



Applications & Implications of Big Data Analytics and AI in Finance

Maschio Fernando¹, Dr. A. Shaji George², Dr. K. Krishnamoorthy³

Datalk V5 Technologies Private Limited, Chennai, India¹

Masters IT Solutions, Chennai, India²

Professor, Department of Computer Science & Engineering, Sudharsan Engineering College, Tamilnadu, India³

Abstract: Over the past two decades, the financial sector has seen a shift in how individuals and businesses operate. The adoption of AI and Big Data analytics has transformed the manner in which financial institutions operate, and how they interact with customers and other institutions. Several researchers opine that the financial sector will see further transformation due to the increased adoption of technologies that enable humans to dedicate more time in innovating and performing sophisticated tasks. On a daily basis, the financial sector tracks billions of market events and generates massive and diverse amounts of data that falls under the category of Big Data. AI models and algorithms that utilize Big Data to spot patterns and glean insights in order to make decisions in areas such as portfolio strategy and fraud detection have become increasingly common in the finance sector. Organizations in the financial sector have a varying level of capability and competency with respect to the adoption and utilization of these technologies. The utilization of these technologies and the implications of their use has become the subject of debate in the financial sector. The goal of this paper is to report on the applications of AI and Big Data analytics in finance, and the ethical, organizational, and legal repercussions of the use of these technologies.

Keywords: Big data analytics, AI in finance, Implications of Big data, Fintech, Challenges of Big data and AI in finance, Applications of Big data and AI in finance.

I. INTRODUCTION

Historically, the financial sector has rapidly adopted new technologies in order to gain a competitive advantage in financial markets. Thomas Alva Edison's improved version of the stock ticker was widely used to transmit stock information over the telegraph [1]. The Federal Reserve also utilized an early electronic system to send and receive payments [2]. The popularization of Big Data by John Mashey [3] in the late 90's has dramatically changed how industries operate. According to a survey of organizations across various industries in 2014, 87% of organizations believed that Big Data analytics would change the competitive landscape of their industry within the next three years [4]. 89% of the surveyed organizations also believed that companies that do not adopt a strategy to utilize Big Data analytics within the next year risked losing market share and momentum. The combination of structured and unstructured data and the volume of data generated has been both an opportunity and a challenge for the financial sector. Financial services such as asset and wealth management, crowdfunding platforms, cryptocurrencies and algorithmic trading have benefited enormously by using insights gleaned from Big Data in order to gain a competitive advantage.

Furthermore, artificial intelligence (AI) is used to perform tasks that would traditionally have required humans to perform to completion. AI is increasingly used in the financial sector to forecast credit risk, handle credit underwriting and portfolio strategies, fraud detection and mitigation [5]. AI can help reduce costs such as fees and commissions, increase the personalization levels of financial products offered to customers, analyze creditworthiness of individuals and enterprises with a minimal credit history and enable quicker and more accurate decision-making [7]. The adoption of Big Data analytics and AI in the financial sector has also posed a lot of challenges and risks. Managing the volume of data generated and used in finance has also been a challenge. Ensuring data security and reliability by having a robust data governance framework is critical. The financial sector is also required to be compliant with regulatory requirements that govern the accuracy and access of financial data used in reports. The use of inaccurate and poor-quality data by an AI has the potential of making biased or discriminatory decision-making that negatively affect financial consumers. The inability or difficulty in explaining the decision-making process of an AI may make it difficult to comply with the current supervisory and regulatory requirements, which will not have taken the increased deployment of AI into consideration. This article aims to provide an overview of the applications of Big Data analytics and AI in finance, and the policy implications of using these technologies in the present day and moving forward. This



article also aims to showcase areas where the use of Big Data analytics and AI could be beneficial, and how future regulations and policies need to take into account the increasing adoption of these technologies.

II. LITERATURE REVIEW

Big Data and AI has been extremely influential across multiple business areas. Research and innovations surrounding these two technologies has been useful in predicting economic patterns, creating sophisticated risk controls, visualizing financial market data and enabling businesses to make better decisions at a much quicker pace.

Big Data has had a major effect in various disciplines such as marketing [8], [9], sales [10], [9], human resources management [11], research and development and finance [12]. Hasan et al.'s [12] study on the state of Big Data in finance discovered that Big Data is widely used in financial institutions, financial markets and digital marketplaces. Billions of data points are generated everyday by various activities such as purchasing, updating personal details or opening new lines of credit. Using the data generated, businesses predict patterns for customers and themselves. For instance, they may use this data to help customers to keep track of their spending habits, and to evaluate the creditworthiness of a business. Schroeder's article [9] expands on Hasan et al.'s study by painting a picture of how Big Data is used, what opportunities have been identified and the various pathways towards their implementation and the challenges that organizations face in the present and the future. A common theme amongst the interviewees for this article was the need to perform a cost-benefit analysis of collecting data, since there is a prevailing belief that the more the volume of data, the better it is. This study also raises the question of legal and regulatory issues, and how the lack of clarity in the regulatory and legal environments surrounding data is emerging as a critical issue. Big data is growing at a much quicker pace than the regulatory bodies governing them. Richard Mason's privacy, accuracy, property and accessibility (PAPA) [13] framework regarding the issues of the information age was re-examined and applied in the context of Big Data analytics by Richardson et al. [14]. Although Mason's framework was defined in a pre-World Wide Web world, it still is relevant in identifying the high-level issues surrounding the use of Big Data analytics in finance. Richardson et. al specifically note that privacy risks, discrimination and negative outcomes from targeted marketing are all possible negative effects that need to be considered.

The recent growth of FinTech companies powered by artificial intelligence (AI) is proof of how AI is impacting the financial sector, the economy and society. Research areas such as business intelligence, data science, economics and finance are becoming an increasing glamorous and critical area for AI applications. AI is being used to provide and power services like automated chat-bots for customer service and payment infrastructure. Cao's review of the state of AI applications in finance identifies a number of high-level applications of AI research in the financial sector such as modelling economic-financial mechanisms, financial market analysis, agent-based economics, intelligent investment and risk management [5]. One area where AI is making huge strides in industry is in environmental, social and governance (ESG) investing. Hughes et al.'s paper on how technological advancements are impacting ESG investing describe how the use of AI has directly led to the creation of a new type of ESG ratings provider that can largely replace the work of human research analysts [15]. This particular instance of AI use has also led to the reduction of human biases and errors, since most research analysts are affiliated with large financial institutions that evaluate ESG investments with the use of proprietary methodologies. The use of AI is steadily increasing in the asset management sector to manage allocations and to bring down operational costs. Kato Yasuyuki [6] reports that there are 10 investment trusts sold in Japan that use AI in their workflows as of 2018. Yasuyuki further states that there will come a time in asset management where it will be impossible to participate in the financial markets without having an AI model. While the adoption of AI in decision-making processes has been transformational, there are some challenges that need to be resolved. An investigation by The Markup revealed that US-based lenders that use algorithms to determine the eligibility of mortgage applicants were more likely to reject applicants of a minority background than white applicants of similar financial characteristics [16]. When an applicant is rejected, lenders are required to tell them why they were rejected [17]. This creates a timeline and paper-trail of decisions so that lenders can be held accountable for their decision-making process. The disclosure of this information also lets the applicant to improve their chances of availing credit in the future. However, a decision made by an AI is hard to explain, since there is no clear delineation of the decision-making process [18]. The exponential adoption of these new technologies has laid bare the issues present in both the social and economic structures [17], [18]. Thus, the existing anti-discrimination frameworks and legislation need to be re-examined and redefined in order to remedy the significant issues in the status quo.



III. METHODOLOGY

The purpose of this paper is to paint a picture of the current landscape of the applications and the policy and legal environment concerning the use of Big Data analytics and AI in finance. Secondary data sources such as Google Scholar, the European Banking Authority and Mendeley Data, and other published works were mainly used. Although there is an established corpus of research for Big Data analytics and AI, research into the applications and policy aspects of these technologies is not particularly extensive. The literature search primarily focused on academic and peer-reviewed works, but news reports and articles were used in some instances. Works relating to portfolio management, risk management, financial analysis and forecasting, digital economies, consumer banking, and fraud detection were found to have a relationship with Big Data analytics and AI. A preliminary search was performed to identify potentially relevant literature. The abstract and the body of the chosen literature were then used to filter the literature for use in this paper.

IV. APPLICATIONS

The financial sector generates a large volume of structured and unstructured data in a very short period of time. The New York Stock Exchange is reported to have been generating 4 to 5 terabytes of data a day in 2013 [19]. With the improvement in information systems and network infrastructure, this figure will only continue to increase. Portfolio management, risk analyses, market valuations, index performance tracking, and option pricing are some examples where Big Data comes into play in the financial sector. With the increasing adoption of Big Data in finance, gleaning insights from historical performances has become a much simpler task. This has also led to an increase in institutions attempting to predict the future in order to maximize gains regarding investment opportunities and outcomes.

Another source of Big Data for the financial sector is the need to stay compliant with current and future regulations. After the 2008 financial crisis, institutions are required to provide more data and detailed reports of risks involving their business to regulatory authorities. This has led to an increased demand for regulatory data, so that institutions can continue to do business. Regulatory requirements have become increasingly data-driven, with routine stress tests being performed to gauge the health of an institution. Thus, it is crucial for institutions to collect and manage the quality and the security of the data that they generate.

The use of Big Data analytics in the financial sector has dovetailed with the increased adoption of AI. AI uses computational and modelling methods to understand market conditions, design trading strategies and forecast risk. A major success of the application of AI is in fraud detection and analysis. By analyzing transaction data, AI can perceive patterns and flag suspicious transactions that may be of a high monetary value or performed by an unexpected entity. This cuts down on the time taken during the review and mediation process.

AI is also increasingly used in providing personalized services to customers and institutions alike. Onboarding a new client is made easier with the use of AI models to process signed documents that are required for regulatory purposes and send the results back to decision-makers. AI-powered chatbots enable customers to access bank accounts, apply for credit lines, tracking spending habits and follow up on applications or requests. Using AI-powered applications in a customer-facing scenario simplifies customer request processing and enables institutions to provide round-the-clock service to their customers.

ESG investing and scoring is another field where the use of Big Data analytics and AI has been successful. Traditionally, ESG research has been the domain of human researchers. The volume and the variety of the data used has led to inconsistencies and how organization report their analyses. ESG analysis and scoring typically involves finding relative differences between organization in the same industry. ESG disclosures have only begun to be regulated at a high level [15]. There is no standardized method of ESG scoring, because different organizations use different frameworks designed by their analysts. Thus, measuring the ESG practices of an organization and assigning a score can lead to inconsistency and a lack of interpretability. The use of Big Data analytics and AI enables organizations to apply multiple frameworks to score organizations on their ESG practices. This also enables one-to-one comparisons to be made between organizations in the same industry.

V. CHALLENGES

The adoption of Big Data analytics and AI has brought about many technical challenges. Storing and manipulating vast volumes of data is a critical issue in many organizations. In an organization, there are various internal and external data sources and various requirements that make it challenging to gather data. Consolidating these disparate sources and



taking legacy information systems into account is a big challenge. Since the financial sector generates large volumes of data on a daily basis, this issue is only exacerbated by the adoption of Big Data analytics and AI. The volume and variety of Big Data means that bringing all the required data together in a way that can be used by analytics needs to be a organization-wide effort. Another challenge with using these technologies is ensuring that the quality of data used is good enough. Training AI models requires large datasets, and if inaccurate data is used, that will lead to poor outputs. Having high-quality data in itself is not helpful if a significant amount of time is taken to process and analyze the data. Organizations often focus on getting the highest quality data possible at the expense of gaining timely insights that will be useful in a real-time situation. Thus, the challenge is to ensure that the quality and quantity of the data used is balanced with the need to make timely decisions.

Similarly, there are ethical issues that need to be addressed with the adoption of these technologies. Using large datasets requires access to large volumes of data with personally identifiable information. Privacy policies that give users control over how their data is used is insufficient, since there is a lot of personal data that is collected, analyzed and improperly accessed. The security and safety of sensitive data and communicating how the data is used and who uses it in a transparent manner is crucial. The lack of explain ability of an AI model, even by the people who create it, poses many questions. AI models cannot be examined to reproduce a series of steps that were taken in the decision-making process to arrive at a decision. This causes a lack of confidence in these models since there needs to be a level of transparency and accountability to ensure that decisions are made equitably. Ensuring that the biases of humans aren't passed down to an AI during the design and implementation phase is another key issue. If incomplete or very narrow data is utilized to train an AI, any applications dependent on AI decision-making is bound to be flawed. This can lead to costly failures and damage to an organization's reputation.

VI. LEGAL AND POLICY IMPLICATIONS

There is a large gap between organizational policy frameworks and the use of Big Data analytics and AI. Ensuring privacy and security is a major concern, since the privacy and security frameworks utilized by organizations may not be well-equipped to handle the scale of data that is being generated today. Traditionally, the approach utilized to handle issues regarding the use of Big Data analytics and AI in finance has been reactive rather than proactive. Consequently, data breaches and leaks are becoming a common occurrence. The traditional models and frameworks of information systems security needs to be revised in order to handle these issues. A data governance structure needs to be created that defines who is responsible for enforcement of policies, and how data is handled and protected throughout organizations. Anonymizing data to remove personal identifiers, communicating how the data is utilized and what frameworks are in place to secure it and protect privacy in a transparent manner with stakeholders such as employees and clients is also vital. There also needs to be an option for stakeholders to withdraw consent and delete their records. Organizations must also have a system of controls that gatekeep employee access to data and AI models that utilize sensitive data. Gatekeeping access and segregation of duties will restrict access to sensitive information based on the user role and job function. If sensitive data such as personal identifiers and financial information is shared with third parties, the onus is on organizations to ensure that the third party's utilization is in line with their own standards and perform periodic audits to ensure that there are no malpractices.

There is a similar gap between current legislation and the use of Big Data analytics and AI. The lag between technological advancement and legislation means that there are currently very few regulatory laws across the world that specifically address the use of AI and Big Data. The existing legislation focus on protecting the privacy of individuals and organizations in specific circumstances and require data to be anonymized. However, achieving full anonymization of data is challenging in practice because of the large volume of data across multiple sources that can be used for cross-referencing. By doing so, the identities of individuals and organizations that are supposed to remain private can be learned and potentially misused. The use of AI models to make decisions to serve customers has also meant that bias in AI has become a legal challenge that needs to be addressed. Like other businesses, financial institutions are legally required to follow anti-discrimination laws across the world. The use of AI models that are flawed or biased means that discriminatory and biased decisions are becoming more commonplace. This has the potential of alienating stakeholders, falling afoul of legislation and harming an organization's reputation and revenue. Since AI models are designed by humans, human biases are easily introduced into the model. The datasets used to train AI models to make decisions can also lead to inequitable decision-making if the datasets are of poor quality. This could have negative impacts on decision-making for situations such as credit line applications, mortgage applications and tax compliance. Thus, the increased use of AI and Big Data analytics means that financial institutions should expect to see increasing legislative and judicial scrutiny with regard to how these technologies are deployed and how they affect consumers.



VII. CONCLUSION

Big Data analytics and AI have revolutionized the manner in which financial services and products are designed and the adoption of these technologies has only grown due to the COVID-19 pandemic. Financial institutions have had to adapt their business models to compete in a market and economy that is becoming ever-increasingly digital. The applications of Big Data analytics and AI in finance include portfolio analysis and strategy, automating repetitive workflows, improving risk analysis and fraud detection and mitigation. These technologies have enabled employees and customers to focus their energies on performing complicated tasks and innovation. However, the use of these technologies has been the subject of debate. The quality of data used and the transparency and impartiality of AI that are used in decision-making processes needs to be ensured. Compliance with regulators and legislative bodies is another issue. The increased exposure to privacy and cybersecurity risks, and the collection and storage of sensitive data and models needs to be addressed in order to project confidence and transparency. There also needs to be organizational-level policy implementations to ensure data stewardship practices are followed and third parties who have access to sensitive information. Despite the financial sector generating vast volumes of quantitative data, there is currently very few empirical studies in the use of Big Data and AI in finance. Further areas of exploration are the equitable access of public sector data, and the explain ability of models used in the finance sector. Reconciling the exponential growth in the financial sector by the use of Big Data analytics and AI and the questions that are posed by their utilization will continue to be an open question.

REFERENCES

- [1]. "Stock Tickers: The Thomas A. Edison Papers at Rutgers University." Accessed: Nov. 04, 2021. [Online]. Available: <http://edison.rutgers.edu/ticker.htm>
- [2]. "Fedwire® and National Settlement Services - FEDERAL RESERVE BANK of NEW YORK." Accessed: Nov. 04, 2021. [Online]. Available: <https://www.newyorkfed.org/aboutthefed/fedpoint/fed43.html>
- [3]. S. Lohr, "The Origins of 'Big Data': An Etymological Detective Story," Bits Blog, Feb. 2013. Accessed: Nov. 03, 2021. [Online]. Available: <https://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/>
- [4]. L. Columbus, "84% Of Enterprises See Big Data Analytics Changing Their Industries' Competitive Landscapes In The Next Year," Forbes. Accessed: Nov. 03, 2021. [Online]. Available: <https://www.forbes.com/sites/louiscolumbus/2014/10/19/84-of-enterprises-see-big-data-analytics-changing-their-industries-competitive-landscapes-in-the-next-year/>
- [5]. L. Cao, "AI in Finance: A Review," Social Science Research Network, Rochester, NY, {SSRN} {Scholarly} {Paper} ID 3647625, Jul. 2020. doi: 10.2139/ssrn.3647625.
- [6]. K. Yasuyuki, "AI/Fintech and Asset Management Businesses," Public Policy Review, p. 28, 2020.
- [7]. M. Macpherson, A. Gasperini, and M. Bosco, "Artificial Intelligence and FinTech Technologies for ESG Data and Analysis," SSRN Electronic Journal, 2021, doi: 10.2139/ssrn.3790774.
- [8]. P. Chintagunta, D. M. Hanssens, and J. R. Hauser, "Editorial—Marketing Science and Big Data," Marketing Science, vol. 35, no. 3, pp. 341–342, May 2016, doi: 10.1287/mksc.2016.0996.
- [9]. R. Schroeder, "Big data business models: Challenges and opportunities," Cogent Social Sciences, vol. 2, no. 1, p. 1166924, Dec. 2016, doi: 10.1080/23311886.2016.1166924.
- [10]. H. Hallikainen, E. Savimäki, and T. Laukkanen, "Fostering B2B sales with customer big data analytics," Industrial Marketing Management, vol. 86, pp. 90–98, Apr. 2020, doi: 10.1016/j.indmarman.2019.12.005.
- [11]. S. Zang and M. Ye, "Human Resource Management in the Era of Big Data," Journal of Human Resource and Sustainability Studies, vol. 3, no. 1, p. 41, Feb. 2015, doi: 10.4236/jhrss.2015.31006.
- [12]. Md. M. Hasan, J. Popp, and J. Oláh, "Current landscape and influence of big data on finance," Journal of Big Data, vol. 7, no. 1, p. 21, Mar. 2020, doi: 10.1186/s40537-020-00291-z.
- [13]. R. O. Mason, "Four Ethical Issues of the Information Age," MIS Quarterly, vol. 10, no. 1, pp. 5–12, 1986, Accessed: Nov. 11, 2021. [Online]. Available: <https://www.jstor.org/stable/248873>
- [14]. S. Richardson, S. Petter, and M. Carter, "Five Ethical Issues in the Big Data Analytics Age," Communications of the Association for Information Systems, vol. 49, no. 1, Nov. 2021, doi: 10.17705/1CAIS.04918.
- [15]. A. Hughes, M. A. Urban, and D. Wójcik, "Alternative ESG Ratings: How Technological Innovation Is Reshaping Sustainable Investment," Sustainability, vol. 13, no. 6, p. 3551, Jan. 2021, doi: 10.3390/su13063551.
- [16]. "The secret bias hidden in mortgage-approval algorithms," AP NEWS. Aug. 2021. Accessed: Nov. 20, 2021. [Online]. Available: <https://apnews.com/article/lifestyle-technology-business-race-and-ethnicity-mortgages-2d3d40d5751f933a88c1e17063657586>
- [17]. A. Klein, "Reducing bias in AI-based financial services," Brookings. Jul. 2020. Accessed: Nov. 20, 2021. [Online]. Available: <https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/>
- [18]. N. Vigdor, "Apple Card Investigated After Gender Discrimination Complaints," The New York Times, Nov. 2019, Accessed: Nov. 20, 2021. [Online]. Available: <https://www.nytimes.com/2019/11/10/business/apple-credit-card-investigation.html>
- [19]. T. Groenfeldt, "At NYSE, The Data Deluge Overwhelms Traditional Databases," Forbes. Accessed: Nov. 23, 2021. [Online]. Available: <https://www.forbes.com/sites/tomgroenfeldt/2013/02/14/at-nyse-the-data-deluge-overwhelms-traditional-databases/>