

Adversarially Robust Machine Learning for Financial Risk Prediction under Distribution Shifts

Shawna Gross

Department of Computer Science, University of North Texas, Denton, TX, USA.
hellosawna@unt.edu

Lieke Elsemulder

Department of Computer Science, University of New Hampshire, Durham, NH, USA.
liekeelsemulder239@unh.edu

Nazario Cases

Department of Computer Science, Colorado State University, Fort Collins, CO, USA.
nazariowork@colostate.edu

Abstract

Financial risk prediction is increasingly mediated by machine learning systems that operate in non-stationary markets, adversarial environments, and highly regulated institutional settings. Although predictive accuracy under historical test distributions remains a dominant evaluation criterion, it offers limited assurance when market regimes shift, data-generating processes change, or strategic actors manipulate inputs and signals. This paper examines adversarially robust machine learning for financial risk prediction through a system-level lens, emphasizing architecture, infrastructure, governance, deployment, sustainability, fairness, and policy. It argues that robustness cannot be treated as a narrow model property added after training; rather, it must be designed into socio-technical pipelines that include data acquisition, feature construction, validation, uncertainty calibration, human oversight, auditability, and incident response. The analysis connects adversarial machine learning, domain adaptation, invariant learning, and distributionally robust optimization to the distinctive properties of financial markets, including heavy tails, volatility clustering, regime breaks, reflexive behavior, and regulatory constraints. It further considers how calibration and uncertainty quantification shape risk decisions, how fairness concerns intersect with systemic stability, and how operational sustainability depends on monitoring, documentation, and governance capacity. Case illustrations from credit scoring, fraud detection, market risk, and anti-money laundering show that robustness failures often emerge from institutional misalignment rather than from isolated algorithmic defects. The paper concludes with a research and policy agenda that treats adversarial robustness as an infrastructural obligation, requiring coordinated advances in model design, validation, accountability, and regulatory supervision.

Keywords

adversarial robustness, financial risk prediction, distribution shift, uncertainty calibration, machine learning governance, systemic risk, responsible artificial intelligence.

1. Introduction

Machine learning has become a central component of contemporary financial risk prediction, supporting credit decisions, fraud detection, market risk estimation, liquidity monitoring, anti-money laundering triage, and operational risk assessment. These applications promise

improvements in predictive discrimination, automation, and scalability, yet they also introduce new failure modes that are poorly captured by conventional backtesting and held-out validation. Financial environments are not stationary laboratories. They are characterized by regime changes, policy interventions, feedback loops, strategic adaptation, and rare but consequential tail events. A model that performs well during a period of low volatility may fail abruptly when volatility clusters, correlations shift, or liquidity evaporates. The resulting errors are not merely technical; they can affect access to credit, institutional solvency, market confidence, and regulatory compliance. This paper therefore approaches adversarially robust machine learning for financial risk prediction as a problem of system design rather than model tuning alone.

The adversarial dimension is especially important because financial actors have incentives to manipulate data, conceal risk, and exploit predictive systems. Fraudsters adapt to detection thresholds, borrowers may strategically alter applications, traders may probe automated risk controls, and malicious actors may inject poisoned data into external feeds. Research on adversarial examples has shown that small, often imperceptible perturbations can cause severe misclassification in neural networks [1], and that such vulnerabilities are not confined to a single architecture or task [2]. Defenses such as adversarial training and robust optimization have improved resilience in controlled settings [3], but they rarely address the full operational pipeline through which financial data are collected, validated, transformed, and acted upon. Robustness in finance must therefore contend with both naturally occurring distribution shifts and deliberately induced perturbations.

Distribution shift is the second organizing concern. In machine learning theory, differences between source and target domains can degrade generalization even when training error is low [4]. Financial markets produce exactly these differences through changes in monetary policy, geopolitical shocks, technological adoption, competitive dynamics, and regulatory reform. Methods such as invariant risk minimization and distributionally robust optimization attempt to learn predictors that remain stable across environments or perform well under worst-case distributions [5], [6]. These approaches are valuable, but their assumptions must be translated into institutional practices. A financial institution cannot simply optimize for worst-case performance without considering capital costs, false positive burdens, model interpretability, and regulatory expectations. Robustness thus becomes a governance trade-off, not a purely computational objective.

The system-level argument advanced here is that adversarially robust financial risk prediction requires integrated attention to four layers. The first is data and feature infrastructure, including provenance, quality controls, and resistance to poisoning. The second is model architecture, including ensembling, regularization, and uncertainty estimation. The third is decision governance, including calibration, thresholding, human review, and escalation. The fourth is regulatory and accountability infrastructure, including documentation, auditing, and redress. Failures at any layer can undermine the others. A well-regularized model may still be deployed on corrupted data, a calibrated model may be used with inappropriate thresholds, and an accurate system may lack a credible appeals process. Robustness is therefore an emergent property of the socio-technical system, not a checkbox attached to a model card.

2. Conceptual Foundations and Threat Landscape

Adversarial machine learning emerged from the observation that many high-performing models exhibit surprising fragility under small input perturbations [2]. Subsequent work demonstrated that adversarial examples can transfer across models and that adversarial

training can improve robustness against certain attack classes [1], [3]. These findings are relevant to finance because financial data are often high-dimensional, noisy, and partially adversarial. However, the financial threat model differs from standard image classification benchmarks. An attacker may not seek a one-off misclassification; instead, they may seek persistent mispricing, regulatory arbitrage, reputational damage, or the erosion of trust in a model-driven process. The attack surface includes transaction records, application forms, market feeds, social media signals, third-party data, and even the labels used for supervised learning.

A key conceptual distinction is between robustness to random noise and robustness to strategic manipulation. Random noise can often be addressed through regularization, data cleaning, and robust loss functions. Strategic manipulation requires anticipating adaptive adversaries who respond to defenses. In fraud detection, for example, adversaries learn which patterns trigger alerts and adjust behavior accordingly. In credit scoring, applicants or intermediaries may discover which features most influence decisions and alter applications without changing underlying creditworthiness. In market risk, manipulative trading can distort price signals used by risk models. These dynamics create an arms race in which static robustness guarantees decay over time. A deployed model must therefore be monitored continuously and updated in response to observed adversary behavior.

Domain adaptation and invariant learning provide conceptual resources for addressing distribution shift, but they also reveal limits. Domain adaptation typically assumes access to source and target data or a stable notion of domains [4]. Financial regimes are not always cleanly separable, and the target domain may be partially unknowable in advance. Invariant risk minimization seeks features whose predictive relationship remains stable across environments [5], yet financial environments are often endogenous and reflexive. A feature that appears invariant in historical data may become unstable once many institutions use similar models. Distributionally robust optimization offers a way to optimize against worst-case distributions within a specified uncertainty set [6], but choosing that set is a substantive judgment involving regulatory, economic, and ethical considerations. It cannot be resolved by algorithmic convenience alone.

The threat landscape also includes data poisoning, backdoor attacks, and model extraction. Poisoning is particularly concerning in financial systems because labels may be delayed, contested, or manipulated. A fraudulent transaction may be labeled legitimate for months before investigation, and an adversarial actor may deliberately generate misleading labels to corrupt future training. Backdoor attacks can cause a model to behave normally except under specific triggers, which is dangerous in automated trading or credit approval. Model extraction allows competitors or attackers to approximate a proprietary risk model and then probe its weaknesses. These threats reinforce the need for provenance tracking, anomaly detection, access control, and adversarial red teaming as routine components of financial machine learning infrastructure.

3. Distribution Shifts in Financial Systems

Financial time series exhibit stylized facts that complicate machine learning: heavy tails, volatility clustering, leverage effects, and non-linear dependence [17]. Volatility models such as autoregressive conditional heteroscedasticity and generalized autoregressive conditional heteroscedasticity capture temporal dependence in variance, but they are not designed to handle abrupt regime changes or adversarial manipulation [18], [19]. Predictive comparisons across models also require careful statistical testing because differences in accuracy may not

be stable across periods [20]. These econometric foundations remind us that financial risk prediction is not a generic supervised learning problem. It is embedded in economic processes that generate structural breaks and feedback effects.

Regime shifts can be gradual or sudden. Gradual shifts include demographic change, technological substitution, and evolving consumer behavior. Sudden shifts include financial crises, pandemics, geopolitical conflicts, and central bank interventions. In each case, the joint distribution of features and outcomes changes. A model trained on pre-crisis data may underestimate tail risk, overestimate liquidity, or misclassify creditworthiness. Robustness methods that assume a bounded shift may fail when the shift is structural rather than perturbational. This is why scenario analysis, stress testing, and expert judgment remain essential complements to machine learning. The goal is not to replace human judgment but to design systems that know when their predictions are unreliable.

Financial distribution shifts are also endogenous. If many institutions adopt similar models, their collective actions can alter market dynamics. A risk model that predicts a sell-off may trigger coordinated selling, thereby amplifying the predicted event. This reflexivity means that predictive performance can degrade precisely because the model is widely used. Adversarial robustness must therefore consider systemic effects, not only individual model behavior. A robust model for one institution may be destabilizing if it ignores crowded trades, correlated exposures, or liquidity constraints. Governance frameworks should encourage diversity of models, transparency about common exposures, and circuit breakers that prevent automated feedback loops from accelerating instability.

Cross-domain comparisons are instructive. In medical imaging, distribution shift may arise from new scanners or patient populations. In autonomous driving, it may arise from weather or road conditions. In finance, shift is compounded by strategic actors, regulatory constraints, and immediate economic consequences. A medical model error may harm a patient; a financial model error may contribute to a market crash or deny credit to an entire demographic group. This asymmetry raises the bar for validation and accountability. Robustness techniques from other domains can inform financial practice, but they must be adapted to the pace, scale, and regulatory intensity of financial decision-making.

4. Adversarial Robustness Architectures for Risk Prediction

Architectural choices influence robustness. Deep ensembles provide a simple and effective way to improve predictive uncertainty by training multiple models and aggregating their outputs [10]. Bayesian approximations such as dropout-based uncertainty estimation can also improve calibration and detection of out-of-distribution inputs [11]. These methods are not sufficient on their own, but they create a foundation for identifying cases where the model is uncertain and human review is warranted. In financial risk, uncertainty estimates are valuable because they allow thresholds to adapt to market conditions. During stable periods, the system may automate more decisions; during turbulent periods, it may escalate more cases to analysts.

Regularization and robust optimization are also central. Adversarial training can improve resilience to bounded perturbations [3], while distributionally robust objectives can improve worst-case performance across groups or environments [6]. However, these methods impose costs. They may reduce average accuracy, increase training complexity, or produce overly conservative predictions. In credit scoring, excessive conservatism can exclude creditworthy borrowers. In fraud detection, it can generate false positives that burden customers and investigators. The design challenge is to balance robustness against operational efficiency and

fairness. This requires explicit governance criteria for acceptable trade-offs, not implicit choices buried in model selection.

Infrastructure matters as much as architecture. Data pipelines must be resilient to poisoning, latency, and missing values. Feature stores should maintain versioning and lineage so that models can be reproduced and audited. Monitoring systems should detect drift in inputs, outputs, and performance metrics. Incident response procedures should define how to disable a model, roll back to a prior version, or switch to manual review. These are not glamorous components, but they determine whether robustness survives contact with production systems. Technical debt in machine learning systems often accumulates through hidden dependencies, undeclared consumers, and unstable data sources [12]. Financial institutions cannot afford such debt in systems that affect capital adequacy or consumer access.

Multi-agent and ensemble governance architectures offer a forward-looking direction. Rather than relying on a single predictive model, institutions can combine specialized models for fraud, credit, liquidity, and market risk with a coordination layer that monitors disagreement and uncertainty. Trust-calibrated multi-agent decision intelligence is relevant here because it frames risk management as a problem of coordinating multiple analytical agents under regime shifts [8]. The system-level implication is that robustness can emerge from redundancy, diversity, and structured disagreement. When models disagree, the system can escalate rather than average away the conflict. This approach aligns with regulatory expectations for challenge, validation, and human oversight.

5. Calibration, Uncertainty, and Decision Governance

Calibration is essential for financial risk prediction because decisions depend on probability thresholds. A model that outputs a risk score of high confidence but is poorly calibrated can cause systematic mispricing of risk. Modern neural networks are often miscalibrated, producing confidence estimates that do not match empirical accuracy [7]. Calibrated regression methods can improve uncertainty estimates for continuous outcomes [9], and deep ensembles provide another route to better calibration [10]. Without calibration, robustness claims are fragile because a model may be accurate on average yet dangerously overconfident in tail cases.

Decision governance translates calibrated probabilities into actions. Financial institutions use thresholds for loan approval, fraud alerts, trading limits, and capital reserves. These thresholds reflect regulatory requirements, risk appetite, and business strategy. If a model becomes more robust but less calibrated, threshold decisions may become inconsistent. If calibration drifts after deployment, previously acceptable thresholds may become unsafe. Governance must therefore include periodic recalibration, threshold review, and monitoring of decision outcomes. Human reviewers should receive uncertainty information and explanations that help them adjudicate ambiguous cases. The goal is not full automation but accountable augmentation.

Uncertainty quantification also supports regulatory communication. Supervisors may ask how a model performs under stress, how it detects regime shifts, and how it responds to adversarial inputs. A robust system should be able to report not only point predictions but also confidence intervals, disagreement metrics, and scenario-specific performance. These reports must be understandable to non-specialists and comparable across models. This is a design requirement for model risk management, not an afterthought. It connects technical uncertainty to institutional accountability.

The required trust-calibrated perspective [8] suggests that calibration and trust are not merely statistical properties but governance constructs. A model earns trust when its uncertainty estimates are reliable, its failure modes are documented, and its decisions can be challenged. Trust is lost when institutions deploy opaque systems that cannot explain errors or provide redress. Robustness therefore includes procedural legitimacy. A financially robust institution may still be socially fragile if its models systematically disadvantage protected groups or if affected individuals cannot appeal automated decisions.

6. Fairness, Stability, and Regulatory Infrastructure

Fairness and robustness are intertwined in financial risk prediction. A model may be robust to adversarial perturbations yet biased against a demographic group. Conversely, a fair model may be fragile under distribution shift if fairness constraints are satisfied only in the training environment. Research on fairness in machine learning emphasizes that statistical definitions of fairness often conflict and that context determines which trade-offs are acceptable [15]. In finance, fairness concerns are compounded by legal frameworks such as fair lending, anti-discrimination, and consumer protection. Robustness must therefore include robustness of fairness guarantees across regimes and subgroups.

Large-scale machine learning systems also raise concerns about environmental and institutional sustainability. The resource intensity of training and deploying large models has been critiqued for its carbon footprint and concentration of power [13]. Financial institutions may not train foundation models from scratch, but they increasingly rely on third-party models and cloud infrastructure. This dependence creates supply chain risks, opacity, and lock-in. A robust system must consider vendor reliability, data residency, model updates, and contractual accountability. Sustainability is not only environmental; it is also institutional and operational.

Accountability infrastructure is necessary to make robustness meaningful. Internal algorithmic auditing frameworks can help organizations document design choices, test for harms, and assign responsibility [14]. External supervision can reinforce these practices through examinations, reporting requirements, and enforcement. However, audits must be technically credible and independent. A checklist approach that treats robustness as a compliance artifact will not detect adaptive adversaries or novel regime shifts. Auditors need access to data, code, models, and decision logs, as well as the authority to challenge assumptions. Without such infrastructure, robustness claims remain performative.

Regulatory policy should encourage robustness without mandating a single technical solution. Principles-based supervision can require institutions to demonstrate that they monitor distribution shift, test adversarial scenarios, calibrate uncertainty, and provide redress. At the same time, regulators must avoid creating a false sense of safety through rigid certification. Financial markets evolve quickly, and static certifications can become obsolete. A more effective approach is continuous supervision, incident reporting, and shared threat intelligence. This aligns with the idea that robustness is a dynamic capability rather than a fixed state.

7. Deployment, Sustainability, and Operational Resilience

Deployment is where robustness is either realized or lost. A model that performs well in research may fail in production due to latency constraints, data quality problems, integration errors, or user behavior. Financial institutions often have legacy systems, siloed data, and complex approval processes. Deploying an adversarially robust model requires cross-functional coordination among data scientists, engineers, risk officers, compliance staff, and

business leaders. It also requires clear ownership. If no one is accountable for monitoring drift or responding to attacks, robustness will erode.

Operational resilience depends on redundancy and graceful degradation. If an external data feed is compromised, the system should be able to fall back to alternative signals or manual review. If a model is attacked, the institution should be able to isolate it without shutting down critical services. If a regime shift occurs, the system should escalate uncertainty rather than continue making high-confidence errors. These capabilities require simulation exercises and red teaming. They also require a culture that treats model failure as a foreseeable event rather than an embarrassment. Financial institutions already conduct stress tests and business continuity planning; adversarial machine learning should be integrated into those exercises.

Sustainability also involves maintaining model performance over time. Models decay as data distributions change and adversaries adapt. Continuous learning can help, but it introduces risks such as feedback loops and label leakage. If the model learns from its own decisions, it may reinforce existing biases or become vulnerable to manipulation. Governance must therefore define when and how models are updated, who approves changes, and how updates are documented. The pace of updating should match the pace of risk. In high-frequency trading, updates may be rapid but tightly controlled. In credit scoring, updates may be slower but more transparent.

Cross-domain comparisons highlight the importance of institutional context. In healthcare, robustness failures can harm patients but may be detected through clinical outcomes. In finance, failures can propagate through markets and affect people who never interacted with the institution directly. This systemic dimension means that robustness must be evaluated at the level of the financial network, not only the individual firm. Supervisors should consider correlated model behavior, common data vendors, and shared infrastructure. A single robust model is insufficient if the entire system is homogeneous and fragile.

8. Case Illustrations and Cross-Domain Comparisons

Credit scoring provides a concrete illustration. A lender may use machine learning to predict default risk. During an economic expansion, the model performs well. When unemployment rises, the relationship between features and default changes. If the model was trained on historical data without stress weighting, it may underestimate risk. If applicants learn which features matter, they may strategically alter applications. If the model is opaque, rejected applicants may not know how to improve their creditworthiness. A robust system would combine distributionally robust training, uncertainty calibration, human review, and adverse action explanations. It would also monitor for disparate impact across protected groups [15].

Fraud detection illustrates adversarial adaptation. Fraudsters continuously probe detection systems. A model that relies on static rules or historical fraud patterns may become obsolete quickly. Robustness requires online learning, anomaly detection, graph-based features, and rapid incident response. It also requires careful management of false positives, because blocking legitimate transactions harms customers and erodes trust. Here, the trade-off between robustness and usability is explicit. The system must decide when to block, when to flag, and when to allow with monitoring. These decisions are governance choices, not merely model outputs.

Market risk and algorithmic trading illustrate systemic reflexivity. A model that predicts volatility may trigger trading that increases volatility. If many firms use similar models, crowded trades can amplify shocks. Robustness therefore includes diversity of strategies,

circuit breakers, and supervisory visibility into aggregate exposures. Financial econometrics provides tools for modeling volatility clustering [18], [19], but these tools must be integrated with machine learning systems that can adapt to new data without destabilizing the market. The challenge is to combine statistical rigor with operational restraint.

Anti-money laundering illustrates the tension between detection and privacy. Robust models may require more data sharing, but privacy laws and civil liberties constrain such sharing. Adversarial actors may exploit these constraints by moving funds through jurisdictions with weaker oversight. A system-level approach would combine privacy-preserving technologies, cross-institutional intelligence, and clear legal safeguards. It would also recognize that robustness is not only about catching more illicit activity; it is about minimizing false accusations and protecting due process. These ethical dimensions are central to sustainable financial AI.

9. Policy Implications and Research Agenda

Policy should treat adversarially robust financial machine learning as critical infrastructure. This implies mandatory documentation of training data, model assumptions, validation procedures, and known failure modes. It implies regular adversarial testing, distribution shift monitoring, and calibration audits. It implies clear accountability when automated decisions cause harm. It also implies investment in supervisory capacity, because regulators cannot oversee what they do not understand. The goal is not to stifle innovation but to ensure that innovation does not externalize risk onto consumers, markets, or the public.

Research should advance in several directions. First, there is a need for robustness methods that account for strategic adaptation and endogenous market feedback. Second, there is a need for calibration techniques that remain reliable under regime shifts. Third, there is a need for fairness-aware robustness that does not sacrifice inclusion for stability. Fourth, there is a need for governance frameworks that connect technical metrics to legal and ethical standards. Fifth, there is a need for shared benchmarks that reflect financial realities rather than generic datasets. These benchmarks should include adversarial scenarios, stress periods, and demographic subgroup analyses.

Interdisciplinary collaboration is essential. Computer scientists, economists, statisticians, legal scholars, and policymakers must work together to define what robustness means in financial contexts. A purely technical definition may ignore regulatory constraints; a purely legal definition may ignore computational feasibility. A system-level approach recognizes that robustness is a collective achievement. It requires model developers to anticipate adversaries, institutions to monitor deployment, regulators to enforce accountability, and society to debate acceptable trade-offs. Trust-calibrated multi-agent decision intelligence [8] and related approaches offer promising architecture, but they must be embedded in broader governance.

10. Conclusion

Adversarially robust machine learning for financial risk prediction under distribution shifts cannot be reduced to a single algorithm or defense. It is a socio-technical challenge that spans data infrastructure, model design, uncertainty calibration, decision governance, fairness, and regulation. Financial markets are non-stationary, reflexive, and adversarial. Models that ignore these properties may perform well in backtests while failing catastrophically in practice. Robustness requires continuous monitoring, diverse architectures, calibrated

uncertainty, human oversight, and accountability. It also requires a policy environment that incentivizes resilience rather than superficial compliance.

The paper has argued that robustness should be treated as an infrastructural obligation. Institutions should design systems that know when they are uncertain, detect when distributions shift, resist manipulation, and provide redress when errors occur. They should document assumptions, test adversarial scenarios, and integrate fairness into robustness evaluation. Regulators should require evidence of these practices without prescribing a single technical solution. Researchers should develop methods that reflect the strategic and systemic nature of financial risk. By combining technical rigor with governance discipline, financial institutions can deploy machine learning in ways that are not only accurate but also trustworthy, sustainable, and socially legitimate.

References

1. Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. arXiv. <https://arxiv.org/abs/1412.6572>
2. Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2014). Intriguing properties of neural networks. arXiv. <https://arxiv.org/abs/1312.6199>
3. Madry, A., Makelov, A., Schmidt, L., Tsipras, D., & Vladu, A. (2018). Towards deep learning models resistant to adversarial attacks. In International Conference on Learning Representations.
4. Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Vaughan, J. W. (2010). A theory of learning from different domains. *Machine Learning*, 79(1-2), 151-175.
5. Arjovsky, M., Bottou, L., Gulrajani, I., & Lopez-Paz, D. (2019). Invariant risk minimization. arXiv. <https://arxiv.org/abs/1907.02893>
6. Sagawa, S., Koh, P. W., Hashimoto, T. B., & Liang, P. (2020). Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. In International Conference on Learning Representations.
7. Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. In Proceedings of the 34th International Conference on Machine Learning (pp. 1321-1330).
8. Liu, T. (2026). Trust-Calibrated Multi-Agent Decision Intelligence for Financial Risk Management under Market Regime Shifts. Available at SSRN 7307503.
9. Kuleshov, V., Fenner, N., & Ermon, S. (2018). Accurate uncertainties for deep learning using calibrated regression. In Proceedings of the 35th International Conference on Machine Learning (pp. 2796-2804).
10. Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. In Advances in Neural Information Processing Systems (pp. 6402-6413).
11. Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In Proceedings of the 33rd International Conference on Machine Learning (pp. 1050-1059).
12. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine

- learning systems. In *Advances in Neural Information Processing Systems* (pp. 2503-2511).
13. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623).
 14. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33-44).
 15. Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities*. fairmlbook.org.
 16. Lopez de Prado, M. (2018). *Advances in financial machine learning*. Wiley.
 17. Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223-236.
 18. Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987-1007.
 19. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
 20. Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-263.