An approach to benchmark fraud detection algorithms in the COVID-19 era

Revista Latinoamericana de Economía y Sociedad Digital

Issue 2, agosto 2021

Autores: Miroslawa Alunowska Figueroa, Daniel Turner-Szymkiewicz, Juan Sebastián Cárdenas-Rodriguez, Ulf Norinder, Edgar Alonso Lopez-Rojas

DOI: 10.53857/RPGD2470

Publicado: 25 agosto, 2021

Recibido: 21 marzo, 2021

Cita sugerida: Alunowska Figueroa, Miroslawa; Turner-Szymkiewicz, Daniel; Alonso Lopez-Rojas, Edgar; Cárdenas-Rodriguez,Juan Sebastián & Norinder, Ulf (2021) "An approach to benchmark fraud detection algorithms in the COVID-19 era" en Revista Latinoamericana de Economía y Sociedad Digital, Issue 2

Licencia: Creative Commons Atribución-NoComercial 4.0 Internacional (CC BY-NC 4.0)

Tipo: Ensayo

Palabras clave: benchmark simulation, COVID-19, Digital payments fraud, machine learning, synthetic data for finance

---

Resumen

Para afrontar los desafíos en la lucha contra los delitos financieros, especialmente en el contexto de la pandemia del COVID-19, este artículo focaliza en los datos sintéticos financieros y en el uso de una herramienta de benchmark confiable para evaluar los algoritmos de detección de fraude. Los departamentos de control de cumplimiento de las instituciones financieras enfrentan el desafío de reducir el número de personas inocentes a las que se las acusa de fraude por error. Para enfrentar este problema, las instituciones financieras están investigando la aplicación de algoritmos de aprendizaje automático para la detección de fraude y la tecnología de análisis de datos para desarrollar un sistema de detección del fraude más exacto y preciso. Sin embargo, los enfoques para la optimización y la automatización del control bancario y los procesos de testeo son desafiantes ya que no existe consenso en un benchmark. Investigamos la importancia de medir la aplicación de un benchmark para detectar los delitos financieros ante un sector financiero digital en

crecimiento, como es el caso de México. Este estudio cobra especial importancia debido a las serias amenazas que enfrenta un sistema financiero que se está desarrollando rápidamente. (Informe del Banco Central de México 2019) Estos riesgos se han empeorado aún más como resultado de los cambios acelerados hacia los pagos digitales como producto de la pandemia COVID-19.

Abstract

To address the challenges in the fight against financial crime, particularly in the COVID-19 pandemic context, this paper focuses on financial synthetic data and the use of a reliable benchmark tool to test fraud detection algorithms. Compliance departments at financial institutions face the challenge of reducing the number of innocent people erroneously accused of fraud. To cope with this problem financial institutions are exploring the application of machine learning fraud detection algorithms and data analysis technologies to develop a more accurate and precise fraud detection system. However, approaches to streamlining and automating banks' monitoring and testing processes is challenging as there is no consensus on a benchmark. We explore the relevance of measuring the applicability of a financial crime benchmark in the presence of a growing digital financial sector, such as in the case of Mexico. This study is particularly important due to serious threats that are faced by a rapidly developing financial system (2019 Mexican Central Bank Report). These risks have been further exacerbated as a result of the COVID-19 pandemic accelerating the shift towards digital payments.

Resumo

Para enfrentar os desafios na luta contra o crime financeiro, particularmente no contexto da pandemia da COVID-19, este artigo se concentra em dados sintéticos financeiros e no uso de uma ferramenta de referência confiável para testar algoritmos de detecção de fraude. Os departamentos de compliance de instituições financeiras enfrentam o desafio de reduzir o número de pessoas inocentes erroneamente acusadas de fraude. Para lidar com esse problema, as instituições financeiras estão explorando a aplicação de algoritmos de detecção de fraude que envolvem aprendizado de máquina e tecnologias de análise de dados para desenvolver um sistema de detecção de fraude mais preciso. No entanto, as abordagens para agilizar e automatizar os processos de monitoramento e teste dos bancos são desafiadoras, pois não há consenso sobre a referência a ser utilizada. Nós exploramos a relevância de medir a aplicabilidade de uma referência de crime financeiro diante de um crescente setor financeiro digital, como no caso do México. Este estudo é particularmente importante devido às sérias ameaças que um sistema financeiro em rápido desenvolvimento enfrenta (Relatório do Banco Central do México de 2019). Esses riscos foram ainda mais exacerbados como resultado da pandemia da COVID-19, que acelerou a mudança para

pagamentos digitais.

1. Introduction

The COVID-19 pandemic has caused a societal shift in persuading more people to adopt digital payment platforms as their primary form of payment (Pandey et al., 2020). Nevertheless, with such a shift in consumer activity, we also see an evolution of fraud occurring that takes advantage of the vulnerabilities in digital payment systems that appear in the wake of such substantial changes (Karpoff, 2020). Financial fraud requires extensive knowledge of protocols and systems to access accounts and transfer services, thus banks have been strained to increase investment in security and fraud protection (van Driel, 2019). Furthermore, these problems require new ways for control systems to rapidly adapt to contemporary circumstances. A novel solution to this issue has been presented as a benchmark that can appropriately measure the performance of a transaction monitoring system (TMS). Nonetheless, this challenge presents one of the largest hurdles financial institutions face today.

Financial institutions tune their control systems according to applicable regulations; these can be broken down into two clear objectives: detect and prevent criminal activity (increasing true positives) and reduce the number of innocent people wrongfully accused of fraud (reducing false positives). The endeavors of financial institutions to achieve these goals are frequently hindered, predominantly because of their inability to adequately assess the actual amount of false negatives (undetected fraud) present in their datasets, rendering conventional metrics of detection with little value of assessing the greater scope of financial crime.

Machine learning (ML) fraud detection algorithms, including recent advances in Deep Learning (DL), have been able to play a more effective role in finding the hidden correlations between user behavior and fraud actions. However, many of these algorithms often evolve with minimum input from an analyst and rely considerably on the data that is provided. This in itself poses an additional layer of depth in the adoption/establishment of technical standards to create best practices and is still an ongoing topic of debate in the industry. This is especially true within the data bias assessment or fairness assessment, as ML models trained on a biased dataset will cause bias/discrimination unintentionally (Das et al., 2020).

The technological race between fraud and security has a large set of ramifications that is often not recognized until much later. To rapidly counter fraud a less-than-robust model is often introduced, which may be prone to bias through hastily made miss-classifications on the incoming data. Moreover, to try to circumvent this approach and organize historical data; the financial institution puts itself at great risk of a privacy leak. We believe that the use of adequately bench-marked synthetic data can help mitigate a great part of this

problem, and thus reduce concerns for bias and privacy in the realm of compliance.

Much of the data that can be used to regulate the training of a proper model is restricted under privacy regulations (Zhang et al., 2020), significantly reducing the possibility of collaboration between different stakeholders to improve fraud control tools that can prevent financial crime, bias, and privacy leaks. A crucial part of the solution to these problems will require a combined response of enriched synthetic dataset generation and proper metrics that can adequately benchmark the performance and effectiveness of machine learning algorithms that operate as part of transaction monitoring systems.

2. Background and context

Mexico is the second-largest economy in Latin America and ranks among the world's 15 largest economies (World Bank, 2020). It has a robust financial sector and a well-capitalized banking sector composed of transnational banks, national banks, and start-ups in fintech (OECD, 2019). Although it is the springboard for internationalization for the region, nearly half of its adult population does not utilize the banking ecosystem. In Mexico, 'cash is king', even with the inclusion of digital solutions to cope with the pandemic only 37% of adults hold a bank account (CNBV, 2020). This is particularly prominent in the south of Mexico (OECD, 2019), where some areas have created their own alternative currency: *Tumin* (Saldivar and Zavaleta, 2020). There are several factors that have led to this situation such as high levels of the informal economy, low income, mistrust in the banking system, high commissions, high-interest rates, abhorrence to taxation, and few products directed to lower-income segments. All of these aforementioned issues serve as barriers to financial inclusion.

Individuals who are part of the informal economy are concerned that part of their earnings will go to taxes, which in their view means having fewer earnings. It is not uncommon for individuals to distrust banks and refer to them as entities that will "steal" their money. This sentiment is reinforced even further by how profitable banks in Mexico are (Estañol, 2019; Juarez, 2015). On average 30% of the Mexican commercial banks' income is attributed to commissions (Condusef, 2017) and interest rates are circa 30% for cards, whereas small enterprises face high borrowing costs (OECD, 2019).

Financial inclusion allows people and businesses to access useful and affordable financial products (World Bank, 2018), which has been a top priority in Mexico for the past 20 years. To achieve the desired impact, regulatory frameworks and policies must be set in place (Naghavi, 2020). One of the more recent government programs that can be considered an advance in Latina America for financial inclusion is the National Financial Inclusion Policy 2020-2024 (OCDE/CAF, 2020). The program aims to increase the number of Mexicans with a bank account from 47% in 2020 to 65% in the next four years (SHCP/Ministry of Public Education, 2020). This is a collaboration between the Bank of Mexico (Banxico), the Ministry of Finance and Public Credit (SHCP), and the Ministry of Public Education built on

top of several actions taken during the past 20 years to foster financial inclusion. Among some of those relevant actions taken are:

– Interbank Electronic Payment System or SPEI® (in 2005), allowing shops (e.g. a popular chain of convenience stores is Oxxo) to be agents for financial institutions and act as intermediaries for paying utilities, withdrawals, or checking bank account balance (in 2008)

– Creation of the Department for Financial Inclusion (in 2009) and its surveys (CNBV/INEGI 2012, 2015, 208)

– 2016 National Financial Inclusion Policy

– Fintech Law (in 2018)

– Digital payment platform or Digital Purchase "CoDi" which is based on QR codes and near-field communication technology (in 2019).

Similar to other countries, digital banking has been considered a key component to achieve financial inclusion (Morales Guzman, 2018), (Alonso Fernández de Lis, Hoyo, López-Moctezuma and Tuesta, 2013), (Maina, 2019). Alipay is referred to by the Mexican Government as a success story of such an achievement (CNBV, 2020).

A number of digital payment solutions are both public and private: Samsung Pay, Apple Pay, CoDi, Rappi Pay, PayPal, Mercado Libre. Figure 1 depicts an overview of the Mexican Mobile Systems Ecosystem; between the Merchants and the Customers, there are several actors, Financial Institutions (banks), regulatory authorities (such as Banxico and SHCP), Internet technology (e.g. fiber optic or 4G technology) and its infrastructure, Mobile service providers (e.g. Telcel), Mobile payment models (e.g. Paypal, Apple Pay), Device (e.g. Samsung) and chip manufacturers, Card networks (e.g. Visa, Mastercard, American Express), Payment processors, PsP/Gateways, Point of Sales (POS) Manufacturers and payment networks.

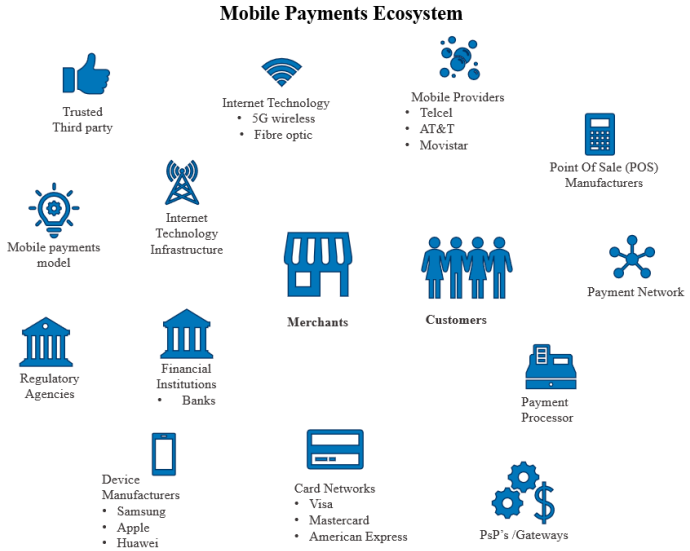


Figure 1. *Overview of the Mexican Mobile Systems Ecosystem. Own visualisation. Between the Merchants and the Customers there are several actors: the Financial Institutions (banks), the regulatory authorities (such as Banxico and SHCP), Internet technology (e.g. fibre optic or 4G technology) and its infrastructure, Mobile service providers (e.g. Telcel), Mobile payment models (e.g. Paypal, Apple Pay), Device (e.g. Samsung) and chip manufacturers, Card networks (e.g. Visa, Mastercard, American Express), Payment processors, PsP/Gateways, Point of Sales (POS) Manufacturers and payment networks.*

The Mexican Financial System is structured with two main actors at play; the financial authorities or regulatory bodies; and the financial institutions. Figure 2 depicts an overview of the Financial System. The two most important regulatory bodies are the Bank of Mexico (Banxico), which has been autonomous since 1994, and the Ministry of Finance and Public Credit (SHCP). It is important to highlight that Mexico is an observer at the Budapest Convention on cybercrime and its regulation is aligned to international regulatory bodies such as The Basel Committee. Furthermore, the banking sector is regulated by the National Banking and Securities Commission (CNBV), which is part of the SHCP, and complies with Banxico's regulation of payment systems and transfers.

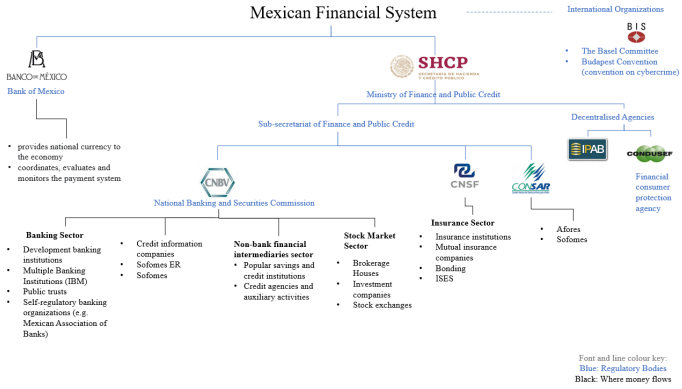


Figure 2. *Overview of the Mexican Financial System. Font and line color key: Blue: Regulatory Bodies, Black: Where money flows.*

The regulatory framework is continuously being updated, one of the most significant reforms occurred in 2014 to boost access to financial products and financial inclusion (Avendaño Carbellido, 2016). Hence, the banking law has adapted to advances in technology as well (Avendaño Carbellido, 2017). Particularly relevant are two legal instruments: The Credit Institutions Law (DOF, 1990) and the Financial Technology Law (DOF, 2018). Article 142 of the Credit Institutions Law legislates banking secrecy, whereas article 115 requests the establishment of measures and procedures to prevent or detect crime as defined by the articles 139, 148 or 400 of the Federal Penal Code.

The fintech law, considered a milestone of legislation in the Latin American financial sector, introduces a regulatory regime for open banking and covers three types of fintech companies: electronic payment institutions, asset management institutions, and crowdfunding institutions. Article 58 of the Financial Technology Law regulates the

establishment of measures and procedures to prevent or detect crime. Responsible for upholding these procedures is the Financial Intelligence Unit (UIF), an institutional body focused on implementing policies to combat money laundering and financial terrorism, originally derived from Mexico's participation in the Financial Action Task Force (FATF-GAFI). The FATF is an intergovernmental organization founded on the initiative of the G7 to develop policies to combat money laundering.

According to the 2019 Mexican Central Bank Report, cyber risks were one of the most serious threats faced by the Mexican financial system (Banco de México, 2019). These risks have been exacerbated as a result of having accelerated the shift towards digital payments brought about by the COVID-19 pandemic (Procuraduría Federal del Consumidor, 2020). Criminals have taken the opportunity to exploit the COVID health emergency by targeting vulnerable individuals through cyber fraud. For instance, an estimated 4 million complaints were registered in the third quarter of 2020. Moreover, it also reveals a greater proportion of total frauds from 32% in 2016 to 69% in 2020 (Condusef, 2020). Likewise, in Q3 2020 e-commerce fraud complaints reached 4.1 million and mobile banking transactions fraud complaints reached 142k. The amount of cyber fraud reached 8 million MXN (UDS 430k). 85% of these complaints were resolved in favor of the user and the financial institutions had to cover the cost of the compensation.

The Financial Consumer Protection Agency (Condusef) data show that on average 463 fraud-related e-commerce and mobile banking transactions occur every hour as a result of scams. The situation is so dire that the legislative body has decided to intervene (Senate of the Republic, 2020). In March 2021, draft legislation was approved by the Senate Economy Committee to combat cyber fraud by increasing user protections and ramping up media campaigns (Senate of the Republic, 2021).

Banks utilize TMS as one of the core systems of the compliance capability areas (Figure 3). Generally, there are three main compliance capability areas in anti-money laundering and counter financing of terrorism (AML/CFT) compliance programs: Human Capital, Control Systems, and Management Process. A TMS monitors customer transactions daily or in real-time, as the volume of the transactions increases, so does the complexity of the monitoring. Accordingly, this complexity was particularly accentuated during the pandemic. An effective TMS should enable financial institutions to assess whether a customers' transactions are legitimate or suspicious when compared with historical information. However, the effectiveness of these systems is often not ideal and can incur a multitude of costs. For example, there is an operational cost in customer service when a customer's transaction is flagged as a fraud when it is legitimate. Whereas when a fraud is approved by the system, the costs are the chargeback and the transaction value. Furthermore, a customer may feel dissatisfied by the service provided and choose to take their capital elsewhere.

A TMS must be updated constantly to adapt to the rapid pace of technological advances, sophistication of criminal activity, and increasingly strict regulatory requirements for

compliance. Financial institutions find it difficult to track development and compare its compliance performance. Indeed, many Chief Compliance Officers (CCOs) are increasingly concerned with upholding compliance regulations, that is, proving to regulators that financial institutions are proactively identifying illicit activity. Although financial institutions have invested heavily into technology that supports the compliance framework, the automation of this technology has proven challenging to implement as it requires constant attention and time-consuming updates to keep it from becoming obsolete.

Traditionally, fraud is detected by rule-based systems that use algorithms to perform several fraud detection scenarios. To approve a transaction, a fraud analyst writes several rules one on top of another that identifies signals, such as unusually large transactions or unfamiliar geographic locations. Thus, rule-based systems need frequent incorporation of scenarios to cohesively detect correlations between user behavior and fraud actions. Moreover, these systems are unable to process real-time digital data streams, a worrying concern when given the recent shift to digital transactions during the course of the pandemic. Currently, most of the fine-tuning assessment of suspicious transactions still relies on compliance officers, indeed, personnel accounts for a significant share of compliance costs. For these reasons, financial institutions are interested in investing in machine learning and adapting deep learning fraud detection algorithms because these can be more effective in finding correlations (Aggarwal, Wareman, and Lehman 2020, 140-166), (Jullum, Loland, Huseby, Anonsen, and Lorentzen 2020). What follows is a brief overview of the current state of artificial intelligence approaches.

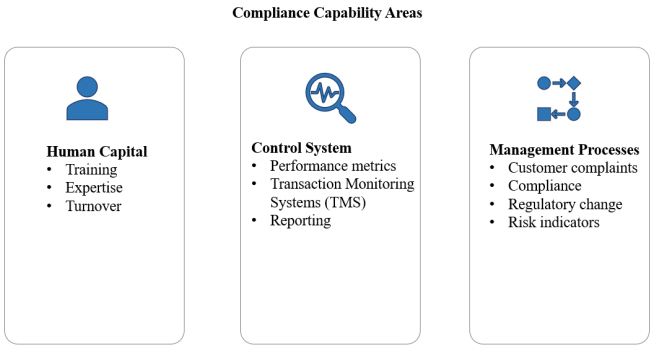


Figure 3. *Types of Compliance Capabilities Areas. Transaction Monitoring Systems mapping within a very simplified compliance capability.*

3. Challenges in financial crime analytics

This section covers the current challenges in the fight against financial crime, in particular, we will discuss the use and generation of financial synthetic data and the use of a reliable benchmark tool to test fraud detection algorithms. We will contemplate the relevance of measuring the effectiveness of financial crime controls in the context of a growing mobile payments market that promotes financial inclusion in Mexico.

3.1 Access to financial datasets

Access to real-world data is a domain agnostic challenge that is present throughout many sectors, whether they are academia, industry, or governance. Financial transactions across different channels, such as credit card payments, insurance claims, or online payments, are generated by financial institutions. These datasets are not publicly available due to privacy regulations. Furthermore, having access to fraud-annotated datasets, that is, transactions that have been proven to be fraudulent, is an even greater endeavor . As such, if an anonymized dataset is publicly available for academic research; the time window shared might not be sufficient for a researcher to adapt a fraud detection algorithm “to learn” what is normal and abnormal consumer behavior.

3.2 Imbalanced datasets

As a result of the imbalanced nature of transactional data that machine algorithms are trained on, their adaptation for fraud detection is more challenging. An imbalanced class distribution refers to a set of data where the number of fraudulent transactions is significantly lower than legitimate transactions i.e. around 1% or less of all total transactions. Conventional ML classification algorithms are designed to improve the overall correct identification of fraudulent and normal transactions by reducing the error range of the algorithm. However, highly imbalanced datasets can heavily skew their performance. As a result, classification algorithms can be biased or inaccurate in the identification of the fraudulent minority class. Nevertheless, undersampling and oversampling techniques, such as Synthetic Minority Oversampling Technique (SMOTE), are often used in addressing these issues.

3.3 Synthetic data

Synthetic data has shown to be a cost-effective value proposition in several domains (Barr *et al*., 2020). In finance, its use case has gained significant interest, in particular in the areas of financial crime detection and compliance (Lopez-Rojas, Axelsson, and Baca, 2018) through addressing challenges posed by accessing financial datasets due to privacy regulation (Barse, Kvarnstrom, Jonsson, 2004). Furthermore, within synthetic data, conditions yet to be encountered can be simulated. Thus, potential strategies that criminals deploy can be taken advantage of, to adapt current control systems for plausible criminal behavior.

Synthetic data can be defined as entirely new data points generated to resemble or keep the statistical properties of the real dataset. A key consideration is the effectiveness of the synthetic data when it is used. For instance, a core principle of validating the usability of synthetic data is to observe if an ML model built from synthetic data performs as well as models built from real data. Therefore, synthetic data must aim at mimicking the statistical properties of the real dataset, however, a synthetic dataset cannot be a replica of real data,

for that would compromise on the privacy we aim to uphold.

The quality of synthetic data relies on two principal factors; a) type of methodology applied to the generation of the data and b) robustness, quality, and feature depth of the original data. To begin with, simply extrapolating the correlations or patterns among the real dataset's attributes or variables is not enough to create quality synthetic data. Also, it is sound to assume that we are starting with a real dataset that is large enough, that has as much information on its attributes as possible, and that is unbiased. Although there is no consensus on the approach to create synthetic data, it is important to note that the chosen method should be explainable i.e. answer questions about the data creation process (Barr *et al*., 2020). Indeed, a challenge that synthetic data faces is of limited use-cases because the quality and trustworthiness of synthetic data have not been accepted by practitioners. Thus, it is advisable to also make a comparison between the synthetic data and the real dataset to fully understand its reliability.

In finance, generative models have gained significant interest, as they can provide privacy preservation *i.e.* the synthetic data cannot be mapped back to the original dataset (Assefa, 2019), although recent work has found that such claims may be unsupported and require further research to understand (Stadler, Oprisanu, and Troncoso, 2020). Regardless of such opinions, generative methods are considered better than the de-identification methods, such as data masking, shuffling, and encryption, given that de-identification generated synthetic data can be mapped back one to one to the original dataset.

3.4 The limitations of ML metrics in financial crime analytics

The implementation of machine learning or deep learning can ensure efficient and effective regulatory compliance programs and save money for the financial institution. Metrics are used to evaluate the performance of fraud detection algorithms. Fraud detection is at the basic level a binary classification problem where it is predicted whether a transaction is a fraud (positive) or non-fraud (negative). There are four possible classification outcomes; when an algorithm correctly classifies a transaction as fraud or non-fraud it is called True Positive (TP) and True Negative (TN), respectively. When the algorithm erroneously identifies a non-fraud transaction as a fraud, it is called Type I error or False Positive (FP). A Type II error, also called False Negative (FN), is when a transaction is erroneously identified as non-fraud given it is a fraud.

Precision, recall, and accuracy are the traditional metrics used to assess the performance of classification algorithms, unfortunately, these metrics are inadequate in fraud detection, albeit there is some preference for precision-recall curves (Saito & Rehmsmeier, 2015). Precision measures the proportion of fraud identifications that were correct. Recall measures the proportion of actual frauds that were identified correctly. F1 Score is a function of precision and recall. Accuracy describes "close-to-true value", however, in the domain of financial fraud it is not very useful since the datasets are imbalanced; that is,

fraudulent transactions are less than 1%. So, if a model predicts that at least 99% of all the transactions are legitimate, this is not a very helpful measure to gauge the performance of the classifier. To address some of these issues, we have seen the rise of Cohen's Kappa and Matthews Correlation Coefficient (MCC), which tries to address a larger scope of variables to gauge classifier performance. Although the use of K has been an area of debate inside ML domains (Delgado & Tibau, 2019) and outside them (Bexkens et al., 2018), we see a much more positive reception for MCC (Chicco & Jurman, 2020).

Nonetheless, we feel that these measures are still not enough to formulate an adequate benchmark with machine learning in the domain of financial fraud and propose that model performance should also be evaluated in terms of the financial amounts of fraud missed and falsely identified normal transactions to determine the appropriate metrics, and thus achieve a much more robust and discriminative classifier.

4. Benefits of a benchmark in financial services

4.1 Fraud detection in financial services

Compliance departments face the challenge of reducing the number of innocent people wrongfully accused. To cope with this problem financial institutions are exploring the use of fraud detection algorithms and data analysis technologies to have a more accurate and precise fraud detection system. However, approaches to streamlining and automating banks' monitoring and testing processes often lead to more bias and privacy concerns (Johnson et al., 2019).

Existing solutions within these domains are tailored to integrate with TMS. These are often modified by their respective vendors to provide some deeper level of analytics and fraud detection over their competitors. Nonetheless, there are also solutions outside TMS that serve as platforms for training sophisticated machine learning models for swift deployment and classification of fraud. These either serve to expand on the current pool of data available through adding more parameters or expanding existing ones to train more robust models. Although these solutions do increase the detection of fraud, they are still under the constraints of data privacy and are severely limited in the data they can utilize.

As shown in Figure 4, financial institutions have the task at hand to develop and maintain controls for their TMS. It is crucial for the development of effective fraud detection to have accurate alerts. The generated alerts are considered 'good' when the number of fraudulent transactions detected is representative of the total fraud, and that the number of innocent people who are mislabeled as fraudsters is as low as possible. Accordingly, the greatest challenge for banks is to understand beforehand, how effective a new control is (an algorithm responsible for classification) in contrast to already active control. This laborious task regularly involves numerous iterations of trial and error, taking up months if not years to perfect. Once the new control has been assessed to be within an adequate range of

classification, it is deployed to make actual decisions in real-time. However, over-time the performance of the control will diminish, as criminals will deviate from their usual behavior to avoid being detected. Thus, this time-consuming process is iterated on repeatedly, leaving small gaps for criminals to exploit and profit from. Ideally, with an adequate benchmark to measure the performance of these controls ahead of time, financial institutions will be able to cut down on the time it takes to deploy their controls, and thus utilize the most appropriate control for current criminal behavior.

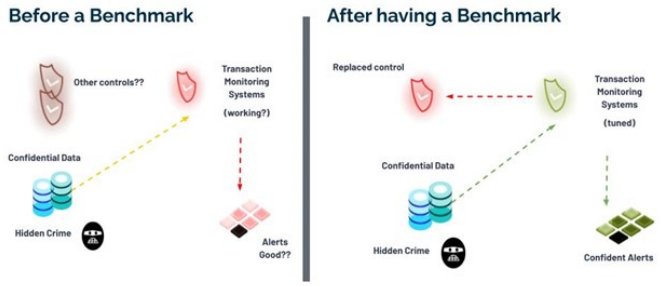


Figure 4. *Impact of assessing financial crime controls (benchmarking).*

4.2 The future of technology in Digital Financial Services

The COVID-19 pandemic drove the Mexican population to change their shopping, buying and paying for services. Consequently, a large proportion of Mexico has moved from physical currency towards digital solutions, such as mobile phones or online payment systems. Nonetheless, frauds have also followed this trend and thereby increased. The digital transformation of many sectors, as in the case of the financial sector, has been accelerated as a result of the Covid19 Pandemic. This suggests, as the profitability of the financial entities rises, financial inclusion will be playing a larger role in these ecosystems.

Despite the policies and regulations set in place, technological advances in finance have been misdirected. If digital banking is to be considered as a tool to help fight inequality and contribute to the economic empowerment of the non-banked population, it is necessary to address several challenges.

First of all, cultural, practical, structural, or political factors behind why some Mexicans seem to be hesitant to adopt technological advances need to be identified and addressed. For some, it was surprising that Samsung Pay decided to not continue operations in Mexico. Hence, it is important to understand the barriers contributing to the low use of financial apps or services such as CoDi or Samsung Pay. High commissions are arguably one of the most daunting hurdles that consumers will face when considering the adoption of digital payments. These are exacerbated even more so for the vulnerable populations who are more prone to economic hardships. Moreover, indigenous or older populations may find it more difficult to adapt to the new technologies, thus great care must be taken to evaluate the app’s accessibility for all types of demographics. Nevertheless, perhaps an important issue to address is building trust in the system: implementation of laws and regulations to protect

consumers would go a long way in establishing the foundation for this. Yet, if this is not addressed, it will perpetuate a perception of impunity, and thus, mistrust in the banks and systems which tie these delicate and intricate systems together.

Secondly, there should be further work regarding how to improve current laws and regulations. As such, in the realm of mobile banking regulatory challenges, we observe a system forced to fit into a rigid framework that does not work for people with low income or credentials. Reasonably, we should think outside of the box to develop fair and inclusive financial services that are able to broaden the scope of inclusion, while still being able to remain profitable and safe. ML-based solutions will have an important role in promoting financial inclusion. Furthermore, one can debate the degree of society who will stay 'cashless'. As such, it is worth noting that there should not be a situation where discrimination against a particular group remains, therefore, we must refine our systems from the ground up.

A robustly designed benchmarking tool that allows for the assessment of how reliable a fraud detection system is performing, is an essential building block in the creation of a safer digital ecosystem that can incorporate a wide range of demographics that may have been previously omitted. Indeed, the confidence that crime is actually being reduced will not only make consumers feel safer about where they store their money; but also smooth out the customer onboarding process, which more often than not, can dissuade or discriminate against potential customers.

The growth of innovative technologies is another essential step in building a fair economical digital society, however, areas of the FinTech Law must be revised to lower the capital entry barriers to make Mexico an appealing place for fintech firms. Other countries, such as the United Kingdom, have already begun to push their policies for innovation. One such example is a digital sandbox of data to speed up the development of companies that may not have access to the necessary resources for their solution. This combination of synthetic data and open source code has found great success in the exploration and scaling of innovative financial inclusion technologies, hence, it has been praised by an independent review of the UK fintech industry (Kalifa, 2021). Consequently, the inclusion of regulatory guidance for other innovations in financial services, for instance, balance sheet lending, numerous investment services, or central bank digital currencies; would further help in the refinement of financial inclusion.

Thirdly, when taken into account that regulatory bodies themselves will use new technologies to modernize or improve operations and processes. The exponential growth of the mobile payment strategy to boost financial inclusion will diversify the communities that partake in the Mexican Financial System. Therefore, regulators must be prepared to deal with a substantial growth of diverse scenarios and adapt accordingly by utilizing cutting-edge technology, or risk becoming overwhelmed and unable to supervise the financial system.

Fourthly, an approach to public policy, laws, regulations, and governance should be created from the ground up with a digital native mentality. Certainly, current governance has evolved following technology advancements, but it is undeniable that they are from times past. For case in point, a classic TMS may be considered an archaic legacy system that has evolved to adapt to new rules, yet its evolution is limited by the direct interaction of the user. Thus, it is becoming much more obsolete over time in contrast to a TMS that has machine learning embedded for their decision-making. There are yet many technological advancements and changes coming from the market that need swift adaptation of current solutions. To name just a few, we have digital banking, decentralization from blockchain technology, digital currencies. All providing a new influx of data with higher complexity than anything we have seen before; something that the current and older framework is most definitely not prepared for.

Conclusively, there is a great challenge to create public policy, laws, regulations, and governance that adhere to the rapid advancement of technology. This is further complicated by a fast-spreading marketplace that aims to foster an environment of competitiveness and innovation, while also trying to boost financial inclusion and protect customer rights. Thus, it is necessary to work on strategies, for instance, an open, continuous, and inclusive dialogue with shareholders can be adopted. This can be complemented by workshops among policymakers, academics, and finance and technology experts; to discuss ways of benefiting from technological advancements, and thus, avoiding the pitfalls that come with complex regulatory environments.

For financial crime and AML, a Triple Helix cooperation model among the academy, financial institutions, and law enforcement agencies have been suggested to have several benefits for the fight against financial crime (Lopez-Rojas, 2019). A major challenge is that institutions are currently uncompromising in shifting away from their respective domains; hence, the lack of mutual understanding and trust between academia and financial institutions. Therefore, hard work is needed to understand each other’s needs and reach a consensus on a collaborative environment for innovation.

5. Conclusion and Future Work

Safety is key to build up and maintain consumer trust in the use of digital banking. A benchmarking tool is extremely useful to help financial institutions manage the regulatory risks posed to them by fraud. In this paper we explored the particular case of Mexico and discussed how benchmarking of financial crime controls can help create a safer environment that will encourage Mexicans to use digital financial services.

During times of instability, such as the recent pandemic, as well as looking ahead to the post-COVID19 era, it is necessary to work on strategies that strengthen the digital financial sector. Thus, it is essential to create a more dynamic infrastructure, one that boosts economic recovery, bolsters inclusion, and can even help grow and maintain itself as the

Latin American fintech hub. Indeed, given the multidimensional nature of the challenges of financial inclusion; the design of such strategies must incorporate knowledge of both the economic and social context. Which unites the stakeholder's (e.g., consumers, Banxico, public and private sector) incentives and interests, with the innovation of technologies (e.g., new business models, products).

The crux of addressing these problems is to acknowledge that the industry currently suffers due to a lack of quality real data, due to privacy restrictions or lack of optimized data storage and cleaning. This is especially a pain point when a set of labeled data is required, as it can take many months of restructuring organizational levels and data infrastructure to make use of real historical data.

Synthetic data does not have such a limitation, because it can be created quickly and integrated with clear labels for machine learning tasks. However, there is currently no clear measure on the quality of synthetic data that everyone can agree on; moreover, the development of synthetic data for sophisticated financial crime, such as money laundering, is still embryonic in its development.

Although further work is required to define such strategies given the evolution of technology and the way trustable finance will be conducted in the future; some key areas that will drive this innovation forward are:

A. increasing the quality and confidence in the use of synthetically generated data

B. further development of benchmarking tools using synthetically generated data for machine learning algorithms, and

C. building the foundations to extend the work from privacy to different aspects of trustworthiness such as explainability and robustness for the generated data

These principles will gear towards the development of high-fidelity synthetic data to better inform data-driven learning, and solutions in financial crime for benchmarking machine learning controls, consequently, paving the way to the creation of a safer and more accessible financial digital ecosystem.

References

– Aggarwal Nikhil, Wareman Sean, and Lehman Rasmus. 2020. "Applications of machine learning in the identification, measurement and mitigation of money laundering." Journal of Financial Compliance 4(2), pp. 140-166.

– Alonso Javier, Fernández de Lis Santiago, Hoyo Carmen, López-Moctezuma Carlos and Tuesta David. 2013. "Mobile banking in Mexico as a mechanism for financial inclusion: recent developments and a closer look into the potential market." BBVA Research Working Papers Number 13/20, June 2013. https://www.rrojasdatabank.info/mobilebanking4.pdf

– Avendaño Carbellido Octavio. 2016. “La reforma financiera y su impacto en el usuario.” Entreciencias: Diálogos en la Sociedad del Conocimiento, vol. 4, núm. 10, 2016 Universidad Nacional Autónoma de México. May 2016, DOI: http://dx.doi.org/10.21933/J:EDSC.2016.10.182 https://www.redalyc.org/jatsRepo/4576/457646537003/html/index.html

– Avendaño Carbellido Octavio. 2018. “The challenges of electronic banking in Mexico.” Revista del Instituto de Ciencias Jurídicas de Puebla. ISSN 1870-2147. New epoch Vol. 12, No. 41: 87-108

– Banco de Mexico. 2019. “Reporte de Estabilidad Financiera.” Banco de Mexico, December, 2019. https://www.banxico.org.mx/publicaciones-y-prensa/reportes-sobre-el-sistemafinanciero/%7B04E197EE-B6FC-7BA1-72A0-32D3E6C9BF28%7D.pdf

– Barr Brian, Xu Ke, Silva Claudio, Bertini Enrico, Reilly Robert, Bruss C. Bayan, Wittenbach Jason D. 2020. “Towards Ground Truth Explainability on Tabular Data” arXiv:2007.10532v1 [cs.LG] 20 Jul 2020

– Barse E.L., Kvarnstrom H., Jonsson E. . 2004, Synthesizing test data for fraud detection systems Conference Paper Conference: Computer Security Applications Conference, 2003. Proceedings. 19th Annual IEEE Xplore DOI: 10.1109/CSAC.2003.1254343

– Bexkens Rens, Claessen Femke MAP, Kodde Izaak F, Oh Luke S, Eygendaal Denise, and van den Bekerom Michel PJ. The kappa paradox. Shoulder & Elbow, 10(4):308–308, October 2018. ISSN 1758-5732. doi: 10.1177/1758573218791813. URL https://doi.org/10.1177/1758573218791813. Publisher: SAGE Publications Ltd.

– Chicco Davide, Jurman Giuseppe. The advantages of the Matthews correlation coefficient(MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics , 21 (1):6, 2020. ISSN 1471-2164. doi: 10.1186/s12864-019-6413-7. URL https://doi.org/10.1186/s12864-019-6413-7

– CNBV. 2020. “Inclusión Financiera en México.” https://www.gob.mx/cms/uploads/attachment/file/590346/Inclusion_financiera_mexico_difusion.pdf

– Condusef. 2017. “La CONDUSEF informa sobre las comisiones bancarias y sus reclamaciones” https://www.condusef.gob.mx/?p=contenido&idc=379&idcat=1

– Condusef. 2021. “2020 Third trimester report.” https://www.condusef.gob.mx/?p=estadisticas

– CNBV, 2020, p.48 https://www.gob.mx/cms/uploads/attachment/file/590346/Inclusion_financiera_mexico_difusion.pdf

– CNBV, INEGI Encuesta Nacional de Inclusión Financiera https://www.inegi.org.mx/programas/enif/2018/

– Das Sanjiv, Donini Michele, Gelman Jason, Haas Kevin, Hardt Mila, Katzman Jared, Kenthapadi Krishnaram, Larroy Pedro, Yilmaz Pinar, and Zafar Bilal. 2020. "Fairness measures for machine learning in finance."

– Delgado Rosario and Tibau Xavier-Andoni. 2019. "Why Cohen's Kappa should be avoided as performance measure in classification." PLOS ONE , 14(9):e0222916, September 2019. ISSN 1932-6203. doi: 10.1371/journal.pone.0222916. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0222916 . Publisher: Public Library of Science.

– DOF (Diario Oficial de la Federacion). 2018. 9 March. http://www.diputados.gob.mx/LeyesBiblio/ref/lritf/LRITF_orig_09mar18.pdf

– DOF. (Diario Oficial de la Federacion). 1990. Ley de Instituciones de crédito. July 18, 1990. https://www.senado.gob.mx/comisiones/finanzas_publicas/docs/LIC.pdf

– Estanol Eduardo. 2019." Los bancos en México logran mejor desempeño que en sus países de origen." Expansion, February 22, 2019.

– Johnson Kristin, Pasquale Frank, and Chapman Jennifer. Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation. FORDHAM LAW REVIEW, 88:32, 2019.

– Juarez Edgar. 2015. "México, una joya para los bancos extranjeros." El Economista, March 17, 2015 https://www.eleconomista.com.mx/sectorfinanciero/Mexico-una-joya-para-los-bancos-extranjeros-20150317-0182.html

– Jullum Martin, Loland Anders, Huseby Ragnar Bang, Anonsen Geir, and Lorentzen Johannes. 2020. "Detecting money laundering transactions with machine learning" Journal of Money Laundering Control, January 4 2020. ISSN: 1368-5201

– Kalifa, Ron.. 2021. "Kalifa Review of UK Fintech" https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/971370/KalifaReviewofUKFintech.pdf

– Karpoff Jonathan M. 2020 "The future of financial fraud." Journal of Corporate Finance , pp. 101694.

– Lopez-Rojas E. A., Axelsson S., and Baca D. 2018. "Analysis of Fraud Controls Using the PaySim Financial Simulator." International Journal of Simulation and Process Modelling. 13 (4), pp. 377-386, ISBN: 1740-2131.

– Lopez-Rojas Edgar Alonso. 2019. "Triple Helix Approach for Anti-Money Laundering (AML)

Research Using Synthetic Data Generation Methods." The 10th International Conference on Society and Information Technologies: ICSIT 2019 At: Orlando, USA. ISSN 00401625. doi: 10.1016/j.techfore.2016.04.024.

– Maina Juliet. 2018. "Manual regulatorio y de politicas de dinero móvil" https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2018/09/GSMA-Mobile-Money-Policy-Handbook_Spanish.pdf

– Morales Guzman, Rafael. 2018. "How Financial Tech Can Aid Financial Inclusion in Mexico." Cornell Policy Review, November 16, 2018. http://www.cornellpolicyreview.com/fin-techmexico/

– Naghavi Nika. 2020. "State of the Industry Report on Mobile Money" https://www.gsma.com/sotir/wp-content/uploads/2020/03/GSMA-State-of-the-Industry-Report-on-Mobile-Money-2019-Full-Report.pdf

– OECD. 2019. "OECD Economic Surveys Mexico" https://www.oecd.org/economy/surveys/Mexico-2019-OECD-economic-survey-overview.pdf

– OCDE/CAF. 2020. "Estrategias nacionales de inclusión y educación financiera en América Latina y el Caribe: retos de implementación." http://www.oecd.org/financial/education/Estrategias-nacionales-de-inclusi%C3%B3n-y-educaci%C3%B3n-financiera-en-America-Latina-y-el-Caribe.pdf

– Pandey Neena, Pal Abhipsa, et al. 2020. "Impact of digital surge during COVID-19 pandemic: A viewpoint on research and practice". International Journal of Information Management , 55:102171.

– Procuraduria Federal del Consumidor. 2020. "Todo a un clic. Compras sin salir de casa (In Spanish)." Mexican Government-Federal Consumer Protection Agency, September 2, 2020. https://www.gob.mx/profeco/articulos/todo-a-un-cliccompras-sin-salir-de-casa?idiom=es

– Saito Takaya and Rehmsmeier Marc. 2015. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLoS ONE, 10(3). ISSN 1932-6203. doi: 10.1371/journal.pone.0118432. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4349800/.

– Saldívar Alejandro, Zavaleta Noé. "Túmin, una moneda que resiste." Proceso. December 28, 2020 https://www.proceso.com.mx/nacional/estados/2020/12/28/tumin-una-moneda-que-resiste-255250.html

– Senate of the Republic. 2021. "Impulsa Comisión de Economía dictamen para proteger a usuarios; aumentan fraudes cibernéticos". March, 2021http://comunicacion.senado.gob.mx/index.php/informacion/boletines/50425-impulsa-comision-de-economia-dictamen-para-proteger-a-usuarios-aumentan-fraudes-ciberneticos.html

– Senate of the Republic. 2020. “Urgen legislar en materia de ciberseguridad ante el incremento de delitos digitales.” October 9, 2020 http://comunicacion.senado.gob.mx/index.php/informacion/boletines/49401-urgen-legislar-en-materia-de-ciberseguridad-ante-el-incremento-de-delitos-digitales.html

– SHCP/Ministry of Public Education. 2020. “Politica Nacional de Inclusión Financiera.” https://www.gob.mx/cms/uploads/attachment/file/585234/PNIF_2020.pdf

– Stadler Theresa, Oprisanu Bristena, and Troncoso Carmela. 2020. “Synthetic Data – A Privacy Mirage” arXiv:2011.07018v2 [cs.LG] 11 Dec 2020

– van Driel Hugo. 2019. “Financial fraud, scandals, and regulation: A conceptual framework and literature review”. Business History, 61(8):1259–1299, 2019. doi: 10.1080/00076791.2018.1519026

– World Bank. 2018. “Financial Inclusion.” https://www.worldbank.org/en/topic/financialinclusion/overview

– World Bank. 2020. “Mexico Overview” https://www.worldbank.org/en/country/mexico/overview

– Zhang Wanrong, Ohrimenko Olga, Cummings Rachel. 2020. “Attribute privacy: Framework and mechanisms.” doi: arXiv:2009.04013.

Acknowledgments

This work was supported by UK’s innovation agency, Innovate UK, under granted projects to EalaX Ltd: FraudSim 82929 and CP-Mark 89039.

We thank the anonymous referees and Dr. Xi (Sisi) Hu (Program Fellow at Harvard Law School) for their useful comments.

Authors’ Bio

Ulf Norinder: Dr Norinder is a recognized researcher in both the pharmaceutical sector and academia. Ulf worked at AstraZeneca, Lundbeck and at the Department of Computer and Systems Sciences, Stockholm University. Dr. Norinder has over 6000 citations, his papers cover topics on machine learning, pattern recognition, computational toxicology and cheminformatics. Ulf is Ealax’s Chief Scientific Advisor.

Edgar Lopez-Rojas: Dr Lopez-Rojas is the lead expert and founder of Ealax. Edgar has a PhD in Computer Science in the development of simulation methods for fighting financial fraud. Edgar has previously worked as Financial Crime Analytics Expert at Simudyne, as a

researcher at the Digital Forensics group at NTNU, and was part of the ISO Standards Norge committee for Fintech SN/K 250 “Bank and Financial Services”.

Juan Sebastián Cárdenas-Rodriguez: Graduate of Mathematical Engineering. Mainly interested in subjects surrounding Mathematics and Computer Science, with works in hypercomputation, discrete simulations and agent-based simulations. Sebastian is co-founder of CannBlock and is an intern software developer at Ealax.

Daniel Turner-Szymkiewicz: Graduate of BSc Pharmacology and MSc Neuroscience with an interest in the analysis and refinement of synthetic data through machine learning in the field of financial crime. Currently working as a Data Scientist for Ealax in the application of machine learning models on synthetic data in fraud detection.

Miroslawa Alunowska Figueroa: PhD in Engineering Science with an interest in machine learning applied to finance. Miroslawa has experience in developing proof of concept models, optimizing procedures and managing projects.