iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

BIAS AND FAIRNESS IN ARTIFICIAL INTELLIGENCE APPLICATIONS WITHIN THE FINANCIAL SECTOR

Syed Moiduddin Rashad
Department of Computer Science, College of Science,
Northern Border University, Saudi Arabia.
syed.rashad@nbu.edu.sa

Mohamed Asim
Department of Computer Science, College of Science,
Northern Border University, Saudi Arabia.
Mohamed.asim@nbu.edu.sa

ABSTRACT
AI-powered automation of credit evaluation, fraud detection, and investment decision-making is revolutionizing the financial sector. Nevertheless, the widespread use of AI in finance is also raising considerable concerns about bias and fairness. Bias is when algorithms, which have been trained on biased or unbalanced data, perform in an unfair or discriminatory manner. In finance, this means that some people would have a limited credit access, make unjustified loan rejections, or a particular demographic group would be excessively over-flagged for triggering of transactions.
These biases are very often invisible and can come from biased datasets, improperly chosen features, or implicitly from human decisions taken during the designing of models. Once such AI is released, it incites feedback loops that may deepen the original inequalities, damage the principle of financial inclusion, and decrease the level of trust of people in the system.
To tackle such issues, companies have decided to incorporate data practices that are sensitive to fairness, perform algorithmic audits, and use explainable AI models like LIME and SHAP. Governance frameworks based on ethics and worldwide rules such as the EU AI Act, which set transparency and accountability as prerequisites, are becoming the norm.
In the end, it is the use of AI in a morally right way that calls for a proper balancing of ethics and innovation. By commitment to fairness, constant oversight, and a multitude of voices in the design of AI, financial institutions will be able to deploy technology that will be a tool for equity, trust, and growth that is sustainable in the long run.

1. INTRODUCTION

With the help of AI, banks and other financial institutions are going to be able to make decisions faster and cheaper, and hence will be able to gain a competitive advantage. A report by McKinsey in 2024 states that more than seventy percent of global banks have already implemented AI in credit scoring, risk modeling, and fraud analytics. What started as banking digitalization merely changed the bargaining power by which banks exercise and the risks have transformed into the permanence of biases in the systems. Discriminatory (embedded) effects - examples can be given from credit limits unequal or unfavorably declined loans - eventually can erode the trust of the public. The absentee of this NGO watchdog, the cognitive bias, represents the underlying issue of the AI dilemma and this paper acts as the call urging the utmost ethical AI principles to serve equitable outcomes in financial systems.

2. Sources and Types of Bias in AI-Driven Finance

Bias of AI in finance is an issue that is buried within different layers - the process of data capture, feature engineering, development of the model, and the model's utilization. Bias might be the result of a visibly skewed sample or, involuntarily, from proxy variables. Below is a thorough table of the major kinds of biases that happen in AI-financial systems.

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

Type of Bias	Stage of Occurrence	Description	Example in Finance
Data Bias	Data Collection	Reflects historical inequalities in datasets.	Credit scores skewed by income group.
Sampling Bias	Data Preparation	Sample underrepresents a demographic.	Fewer low-income borrowers in data.
Algorithmic Bias	Model Design	Algorithm learns spurious correlations.	Penalizing certain ZIP codes.
Proxy Bias	Feature Engineering	Variables indirectly encode sensitive traits.	Using postal codes as income proxies.
Feedback Bias	Post Deployment	Models reinforce existing inequities over time.	High-risk tags leading to fewer futur e loans.

Figure 1: Bias in the AI Lifecycle for Financial Systems (Data → Model → Decision → Feedback → Retraining).

3. Case Studies: Credit Scoring, Fraud Detection, and Trading Bias

Bias is a major challenge in all financial applications to which the technology is applied. In the case of the Apple Card in 2019, one of the major criticisms was that the credit limits given to women were dramatically lower than those given to men with similar profiles. The incident highlights how AI systems can function as black boxes, which, as a consequence, can be risky. Due to the imbalance in the data, fraud detection models are more likely to flag minority transactions as more fraudulent. Algorithmic trading can lead to market bias as well, which happens through cyclical feedback loops where models reinforce past market behavior.

Figure 2: Illustration of Bias in Various Financial AI Applications (Credit, Fraud, and Trading Systems).

4. Impact of Bias on Financial Equity and Consumer Trust

The presence of bias within automated financial decision-making structures is not solely a data issue, rather, it threatens the core qualities that modern finance is built on. Equity, accessibility, and institutional integrity are the major pillars that regulate, sustain, and foster the financial systems of the world. When AI models that are based on skewed or partial data sets generate biased results, then, by a chain of unintended consequences, the disadvantage of certain individuals or communities follows. The repercussion of this may be that unfair credit scoring, unequal loan approvals, or misclassified risk groups come out of the decisions, thus causing financial exclusion of those who may be fully creditworthy. The latter not only causes the direct harm of individuals, but also is a hurdle of achieving broader goals, such as financial inclusion and social mobility that are interlinked.

If bias is left unchecked it will create systemic risks that are woven into financial ecosystems. In the scenario where a certain group of people is always underserved, thus not practically visible enough, capital allocation will be inefficient and even unbalanced. The absence of the diversity of the financial audience will not only limit the growth of the market but also lead to economic instability in the long run. As per the report of the U.S. Federal Reserve (2023), more than twenty large banks have been under the surveillance of regulators due to incidents connected with algorithmic bias which has become widely spread and has a significant impact. The result of this plague, however, is not only reputational– institutions found using biased AI systems may experience penalties, lawsuits, loss of customer trust, among others.

Could it be that one of the most serious problems with bias is that it leads to the perception of this phenomenon that, thus, discourages participation the most? Groups of people that had been marginalized in the past and, thus, excluded by systems operated by humans, would most probably now perceive systems based on digital platforms in the same way though with different degrees of skepticism. When there is a belief prevalent among the members of a community that AI “decides against them,” they stop engaging with formal financial systems and prefer informal or cash-based transactions instead. Thereby such a phenomenon works as a double-edged sword in the sense that the economic gap increases by these actions becoming less visible to the credit that is the main channel of access to growth opportunities. In consequence, the matter of fairness in AI is far beyond the issue of compliance: it is linkage to public trust, inclusivity, and long-term stability of the global financial ecosystem.

5. Mitigation Frameworks and Ethical AI Strategies

Efforts to reduce bias in decision-making algorithms require implementation of design features that anticipate and address the issue as well as a proper structure for management. Approaches to the problem work on three

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

steps: pre-processing, in-processing, and post-processing. Table 2 provides a concise overview of the performing methods in financial organizations.

Stage	Technique	Financial Application Example
Pre-processing	Data reweighing, resampling	Balancing demographic representation in credit datasets.
In-processing	Fairness constraints, adversarial debiasing	Ensuring equal opportunity in credit approvals.
Post-processing	Threshold adjustment, recalibration	Equalizing false positive rates in fraud detection.

6. Comparative Fairness Metrics in Finance

Fairness metrics are used to quantify the degree of fairness of a model's performance. Regulators in finance are progressively requiring these measures as part of model validation. Table 3 lists the most commonly used fairness metrics together with their importance for financial systems.

Fairness Metric	Definition	Financial Use Case
Demographic Parity	Equal selection rates across demographic groups.	Credit approval fairness.
Equal Opportunity	Equal true positive rates across groups.	Loan repayment prediction.
Predictive Parity	Equal precision rates across demographics.	Fraud prediction parity.
Equalized Odds	Equal false positive and true positive rates.	Credit scoring systems.
Calibration	Predicted scores match actual outcomes.	Risk scoring validation.

7. FUTURE DIRECTIONS FOR RESPONSIBLE AI IN FINANCE

AI systems for finance to follow will have as their integral features the continuous bias checking and the availability of fairness dashboard. Both EU AI Act and the U.S. National AI Initiative are examples of global efforts that promote openness, accountability, and clear understanding. Next steps entail federated learning as a privacy-friendly mode of data analysis, synthetic data to make training sets balanced, and baking fairness-by-design in each step of development.

8. CONCLUSION

The finance world powered by AI is an exciting vision to materialize with lots of benefits for the efficiency, accuracy, and innovation; however, fairness ought to be retained as one of the main pillars of the enterprise. By ethics incorporated in design, diligent auditing, and presence of divergent viewpoints, institutions are enabled to build systems that are able to deliver democratic results. The implementation of responsible AI should be considered as more than just compliance, and rather as a moral and a strategic necessity for long-term financial stability.

REFERENCES

1) Kagalkar, Anup, Akshay Sharma, Bhushan Chaudhri, and Satish Kabade. 2024. “AI-Powered Pension Ecosystems: Transforming Claims, Payments, and Member Services”. *International Journal of AI, BigData, Computational and Management Studies* 5 (4): 145-50.
   a. https://doi.org/10.63282/3050-9416.IJAIBDCMS-V5I4P115.
2) *Fnu, H., & Murri, S. (2025, April). Algorithmic Approach for Fraudulent Transaction Detection using Market Basket Analysis with Big Data. In 2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) (pp. 1-7). IEEE.*

iJETRM
International Journal of Engineering Technology Research & Management
(IJETRM)
https://ijetrm.com/

3) M. Kamballi, S. Sanghi, A. Kagalkar, S. C. G. Varma and S. Gupta, "AI and Predictive Analytics in Financial Process Engineering," 2025 International Conference on Sustainability, Innovation & Technology (ICSIT), Nagpur, India, 2025, pp. 1-5, doi: 10.1109/ICSIT65336.2025.11293873.
4) *Fnu, H., Kaushik, K., Gaur, P., Saratchandran, D. V., Thapliyal, A. G., Sidhu, K. S., & Singh, V. (2025, April). Multi-Agent Systems for Collaborative Financial Decision-Making Over Distributed Network Architectures. In 2025 3rd International Conference on Advancement in Computation & Computer Technologies (InCACCT) (pp. 939-944). IEEE*
5) Kannan, S., 2022. The Role Of AI And Machine Learning In Financial Services: A Neural Networkbased Framework For Predictive Analytics And Customercentric Innovations. *Migration Letters*, *19*(6), pp.985-1000
6) Gopal Varma, S. C. (2023). Bridging the Gap Between Traditional Software Engineering and Modern AI Development Practices. *International Journal of Emerging Trends in Computer Science and Information Technology*, *4*(3), 32-40. https://doi.org/10.63282/3050-9246.IJETCSIT-V4I3P104
7) Borra, C. R., Cheekati, S., Pareek, P. K., Rayala, R. V., Kothuri, S. R., & Kumar, S. V. (2025, May). An Intelligent IoT Attack Detection Model Using Weighted ELM and Educational Achievement Guided Optimization. In *2025 13th International Conference on Smart Grid (icSmartGrid)* (pp. 789-796). IEEE.
8) Rayala, R. V., Borra, C. R., Pareek, P. K., & Cheekati, S. (2024, November). Enhancing Cybersecurity in Modern Networks: A Low-Complexity NIDS Framework using Lightweight SRNN Model Tuned with Coot and Lion Swarm Algorithms. In *2024 International Conference on Recent Advances in Science and Engineering Technology (ICRASET)* (pp. 1-8). IEEE. (**23**)