Risks and Vulnerabilities of AI in Financial Institutions

Darren Pritchard

IDS300W

UID: 01241796

Old Dominion University

Norfolk, VA, USA

dprit002@odu.edu

VULNERABILITIES AND RISKS OF AI IN FINANCIAL INSTITUTIONS

ABSTRACT

As artificial intelligence (AI) becomes increasingly prevalent in financial institutions, it brings with it both benefits and risks. On one hand, AI can improve efficiency, accuracy, and decision-making processes. On the other hand, it can also introduce new vulnerabilities and risks. This paper explores the major vulnerabilities and risks posed by AI in the context of financial institutions, with a focus on the intersection of finance, cybersecurity, and law. Through a literature review and analysis, we identify several key areas of concern, including data privacy and security, bias and discrimination, accountability and responsibility, and regulatory compliance. We also discuss potential strategies and solutions for mitigating these risks and ensuring that AI is used responsibly and ethically in financial institutions. Overall, this paper aims to provide a comprehensive overview of the risks and challenges that come with the use of AI in finance, and to promote a thoughtful and informed approach to its implementation.

Keywords: artificial intelligence, finance, financial institutions, risk, vulnerability, cybersecurity, law, regulation, ethics, decision-making, algorithmic bias

VULNERABILITIES AND RISKS OF AI IN FINANCIAL INSTITUTIONS

Introduction

Artificial Intelligence (AI) has been widely adopted by financial institutions due to its potential to revolutionize operations and drive profits. However, with the rapid advancement of AI, there are growing concerns about the major vulnerabilities and risks it poses in this context. This research paper aims to explore and analyze the major vulnerabilities and risks posed by AI in financial institutions. The research question that is addressed is, “What are the major vulnerabilities and risks posed by artificial intelligence in the context of financial institutions?” To address this question, an interdisciplinary approach is employed, drawing on the fields of finance, cybersecurity, and law. Through this interdisciplinary lens, the unique risks and challenges posed by AI in financial institutions are examined and the best practices and strategies to mitigate these risks are explored. An interdisciplinary approach allows us to analyze the complex issues related to AI in financial institutions from different perspectives, enabling us to gain a more comprehensive understanding of the risks and vulnerabilities.

In this paragraph, key terms used in this paper are defined, including AI, risks, vulnerabilities, training data, the various phases of AI development, and explainability and interpretability. AI refers to "a field of computer science and engineering devoted to creating machines that can perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and language translation" (Russell & Norvig, 2020, p. 2). Risks, as defined in cybersecurity, are the potential for an adverse outcome resulting from an event or activity, and its probability (European Union Agency for Cybersecurity, 2020,

p. 4), while vulnerabilities are weaknesses in an information system, system security procedures, internal controls, or implementation that could be exploited by a threat source (Stine, 2014). However, in the financial context, "risk" refers to the probability of loss or negative deviation from an expected outcome. This can be influenced by various factors such as market fluctuations, economic conditions, and unexpected events. One commonly used definition is provided by the International Organization for Standardization (ISO), which defines risk as "the effect of uncertainty on objectives." (ISO, 2018). For clarity, this paper details each as “cybersecurity risk” and “financial risk,” respectively. Training data are "data used to teach or train an AI model, which can be labeled (supervised) or unlabeled (unsupervised)" (Géron, 2019, p. 40), and the various phases of AI development include problem formulation, data collection and preparation, algorithm selection, model training, and evaluation (Russell & Norvig, 2020). Explainability refers to the ability of an AI system to explain its decision-making process to humans, while interpretability involves making an AI system's algorithms and processes more transparent and understandable to humans for analysis and auditing. Both concepts ensure the fairness, transparency, and trustworthiness of AI systems. (Brown & Fisch, 2021).

CYBERSECURITY

Cybersecurity is a crucial aspect of the study of artificial intelligence (AI) due to the numerous security risks and vulnerabilities associated with AI development and deployment. As noted by the European Union Agency for Cybersecurity (2020), the increasing adoption of AI in critical sectors such as healthcare and finance has created new opportunities for cyber attackers. This is especially true during the COVID-19 pandemic, where cyber attacks have increased

significantly due to the shift to remote work and the heightened demand for online services (Eian et al., 2020; Rameem Zahra et al., 2022). Therefore, cybersecurity measures such as data protection, threat detection, and incident response are vital to the safe and secure implementation of AI systems. As noted by the National Institute of Standards and Technology (2014), a robust cybersecurity framework can help organizations identify, protect against, and mitigate cyber threats. Ultimately, a comprehensive understanding of cybersecurity is necessary for the effective development and deployment of AI systems (Russell & Norvig, 2020).

FINANCE

The field of finance also contributes to the study of the risks and vulnerabilities of AI from a financial point of view. A dynamic capacity analysis during the COVID-19 pandemic conducted by Drydakis (2022) examined how artificial intelligence can reduce business risks for small and medium-sized enterprises (SMEs). Another study by Hoang et al. (n. d.) explored the use of federated artificial intelligence for a unified credit assessment, which reduces financial risk. ISO 31000:2018 (ISO, 2018) provides guidelines for risk management in the financial sector, which can be applied to the management of financial risks related to AI. The article by Vučinić and Luburić (2022) emphasizes the importance of risk-based thinking in financial, financial and cybersecurity technologies, emphasizing the need to take into account cybersecurity risks in financial technologies. Finance ultimately contributes to the study of AI risks and vulnerabilities by providing frameworks and information for risk management and mitigation. Hoang et al. notes the possibility of discrimination caused by biased training data, stating that credit reporting systems must meet this challenge to ensure the accuracy and

reliability of credit scores. In addition, the authors emphasize the importance of protecting the privacy and security of individuals while ensuring the transparency and explainability of AI-based decisions to strengthen trust and engagement with consumers. The authors propose a unified credit score system that integrates financial, social, contextual and technological characteristics, as shown in Figure 1. This system aims to provide a complete representation of each individual by incorporating both hardware data (financial and contextual data) and software data (social and technological information) in the calculation of credit scores. The inclusion of a wide range of data sources, including semantic, visual, mobile and tracking features, promises to improve credit scoring and better predict the likelihood of default behaviors.

Figure 1. Types and dimensions of information to be used for an AI-derived unified credit score. (Hoang et al.)

Type	Dimension	Description
Financial	Demographic	Personal details, including consumer profile, backgrounds, and socioeconomic measures
	Transaction	Transactional records, including purchases, inquiries and transfers
	Credit	Behavioural features related to historical credit activities
	Tenure	Time analysis features related to banking activities
Social	Behaviour	Social features related to how an individual conducts oneself in social networking sites
	Preference	Personal preferences and habits based on social profile
	Perception	Data features related to emotions, honesty, integrity, etc
	Connectivity	Social relationships including social networks associated with credit ratings and communications
	Content	Unstructured data features generated by consumers in social networking sites
Contextual	Geolocation	Geographical features, including location-based conditions and networks
	Time	Temporal and seasonal features
	Environment	Environmental features, including epidemiological and pollution conditions
	Community	Physical context features, including neighbourhood, groups and businesses
Technological	Semantic	Linguistic and philosophical features based on natural language modelling
	Vision	Visual features, including facial features and user-generated images/videos
	Mobile	Activity data features based on mobile usage and browsing behaviours
	Tracking	Movement-based features including sensor-based movements and positioning measures

LAW

Law plays a crucial role in the study of AI and its associated risks, especially in the financial market where one of the main challenges is to identify threats to consumers and implement the necessary changes to consumer protection legislation to address these risks (Nizioł, 2021). Regulations and guidelines are needed to ensure that AI systems are developed and used responsibly, ethically and transparently. The European Union Agency for Cybersecurity (ENISA) has published recommendations for Europe in the field of AI cybersecurity, highlighting the need for a comprehensive approach covering the entire AI development process. In addition, the General Data Protection Regulation (GDPR) provides a legal framework for data protection that is relevant for the development and use of AI systems. The National Institute of Standards and Technology (NIST) has also published a framework for improving the cybersecurity of critical infrastructures, which includes guidelines for managing cybersecurity risks associated with AI systems. The legal implications of AI and its associated risks are also explored by lawyers and practitioners. For example, Russell and Norvig's book (2020), "Artificial Intelligence: A Modern Approach", provides an overview of the legal and ethical issues related to AI, including issues of confidentiality, accountability and transparency. As AI continues to progress, it is essential that the legal framework keeps pace with technological developments and addresses the risks associated with AI.

COMMON GROUND

While the fields of cybersecurity, finance, and law differ in their approach to AI, they all share a common goal of mitigating the risks associated with AI development and deployment.

VULNERABILITIES AND RISKS OF AI IN FINANCIAL INSTITUTIONS

In the context of AI, risks are defined as the potential harm that can arise from the use of AI, including privacy violations, cyber attacks, and financial losses. To address these risks, all three fields emphasize the importance of developing and implementing effective risk management strategies. Additionally, they recognize the need for robust data governance practices and for ensuring that AI systems are transparent, explainable, and accountable. The National Institute of Standards and Technology (NIST) Cybersecurity Framework and the European Union Agency for Cybersecurity's (ENISA) recommendations for cybersecurity for AI both provide frameworks for managing risk that are widely adopted in these fields. Finally, all three fields acknowledge the importance of interdisciplinary collaboration to address the challenges of AI development and deployment, and recognize the need for ongoing dialogue and cooperation to ensure that AI is developed and deployed responsibly and ethically.

DISCIPLINARY CONFLICTS

Despite the common ground in their concerns, the different disciplines involved in the study of AI often approach the topic from different angles, which can lead to conflicts. One such conflict arises between computer science and law. Computer scientists tend to focus on the technical aspects of AI, including its design and development, while legal scholars focus on the ethical and legal implications of AI use. This difference in focus can lead to a lack of understanding and communication between the two disciplines, resulting in challenges in regulating and managing AI. For instance, computer scientists may design AI systems that violate privacy or perpetuate bias, while lawyers may not fully understand the technical limitations and capabilities of AI. To address these conflicts, interdisciplinary collaboration is

crucial, and policymakers must work to create a shared language and understanding between the different disciplines involved in the study of AI.

One way to construct a more comprehensive understanding of the topic expressed in our research question is to engage in more interdisciplinary research. By bringing together experts from different fields, such as law, cybersecurity, and finance, we can develop a more holistic view of the issues surrounding AI development and implementation. Additionally, further research could focus on identifying areas of overlap and divergence in the perspectives of different disciplines. This would allow us to identify potential gaps in our understanding and develop more comprehensive solutions that take into account the perspectives of multiple stakeholders. Finally, it is important to recognize that the development and implementation of AI is an ongoing process, and that continued interdisciplinary research is necessary to address emerging issues and adapt to new challenges. By working together, researchers from different disciplines can help ensure that AI is developed and deployed in a way that is safe, ethical, and beneficial for society as a whole.

To further develop and refine the understanding and theories presented in this interdisciplinary study, future research could consider several avenues. For instance, including additional disciplinary perspectives could enrich the findings and provide a more nuanced understanding of the topic at hand. Moreover, conducting further fieldwork in specific regions or with particular groups could reveal additional insights and potentially broaden the applicability of the theories presented. Additionally, testing the theories through a variety of methods, such as experimental studies or longitudinal analyses, could strengthen the evidence base for the

interdisciplinary findings. Furthermore, communicating the insights and theories to a broader audience, including policymakers, and the general public, could help bridge the gaps between the different disciplinary approaches and facilitate a more holistic and effective response to the issues at hand.

CONCLUSION

In conclusion, the development and deployment of AI raise complex ethical, social, and legal issues that need to be addressed in a collaborative and multidisciplinary manner. The research has shown that AI technology can provide numerous benefits across various domains, such as healthcare, finance, and education. However, it also presents various risks and vulnerabilities, such as data breaches, cyber attacks, and biases that can affect individuals and society. Different disciplines, such as law, computer science, and philosophy, have contributed to the study of AI and its ethical implications. While each discipline brings its unique perspective, there is also common ground, such as the importance of transparency, accountability, and human-centered design. Through a dynamic capabilities analysis, Drydakis (2022) demonstrates the value of artificial intelligence in mitigating risks for small and medium-sized enterprises. Hoang et al. (n.d.) illustrate the potential for federated artificial intelligence to improve credit assessment, while Vučinić and Luburić (2022) argue that fintech and risk-based thinking are essential in managing cyber risks.

Despite disciplinary conflicts, the research shows that interdisciplinary collaboration can lead to significant common ground and a more comprehensive understanding of the topic at hand. As we reflect on this research, it becomes clear that future studies could benefit from

incorporating additional disciplines or testing the theories presented in new contexts. At the same time, disciplinary conflicts arise, such as different views on the level of regulation, responsibility, and the role of AI in society. As AI continues to advance, it is crucial to have ongoing interdisciplinary dialogue and collaboration to ensure that the technology is developed and deployed in a responsible and ethical manner.

REFERENCES

Drydakis, N. (2022). Artificial Intelligence and reduced smes' business risks. A dynamic capabilities analysis during the COVID-19 pandemic. *Information Systems Frontiers*, *24*(4), 1223–1247. https://doi.org/10.1007/s10796-022-10249-6

Eian, I. C., Yong, L. K., Li, M., Qi, Y. H., & Z, F. (2020). Cyber attacks in the era of COVID-19 and possible solution domains. https://doi.org/10.20944/preprints202009.0630.v1

Hoang, M.-D., Le, L., Nguyen, A.-T., Le, T., & Nguyen, H. D. (n.d.). *Federated Artificial Intelligence for Unified Credit Assessment*. https://doi.org/10.48550/arXiv.2105.09484

Rameem Zahra, S., Ahsan Chishti, M., Iqbal Baba, A., & Wu, F. (2022). Detecting covid-19 chaos driven phishing/malicious URL attacks by a fuzzy logic and data mining based Intelligence System. *Egyptian Informatics Journal*, *23*(2), 197–214. https://doi.org/10.1016/j.eij.2021.12.003

Vučinić, M., & Luburić, R. (2022). Fintech, risk-based thinking and cyber risk. *Journal of Central Banking Theory and Practice*, *11*(2), 27–53. https://doi.org/10.2478/jcbtp-2022-0012

Nizioł, K. (2021). The challenges of Consumer Protection Law connected with the development of artificial intelligence on the example of Financial Services (chosen legal aspects). *Procedia Computer Science*, *192*, 4103–4111.

https://doi.org/10.1016/j.procs.2021.09.185

Cybersecurity of AI and standardisation. ENISA. (2023, March 15). Retrieved April 5, 2023, from https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation

Géron, A. (2019). *Hands-on machine learning with scikit-learn, Keras, and tensorflow: Concepts, tools and techniques to build Intelligent Systems*. O'Reilly.

Stine, K. (2014). Framework for improving critical infrastructure cybersecurity, version 1.0 . https://doi.org/10.6028/nist.cswp.1

Russell, S. J., & Norvig, P. (2020). Artificial intelligence: A modern approach. Pearson.

ISO. (2018). Risk management – Guidelines. ISO 31000:2018. Retrieved from https://www.iso.org/standard/65694.html

Brown, T. B., & Fisch, A. (2021). Language models are few-shot learners. In Advances in Neural Information Processing Systems (Vol. 34, pp. 18732-18742).