*Explores PayPal and H2O.ai’s use of automated machine learning to reduce financial fraud*

**Assignment**

**5**

A5

ALY6060 Decision Support & Business Intelligence

Assignment 5 – Automated Machine Learning

**PREPERATION:**

By: John DiSessa

For: Professor Ellis

On: June 26th, 2022

Introduction

Financial fraud is rampant in any industry. Whether it be via traditional banking, cryptocurrency, or financial technology, hackers continue to evolve and steal billions of dollars from consumers every year. As a fast-growing company, PayPal handles an increasing volume of consumer capital which increases the incentives for hackers to fraud PayPal and their customers. PayPal recognized that their fraud prevention and detection was not good enough and consumer trust was hurting. They worked with H2O.ai, who used automated machine learning, to improve their fraud prevention and detection by 6%.

Analysis

**Business Need for Artificial Intelligence**

Fraud has been and continues to be the Achilles heel for the financial services industry. In 2021, $5.8 billion was lost to hackers, scams, and fraud in the United States. Some of this money results from human error, but some of it also results in failure by financial service companies to protect their consumers. They have had their servers hacked or allowed users to interact with fraudulent accounts. In these instances, these companies owe their customers their money, have to pay insurance claims, and lose millions of dollars every year in reparations. Traditional banks easily recover from these reparations since they are well-funded and have a proven history of making it right with consumers. Of course, it still is in their best interest to avoid fraudulent transactions, but they have deep roots in the industry where consumer trust and merchant adoption can be unaffected by fraud. PayPal on the other hand, is still trying to grow their user base and global adoption of their service. Many merchants, both with physical stores and online-only shops, don’t allow customers to pay with PayPal but still allow Visa and Mastercard transactions who partner with all of the traditional financial instructions with credit and debit cards. Despite having superior technology, merchants and customers are hesitant to use PayPal for fear of fraud. PayPal suffers from network effects compared to their rivals. Visa and Mastercard increase adoption because each new user knows that their payments will be accepted everywhere, which only continues to grow their customer base. Many people ignore PayPal simply because they are concerned that merchants won’t accept their payments. Reducing fraud will allow consumers to grow their trust in PayPal, which will increase their user base, which will then attract more consumers.

Even though PayPal has seen incredible growth, they recognized fraud as their biggest hurdle back in 2014. PayPal was created in 1998, went public in 2002, and by 2010, earned $3.5 billion in revenue and had over 87 million active accounts. After the infamous iTunes/PayPal hack in 2010, PayPal invested heavily into machine learning to understand how fraudulent accounts were created and how to stop them. Their previous investments in software engineers and data analysts were not sufficient. Hackers continued to evolve and were able to keep pace with PayPal’s improvements. In 2014, PayPal turned to H2O.ai to use automated machine learning to process large amounts of data and scale their algorithms to adapt to real-time changes. PayPal needed to find ways to increase fraud detection accuracy and decrease detection time while innovating fast enough to outpace hackers.

**Artificial Intelligence Results**

PayPal first approached this problem by conducting machine learning on their own, but were unsatisfied with the accuracy of their models. Their approach, at least, was actually helpful for the final solution. Rather than analyzing each account one at a time, assessing their transaction history, and determining the likelihood the account was fraudulent or that at least one of their transactions was fraudulent, they looked for patterns. Using confirmed cases of fraud, they were able to map common activity and assess their entire network for similar activity. “Once a bad account is identified based on payment transaction data, other bad accounts sharing the same network structure can be located.” If a fraud account committed 30 transactions over a 5-minute span, they looked for all accounts who also committed 30 or more transactions in a short time-frame. While this did work, it was not consistently accurate enough for them to be confident in. Their biggest issue was feature identification. Every transaction in their database had almost 600 recorded features. PayPal could not figure out the optimal combination of features or create algorithms that updated these optimal features as more data was tracked. H2O.ai used Driverless AI (automated machine learning) to optimize feature identification. “One of the company’s immediate goals is to evaluate Driverless AI directly with raw data by plugging it into the data stream using timeseries functionality to eliminate manual feature engineering on new data.” H2O.ai used 1 years’ worth of PayPal transactions to train their models and 3 months of data to test their models. They also used an IBM Power 8 GPU server which simply provided a necessary increase over PayPal’s raw computing power. H2O.ai trained their models 6 times faster than what PayPal had tried on their own. Using automated machine learning and advanced computer power, they gave PayPal the results they were looking for. H2O.ai was able to identify the top 5 features in identifying fraud and improve upon PayPal’s model accuracy. “The top 5 features extracted by H2O Driverless AI bested 10 years’ worth of expert engineered features. At the same time, H2O Driverless AI increased model accuracy increased from .89 to .947.”

**Ethical Considerations**

With any implementation of artificial intelligence, I believe a strong ethical evaluation is required. Since artificial intelligence can affect human behavior, companies must make sure they obtain data ethically, experiment ethically, and use their findings to benefit society. PayPal and H2O.ai have improved society in an ethical manner. The data H2O.ai used to train and evaluate their models came solely from PayPal’s transaction and account history. The transaction data contained only necessary information about the transaction amounts, involved parties, and timestamps. When they looked at user account information, they only used information that users inputted upon their account creation. PayPal and H2O.ai did not buy data from outside sources in any way. Their experiments only used the data to create and evaluate their models. There was no human intervention so we can conclude they built their models ethically. The most important ethical dilemma for implementing artificial intelligence is how it will affect society. PayPal’s goal had always been to reduce fraud in order to save money and increase consumer trust. PayPal did not take these results and use them in any other way besides improving their own service. Fraud reduction is universally accepted as a benefit to PayPal users.

Summary

H2O.ai was able to use automated machine learning to improve PayPal’s ability to detect fraud. The 6% increase in model accuracy may not seem like a lot, but because PayPal’s transaction volume is so significant ($936 billion in 2020), 6% could results in millions of dollars in savings and an increase in consumer trust. PayPal has been able to use their algorithms to reduce fraud, improve consumer trust, and grow their business. As one of the largest financial technology firms, they are obligated to operate in the best interest of their consumers. PayPal and H2O.ai ethically used automated machine learning to improve the customer experience.

Citations

“0Xdata Is Now H2O.Ai, Announces Production Customers Including PayPal, Nielsen, and Cisco.” *0Xdata Is Now H2O.Ai, Announces Production Customers Including PayPal, Nielsen, and Cisco | Business Wire*, 18 Nov. 2014, https://www.businesswire.com/news/home/20141118005268/en/0Xdata-is-Now-H2O.ai-Announces-Production-Customers-Including-PayPal-Nielsen-and-Cisco.

0Xdata, director. *YouTube*, YouTube, 11 Dec. 2017, https://www.youtube.com/watch?v=r9S3xchrzlY. Accessed 25 June 2022.

0Xdata, director. *YouTube*, YouTube, 2 Dec. 2014, https://www.youtube.com/watch?v=RqkheMI3Ciw. Accessed 25 June 2022.

*2018.08.15 H20 Customer Case Study Paypal - h2o.Ai*. https://h2o.ai/content/dam/h2o/en/marketing/documents/2018/06/H20-Customer-Case-Study-PayPal-5302018-FINAL-NEW-2-1.pdf.

*AutoML*, https://www.automl.org/automl/.

Best, Raynor de. “PayPal Accounts 2022.” *Statista*, 9 May 2022, https://www.statista.com/statistics/218493/paypals-total-active-registered-accounts-from-2010/.

“Fraud Is Becoming the Biggest Headache in Financial Services.” *Tearsheet*, https://tearsheet.co/security-2/fraud-is-becoming-the-biggest-headache-in-financial-services/.

Iacurci, Greg. “Consumers Lost $5.8 Billion to Fraud Last Year - up 70% over 2020.” *CNBC*, CNBC, 22 Feb. 2022, https://www.cnbc.com/2022/02/22/consumers-lost-5point8-billion-to-fraud-last-year-up-70percent-over-2020.html.

“ITunes Hasn't Been Hacked, Though Your PayPal Might Have Been Phished (Updated).” *The Guardian*, Guardian News and Media, 24 Aug. 2010, https://www.theguardian.com/technology/blog/2010/aug/24/itunes-paypal-scams-phishing.

Kauflin, Jeff. “PayPal Admits 4.5 Million Accounts Were Illegitimate as Fintech's Fraud Problem Grows.” *Forbes*, Forbes Magazine, 9 Feb. 2022, https://www.forbes.com/sites/jeffkauflin/2022/02/02/paypal-admits-45-million-accounts-were-illegitimate-as-fintechs-fraud-problem-grows/?sh=6418b09836b9.

“PayPal Revenue and Usage Statistics (2022).” *Business of Apps*, 4 May 2022, https://www.businessofapps.com/data/paypal-statistics/.

“Top 22 Automl Case Studies / Examples: In-Depth Guide [2022].” *AIMultiple*, https://research.aimultiple.com/automl-case-studies/.