Undergraduate Thesis Prospectus

Multivariate Anomaly Detection: Evaluating Isolation Forest

(technical research project in Computer Science)

Regulating Financial Institutions: Different for FinTechs

(sociotechnical research project)

by

Alan Phlips

October 27, 2022

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Alan Phlips

Technical advisor:	Rosanne Vrugtman, Department of Computer Science
STS advisor:	Peter Norton, Department of Engineering and Society

General research problem

How can technological innovations improve the Financial Services industry?

In recent decades, digital technology has advanced rapidly across industry, including financial services. This can be seen in mobile banking, peer-to-peer payment services, and online trading platforms (Columbia, 2022). However, finance's far reaching impacts has in certain cases prevented technology from being available and in others helped it to grow to global importance without regulation. Regulations can prevent traditional banks from providing services to customers, while the lack thereof on Financial Technology (FinTech) companies can help them do the same. This discrepancy concerns consumers, regulators, and firms alike. Technology now also exists to increase capabilities within firms. Applications of machine learning can reduce costs, enable new business opportunities, or improve existing services (CFI, 2022).

Multivariate Anomaly Detection: Evaluating Isolation Forest

How can multidimensional abnormalities in Credit Bureau data be more effectively detected to reduce errors in card approval and line assignments?

This will be a solo capstone project in the Computer Science department under Rosanne Vrugtman for CS 4991 in Spring of 2023. Data provided to banks from Credit Bureaus includes fields whose values depend on that of other fields. Starting in 2020, contradictory values in several such attributes appeared in the data. Individual attribute values were acceptable, but in context were anomalous. Anomalies like these make it harder to evaluate an individual's qualification for a new line of credit or a line increase. This particular instance was costly in terms of company resources and potential loss of business.

There are several machine learning algorithms that may detect such abnormalities, including Isolation Forest (Lewinson, 2018). Isolation Forest has been implemented at scale to detect abnormalities including preventing abuse in social networks (Verbus, 2019). This project seeks to evaluate Isolation Forest's ability to detect anomalies in samples of monthly bureau data. The project is constrained in access to data needed for the algorithm and in opportunity to test against other algorithms, limiting ability to compare alternatives for the same task.

Currently, anomaly detection is done with a proprietary technique that considers only values of each attribute alone. This technique's lack of multivariate consideration necessitates a new solution. This project will be a proof of concept to address this issue by experimenting with data within a credit cycle. The cycle data will train an Isolation Forest and the resulting model will be evaluated on what records it considers most anomalous. If these records are considered appropriately identified, subsequent experiments include identifying more specific anomalies, using data from different cycles, or different datasets altogether. If successful, an Isolation Forest based model could be implemented to preemptively identify all types of anomalies in credit bureau data, thus reducing errors and saving company resources. If unsuccessful, other algorithms could be explored to address the same issue.

Regulating Financial Institutions: Different for FinTechs

In the U.S., how do FinTech firms, banks, other financial service companies, and their clients compete to influence the regulatory regime to which FinTech companies are subject?

FinTech firms provide financial opportunities previously unavailable to most Americans. However, due to short lifespans and innovative services, these firms are not regulated like older

institutions. This raises concerns on how they will fare in a financial crisis and how their consumers will fare long term.

Participants include the American Bankers Association and the Financial Technology Association, two U.S. trade associations representing traditional financial institutions and FinTech companies, respectively. The American Bankers Association argues exemptions from regulations give these new companies an unfair competitive advantage and introduce cybersecurity risks (Williams, 2022). They support "responsible innovation" in financial services conditional on consumers getting the "protections they deserve" (Williams, 2022). This reflects less a desire for consumer safety and more their own desire to retain profits through a more consistent regulatory environment. The Financial Technology Association claims to be for security, responsible use, and financial inclusion but only wants regulation that enables their services. They describe their member's products as creating a "more equitable and inclusive financial system" to appear as proponents of these ideals while portraying traditional banks and regulators as opponents (FTA, 2022).

Participants also include consumer advocates, who seek increased disclosure of data use, costs, and more accountability (Saunders, 2019). These advocates note how FinTechs "have the potential to provide important benefits to consumers," but stress that the innovative approaches are "not invariably positive" (Saunders, 2019). They desire for both convenience and safety for consumers in handling finances. Finally, participants include federal regulators, who acknowledge FinTechs have increased consumer access to financial services, but have not hugely changed their regulation. Powell (2017) highlights the "balance that needs to be achieved in this innovative environment" to remain neutral between banks and FinTechs. He also mentions that

the convenience and speed of services must not “undermine the safety, security, and reliability” of those same services as regulators are meant to enforce such standards (Powell, 2017).

Researchers have investigated regulating banking in an environment of innovation. For example, Busch (2001) examined managing the banking in six European nations during a period of innovation around 20 years ago. The World Bank (2021) recently researched regulatory approaches to address global consumer risks with FinTech. The World Bank also previously examined old banking regulations for lessons to improve regulating Pension Funds (Rocha, 1999).

References

Busch, A. (2001). Managing Innovation: Regulating the Banking Sector in a Rapidly Changing Environment. *Success and Failure in Public Governance: A Comparative Analysis*. 311-25.

CFI (2022, January 15). Corporate Finance Institute. Machine Learning (in Finance). www.corporatefinanceinstitute.com/resources/knowledge/other/machine-learning-in-finance/

Columbia (2022). What is Financial Technology (FinTech)? A Beginner's Guide for 2022. *Columbia Engineering Boot Camps*. www.bootcamp.cvn.columbia.edu/blog/what-is-fintech/

FTA (2022). Financial Technology Association. FinTech Policy Priorities. www.ftassociation.org/sound-policy/

Lewinson, E. (2018, June 2). Outlier Detection with Isolation Forest. *Towards Data Science*. www.towardsdatascience.com/outlier-detection-with-isolation-forest-3d190448d45e

Powell, J. H. (Presenter). (2017, October 18). Financial Innovation: A World in Transition. Speech presented at 41st Annual Central Banking Seminar, Federal Reserve Bank of New York, New York, NY, United States.

Rocha, R., Hinz, R., & Gutierrez, J. (1999, December). Improving Regulations and Supervision of Pension Funds: Are there Lessons from the Banking Sector? (English). World Bank Group.

Saunders, L. (2019, March). Fintech and Consumer Protection: A Snapshot. www.nclc.org/images/pdf/cons-protection/rpt-fintech-and-consumer-protection-a-snapshot-march2019.pdf

Verbus, J. (2019, August 13). Detecting and preventing abuse on LinkedIn using isolation forests. *LinkedIn Engineering*. www.engineering.linkedin.com/blog/2019/isolation-forest

Williams, W. (2022). Most Consumers Stick to Traditional Banking Over Digital Alternatives. *ABA Banking Journal*. www.bankingjournal.aba.com/2022/09/most-consumers-stick-to-traditional-banking-over-digital-alternatives/

World Bank (2021). Consumer Risks in Fintech: New Manifestations of Consumer Risks and Emerging Regulatory Approaches. World Bank. https://openknowledge.worldbank.org/handle/10986/35699