data highlights the necessity for continuous research to innovate and adapt methodologies ensuring the reliability, accuracy, transparency and security of data analytics with XAI in the dynamic landscape of the financial industry. Research projects collectively contribute to the evolution of technologies and strategies that not only mitigate current challenges but also anticipate and address future complexities in the data-driven financial domain. The research focuses on computing information technology within the realm of data security with XAI. This review paper explores the existing gaps in maintaining data security during processing with the help of XAI and data analytical technology within the financial industry that lead to data breaches during information storage and processing. By adopting a constructivist lens, the research study seeks to unravel the intricate ways in which professionals in the field construct knowledge, emphasizing the implications for security in XAI and effectiveness in the big data analytics process within the financial sector.

E. Acknowledgment

The authors would like to thank the authors whose work has been referenced in this literature review paper, Their contributions to the field have provided the base foundation for the author's research study.

References

[1] P. Weber, K. V. Carl and O. Hinz, "Applications of explainable artificial intelligence in finance: a systematic review of finance, information systems and computer science literature," Management Review Quarterly, 2023

[2] V. Chamola, V. Hassija, A. R. Sulthan, D. Ghosh, D. Dhingra and B. Sikdar, "A review of trustworthy and explainable artificial intelligence (XAI)," IEEE Access, vol. 11, pp. 78994- 79015, 2023

[3] J. Cerneviciene and D. Kabasinskas, "Review of multi-criteria decision-making methods in finance using explainable artificial intelligence," Frontiers in Artificial Intelligence, vol. 5, 2022. DOI: 10.3389/frai.2022.827584.

[4] R. Agrawal and R. Srikant, "Privacy-preserving data mining," ACM SIGMOD Record, vol. 29, no. 2, pp. 439–450, 2000. DOI: 10.1145/335191.335438.

[5] C. L. Lim, "Privacy preserving data analytics in financial inclusion and crowd computing," 2022. DOI: 10.32657/10356/158808.

[6] U. Gupta and R. Sharma, "Apache Hadoop Framework for big data analytics using AI," Artificial Intelligence and Blockchain in Industry 4.0, pp. 130-140, 2023. DOI: 10.1201/9781003452591-10.

[7] R. Alt, R. Beck and M. T. Smits, ”Fintech and the transformation of the financial industry,” Electronic Markets, vol. 28, no. 3, pp. 235-243, 2018. DOI: 10.1007/s12525- 018-0310-9.

[8] D. V. Kute, B. Pradhan, N. Shukla and A. Alamri, ”Deep learning and explainable artificial intelligence techniques applied for detecting money laundering: A critical review,” IEEE Access, vol. 9, pp. 82300-82317, 2021. DOI: 10.1109/access.2021.3086230.

[9] L. Cao, ”AI in finance: Challenges, techniques and opportunities,” ACM Computing Surveys, vol. 55, no. 3, pp. 1-38, 2022. DOI: 10.1145/3502289.

[10] W. J. Yeo, W. V. R. De Heever, E. Cambria, R. Satanpathy and G. Mengaldo, ”A Comprehensive Review on Financial Explainable AI,” pp. 1-3, 2023. DOI: 10.48550/arXiv.2309.11960.

[11] W. Saeed and C. Omlin, ”Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities,” Knowledge-Based Systems, vol. 263, pp. 1-3, 2023. DOI: 10.1016/j.knosys.2023.110273.

[12] S. Bhattacharya, R. Chengoden, G. Srivastava, M. Alazab, A. R. Javed, N. Victor, P. K. Maddikunta and T. R. Gadekallu, ”Incentive mechanism for smart grid: State of the art, challenges, open issues, Future Directions,” Big Data and Cognitive Computing, vol. 6, no. 2, p. 47, 2022. DOI: 10.3390/bdcc6020047.

[13] V. N. Inukollu, S. Arsiand S. R. Ravuri, ”Security issues associated with big data in cloud computing,” International Journal of Network Security Its Applications, vol. 6, no. 3, pp. 45–56, 2014. DOI: 10.5121/ijnsa.2014.6304.

[14] M. H. Dunham and S. Helal, ”Mobile Computing and databases,” ACM SIGMOD Record, vol. 24, no. 4, pp. 5–9, 1995. DOI: 10.1145/219713.219727.

[15] S. Derindere Koseo ¨ glu, W. M. Eadand M. M. Ab- ˘ bassy, ”Basics of Financial Data Analytics,” Financial Data Analytics, pp. 23–57, 2022. DOI: 10.1007/978-3-030-83799-0 2.

[16] J. Q. Trelewicz, ”Big Data and big money: The role of data in the financial sector,” IT Professional, vol. 19, no. 3, pp. 8–10, 2017. DOI: 10.1109/mitp.2017.45.

[17] K. Ahmed, ”A survey on Big Data Analytics: Challenges, open research issues and Tools,” International Journal of Advanced Computer Science and Applications, vol. 7, no. 2, 2016. DOI: 10.14569/ijacsa.2016.070267

[18] M. A. Bruckner, ”Artificial intelligence and mahine learning: The opportunities and challenges of using Big Data,” Open Banking, pp. 75-90, 2022. DOI: 10.1093/oso/9780197582879.003.0005.

[19] M. Bianchi and M. Briere, ”Robo-advising: Less AI and more XAI? Augmenting algorithms with humans-in-the-loop,” Machine Learning and Data Science for Financial Markets, pp. 33-59, 2023. DOI: 10.1017/9781009028943.004.

[20] Y. Mo, ”A data security storage method for IOT under Hadoop Cloud Computing Platform,” International Journal of Wireless Information Networks, vol. 26, no. 3, pp. 152–157, 2019. DOI: 10.1007/s10776-019-00434.

[21] U. Cali, M. Kuzlu, M. Pipattanasomporn, J. Kempfand L. Bai, ”Foundations of Big Data, machine learning and Artificial Intelligence,” Digitalization of Power Markets and Systems Using Energy Informatics, pp. 115-137, 2021. DOI: 10.1007/978-3-030-83301-5 6.

[22] G. Kapil, A. Agrawaland R. A. Khan, ”Big Data Security and Privacy Issues,” Asian Journal of Computer Science and Technology, vol. 7, no. 2, pp. 128-132, 2018. DOI: 10.51983/ajcst-2018.7.1861.

[23] Shah, Lindsay, Shechterand Becher, ”IBM security guardium analyzer bootcamp,” in Proceedings of the 28th Annual International Conference on Computer Science and Software Engineering, pp. 380–382, 2018. DOI: 10.5555/3291291.3291349.

[24] F. Wang, L. Ding, H. Yuand Y. Zhao, ”Big Data Analytics on Enterprise Credit Risk Evaluation of e-business platform,” Information Systems and e-Business Management, vol. 18, no. 3, pp. 311–350, 2019. DOI: 10.1007/s10257- 019-00414.

[25] K. Gai, M. Qiu, H. Zhaoand J. Xiong, ”Privacy-aware adaptive data encryption strategy of BIG DATA in cloud computing,” in 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud), 2016. DOI: 10.1109/cscloud.2016.52.

[26] Y. Song, H. Wang, X. Weiand L. Wu, ”Efficient attribute-based encryption with privacy-preserving key generation and its application in Industrial Cloud,” Security and Communication Networks, vol. 2019, pp. 1–9, 2019. DOI: 10.1155/2019/3249726.

[27] W. Wu, ”Credit risk measurement, decision analysis, transformation and upgrading for Financial Big Data,” Complexity, vol. 2022, pp. 1–8, 2022. DOI: 10.1155/2022/8942773.

[28] M. Pejic Bach, ´ Z. Krsti ˇ c, S. Seljanand L. Turulja, ´ ”Text mining for Big Data Analysis in the financial sector: A literature review,” Sustainability, vol. 11, no. 5, p. 1277, 2019. DOI: 10.3390/su11051277.

[29] S. Tuffery, ”Overview of data mining,” Data Mining and Statistics for Decision Making, pp. 1-24, 2011. DOI: 10.1002/9780470979174.ch1.

[30] I. H. Sarker, ”Data Science and Analytics: An overview from data-driven smart computing, decisionmaking and applications perspective,” SN Computer Science, vol. 2, no. 5, 2021. DOI: 10.1007/s42979-021-00765-8.

[31] J. Goldenfein, ”Algorithmic Transparency and Decision-Making Accountability: Thoughts for buying machine learning algorithms in office of the Victorian Information Commissioner (ed),” pp. 8-11, 2019. DOI: https://ssrn.com/abstract=3445873.

[32] Y. Duan, J. S. Edwardsand Y. K. Dwivedi, ”Artificial Intelligence for decision making in the era of big dataevolution, challenges and research agenda,” International Journal of Information Management, vol. 48, pp. 63-71, 2019. DOI: 10.1016/j.ijinfomgt.2019.01.021.

[33] R. R. Hoffman, G. Kleinand S. T. Mueller, ”Explaining explanation for ’explainable AI,’” in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 62, no. 1, pp. 197-201, 2018. [34] D. Gunning, E. Vorm, Y. Wangand M. Turek, ”DARPA’s explainable AI (XAI) program: A retrospective,” 2021.

[34] J.J. Smith, "Annual Report on Financial Data Breaches," Cybersecurity Institute, New York, USA, Rep. FDB-2023, Mar. 2024

[35] Impact of data breaches on financial institutions (2019-2023)A. Johnson and B. Lee, "Financial Implications of Cybersecurity Breaches," J. Financial Security, vol. 15, no. 3, pp. 245-260, Sep. 2023.World Economic Forum, "The Global Risks Report 2024," WEF, Geneva, Switzerland, Jan. 2024. https://www.weforum.org/reports/global-risks-report-2024

[36] H. B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. AISTATS, 2017, pp. 1273-1282.C. Dwork, "Differential Privacy," in Proc. ICALP, 2006, pp. 1-12.C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," in Proc. STOC, 2009, pp. 169-178.

[37] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, May 2015.

[38] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," IEEE Access, vol. 6, pp. 52138-52160, 2018.

[39] F. Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge, MA, USA: Harvard University Press, 2015.

[40] J. Kroll et al., "Accountable Algorithms," Univ. Pennsylvania Law Rev., vol. 165, no. 3, pp. 633-705, 2017.

[41] D. Gunning and D. Aha, "DARPA's Explainable Artificial Intelligence (XAI) Program," AI Magazine, vol. 40, no. 2, pp. 44-58, 2019.R. Guidotti et al., "A Survey of Methods for Explaining Black Box Models," ACM Comput. Surv., vol. 51, no. 5, pp. 1-42, Aug. 2018.

[42] R. Guidotti et al., "A Survey of Methods for Explaining Black Box Models," ACM Comput. Surv., vol. 51, no. 5, pp. 1-42, Aug. 2018