

Delfau, Emiliano

**Working Paper**

## Propuestas para la bancarización e inclusión financiera en Argentina

Serie Documentos de Trabajo, No. 716

**Provided in Cooperation with:**

University of CEMA, Buenos Aires

*Suggested Citation:* Delfau, Emiliano (2020) : Propuestas para la bancarización e inclusión financiera en Argentina, Serie Documentos de Trabajo, No. 716, Universidad del Centro de Estudios Macroeconómicos de Argentina (UCEMA), Buenos Aires

This Version is available at:

<https://hdl.handle.net/10419/238341>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**UNIVERSIDAD DEL CEMA  
Buenos Aires  
Argentina**

Serie  
**DOCUMENTOS DE TRABAJO**

**Área: Finanzas**

**PROPUESTAS PARA LA BANCARIZACIÓN E  
INCLUSIÓN FINANCIERA EN ARGENTINA**

**Emiliano Delfau**

**Febrero 2020  
Nro. 716**

**[www.cema.edu.ar/publicaciones/doc\\_trabajo.html](http://www.cema.edu.ar/publicaciones/doc_trabajo.html)  
UCEMA: Av. Córdoba 374, C1054AAP Buenos Aires, Argentina  
ISSN 1668-4575 (impreso), ISSN 1668-4583 (en línea)  
Editor: Jorge M. Streb; asistente editorial: Valeria Dowding <jae@cema.edu.ar>**



# Propuestas para la Bancarización e Inclusión Financiera en Argentina<sup>1</sup>

Emiliano Delfau<sup>2</sup>

Universidad del CEMA,  
PhD en Finanzas,  
Argentina  
Febrero 2020

## Resumen

Es sabido que en Argentina, contrariamente a la región, el crédito interno sobre PBI históricamente se encontró en niveles bajos. Analizando los últimos diez años podemos ver que el país fluctuó siempre entre un 12 y 15% de bancarización respecto al PBI (Banco Mundial, 2017). Asimismo, encontramos que el sector micro informal<sup>3</sup> de la estructura productiva alcanzó el 49,3% hacia fines de 2018 (UCA, 2018).

Esto nos muestra que el desafío no es solo bancarizar a la población no bancarizada sino, además, lograr bancarizar a parte de la población del sector micro informal. De esta manera estaríamos no solo abordando un proyecto de inclusión financiera, sino que asimismo trataríamos de minimizar el uso de efectivo por otros medios de pago y transacciones.

No obstante las características mencionadas anteriormente, la Argentina si se encuentra dentro de las tendencias tecnológicas mundiales y, por lo tanto, el país cuenta con un "ecosistema" tecnológico que le permitiría afrontar los desafíos antes mencionados. Bajo la premisa de "*todo dato es dato crediticio*" se plantea la creación de una plataforma o banco con abordaje 100% digital cuyo motor principal sea un *score* de crédito basado en análisis de *Big Data* y técnicas de *Machine Learning*.

Finalmente se enumeran algunos casos de éxito sobre este nuevo modelo de negocio mediante la aplicación de *Big Data* y técnicas de *Machine Learning*

**Palabras Claves:** Inclusión Financiera, Big Data, Fintechs, Datos no Estructurados, Analytics, Scoring, Recnología Digital, Banca Digital

---

<sup>1</sup> El presente trabajo participó del concurso académico "PREMIO ADEBA 2019" organizado por la Asociación de Bancos Argentinos ("ADEBA"), CUIT 30-70831616-0 bajo el tema "Propuestas para la bancarización e inclusión financiera en Argentina " obteniendo el 2do lugar.

<sup>2</sup> Las opiniones vertidas en este trabajo son personales del autor y no reflejan necesariamente los puntos de vista de la UCEMA o CEBaFi. Comentarios son bienvenidos en: emiliano.delfau@hotmail.com

<sup>3</sup> Actividades laborales autónomas no profesionales o llevadas a cabo en pequeñas unidades productivas de baja productividad, alta rotación y baja o nula vinculación con el mercado formal.

## Tabla de contenido

<b>BANCARIZACIÓN E INCLUSIÓN FINANCIERA EN ARGENTINA.....</b>	<b>3</b>
INTRODUCCIÓN .....	3
SITUACIÓN ACTUAL Y EVOLUCIÓN HISTÓRICA .....	3
POBLACIÓN OBJETIVO.....	4
SERVICIOS FINANCIEROS.....	6
SERVICIOS FINANCIEROS Y TRANSFORMACIÓN ELECTRÓNICA .....	7
ACCESIBILIDAD.....	9
ANÁLISIS DE PREFERENCIAS DEL CONSUMIDOR.....	11
DEMOGRAFÍA DIGITAL: ENTORNO TECNOLÓGICO Y DE COMUNICACIÓN .....	12
<b>MODELOS TRADICIONALES DE SCORE VERSUS MODELOS BASADOS EN BIG-DATA .....</b>	<b>13</b>
EL ECOSISTEMA EN <i>MACHINE LEARNING</i> Y <i>BIG-DATA</i> .....	15
TERMINOLOGÍAS Y CONCEPTOS .....	16
ANTECEDENTES EN <i>BIG-DATA</i> .....	17
<b><i>MACHINE LEARNING</i> PARA EL DESARROLLO DE UN MODELO SCORE BASADO EN <i>BIG-DATA</i> .....</b>	<b>18</b>
EL PROCESO DE <i>ANALYTICS</i> EN <i>BIG-DATA</i> .....	18
Recolección de Datos .....	19
Agregación de Datos.....	20
Análisis de Datos.....	20
DISEÑO FUNCIONAL: CASO PRÁCTICO .....	20
Definición del problema y determinación de la variable objetivo .....	22
Recopilación y transformación de datos .....	22
Modelo final.....	23
<b>OPORTUNIDADES PARA LA INCLUSIÓN FINANCIERA A PARTIR DE LA TECNOLOGÍA DIGITAL, TÉCNICAS DE <i>MACHINE LEARNING</i> Y <i>BIG-DATA</i>.....</b>	<b>24</b>
EL PANORAMA DE LOS PAGOS DIGITALES. ....	25
APROVECHANDO LA TECNOLOGÍA DIGITAL ENTRE LOS NO BANCARIZADOS.....	26
OPORTUNIDADES PARA AMPLIAR LA APERTURA DE CUENTAS ENTRE LOS NO BANCARIZADOS .....	28
AUMENTO DE UTILIZACIÓN DE PRODUCTOS FINANCIEROS Y DISMINUCIÓN DE EFECTIVO.....	28
POTENCIAR LOS PDA A NIVEL NACIONAL .....	29
PERMITIR O FACILITAR EL ACCESO AL CRÉDITO A AQUELLAS PERSONAS QUE SE ENCUENTRAN EXCLUIDAS.....	30
LA IMPORTANCIA DE TENER UN SCORE DE CRÉDITO .....	30
<b>ANEXO: CASOS DE ESTUDIO EXITOSOS EN <i>BIG-DATA</i>.....</b>	<b>31</b>
INDIA 1 .....	31
INDIA 2 .....	32
ALEMANIA .....	32
USA .....	33
HONG KONG .....	33
<b>REFERENCIAS .....</b>	<b>34</b>

## Bancarización e Inclusión Financiera en Argentina

### Introducción

Como regla generalizada podríamos decir que, para cualquier país del mundo, independientemente de su grado de desarrollo, el acceso al crédito es un requerimiento esencial para el crecimiento, la movilidad social y el éxito financiero de sus ciudadanos.

Cuando hablamos del grado o nivel de bancarización de un país respecto a su Producto Bruto Interno estamos manifestando implícitamente la siguiente pregunta: ¿cuál es el grado de penetración que tiene el sistema de *scoring* (es decir un buen puntaje o *rating*) de crédito de un país para la sociedad al momento de que cualquier ciudadano desee adquirir o alquilar una vivienda, comprar un auto, acceder a empleos o comenzar un nuevo negocio, acceder a niveles de educación superior o alcanzar otras metas superadoras?.

### Situación Actual y Evolución Histórica

Si realizamos un análisis estático comparativo del Crédito Interno al Sector Privado como porcentaje del PIB de Argentina respecto a respecto a otros países del mundo (desarrollados o en desarrollo) podemos ver que la Argentina ocupa, lamentablemente, el primer lugar en lo que respecta al bajo nivel de bancarización como porcentaje del PBI (Banco Mundial, 2017)<sup>4</sup>.

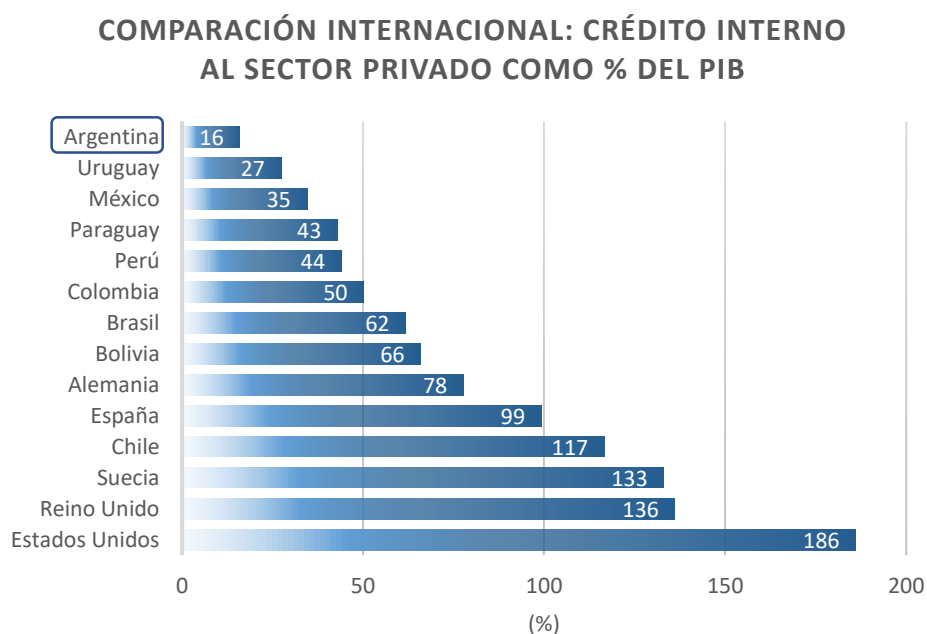


Figura 1

Es interesante destacar que el análisis estático comparativo realizado anteriormente es un problema que persiste en la Argentina desde siempre. A modo de ejemplo en la

<sup>4</sup> Argentina información al 2017, resto de los países información al 2018

Figura 2 analizamos la evolución del Crédito Interno al Sector Privado de Argentina y el Mundo (incluyendo Argentina). Podemos ver que en los últimos 57 años Argentina estuvo siempre muy rezagada respecto a nivel y la evolución mundial. Si además observamos el último período de 10 años podemos ver que el país fluctuó siempre entre un 12 y 15 por ciento de bancarización respecto al PBI (Banco Mundial, 2017). Si por otra parte analizamos la evolución del Mundo, podemos ver que este indicador crece sostenidamente de un nivel inicial de 54 por ciento hasta alcanzar una relación de casi 129 por ciento para el año 2018. Es importante destacar la aceleración que ocurre a principios de los 80'. A partir del 70' se introducen los primeros modelos de *score* de crédito, luego a partir de los 80' y 90' gracias a la tecnología esta técnica de calificación crediticia se generalizó<sup>5</sup>.

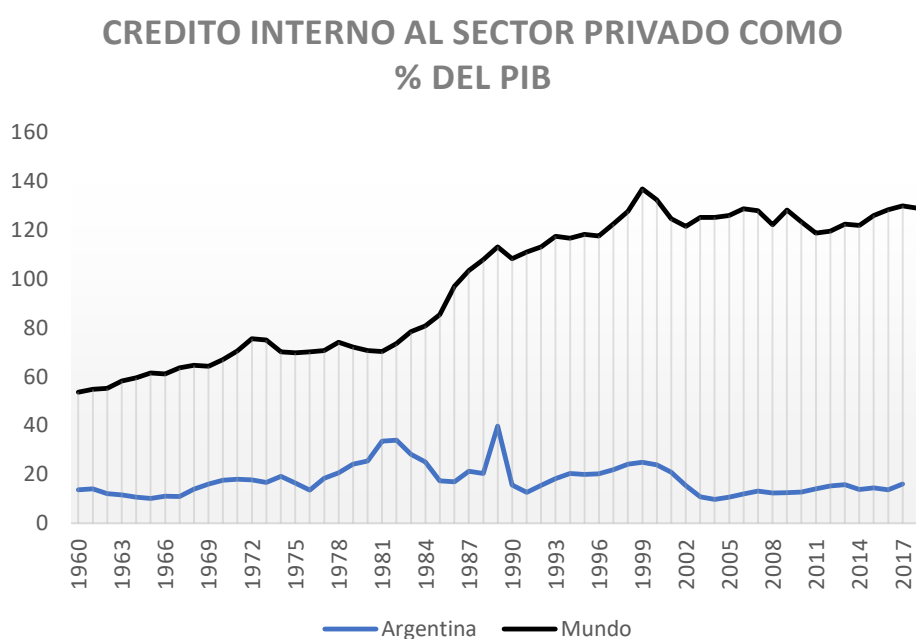


Figura 2

Si bien existen aspectos “tecnológicos” que ayudan al ecosistema del crédito para alcanzar su desarrollo. Lo cierto es que, en contraposición, la Argentina se encontró siempre con una evolución estancada respecto a dicho índice mundial. Podemos, en términos generales, señalar dos causas principales que a priori explican gran parte del comportamiento que hace que dicho nivel se mantenga históricamente bajo. El primer motivo es el alto grado de inestabilidad económica y financiera que sufre el país de manera recurrente. El segundo motivo se haya en el hecho de que la Argentina tiene un alto nivel de impuestos y/o comisiones sobre transacciones bancarias que, en su conjunto, alientan a la informalidad y al uso de efectivo.

### Población Objetivo

Entrando más en detalle respecto a la población objetivo podemos mencionar que la Población Económicamente Activa (PEA) al primer trimestre de 2019 era de 12.989.801

<sup>5</sup> <http://www.bcra.gov.ar/pdfs/investigaciones/Credit%20Scoring%20%28sum%29.pdf>

personas (INDEC, 2019). Si a esa misma fecha consideramos el porcentaje de desocupación del 9,1 por ciento, podemos decir que tenemos una potencialidad de crédito que alcanza a 11,8 millones de personas.

Asimismo, la población de ocupados en el sector micro-informal<sup>6</sup> de la estructura productiva alcanzó el 49,3 por ciento hacia fines de 2018 (UCA, 2018). Esto nos muestra que el desafío no es solo bancarizar a la población no bancarizada, sino que, además lograr bancarizar a parte de la población del sector micro-informal. De esta manera estaríamos no solo abordando un proyecto de inclusión financiera, sino que además trataríamos de minimizar el uso de efectivo por otros medios de pago y transacciones. La Figura 3 a continuación muestra esto.

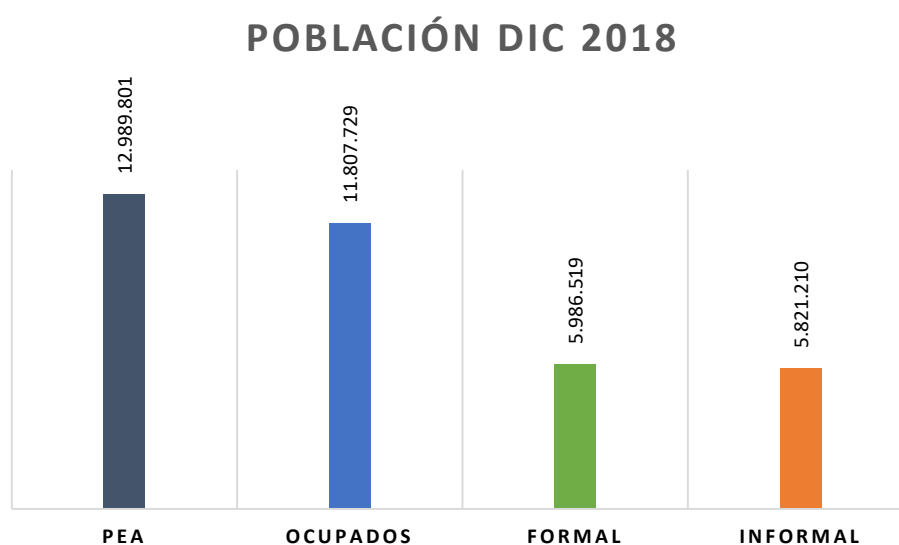


Figura 3

Del análisis anterior podemos establecer que aproximadamente 5 millones de ciudadanos, y también consumidores, en el país seguramente no posean historial crediticio alguno o carezcan de historial crediticio suficiente para generar un *score* de crédito, situación la cual impide el acceso a los servicios bancarios tradicionales. Es por eso que encontrar una forma de obtener acceso asequible al crédito es de vital importancia no sólo para el bienestar económico de esta población sino porque además, también representa un mercado sin explotar con el potencial de grandes ganancias.

Debido a las condiciones actuales del país y a las “ineficiencias” del sistema financiero tradicional en conjunto con los *Bureau* de crédito no es sorprendente que en esta era de *Big-Data*, donde la información obtenida a partir de las búsquedas en Internet, las redes sociales y las aplicaciones móviles, esta población excluida sea cada día más explotada a modo de lograr alcanzar bancarizar la misma.

---

<sup>6</sup> Actividades laborales autónomas no profesionales o llevadas a cabo en pequeñas unidades productivas de baja productividad, alta rotación y baja o nula vinculación con el mercado formal.



Particularmente podemos destacar que desde el año 2008/2009 a nivel global los prestamistas existentes y los nuevos (*Fintechs*) se encuentran intensificando el uso de técnicas de armado de perfiles del consumidor mediante estas técnicas que denominamos de “*Big-Data*”. Tal como desarrollaremos a continuación, este movimiento masivo viene explicado principalmente por el “facilitamiento” del acceso a los *smartphones*, la utilización de redes sociales y medios electrónicos de pago. Fenómenos los cuales, a su vez, acentúan cada vez más la “huella digital”<sup>7</sup> que los consumidores van dejando y son explotados por los prestamistas no solo para conseguir un mejor perfil del cliente sino para predecir el comportamiento y las preferencias del consumidor.

En lo que respecta al proceso de calificación crediticia, la tendencia más prometedora es la inclusión de datos alternativos convencionales en los archivos de crédito. Esta práctica podría mejorar el acceso a nivel nacional a las oportunidades financieras al hacer que más consumidores sean "calificables" dentro del sistema financiero convencional. Cada día es más evidente el planteo que se realiza sobre los modelos de *scoring* tradicionales. Existe un cierto grado de preocupación de que estas herramientas de calificación, en un punto, puedan perjudicar de manera injustificada a ciertos consumidores. Un estudio realizado en Estados Unidos muestra que en ese país aproximadamente el 20 por ciento de la población son clasificados como "no calificables", lo que se traduce en un no acceso al sistema tradicional de crédito. La situación mencionada anteriormente es extrapolable a la mayor parte del mundo y la Argentina no es la excepción. Particularmente podemos decir que si en Argentina tenemos aproximadamente 5 millones de ciudadanos categorizados como PEA en el sector micro-informal el porcentaje de personas “no calificables” ascendería al 45 por ciento.

Cabe señalar que si bien existe una tendencia creciente principalmente impulsadas por las *Fintechs* a la explotación de *Big-Data* para desarrollar nuevos modelos de negocio a estratos poblacionales hoy no atendidos. También corresponde realizar un análisis detallado si este camino será realmente más beneficioso para el consumidor de bajos ingresos. Estos productos pueden llenar un vacío y proporcionar acceso asequible al crédito a estas poblaciones desatendidas o pueden ser un medio de aprovecharse de las comunidades vulnerables.

## **Servicios Financieros**

En los últimos años la Argentina ha avanzado en lo que respecta al ámbito del acceso a cuentas de ahorro y medios de pago electrónicos mediante la acreditación de salarios, jubilaciones, programas sociales, etc. A la fecha el 80 por ciento de la población activa posee una caja de ahorro y una tarjeta de débito asociada. No obstante este logro, el

---

<sup>7</sup> El término “huella digital” o a veces “sombra digital” se refiere al conjunto único de actividades, acciones, contribuciones y comunicaciones digitales rastreables, que se manifiestan en Internet o en dispositivos digitales. Es decir, es la información que queda como resultado de la navegación web de un usuario y se almacena de diferentes formas. Podemos encontrar dos tipos de “huella digital” una denominada “activa” que se crean de manera deliberada por el usuario y otras denominadas “pasivas” que son recopiladas sin que el usuario tenga conocimiento.

nivel de utilización es bajo, dado que los medios de pagos más frecuentes continúan siendo en efectivo. Según estudios realizados, el 51 por ciento de la Población Económicamente Activa con empleo cobran su sueldo en efectivo<sup>8</sup>, este porcentaje queda explicado por el sector micro-informal mencionado en el apartado anterior (ver Figura 3).

Asimismo, es interesante analizar el hábito sobre la población que cobra sus haberes por medio de depósito bancario en lo que respecta a la extracción de dinero y a la utilización de canales electrónicos para pagos. Del trabajo realizado por el BCRA-CAF en 2017<sup>9</sup> se puede apreciar que solo un 24 por ciento utiliza el saldo en la cuenta bancaria para realizar compras y pagos con la tarjeta de débito, crédito y o transferencia electrónica. También existe un segmento importante (alcanza al 21 por ciento de la población) que extrae el 100 por ciento de su depósito. A este último segmento se lo señala como “semi-bancarizado” y es interesante notar que sobre este grupo un 76 por ciento, dice que retira todo su depósito porque prefiere hacer los pagos en efectivo, mientras que un 8 y 7 por ciento afirman que el motivo del retiro de depósito corresponde ya sea a: lejanía de la sucursal o por que consideran que es una pérdida de tiempo ir a las sucursales, respectivamente. Finalmente, en este análisis de comportamiento podemos encontrar el punto más crítico que es que tan solo un 4 por ciento de la población afirma dejar dinero en la cuenta para ahorro, inversión o para realizar pagos electrónicos. La Figura 4 muestra este comportamiento.

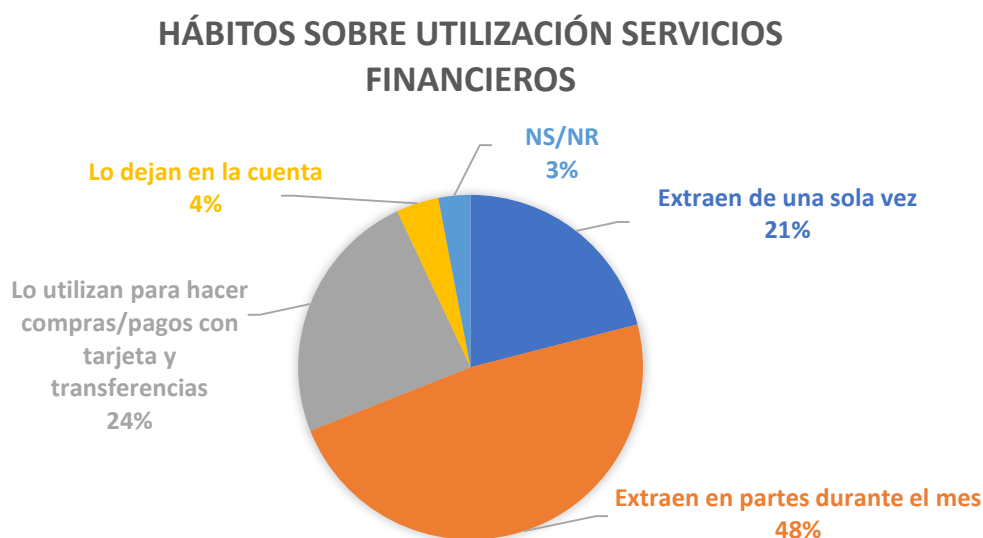


Figura 4

### Servicios Financieros y Transformación Electrónica

Respecto a la utilización de productos financieros podemos destacar que las compras en comercios con tarjeta de débito y crédito en los últimos 2 años<sup>10</sup> (período de análisis

<sup>8</sup> BCRA-CAF: Encuesta de medición de capacidades financieras en Argentina, 2017

<sup>9</sup> Ídem anterior.

<sup>10</sup> Ídem anterior.

2017) alcanzaron un 43 y 34 por ciento respectivamente. Cabe aclarar que el 34 por ciento de las compras con tarjeta de crédito corresponden a transacciones realizadas en un pago, cuando analizamos las compras en cuotas este indicador alcanza un 46 por ciento. Por otra parte, podemos destacar que los depósitos en cuenta fueron de tan solo de un 23 por ciento y las compras de moneda extranjera alcanzó un nivel de 11 por ciento.

Entrando en el universo “digital” podemos destacar que menos de un quinto de las personas (16 por ciento) realizó pagos de servicios por *HomeBanking* o mediante alguna plataforma de telefonía móvil, mientras que sólo el 13 por ciento realizó alguna transferencia electrónica, es decir, por *HomeBanking* o alguna plataforma de telefonía móvil. Es interesante destacar que si uno apertura este 13 por ciento encuentra que en las zonas rurales este indicador llega tan sólo a un 7 por ciento<sup>11</sup>. Por otra parte, tan sólo un 20 por ciento de la muestra realizó alguna compra *on-line* con tarjeta de crédito. Finalmente podemos encontrar que un 25 por ciento pagó algún servicio por débito automático.

La Figura 5 de abajo ilustra el comportamiento antes mencionado. Queda, además, en evidencia que existe una potencialidad enorme en lo que respecta al grado de utilización de los servicios financieros en canales electrónicos (como *HomeBanking*). Por lo que creemos que se debe llevar a cabo una estrategia integral que fomente el uso activo de estas cuentas en función del ahorro, para gestionar el riesgo, y para pagos y cobros. Luego desarrollaremos algunas ideas respecto a como abordar el grado de utilización de dichos servicios financieros.

### COMPORTAMIENTO SERVICIÓN FINANCIEROS (ÚLTIMOS 2 AÑOS)

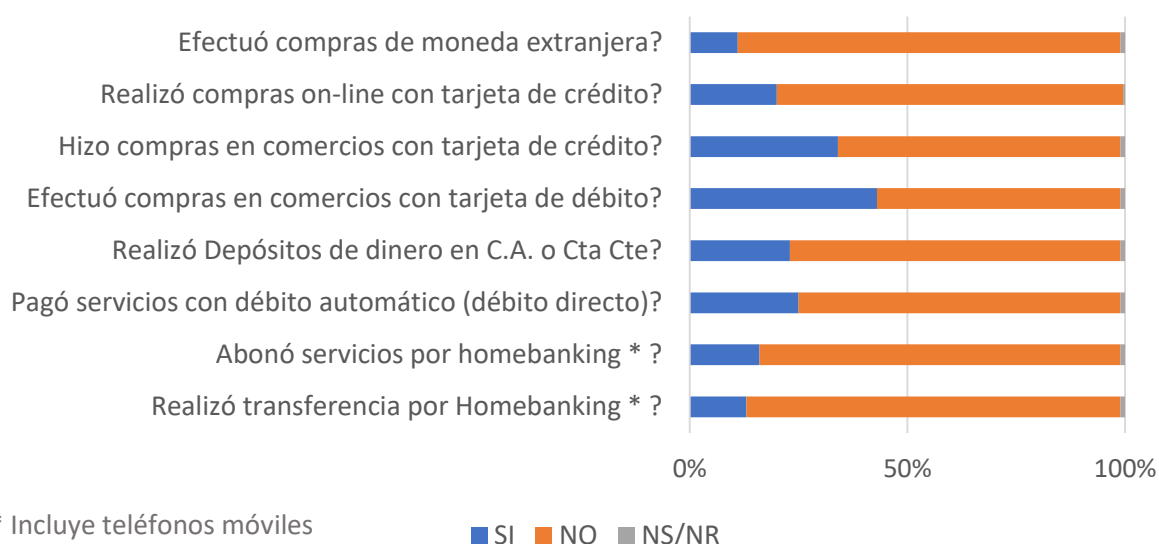


Figura 5

<sup>11</sup> Fuente: BCRA-CAF. Encuesta de medición de capacidades financieras en Argentina (2017)

## Accesibilidad

El Informe de Estabilidad Financiera del Segundo Semestre 2017<sup>12</sup> detalla en el recuadro 7 que: “El Banco Central, en el marco de su finalidad de promover el desarrollo económico con equidad social, tiene como uno de sus objetivos principales fomentar la “inclusión financiera” de la población; es decir, ampliar los niveles de acceso y uso a los servicios financieros. Para ello, como punto de partida es necesario realizar un análisis del estado de la inclusión financiera en el país a través del desarrollo de métricas específicas para sus diferentes componentes (de acceso, uso y calidad).”

Particularmente si observamos los puntos de acceso que ofrece el sistema financiero actualmente (puntos de acceso hacen referencia a puntos de depósito y extracción de efectivo como sucursales o cajeros automáticos), éstos se encuentran por debajo de la región. Tal como muestra la Figura 6 más abajo, la Argentina presenta tan solo 7,9 PDA por cada 10,000 adultos.

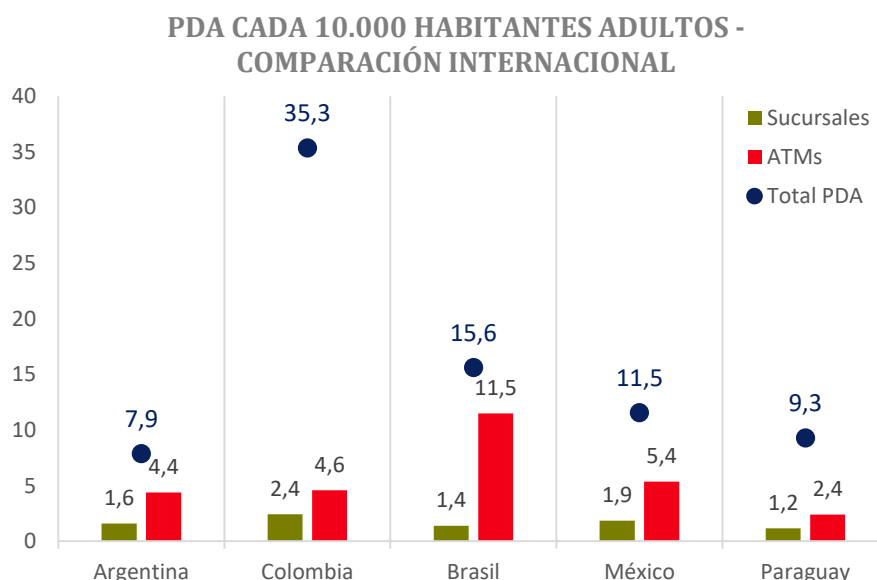


Figura 6

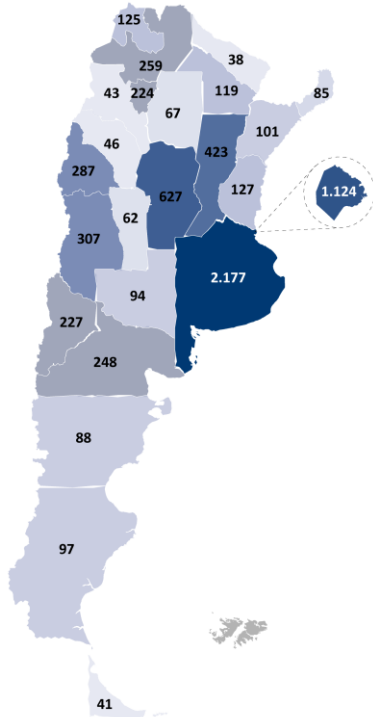
Si nos adentramos a analizar la distribución geográfica del país, los PDA (sucursales bancarias, cajeros automáticos o ATM, terminales de autoservicio y las dependencias automatizadas) del sistema financiero alcanzaba un total de 26.037 a nivel nacional hacia 2017. Lo interesante respecto a esta capacidad es que de las 23 provincias que conforman al país casi la mitad del total de PDA se encontraba en dos jurisdicciones: Provincia de Buenos Aires y CABA lo que evidencia marcadas diferencias en su distribución. Dicha situación se presenta en la Figura 7<sup>13</sup> y en la Tabla 1<sup>14</sup> a continuación.

<sup>12</sup> Fuente: [https://www.bcra.gob.ar/PublicacionesEstadisticas/IEF\\_0217.asp](https://www.bcra.gob.ar/PublicacionesEstadisticas/IEF_0217.asp)

<sup>13</sup> Fuente: [https://ideasdepeso.com/2018/07/06/mas-puntos-de-acceso-menos-efectivo/#\\_ftn9](https://ideasdepeso.com/2018/07/06/mas-puntos-de-acceso-menos-efectivo/#_ftn9)

<sup>14</sup> Fuente: [https://www.bcra.gob.ar/PublicacionesEstadisticas/IEF\\_0217.asp](https://www.bcra.gob.ar/PublicacionesEstadisticas/IEF_0217.asp)

**Puntos de extracción**



**Sucursales bancarias + ATM**

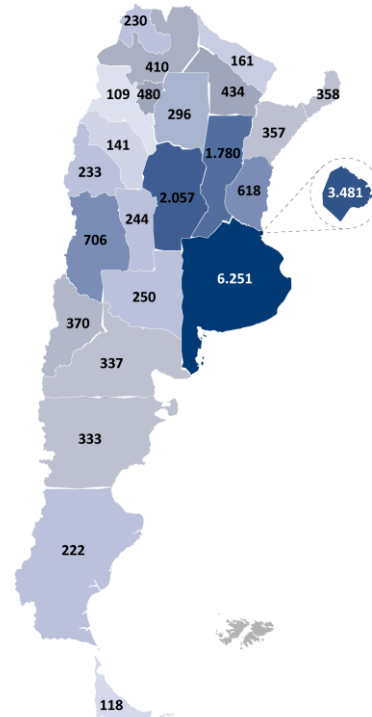


Figura 7

Provincia	% PDA	% población	% superficie
Buenos Aires	30,3%	38,7%	11,3%
CABA	19,1%	8,6%	0,0%
Santa Fe	10,2%	8,0%	4,9%
Córdoba	9,7%	8,3%	6,0%
Mendoza	3,6%	4,3%	5,4%
Entre Ríos	3,1%	3,0%	2,9%
Tucumán	2,4%	3,5%	0,8%
Salta	2,0%	2,9%	5,8%
Chaco	1,8%	2,5%	3,6%
Misiones	1,8%	2,5%	1,1%
Neuquén	1,7%	1,4%	3,4%
Río Negro	1,7%	1,6%	8,0%
Chubut	1,6%	1,3%	7,5%
Corrientes	1,6%	2,3%	3,2%
Santiago del Estero	1,3%	2,0%	5,0%
San Luis	1,2%	1,1%	2,8%
San Juan	1,2%	1,6%	3,2%
Jujuy	1,2%	1,6%	2,0%
La Pampa	1,2%	0,8%	4,5%
Santa Cruz	1,0%	0,7%	8,6%
Formosa	0,7%	1,2%	2,6%
Tierra del Fuego	0,6%	0,4%	0,8%
La Rioja	0,6%	0,8%	3,2%
Catamarca	0,5%	0,9%	3,3%

Tabla 1

Finalmente, otro factor de análisis importante, es la distribución del uso de productos de crédito. Como hemos mencionado anteriormente aproximadamente el 50 por ciento de la población adulta posee algún producto. Lo que es interesante destacar es que, al igual que lo que ocurre con los PDA es que existe una evidenciada concentración principalmente en CABA donde el 82 por ciento de la población adulta posee un producto financiero (en entidad registrada) mientras que en el resto de las provincias dicho porcentaje se reduce a la mitad. La siguiente Figura 8 a continuación ilustra esto<sup>15</sup>.

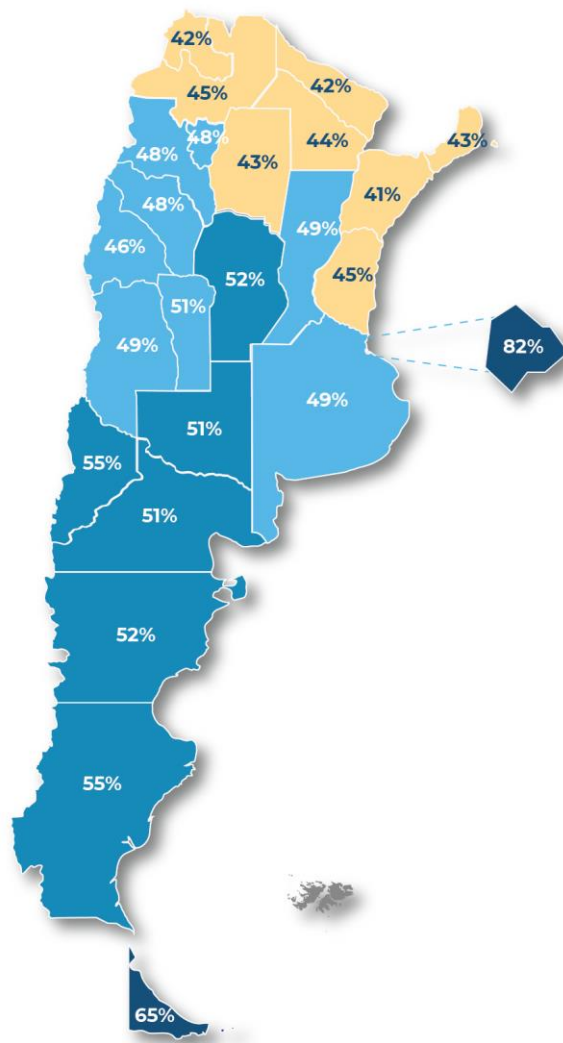


Figura 8

### Análisis de Preferencias del Consumidor

No obstante la evolución señalada anteriormente respecto al bajo nivel de bancarización del país, existen otras mega-tendencias sobre las cuales Argentina no es la excepción respecto al resto del mundo. La primer mega-tendencia que no le es ajena al país es el cambio generacional y la evolución sobre la percepción de la “banca

<sup>15</sup> Fuente: BCRA e INDEC. Estrategia Nacional de Inclusión Financiera (2018).

tradicional” y el consumidor financiero en la nueva era digital. Según estudios realizados (CGI Group Inc, 2014) los consumidores presentan ciertas cualidades sobre las cuales la banca tradicional se ve forzada a cambiar, principalmente en aspecto de atención y ofertas especializadas y accesibilidad. La Tabla 2 a continuación resume el “grado de satisfacción” del consumidor respecto a su banco operativo:

<b>Demanda de los consumidores</b>	<b>Grado de satisfacción</b>
Los recompensen por ser clientes	34%
Les den acceso a sus saldos en cualquier momento y lugar	84%
Los vean como una persona	55%
Les ofrezcan consejos para invertir	37%
Puedan consultar gastos y reciban consejos sobre ahorros	34%
Tengan acceso a asesores independientes cuando los necesite	50%
Acceso al crédito hasta límite acordado en cualquier momento	49%
Pagar bienes y servicios mediante cualquier dispositivo	49%

Tabla 2

Otro punto importante de la encuesta es que aproximadamente el 60 por ciento de la población prefiere que los servicios sean digitales.

### **Demografía Digital: entorno Tecnológico y de Comunicación**

La otra mega-tendencia que Argentina también comparte es el nuevo ecosistema de telefonía móvil que actualmente rige en el planeta. Argentina, con 61 millones de conexiones móviles, es el tercer mercado de telefonía móvil más grande de América Latina. La penetración de suscriptores únicos (es decir, por usuario humano) tiene una madurez del 90 por ciento, más alta que Europa (85 por ciento), EE. UU. (80 por ciento) y el promedio regional (68 por ciento). Por lo tanto, las necesidades actuales no se encuentran en el hecho de agregar nuevos suscriptores, sino en que los usuarios existentes de teléfonos inteligentes migren de 3G y 4G. Asimismo se espera que para el año 2020 el nivel de conexiones SIM alcance el 159 por ciento (73 millones), el nivel de adopción de *smartphones* llegue al 71 por ciento y un porcentaje de cobertura del 85 por ciento (GSMA, 2016).

Según una encuesta realizada por *GSMA Intelligence*<sup>16</sup>, “pese a que alrededor del 97 por ciento de los argentinos vive o trabaja en zonas cubiertas por servicios de 3G o 4G, casi un tercio (27 por ciento) de la población no accede a internet a través de su móvil”. Según el informe, el 72 por ciento de los desconectados argumentan que el motivo principal es por la falta de contenido local relevante, seguido por la falta de alfabetización digital y aptitudes digitales (19 por ciento). Esta evidencia es relevante dado que refuta la hipótesis del “precio del servicio” como barrera más importante para más del 80 por ciento de aquellos que no acceden a Internet. El informe de GSMA destaca como conclusión la importancia de “promover la adaptación de contenidos y servicios tradicionales como la radio o la TV al canal móvil, aprovechando la ubicuidad,

<sup>16</sup> *Country overview: Argentina - Impacto del ecosistema móvil: perspectivas y oportunidades (2016)*

personalización e interactividad, teniendo en cuenta aquello que interesa a los usuarios según su ubicación, costumbres, e idiosincrasia local”<sup>17</sup>.

Entendemos que alcanzar este potencial requiere como condición necesaria una economía estable que promueva un entorno empresarial propicio para la industria móvil y poder ampliar así el ecosistema creando contenido y servicios en nuevas aplicaciones a modo de brindar nuevas oportunidades a las PyMes de inserción a la cadena de valor del ecosistema digital. El informe de GSMA destaca que si bien los rubros de entretenimiento y video tiene una fuerte presencia existen otras áreas en donde la telefonía móvil podría ayudar a generar impacto tales como agricultura, educación y salud.

Finalmente, al igual que se evidencia cierta disparidad regional en los PDA y el uso de productos financiero entre CABA y Buenos Aires, también es evidente en los usos de servicios móviles e internet, encontrando que CABA tiene diferencias de 19 y 10 puntos porcentuales por encima de Buenos Aires respectivamente. La Figura 9 a continuación muestra este punto.

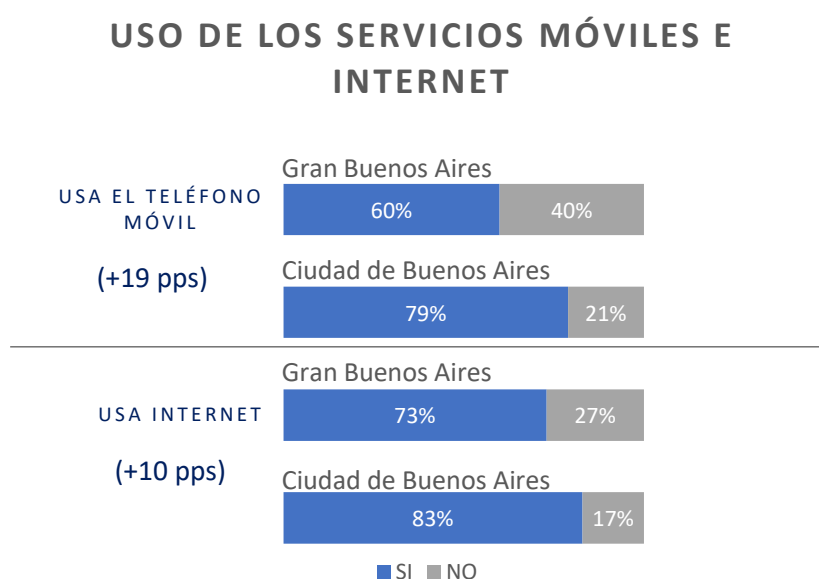


Figura 9

## Modelos tradicionales de *score* versus modelos basados en *Big-Data*

Los modelos tradicionales de *scoring* consideran un conjunto limitado de datos. Por ejemplo, *The Fair Isaac Corporation* (FICO)<sup>18</sup> ha utilizado soluciones algorítmicas

<sup>17</sup> Ídem anterior.

<sup>18</sup> FICO es una compañía americana creada en 1989 focalizada en servicios de desarrollo de modelos de *score* de crédito. El modelo de *score* FICO es utilizado por la mayoría de los bancos y compañías financieras que realizan oferta de crédito. Para el desarrollo de sus modelos se basan en información tres bases de bureau nacionales: *Experian*, *Equifax* y *TransUnion*.



relativamente simples que integran un número limitado de categorías de datos. La Figura 10 de abajo<sup>19</sup> muestra que el *score* se encuentra compuesto por cinco grupos de datos:

- **Historial de pagos (35%):** abarca tanto información positiva como negativa, las positivas contemplan: cancelaciones de líneas de financiamiento en tiempo y forma, consumo de límites dentro de los parámetros establecidos. La información negativa (es decir, que resta puntaje al *score*) puede ser información de atraso en las obligaciones, quiebras, juicios, o asimismo la “falta de comportamiento”.
- **Nivel de Endeudamiento (30%):** esta categoría considera ciertas mediciones específicas tales como: relación deuda sobre límite, el número de cuentas con saldos acreedores, relación cuota ingreso, etc.
- **Historial Crediticio (15%):** esta categoría incluye básicamente el impacto que tiene un historial de crédito a medida que pasa el tiempo, es decir, a medida que el crédito madura, éste puede tener un impacto positivo en el *score* de FICO. Básicamente tenemos dos categorías: antigüedad promedio de las cuentas y la antigüedad de la cuenta máxima.
- **Tipos de productos utilizados (10%):** Incluye: desembolso único, créditos revolventes, hipotecas, etc. Mientras más diversificación haya, mejor será el puntaje de *score*.
- **Consultas de crédito (10%):** considera las consultas realizadas por los clientes al momento de solicitar productos financieros. Si se realiza una gran cantidad de consultas, esto puede afectar negativamente al *score* de FICO. Es interesante destacar que dependiendo del producto que se esté consultando como, por ejemplo, hipotecas, préstamos para automóviles, o préstamos estudiantiles durante un período corto (dos semanas o 45 días) no afectarían de manera negativa al *score* de crédito dado que es más aceptado que en estos casos los consumidores estén buscando mejores precios.

### FACTORES UTILIZADOS POR FAIR ISAAC

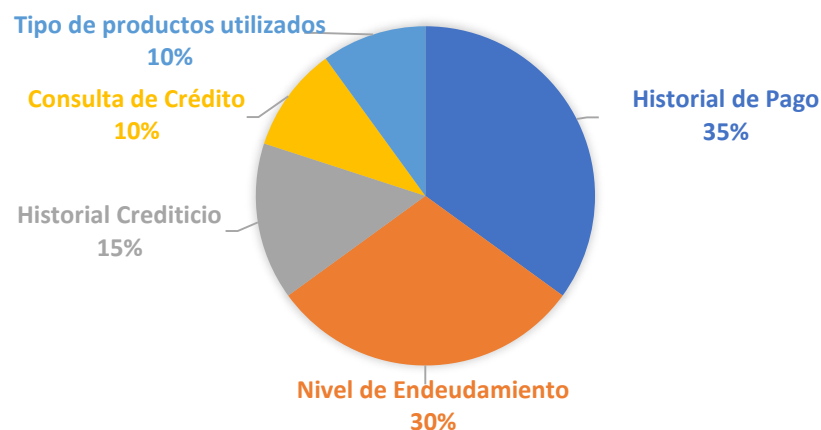


Figura 10

<sup>19</sup> Fuente: <https://www.doughroller.net/credit/a-rare-glimpse-inside-the-fico-credit-score-formula/>

Como consecuencia práctica, las herramientas tradicionales de calificación crediticia pueden, en ocasiones, perpetuar la injusticia al negar a ciertos grupos el acceso favorable al crédito simplemente porque han sido excluidos del mercado crediticio en el pasado.

Las debilidades implícitas que surgen de los modelos tradicionales de *score* han provocado la aparición de nuevas herramientas y por ende nuevos competidores en la industria del crédito. Cabe destacar que los factores tradicionales utilizados por los actuales modelos de *score* de crédito siguen siendo centrales para la toma de decisión automatizada o masiva de crédito, la industria de calificación crediticia se encuentra en un rápido proceso de migración hacia nuevas herramientas alternativas. Incluso las agencias tradicionales de *Bureau* están en proceso de desarrollar herramientas alternativas con datos complementarios a los tradicionales.

Quizás la metodología más comunmente desarrollada y patentada en los últimos años fue la de algoritmos de *Machine Learning*. Es decir, varias empresas desarrollan y patentan modelos o algoritmos de *Machine Learning* para seleccionar y clasificar miles de puntos de datos disponibles a nivel de consumidor. Estas nuevas empresas (*Fintechs*) tratan a sus herramientas de *Machine Learning* como secretos comerciales estrechamente guardados situación la cual no permite obtener una imagen completa o representativa general de la industria. Sin embargo, es posible obtener parte de la información mediante la divulgación de solicitudes de patentes, dichas solicitudes ofrecen información valiosa sobre cómo funcionan las herramientas de *Machine Learning* y los riesgos que éstas pueden presentar.

Dado que el presente trabajo no pretende ser estrictamente técnico a continuación daremos una descripción general de cómo funcionan las técnicas de *Big-Data* para el desarrollo, testeo e implementación de modelos de *scorings*. También se presentarán algunos conceptos básicos y terminologías, luego describiremos en qué difieren las herramientas tradicionales de *credit scoring* con las herramientas de *Machine Learning* aplicadas y, finalmente, trataremos de presentar una descripción lo más detallada posible (paso a paso) de cómo se podría diseñar e implementar una herramienta de *score* basada en *Machine Learning* tomando como modelo de negocio un ejemplo real de una de las *Fintechs* más innovadoras respecto a la materia.

### **El Ecosistema en *Machine Learning* y *Big-Data***

Podemos definir el problema de desarrollo de un modelo de *score* para determinar la probabilidad de *default* de un cliente o consumidor como un problema “no estructurado”. Es decir, no existe una única regla para lograr estimar esta probabilidad. De manera frecuente las compañías de *Bureau* tradicionales utilizan soluciones relativamente simples que abarcan un número limitado de datos tal como el ejemplo presentando anteriormente sobre el *score* de crédito de FICO.

Si bien este tipo de modelo tradicional de *score* de crédito puede parecer simple de aplicar y relativamente fácil de entender para un solicitante de préstamo. Estas

características de simpleza también pueden llevar a decisiones de crédito que pueden ser “poco inclusivas” y que, como hemos mencionado anteriormente, pueden perjudicar a aquel individuo que previamente no ha tenido acceso al sistema crediticio.

En contraposición al modelo tradicional de *score* encontramos nuevos modelos de *score* basados en *Big-Data* y algoritmos de *Machine Learning*. Como ejemplo antagónico al modelo de *score* de FICO podemos mencionar a la *Fintech* ZestFinance<sup>20</sup>. Dicha compañía fue fundada en 2009 por (entre otros) el ex Director de Información de Google (CIO), Douglas Merrill. Esta *Fintech*, es capaz de analizar una gran cantidad de datos y logra calificar a más del doble de la población estadounidense que los modelos tradicionales y, además, logrando reducir la tasa de mora en aproximadamente un 60 por ciento en promedio con la industria.

El mismo Merrill señala que “... El crédito actual se basa en el comportamiento pasado: por esto, el buen crédito genera más crédito bueno y el mal crédito genera más crédito malo. Creemos que hemos encontrado una mejor manera de obtener una imagen completa de los solicitantes de crédito. En 10 a 15 años, esperamos que esta sea la única forma en que se emita el crédito”<sup>21</sup>.

Si observamos los departamentos de ZestFinance podemos ver que, en esencia, es una compañía de “matemática”. Es decir, se encuentra conformado por profesionales de *data scientists*. Por otra parte, a diferencia de las compañías tradicionales que estiman un único modelo de *score*, el modelo de negocio de ZestFinance, es decir el cálculo de *score*, se para sobre el desarrollo de 10 modelos analíticos basados en *Machine Learning* que se ejecutan en paralelo para lograr interpretar más de 10.000 puntos de datos por solicitante y, así, poder llegar a conseguir más de 70.000 señales.

## Terminologías y Conceptos

Para dar comienzo a este apartado primero definiremos el concepto de algoritmo. Podemos describir un algoritmo como todo proceso o función de transformación de valor computacional definido que requiere de un conjunto de datos de entrada (*inputs*) y produce un conjunto de valores de salida (*outputs*). Es decir, podemos definir el término algoritmo como una secuencia de pasos computacionales que transforman una entrada de datos en una salida. Obviamente, los algoritmos pueden variar en su complejidad, una forma de distinguir estos es mediante la clasificación del problema a resolver. Por ejemplo: a) si el problema es simple o se encuentra bien definido o, b) si el problema es complejo o no se encuentra bien definido, es decir, no tiene una clara descripción o límite. Los problemas “bien definidos” también pueden clasificarse como problemas “estructurados” mientras que los problemas “no definidos” como “problemas no estructurados”. Esta clasificación es importante porque los problemas estructurados tienen generalmente una única respuesta para un conjunto de datos de entrada (por ejemplo, la media geométrica), mientras que los problemas no estructurados pueden tener múltiples respuestas “correctas”, y estas múltiples

---

<sup>20</sup> <https://www.zestfinance.com>

<sup>21</sup> <https://pando.com/2013/07/31/flush-with-20m-from-peter-thiel-zestfinance-is-measuring-credit-risk-through-non-traditional-big-data/>

respuestas correctas, a su vez, pueden ser unas mejores que otras. A modo de dar un ejemplo no estructurado podemos utilizar la analogía de un cocinero y una receta particular. Cada vez que éste haga el mismo plato, no tendremos un resultado único y correcto, dado que dependerá de los ingredientes (o datos) y tampoco existe un único tiempo de cocción, proporciones, condimentos, etc.

El término *Machine Learning* (que se encuentra relacionado, pero difiere del término *Big-Data*) determina el conjunto de métodos que tienen como objetivo detectar algún patrón en los datos de manera automática para luego predecir comportamientos o patrones futuros o predecir otros comportamientos bajo incertidumbre.

La metodología *Machine Learning* puede tener dos tipos de aprendizaje, el que se denomina “aprendizaje supervisado” y el que se denomina “aprendizaje no supervisado”. En el caso del aprendizaje supervisado, se tiene una salida o variable objetivo (denominada comúnmente variable dependiente) conocida o deseada, y lo que se desea conocer es la relación entre esta variable deseada y otras variables que se denominan comúnmente explicativas. En el caso de aprendizaje no supervisado, no se tiene algo específico que se desee predecir o determinar, lo que significa que el proceso no se centra en comprender una variable objetivo conocida. Sin embargo, el aprendizaje no supervisado puede encontrar relaciones entre los datos de que pueden ser de utilidad para el futuro.

### **Antecedentes en *Big-Data***

El ya conocido crecimiento exponencial que la rápida evolución de la tecnología ha dado paso al surgimiento de lo que algunos analistas de la industria denominan "la década del *Big-Data*". El término *Big-Data* puede tener varias definiciones en mayor o menor medida técnicas. No obstante, en el uso común de la palabra (y para delinear el presente apartado) podemos definir el término *Big-Data* como aquella gran cantidad de datos que los consumidores generan en su vida cotidiana a través de transacciones comerciales, mensajes de correo electrónico, fotos, videos, tráfico web, registros de actividades almacenados en grandes bases de datos estructuradas, o texto no estructurado publicado en la web como por ejemplo blogs y redes sociales.

En la última década, la cantidad de datos generados ha crecido exponencialmente, en parte debido al auge de las técnicas de rastreo de *web* (*web tracking*) y al mayor uso de dispositivos móviles con acceso a Internet. Asimismo, las innovaciones en la capacidad informática, el menor costo de almacenamiento de datos y los avances en el análisis estadístico hacen, en su conjunto, que sea más fácil interpretar y monetizar los datos. La Figura 11 a continuación muestra dicha evolución (2018\* estimado)<sup>22</sup>.

---

<sup>22</sup> NOTA: 1 Zettabyte (ZB) es equivalente a 1.000 millones de terabytes (TB). Y donde 1 Terabyte (TB) es equivalente 1.000 Gigabytes (GB), es decir un ZB es igual a 1.000.000.000.000 (un trillón) de GB.

## VOLUMEN DE DATOS/INFORMACIÓN MUNDIAL CREADA (en zetabytes)

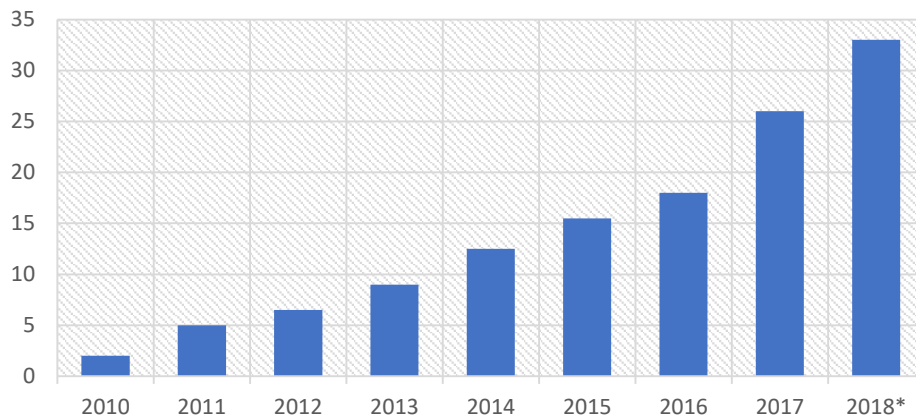


Figura 11

### ***Machine Learning* para el desarrollo de un Modelo Score basado en *Big-Data***

#### ***El proceso de Analytics en Big-Data***

Es probable que nos sorprendamos si analizamos los tipos de datos que se utilizan para predecir la probabilidad de *default* de un cliente. Si bien los datos tradicionales de comportamiento crediticio siguen siendo útiles (siempre y cuando el cliente tenga un historial crediticio) otros puntos de datos no convencionales pueden, a priori, parecer tener poca relación con la solvencia del cliente o la probabilidad de *default*. Por ejemplo, dentro de los datos no estructurados más conocido encontramos que la rapidez con la que un cliente (solicitante) se desplaza a través del apartado de “divulgación de términos y condiciones” que se encuentra en la página que ofrece el producto, o la hora en que se conecta a internet (diurno o nocturno), o la evidencia de abandonar las redes sociales, el hábito de gasto del individuo dado un contexto de ubicación geográfica, es decir, por ejemplo: no es lo mismo que una persona destine la mitad de su salario para alquilar de vivienda en una zona cara que en una zona mucho más barata.

Muchas empresas se dedican y/o se especializan en el tratamiento de datos (empresas relacionadas con el concepto de *Analytics*). En este ámbito podemos destacar tres funciones principales diferentes:

- Recolección de Datos,
- Agregación de Datos y
- Análisis de Datos.

## Recolección de Datos

Antes de sumergirnos en la comprensión de la recolección de datos es importante comprender cómo se crean los datos. Como ya mencionamos anteriormente, con la introducción de más dispositivos habilitados para acceder a Internet (ya sean computadoras en todas sus variantes, teléfonos móviles, relojes inteligentes, consolas de video juegos, *smart tvs* y *tablets*), la cantidad de datos que genera un consumidor diariamente es enorme. A modo de entender que significa “enorme” cabe señalar que entre 2010 y 2016 por ejemplo, la cantidad de datos generados aumentó en un factor de nueve a 18 zettabytes (18 billones de gigabytes)<sup>23</sup>. Cada vez que un consumidor visita un sitio web, realiza una compra o indica una preferencia en *Facebook* u otros sitios de redes sociales, se crean datos.

A modo de describir como es el proceso de creación de datos supongamos que, por ejemplo, una mujer interesada en comprar un libro (novela de misterio) se sentara en su computadora y abrirá un navegador *web*. A continuación, escribe "*Amazon.com*" en la línea URL. Al escribir la URL, su computadora solicita la página del servidor de Amazon. La computadora transmite su dirección de Protocolo de Internet (IP) a la página web. Una dirección IP es similar a una dirección física, ya que cada dirección es única. Según la dirección IP de la mujer, el servidor del sitio *web* puede predecir su código postal (con distintos grados de precisión). El servidor de Amazon envía la página *web* y descarga una "*cookie*" (línea de texto) en el disco duro de la mujer. Es importante destacar que, además, varias otras empresas de *marketing* de terceros que tienen contrato con Amazon también pueden descargar sus *cookies*. Una *cookie* puede contener varios tipos de información, incluidos (entre otros) el momento de su visita, las subpáginas que visitó y los artículos que compró. Las *cookies* también suelen designar una identificación única para la computadora que uno utiliza. Al asignar una identificación única, las compañías que se dedican al seguimiento de terceros o consumidores pueden ver otras páginas que la persona ha visitado, intuyendo así sus preferencias. Si bien las *cookies* son conocidas, las compañías de seguimiento también pueden incorporar un software llamado "*web beacon*" que no solo puede rastrear qué páginas web visita una persona, sino que también registra el texto ingresado. Por ejemplo, si una página web tiene una "*web beacon*", entonces cuando la persona usa la función de "búsqueda" en una página web, como la de Amazon, esa información se transmite a un vendedor externo. Las páginas subsiguientes que la persona visita también se rastrean. Si nuestra mujer del ejemplo anterior compra algunas novelas de misterio en Amazon, reserva vuelos y hotel para vacacionar, navega por la web para obtener noticias y "busca en Google" las mejores tarifas para seguro de automóvil, las compañías de seguimiento pueden capturar cada movimiento que hace.

Finalmente existe una última técnica de rastreo denominada "*web crawling*" que las empresas pueden usar sin desarrollar una relación con una página de host. Básicamente esta técnica involucra un proceso de duplicación y categorización de información de sitios web, generalmente por medios automatizados. Es decir, los programadores

---

<sup>23</sup> Fuente: <https://www.statista.com/statistics/871513/worldwide-data-created/>

pueden escribir *software* que escanea sitios web y clasifica publicaciones, no solo se pueden rastrear datos sino también para analizar perfiles e imágenes.

### Agregación de Datos

De manera abreviada podemos decir que la agregación de datos es el proceso de combinar una matriz de datos o fuentes de datos, trabajarlos para que sean uniformes y finalmente compilar un retrato completo de una persona, su comportamiento o características. Por ejemplo, se pueden combinar datos de agencias de crédito alternativas con datos obtenidos del rastreo *web* para tomar una decisión sobre si prestar dinero a las personas.

También existe otro proceso interno en este punto que se denomina habitualmente creación de “Meta-Variables”. Este tipo de variable surge de submodelos que tienen como objetivo crear otras variables que expliquen el comportamiento de la persona. Por ejemplo, si un solicitante de crédito carga información sobre gastos de servicios básicos como alquiler, este puede ser comparado con el precio de alquiler promedio de la zona y crear así una variable que marque el grado de veracidad de los gastos.

### Análisis de Datos

Los datos mencionados anteriormente, procesados o agregados, se ejecutan a través de una serie de modelos (generalmente llamados algoritmos) para revelar patrones o probar hipótesis.

Cabe mencionar que, si bien nosotros separamos los “procesos” de recopilación, agregación y el análisis de datos en el uso de *Big-Data*, esto no necesariamente se realiza por separado. Existen compañías que compran datos pero que también recopilan datos a través del rastreo web.

Finalmente podemos mencionar que las técnicas de algoritmos utilizadas frecuentemente se denominan técnicas de Inteligencia Artificial (IA) o técnicas de *Machine Learning* (ML). Estas técnicas básicamente lo que hacen es ejecutar una gran cantidad de modelos simultáneamente y los hacen competir de manera “democrática” para tomar una decisión.

### Diseño Funcional: Caso práctico

A modo de ejemplo trataremos dos modelos presentados. En primer lugar, tenemos a la bien conocida empresa de redes sociales *Facebook* que en el año 2012 presentó recientemente una solicitud de patente relativa a un método para “Autorización y autenticación basada en la red social de una persona”. La patente indica que una de sus aplicaciones podría ser para usarse como *score* de crédito<sup>24</sup>. El texto dice: “...Cuando una persona solicita un crédito, el prestamista examina los *credit ratings* de las personas que conforman la red social del solicitante, es decir, aquellas personas que están

---

<sup>24</sup> Fuente: Columna 2 párrafo número 10: <https://patents.google.com/patent/US9100400B2/en>

conectadas al solicitante a través de nodos autorizados de la red social. Si el *credit scoring* promedio de estos miembros alcanza un umbral mínimo, el prestamista continúa procesando la solicitud de préstamo. De lo contrario, la solicitud es rechazada”. En este apartado trataremos de poner en contexto el diseño estructural general de un modelo de *Machine Learning* “no estructurado”. Si bien no existe una única metodología para el desarrollo de una solución con *Big-Data* creemos es útil tratar de delinear un sistema que permita sentar las bases para el desarrollo del mismo.

A modo de presentar una metodología generalizada trabajaremos en un modelo de *score* basado en tres etapas:

- 1) **Identificar la variable objetivo:** es decir definir el problema a resolver.
- 2) **Establecer el repositorio de Datos:** incluyendo las transformaciones necesarias.
- 3) **Implementación y calibración del modelo:** proceso de implementación y ajuste del modelo a través de la utilización de datos de “entrenamiento” y selección de características.

Estos tres pasos son genéricos a cualquier proceso de desarrollo de modelos y refleja, además, el proceso que ZestFinance describe en su solicitud de patente para su herramienta “alternativa” de *score* crediticio<sup>25</sup>. En la Figura 12, a continuación, presentamos un esquema de dicho modelo:

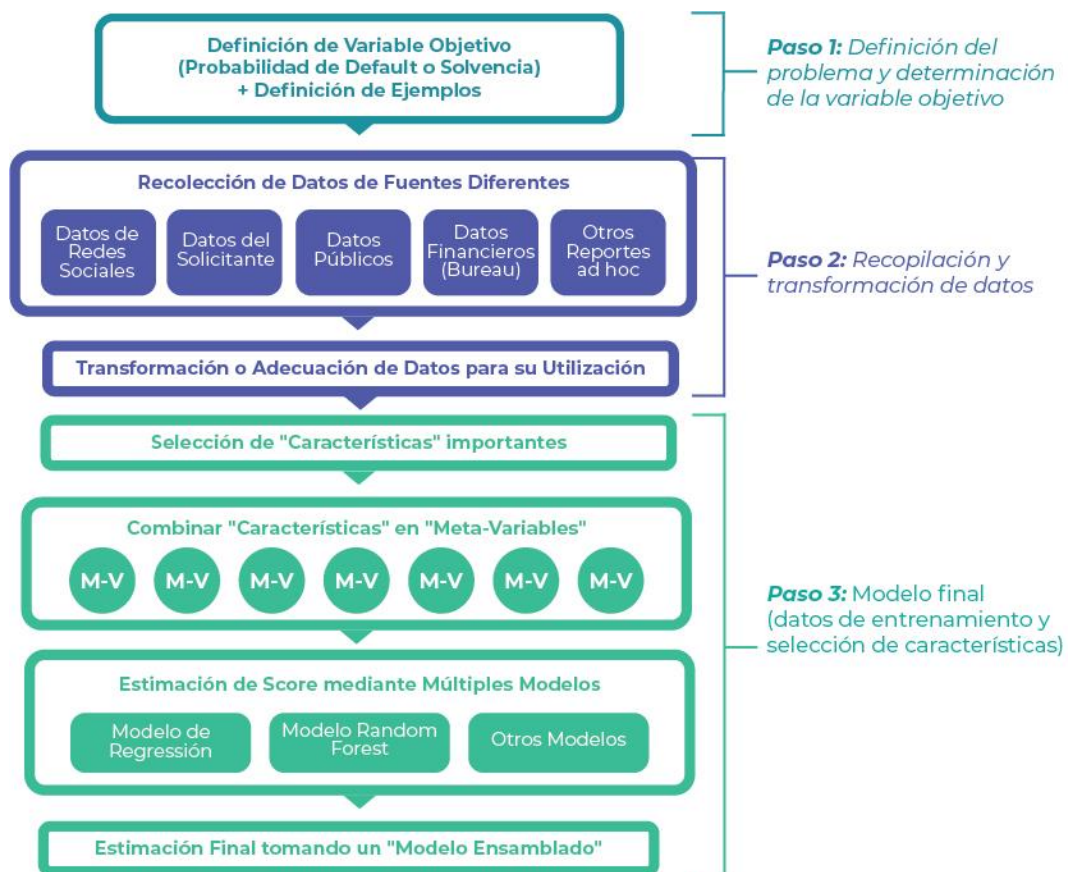


Figura 12

<sup>25</sup> Fuente: <https://patents.google.com/patent/US20150019405A1/en?q=14%2f276%2c632>



## Definición del problema y determinación de la variable objetivo

Antes de desarrollar el modelo “supervisado” debemos definir el problema a resolver, es decir, la variable objetivo. En nuestro caso queremos estimar ya sea la probabilidad de *default* o el nivel de solvencia del cliente. Si bien esto parece ser obvio, al momento de trabajar en un problema con datos “no estructurados” podemos encontrarnos con que no tenemos una única respuesta. A modo de dar un ejemplo, si queremos desarrollar un modelo que segmente (es decir que realice clusters) clientes para poder determinar un subconjunto de éstos que sean más propensos a responder y aceptar una promoción (publicidad dirigida), nos encontramos con que no existe una fórmula o regla predefinida que nos diga por qué ciertos clientes responden a los anuncios dirigidos, mientras que otros no.

Entrando un poco más en detalle sobre lo que es la “variable objetivo” podemos decir que ésta representa un estado o resultado de interés. Comunmente la variable objetivo se define en base resultados pasados o características de variables. Estos resultados o características también se describen como “atributos de clase”. Es decir, supongamos que una cadena de *retail* en algún momento envió una promoción de descuento por mail. El email contenía un código de descuento para utilizar en la página *web* (compra en línea). Al momento de finalizar la promoción la cadena de retail puede confeccionar una lista de aquellos clientes que si respondieron a la oferta promocional y otra lista con aquellos clientes que no respondieron. Ahora bien, cada una de estas listas se corresponderá a una de las dos clases que representan la capacidad de respuesta ante la promoción, que es nuestra variable de interés. Nuestro “atributo de clase” para el primer grupo sería “permeable a la campaña dirigida”, mientras que el segundo grupo sería “sin respuesta a la campaña dirigida”.

## Recopilación y transformación de datos

Una vez definida la variable objetivo y las clases establecidas, se procede a recolectar la información necesaria. Esta información tiene un carácter de “entrenamiento” (training data set) que será utilizada durante el proceso de aprendizaje del modelo. Como ya hemos mencionado anteriormente, en la era del *Big-Data* siempre se trata de aprovechar todo tipo de información, desde el historial de compras en la web, antecedentes crediticios, historial de navegación en Internet (*cookies*), y hasta los grupos de “amigos” de alguna plataforma de redes sociales.

Siempre es conveniente clasificar los datos según sus características. Por ejemplo, podríamos establecer tres grupos:

- Datos del solicitante: esta categoría incluye los datos proporcionados por el solicitante como los datos de la actividad de su navegador *web*.
- Datos Privados: datos gubernamentales o de bases administradas por organismos (de identificación y bienes por ejemplo).
- Datos Públicos: puede abarcar desde datos de compras *online*, datos fuera de línea y hasta historial médico o de seguros.
- Datos de Redes Sociales: consiste básicamente en actividades en redes sociales y se incluye información agregada de las publicaciones y su red social.

El último paso de este módulo es la transformación de datos. Es decir, se trata de construir unidades de medida que sean utilizables para el procesamiento del modelo. Por ejemplo, en lugar de tomar el salario bruto de una persona de manera nominal, se trata de traducir éste en un puntaje *score* (percentil) contra salarios brutos de otras personas que trabajan en la misma industria y región geográfica. Otro ejemplo es el tipo de dato que mencionamos anteriormente que hacía referencia a la cantidad de tiempo que la persona dedica a leer los “terminos y condiciones” del producto que se puede ver como un indicador de nivel de responsabilidad del solicitante. Este dato no ingresa en formato de segundos, minutos, etc. Sino más bien se trata de crear una variable del tipo categórica u ordinal en una escala, por ejemplo, de 0 a 2 donde 0 indica poco o ningún “cuidado” durante el proceso de solicitud y 2 indica “atención meticulosa” a los detalles durante el proceso de solicitud.

Sin embargo, como hemos presentado en la imagen de patente de ZestFinance, el proceso de transformación de datos puede no necesariamente terminar aquí. Después de transformar las variables, ZestFinance procesa las variables resultantes “utilizando uno o más algoritmos” (que pueden ser econométricos, financieros o *Machine Learning*) para generar una pluralidad de conjuntos de decisiones independientes que describen aspectos específicos de un solicitante, que la patente se refiere como “meta-variables”<sup>26</sup>. Las meta-variables de ZestFinance parecen “ubicar” a los solicitantes en categorías o clases mediante inferencia a partir de uno o más datos previamente transformados. Por ejemplo, una meta-variable podría comparar el ingreso informado por un solicitante con el ingreso promedio de aquellos individuos con profesiones similares y que viven en la ciudad del solicitante, y luego generar un “control de veracidad” que representa la probabilidad de que el solicitante esté adulterando su salario. Como otro ejemplo, ZestFinance podría calificar al solicitante en un grado de escala “estabilidad personal”, en función de la cantidad de tiempo que “ha sido accesible de manera constante en un pequeño número de direcciones o números de teléfono”. La patente además explica que las meta-variables son “muy útiles en la etapa intermedia de construcción de una función de *score* crediticia”, y pueden usarse para determinar “...qué señales se medirán en el proceso de calificación final, y qué peso es para ser asignado a cada señal”.

#### Modelo final

Una vez finalizado el proceso de recopilación y transformación de datos se puede dar comienzo al proceso de *Machine Learning*. Es importante destacar que es factible que a medida que el algoritmo “aprenda” comience a descartar datos de entrada por no ser relevantes en función de la variable objetivo. Habitualmente los procesos de *Machine Learning* conllevan una rutina de optimización que intenta identificar aquellas variables de entrada que son más significativas y qué peso asignarle a cada una. Este proceso iterativo de identificar aquellas variables de entrada relevantes y descartar las irrelevantes se lo denota como “selección de características”.

---

<sup>26</sup> Ver punto [0043] United States Patent Application Publicatio Pub.No.: US 2015/0019405 A1

Finalmente, el último paso es obtener un puntaje final el cual se obtiene de un proceso “democrático” entre varios modelos. Es decir, cada modelo “vota” a modo de lograr obtener un *score* final. Existen varias formas de combinar múltiples modelos en una sola decisión agregada utilizando *Machine Learning*. A modo de ejemplo podemos suponer que cada modelo de *score* arroja un puntaje en escala de percentil y finalmente se realiza un promedio de todos los puntajes. Supongamos que tenemos 3 modelos. El “Modelo I” basado en la técnica de clasificación de árbol “Random Forest”, el “Modelo II” basado en Regresión Logística y el “Modelo III” basado en Redes Neuronales. Asumamos que el “Modelo I” arroja un valor de 0,78, el “Modelo II” arroja un valor de 0,20 y el “Modelo III” arroja un valor de 0,95, simplemente se realiza un promedio de estos valores.

## **Oportunidades para la Inclusión Financiera a partir de la Tecnología Digital, técnicas de *Machine Learning* y *Big-Data***

Como hemos mencionado a lo largo del trabajo, existe una gran proporción de la población de clase media con ingresos estables que si bien puede tener caja de ahorro (el 80 por ciento de la población posee una caja de ahorro) no se encuentra bancarizada lo cual no le permite tener acceso a productos de crédito como tarjetas, descubiertos, préstamos o chequeras. Asimismo, al no tener un historial formal de comportamiento de crédito la banca tradicional, que utiliza modelos de *scoring* basados en el comportamiento crediticio o financiero<sup>27</sup>, no los considera aptos de ser asistidos crediticiamente. Por lo tanto, este gran segmento de clientes no bancarizados se encuentra “atrapados” en un bucle continuo de exclusión financiera.

Dadas las tendencias tecnológicas mundiales que hemos presentado, y sobre las cuales la Argentina no es la excepción, una forma viable de romper el paradigma de la exclusión financiera es crear una solución que haga más robusto al ecosistema financiero actual permitiendo, entre otras cosas:

- Potenciar o ampliar los PDA a nivel nacional a modo de satisfacer la demanda y facilitar las transacciones electrónicas y medios de pago.
- Aumente la utilización de productos financieros a nivel nacional sobre aquellas personas que ya poseen una cuenta.
- Incentivar el uso de medio electrónico de pagos y, por ende, alentar la disminución del uso de efectivo,
- Permitir o facilitar el acceso al crédito a aquellas personas que se encuentran excluidas.
- Acompañar las nuevas necesidades de los clientes en lo que respecta a servicios financieros.

Antes de comenzar a detallar cada una de las iniciativas. Entendemos que dicho ecosistema sería más propicio bajo un modelo o iniciativa de carácter “Público-Privado”. Para esto se propone trabajar con algún modelo de Microfinanzas ya existente. En

---

<sup>27</sup> Los Bureaus tradicionales del sistema financiero argentino son Veraz y Nosis.

América Latina los elementos más conocidos para diferenciar los modelos de Microfinanzas se dan por medio de los conceptos de “*upgrading*” y “*downscaling*”.

El “*upgrading*” se refiere a la transformación de organizaciones de Microfinanzas no gubernamentales (es decir ONG) en entidades formales supervisadas por la autoridad bancaria pertinente. A su vez, el “*downscaling*” es el proceso a través del cual las instituciones financieras formales, que generalmente se encuentran fuera del ámbito de las Microfinanzas, se involucran en este sector.

Si bien en Latinoamérica el proceso de microfinanzas de “*upgrading*” no es muy conocido, ha tenido un factor relevante al momento de impulsar el desarrollo. Quizás una de motivaciones principales de éste ha sido lograr ampliar las fuentes de financiamiento para lograr conseguir un mayor volumen de préstamos. A modo de alistar algunos casos de éxito encontramos: Banco ProCredit Los Andes (Bolivia), Banco Sol (Bolivia), Finamérica (Colombia), MiBanco (Perú) y Compartamos (México).

Por otra parte, el “*downgrading*” es un proceso un poco más conocido y se ha sabido destacar por la rentabilidad y la diversificación de productos. Podemos decir que en este modelo lo que se desarrolla es el ingreso de entidades dedicadas al crédito de consumo sobre aquellos segmentos de bajos ingresos. Quizás el mayor riesgo de este modelo de microfinanzas es la diferencia de modelo de negocio entre los segmentos de personas y PyMes (o Microempresas). Entre los casos destacados encontramos al: Banco Santander/Banefe (Chile), Banco de Crédito (Perú), Banco Caja Social (Colombia) y también algunas entidades públicas que entraron al sector como: Banco Estado (Chile) y Banco do Nordeste (Brasil)

### **El panorama de los pagos digitales.**

Cambiar los pagos de efectivo a medios digitales puede tener beneficios más allá de expandir la apertura de cuentas y aumentar su uso. Creemos que una buena estrategia de digitalización de pagos puede además aumentar la velocidad de los pagos y reducir los costos sobre los créditos y débitos. Otra externalidad positiva es que éste proceso también puede mejorar la seguridad de los pagos y, por lo tanto, lograr disminuir la incidencia de delitos asociados a los mismos. Asimismo, también se ha demostrado que reemplazar pagos en efectivo por canales digitales aumenta la transparencia y reducen la corrupción. Finalmente, también hay que considerar que, al proporcionar un primer punto de entrada importante en el sistema financiero formal, la movilidad hacia los pagos digitales puede conducir a aumentos en el ahorro y disminuir así, el ahorro informal.

Los teléfonos móviles e Internet han dado lugar a una nueva generación de servicios financieros. El uso de estos servicios no requiere necesariamente dispositivos sofisticados. En África, los teléfonos móviles basados en texto relativamente simples han impulsado la propagación de las cuentas de dinero móvil. Servicios similares están disponibles en otras partes del mundo en desarrollo. Y la tecnología de teléfonos inteligentes se utiliza cada vez más para realizar transacciones a través de cuentas de instituciones financieras en algunas economías en desarrollo.

Los teléfonos móviles e Internet pueden impulsar la inclusión financiera solo si están respaldados por la infraestructura necesaria. La infraestructura física y las redes móviles, son clave. También se necesita infraestructura financiera. Esto incluye tanto un sistema de pagos adecuado como una red física para entregar pagos a todos los rincones de una economía, tanto urbana como rural.

Si bien es posible que las instituciones financieras no encuentren rentable abrir una sucursal física en cada lugar que tenga una gran población no bancarizada, pueden utilizar la banca de agentes, formando asociaciones con oficinas de correos o tiendas minoristas para ofrecer servicios financieros básicos a los consumidores. Las personas que utilizan pagos digitales deben poder depositar y retirar efectivo de manera segura, confiable y conveniente en los puntos de entrada y salida de efectivo, ya sea que se trate de un agente bancario, un agente de dinero móvil o un cajero automático (ATM). Idealmente, las personas que reciben pagos digitales deberían mantener sus fondos en forma digital y realizar compras y pagar facturas electrónicamente. Pero, como hemos visto, en muchos lugares los pagos digitales aún no son ampliamente aceptados para compras diarias en tiendas minoristas y mercados locales, especialmente en economías en desarrollo. Por lo tanto, la mayoría de las personas deben poder retirar al menos parte del dinero que reciben a través de pagos digitales. De hecho, una experiencia de retiro confiable es clave para el éxito de los pagos digitales.

Si bien la tecnología y la infraestructura son una parte importante del desarrollo de los pagos digitales, para garantizar que las personas se beneficien de estos servicios financieros, los Gobiernos deben tener en cuenta las normas establecidas y proteger al consumidor. Entendemos de la tecnología utilizada, los servicios financieros deben tener las necesidades de los grupos desfavorecidos que pueden tener bajos conocimientos de alfabetización y aritmética. También es importante analizar quién tiene acceso a la tecnología digital necesaria para usar los servicios y quién no.

### **Aprovechando la tecnología digital entre los no bancarizados**

En muchas economías de altos ingresos, las tarjetas de débito y crédito utilizadas en los terminales de punto de venta (POS) dominan el panorama de los pagos digitales. En la mayoría de las economías en desarrollo, por el contrario, pocas personas tienen esas tarjetas. Pero muchos tienen un teléfono móvil, lo que podría permitir a estas economías saltar directamente a los pagos móviles.

Simplemente tener un teléfono móvil puede permitir el acceso a cuentas de dinero móvil y otras cuentas financieras basadas en texto o aplicaciones. Tener acceso a Internet también amplía las posibilidades. De hecho, los datos de *Global Findex* sugieren que los teléfonos móviles e Internet podrían ayudar en gran medida a superar algunas de las barreras que los adultos no bancarizados dicen que les impide acceder a los servicios financieros. En el año 2017 Argentina tenía aproximadamente 13,1 millones de argentinos sin cuenta bancaria, pero con teléfono móvil.

Los servicios financieros digitales pueden reducir la distancia entre las instituciones financieras y sus clientes. Y al reducir el costo de proporcionar servicios financieros, la

tecnología digital podría ser útil para aquellos que citan los altos costos como una razón para no tener una cuenta en una institución financiera.

Las tarjetas de identificación biométricas basadas en tecnología digital proporcionan otra forma de reducir las barreras para la titularidad de la cuenta. En India, la estrategia base de inclusión financiera se ha basado en tres ejes principales: 1) un programa de inclusión financiera del Gobierno (*Jan-Dhan*), 2) la creación de un número de identidad único de 12 dígitos (denominado *Aadhaar*) que los residentes de la India pueden obtener de forma voluntaria, este número se genera sobre sus datos socio demográficos o personales como: nombre, teléfono, dirección, etc. y biométricos tales como: huella dactilar y el iris, 3) Telefonía móvil.

Esta iniciativa tuvo inicio en enero 2009 por el Gobierno de la India y a la fecha *Aadhaar* es el sistema de identificación biométrica más grande del mundo. Paul Romer, Economista Jefe del Banco Mundial describió a *Aadhaar* como "...el programa de identificación más sofisticado del mundo".

Esta trinidad (comúnmente llamada "*JAM*" por el acrónimo de "*Jan-Dhan*", "*Aadhaar*" y "*Mobile*") tuvo como objetivo principal integrar el sistema de identificación y las aplicaciones de pago con el fin de lograr no solo ser más rentable, sino también mas eficiente en el proceso de apertura y acceso a cuentas bancarias.

El sistema de identificación de 12 dígitos incluye un servicio electrónico *KYC* (*e-KYC*<sup>28</sup>) para acelerar la verificación de la identidad del cliente. El *e-KYC* otorga la posibilidad a una persona que posea un número de *Aadhaar* permita a la Autoridad de Identificación Única de la India (UIDAI) revelar su información personal a cualquier proveedor de servicios que éste individuo desee activar de manera instantánea, ya sean servicios de conexiones móviles como cuentas bancarias.

Lo interesante es que el *e-KYC* es un proceso "sin papel", basado puramente en el consentimiento, de carácter privado e instantáneo. El *e-KYC* comparte 62 datos de Diligencia Debida del Cliente (DDC) confiables con la entidad informante en tiempo real. Además, dado que los datos de *KYC* se divulgan directamente a los proveedores de servicios si y sólo si con el consentimiento del cliente, su privacidad permanece protegida siempre que se apliquen medidas sólidas de protección de datos. Al 2018 aproximadamente se han realizado un total de 4.900 millones de transacciones de *e-KYC* a través de *Aadhaar*.

A raíz de esta iniciativa los bancos y los operadores de redes de pago han incorporado la autenticación de *Aadhaar* en micro-ATM para proporcionar servicios bancarios sin sucursales en cualquier parte del país de manera interoperable, escalable y en tiempo real. Desde el punto de vista del FSP, ofrece enormes beneficios en términos de la casi eliminación del papel y la consiguiente carga de mantener registros y facilitar la auditoría y el análisis forense a través del almacenamiento electrónico de información<sup>29</sup>.

---

<sup>28</sup> Por las siglas de "*Know Your Customer*" en inglés.

<sup>29</sup> G20 *Digital Identity Onboarding* – Argentina 2018

## Oportunidades para ampliar la apertura de cuentas entre los no bancarizados

Un fenómeno común a lo largo del planeta, y que Argentina no es la excepción, es que millones de adultos no bancarizados todavía reciben pagos regulares en efectivo, es decir, salarios. Estos pagos de efectivos pueden provenir del Gobierno, instituciones privadas, etc. Digitalizar dichos pagos es una forma comprobada de aumentar la titularidad de cuentas de ahorro. Según el informe del *Global Findex database* del año 2017 a nivel mundial, el 9 por ciento de los adultos, o el 13 por ciento de los titulares de cuentas de ahorro, abrieron su primera cuenta específicamente para recibir salarios del sector privado o del Gobierno. La proporción es mayor en muchas economías como Argentina donde el valor alcanza aproximadamente el 25 por ciento.

Cabe mencionar que los Gobiernos hacen varios tipos de pagos a las personas, es decir, no solo pagan salarios a los empleados públicos, sino también además distribuyen pensiones y beneficios sociales. En Argentina los pagos digitales que hacen a las transferencias gubernamentales han tenido un impacto significativo dado que para el 2017 alrededor del 11 por ciento de los beneficiarios abrieron su primera cuenta para recibir la transferencia del Gobierno. Según el *Global Findex database* del año 2017, más importante aún es el impacto entre mujeres y adultos más pobres dado que estos pueden beneficiarse más que proporcionalmente cuando el Gobierno digitaliza los pagos de transferencia. En Argentina particularmente, casi una cuarta parte de los titulares de cuentas bancarias en el 40 por ciento más pobre de los hogares abrieron su primera cuenta por esa razón.

Entrando en el sector privado la digitalización de los pagos salariales en Argentina también podría reducir el número de adultos no bancarizados hasta en un quinto. Esto se ve potenciado con el hecho de que la mayor parte de estos asalariados ya tienen un teléfono móvil<sup>30</sup>. En Argentina hemos visto que aproximadamente el 50 por ciento de la población percibe parcial o totalmente su salario en efectivo, la propiedad de teléfonos móviles de esta población alcanza el 90 por ciento.

## Aumento de utilización de productos financieros y disminución de efectivo

Es importante destacar que la entidad debería ofrecer no solo asistencia crediticia, sino que además debería tener la funcionalidad de “billetera virtual”, es decir, de ser un medio digital de pago de servicios. Cuando hacemos referencia a una “billetera virtual” nos referimos a un servicio de compra simplificado (es decir con un clic) apalancado en la tecnología *Near Field Communication* (NFC) en conjunto con un sistema de encriptado.

Una de las billeteras virtuales mundialmente conocidas es la “*Apple Pay*”. Este modelo innovador (lanzado en 2014) se fundamenta principalmente en dos principios: que sea confiable y sencillo. El modelo de “*Apple Pay*” se monta sobre chips de tecnología NFC que interactúan con el lector de tarjetas de crédito y, más recientemente, han incorporado un chip separado llamado “elemento seguro” (SE, por sus siglas en inglés).

---

<sup>30</sup> *Global Findex database* del año 2017.

Este nuevo chip genera por cada transacción un código único al azar encriptado en lugar de transmitir el número de débito o tarjeta de crédito del usuario (parecido a lo que se conoce como *Blockchain*). De esta manera la “billetera virtual” oculta los detalles bancarios tanto al comerciante como a Apple, minimizando el riesgo de robo de datos de tarjeta de crédito.

Otra manera de incentivar el uso de la cuenta bancaria sería digitalizar más completamente los pagos de agua, electricidad, gas y otras facturas de servicios públicos. Para el año 2017 en Argentina aproximadamente 17,1 millones de argentinos han abonado servicios en efectivo durante el último año<sup>31</sup>. Es importante destacar los motivos por los cuales prefieren el pago en efectivo en lugar de los canales digitales. En ocasiones las personas eligen no optar por dicho canal debido a costos, falta de comprobante de pagos, seguridad u otras preocupaciones. Es por eso que adoptar un enfoque de incentivos para pagos digitales y mejorar la eficiencia puede ocasionar externalidades positivas en la utilización de dichos medios.

### **Potenciar los PDA a nivel nacional**

Hemos visto en los apartados anteriores existe un problema de acceso a los PDA a nivel nacional. Salvo dos jurisdicciones: Provincia de Buenos Aires y CABA, que poseen un 30 y 19 por ciento de área cubierta por cada 10.000 habitantes respectivamente. El resto de las provincias poseen valores que promedian el 2 por ciento del área cubierta, salvo Santa Fe y Córdoba que tienen un 10 por ciento aproximadamente.

Nuestra propuesta es crear, mediante una iniciativa público-privada, una entidad con abordaje 100 por ciento móvil que provea asistencia crediticia y medios digitales de pagos de servicios que tenga como fin avanzar en la inclusión financiera. Dado que Argentina posee una gran parte de su población económicamente activa bajo la modalidad de trabajo “micro-informal” se propone una iniciativa público-privada a modo de poder atender a este público mediante la posibilidad de depositar dinero en efectivo en este nuevo banco digital (suponiendo que parte o la totalidad de su salario es abonado en efectivo) mediante la red de sucursales del Correo Argentino. Es decir, dado que el banco tiene un abordaje 100 por ciento digital, la idea es que las personas posean o no una Caja de Ahorro puedan acreditar su dinero (efectivo) en estas nuevas “sucursales” que servirían como canales de depósitos, además de cualquier otro canal tradicional. Cabe destacar que la red comercial del Correo Argentino se encuentra conformada por más de 3.300 puntos de venta en todo el país. Por lo tanto, es la red más extensa del ámbito nacional y permitiría a los clientes del banco digital acceder a depositar su dinero prácticamente en cualquier lugar y de manera segura. No obstante, también sería interesante abordar la idea de poder acreditar el dinero mediante otros canales, como podría ser estaciones de servicio de carga de combustible, por ejemplo, YPF cuenta con más de 1.520 estaciones a nivel nacional.

---

<sup>31</sup> *Global Findex database: Adults with an account paying utility bills in the past year in cash 2017*



## **Permitir o facilitar el acceso al crédito a aquellas personas que se encuentran excluidas.**

Bajo la premisa de “todo dato es dato de crédito” se plantea la creación de una plataforma o banco con abordaje 100 por ciento digital cuyo motor principal sea un *score* de crédito basado en análisis de *Big-Data* y técnicas de *Machine Learning*. Es decir, la manera de lograr bancarizar a las personas no bancarizadas es mediante el desarrollo de un *scoring* que no sólo utilice información tradicional de *Bureau* (financiera) sino que se pretende adoptar un enfoque integral de información crediticia no convencional mediante puntos de datos extraídos de actividades “en línea” (*online*) y “fuera de línea” (*offline*) de los consumidores. Es decir, cada vez que estas personas no bancarizadas, que poseen telefonía móvil, hacen una llamada, envían un mensaje de texto, navegan por Internet, utilizan las redes sociales en sus teléfonos o recargan sus tarjetas prepagas, profundizan las “huellas digitales” que dejan atrás. Más precisamente, los datos de registros de teléfonos móviles, recargas prepagas, pagos de facturas móviles y navegación móvil o historial de descargas de aplicaciones se pueden utilizar para evaluar el riesgo del consumidor y determinar la solvencia de los clientes no bancarizados o desatendidos. En cierto punto un *score* de crédito mediante la aplicación de *Big-Data* y técnicas de *Machine Learning* puede basar la decisión de crédito analizando variables como en “dónde compran” las personas, “las compras” que hacen, sus “redes sociales” y otros factores que no están intuitivamente relacionados con la solvencia crediticia del individuo. Bajo este *score* los prestamistas pueden usar el resultado de su calificación crediticia para ofrecer crédito no garantizado, de corto plazo y a un costo mucho menor que los préstamos tradicionales.

### **La importancia de tener un Score de Crédito**

Al comienzo del presente trabajo hemos mencionado cuan importante es tener un *score* de crédito. Poseer un buen *score* de crédito le permite a las personas y empresas tener no solo acceso al crédito y financiamiento, sino que, además es necesario al momento de adquirir una vivienda, a una educación superior, o establecer un negocio.

También hemos mencionado que, hasta el momento, la industria bancaria tradicional no ha mostrado mucho interés en esta gran población de prestatarios no bancarizados. Por lo tanto, este segmento de la población se ha visto forzada a obtener préstamos con un alto costo (tasas de interés usureras) de prestamistas no tradicionales.

La idea central es que cada línea de crédito otorgada por esta entidad (principalmente línea de microcréditos) sea luego monitoreada e informada a la Central de Deudores del Banco Central de la Republica Argentina. De esta manera se estaría creando crédito a este grupo desatendido y generando información “tradicional” que utilizan los *Bureaus* de crédito para el desarrollo de sus modelos de *score* que luego son consumidos o utilizado por los bancos tradicionales a modo de poder romper con el bucle continuo de exclusión financiera.

Algunos casos de éxito sobre este tipo de iniciativas han demostrado que una persona sin historial crediticio posee un *score* de entre 350 y 450. A partir de que se otorga un

crédito y la persona va realizando sus pagos en tiempo y forma, el *score* de este grupo ha sufrido aumentos de, en promedio, 170 puntos<sup>32</sup> tal y como se ilustra en la Figura 14 de abajo. Obtener un *score* por encima de 550 puede ser una gran oportunidad para ingresar al sistema financiero tradicional y así obtener más oferta de productos.

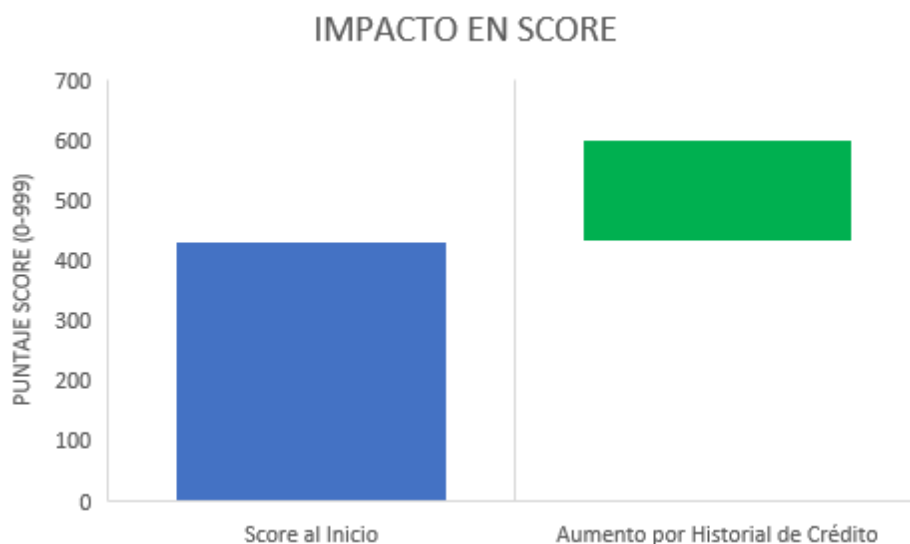


Figura 14

En la Argentina, el potencial de crecimiento de este segmento parece no tener techo. Un país de ingresos medio-altos, con una economía considerablemente desarrollada, bajas tasas de bancarización (50 por ciento de la población, según el Banco Mundial) y acceso a tarjetas de crédito (25 por ciento, según el Informe Ecosistema Fintech Argentino) presentan no solo un desafío sino también una oportunidad única para el desarrollo y la innovación financiera.

## ANEXO: Casos de estudio exitosos en *Big-Data*

Finalmente, a modo de enumerar algunos casos de éxito sobre esta nueva manera de calificar a los clientes mediante *Big Data* y técnicas de *Machine Learning* se presentan algunos ejemplos internacionales que se encuentran operativos hace varios años:

### India 1

- Encontramos<sup>33</sup> un acuerdo de inclusión financiera mediante la asociación comercial innovadora entre un proveedor de servicios móviles (Vodacom) y una compañía (First Access) centrada en la explotación de datos (analytics) que utiliza datos móviles prepagos para predecir el riesgo de crédito para aquellos consumidores que nunca han tenido una cuenta bancaria. En este acuerdo First Access ofrece un *score* en tiempo real para clientes de bajos ingresos aprovechando datos demográficos,

<sup>32</sup> <https://missionassetfund.org>

<sup>33</sup> Segundos usuarios más grandes de teléfonos móviles en el mundo.

geográficos, financieros y de redes sociales de los registros de los teléfonos móviles de los suscriptores.

El proceso de esta plataforma es relativamente sencillo: los suscriptores autorizan el *score* por mensaje de texto, éste es enviado a las instituciones financieras participantes en tiempo real junto con una recomendación sobre el monto del préstamo y el criterio de aceptación para el desembolso instantáneo.

## India 2

- En 2009 el Gobierno de India lanzó un sistema de tarjetas de identificación biométricas basadas en tecnología digital que proporcionan otra forma de reducir las barreras para la titularidad de la cuenta. La estrategia base de inclusión financiera se ha basado en tres ejes principales: 1) un programa de inclusión financiera del Gobierno (*Jan-Dhan*), 2) la creación de un número de identidad único de 12 dígitos (denominado *Aadhaar*) que los residentes de la India pueden obtener de forma voluntaria, este número se genera sobre sus datos socio demográficos o personales. A la fecha *Aadhaar* es el sistema de identificación biométrica más grande del mundo.

Este sistema tuvo como objetivo principal integrar el sistema de identificación y las aplicaciones de pago con el fin de lograr no solo ser más rentable, sino también más eficiente en el proceso de apertura y acceso a cuentas bancarias.

El sistema de identificación de 12 dígitos incluye un servicio electrónico *KYC* (*e-KYC*<sup>34</sup>) para acelerar la verificación de la identidad del cliente. El *e-KYC* otorga la posibilidad a una persona que posea un número de *Aadhaar* permita a la Autoridad de Identificación Única de la India (UIDAI) revelar su información personal a cualquier proveedor de servicios que éste individuo desee activar de manera instantánea, ya sean servicios de conexiones móviles como cuentas bancarias. Al 2018 aproximadamente se han realizado un total de 4.900 millones de transacciones de *e-KYC* a través de *Aadhaar*.

## Alemania

- Encontramos el caso de un prestamista 100 por ciento online (Kreditech) que mediante técnicas de *Machine Learning* desarrolla un *score* analizando más de 15.000 datos. La información utilizada es recopilada mediante *cookies*, ubicación (GPS), el comportamiento del navegador y las redes sociales. A la fecha tiene más de 800.000 clientes con más de 2 millones de préstamos otorgados en distintos países como Rusia, República Checa, España, México y Polonia.

---

<sup>34</sup> Por las siglas de "Know Your Customer" en inglés.

## USA

- ZestFinance (anteriormente ZestCash), una compañía de tecnología de servicios financieros con sede en Los Ángeles que utiliza técnicas de *Machine Learning* para ayudar a las empresas a tomar decisiones de crédito más precisas.

Alguno de los impactos más conocidos de sobre los modelos de esta compañía son que han logrado disminuir un 33 por ciento la tasa de incumplimiento en una financiera de automóviles y aumentar la tasa de aprobación en un 14 por ciento (*Prestige*), han logrado un aumento de USD 800 millones en expansión de crédito para el grupo top 10 de proveedores de tarjetas de crédito, etc.

## Hong Kong

- Lenddo es una empresa exitosa con abordaje digital radicada en Hong Kong que se dedica al desarrollo de modelos predictivos de evaluación de crédito analizando redes sociales en múltiples geografías como Filipinas, Colombia y Brasil. El modelo de *scoring* (algoritmo patentado) se realiza sobre la base de miles de puntos de datos recopilados de la actividad de las redes sociales en múltiples plataformas. Algunos de los puntos de datos utilizados incluyen la cantidad de cuentas de redes sociales vinculadas al cliente, la cantidad de amigos y seguidores de las redes sociales, la duración del tiempo activo en las redes sociales y la fortaleza de la red social del cliente (el último de los cuales es evaluado por cuántos amigos avalan la solvencia del cliente).

La ventaja de esta compañía es que es muy fácil de usar para las instituciones financieras. Funciona a través de una API basada en la nube. Una institución financiera puede agregar un poco de código a su flujo de trabajo de la aplicación de préstamos que luego les permite a los clientes optar y compartir sus redes sociales. Lenddo también ofrece un kit de desarrollo de software (SDK) que los bancos pueden agregar a una aplicación móvil para que los usuarios compartan las redes sociales.

## REFERENCIAS

- Banco Mundial, Banco de datos. Indicadores del desarrollo mundial, 2018
- BCRA. Ideas de Peso. Más Puntos de Acceso, Menos Efectivo, 2018.
- BCRA. Informe de Estabilidad Financiera, segundo semestre, 2017
- BCRA y CAF. Encuesta de medición de capacidades financieras en Argentina, 2017
- CGI Group Inc. Understanding Financial Consumers in the Digital Era. A Survey and Perspective on Emerging Financial Consumer Trends, 2014.*
- F. Nogarei y I. Carballo. Inclusión Financiera, Microfinanzas y Fintech: los casos de México, Chile y Perú, 2018
- G20 Digital Identity Onboarding – Argentina 2018*
- GSMA. *Country Overview: Argentina. Impacto del ecosistema móvil: perspectivas y oportunidades, 2016*
- INDEC. Mercado de trabajo. Tasas e indicadores socioeconómicos (EPH). Primer trimestre de 2019.
- M. Hurley and J. Debayo. *Credit Scoring in the Era of Big Data, 2017*
- Ministerio de Hacienda. Estrategia Nacional de Inclusión Financiera, 2019
- UCA. Documento de Trabajo. Heterogeneidad y Fragmentación del Mercado de Trabajo (2010-2018).
- World Bank Group. The Global Findex Database, 2017
- ZestFinance, *Our Story*, <http://www.zestfinance.com/our-story.html>