INSTITUTO TECNOLÓGICO DE BUENOS AIRES – ITBA

MAESTRIA EN DIRECCION ESTRATEGICA Y TECNOLOGICA

ESCUELA DE POSTGRADO

Utilización de bases de datos biométricas en Brasil para reducción de fraude de identidad en una fintech

AUTOR: Casal, Eduardo Enrique (Leg. N° 46465)

TUTOR DE TESIS: Fahnle, Pablo

TESIS PRESENTADA PARA LA OBTENCIÓN DEL TÍTULO DE MAGISTER EN DIRECCION ESTRATEGICA Y TECNOLOGICA (ARGENTINA) Y MASTER EJECUTIVA EN DIRECCIÓN ESTRATEGICA Y TECNOLOGICA (ESPAÑA)

BUENOS AIRES

SEGUNDO CUATRIMESTRE, 2020

No autorizo al Instituto Tecnológico de Buenos Aires (ITBA) a publicar y/o difundir en medio alguno el contenido de este trabajo el cual posee fines exclusivamente académicos correspondiente a la maestría cursada en esta Institución. Su uso y difusión queda limitado a jurados de tesis y autoridades de la universidad.

Índice

Índice de figuras

Dedicatoria

A mis padres que me ayudaron a lograr mis estudios universitarios, educando con su trabajo, su esfuerzo y su ética. Sin ellos no sería quién soy hoy.

A mi pareja que me empujó a estudiar la maestría, con quién compartí toda la cursada.

Introducción

FacilPay es una empresa tecnológica brasilera dedicada a realizar micro-créditos a individuos con poca o nula bancarización. Al ser una empresa joven, los sistemas de validación de identidad fueron construidos a medida que se encontraron problemas, teniendo en consideración la disponibilidad de datos del público objetivo. En este trabajo se analizará la implementación de un sistema de validación biométrica con el objetivo de reducir el fraude de identidad en dicha startup.

La biometría facial se refiere a las mediciones de las características de los rostros de los individuos. Los sistemas de biometría facial se especializan en identificar y diferenciar personas a partir de estos valores.

Uno de los métodos más conocidos de biometría son las huellas dactilares. Si bien las impresiones de manos se han usado de modo artístico hace casi 40.000 años (de acuerdo a un estudio publicado en Nature[1]), su uso a modo identificatorio, es decir para establecer la identidad de una persona, data al menos del 221 AC. Como escribe Jeffery G. Barnes en "The Fingerprint Sourcebook": *Los chinos son la primera cultura que se conoce que usaran huellas dactilares para la identificación. El ejemplo más antiguo viene de un documento chino titulado "El Volumen de una Investigación de la Escena del Crimen - Robo", de la Dinastía Qin (221 a 206 B.C.). El documento contiene una descripción de cómo se utilizaban las impresiones de las manos como un tipo de evidencia (Xiang-Xin y Chun-Ge, 1988, p 283).*

El autor continúa explicando cómo las huellas dactilares se utilizaban en sellos de arcilla, en conjunto con un nombre, para establecer la identidad del signatario de forma fehaciente y evitar que se alterase el contenido del mensaje.

Esta noción de que las huellas dactilares identifican a un individuo se presenta en distintas culturas desde entonces como método para agregar autenticidad a documentos. En Asia pueden encontrarse referencias al acto de firmar con dedos o la palma de la mano en Japón e India. No es hasta el año 1788 que se reconoce la unicidad de la huella dactilar en Europa.

[1] Aubert, M., Brumm, A., Ramli, M. *et al*. Pleistocene cave art from Sulawesi, Indonesia. *Nature* 514, 223–227 (2014).

La impresiones dactilares no fueron siempre el método utilizado para identificar a un individuo. En un principio se utilizaban identikits (y luego fotografías) de personas, junto con un conjunto de mediciones de un individuo que, mediante la aplicación de una fórmula, debían identificar de manera única al sujeto. Dicho sistema, inventado por el antropólogo francés Alphonse Bertillon, estuvo en uso durante más de 30 años, hasta que dos hermanos gemelos fueron condenados a la misma prisión, donde se descubrió que ambos tenían el mismo conjunto de mediciones, pero huellas dactilares distintas.[2]

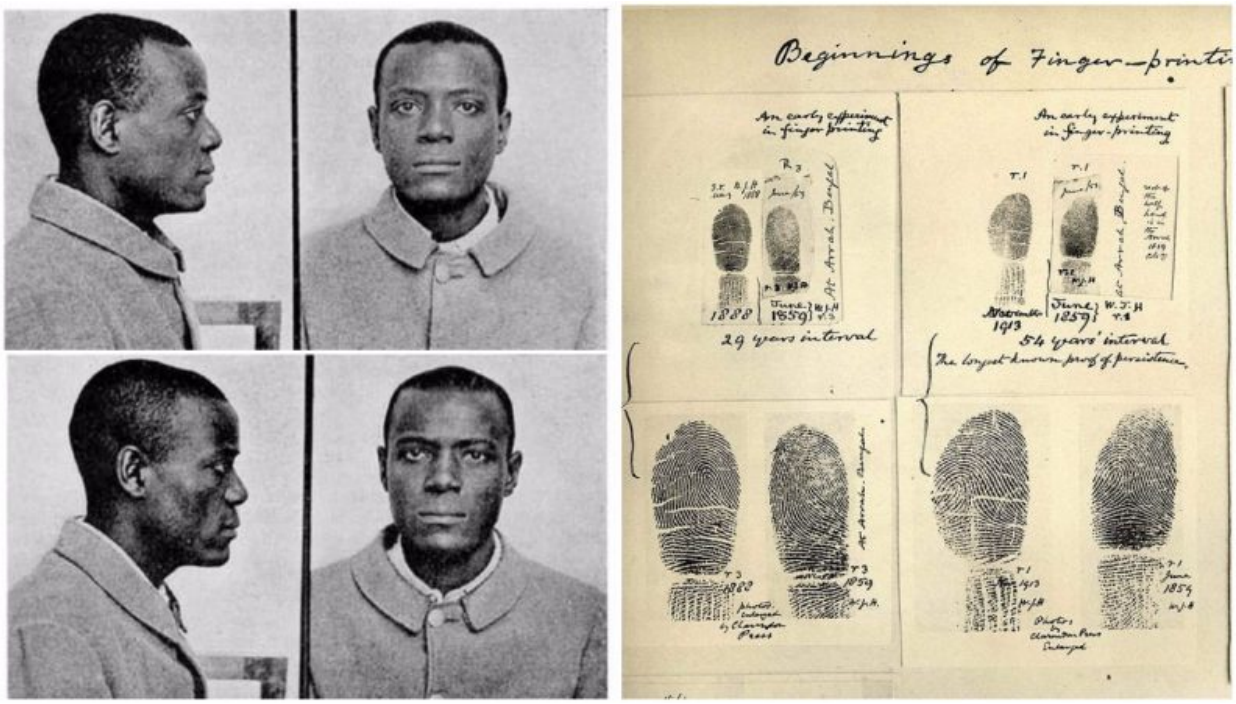

Figura 1: Will West y William West

[2] Browne, Douglas G., Alan Brock, *Fingerprints: Fifty Years of Scientific Crime Detection*, George G. Harrap & Co., Ltd., London, 1953, 105–106

Gracias a avances en tecnología, los métodos de biometría se han vuelto cada vez más comunes a la hora de implementar medidas de seguridad. Si bien la mayoría de los móviles hoy en día cuentan con la capacidad de identificar a un usuario a través de sus facciones, las bases para las implementaciones que existen en la actualidad están presentes hace tiempo: La idea de identificar un individuo a través de sus retinas fue propuesta en 1936 por Frank Burch[3], en 1964 Woodrow Bledsoe desarrolló un método semi-automático de reconocimiento facial, y ya en los años 70 aparecieron los primeros sistemas comerciales de reconocimiento de manos.

Sin embargo, los dispositivos de seguridad biométrica para usuarios finales no fueron introducidos hasta el 2004 en computadores personales. IBM fue el primero en introducir el lector de huellas en su línea Thinkpad[4] y Microsoft lanzó su propio lector de huellas para ordenadores que funcionasen sobre sus sistemas operativos[5]. En móviles, no fue hasta el 2011 que Motorola disponibilizó el Atrix, el primer smartphone con lector de huellas dactilares. Si bien el reconocimiento facial como medida de seguridad biométrica estaba disponible ese mismo año en la versión 4 del sistema operativo Android, no fue hasta el 2017 que adquirió popularidad de la mano de la implementación de FaceID de Apple que utiliza sensores de profundidad para evitar ser engañado.

En otras palabras, si bien existen rasgos únicos que identifican a las personas y existe la capacidad de medirlos hace casi 200 años, sólo hace unos pocos es que fueron implementados de forma que usuarios finales puedan utilizarlos de forma segura.

Actualmente se utilizan métodos de biometría digital para restringir acceso a zonas de alta seguridad[6], complementando métodos físicos (como tarjetas inteligentes) que podrían ser robados o clonados.

---

[3] Mohammad S. Obaidat, Issa Traore, Isaac Woungang, *Biometric-Based Physical and Cybersecurity Systems*, Springer, 2018, 45-46

[4] *IBM Introducing Fingerprint Reader into Laptop*, TechNewsWorld, 2004 https://www.technewsworld.com/story/37017.html, recuperado 13/10/20

[5] *Researcher Hacks Microsoft Fingerprint Reader*, PCWorld, 2006 https://www.pcworld.com/article/124978/article.html, recuperado 13/10/20

[6] *Understanding biometrics: How to choose the right biometric technology for your organisation, Argus Global* https://www.planetbiometrics.com/creo_files/upload/article-files/how_to_choose_the_right_biometric.pdf

El avance de la digitalización también ha permitido instalar biometría facial como complemento para identificar posibles sujetos de interés mediante cámaras de seguridad. Estos métodos generan controversia en la población, ya que la misma tecnología puede ser utilizada para identificar criminales, o para identificar individuos sin antecedentes, como ser manifestantes en una protesta[7].

Estado actual de las fintech

Desde el 2010 las inversiones en fintechs crecieron de 2 mil millones de dólares a 242 mil millones de dólares en 2019, según estudios de CBInsights[89]. El crecimiento del mercado se debe, en parte, a la entrada de nativos digitales: jóvenes milennials que sienten que sus necesidades son mejor atendidas en nuevas startups. Esto es en parte por la mentalidad de pensar primero en el móvil, según los analistas de *VisionCritical*[10].

[7] *In Hong Kong Protests, Faces Become Weapons,* New York Times, 2019, https://www.nytimes.com/2019/07/26/technology/hong-kong-protests-facial-recognition-surveillance.html, recuperado 13/10/20

[8] *Global Fintech Investments Surged in 2018 with Investments in China Taking the Lead, Accenture Analysis Finds; UK Gains Sharply Despite Brexit Doubts*, Accenture, 2019 https://newsroom.accenture.com/news/global-fintech-investments-surged-in-2018-with-investments-in-china-taking-the-lead-accenture-analysis-finds-uk-gains-sharply-despite-brexit-doubts.htm, recuperado 13/10/20

[9] *The State Of Fintech: Investment & Sector Trends To Watch 2019Q4 Report*, CBInsights, 2019 https://www.cbinsights.com/research/report/fintech-trends-q4-2019/

[10] *3 factors driving fintech startups*, VisionCritical, 2019 https://www.visioncritical.com/blog/3-factors-driving-fintech-startups, recuperado 13/10/20

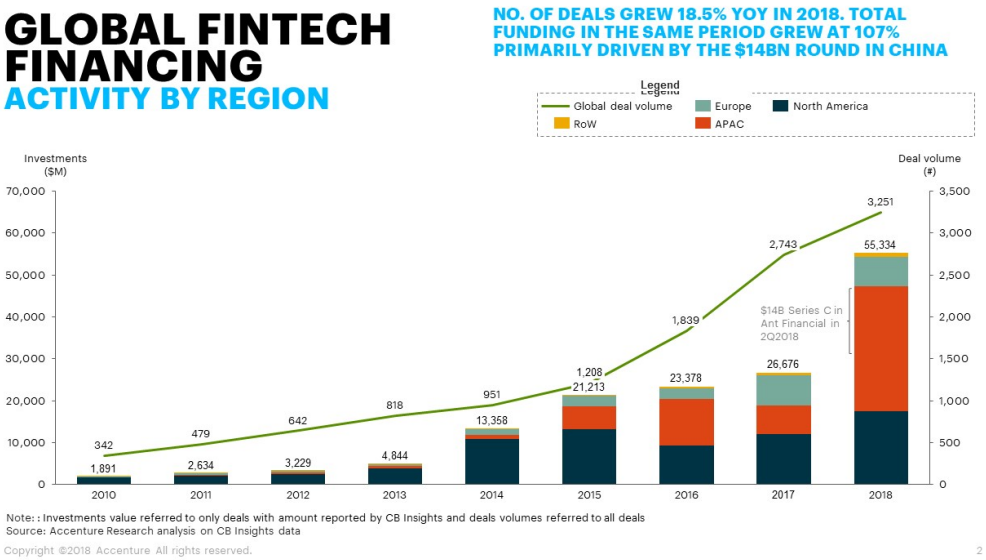

Figura 2: Montos de financiamiento de fintechs de 2010 a 2018

Entre los sectores que más crecieron se ubican las industrias de pagos y de préstamos, según un reporte de *Accenture*. Ejemplos de éstas son NuBank, un banco digital fundado en 2018, y Stripe, un procesador de pagos (en la categoría de pagos) y Affirm que presta dinero a usuarios jóvenes para comprar en cuotas, fundado en 2017.

Sin dudas, los bancos digitales son las startups que generaron más impacto. Cambiando sucursales físicas por aplicaciones a través de las cuales los usuarios pueden resolver todos sus trámites, y dónde pueden ver sus gastos y analizar sus hábitos de consumo. Este cambio de lo físico a lo digital no podría darse si los usuarios no se sintiesen cómodos con entidades completamente virtuales, y habituados al uso de servicios online.

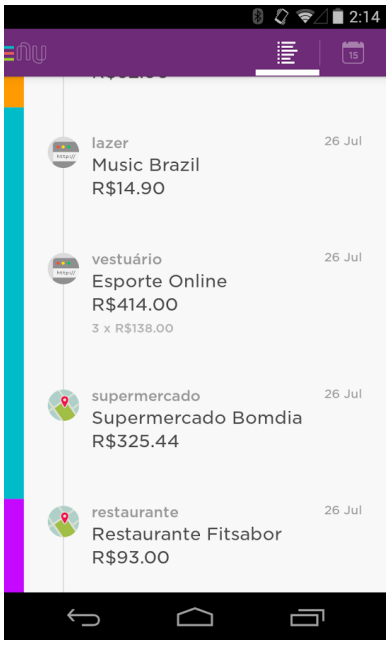

Figura 3: La app de Nubank muestra los gastos recientes

El acceso masivo a dispositivos móviles también ha permitido que entrantes acerquen nuevas propuestas de valor a los usuarios. Las fintechs han empujado la demanda de microcréditos personales, y desde el 2016 son las responsables por emitir el mayor volumen de préstamos[11].

Al contar con aplicaciones, las fintech tienen una ventaja frente a los bancos y casas de préstamos tradicionales: están disponibles las 24 horas y en cualquier lugar. Los usuarios no tienen que salir de sus casas, ni hacer colas, o interactuar con vendedores. Al ser nativos digitales, las nuevas generaciones se sienten más cómodas con préstamos online: según una encuesta realizada por Experian, un 25% de los participantes entre 18 y 44 años indicaron haber realizado un préstamo online, contra un 14% de los mayores de 44 años[12].

[11] *Personal Loan Market Overview*, Transunion, 2019 https://www.transunion.com/resources/transunion/doc/insights/articles/tu-personal-loan-market-2019.pdf

[12] *Survey: Consumers Want Personal Loans for Large Purchases and Debt Consolidation*, Experian, 2019 https://www.experian.com/blogs/ask-experian/survey-consumers-want-personal-loans-for-large-purchases-and-debt-consolidation/, recuperado 13/10/20

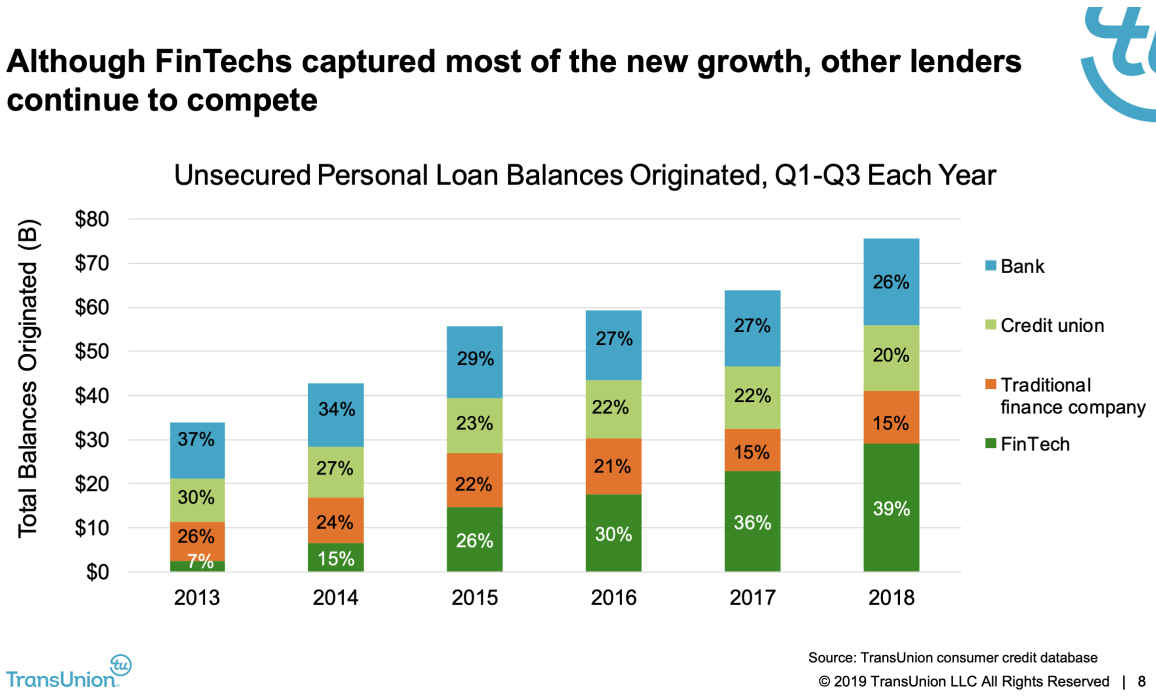

Figura 4: Las fintech lideran el mercado de préstamos personales

Al mismo tiempo, las fintechs pueden apalancarse en nuevos modelos de scoring: es decir, cómo calcular el riesgo financiero de un usuario. Las empresas tradicionales acuden a bases de datos de deudas (como ser el Veraz en Argentina), a históricos de deuda con el prestamista, o a comprobantes (como ser recibos de sueldo, pagos o títulos de bienes). Los bancos, adicionalmente, pueden acceder a datos de los productos que consumen sus clientes (como ser consumos en tarjetas de crédito, servicios adheridos, o incluso los salarios que percibe). Las fintech, al realizar la operación desde una app, cuentan además con los datos digitales.

Este acceso a los datos del dispositivo permite estimar con mayor precisión la probabilidad de default, es decir, el riesgo de no pago de un usuario, a partir de datos como llamadas, mensajes de texto, contactos o aplicaciones instaladas. Un estudio demostró que utilizando estos datos, es posible reducir el atraso en pagos un 200% y aceptar 250% más clientes, cuando se comparaba el método desarrollado con métodos tradicionales de scoring[13].

[13] Ruiz, Saulo & Gomes, Pedro & Rodrigues, Luís & Gama, João, *Credit Scoring in Microfinance Using Non-traditional Data*, 2017, 447-458

FacilPay, la empresa que será estudiada en este documento, toma el celular del cliente como colateral a la hora del préstamo. Para reducir el riesgo de no pago, utiliza una aplicación instalada en el dispositivo que, en caso de atraso de pago, bloquea el celular e impide su uso para cualquier otra función que no sea la de llamar a servicios de emergencias, soporte o pagar la deuda. De esta forma, es capaz de ofrecer préstamos sin penalizaciones por atraso, mejorando aún más la probabilidad de repago.

Las demandas de los usuarios digitales también representan un desafío para las fintechs. Estos clientes esperan que sus demandas sean satisfechas mucho más rápido. Mientras que prestamistas tradicionales tardan días o semanas en disponibilizar el dinero (ciertos prestamistas ofrecen tiempos desde 48 a 72 horas después del análisis de aprobación[14] [15]), las empresas digitales deben entregarlos en 24 horas (y en algunos casos, en el día[16]). Esto representa un desafío para las fintech, dado que debido a regulaciones desactualizadas, deben acoplar sus procesos automáticos a procesos manuales de ciertos proveedores necesarios para cumplir con las leyes.

Por otro lado, estas startups deben mantener una excelente calidad de servicio: cualquier problema que los clientes pudieren tener lo reflejarán en los comentarios de sus aplicaciones, que será lo primero que verán otros potenciales usuarios. Las empresas digitales pueden operar con costos mucho menores a empresas tradicionales que requieren de locales y su clientela depende de la cercanía de los mismos.

Pero al estar online, las fintech se exponen a nuevos tipos de riesgos, o mayor cantidad de los mismos. La misma comodidad que permite que el usuario realice un pedido sin moverse de su hogar, da un anonimato adicional a aquellas personas decididas a cometer fraude.

[14] *EMPRÉSTIMO CONSIGNADO: QUANTO TEMPO DEMORA PARA CAIR NA CONTA?*, Credito Folha, 2019 https://creditofolha.com/emprestimo-consignado-quanto-tempo-demora-para-cair-na-conta/, recuperado 13/10/20

[15] *Em quanto tempo ocorre a liberação do Empréstimo Consignado?*, QualiConsig, 2018 https://qualiconsig.com.br/emprestimo-consignado-demora-pra-cair-na-conta/, recuperado 13/10/20

[16] *Perguntas frequentes*, Simplic https://www.simplic.com.br/faq, recuperado 13/10/20

En el caso particular del fraude de identidad, la posibilidad de realizar todo el pedido complementamente en línea significa que el defraudador no debe exponerse en ningún momento. El criminal puede realizar el pedido del préstamo utilizando una identidad falsa, con un dispositivo robado, desde una conexión pública y de esa manera proteger su identidad real, enviando dinero a cuentas que cierra después de retirar el efectivo. Es decir, no hay ninguna repercusión para los defraudadores, aun cuando se logra identificarlos. Y dado que el dinero es transferido a cuentas bancarias y el defraudador lo puede extraer, hay un incentivo adicional para realizar fraude.

Brasil es un país con un número elevado de casos de fraude: En el 2018 se detectaron más de 1.8 millones de casos de fraude de identidad, según un estudio de *Serasa Experian*[17], una empresa brasilera dedicada a brindar servicios en base a los datos que centraliza de tiendas, bancos y financieras. Más de 400 mil de estos intentos de fraude se dieron en bancos y financieras.

Para realizar fraude, los individuos roban documentos o roban los datos del cliente mediante sitios falsos o con bases de datos que son expuestas[18]. Luego utilizan programas de edición de imagen para reemplazar la foto de la persona por la del defraudador, y con el nuevo documento obtienen distintos productos a nombre del individuo.

Estado del arte y del conocimiento

La existencia de bases de datos que almacenan datos biométricos de los ciudadanos es una práctica cada vez más común en muchos países[19]. Estas bases de datos asocian Identificadores Nacionales (como ser el DNI) y características físicas tales como huellas dactilares, irises, impresiones de palmas, fotos, patrones de voz, de caminar (*Gait*) y ADN. Esta información permite identificar de diversas maneras a un individuo a través de varios sistemas.

---

[17] *A cada 16 segundos, uma tentativa de fraude acontece no Brasil, revela Serasa*, Serasa, 2018 https://www.serasaexperian.com.br/sala-de-imprensa/a-cada-16-segundos-uma-tentativa-de-fraude-acontece-no-brasil-revela-serasa, recuperado 13/10/20

[18] *Saiba o que é fraude e quais os tipos mais comuns*, SerasaConsumidor, https://www.serasaconsumidor.com.br/ensina/seu-cpf-protegido/o-que-e-fraude/, recuperado 13/10/20

[19] *Mandatory National IDs and Biometric Databases*, Electronic Frontier Foundation https://www.eff.org/issues/national-ids, recuperado 13/10/20

El almacenamiento y utilización de datos biométricos ha sido cuestionado en numerosas oportunidades. En 2011, 80 organizaciones e individuos de 27 países firmaron un petitorio destinado al Consejo Europeo para que se investigue a fondo la recolección y almacenamiento de los datos biométricos por parte de los estados miembros[20]. En 2014, una nueva ley que buscaba facilitar las capacidades de recolección de datos encontró amplia resistencia al entrar en debate[21]. La misma permitía almacenar huellas dactilares, irises y rostros de personas entrando y saliendo de Australia, así como también cambiar los requerimientos para recolectar más datos, de modo que su aprobación fuese más simple. En 2016, un órgano de control estatal francés pidió la suspensión de una base de datos que contenía información de 60 millones de personas, bajo la sospecha de que podía ser abusada por organismos de inteligencia[22]. E incluso recientemente, en 2020, continúan las preocupaciones por la utilización de datos biométricos, incluso en países que ya las han implementado ampliamente[23].

En Brasil existen múltiples bases de datos que almacenan datos relacionados a los ciudadanos. En ellas se distinguen las no biométricas y las biométricas. Entre las no biométricas se encuentran las bases de la *Receita Federal* (que almacena datos que están disponibles en el documento como ser fecha de nacimiento, nombre de los padres y sexo) y la *Policía Federal* (que almacena antecedentes criminales) y dentro de las biométricas se encuentra la del *Serviço Federal de Processamento de Dados* (o SERPRO, que contiene las imágenes de los individuos con CNH, la *Carteira Nacional de Habilitação* o permiso para conducir).

Además de las bases de datos provistas por el estado nacional, existen empresas que proveen el servicio de interface con dichas bases y en muchos casos enriquecen los datos estatales con información privada propia o de terceros. Entre la información que agregan generalmente se

[20] *Alliance appeals to Council of Europe to address biometrics privacy*, Privacy International, 2011 https://privacyinternational.org/blog/1580/alliance-appeals-council-europe-address-biometrics-privacy, recuperado 13/10/20

[21] Opposition grows to storage of photo and biometric data, The Sydney Morning Herald, 2014 https://www.smh.com.au/politics/federal/opposition-grows-to-storage-of-photo-and-biometric-data-20141015-116lur.html, recuperado 13/10/20

[22] *French privacy row over mass ID database*, BBC, 2016 https://www.bbc.com/news/37894968, recuperado 13/10/20

[23] *Despite public concerns, facial recognition gets traction in Congress*, Federal Computer Week, 2020 https://fcw.com/articles/2020/02/09/congress-facial-recognition.aspx, recuperado 13/10/20

encuentran deudas con distintos entes (como ser supermercados, bancos y prestamistas), estimaciones salariales y direcciones conocidas cuando se trata de información no biométrica. Entre estos proveedores se encuentran IdWall y BigDataCorp. También existen bases de datos biométricas, más precisamente enfocadas en almacenamiento de rostros, que cumplen el mismo propósito. Entre ellas se encuentran Certibio y Acesso Digital, que a diferencia de otras que funcionan como interfaz con una base gubernamental, cuenta con rostros almacenados por distintos entes privados y es capaz de decir si un rostro ya se encuentra bajo un cierto identificador, permitiendo identificar si un mismo individuo está intentando utilizar varios documentos.

Para poder cumplir con las leyes de protección de datos personales en Brasil, estas bases de datos sólo pueden retornar la probabilidad de que la foto provista corresponda al usuario que se está consultando. Es decir, los clientes de dichos servicios no tienen acceso a las fotos de los individuos.

Marco Teórico

Bases de la biometría facial

Los sistemas de biometría facial utilizan rasgos distintivos de los rostros que se identifican mediante puntos elevados y hundidos. Estos puntos son llamados nodos y cada rostro tiene alrededor de 80 de ellos[24]. A partir de estos puntos se calculan ciertas características del rostro como ser:

- Distancia entre ojos
- Ancho de la nariz
- Profundidad de los ojos
- Forma de los pómulos
- Longitud de la mandíbula

Estos valores son codificados y guardados en la base de datos, representando la cara. Además de almacenar esta información, los sistemas de reconocimiento facial pueden requerir varias

[24] *How Facial Recognition Systems Work*, How Stuff Works https://electronics.howstuffworks.com/gadgets/high-tech-gadgets/facial-recognition.htm/printable, recuperado 13/10/20

imágenes de una persona para tomar un perfil tridimensional de la misma. En este caso se utilizan ciertos puntos rígidos del rostro que son únicos y no cambian con el tiempo, como ser la curvatura de los ojos, la nariz y el mentón.

Estos análisis en tres dimensiones también permiten crear sistemas de prueba de vida: programas que a través de un proceso de recolección de imágenes permiten asegurar que una persona es real (y no se trata de un muñeco o una foto). Entre las empresas que proveen este servicio se encuentra ZoOm 3D, un servicio que provee identificación biométrica: su sistema provee un análisis de las imágenes, generando un mapa en tres dimensiones del rostro a partir de las imágenes, permitiendo validar que se trata de una persona real[25].

Una vez que se obtienen los valores de las características, se comparan contra los valores almacenados en las bases de datos. En casos de comparaciones 1:1 se toma la foto del individuo contra el que se está comparando, se calculan los valores de aquel y se retorna la comparación contra el individuo analizado. En casos de comparaciones 1:N, se busca por los valores en la base de datos, y se retornan todas las ocurrencias de fotos que podrían corresponder a la misma persona.

Usos de bases biométricas 1:1 y bases 1:N

En los casos en los que es posible, es deseable verificar que una foto corresponde con certeza a un individuo. Las bases biométricas estatales dan esa seguridad, pudiendo ser consultadas por el identificador nacional (en el caso de Brasil, llamado CPF) junto a una imagen del individuo que se desea validar. En estos casos es posible confiar en el resultado de dicha comparación: es decir, el ente que provee el dato puede asegurar con certeza que la foto almacenada corresponde al individuo.

En el caso de bases biométricas privadas, no existen datos del individuo hasta que se realiza la primera consulta. Estos sistemas suelen registrar las fotos de los individuos con cada consulta, almacenándolas y retornando un valor dependiendo de si el individuo se asemeja o no a las fotos almacenadas para esa persona. Dado que la validación se basa en un supuesto de que esa persona

[25] *Meet Zoom*, Zoom https://www.zoomlogin.com/#page-blk-meet-zoom, recuperado 13/10/20

corresponde al identificador contra el que se lo está comparando, estos sistemas retornan también un valor representando cuántas veces apareció este individuo en distintos clientes. Estas bases suelen realizar una comparación 1:N entre los individuos almacenados para asegurarse que cuando se da de alta una nueva persona, no se trata de una persona intentando realizar fraude de identidad. Además, dado que suelen ser bases de datos de diversos grupos privados, algunas permiten que los miembros aporten anotaciones (como ser declarar que un individuo no pagó una deuda).

Dilemas éticos en la utilización de sistemas biométricos

La utilización de sistemas biométricos para identificación y toma de decisiones en base a características físicas de los individuos debe ser sometido a análisis en base a distintos enfoques.

Un primer enfoque es la privacidad: Las fotos de los usuarios son considerados datos sensibles de acuerdo a diversos sistemas de protección de datos personales[26]. Dado que estos datos se utilizan para validación de individuos, no pueden ser anonimizados. Eso significa un alto riesgo de seguridad en el almacenamiento y procesamiento de estos datos.

También se deben considerar los resultados que se obtienen a partir de relacionar datos biométricos con comportamientos. La utilización de rasgos como ser color de piel o rasgos faciales para la toma de decisiones podría generar un bias discriminatorio que es reforzado por el mismo sistema. Por ejemplo, el limitar a un individuo el acceso a productos financieros dado el historial de personas con rasgos físicos similares podría generarle desventajas a todos los individuos con esos rasgos, independientemente si el resultado particular pudiera ser distinto de la media.

El último enfoque es el legal, dónde se debe considerar la validez de las decisiones tomadas a partir de los resultados de la utilización de sistemas biométricos. La posibilidad de falla del sistema pone en cuestión la legalidad de la asociación entre un individuo y el resultado del sistema[27], y las consecuencias que pudieren surgir de la asociación entre quien utiliza el sistema y quien dice ser.

---

[26] *What is GDPR?*, idStation https://www.idstation.eu/Home/GDPR, recuperado 13/10/20

[27] Pato, Joseph N. & Millett, Lynette I., *Biometric Recognition: Challenges and Opportunities*, NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, THE NATIONAL ACADEMIES PRESS 2010 96-98

Ventajas de la utilización de sistemas biométricos

En el caso de estudio (una fintech dedicada a los micro-créditos personales), la utilización de sistemas de biometría facial presenta diversas ventajas. En primer lugar, al utilizar sistemas capaces de validar con seguridad la identidad de un cliente, se obtienen clientes de mejor calidad. Es decir, clientes que con seguridad son quienes dicen ser.

Esto afecta directamente la reducción de préstamos que no son repagados, en particular aquellos que no reciben ningún pago antes de ser enviados a pérdida denominados First Payment Default (FPD). Esto se debe a que los usuarios que realizan fraude de identidad para obtener un préstamo no realizan ningún pago. Estos préstamos generan pérdida adicional ya que para realizar los préstamos, la empresa toma capital prestado a una tasa menor y luego debe cubrir los intereses del periodo desde que se otorgó el préstamo.

Al tener menores tasas de default, se aumenta la contribución marginal de cada préstamo. Esto permite que la empresa reduzca la tasa de interés para todos los préstamos, manteniendo la misma ganancia por préstamo.

Una menor tasa de interés, a su vez, resultará en más clientes aceptando las condiciones del préstamo. De la misma forma, clientes que podrían obtener mejores tasas de interés en otras entidades financieras con más información para su proceso de scoring (es decir la determinación del riesgo del cliente), se verían más dispuestos a aceptar la nueva tasa más baja. Estas dos variables resultan en una mayor cantidad de préstamos y también en préstamos de mayor calidad, ya que los usuarios que los aceptan tienen otras alternativas similares.

Por último, la utilización de un sistema de identificación biométrica significa que individuos (tanto aquellos que utilizan el sistema como aquellos que no) están mucho menos expuestos a que se utilice su identidad fraudulentamente, evitando quedar negativizados (es decir, declarados deudores) por un producto financiero que no adquirieron.

El mercado mobile

Una de las diferencias entre casas financieras tradicionales y fintech es que mientras las primeras dependen de locales para operar, las segundas funcionan completamente online muchas veces. Si la decisión de un cliente de una financiera tradicional de utilizar sus productos o no se podría formar al visitar la tienda, los clientes de una fintech muchas veces ven sus decisiones formadas por las opiniones de otros usuarios y la experiencia que tienen al utilizar las aplicaciones.

Si bien no hay locaciones físicas, a la hora de competir las fintech compiten por el posicionamiento en palabras clave, como ser "préstamos". A la hora de buscar servicios, los potenciales clientes utilizarán estas palabras clave en buscadores y en la tienda de aplicaciones para investigar ofertas, eligiendo la que más los convenza. Estas posiciones no están directamente relacionadas al gasto en publicidad de una empresa. Si bien puede invertirse más dinero para adquirir más instalaciones o aparecer en los motores de búsqueda, las posiciones están relacionadas a distintas variables. Entre los factores que influyen se encuentran algunos dentro del control de quien publica la app (como ser nombre, descripción y frecuencia de actualizaciones), y otros fuera del control que refieren a la reputación (como ser links a la app, reviews y tiempo de utilización por los usuarios)[28].

Es decir, para poder distinguirse por encima de otras aplicaciones y atrapar la atención de un usuario una fintech debe tener una aplicación relevante, con buenas calificaciones por parte de los usuarios, con referencias a ella en medios populares y que los usuarios utilizan a menudo.

CPI y CAC

Para medir la contribución marginal de un producto, no basta con sólo calcular los costos involucrados en la utilización de servicios (por ejemplo, el costo de emisión de un contrato), sino que se debe calcular el costo de alcanzar ese usuario. CPI (cost per install, o costo por install) refiere al costo de lograr que un usuario descargue la aplicación. Cuando se contrata una campaña por CPI, se establece un valor que se le pagará al publicista por cada usuario que haga click en la publicidad y cuyo click se convierta en una descarga. En estas campañas, el anunciante establece

[28] *App Ranking Factors » How to Improve App Store Search Rankings*, AppRadar, 2019 https://appradar.com/academy/bonus-chapters/app-store-ranking-factors/, recuperado 13/10/20

un tope que está dispuesto a pagar por instalación y los publicistas eligen qué anuncios mostrar (generalmente se muestran aquellos que más pagan primero).

El Costo de Adquisición (Customer Acquisition Cost o CAC) es el costo que surge de lograr que un usuario complete todo el proceso desde la descarga hasta realizar la acción deseada (en este caso, adquirir un préstamo). Este costo se obtiene de sumar todas las tarifas a lo largo del proceso y también debe considerar los gastos realizados en usuarios que no completaron el proceso. Por ejemplo, si por cada 10 usuarios que instalan la aplicación sólo 1 termina el proceso, si el CPI es 1 USD, el CAC tendrá un costo de 10 dólares representando las 9 instalaciones no finalizadas.

Dependiendo del tipo de aplicación, el grado de utilización y permanencia del usuario (denominado *engagement*) resulta de vital importancia. Los clientes que se encuentran satisfechos con la aplicación y el servicio a menudo lo recomiendan a otros: estas instalaciones no pagas (referidos) reducen el CAC al traer nuevos usuarios sin costo adicional.

También, los usuarios más comprometidos suelen dejar calificaciones positivas en las tiendas de los distintos sistemas operativos (*Google PlayStore* y *Apple AppStore*). El rendimiento de estas métricas (calificaciones y engagement) es utilizado por las tiendas a la hora de recomendar una aplicación a un usuario cuando realiza una búsqueda, posicionándola antes de otras que podrían ser competidoras. Los usuarios adquiridos de esta forma son llamados *orgánicos* y reducen aún más el CAC al no provenir de campañas de marketing.

Publicidad online y publicidad offline

La publicidad online se refiere a las campañas publicitarias iniciadas en medios digitales. A diferencia de las publicidades tradicionales, las publicidades digitales presentan una gran facilidad y precisión a la hora de medir el impacto: es muy fácil observar qué segmentos interactúan con ellas y optimizar para conversiones. La atribución en medios físicos es difícil de calcular y a menudo no puede relacionarse a una única campaña.

Por otro lado, la publicidad tradicional tiene ciertas ventajas. Los anuncios radiales y folletos pueden ser enfocados en ciertas áreas que son de importancia para el negocio. Adicionalmente, un estudio de 2015 descubrió que las personas tenían mejor retención de las marcas de las cuales veían un anuncio en papel que uno digital[29].

Es decir, lo importante es distinguir el objetivo de la campaña publicitaria. Las startups utilizarán mayormente anuncios digitales para poder optimizar sus costos, validar el negocio y obtener contribuciones marginales positivas, pudiendo entender los gastos de cada paso del proceso. Las empresas maduras que intentan crear conocimiento de marca o dar a conocer un nuevo producto lo realizan a través de anuncios en la vía pública.

Por último, las publicidades digitales tienen costos de ingreso muy bajos y permiten direccionar sus públicos objetivos a partir de las mismas cualidades que los anuncios tradicionales (ubicación, horario de interés) y a través de propiedades más complejas (como gustos, interés en temas, dispositivos electrónicos, edad y sexo).

Utilizando herramientas de atribución, es posible entender en qué momento un usuario tomó la decisión de realizar la instalación. De esta forma, se pueden optimizar estos canales de adquisición, dirigiendo los esfuerzos no sólo a los medios que más conversiones generen, sino a los que generen una mejor calidad de clientes (por ejemplo, clientes recurrentes).

[29] Paper Beats Digital In Many Ways, According To Neuroscience, Forbes, 2015 https://www.forbes.com/sites/rogerdooley/2015/09/16/paper-vs-digital/, recuperado 13/10/20

Resumen Ejecutivo

En este documento se describe la implementación de un método de validación biométrica para los clientes de *FacilPay*, una startup en el sector fintech de Brasil. La misma se dedica a entregar préstamos a personas de sectores C, D y E, negativizadas[30] o no.

El modelo de negocio de la empresa es permitir que el usuario descargue una aplicación mediante su celular y realice un pedido de un microcrédito de manera totalmente virtual (hasta 2.000 reales, alrededor de 500 USD, a devolver en hasta 12 meses). Debido a la facilidad para aplicar y el anonimato que provee el uso de una aplicación de manera remota, se detectaron un número elevado de casos de fraude de identidad. Esta hipótesis fue elaborada a partir de observar un número elevado de préstamos enviados a pérdida sin ningún pago, y fue confirmada al realizar pruebas con potenciales proveedores de servicios de validación de identidad, con los datos ya obtenidos de los usuarios (documentos de identidad y fotos de sus rostros).

Este diagnóstico es explicado de manera extensiva, cubriendo la metodología, las limitaciones, el funnel de conversión de la aplicación y el resultado de dichas pruebas. También se describen los proveedores de validación biométrica utilizados, haciendo la distinción entre proveedores que utilizan bases de datos gubernamentales y proveedores que dependen de los datos aportados por sus clientes para intentar definir la validez de la identidad de un individuo. Adicionalmente se hace referencia a las restricciones de cada uno, y los parámetros de filtro establecidos cuando es imposible determinar con seguridad la identidad del cliente.

A continuación se detalla la propuesta de valor y el plan de implementación para incorporar estas mejoras al proceso de negocio de la empresa, de forma de obtener el objetivo deseado reduciendo el impacto negativo que podría tener en los usuarios, observado en la dificultad de los mismos para completar el proceso de solicitud. Se incluye el cronograma de trabajo de todas las áreas involucradas.

[30] Una persona negativizada es aquella que, por deudas previas, se encuentra en un registro de deudores. Usualmente no pueden acceder a otros servicios financieros.

La estimación en costos y las proyecciones se realizan a partir del funcionamiento histórico de la startup. Si bien se calculan los flujos de fondos resultantes del proyecto, no se cuantifica la percepción de una cartera de créditos más estable y con menor fraude frente a posibles inversores.

Luego se detalla el resultado de la implementación, la reducción en préstamos enviados a pérdida y la performance de cada uno de los proveedores utilizados, realizando una consideración de aquellos clientes que fueron rechazados por el sistema.

Finalmente se detalla el impacto de las implementaciones en el bottom line de la empresa, a través de los cambios producidos en la contribución marginal de un préstamo promedio: el aumento en el costo de validación, las ganancias adicionales por mayor cantidad y montos de pagos, y los incrementos en costos de adquisición a partir del aumento en el número rechazado.

Se ofrecen también sugerencias sobre posibles mejoras al sistema, y observaciones sobre los aciertos y fallas de la implementación del mismo.

Diagnóstico

FacilPay inicialmente prestaba servicios de financiación de compra de celulares. Para eso, los clientes debían asistir a una tienda física de una empresa con la que se había realizado un convenio: Los vendedores de esa tienda eran entrenados, y realizaban el proceso de alta del cliente a través de un portal de la empresa cuando este solicitaba el beneficio. Dado que era un sistema offline, se debía capacitar a los vendedores de cada sucursal. Ese proceso exigía muchos gastos y horas hombre, además de tener un alcance limitado, por lo que se modificó el sistema para realizar préstamos en línea.

Mediante el nuevo sistema, los clientes descargan una aplicación a sus teléfonos celulares y realizan el proceso de alta ellos mismos. Durante el proceso, el cliente debe proveer sus datos, prueba de identidad y una cuenta destino para los fondos. Al terminarlo, el pedido es revisado y los fondos liberados.

Poco después del lanzamiento de la nueva modalidad, se observó un porcentaje inusualmente alto de préstamos en los que el individuo no realizaba ningún pago, comparado con el promedio de los valores de mercado para los mismos segmentos de consumidores. Debido a esto, se decidió detener las campañas de adquisición de usuarios y realizar un análisis para determinar si existía fraude de identidad.

Definición y alcance del problema

En este trabajo se pretende estudiar el impacto de la implementación de un método biométrico de validación de identidad. Se distinguen dos tipos de fraude de identidad: el fraude debido al robo de documentos y el auto-fraude, es decir una persona que utiliza sus propios datos para adquirir un producto financiero sin intención de cumplir con las obligaciones. Ambos tipos de fraude se ven reflejados en la métrica de First Payment Default (FPD, o Default en Primer Pago, es decir aquellos préstamos que no devolvieron ningún monto antes del tiempo determinado para enviar el préstamo a pérdida o *write-off*).

Cuando un préstamo no es devuelto, es decir entra en default, el costo de la pérdida (es decir, los intereses pagados por el monto apalancado y el total del préstamo) es prorrateado entre todos los préstamos. Es decir, es importante para una fintech dedicada a préstamos mantener una tasa de default baja, ya que afecta la contribución marginal de cada préstamo.

Cuando un usuario no cumple con una obligación financiera, su CPF (Cadastro de Pessoas Físicas, o Registro de Personas Físicas en Brasil, el equivalente al DNI argentino) es agregado a un registro de deudores. La inclusión en ese registro inhabilita al usuario a adquirir ciertos productos financieros. Cuando un usuario está en dicho registro, se dice que fue negativizado. El robo de identidad genera pérdidas dado que quien está realizando el fraude no se ve disuadido por la penalidad de ser negativizado.

Se esperaba que al implementar el sistema de validación biométrica se eliminasen los casos de fraude de identidad. Se consideraba que al reducir los casos de fraude de identidad, la métrica de First Payment Default debería disminuir. Se consideraba que existían, al momento previo de la implementación, un cierto número de casos de fraude por robo de identidad. No se considerarán los casos en los que el usuario está dispuesto a ser negativizado (es decir, aquellos casos donde el usuario utiliza su verdadera identidad para adquirir un préstamo sin intención de pagarlo). El alcance de este trabajo es evaluar el impacto de la implementación del nuevo sistema en la métrica mencionada.

Limitaciones

El sistema de verificación biométrica utiliza una base de datos faciales del registro nacional de conductores de Brasil, o CNH (Carteira Nacional de Habilitação). Se asume que los datos en las bases biométricas consultadas son verídicos. Existe la posibilidad de que un criminal elabore un tipo de fraude de robo de identidad complejo: mediante un certificado de nacimiento falsificado, puede presentarse ante un ente público y apropiarse de la identidad de una persona, a partir de la cual podría registrar su rostro en la base de datos mencionada anteriormente. Escapa al proyecto la consideración de estos casos.

Asimismo, se consideran en esta base las personas que tienen una habilitación para conducir vehículos. Aquellas personas que no se encuentren en el registro serán rechazadas. Este rechazo reduce la cantidad de personas que pueden acceder a un préstamo (es decir, baja la tasa de aprobación) pero a cambio da mayor seguridad sobre los préstamos emitidos.

A raíz de dicha limitación se adicionó una segunda base de datos biométrica privada, con restricciones particulares del sistema. Se entiende que los datos no garantizan identidad con la misma rigurosidad que la base de datos estatal.

Durante el período de implementación estudiado, no hubo otros cambios que pudiesen afectar la tasa de aprobación o la calidad de los préstamos emitidos.

Relevancia

Brasil es un país con un número elevado de casos de fraude de identidad. Según un informe de la *Confederação Nacional de Dirigentes Lojistas (CNDL)* de Agosto de 2019, "más de 12 millones de consumidores sufrieron algún tipo de fraude en los últimos 12 meses (y) las pérdidas generadas ascienden a 1.8 billones de reales"[31].

Para obtener un documento falsificado, un defraudador puede pagar entre 40 y 1200 reales y puede obtener dichos documentos, con datos de personas verdaderas en lugares públicos como una plaza[32]. El sistema particular de la empresa para reducir los atrasos en las cuotas del préstamo también resulta poco efectivo en estos casos: los criminales a menudo utilizan teléfonos robados (que suelen sustraer junto con los documentos).

[31] *FRAUDES FINANCEIRAS NO BRASIL*, Confederação Nacional de Dirigentes Lojistas, 2019 http://www.cndl.org.br/upload/PP40/materiais/pesquisas/Fraudes%20Financeiras/1/SPC%20Analise%20Fraudes%20Financeiras%20no%20Brasil.pdf

[32] *Falsificação de documento representa 75% das fraudes registradas em MG*, Jornal Hoje, 2014 http://g1.globo.com/jornal-hoje/noticia/2014/02/falsificacao-de-documento-representa-75-das-fraudes-registradas-em-mg.html, recuperado 13/10/20

Antes de comenzar el análisis, se contaba con un porcentaje de FPD por encima de la media esperada para este tipo de negocios. Distintas conversaciones con competidores revelaron que la problemática analizada (el fraude de identidad) era muy común en el sector fintech, y en el segmento al que pertenecen los usuarios. Se llegó a la sospecha de que existía un alto riesgo de ser víctimas de fraude de identidad por parte de clientes maliciosos.

Se discutieron diversas alternativas como ser distintos sistemas de identificación biométrica, validación a través de teléfonos o domicilio físico y se eligió la de utilizar una selfie con prueba de vida como la que menor impacto en el funnel de conversión (la métrica que mide qué porcentaje de usuarios finaliza el proceso de obtención de préstamos) tendría.

Metodología

Para validar el supuesto de que la empresa era víctima de fraude de identidad, se contactaron distintos proveedores de biometría facial y se realizó una evaluación, utilizando las selfies de los usuarios y los identificadores nacionales de quienes decían ser. Los resultados fueron utilizados tanto para evaluar el rendimiento de cada proveedor como para validar la hipótesis de que los casos de fraude de identidad resultaban en préstamos no devueltos.

Para el estudio del impacto en el negocio se tomará el porcentaje de préstamos en First Payment Default (préstamos a pérdida sin capital devuelto) para las cohortes anteriores a la implementación y se compararán con las cohortes posteriores. Una cohorte es un grupo de préstamos emitidos en un período de tiempo (el período utilizado en este caso es 1 mes) y que tienen las mismas características.

Durante el período analizado no se realizaron otros cambios que pudieren afectar la calidad de los préstamos emitidos, ni se observaron factores macro económicos que pudieren afectar la capacidad de repago de los individuos.

Funnel de conversión

El funnel para obtener un préstamo está dividido en dos partes: un proceso de validación de identidad y un proceso de pedido del préstamo.

El proceso de validación consiste en los siguientes pasos:

- Registro: el usuario debe utilizar su CPF y correo electrónico.
- Datos personales: se le pide al usuario su nombre completo, nombre de la madre, y otros datos que están disponibles en el documento.
- Selfie: Se le pide al usuario una selfie (foto simple en un principio y prueba de vida luego de la implementación del sistema).
- Fotos del documento: Se le pide una foto del dorso y el reverso del documento de identidad del usuario.

Figura 5: Ejemplo de CPF falso

Una vez aprobado el proceso de identidad, el usuario debe completar la solicitud de préstamo que consiste en los siguientes pasos:

- Selección de monto y plazo: el usuario selecciona el monto del préstamo y el plazo a retornarlo.
- Datos de la cuenta: se pide la cuenta bancaria a la que se transferirá el dinero. La misma debe estar a nombre del usuario.
- Confirmación de los montos y valores del préstamo: se le exponen al usuario todos los datos relacionados como impuestos, tasas, cuota mensual y valor total. Si el valor del

préstamo aprobado para el usuario es menor al pedido por él, se le hace una contraoferta en este paso.

- Firma del contrato: el usuario firma digitalmente un contrato. Esta firma se realiza mediante un proveedor y, dado que depende de un número de teléfono asociado, tiene una validez baja[33].

Trade-off

La implementación del método de validación biométrica elegido trae ciertas desventajas. En primer lugar, dado que se está utilizando un nuevo servicio como paso intermedio para validar al usuario, se incrementa el costo de emisión del primer préstamo (ya que una vez que el usuario está validado, no se vuelve a pasar por este proceso).

El paso adicional en el proceso para obtener un préstamo también significa que el usuario debe realizar más acciones antes de obtenerlo. Generalmente, cada paso del funnel incrementa la posibilidad de que el usuario abandone el pedido. Para minimizar el impacto, se reemplazó el paso donde el usuario debía tomar su selfie por una prueba de vida (llamada "liveness test"): la misma consiste en pedir al usuario que acerque y aleje el dispositivo encuadrando su rostro. Durante la prueba de vida se toma uno de los cuadros como selfie para evitar un paso adicional. Si bien el proceso exige mejores condiciones de luz y no es tan rápido como una selfie, no se vieron diferencias en la caída en este paso (*churn*) para usuarios que buscaban obtener un préstamo. Aquellos que intentasen utilizar una fotografía no pasarían este paso, pero serían usuarios fraudulentos.

A la vez, para que un usuario pueda acceder a un préstamo este debe contar con un rostro para validar ya sea en el registro de conductores o en una base privada, ya aceptada con anterioridad por otros privados. Por esto, los usuarios que no tengan registros en ninguna de estas dos bases,

[33] La firma digital se realiza mediante *Clicksign* (https://www.clicksign.com/), un proveedor de firma digital de Brasil. A la hora de firmar el contrato, *Clicksign* le envía un código mediante un mensaje de texto al número asociado al cliente, que este luego ingresa en la aplicación para confirmar que leyó y está de acuerdo con el contrato. Este registro, y la copia del contrato firmado, quedan guardados.

serán rechazados, incrementando el costo de emisión del préstamo promedio aún más. Estos usuarios podrían ser legítimos.

También serán rechazados aquellos cuyos rostros no cumplan un umbral de similitud con la persona que dicen ser. A pesar de que esos casos también impactan en el costo de emisión (ya que significan más usuarios rechazados), reducen las pérdidas por préstamo y por lo tanto el impacto en el resultado es positivo.

La ventaja obtenida de utilizar una validación más restrictiva es que se cuenta con una cartera de usuarios validados que, en caso de falta de pago, puede asumirse que no tenía intención de devolver el dinero.

Individuos

El público objetivo de la empresa estudiada son las personas no bancarizadas. En Brasil, el 40% de la población, es decir más de 80 millones de personas, no tienen acceso a servicios financieros formales, según un estudio de *Itaú*[34]. Esta población pertenece a los segmentos C, D y E, es decir, los segmentos más vulnerables y con menor educación.

Esta población fue tradicionalmente atendida por financieras: prestamistas privados que entregan montos bajos de dinero a tasas elevadas. En los últimos años han aparecido fintechs dedicadas a servir estos segmentos que resultan en competidores directos de la empresa. La competencia en este segmento se da por precio, y los clientes suelen obtener préstamos a tasas más bajas para cancelar deudas más caras[35].

[34] *Itaú Unibanco: Banking the unbanked in Brazil*, Itaú, 2019 https://www.businesscalltoaction.org/wp-content/files_mf/bcta_casestudy_itau_web.pdf

[35] *40% das pessoas que pedem crédito consignado usam dinheiro para pagar dívidas, diz pesquisa*, Globo Economía, 2017 https://g1.globo.com/economia/seu-dinheiro/noticia/40-das-pessoas-que-pedem-credito-consignado-usam-dinheiro-para-pagar-dividas-diz-pesquisa.ghtml, recuperado 13/10/20

	CET (%a.m.) - 9 meses				
	R$500	R$1.000	R$1.500	R$2.000	R$2.500
flexipag	12,93%	11.70%	11.28%	11.07%	10.94%
noverde	11,51%	11,51%	11,51%	11,51%	11,51%
Simplic	22,11%	19.35%	16.49%	16.03%	15.76%
MoneyMan	18,74%	16.87%	16.24%	15.92%	N/A

Figura 6: Comparación de tasas con principales competidores

Sin embargo, debido a las restricciones en la tecnología de bloqueo utilizada, la empresa sólo puede atender a clientes que posean un smartphone Android fabricado por Samsung. El market share de Samsung en el mercado de Brasil es de 45%, muy por encima del segundo (Motorola, con 20%) según *Statista*[36].

[36] *Distribution of smartphone shipments share in Brazil in 1st quarter 2017 and 2018*, by brand, Statista, 2018 https://www.statista.com/statistics/693317/smartphone-market-share-vendor-in-brazil/

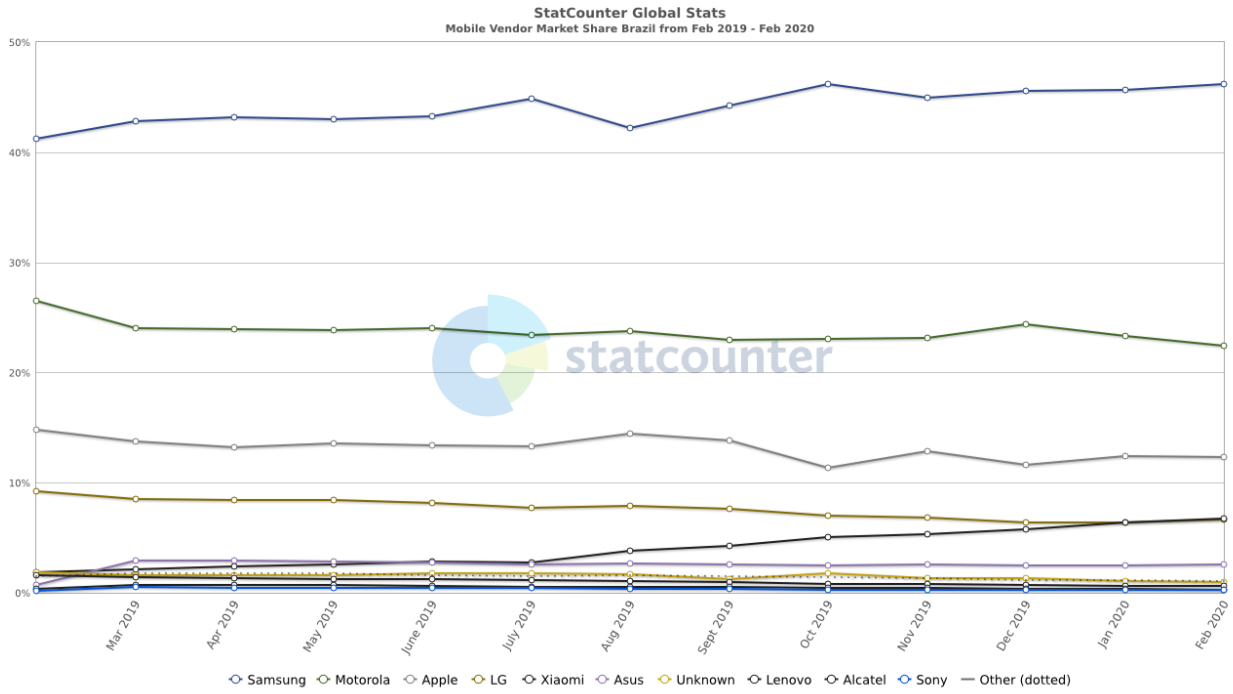

Figura 7: Distribución del Mercado de Brasil por Fabricante

Es decir, la empresa puede servir aproximadamente un total de 38 millones de personas en todo Brasil. Dado que actualmente la empresa sólo atiende a clientes en São Paulo, el público objetivo estimado es de 8 millones de personas.

Estos clientes a menudo no tienen un ingreso fijo y viven de la economía informal. Los jefes de familia en estos casos suelen tener trabajos poco calificados como repartidores, conductores y personal de call-centers. El uso de los préstamos tomados, cuando no se trata de cancelar otra deuda, suele ser para llegar al siguiente cobro, momento en el que generalmente devuelven la totalidad del préstamo.

Backtesting

Backtesting es el proceso de probar un sistema con datos antiguos, para intentar predecir qué calidad de resultados se obtendrían una vez que el sistema se encuentre implementado. Este paso es una prueba para evaluar el valor del producto. Para las bases de datos biométricas, la prueba del

sistema consiste en enviar las capturas de rostros junto con los CPF, y evaluar la cantidad de resultados con identidad validada, falsa y sin poder validar, comparando dichos resultados con la calificación del préstamo (pagado o con FPD).

Al momento de comenzar el estudio, el porcentaje de préstamos enviados a pérdida total era mayor al 20%. Se iniciaron las evaluaciones con proveedores en Octubre de 2018 y se decidió eliminar los gastos en publicidad para reducir la emisión de préstamos. Durante este período, hasta Marzo de 2019 se emitieron un promedio de 50 préstamos mensuales. A pesar de que el porcentaje de FPD se redujo al adquirir únicamente usuarios orgánicos, el porcentaje continúa siendo elevado, alrededor de un 18%, aunque la muestra se considera no representativa.

El costo de validación de un usuario al momento de comenzar era nulo. La única comparación que se hacía era una validación entre el rostro del usuario y el rostro del documento de identidad. Para realizar esta acción se utilizaban los servicios de Microsoft Azure, en su gama gratuita.

Se contactaron distintos proveedores de biometría facial, y se evaluaron los resultados. Se eligió *Certibio* como proveedor por la capacidad de comparar contra una base de datos oficial. Se enviaron 464 fotos para evaluar junto a su identificador de CPF, de las cuales sólo 135 contenían un registro en la base de SERPRO, que almacena las fotos del registro nacional de conductores, o CNH. De estos registros, 4 fueron identificados como individuos falsos. Es importante destacar que, dado que al tomar estas selfies no había una prueba de vida, existía la posibilidad de que el usuario hubiese enviado una foto de foto de la persona.

Se analizaron las fotos que pasaron con éxito la validación y se encontraron 3 casos que pertenecían a fotos de redes sociales, y otros 3 casos que eran sospechados de haber tomado una foto de una foto, pero no había certeza.

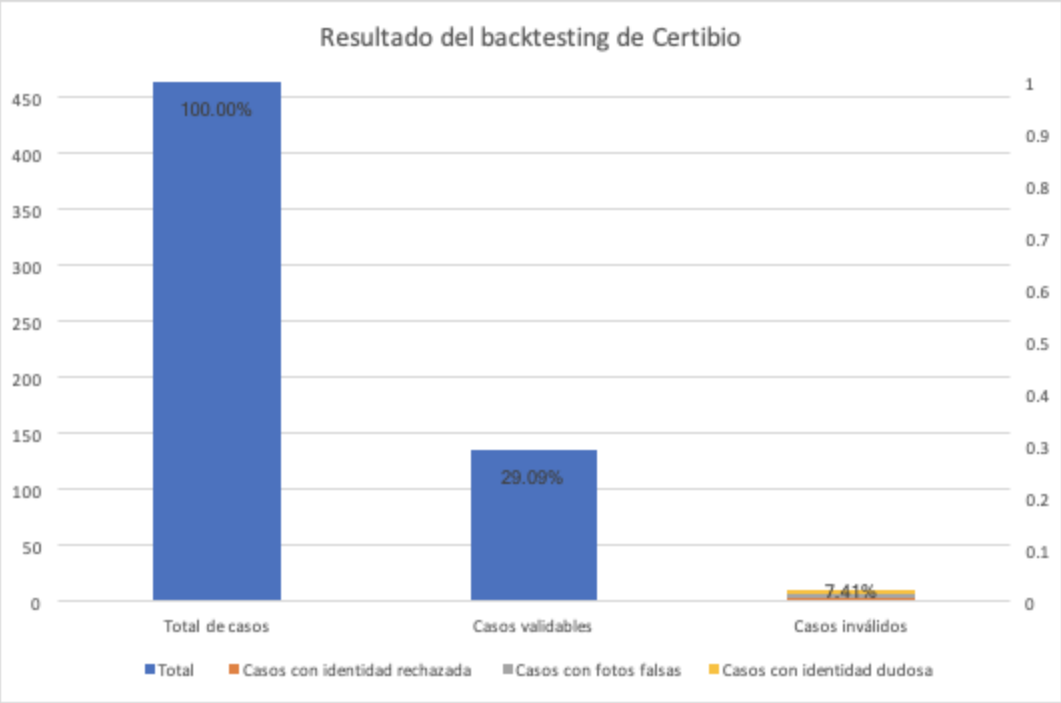

Figura 8: Resultado del backtesting de Certibio

En otras palabras, la implementación de la solución de Certibio suponía una reducción de alrededor del 70% en préstamos emitidos, obteniendo a cambio una mejora de entre 5% y 7% en la reducción de préstamos que serían declarados en pérdida.

Es decir, dado que se rechazaría cualquier usuario que no tuviese un registro en la base del registro nacional de conductores, el costo de aprobar a un usuario sería mayor. Se llegó a un acuerdo donde sólo se cobraría un valor de 0,2 reales por validación, sólo en los casos que existiese una foto en la base de datos asociada a ese identificador. Por lo tanto, sólo se incrementaría el costo de este paso por los usuarios inválidos, es decir, un 7% en el peor caso.

A pesar de este acuerdo, dado que se rechazan más préstamos se incrementa el costo por adquisición (CAC). Para reducir esto, es posible ajustar las campañas de adquisición de usuarios a personas que puedan tener más probabilidades de tener un CNH, aumentando el costo por instalación (CPI) pero reduciendo el CAC frente a la alternativa de no cambiar la campaña.

Dado que el público objetivo corresponde a los segmentos C, D y E, existía la posibilidad de que gran parte de los clientes no contase con un vehículo y por lo tanto no tuviese necesidad de un registro de conducir. Esta dificultad fue confirmada con el elevado número de casos que no pudieron ser validados por el proveedor biométrico- 329 casos o más de un 70%.

Para reducir la tasa de rechazo, se implementó un segundo proveedor biométrico de bases privadas con validación 1:1 y 1:N. Dado que este proveedor no tiene acceso a fotos validadas por el estado, se le dio menor prioridad y se establecieron reglas más estrictas. Las restricciones aplicadas fueron:

- El rostro de la persona no podía pertenecer a ningún otro identificador.
- El rostro debía coincidir con el identificador especificado.
- El identificador especificado debía tener registros en la base de datos de al menos 6 meses.
- El identificador debía haber estado presente en al menos 3 instancias anteriores y no poseer declaraciones de fraude.

Se contactó a *Acesso* con 485 fotos para validar la identidad de los usuarios (dada la diferencia de tiempo se contaba con más casos para validar). Del total de casos enviados, un total de 327 tenían algún tipo de registro en la base de datos. Es importante mencionar que la base de datos del proveedor no asegura la validez de la identidad de una persona, sino que retorna un valor dependiendo de cuántas veces fue registrado esa persona con ese CPF, si esa persona aparece vinculada a otro CPF y si fue denunciado. Del total de casos con identidad, 60 hubiesen sido rechazados por las políticas de aprobación.

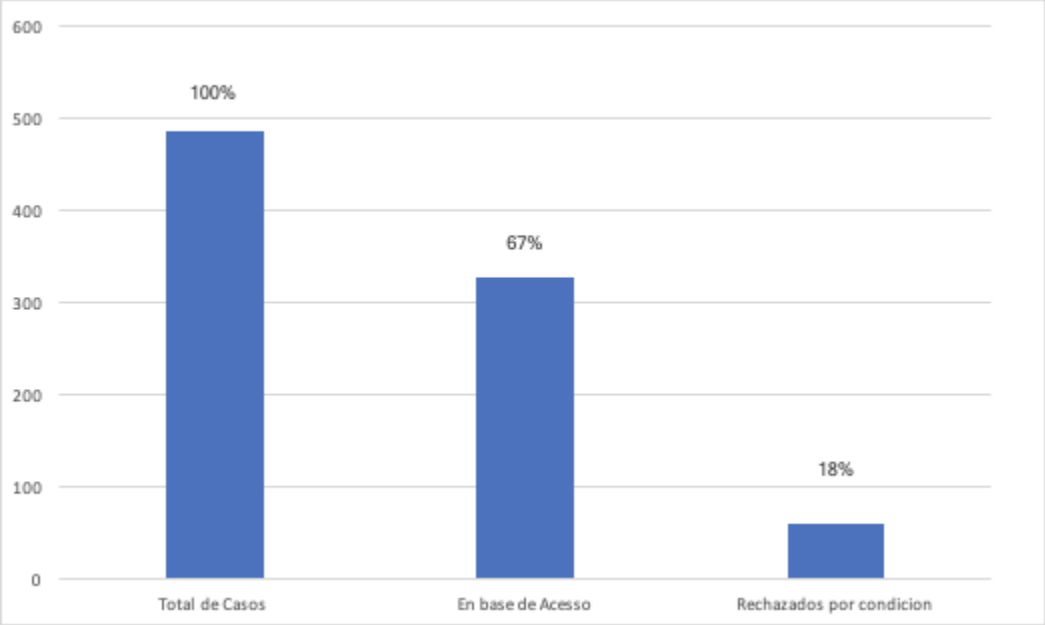

Figura 9: Resultado del backtesting de Acesso

A diferencia del caso con *Certibio*, este proveedor encontró casi un 70% de los casos entre sus registros. Sin embargo, de la cantidad de casos con registros, un 18% hubiesen sido rechazados debido a las condiciones impuestas. Ese porcentaje resulta similar al número de préstamos en default, sin embargo, como se verá después, no incluye únicamente clientes con fraude de identidad.

Plan

A partir de la hipótesis de que los individuos tomando préstamos bajo una identidad falsa no tienen ninguna intención de repago, se asume que cualquier préstamo emitido a ellos irá a pérdida. El objetivo de la implementación será lograr una reducción en el número de personas que acceden a un préstamo realizando fraude de identidad, y por lo tanto, un aumento en el resultado final de la empresa.

Por lo tanto, se realizará el agregado de un proceso de validación biométrica que permita, mediante una base de datos de un proveedor externo, verificar la identidad del cliente que está solicitando el préstamo.

Ventajas de la reducción de fraude

Como se mencionó anteriormente, agregar un método de validación de identidad permitirá reducir el fraude en la emisión de préstamos, aumentando así la ganancia promedio por cada préstamo emitido.

Tener menores tasas de default permiten a la empresa ofrecer mejores tasas de interés a los clientes, y esto a su vez aumenta la tasa de conversión del proceso de aplicación ya que los clientes se ven favorecidos por intereses más bajos. Los menores montos también hacen que sea más fácil devolver el dinero prestado.

A la vez, mejorar el rendimiento de la cartera hace posible obtener financiación a mejores tasas. El principal modo de negocio de un ente emisor de créditos consiste en tomar dinero a una tasa, y ofrecerlo a una mayor, obteniendo la diferencia como ganancia. Al no tener historial crediticio, inicialmente es difícil poder emitir deuda a tasas bajas: Los prestamistas asumen que el riesgo es alto, y esto se ve reflejado en los intereses. Lograr que más clientes devuelvan los préstamos genera confianza en los prestamistas, y se pueden obtener tasas más bajas.

Propuesta de valor

La solución propuesta es un sistema de validación de identidad que incluye un sistema de prueba de vida. El objetivo del sistema de prueba de vida es evitar que un tercero utilice fotos de un individuo para obtener un préstamo a nombre de él- es decir, que utilice información biométrica correcta, un tipo de fraude de identidad más complejo.

Para la prueba de vida se utilizará el sistema de un proveedor externo llamado Zoom. El algoritmo del mismo consiste en un sistema de reconocimiento facial que, mediante análisis de las imágenes

capturadas por la cámara en tiempo real, es capaz de reconocer un rostro y seguir sus movimientos de modo de aseverar que se trata de una persona; no una imagen o una máscara.

Por lo tanto, el producto a implementar consiste en:

- Una pantalla de captura de foto y prueba de identidad en la aplicación.
- Una serie de modificaciones en el servidor para contactar con los proveedores de identidad y almacenar el resultado.

Dado que se trata de integrar estas mejoras en un flujo existente, se deberá modificar lo menos posible el proceso de alta del usuario. Actualmente el proceso de alta consiste de los siguientes pasos:

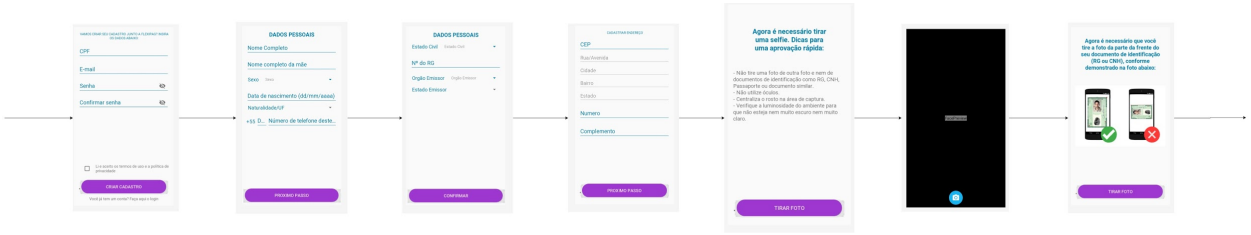

Figura 10: Flujo actual de pantallas para la validación de identidad del usuario

Para evitar agregar complejidad, se reemplazará el paso de la toma de selfie por la pantalla de prueba de identidad. Durante la misma, se tomará una foto al azar y se utilizará dicha imagen como selfie.

Dado que la mayoría de los usuarios no respetaban las instrucciones, que el algoritmo de reconocimiento requiere buenas condiciones de iluminación y que los lentes pueden dificultar el reconocimiento facial, se reemplazarán las instrucciones escritas por un video con instrucciones.

El nuevo flujo de pantallas es el siguiente:

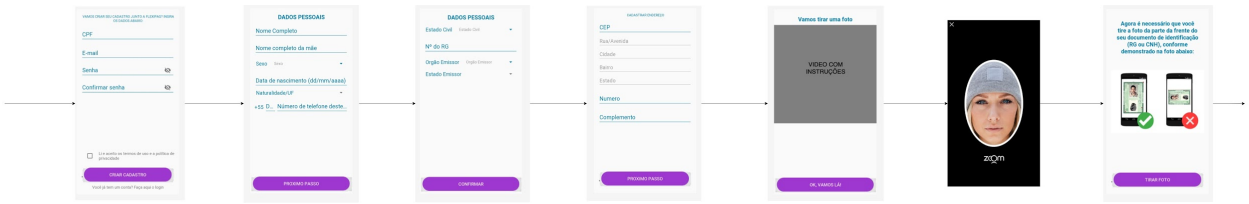

Figura 11: Flujo de pantallas a desarrollar para la validación de identidad biométrica

Se puede observar que el nuevo flujo de pantallas es idéntico al anterior, no agregando complejidad que podría resultar en usuarios que abandonen la aplicación.

Factores claves

El éxito del proyecto dependerá de la cantidad de usuarios que cuenten con la capacidad de ser validados de manera positiva. Es decir, dado que la aprobación del usuario depende de la existencia de sus datos biométricos en las bases de datos consultadas, es importante elegir proveedores de biometría que contengan datos de los segmentos a los que se pretende servir para el servicio de préstamos personales.

Dado que aquellos usuarios que no logren pasar la prueba de vida no podrán continuar con el proceso de validación de identidad, es importante que los usuarios puedan realizar este paso sin dificultades. A partir de la inspección de las selfies capturadas hasta el momento se determinó que estas eran de mala calidad, siendo tomadas en el transporte público, con un mal ángulo o en habitaciones oscuras. Por esto, se decidió cambiar las instrucciones textuales por un video explicativo que comunicase mejor los requerimientos. Si los usuarios no logran completar la prueba de vida con facilidad, se dará un número elevado de usuarios que abandonan el proceso (*churn*).

Cronograma de Trabajo

Antes de comenzar la implementación del sistema, se realizó un relevamiento de los proveedores existentes. Habiendo seleccionado aquellos proveedores que mejor se ajustaban al público objetivo de la empresa, se procedió a un proceso de backtesting que demoró 20 días.

El proceso de backtesting consiste en una prueba del sistema con datos reales. En este caso, se enviaron fotos históricas de los clientes para verificar si correspondían con las identidades que informaron al momento de registro.

A partir de los resultados, se eligieron *Acesso Digital* y *Certibio* como proveedores y se estimó el tiempo de implementación en 3 meses, con 15 días de prueba y lanzamiento.

El calendario de implementación fue el siguiente:

Noviembre 2018	Relevamiento de proveedores de biometría
Diciembre 2018	Backtesting de proveedores
Enero 2019	Comienzo de implementación de proveedores
Marzo 2019	Fin de implementación y pruebas
Abril 2019	Re-lanzamiento de la aplicación

Figura 12: Calendario de desarrollo

Durante el periodo de implementación se suspendieron las campañas de marketing para evitar adquirir más usuarios fraudulentos, pero no se suspendió la emisión a modo de seguir generando una cartera.

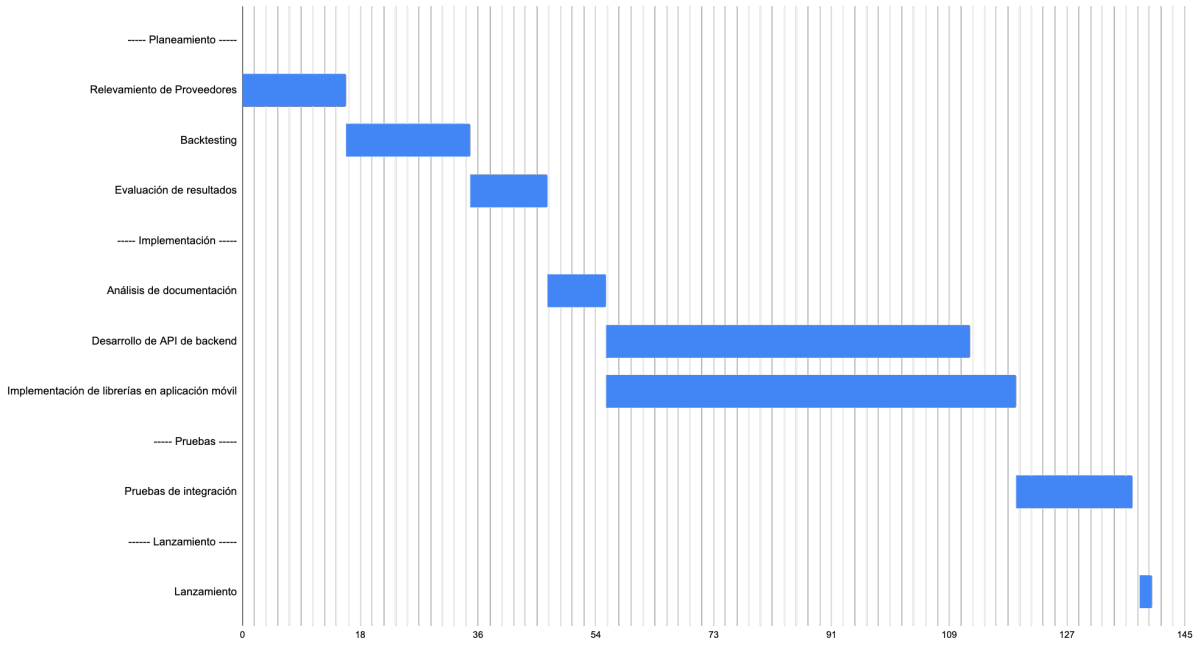

Figura 13: Gantt de desarrollo por tareas

El plan de desarrollo fue elaborado considerando el equipo disponible de tres personas.

Costos

Dado que todo el desarrollo se realizó de manera interna, el costo de desarrollo se estimó en horas de trabajo del equipo de tecnología.

A lo largo de cuatro meses, desde Diciembre de 2018 a fines de Marzo de 2019, se dedicaron un total de 633 horas de desarrollo al proyecto de implementación del sistema de identidad biométrica. Las mismas correspondieron a las áreas de backend y mobile, para realizar las pruebas de backtesting, desarrollar las integraciones con los proveedores, y modificar los flujos de la aplicación mobile para utilizar las nuevas funcionalidades.

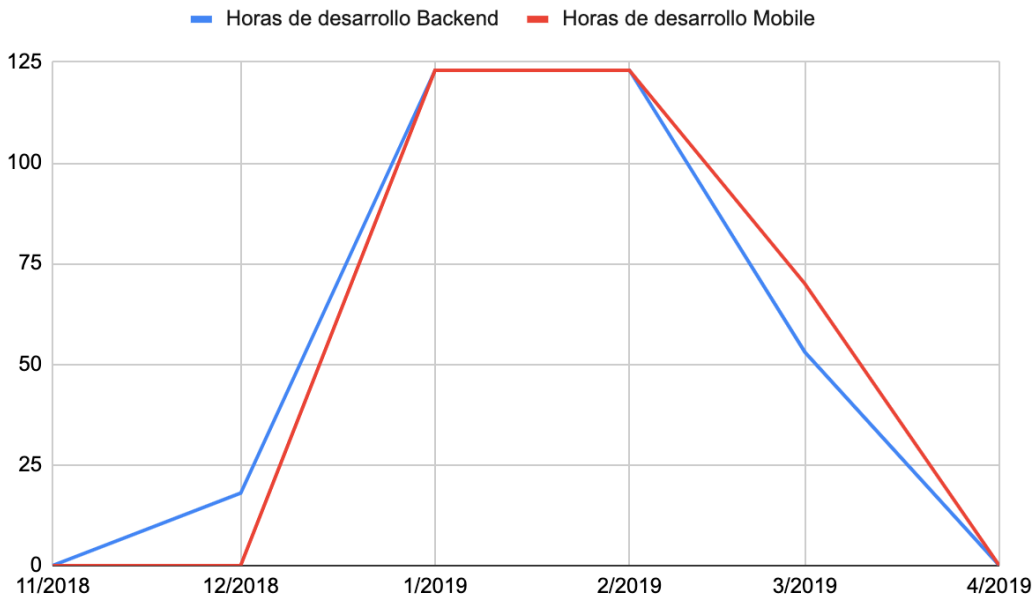

Figura 14: Inversión en horas de desarrollo por área

De acuerdo al salario al momento de la implementación, el costo total del desarrollo fue de 4674 dólares.

Los proveedores de biometría facial no tienen un costo de ingreso, sino que cobran por cada consulta que se realiza a sus sistemas y retorna un resultado. Es decir, aquellas consultas de personas que no estén en la base de datos no incurrirán en costos por la implementación del sistema El costo por cliente identificado (ya sea con validación positiva o negativa) es de 0,2 Reales, incluyendo en el costo la prueba de vida.

Un costo no estimado es el costo de oportunidad de perder clientes que no están intentando cometer fraude, pero cuyos rostros no se encuentran en las bases de datos biométricas. Al no poder ser validados, estos individuos serán rechazados por el sistema resultando en un mayor costo de adquisición. Con suficiente volumen, se podría dejar pasar un porcentaje de estos usuarios no validados para evaluar su performance. De esta manera, sería posible calcular el riesgo, y el interés necesario para que dicho segmento sea rentable.

Si bien se espera un aumento de la métrica del costo de adquisición por los usuarios rechazados por validación, se espera que la mejora en pérdidas y resultados por intereses sea mayor.

Proyección

Analizando los resultados de las evaluaciones de backtesting pedidas a los proveedores elegidos, *Certibio* y *Acesso*, se observó que al analizarlos en conjunto sólo un 24% de los clientes serían rechazados por falta de registro en las bases biométricas. Al mismo tiempo, del total de casos con identificación, un 14% sería rechazado por suponerse fraude de identidad.

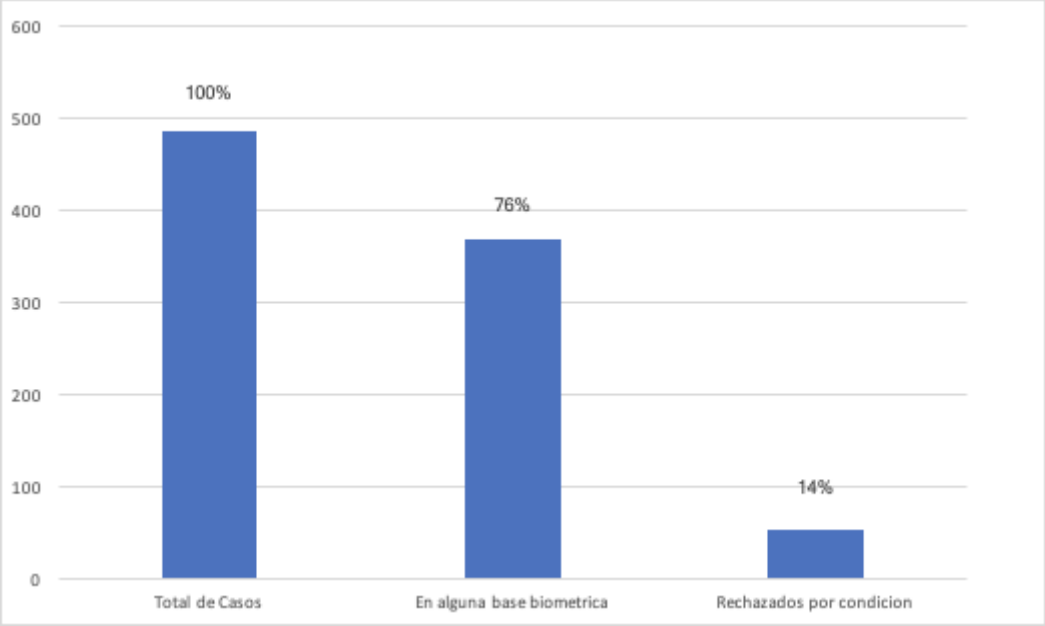

Figura 15: Resultado de los backtesting considerados en conjunto

La reducción en la cantidad de fraude sospechado al compararlo con el backtest de *Acesso Digital* (14% contra 18%, una reducción de 4%) se debe a la forma en la que se determina si un usuario está validado o no. Dado que hay dos fuentes de datos biométricos de distintos niveles de validez, se da prioridad a los resultados obtenidos de *Certibio*, ya que sus datos son provistos por bases de datos gubernamentales, mientras que los datos de *Acesso Digital* son provistos por las entidades

que utilizan el sistema. Por lo tanto, el flujo de decisión para aprobar o rechazar un individuo a partir de su reconocimiento facial resulta:

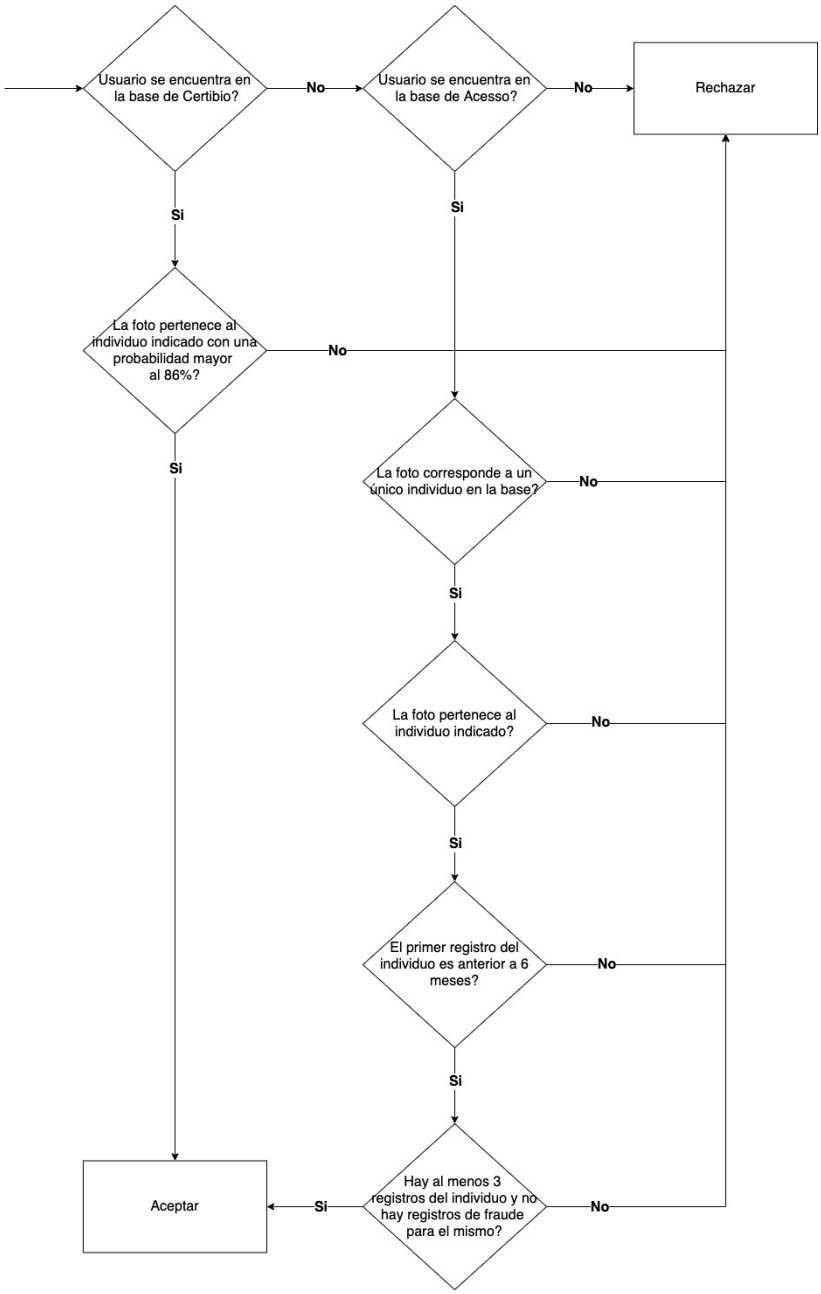

Figura 16: Árbol de decisión para determinar si la identidad de un individuo es válida

A partir de los resultados, se estima que el Costo de adquisición por usuario aumente un 35% (24% por individuos que no serán encontrados en las bases de identidad biométricas y 11% por casos

rechazados por sospecha de fraude de identidad). Al mismo tiempo, se espera que las pérdidas de capital se reduzcan en un 11% y se incrementen los intereses capitalizados en el mismo porcentaje.

	Previo a la implementación	Proyectado
Intereses recibidos	BRL 52,27	BRL 58,00
Intereses pagados	-BRL 13,76	-BRL 13,76
Tasa de Servicio	BRL 0,50	BRL 0,50
Pérdidas de Capital	-BRL 7,20	-BRL 6.40
CAC	-BRL 10,00	-BRL 13,50
Costos de Scoring	-BRL 11,55	-BRL 14,30
Costos de Formalización	-BRL 1,65	-BRL 1,65
Costos de Recaudación	-BRL 2,90	-BRL 2,90
Contribución Marginal Neta	**BRL 5,71**	**BRL 5.99**

Figura 17: Proyección sobre la contribución marginal neta por primer préstamo

TIR del Proyecto

A partir de los gastos en horas de desarrollo y la mejora en la contribución marginal neta, se puede estimar una TIR utilizando 150 préstamos mensuales con un crecimiento del 10% mes a mes durante dos años.

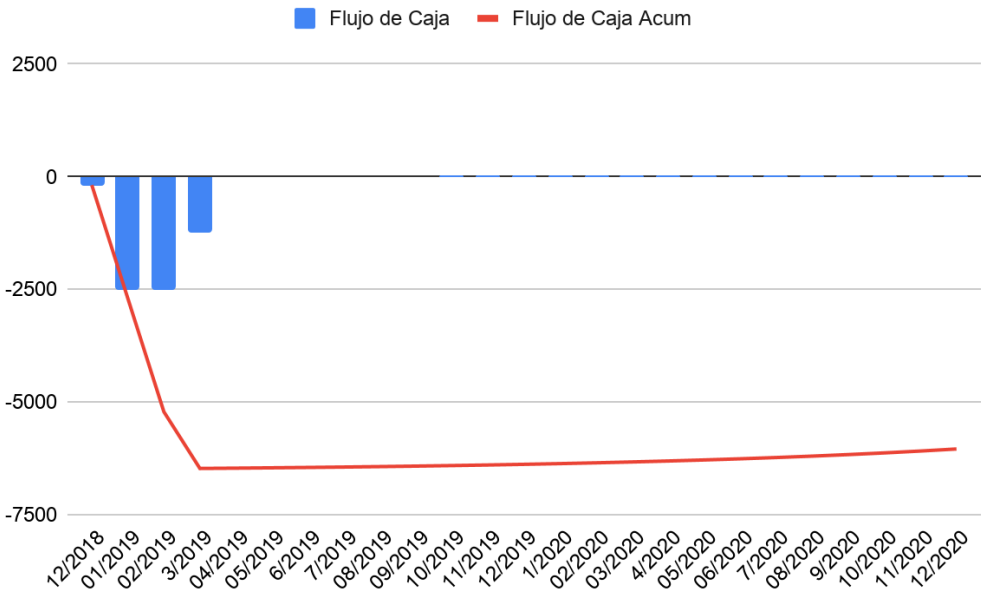

Figura 18: Flujos de caja proyectados a lo largo de un año en USD

El VAN del proyecto es negativo con estos flujos de caja, resultando en -5.938,68 USD, con una TIR negativa.

Consideraciones sobre TIRs negativas en una startup

Se podría argumentar que dado que ejecutar el proyecto resta valor, no debería realizarse, en cambio continuando la emisión de préstamos bajo las mismas condiciones. En este análisis no se está considerando el impacto de una cartera con FPD alto y de alta variabilidad. Estabilizar el porcentaje de préstamos que van a pérdida sin realizar pagos permite dar previsibilidad a las métricas de contribución marginal en escala.

Por otro lado, las startups se rigen por rondas de inversión. Dependiendo de la industria a la que pertenece la empresa, hay métricas benchmark que deben ser alcanzadas para considerar un negocio escalable. En el caso particular de una fintech de créditos estas son:

- Porcentaje de préstamos enviados a pérdida sin pagos (First Payment Default)
- Porcentaje de préstamos con 30, 60 y 90 días de atraso

- Duración y valor promedio de la cartera de préstamos
- Tamaño de la cartera de préstamos

El nivel de FPD que tenía la cartera previa a la implementación iría a dificultar la siguiente ronda y por lo tanto, era necesario mejorar esa métrica antes de comenzar a contactar inversores.

Resultados

Incidencia en la métrica de FPD

Antes de la implementación de los sistemas de validación biométrica, el porcentaje de préstamos que eran enviados a pérdida sin haber recibido pago alguno era mayor al 18%. Es decir, casi 1 en 5 usuarios no devolvían ninguna parte de su deuda.

Luego de implementar el nuevo sistema, se observó una baja significativa de los préstamos en First Payment Default, reduciendo el porcentaje de préstamos no devueltos a un 10%, o 1 en 10. Esto significa una reducción de casi el 50% en las pérdidas ocasionadas por usuarios fraudulentos.

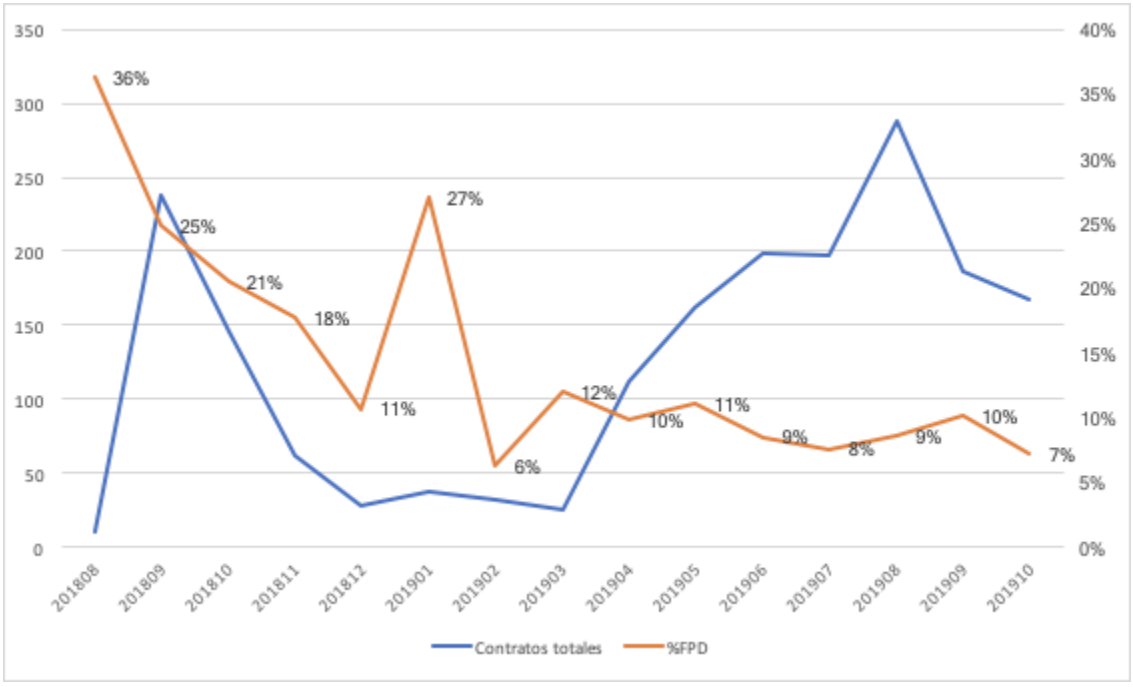

Figura 19: Incidencia de la implementación del nuevo sistema de validación biométrica

En el gráfico se puede observar como el FPD se reduce significativamente a partir de Abril del 2019, el primer mes que se emitieron el 100% de los préstamos con el nuevo sistema. Las variaciones elevadas en los porcentajes de Default en Primer Pago de los meses de Agosto del 2018 y de Noviembre de 2018 a Marzo de 2019 se debe a la baja emisión de préstamos de ese período. Es decir, en períodos con pocos contratos emitidos, es difícil estimar con precisión las métricas: por ejemplo, en Diciembre de 2018 se emitieron 28 contratos, y sólo tres fueron enviados a pérdida, mientras que el mes siguiente, se enviaron 10 préstamos sobre un total de 37. O sea, 200% más que el mes anterior, o 3 veces más en cantidad absoluta, sin realizar ningún cambio.

A partir de Abril se puede ver como con mayor número de préstamos el valor se estabiliza alrededor del 10%.

Esta reducción en el FPD es acorde al número estimado a partir del backtesting realizado con el proveedor, en el mejor caso.

También se observa que se logró alcanzar la cantidad de préstamos emitidos anterior a la reducción en el gasto de publicidad. Esto se debe a la implementación del segundo sistema de validación para los casos que no tuviesen CNH.

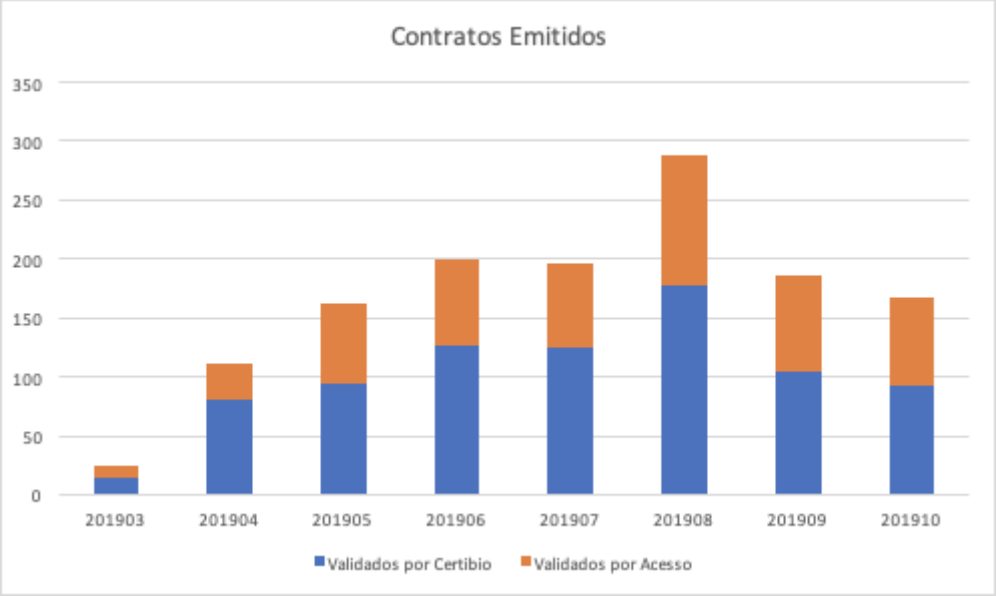

Figura 20: Contratos emitidos por validación de cada proveedor

En el gráfico se muestran los contratos emitidos desde Marzo de 2019 y se observan los contratos emitidos de acuerdo a cada validador. Se puede observar que la cantidad de préstamos validados por *Certibio* es entre un 20% y un 50% superior a los validados por *Acesso*. Esto es posible que sea debido a que la base de datos del proveedor biométrico es incompleta y sólo tiene registros de personas que ya hayan sido registradas en las tiendas de sus clientes.

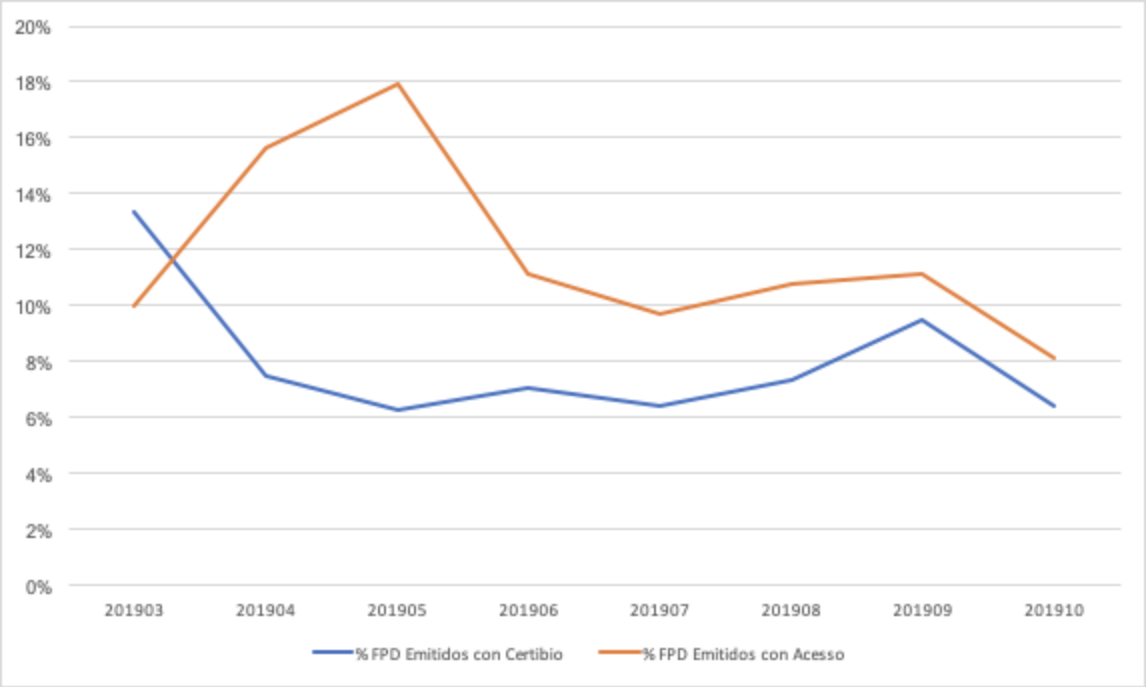

Figura 21: Porcentaje del total de FPD por proveedor biométrico

En el gráfico se muestra el porcentaje de préstamos declarados como perdidos sin pagos de acuerdo al proveedor de biometría utilizado. Se puede observar que a partir de Abril, mes en el que se emitieron el 100% de los préstamos validados por proveedores de biometría, el porcentaje de préstamos en FPD validados por *Acesso* es ligeramente superior al porcentaje de préstamos en FPD validados por *Certibio*. Esto puede deberse en parte a que hay menos préstamos emitidos por ese proveedor y por lo tanto es un error por un número pequeño de muestras, o también puede deberse al hecho de que el proveedor no utiliza datos validados por el gobierno y depende de validaciones de privados.

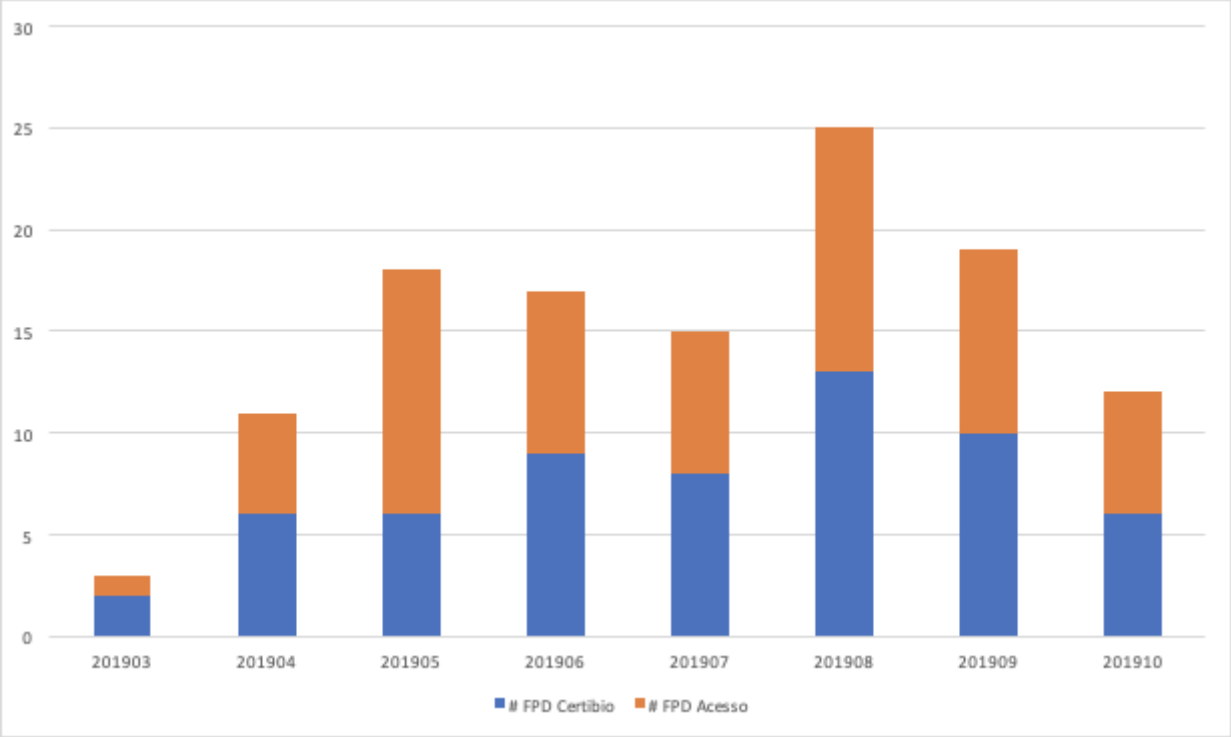

Figura 22: Cantidad de préstamos en FPD de acuerdo al proveedor biométrico

En el gráfico se muestran la cantidad de contratos en First Payment Default, segmentado por proveedor de biometría. Del gráfico se obtiene que la cantidad es similar mes a mes, a pesar de ser emitidos más contratos con *Certibio* que con *Acesso*. Otra explicación posible es que aquellos clientes que tienen registro de conducir, tienen acceso a un vehículo y pertenecen a un segmento de mayores ingresos y mejor capacidad de repago que aquellos que no tienen.

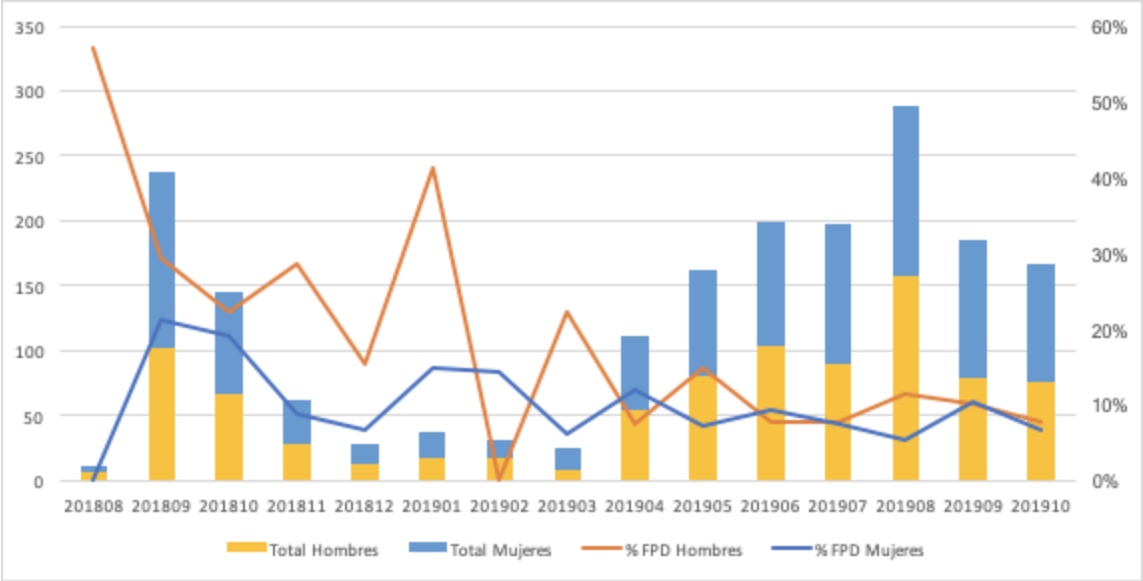

Figura 23: Comparación de FPD entre hombres y mujeres

Al comparar la cantidad de préstamos emitidos por sexo, se puede observar que previo a la implementación del sistema el porcentaje de préstamos emitidos que finalizaron como default en primer pago es mayor en el caso de los hombres. Posterior a la implementación del sistema, el porcentaje de FPD es similar para ambos sexos.

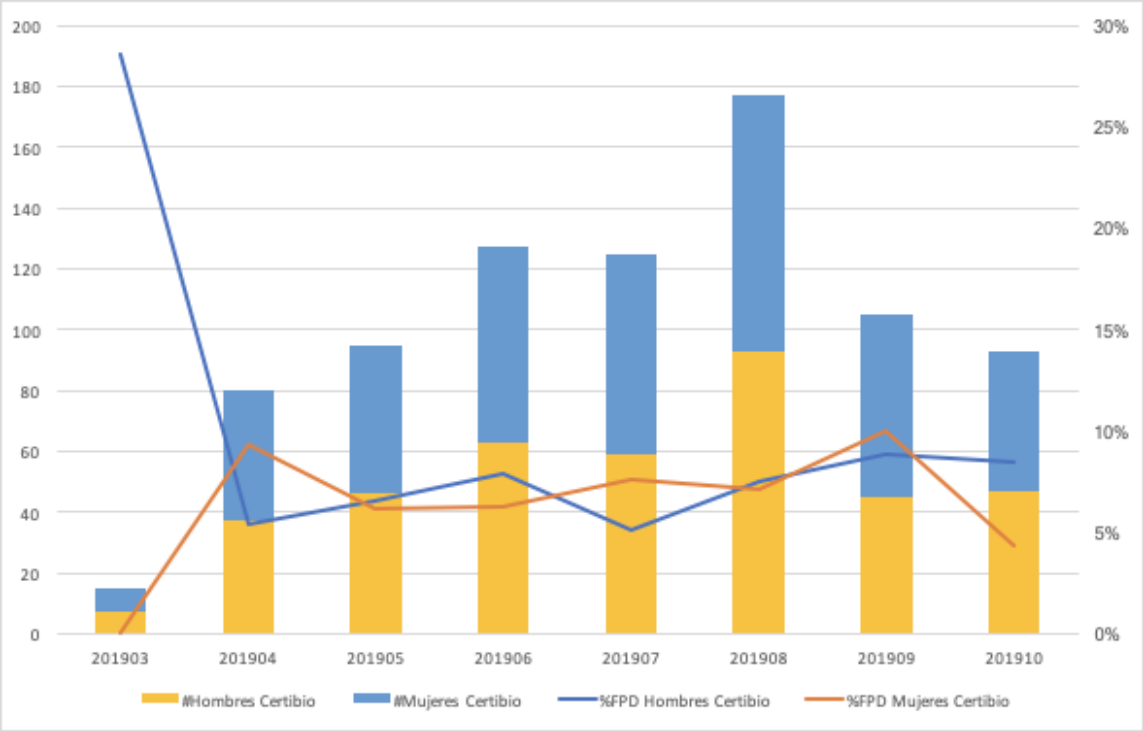

Figura 24: Distribución de préstamos emitidos por sexo para Certibio

El gráfico muestra que el porcentaje de préstamos validados por *Certibio* y cayeron en default antes del primer pago es similar en el caso de hombres y mujeres, alrededor de un 8%. Dado que estas personas se encuentran validadas contra bases de datos estatales, esto establece una métrica base de personas que entran en FPD a pesar de ser quienes dicen ser, es decir, personas que no tienen interés en (o capacidad para) repagar el préstamo.

Figura 25: Distribución de préstamos emitidos por sexo para Acesso

Graficando los mismos valores para *Acesso*, se puede observar que los porcentajes de FPD para hombres y mujeres tienen diferencias de hasta un 15%, con variaciones muy altas mes a mes. Esto puede deberse en parte al número reducido de contratos emitidos con este proveedor biométrico. Los valores de FPD para este proveedor rondan en torno al 10%. Podría deberse a un error debido al tamaño de la muestra.

Con ambos proveedores de biometría se aprobaron un número similar de hombres y mujeres, por lo que no parece haber una tendencia.

Potenciales clientes rechazados

Para analizar el impacto de la implementación del método de validación biométrica también se deben analizar los clientes que fueron rechazados a partir de los nuevos requerimientos. Entre ellos se encuentran clientes sin registros y clientes que no cumplen con las condiciones de validación.

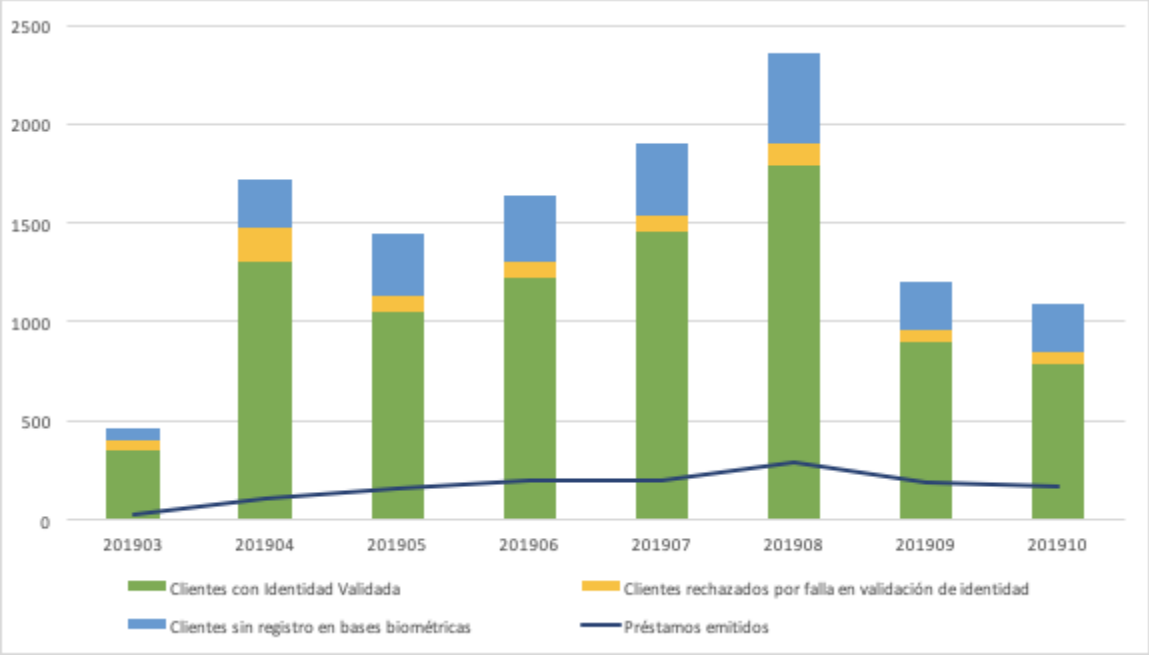

Figura 26: Resultado de la validación de identidad por mes

En el gráfico se puede observar mes a mes la evolución de clientes validados y préstamos emitidos, con un gran porcentaje de personas que logran validar su identidad de manera satisfactoria. Sin embargo, debido a restricciones de historia crediticia u otras particularidades (como usuarios rechazando los préstamos por una tasa de interés elevada o no tener una cuenta a su nombre), sólo un pequeño porcentaje termina el proceso.

En promedio, un 15% de los individuos que logran validar su identidad terminan y confirman la solicitud de préstamo.

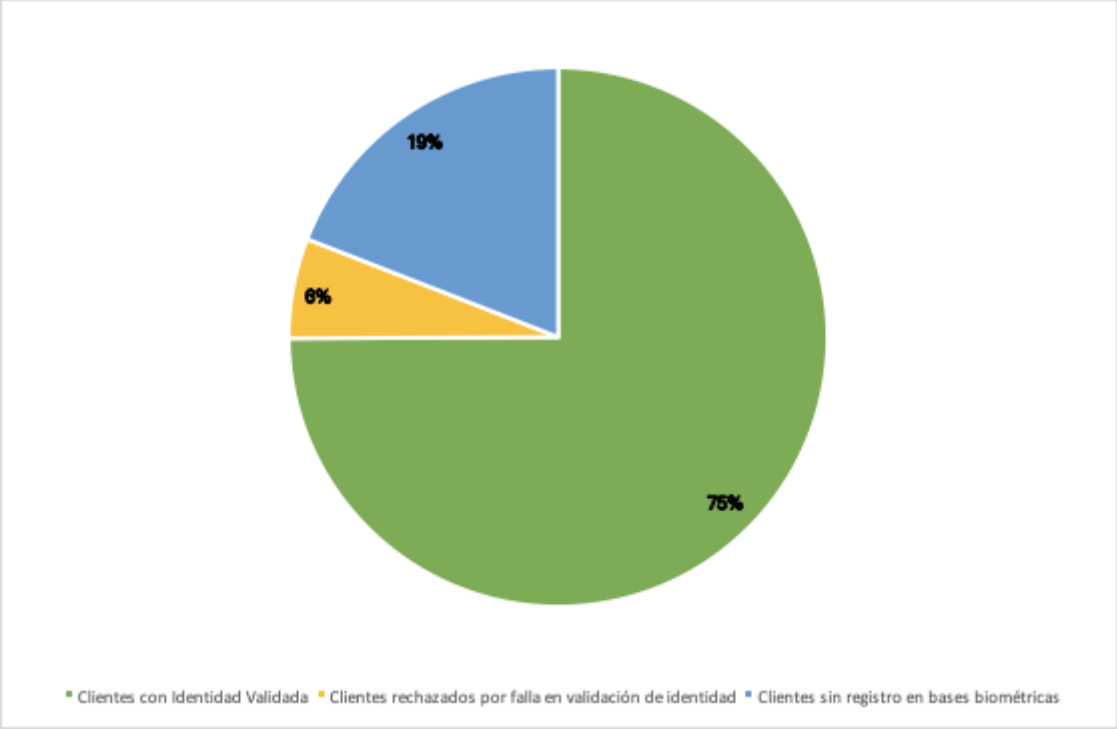

Figura 27: Resultados sobre el total de casos validados

Del total de casos validados, un 75% correspondían a individuos que eran efectivamente quienes decían ser. Es decir, que había una correlación positiva entre la foto almacenada en las bases biométricas y la selfie provista por el usuario. Sólo un 19% de los individuos que intentaron validar su identidad no encontraron un registro en las bases de datos biométricas. Es decir, frente al 70% de usuarios que se esperaban perder por falta de identidad en la base de *Certibio*, y al 30% que se esperaba perder con *Acesso*, el porcentaje resultó inferior al 20%.

El porcentaje de individuos rechazados por no cumplir con las condiciones de validación de identidad fue de 6%, un valor similar a la mejora en porcentaje de préstamos en FPD registrado. Si bien la mejora registrada es superior, se supone que una porción de las personas que obtenían préstamos con identidades falsas no se encontraban en bases biométricas. También se supone que un porcentaje de usuarios legítimos no se encontraban en las bases biométricas, pero no se pudo cuantificar.

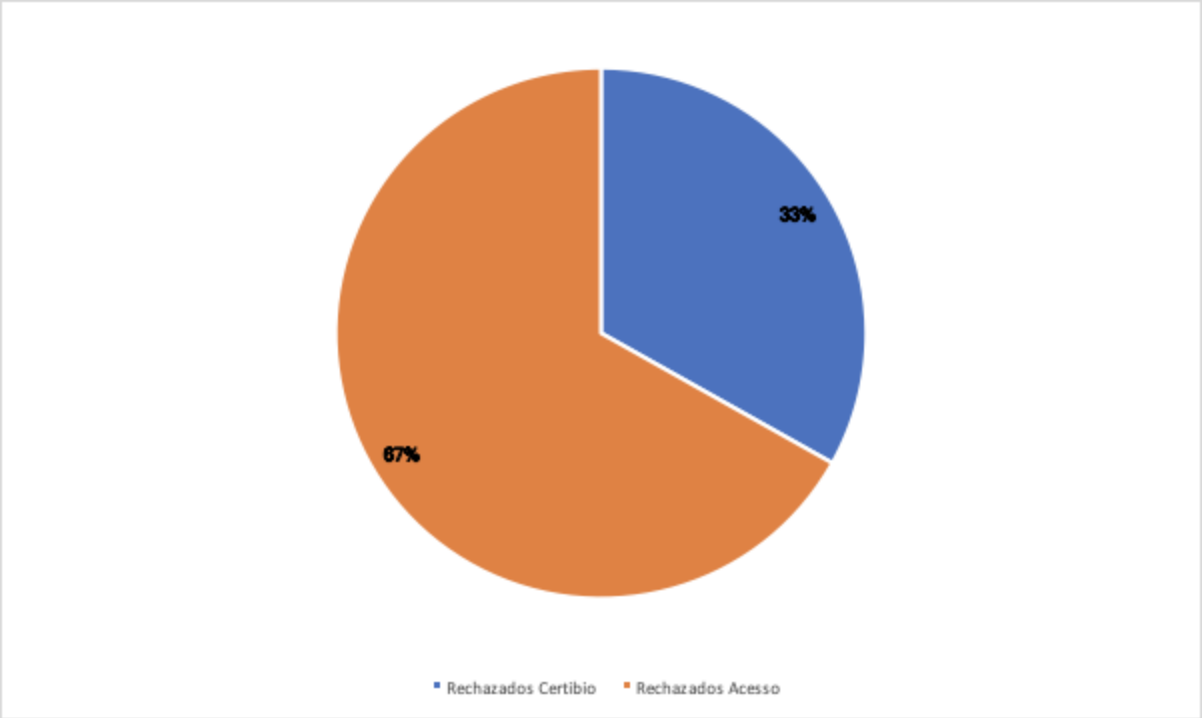

Figura 28: Porcentaje de rechazos por proveedor de biometría

Si bien hay más individuos rechazados por *Acesso*, es importante destacar que dado que *Acesso* es un proveedor privado, no hay certeza sobre la identidad de los individuos que son registrados, por lo que se utilizaron reglas que asegurasen cierta calidad sobre los individuos como se mencionó en la sección de *Backtesting*.

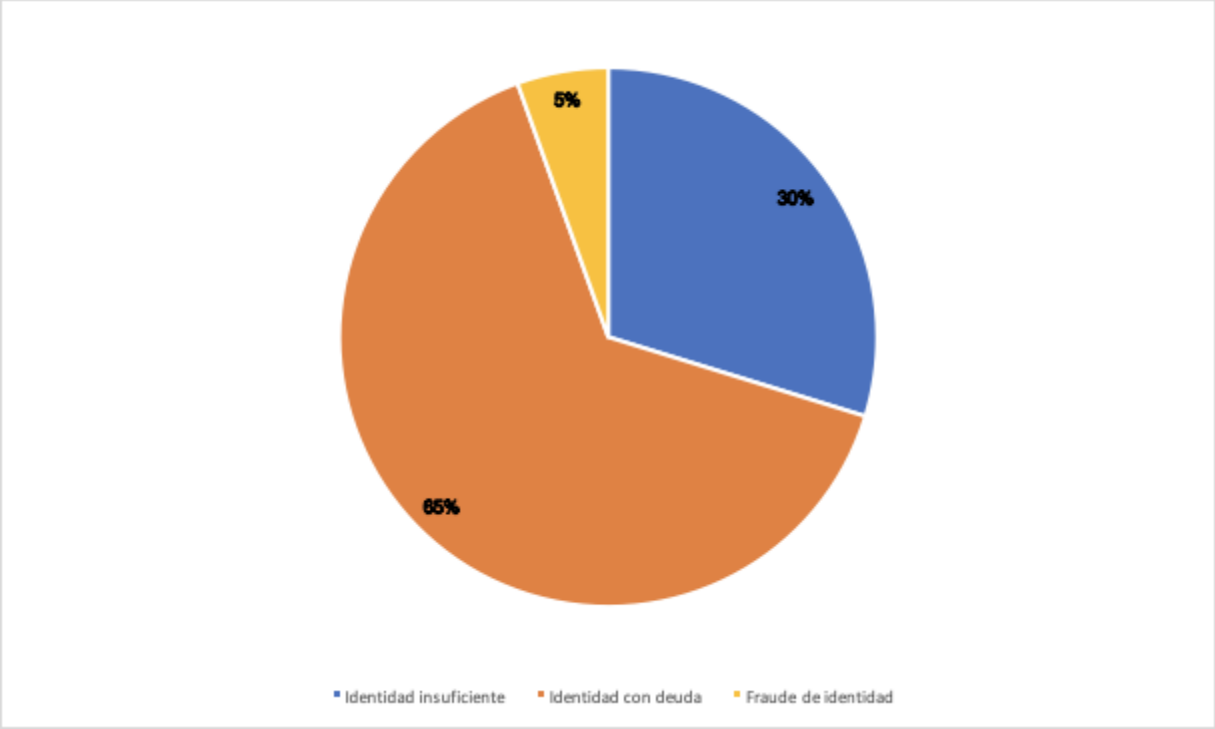

Figura 29: Motivos de rechazo de Acesso

El principal motivo de rechazo cuando se trata de *Acesso* se debe a que el usuario fue registrado antes y su identidad cuenta con algún tipo de deuda o advertencia agregada por parte de las diversas tiendas que utilizan los servicios del proveedor biométrico. El segundo motivo es que si bien se encontraron registros del individuo, no cumplían con los requerimientos de antigüedad. Sólo un 5% de los casos correspondían a individuos que cometían fraude de identidad, con sus rostros identificados en más de un registro.

En ambos proveedores se pudo observar un porcentaje similar de usuarios que estaban intentando realizar fraude de identidad de acuerdo a los datos en las bases de datos de cada uno de los proveedores. El porcentaje similar en *Acesso* y *Certibio* da un grado de seguridad a la base privada, que no cuenta con datos validados por el estado de Brasil.

Resultado en el costo

La implementación de la verificación biométrica se realizó luego de la validación de datos básicos del documento. Se evaluaron los costos de la utilización de este servicio y se concluyó que dado el flujo de usuarios actual, para reducir los costos, se implementaría después de la validación de datos básicos.

	Costo
Verificación de datos básicos	BRL 0,1
Verificación de antecedentes penales	BRL 0,1
Prueba de vida y validación biométrica	BRL 0,2
Análisis crediticio	BRL 0,65

Figura 30: Costos de cada servicio utilizado

Los costos en el funnel de usuarios se ordenaron por mayor incidencia sobre el costo final del préstamo. Es decir, ordenar las validaciones de forma que el costo final de validación por préstamo (considerando los clientes que no terminan el proceso) sea el menor.

Luego de la implementación, se incrementaron en un 20% los costos de Scoring que incluyen las validaciones mencionadas anteriormente debido a la implementación del nuevo método y la mayor tasa de rechazo. A su vez, se redujeron en un 45% las pérdidas de capital (es decir, capital perdido por préstamos no devueltos) y la reducción en cantidad de préstamos en FPD significó también un aumento del 10% en intereses recibidos.

	Previo a la implementación	Post Implementación
Intereses recibidos	BRL 52,27	BRL 57,37
Intereses pagados	-BRL 13,76	-BRL 13,76
Tasa de Servicio	BRL 0,50	BRL 0,50
Pérdidas de Capital	-BRL 7,20	-BRL 4,00
CAC	-BRL 10,00	-BRL 12,50
Costos de Scoring	-BRL 11,55	-BRL 13,37
Costos de Formalización	-BRL 1,65	-BRL 1,65
Costos de Recaudación	-BRL 2,90	-BRL 2,90
Contribución Marginal Neta	**BRL 5,71**	**BRL 9,69**

Figura 31: Impacto en la contribución marginal por préstamo

El resultado final en la Contribución Marginal Neta del préstamo promedio fue de un incremento del 67%.

Conclusiones

En el presente trabajo se analizó el impacto de la implementación de un método de validación biométrica en la contribución marginal del producto principal de una fintech: los micro-créditos personales.

En un entorno de fraude de identidad elevado como es el mercado de Brasil, región en la que se desarrolla el negocio de la startup, la implementación de un sistema de validación de identidad resultó en una mejora importante en el número de préstamos que fueron enviados a pérdida sin recibir ningún pago. Este tipo de préstamos, declarados en default de primer pago (o First Payment Default), tiene una componente de personas que utilizan identidades falsas para adquirir un préstamo y personas legítimas que adquieren un préstamo sin capacidad o intención de repago.

Se implementó un sistema que utilizaba una selfie con prueba de vida (para prevenir fotos de fotos) y se utilizaron dos proveedores: uno de ellos utilizaba datos provistos por una base de datos de fotos estatal, y otro de una base de datos de entes privados asociados. El sistema de validación compara la selfie contra una foto almacenada en la base de datos y devuelve la probabilidad de que dicha persona sea el individuo, o se esté cometiendo algún tipo de fraude.

Si bien se lograron evitar casos en los que un individuo estaba realizando fraude de identidad, la utilización de estos sistemas también significó un menor porcentaje de individuos obteniendo un préstamo a causa de la incapacidad de validar su identidad (de manera de determinar si se estaba cometiendo fraude o no). A pesar de esto, el porcentaje de individuos que fueron rechazados por este motivo fue menor al esperado.

Si bien aumentó el costo de emisión por préstamo debido a un porcentaje mayor de clientes rechazados, la reducción en préstamos en default redujo los valores de pérdida de capital, y aumentó los montos de intereses recibidos. El costo de utilizar el sistema fue menor que las ganancias obtenidas a partir de las mejoras en la métrica de FPD.

El rendimiento del proveedor privado a la hora de evitar fraudes de identidad fue peor que el que utilizaba bases de datos estatales, sin embargo fue mejor que la alternativa de no usarlo: Es decir,

aun cuando era difícil aseverar la identidad de un individuo, se redujo la cantidad de personas que no tenían interés en pagar. Esto resulta una alternativa posible en mercados que no cuentan con bases de datos biométricas estatales, pudiendo armar una alternativa con datos recolectados por privados.

Aciertos y fallas

Mejora en default

Al aplicar las medidas de validación biométricas se logró una reducción importante de los préstamos enviados a pérdida, aumentando la ganancia por préstamo otorgado. Esto demuestra una clara correlación entre el fraude y el bottom line de la empresa.

Como parte del proceso de aprobación del préstamo, se verifican deudas existentes del usuario en bases de datos financieras (similar al Veraz en Argentina). Si estas superan una cierta cantidad o un monto, el préstamo es rechazado.

Dado que el modelo de negocio permite bloquear el celular del usuario, una vez establecido que se eliminó el fraude, se podría haber intentado reducir las restricciones aplicadas sobre los filtros de deudas para mejorar la tasa de aprobación y de esa manera reducir el costo de adquisición por préstamo.

Personas no validadas no atendidas

Como consecuencia de la introducción del sistema de validación biométrica, si un usuario no puede ser validado, es rechazado. Si bien se logró aumentar las ganancias, se están descartando potenciales buenos clientes por los que se está pagando una instalación.

Dado que la mayoría de los gastos se realizan para validar un usuario (su identidad, los datos del documento, sus deudas), los préstamos recurrentes dejan un mayor margen de ganancias. Una

posibilidad para no perder potenciales clientes legítimos sería ofrecer montos más pequeños a usuarios nuevos no validados y observar su comportamiento.

Este mecanismo de realizar préstamos reducidos a nuevos clientes y luego aumentar el límite es común en la mayoría de los prestamistas. Por esto, los defraudadores intentan obtener un préstamo y pagarlo inmediatamente para aumentar el límite. Ya que el comportamiento del usuario está completamente digitalizado, es posible detectar estos patrones y configurar el sistema para que los excluya.

Potenciales mejoras

Integración con más proveedores biométricos

Dado que la implementación del sistema de biometría resultó en una mejora en la contribución marginal unitaria, una posibilidad para reducir los individuos rechazados por falta de identidad es la implementación de nuevos proveedores. Esto reduciría los costos de adquisición por usuario.

Sin embargo, la implementación de un nuevo proveedor incrementaría los costos de scoring promedio y debe evaluarse el rendimiento de los usuarios aprobados con el nuevo servicio. Además, el proveedor podría exigir un monto mínimo (equivalente a una cantidad base de consultas) que podría ser mayor al costo de no implementarlo. Dado que actualmente ningún proveedor contiene los datos biométricos de la totalidad de los ciudadanos de Brasil, ni se puede asegurar que existan varios que sumados cubran el total, cada nuevo proveedor implementado tiene retornos decrecientes.

Características de fotos de defraudadores (Tomadas a distancia, a escondidas)

Dado que se cuenta con las fotos individuales tomadas al momento de solicitar el préstamo, es posible intentar aplicar algoritmos de computación visual para entender el contexto en el que fueron sacadas. Mediante inteligencia artificial es posible clasificar la situación en la que se tomó

la imagen, y luego asociar dichas situaciones a resultados en el comportamiento del usuario (buen pagador, o mal pagador).

Al analizar las fotos manualmente se distinguieron distintas situaciones particulares de cómo era tomada la imagen:

- En el lugar de trabajo
- En el transporte público
- A escondidas
- Alguien más tomaba la foto

Es posible que exista una correlación entre la situación en la que se tomó la foto y el comportamiento del individuo al momento de devolver el préstamo. Por ejemplo, hay personas que prestan su identidad para realizar préstamos fraudulentos voluntariamente, y es esperable que en estos casos la foto sea tomada a la distancia.

Utilización de datos adicionales (IP, geolocación)

Dado que los usuarios realizan todo el pedido de préstamos mediante la aplicación, se cuenta con los datos de su comportamiento durante todo el proceso. En particular, se conocen el IP (la dirección de internet) y los datos de geolocalización del dispositivo desde el que está realizando el pedido. A partir de estos, se puede obtener:

- Si el usuario está utilizando WiFi o datos móviles.
- Si la red WiFi que está utilizando es una red pública (como un café) o privada.
- Si el usuario se movió durante la solicitación.
- Si el usuario intentó utilizar varios dispositivos.
- Si más de un usuario intentó con el mismo dispositivo.

Todos estos datos pueden utilizarse para caracterizar préstamos ya otorgados e intentar predecir el resultado de un nuevo préstamo. Si la identificación fuese posible, sería posible identificar préstamos con más probabilidades de ser fraudulentos y reducir de esa forma las pérdidas.

En un análisis preliminar, se detectó que varios de los préstamos que resultaban enviados a pérdida sin pagos eran realizados desde las cercanías de una plaza que frecuentemente era un punto de venta para documentos falsificados.

Utilización de sistemas de OCR (Reconocimiento óptico de caracteres) para validar documentos y documentoscopía

Actualmente no se utiliza ningún tipo de validación sobre los documentos presentados por los usuarios. Si bien conseguir documentos falsificados no es difícil en Brasil, la utilización de reconocimiento de texto para validar los datos ingresados por el usuario podría disuadir los casos más simples de fraude (aquellos que no intentan obtener un documento con los datos del usuario que intentan impersonar).

Como ventaja adicional, un sistema eficiente de OCR debería simplificar el proceso de alta del usuario, pudiendo tomar los datos del documento en lugar de requerir que sean introducidos manualmente.

Adicionalmente, existen servicios de documentoscopía que permiten validar que el documento sea auténtico. Estos servicios verifican ciertos detalles que son más difíciles de falsificar, además de que los datos del documento correspondan al individuo. Entre estos detalles se encuentran:

- Que el oficial firmante haya estado asignado al momento de emitir.
- Que el número del documento cumpla con el formato del estado emisor al momento de ser emitido.
- Que el escudo de armas utilizado de fondo corresponda al estado emisor al momento de ser emitido.

Figura 32: Documento falso. En este caso, el error está en el número de Registro Geral (RG) donde el defraudador colocó un dígito verificador (pero los emitidos en Pernambuco no llevan)[37]

Sin embargo, la utilización de estos elevaría considerablemente el costo de emisión, además de aumentar el tiempo promedio entre la finalización de la solicitud del préstamo y el depósito del dinero. No se realizó un estudio de backtesting que pudiese validar el impacto que podría tener dicho servicio.

[37] Alertas de Fraude, ACEV - Associação Comercial e Empressarial de Votorantim, http://www.acev.com.br/alertas-de-fraude, recuperado 13/10/20

Bibliografía

National Institute of Justice (U.S.). The Fingerprint Sourcebook. Washington, DC: U.S. Dept. of Justice, Office of Justice Programs, National Institute of Justice, 2011.

Aubert, M., Brumm, A., Ramli, M. et al. Pleistocene cave art from Sulawesi, Indonesia. Nature 514, 223–227 (2014).

Browne, Douglas G., Alan Brock, Fingerprints: Fifty Years of Scientific Crime Detection, George G. Harrap & Co., Ltd., London, 1953, 105–106

Mohammad S. Obaidat, Issa Traore, Isaac Woungang, Biometric-Based Physical and Cybersecurity Systems, Springer, 2018, 45-46

IBM Introducing Fingerprint Reader into Laptop, TechNewsWorld, 2004 https://www.technewsworld.com/story/37017.html, recuperado 13/10/20

Researcher Hacks Microsoft Fingerprint Reader, PCWorld, 2006 https://www.pcworld.com/article/124978/article.html, recuperado 13/10/20

Understanding biometrics: How to choose the right biometric technology for your organisation, Argus Global https://www.planetbiometrics.com/creo_files/upload/article-files/how_to_choose_the_right_biometric.pdf

In Hong Kong Protests, Faces Become Weapons, New York Times, 2019, https://www.nytimes.com/2019/07/26/technology/hong-kong-protests-facial-recognition-surveillance.html, recuperado 13/10/20

Global Fintech Investments Surged in 2018 with Investments in China Taking the Lead, Accenture Analysis Finds; UK Gains Sharply Despite Brexit Doubts, Accenture, 2019 https://newsroom.accenture.com/news/global-fintech-investments-surged-in-2018-with-

investments-in-china-taking-the-lead-accenture-analysis-finds-uk-gains-sharply-despite-brexit-doubts.htm, recuperado 13/10/20

The State Of Fintech: Investment & Sector Trends To Watch 2019Q4 Report, CBInsights, 2019 https://www.cbinsights.com/research/report/fintech-trends-q4-2019/, recuperado 13/10/20

Personal Loan Market Overview, Transunion, 2019 https://www.transunion.com/resources/transunion/doc/insights/articles/tu-personal-loan-market-2019.pdf

Survey: Consumers Want Personal Loans for Large Purchases and Debt Consolidation, Experian, 2019 https://www.experian.com/blogs/ask-experian/survey-consumers-want-personal-loans-for-large-purchases-and-debt-consolidation/, recuperado 13/10/20

EMPRÉSTIMO CONSIGNADO: QUANTO TEMPO DEMORA PARA CAIR NA CONTA?, Credito Folha, 2019 https://creditofolha.com/emprestimo-consignado-quanto-tempo-demora-para-cair-na-conta/, recuperado 13/10/20

Em quanto tempo ocorre a liberação do Empréstimo Consignado?, QualiConsig, 2018 https://qualiconsig.com.br/emprestimo-consignado-demora-pra-cair-na-conta/, recuperado 13/10/20

Perguntas frequentes, Simplic https://www.simplic.com.br/faq, recuperado 13/10/20

A cada 16 segundos, uma tentativa de fraude acontece no Brasil, revela Serasa, Serasa, 2018 https://www.serasaexperian.com.br/sala-de-imprensa/a-cada-16-segundos-uma-tentativa-de-fraude-acontece-no-brasil-revela-serasa, recuperado 13/10/20

Saiba o que é fraude e quais os tipos mais comuns, SerasaConsumidor, https://www.serasaconsumidor.com.br/ensina/seu-cpf-protegido/o-que-e-fraude/, recuperado 13/10/20

Alliance appeals to Council of Europe to address biometrics privacy, Privacy International, 2011 https://privacyinternational.org/blog/1580/alliance-appeals-council-europe-address-biometrics-privacy, recuperado 13/10/20

Opposition grows to storage of photo and biometric data, The Sydney Morning Herald, 2014 https://www.smh.com.au/politics/federal/opposition-grows-to-storage-of-photo-and-biometric-data-20141015-116lur.html, recuperado 13/10/20

French privacy row over mass ID database, BBC, 2016 https://www.bbc.com/news/37894968, recuperado 13/10/20

Despite public concerns, facial recognition gets traction in Congress, Federal Computer Week, 2020 https://fcw.com/articles/2020/02/09/congress-facial-recognition.aspx, recuperado 13/10/20

How Facial Recognition Systems Work, How Stuff Works https://electronics.howstuffworks.com/gadgets/high-tech-gadgets/facial-recognition.htm/printable, recuperado 13/10/20

Meet Zoom, Zoom https://www.zoomlogin.com/#page-blk-meet-zoom, recuperado 13/10/20

What is GDPR?, idStation https://www.idstation.eu/Home/GDPR, recuperado 13/10/20

Pato, Joseph N. & Millett, Lynette I., Biometric Recognition: Challenges and Opportunities, NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES, THE NATIONAL ACADEMIES PRESS 2010 96-98

App Ranking Factors » How to Improve App Store Search Rankings, AppRadar, 2019 https://appradar.com/academy/bonus-chapters/app-store-ranking-factors/, recuperado 13/10/20

Paper Beats Digital In Many Ways, According To Neuroscience, Forbes, 2015 https://www.forbes.com/sites/rogerdooley/2015/09/16/paper-vs-digital/, recuperado 13/10/20

FRAUDES FINANCEIRAS NO BRASIL, Confederação Nacional de Dirigentes Lojistas, 2019 http://www.cndl.org.br/upload/PP40/materiais/pesquisas/Fraudes%20Financeiras/1/SPC%20Analise%20Fraudes%20Financeiras%20no%20Brasil.pdf

Falsificação de documento representa 75% das fraudes registradas em MG, Jornal Hoje, 2014 http://g1.globo.com/jornal-hoje/noticia/2014/02/falsificacao-de-documento-representa-75-das-fraudes-registradas-em-mg.html, recuperado 13/10/20

Itaú Unibanco: Banking the unbanked in Brazil, Itaú, 2019 https://www.businesscalltoaction.org/wp-content/files_mf/bcta_casestudy_itau_web.pdf

40% das pessoas que pedem crédito consignado usam dinheiro para pagar dívidas, diz pesquisa, Globo Economía, 2017 https://g1.globo.com/economia/seu-dinheiro/noticia/40-das-pessoas-que-pedem-credito-consignado-usam-dinheiro-para-pagar-dividas-diz-pesquisa.ghtml, recuperado 13/10/20

Distribution of smartphone shipments share in Brazil in 1st quarter 2017 and 2018, by brand, Statista, 2018 https://www.statista.com/statistics/693317/smartphone-market-share-vendor-in-brazil/, recuperado 13/10/20

Alertas de Fraude, ACEV - Associação Comercial e Empressarial de Votorantim, http://www.acev.com.br/alertas-de-fraude, recuperado 13/10/20

Anexos

Se anexan dos tablas con la datos utilizados y anonimizados de los rendimientos de los préstamos.

La tabla Anexo 1 contiene los usuarios validados a través del proceso de backtesting y su condición de First Payment Default para sus préstamos. Los datos de esta tabla están reflejados en los gráficos de la sección *Backtesting*.

La tabla Anexo 2 contiene los usuarios provenientes luego de la implementación del sistema de validación biométrica, sus respectivos puntajes en cada proveedor y el resultado de sus préstamos. Esta tabla también incluye una columna para indicar si fue aprobado o rechazado.