MLops in the Financial Industry

Navigating Challenges and Leveraging Opportunities

Jon Doretti
SE 489
DePaul University
Chicago, Il
jdoretti@depaul.edu

ABSTRACT

Machine Learning Operations (MLOps) within the finance industry is influential within the shift of data handling, decision-making, and innovation. This research examines the challenges and opportunities presented by MLOps in financial institutions. Key challenges include: (1) regulatory compliance, (2) model governance and validation, and (3) data security. On the other hand, MLOps offers significant benefits: (1) enhanced decision-making, (2) increased efficiency, and (4) competitive advantage. While using a single method approach, a literature review, this study identifies critical factors for successful MLOps implementation in finance with. The findings underscore the necessity for: (1) robust governance frameworks, and (2) advanced security measures. As so to harness the full potential of MLOps; while ensuring regulatory adherence and ethical standards. The paper concludes with recommendations for future research, aiming to guide financial institutions in navigating the complexities of MLOps deployment.

CCS CONCEPTS

- Computing methodologies → Machine learning → Machine learning approaches → Machine learning operations (MLOps)
- Applied computing → Enterprise computing → Business process management → Business process modeling
- Security and privacy → Information security → Data security → Database and storage security
- Information systems → Information systems applications → Data mining → Data analytics

KEYWORDS

MLOps**,** Finance Industry**,** Regulatory **Compliance,** Model Governance**,** Data Security**,** Bias and Fairness**,** Operational Efficiency**,** Personalized Customer Experiences**,** Innovation

1 Regulatory Compliance

Regulatory compliance is a critical part of the MLops process. This is due to the many regulatory bodies within the U.S. government. The finance industry is heavily regulated, with stringent requirements designed to ensure the stability of financial systems, protect consumer data, and prevent fraudulent activities. Thus, the implementation of MLops in such a heavily regulated environments requires a deep understanding of these regulations through the development of robust frameworks to maintain compliance.

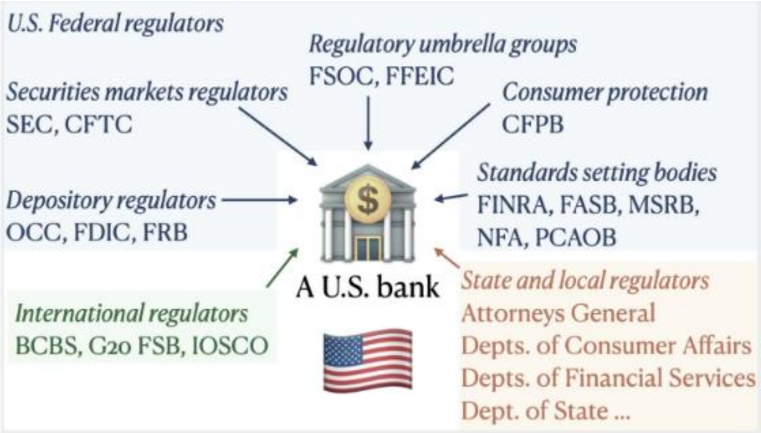
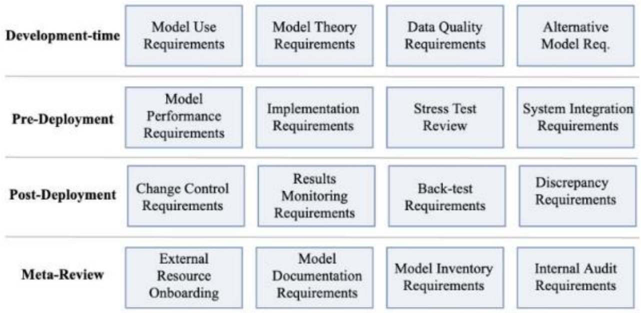


Figure 1: Financial institutions in the US are regulated by a number of regulatory entities at local(yellow), state (yellow), federal (blue) and international levels (green) [3 & 4]

In figure 1, it shows all of the regulatory bodies that would govern a U.S. bank and what area the govern. In the words of Kurshan et al., these robust frameworks were made to address the requirements of the complex governing systems that financial institutions must work within [3].

2 Model Governance and Validation

There is not set guidelines for the model governance process within the financial industry. All companies and financial institutions must set their own process. On the other had there are common themes and similarities between each process. These include: (1) Model risk rating, (2) Initial model validation, and (3) on-going monitoring [3].



Development-time	Model Use Requirements	Model Theory Requirements	Data Quality Requirements	Alternative Model Req.
Pre-Deployment	Model Performance Requirements	Implementation Requirements	Stress Test Review	System Integration Requirements
Post-Deployment	Change Control Requirements	Results Monitoring Requirements	Back-test Requirements	Discrepancy Requirements
Meta-Review	External Resource Onboarding	Model Documentation Requirements	Model Inventory Requirements	Internal Audit Requirements

Figure 2: Regulatory requirements from the Office of the Comptroller of the Currency (OCC) over the model's lifecycle [3]

A model will go thought more rigorous reviews for more risking models. In other words, models that score higher risk assessments with must face stricter requirements than those laid out in figure 2. The initial validation is where the model assessments come from. Here you will also see process validation; this is where model code, data sources, and variable selection are all verified. Lastly with the initial validation you have the results analysis. This is the model's performance, back-testing, benchmarking, and other means of evaluation. The final step that all financial institutions will use is monitoring. Not only is this needed for models in production to protect the financial institution, it is also required by regulators. This includes periodic reviews of model outputs. These reviews are planned based on model risk assessments. It is common for less risky models to only be reviewed once a year. On the other hand, the riskier models are more likely to be reviewed one a quarter or more frequently [3].

2.1 Challenges

Model governance faces many challenges, some include: (1) Model governance is designed for traditional financial models, (2) each review is unique, (3) reviews are time consuming, (4) complexity is making reviews harder. As model reviews were apart of financial institutions before AI models were implemented, current reviews are still using outdated frameworks that are not specifically fitted for AI. Most financial institutions still use model review frameworks that fit with just financial models. In other words, current models were made to look at explicit programs, established relationships among variables, and had clear assumptions. This contradicts AI characteristics of being explicitly programmed and having opacity. Thus, resulting in the effectiveness of reviews being diminished. That being said, each review is unique. Most institutions offer high-level guidance to reviews; making the actual process and requirements for each case to be created on a step-by-step basis. This often makes the review process more time consuming. This is due to complex flow diagrams produced by the AI. These flow diagrams can include several review stages, interdependencies, risk management, and compliance. As AI models become more complex this also adds more steps to the process to the review process [3].

3 Data Security

There are several security practices the protect and secure data. These security practices include architecture schemes such as zero-trust architecture. This core concept relies on the principle of always verifying. This is done through continues authorization and authentication [1]. Another example of data security would be encryption. This would prevent any breachers form accessing the data itself.

Banks often adopt the zero-trust framework s this would reduce the likelihood of a breach. There are three core-principles of zero-trust: (1) verify explicitly, (2) least privilege access, and (3) assume breach. Always authenticate and authorize based on all available data points. This includes user identity, location, device health, and service or workload. This principle emphasizes the need for continuous verification rather than a one-time check at the network perimeter. Next is to limit user and device access to the minimum necessary to perform their job functions. This reduces the attack surface and mitigates the risk of insider threats and lateral movement within the network. Lastly, Banks must operate with the assumption that the network is already compromised. This mindset drives proactive monitoring, threat detection, and response strategies to identify and mitigate threats quickly.

3.1 Challenges

The financial industry faces lots of challenges when it comes to data security. Many financial institutions are still operating on legacy systems. As we collect more and more data, financial institutions take on subsequently more risk.

WARNING
Data theft 80%
Legacy systems struggle to keep up
The bigger the data, the higher the risk

Figure 3: Big data challenges [1]

These challenges can be broken down into four groups: (1) data volume, (2) cyber threats, (3) regulatory compliance, and (4) data underutilization. These financial institutions collect mass amounts of data to perform a multitude of tasks and services. According to Hasan et al., only .5% of all data collected by these financial institutions is utilized. This then results in mass amounts of data taking up storage space and utilization important recourses while not being used. This then allows for cyber criminals to access data that may not be as actively monitored due to the assumption that those data points are not useful. This then create a whole slew of security issues for financial institutions. Data breaches are the most harmful to a financial institution as if they are unable to protect their data, they will lose public trust in protection of other financial instruments or assets under their protection [1].

4 Capabilities

In the financial industry, where data-driven decision-making is critical, MLOps enhances operational efficiency, regulatory compliance, risk management, and customer experiences. This enhances decision-making with: (1) predictive analysis, (2) fraud detection, and (3) risk management. In the financial industry, where data-driven decision-making is critical, MLOps offers powerful capabilities to enhance operational efficiency, regulatory compliance, risk management, and customer experiences. MLOps enables financial institutions to deploy predictive models that can forecast market trends, credit risks, and customer behaviors based on the data they collect and how they train their models [2]. These models provide valuable insights that help in making informed strategic decisions, optimizing investment portfolios, and mitigating risks. Models can analyze a customer's transaction patterns in real-time to identify potentially fraudulent activities. MLOps ensures these models are continuously updated and monitored, improving the accuracy and responsiveness of fraud detection systems. Lastly, models can be trained a deployed to assess real-time insights and perform proactive risk mitigation strategies, enhancing the overall risk management framework [2].

To enhance efficiency in a financial institution, MLOps allows for continuous integration and continuous development (CD/CI) principles within the ML process. Thus, the implementation of CD/CI pipelines ensures that updates and improvements are seamlessly integrated into the production environments. In other words, this allows for the rapid integration and improvement of models ensuring they are effective and up-to-date and regulation. This in turn will improve the scalability of a project as MLOps leverages cloud-based solutions and containerization technologies to scale ML models efficiently. Scalability is crucial for handling large volumes of financial data and supporting high-performance analytics in real-time [2].

Lastly, to assist with regulatory compliance, MLOps supports: (1) model governance, (2) transparency and accountability, and (3) data security and privacy. MLOps allows for the ability to define and implement numerous frameworks for model development, validation, deployment, monitoring, and retirement. Thus, ensuring that all stages of the ML lifecycle adhere to industry regulations. This then, in turn, enhances by implementing explainable AI techniques that provide insights into model decision-making processes. This transparency is crucial for regulatory compliance and builds trust with stakeholders and regulators. Lastly, with such a robust MLOps framework, data security measures such as encryption, access controls, and anonymization techniques are easily implemented. These measures ensure that sensitive financial data is protected, aligning with data privacy regulations like GDPR and CCPA.

5 Conclusion

MLOps offers a comprehensive suite of capabilities that can significantly enhance the operational efficiency, regulatory compliance, risk management, and customer experiences of financial institutions. Adopting robust model governance frameworks, ensuring data security and privacy, and embracing scalable infrastructure solutions, financial institutions can enhance decision-making, improve operational efficiency, and deliver personalized customer experiences. By using CD/CI princilpes and streamlining the deployment and management of ML models, MLOps enables financial institutions to use machine learning; driving innovation and maintaining a competitive edge in a rapidly evolving industry. Future advancements in MLOps technology and practices will further expand these capabilities, providing even greater value to the financial sector. Future research should focus on developing standardized best practices for MLOps implementation in finance and exploring advanced techniques for bias mitigation and model interpretability.

REFERENCES

[1] Hasan, Mahmudul, et al. "ADVANCING DATA SECURITY in GLOBAL BANKING: INNOVATIVE BIG DATA MANAGEMENT TECHNIQUES." International Journal of Management Information Systems and Data Science, vol. 1, no. 2, 1 May 2024, pp. 26–37, www.globalmainstreamjournal.com/index.php/IJMISDS/article/view/133, https://doi.org/10.62304/ijmisds.v1i2.133. Accessed 17 May 2024.

[2] Jan, Chyan-Long. "Detection of Financial Statement Fraud Using Deep Learning for Sustainable Development of Capital Markets under Information Asymmetry." Sustainability, vol. 13, no. 17, 2 Sept. 2021, p. 9879, https://doi.org/10.3390/su13179879.

[3] Kurshan, Eren, et al. "Towards Self-Regulating AI." Proceedings of the First ACM International Conference on AI in Finance, 15 Oct. 2020, https://doi.org/10.1145/3383455.3422564.

[4] M. Labonte. 2020. Who Regulates Whom? An Overview of the U.S. Financial regulatory Framework. Congressional Research Service R44918 (2020).

[5] Moritz, Philipp, et al. "Ray: A Distributed Framework for Emerging {AI} Applications." Www.usenix.org, 2018,www.usenix.org/conference/osdi18/presentation/nishihara. Accessed 6 June 2024.