

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE ECONOMIA

TESIS

Aplicaciones del *Big Data*, en el análisis y administración de riesgos inherentes a los consumidores del sector bancario



Imagen generada por IA de freepik.com, empleando el título de la tesis para su creación

PARA OBTENER TITULO DE LICENCIATURA EN ECONOMIA

P R E S E N T A :

RIVERA FRANCO JOSE RUBEN

ASESOR:

MIGUEL ANGEL JIMENEZ VAZQUEZ

Ciudad Universitaria CD. MX., 20 de Junio del 2023



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

| | |
|---|------------|
| 1 Empleo de la Información en el Sistema Financiero..... | 9 |
| 1.1 Sistema Financiero, Su Efecto Intermediación..... | 9 |
| 1.2 La Importancia de la Información y del Riesgo; Base del Comportamiento de los Agentes. | 15 |
| 1.3 Intermediación financiera y el empleo de la información..... | 20 |
| 1.4 La naturaleza de los Riesgos y su Administración | 30 |
| 2 La Información del Big Data..... | 45 |
| 2.1 Definición del Big Data | 47 |
| 2.2 Características de la información..... | 54 |
| 2.3 El Beneficio Del Análisis Del Big Data..... | 61 |
| 2.4 Herramientas para el Análisis..... | 73 |
| 2.4.1 “Bussines understanding” o Características Empresariales. | 75 |
| 2.4.2 “Data understanding” o Entendimiento de la información. | 78 |
| 2.4.3 Visualización de la información. | 83 |
| 2.4.4 Análisis y Modelación de la información..... | 87 |
| 2.5 Dificultades con el uso del Big Data..... | 91 |
| 2.5.1 Mala praxis; Oportunismo e invasión a la privacidad..... | 96 |
| 3 Empleando el Big Data en el Sistema Bancario..... | 104 |
| 3.1 Riesgo Externos, Existencia del Mercado | 109 |
| 3.1.1 Análisis de Sentimientos (Costumer Understanding) | 121 |
| 3.1.2 “Scoring” en base a información alternativa..... | 132 |
| 4 Conclusiones..... | 137 |
| 5 Referencias | 142 |

INTRODUCCIÓN

Con el continuo progreso en las tecnologías de la información y la masificación de los dispositivos electrónicos, el internet se ha convertido en una herramienta vital para la conexión e interacción entre los agentes económicos. Últimamente, se han creado en el internet diversos ecosistemas virtuales que ayudan a dicha interacción, tales como: tiendas virtuales, redes sociales, blogs de temas específicos, servicios electrónicos automatizados, etc. Cambiando la forma, el ritmo y los mecanismos con los cuales, tanto el oferente como el demandante intercambian información para la composición y ejecución de las transacciones económicas.

Dichas innovaciones tecnológicas, traen consigo distintos beneficios económicos, tal como la reducción de los costos de transacción para la resolución de acuerdos económicos, reduciendo en tiempo y recursos requeridos para la exploración de los precios del mercado; las plataformas y herramientas virtuales le permiten al consumidor acceder a una amplia variedad de información acerca del mercado (tal como precios, productos y proveedores). Al mismo tiempo, la interconexión a través de medios digitales permite a los usuarios negociar y establecer acuerdos entre los agentes, los cuales pueden estar ubicados en distintas regiones del mundo. Así mismo, al emplear dichos mecanismos electrónicos, para la interacción y negociación; los agentes económicos dejan una estela de información, la cual es capturada dentro en distintos medios de la red, tales como: correos electrónicos, documentos; opiniones, comentarios y reseñas sobre el bien consumido, información capturada en los formularios de registro o venta, dirección IP, fecha y hora de acceso, etc.

De acuerdo con Oracle, empresa dedicada al almacenamiento digital, define como Big Data a toda la información almacenada que sobrepasa la capacidad de los sistemas tradicionales (Oracle Corporation, 2019), dicha información se

encuentra contenida en diferentes servidores, la cual, con la constante interacción de los nodos o dispositivos en red; se crean grandes cantidades de información que es actualizada continuamente y en tiempo real. Debido a que cada dispositivo es operado por un agente distinto, este da al dispositivo una determinada finalidad en base a sus necesidades; por lo tanto, cada dispositivo conectado juega un determinado rol dentro del sistema, la información generada por estos, posee características distintas en cuanto su estructura, su contenido, su lenguaje y su finalidad.

Estas peculiares fuentes de información, posee datos con distintas características distintas a las obtenidas tradicionalmente por encuestas o registros. De tal modo que, para su comprensión y empleo en distintas áreas, se requiere de la creación de nuevas herramientas que permitan el procesamiento y su análisis; a este tipo de información es usualmente conocida como: "información alternativa" (tal como fotos, videos, e-mails, textos, audio, llamadas, etc.)

Los datos contenidos en este tipo de registros, se le conoce como "No-Estructurada", la cual se deja de lado para el análisis estadístico, ya que el procesamiento de cada registro u observación requiere de su interpretación individual, para ello son requeridas las capacidades cognitivas humanas para la extracción de información, requiriendo también el tiempo y los costos de dichas acciones; lo cual hace que el valor de dicha información sea opacado por el alto costo que implica su análisis manual. El desarrollo de modelos analíticos denominados "Data Mining" que permiten minimizar dichos costos, haciendo viable el empleo de la información No-Estructurada para la toma de decisiones.

Con el desarrollo de estas herramientas que permiten tanto el análisis como la compilación, de la información "estructurada" como la "No-Estructurada", le da los agentes nuevas capacidades para el desarrollo de conocimientos, requeridos para la toma de decisiones y colocación eficiente de los recursos.

Dentro del sistema bancario, la implementación de dicha tecnología puede complementar la información que es obtenida de los clientes, mejorando la capacidades que poseen los bancos para el análisis y administración de los riesgos que los clientes traen consigo, especialmente de aquellos que poseen muy poca información "Tradicional", tal como clientes con corto o nulo historial crediticio (Jagtiani & Lemieux, 2018); De acuerdo con la compilación elaborada por Evry (2014), empresa dedicada al desarrollo de soluciones de Big Data, existen varios casos en donde la implementación de dicha tecnología, ha ayudado a mejorar el entendimiento de los clientes y sus preferencias individuales, mejorando así las capacidades de las compañías para el diseño de productos, promociones y/o sugerencias que se ajustan a las necesidades, incrementando la satisfacción de sus clientes y al mismo tiempo reduciendo los factores de riesgo.

Al mejorar la capacidad productiva del conocimiento, también, el agente puede perfeccionar sus propias capacidades operativas. Tal como la revista Mckinsey (2018) informa, esta tecnología puede ser empleada para la creación perfiles psicológicos, con la finalidad de agrupar los clientes en dichos perfiles, de los cuales cada grupo pose distintas características y niveles de riesgo. Dando herramientas, que permiten mejorar los procesos y las acciones requeridas para el manejo de los riesgos de contrapartida; los cuales, son originados ante la posibilidad de la selección adversa de los clientes.

Con las herramientas que se diseñan para el análisis de la "información alternativa", en el sistema bancario, es posible a minimizar los efectos generados por la asimetría en la información y las fricciones económicas que limitan la entrada de potenciales clientes. Lo cual, al mismo tiempo limita tanto a la congregación de proyectos de inversión como a los planes de consumo; restringiendo a la fluidez y el uso de los recursos que el banco administra.

Debido a la aversión del riesgo y las características de los mercados de información asimétrica, la fijación de la tasa de interés, el monto de los créditos otorgado y el nivel de colateral; son colocados en niveles subóptimos en donde el equilibrio de dicho nivel incrementa los costos de los clientes para poder acceder al crédito, en donde el banco delega parte de los riesgos al consumidor por la asimetría del mercado; afectando al comportamiento de los agentes al usar dichos recursos de forma ineficiente, tal como la inversión en proyectos de alto riesgo para compensar el sobrepago de la tasa de interés. como indica Stiglitz y Weiss:

“However, the interest rate a bank charge, may itself affect the riskiness of the pool of loans by either: sorting potential borrower (adverse selection effect); or affecting the actions of borrowers (incentive effect). both effects derive directly from the residual imperfect information which is present after banks have evaluated loan applications” ... “The amount of the loan and the amount of collateral the bank demands of loan applicants will also affect both of the behavior of borrowers” ... “increasing interest rate or collateral requirements could increase riskiness of the bank’s loan portfolio, either by discouraging safe investor or inducing borrowers to invest riskier projects.” (Stiglitz & Weiss, 1981 pp. 393).

Ante dichas imprecisiones que el intermediario financiero posee en la administración de los riesgos; sus actividades y resultados perjudican a la economía en general. Al establecer la tasa de interés en un nivel superior al riesgo, se inducen a conductas con un mayor riesgo por parte del cliente, con lo cual aumenta la probabilidad de la materialización del riesgo, incrementando las pérdidas económicas en general, de las cuales son contagiadas entre los agentes relacionados.

Parte sustancial en las funciones de un intermediario financiero, radica en el manejo y la planificación del riesgo. (Merton, 1989) Al tener distintas clases de riesgos en los diferentes productos financieros que son manejados por el intermediario, el proceso de intermediación se vuelve cada vez más intrincado al incrementar el número de participantes dentro del sistema en el cual el intermediario interactúa (Allen & Santomero, 1997). De tal forma, los procesos que

permiten diluir y administrar los riesgos absorbidos, se vuelve un proceso cada vez más costoso ante la continua complejidad y variedad de transacciones. Por lo que, para su gestión efectiva es requerido de un monitoreo continuo de dichos riesgos, siendo necesario un análisis constante y diligente de un flujo incesante de información, con la finalidad de poseer el conocimiento actualizado de los riesgos existentes en los mercados, esto, para poder resultar en una óptima colocación de los recursos.

Dentro de las innovaciones tecnológicas, se han desarrollado procedimientos que emplea modelos estadísticos automatizados, de los cuales se presume es posible obtener un inmenso conocimiento en tiempo real, en base al análisis de toda clase de información contenida en la red, incluyendo la información que no había sido posible utilizar con los procedimientos anteriores. Aunado a ello, son mejoradas las capacidades de los agentes de recopilar información a causa del desarrollo de los canales de comunicación digitales (páginas web, aplicaciones, etc.), de esta forma, es posible recolectar masivamente información a partir de algoritmos automatizados que pueden buscar, recopilar y unificar la información contenida en medios digitales.

Dichas innovaciones tecnológicas cambian la toma de decisiones de los agentes, ya que es mejorado el proceso productivo de conocimiento, al tomar un mayor volumen de información a partir de nuevas fuentes. En el sistema bancario, brinda a los intermediarios financieros ventajas competitivas al reducir los tiempos de respuesta en sus procesos operativos y así mismo, les permite desarrollar herramientas que brindan un vasto conocimiento sobre el ambiente y sobre la sinergia que los activos poseen en su desenvolvimiento dentro del sistema, modificando de esta forma los procesos con los cuales son analizados los riesgos.

De acuerdo con la teoría evolutiva (Nelson & Winter, 1982), es destacada la utilidad de la información, el flujo que esta posee y la capacidad del agente de

obtener el conocimiento a partir del ambiente; dichas características moldean el comportamiento óptimo de los agentes, tanto en los procesos productivos como en los procesos de consumo, los cuales son expresados como planes de consumo y planes de producción; se definen como planes ya que el agente debe de elegir la distribución de sus recursos de acuerdo con sus expectativas futuras (de tal modo que no es posible conocer la utilidad o el beneficio del plan, hasta que una vez se haya concretado).

En base a la información, la cual es transformada en conocimiento, los agentes modifican y adaptan su comportamiento de acuerdo a las circunstancias del ambiente; De tal modo, que los agentes están en un proceso continuo de adaptación, en el cual su comportamiento cambia paulatinamente con la finalidad de obtener la mayor utilidad posible en un sistema altamente dinámico. Bajo estas premisas y ante la continuidad del comportamiento individual, se le pueden ser atribuido ciertas características de predictibilidad en la conducta de los agentes, basados en su comportamiento anterior, en la información que es consumida por los agentes, sus experiencias y de sus rasgos individuales.

La presente investigación tiene como objetivo general analizar las innovaciones tecnológicas relacionadas con el término "Big Data", explorando a nivel mundial, los usos y desarrollos de dicha tecnología, con la finalidad de describir sus capacidades y alcances; desde su origen en el 2006 hasta la fecha en la que se ha sido realizado esta tesis.

Como se ha descrito, se espera que dichas innovaciones mejoren las capacidades de la toma de decisiones de los agentes, lo cual deberá verse reflejado en su funcionalidad y el servicio que brinda al cliente.

Para poder analizar dicho fenómeno, en el primer capítulo se ha realizado una exploración teórica que permitirá describir la función o el papel que juega la información dentro de las decisiones económicas de la firma, enfocado

principalmente a las instituciones bancarias; de tal modo, se espera que sea posible especificar los aspectos relevantes que son afectados por el desarrollo de la tecnología "Big Data" e identificar sus impactos en dichos aspectos.

En el segundo capítulo, se examinarán algunos elementos básicos y sustanciales de las recientes innovaciones tecnológicas que permiten manejar el "Big Data", recabando información técnica y ejemplos de cómo ha sido empleada dicha tecnología internacionalmente, observando los resultados que se han obtenido. Para finalmente, en el tercer capítulo se hace discusión de como dicha tecnología es empleada en la administración de los riesgos de contrapartida (riesgos provenientes del consumidor) e infiriendo teóricamente el impacto de su implementación y desarrollo.

Capítulo I

1 Empleo de la Información en el Sistema Financiero.

1.1 Sistema Financiero, Su Efecto Intermediación.

El sistema financiero es un mecanismo basado en transacciones, el cual está conformado por un conjunto de mercados, reglamentos e instituciones; en el cual, los recursos y valores producidos dentro de la economía real (como los bienes y servicios) pueden ser negociados mediante el empleo de contratos estandarizados o elaborados a la medida, los cuales forman canales o vías de comunicación por los cuales es transferido el valor de diversas mercancías. (Acevedo Balcorta, 2012)

Las instituciones financieras, desempeñan la labor de intermediación dentro del mercado financiero; son estas, las que poseen la capacidad de transmutar los valores del público a contratos de propiedad (o títulos financieros). Para ello, es requerido de un sistema regulador el cual le permita dar confianza al público general sobre la legitimidad tanto del título financiero como el de su emisor, funcionando el agente regulador como un agente neutro en el caso de disputas entre las partes involucradas, a causa de algún desperfecto dentro del mecanismo de conversión del título; Existiendo de esta manera: riesgos en la liquidez del título, así como la existencia de fricciones al emplear el mercado financiero.

El sistema financiero puede ser ilustrado como un plano virtual, en el cual es proyectada la economía real. En donde el valor de los bienes es trasladado por medio de acuerdos, que desde un plano tangible es enviado a otro plano en donde la "palabra" (la promesa o título) es aquel que representa el valor del activo real, convirtiéndose en un valor fiduciario, el cual deberá de ser reconvertido a su valor de uso en algún momento futuro.

La transmutación del plano de la economía real al plano de la economía financiera, surge ante la necesidad de reducir los costos generados por la

administración de los activos. por ejemplo: el dinero agiliza el proceso de transferencia de valor, permitiendo el intercambio de una mercancía definida por otra que no está definida hasta en un momento del futuro, cuando vuelve a ser convertido en mercancía. debido que la moneda es fácilmente portable, divisible, de valor durable y reconocible por la sociedad; las monedas poseen un costo menor de administración que cualquier otro bien, por lo que es preferible almacenar una moneda que cualquier unidad de algún otro bien.

Los activos generan a su portador, costos de su administración por su simple adquisición y custodia; los diversos costos están compuestos por las distintas características del activo: generados por el tiempo de su preservación y al ejercer los derechos de poseer dicho bien.

Existiendo costos asociados por el deterioro o modificación en el valor del activo, siendo esta causada por la depreciación o cambios en el valor comercial del activo; costos generados por el almacenamiento, su movilidad y la protección o custodia del activo; costos asociados al uso y transferencia de los derechos de propiedad, generando costos por el empleo del mercado a causa de requerir información sobre el mercado, requiriendo conocer las existencias de la mercancía y los precios de esta, con la finalidad de llevar a cabo procesos de negociación para la adquisición o liquidación del activo, así también, se tienen costos por los procesos legales en caso de disputas relacionadas ante el incumplimiento del convenio previamente realizado; siendo estos dos últimos conocidos como los costos asociados por el empleo de un mecanismo de precios. (Coase, 1937)

En síntesis, los costos asociados a la administración de un activo son: el costo de la preservación de su valor, el costo de información, el costo de transacción (o negociación) y el costo por la custodia del activo. Ante la existencia de estos costos, el sector financiero es empleado para reducir tanto los costos

mencionados como los riesgos individuales que envuelven a cada una de las transacciones.

*Def. {Genérica} **Riesgo (Risk)** Probabilidad de sufrir pérdidas, al proseguir con los objetivos previamente definidos, el riesgo puede provenir de factores que se encuentran bajo nuestro control o por factores externos.*

Por medio de los títulos financieros, es unificado un complejo sistema económico de características competitivas; de acuerdo con Hayek (1948) una sociedad es un sistema en el cual está conformado por una amplia gama de agentes (hogares y empresas); en donde cada agente (competitivo) se encuentra en un estado de planeación descentralizada, es decir, cada individuo ajusta sus "planes" a las condiciones de su ambiente, creando un ambiente de discordancia (o competitivo). Dentro del sistema financiero, los agentes participan simultáneamente por medio de los distintos productos financieros existentes (como: opciones, swaps, forwards, seguros, créditos, fianzas, fondos de inversión, bonos, stocks, etc.) que poseen características específicas, los cuales absorben los rasgos y características de la economía real que reflejan (Allen & Gale, 1998).

Para cada mercado y para cada necesidad individual: existen distintos esquemas de transacción por los cuales los participantes adquieren y transfieren información, por medio de su interacción y coordinación con el mercado (Hayek, 1958); Creando distintas clases de productos financieros y distintas clases de señales, que corresponden a la necesidad de los distintos "planes"; en donde los planes que los agentes crean, se encuentran determinados por el conocimiento y la habilidad que estos poseen¹, tanto para la elaboración de los planes como para

¹ Entiéndase por "habilidad", como la secuencia sucesiva, automatizada y coordinada de acciones que permiten cumplir objetivos determinados; el conocimiento le da estructura a las rutinas, las cuales, en términos económicos, poseen un bajo costo debido a su característica involuntaria (no es necesario pensar) y un pago conocido el cual se espera que sea el más alto. El conocimiento

el aprovechamiento del valor de los activos. (Grant, 2001) De esta forma, dentro del mercado financiero, la información que se posee es una herramienta para la creación de valor, ya que su posesión permite obtener el máximo rendimiento de los activos que se administran.

Parte importante del comportamiento del agente (para evitar su extinción), se centra en las decisiones financieras, las cuales determinan el compás de su consumo dentro de un ambiente de incertidumbre. Los agentes económicos emplean sus recursos en el momento que se considera que la utilidad de consumir el bien "X" sea la mayor; en términos de la psicología evolutiva (Tooby & Cosmides, 1997): el comportamiento de los agentes económicos es adaptado en base a la estructura cognitiva que el agente posee, la cual se desarrollada mediante la toma rutinaria de decisiones, que le permiten saciar sus necesidades en el momento más oportuno con la finalidad de preservar su especie.

Asumiendo que tanto el mercado real como el mercado financiero son bienes sustitutos perfectos; al asumir, que los bienes que se adquieren en ambos mercados son bienes homogéneos y cuya diferencia radica solamente en el tiempo de su consumo. En consecuencia, dicho bien posee un mismo uso y las mismas cualidades en los distintos tiempos ($I^{t_0} = I^{t_i} = I$), de esta forma el consumidor optara por elegir aquel mercado que le produzca mayores beneficios (β) y menores costos (C) bajo un criterio basado en su propia percepción temporal: eligiendo con base al momento más oportuno para el consumo del bien y con base en su tolerancia a la postergación de las transacciones, a esto último se le denomina como factor de descuento ($\delta^{t_i-t_0}$), (Frederick, Loewenstein, & O'Donoghue, 2002)

brinda al agente la capacidad de diversificar, mejorar y adaptar su comportamiento. (Nelson & Winter, "An Evolution Theory of Economic Change", 1982)

Tabla 1, Costos Asociados al mercado Real y Financiero:

| | Mercado Real (Spot) | Mercado Financiero (Futuro) |
|------------------------|-----------------------------|---|
| Preservación del valor | Tasa de Depreciación | Tasa de descuento (interés e inflación) |
| Información | Búsqueda del valor correcto | Monitoreo del valor |
| Transacción | Proceso de Negociación | Proceso de Negociación y Liquidación |
| Custodia | Almacenamiento | Transferido al intermediario |

Al emplear mercado financiero, este posee sus propios costos de administración, los cuales están relacionados por el cumplimiento de las transacciones que serán llevadas a cabo en algún momento futuro, estos costos son: las tasas de descuento, el monitoreo del precio en el mercado y los costos de transacción asociados al proceso de liquidación.

$$U_{Real} = U_{Financiero}$$

$$I(\beta_0 - C_{Real})\delta^0 = I(\beta_i - C_{Financiero})\delta^{t_i}$$

$$\delta^{t_i-t_0} = \frac{\beta_i - C_i}{\beta_0 - C_0} := \text{Preferencia Temporal}$$

La decisión del agente de elegir entre el consumir en el "ahora" o en el "mañana", se centra en la anticipación de la utilidad ($\frac{\beta_i}{\beta_0}$), la previsión de los costos ($\frac{C_i}{C_0}$) y de su preferencia temporal ($\delta^{t_i-t_0}$), también conocido como "factor de descuento"; la decisión entre estos mercados, se basa en el trade off entre la utilidad obtenida de consumir el mismo bien "X" ahora o en un tiempo futuro "i". En donde la percepción intra-temporal de la utilidad del bien "X" está condicionada tanto por las necesidades actuales del agente, los hábitos que este posee y así también por la capacidad del agente de percibir la incertidumbre. De acuerdo con John Rae (1834), los hábitos están determinados por las facultades intelectuales que el agente posee, lo cual delimita la capacidad del agente de obtener el mayor

provecho de sus activos; siendo esta capacidad definida por Rae como "el deseo efectivo de acumulación"².

Para dar un ejemplo sencillo al *trade-off* mencionado, uno puede referirse al uso del paraguas; el cual, el paraguas solo es útil en el día y momento en que llueve, sin embargo, su costo siempre está presente al transportarlo diariamente. Imaginando que existe un producto financiero para este mercado: el agente compra un "forward", el cual limita el consumo del día "0" por el costo de adquirir este título y pasar el consumo al futuro; por lo cual este título será empleado por el agente para obtener un paraguas en el momento más oportuno (t_i) sin incurrir en el costo de acarrear el paraguas diariamente y protegerse del oportunismo de los vendedores de paraguas en los días lluviosos.

Con el constante desarrollo de los mercados financieros, se conforma un mecanismo complejo, lo cual dificulta tanto la capacidad de aprovechar el mercado como la capacidad de la diversificación (como estrategia de inversión); el uso del mercado se convierte en un proceso intricado debido al incremento de las combinaciones posibles y los costos de evaluar cada opción; de tal forma que se

² Es decir, a medida que el agente posee mayor inteligencia, este puede visualizar con prudencia los distintos beneficios que puede obtener al consumir el mismo bien en distintas fases del tiempo. Por el otro lado, el agente que posee un nivel bajo de inteligencia, sus acciones son dominadas por el instinto, la exaltación hacia los deseos y la necesidad de su inmediata satisfacción. De acuerdo con la neuroeconomía, las actitudes inmediatas estas siendo valoradas por características emocionales (sistema dopaminico: neuronas de corteza orbitofrontal y amígdala) mientras, que las recompensas futuras son previstas y calculadas por características "racionales" (neuronas localizadas en la corteza dorsolateral prefrontal), dando un sistema múltiple en la percepción de recompensas para la toma de decisiones. (Schultz, 2000), a lo cual Kahneman y Tversky (2011) lo denominan como el "sistema dual" (lento o rápido)

vuelve cada vez más costoso el monitorear por cuenta propia los precios de los "n" productos financieros. (Allen & Gale, 1998)

Así mismo los costos individuales que son asumidos por la participación en los mercados, se vuelven cada vez mayor ante el incremento continuo de los participantes; siendo que, para la correcta función del mercado es necesario la creación de canales de comunicación eficientes entre los participantes, que permita la mejor distribución de los recursos. (Allen & Santomero, 1993) es decir: Individualmente es difícil establecer y mantener los suficientes vínculos dentro del mercado, que permitan el aprovechamiento o beneficio de la negociación directa.

Por ejemplo: un solo inversionista no podrá acaparar o soportar la demanda de los créditos de alguna empresa de gran magnitud; al mismo tiempo, para la empresa sería costoso localizar individualmente a los inversionistas necesarios y establecer contratos distintos con cada uno de ellos. El costo de encontrar y mantener los vínculos será cada vez mayor a medida que se administre un creciente volumen de recursos e inversionistas. De esta forma, el intermediario es empleado como una herramienta con la cual se centraliza los costos de la dicha administración, siendo asignado como un "monitor delegado". (Diamond, 1996)

1.2 La Importancia de la Información y del Riesgo; Base del Comportamiento de los Agentes.

Dentro de un mercado de información perfecta (o en el ambiente del perfecto Estatus Quo), " β " se define como la expectativa o el beneficio esperado que produciría el consumir el bien "X" en un determinado tiempo bajo las características específicas del ambiente en dicho momento; El agente crea información a través de la experiencia al consumir determinado bien, siendo esta experiencia la satisfacción o el desagrado experimentado por el agente, la cual modifica la conducta del agente (Freud, 1940). Con esta información, el agente puede determinar el más conveniente para consumir el bien "X", con lo cual se

forman patrones en su comportamiento con la finalidad de maximizar su beneficio, siendo estos patrones conocidos como: "rutinas" (Nelson & Winter, 1982).

El aprendizaje permite evaluar y anticipar la utilidad de consumir determinado bien en determinado momento ($\vartheta(x)_{t_i}$) de tal modo que el factor que determina la maximización de su beneficio, se encuentra en su tasa de aprendizaje del agente; ya que el agente determinará sus acciones habituales que le permitan obtener el mayor beneficio de su consumo inmediato ($\vartheta(x)_{t_i}$), siendo este beneficio esperado formado por las recompensas y aprendizaje obtenido del sistema mesolímbico. (John P. O'Doherty, 2003) Mientras que la corteza dorsolateral (dlPFC) permite desarrollar la mejor estrategia al planificar el consumo futuro a expensas del consumo inmediato, de tal modo que la activación de las neuronas dlPFC inhibe los estímulos del consumo inmediato (Ballard K, 2009), delimitando la preferencia temporal (δ^{t_i}).

De tal modo, el agente escogerá el tiempo adecuado que le permita saciar sus necesidades con el menor esfuerzo (costo) anticipando los costos de sus acciones ($C(x)_{t_i}$) y las posibles consecuencias de dichas acciones (riesgos). Por lo que el agente evaluara el beneficio de dos ventanas temporales delimitadas por estímulos emocionales de corto plazo (núcleo accumbens) y las capacidades cognitivas que permita la planeación y postergación del consumo (dlPFC)³:

$$\beta = \prod_{i=t-\infty}^{t_0} \vartheta(x)_{t_i} \delta^{t_i}$$

³ De acuerdo con Antonio Damásio y Bechara (2005); una conducta basada en la planeación (dlPFC) no necesariamente es la decisión óptima, ya que en esta se incluyen los posibles riesgos que se deberán de afrontar en la acción realizada. Es decir, el agente al ser adverso al riesgo puede sobre ponderar dichos riesgos. Y así mismo, una conducta basada en el factor emotivo, no siempre es la decisión óptima al subvalorar dichos riesgos y no corregir la conducta ante la presencia o conocimiento de dichos riesgos.

En estudios realizados en la universidad de Sudáfrica (Mudzingiri, Muteba Mwamba, & Nicolaas Keyser, 2018), se concluyeron que los componentes importantes que determinan el factor de descuento, son los conocimientos que uno posee sobre la educación financiera y el grado máximo de estudios; en los resultados obtenidos, se muestra que los individuos con "habilidades" financieras han modificado sus "rutinas" o decisiones, por lo cual logran obtener un mayor beneficio (en el experimento) al mantener una mayor paciencia, al aceptar la postergación de los pagos, con la finalidad de poder obtener bonificaciones.

Sin embargo, a causa de la simplicidad del ejercicio, dentro de un ambiente sintético; el experimento no es posible capturar el factor que existe dentro de un mercado de información imperfecta; en donde la "incertidumbre", el ambiente y exposición al riesgo afectan a los procesos que los agentes realizan para la toma de decisiones intratemporales:

$$E_t(X|\kappa) = U_{t_{i+1}} = \prod_{i=t_{-\infty}}^{t_0} \vartheta(x)_{t_i} \delta^{t_i} - \prod_{i=t_{-\infty}}^{t_0} C(x)_{t_i} \delta^{t_i} - P(\text{Risk}|\kappa)\epsilon$$

Basando el análisis del riesgo bajo los axiomas probabilísticos de Колмогорова (Kolmogorov); Se tiene la probabilidad de que algún evento adverso sea detonado por una acción ejecutada por el agente, para la realización de un determinado plan. De tal modo, dicha acción posee un posible desenlace sea planeado o no, del cual, dicho estado final es denominado "posible estado de la naturaleza" $P(\Omega) = 1$.

Dentro de un ambiente de Estatus Quo (sistema determinístico), los componentes del sistema (su ambiente) permanecen totalmente estáticos, estos componentes son perfectamente predecibles con la información que se posee (información perfecta) $P(\Omega|\text{Risk}) = \emptyset$ y $P(\Omega) = P(Z|\kappa) = 1$; Ante un mercado de información perfecta no existe la probabilidad de que pueda surgir algún costo

adicional a causa de algún imprevisto o complicación en la decisión tomada; por lo cual el plan del agente es ejecutado a la perfección $\therefore P(Risk|\kappa)\epsilon = 0$. Es decir, la probabilidad de materialización del riesgo es nula.

Debido a las condiciones "caóticas" del sistema, en el que se encuentra el agente, en cada posible plan de consumo o inversión que pueda ser llevado a cabo; posee en cada una de estos un estado inicial y un objetivo determinado (estado final), el cual, posee un conjunto de trayectorias posibles, las cuales son amoldadas por el ambiente en el que se lleva al cabo sus planes. El ambiente, en el estado de información imperfecta está conformado por componentes conocidos y componentes desconocidos (incertidumbre epistemológica), así mismo, los componentes conocidos pueden tener comportamientos desconocidos (incertidumbre aleatoria) los cuales modifican la trayectoria o posibles desenlaces de los planes elaborados por el agente, a causa de una secuencia de comportamientos no esperados, de los elementos que componen el ambiente. (Helton, 1996)

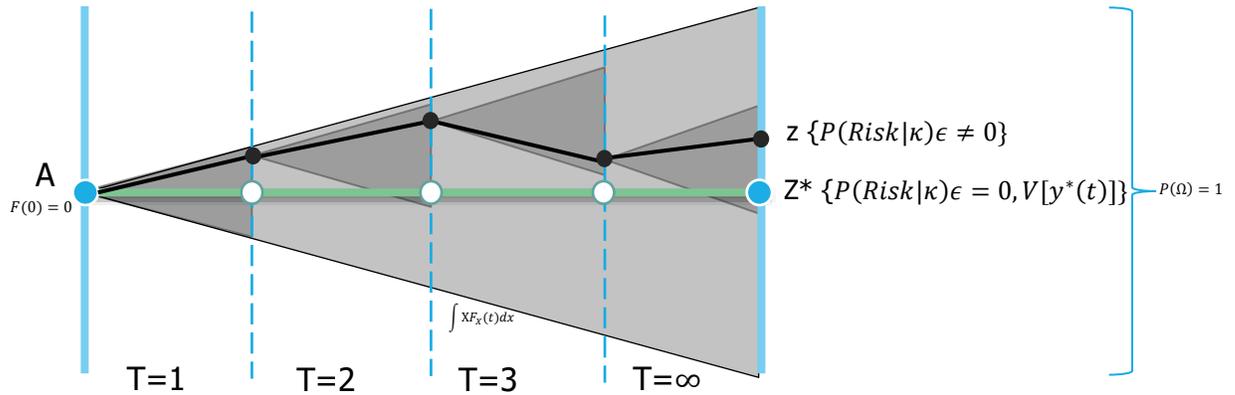
Debido a la naturaleza "caótica" de los elementos no conocidos, estos solo pueden ser estimados por métodos aleatorios, en donde existe siempre un factor o "variable no conocida", que bajo las mismas circunstancias altera el resultado obtenido, de esta forma las variables no conocidas poseen una función de distribución de probabilidad aleatoria, definida dentro de un modelo "lifetime function", como: $F_X(t)$

$$P_X(\Omega) = P_X(Z|\kappa) \cup P_X(Risk|\kappa) \cup F_X(t) = 1$$

- $P_X(Z|\kappa) :=$ Componente Incertidumbre Aleatoria, Z "Success"
- $P_X(Risk|\kappa) :=$ Componente Incertidumbre Aleatoria, Riesgo o Perdida
- $F_X(t) :=$ Componente Incertidumbre Epistemologica
- $\left[\lim_{t \rightarrow \infty} F_X(t) = 1 \right]$ No se posee ninguna certeza de lo que ocurrirá en el largo plazo

Ilustración 1, Representación gráfica de: Plan objetivo del agente y sus posibles resultados;

En donde la línea del centro (Z^*), representa la línea óptima bajo información perfecta; cualquier desviación de la línea representa la materialización del riesgo y el área sombreada representa las posibles conclusiones de las etapas (T) o "posibles estados de la naturaleza". Elaboración propia.



$$E_t(X|\kappa) = \left[X \sum P_X(Z|\kappa) - X \sum P_X(Risk|\kappa) \right] - X \int F_X(t)dt$$

Al ser la utilidad esperada, una variable subjetiva sujeta al conocimiento adquirido por el agente ante su constante participación con el mercado, la probabilidad de un evento es codificada ante los resultados obtenidos en la realización de planes anteriores (conocimiento previo " $\prod_{i=t-\infty}^{t_0} \vartheta(x)_{t_i} \delta^{t_i}$ "), cuando se tiene la creencia de un posible beneficio ante determinada eventualidad (evento pavloviano), el agente es estimulado por los sucesos del ambiente para realizar determinadas acciones que se creen convenientes ante la probabilidad de su beneficios o satisfacción (Day, 2007)

De esta forma, el conocimiento obtenido ante la observación y aprendizaje de los eventos pasados, determina las creencias del agente y por consiguiente su conducta habitual (Kaplan & Garrick, 1981); De acuerdo a los descrito con la ecuación de la utilidad esperada ($E_t(X|\kappa)$) se es conformado de esta forma un acervo de conocimiento (κ); el cual reduce el nivel de incertidumbre epistemológica (Arrow, 1973), pero siempre se mantiene presente la incertidumbre aleatoria. (ya

que permanecen conductas impredecibles de los factores de riesgo, aun después de haberlos conocido) (Flage & Aven, 2009) A medida que el agente que posee mayor conocimiento, tiene una mayor probabilidad de tomar decisiones acertadas que minimicen la diferencia entre la utilidad esperada y la utilidad obtenida (Hayeck, 1948)

$$\max_{k \rightarrow \infty} E_{t+1}(X|\kappa) = U_{t+1} \Rightarrow U^*$$

Bajo el sistema de información asimétrica (y distorsionada), la búsqueda por la información se vuelve una tarea cotidiana el cual posee un costo; dentro de un ambiente dinámico, la nueva información (relevante y específica) posee un valor el cual se atañe al incrementar la utilidad de cada plan elaborado por el agente, lo que le permite realizar planes más efectivos. (Ribeiro, 2004); de tal forma, el costo de adquirir una "unidad" de información adicional deberá de ser menor a la utilidad adicional que pueda brindar; solo así la unidad adicional de información permitirá al agente "actualizar" sus hábitos (o adaptar su comportamiento) y mejorar los beneficios obtenidos de sus planes. De esta forma, el agente posee el incentivo suficiente para pagar por la información adicional y así mismo, el agente posee el incentivo necesario para guardarse el conocimiento para sí mismo (como herramienta de competencia). (Demsetz, 1969)

$$\Delta C_K < \left(U_{t_0} - \sum_{i=1}^{\infty} E_{t_i} \delta^{t_i} \right)$$

1.3 Intermediación financiera y el empleo de la información.

Las instituciones financieras que ejercen la labor de la intermediación, tal como los bancos, poseen entre sus funciones la creación de información, la cual le permita ejercer una administración eficiente; que, con tal de maximizar su utilidad obtenida, el intermediario busca situaciones oportunas en donde el mercado sea capaz de valorizar los recursos que serán asignados por el intermediario.

Desarrollando de esta forma un conjunto de señales y mecanismos, los cuales le permitan al intermediario crear una organización interna que simplifique el traslado de los valores, tanto internamente como externamente, minimizando los costos generados por las medidas tomadas para la cobertura y el manejo de los riesgos; así también, de los costos y las fricciones que son generados por el monitoreo de los acreedores, los deudores y los elementos relacionados al mercado en el cual los valores asignados interactúan.

Los recursos recaudados mediante el fondeo, para poder ser utilizados, requieren de un proceso de vinculación con el mercado: en el cual los recursos son canalizados a diversos planes de consumo y planes de producción; estimulando la economía a través de la formación y acumulación de capital, materializando la innovaciones tecnológicas que se encuentran almacenadas en el stock del conocimiento (en el plano de las ideas), dichas innovaciones se encuentran plasmadas en los planes de inversión que los agentes formulan, materializando proyectos con diversas características y mecanismos de competencia (Schumpeter, 1911). De una forma poética, tal como Allan y Galle (1998) describen: el intermediario se convierte en una figura importante dentro de la sociedad, al transformarse en la voz de una sociedad, aprobando o desaprobando los proyectos de inversión que son propuestos, impulsando el desarrollo económico al financiar los proyectos de alta calidad, con las mejores perspectivas de retorno. (King & Levine, 1993).

Por su parte, la participación del banco en la colocación de recursos en los planes de consumo; el intermediario se convierte en un medio por el cual las transacciones pueden ser llevadas a cabo con mayor facilidad, permitiendo adquirir bienes que son difícilmente liquidados con efectivo, y así mismo crean mecanismos de comunicación que facilitan las transacciones comunes. Esto favorece los flujos de consumo, además de brindar la posibilidad de liquidar y/o garantizar la

liquidación futura de los productos que se han fabricado con los proyectos de inversión. (Levine & King, 1993)

Al ser los bancos, monitores delegados, estos asumen los costos asociados a la producción de información, con la finalidad de poder evaluar asertivamente los proyectos de inversión y buscar las coberturas necesarias sobre los riesgos que puedan existir en el ambiente en el que se desarrollaran; Al delegar dicha tarea al intermediario, reduce los costos individuales que cada inversionista debería de incurrir al participar directamente en el mercado financiero. (Ross Levine, 1997 y 2015). Por lo que su participación es sustancial, para crear un empuje en la economía para desarrollo de innovaciones tecnológicas, al agrupar y promediar los distintos riesgos inherentes a los distintos proyectos de inversión. (Schumpeter, 1912)

La implementación de innovaciones tecnológicas dentro de los procesos administrativos, ocasiona un mejor desempeño en la operatividad, ejecutando procesos con el menor tiempo y costo posible, reduciendo las fricciones creadas dentro de la firma (por los procesos en la tramitología y la eficacia en la transmisión de la información intra-empresa) (Bourini, 2013)⁴. Trasfiriendo sus beneficios a la economía real por medio del efecto intermediación, de tal manera, el intermediario mejora su capacidad de emplear y movilizar los recursos existentes con mayor eficiencia y flexibilidad.

Y por su parte, las innovaciones en la creación de instrumentos financieros, mejoran la efectividad en los canales por los cuales los recursos son trasladados, obteniendo de ello nuevas opciones tanto en la colocación de recursos como en su

⁴ Proceso conocido como "Knowledge Management Process", el cual, es un proceso técnico que describe y mejora los procesos por el cual la empresa captura, procesa, transmite y aplica la información dentro de ella para maximizar la productividad de esta.

recolección. De tal forma, la creación de dichos canales de comunicación transforma la estructura financiera; ergo la estructura económica real. (Qamruzzaman & Jianguo, 2017)

Tal como se ha descrito, las teorías de la intermediación afirman que un intermediario posee funciones trascendentales, ante la presencia de la asimetría de la información. Estas funciones son explícitamente: ser el administrador de los recursos y el monitor delegado, por ello, poseen inherentemente funciones dentro de su marco operativo la administración de los riesgos; Siendo esta función, una parte sustancial de su operatividad, la cual no está cubierta por la teoría de la intermediación actual. (Scholtens & Van-Wensveen, 2003)

Ante la existencia de la información imperfecta, a causa de la imposibilidad de los agentes de observar y computar todos los factores que afectan directamente a los precios; existen en el precio factores observables (η) y otras que no observables (aleatorias " ϵ ") (Grossman & Stiglitz, 1976). En donde existe un valor esperado (\bar{V}) por aquellos que poseen información y un valor aceptado por aquellos que no la poseen ($\bar{V} + \sigma_A$):

$$r = \eta + \epsilon$$

$$p = E(p) + \int F_p(t) dt$$

$$p = \bar{V} + \sigma_A$$

$$r \neq \eta, \text{En condiciones de informacion asimetrica} \therefore \exists \sigma_A$$

El incentivo del intermediario de existir, de acuerdo a la teoría de la intermediación, se encuentra en la diferencia entre el valor intrínseco (Valor esperado) y el valor comercial de un activo (Valor aceptado):

$$\text{(Ecuación 1)} \quad W_i \frac{(V_A - \bar{V})}{\bar{V}} > C_K$$

$$\frac{(V_A - \bar{V})}{\bar{V}} := \text{Valor exedente sobre valor medio "}\bar{V}\text{" del activo}$$

$$[(V_A - \bar{V})] > 0 \Leftrightarrow \sigma_A \neq 0$$

La ecuación 1, propuesta originalmente por Tim Campbelle & William Kracaw en "Information Production, Market Signalling, and the Theory of Financial Intermediation." (1980, pp. 869), sugiere la existencia de un incentivo a la creación de información, la cual le permita identificar la diferencia entre $V_A - \bar{V}$. De tal modo que el inversionista "i" posee suficiente incentivo para exponer un monto (W_i) ante una expectativa en la modificación en el valor del activo "A" (V_A) durante un determinado tiempo. Dicho incentivo se encuentra presente en el mercado, solamente cuando existe variaciones en el precio del activo "A" (σ_A), siendo estas variaciones formadas por los sesgos en la información del mercado, que ocasionan sub-valoración o sobrevaloración de los activos en el mercado.

Bajo la teoría de Stiglitz y Grossman (1976), existen en el mercado un segmento de agentes desinformados que solo poseen la información del precio; bajo el supuesto que existe un segmento mayor de agentes informados, la información es transferida a los agentes desinformados completamente por el precio (eficiencia del mercado), bajo esta asunción se puede percibir que existe, aparentemente, el problema del "Free riding" (o del polizón) en donde el costo de conseguir información por los agentes desinformados depende únicamente de observar y tomar los precios fijados por la información plasmada en el mercado.

De esta forma el equilibrio que los agentes informados forman: $p^* = Cmg + \Delta C_K$, ante el incremento del costo marginal de la información y los rendimientos decrecientes de esta; el agente informado no posee los incentivos suficientes para plasmar completamente la información del mercado en el precio. De tal modo, si los costos de adquirir dicha información son cada vez mayores, la brecha entre el valor esperado y el valor aceptado, será también mayor.

Generalizando la teoría de Campbell y Kracaw (1980) sobre los incentivos del intermediario de participar e implementar su conocimiento dentro del mercado financiero. Asumiendo que el banco lleva a cabo procesos, con los cuales valoriza todos los activos que posee y administra; de tal forma, siempre busca maximizar la ganancia de su operatividad dentro del amplio espectro de actividades que tiene la capacidad de realizar; las cuales, pueden ser categorizadas dentro de cinco polos principales: (Bessis, 2002)

- **Banca Comercial:** Ofrece créditos a través de depósitos del público.
- **Banca de Inversión:** Servicios de asesoramiento, adquisiciones apalancadas, financiamiento de activos (fideicomisos).
- **Comercio (Trading):** mercados de deuda, derivados, instrumentos de renta fija.
- **Banca Privada:** gestión de activos.
- **Otros:** Custodia de valores.

Suponiendo que no existe la intermediación, es evidente la existencia de costos fijos por la administración de los activos de cada inversionista, asumiendo que participan de manera homogénea en el mercado financiero: $(\sum i = ni)$, cada uno de los agentes deberá de tomar los costos asociados de adquirir información, decodificarla y ejercer individualmente su interacción con el mercado (C_k^P) . Formando expectativas que difieren a causa de la variación de la información que cada agente puede obtener por sus propios medios; y así mismo, por discrepancias entre las capacidades analíticas y cognitivas de cada individuo:

(Ecuación 2) Sin efecto intermediación

$$\text{definiendo: } \frac{(V_A - \bar{V})}{\bar{V}} = [Y_A \neq 0] \Leftrightarrow \exists C_k^P$$

$Y_A := \text{Valorización potencial (\%)} \text{ del activo } A \quad \{\text{Expectativa - riesgo}\}$

$$\{\text{Para la economía}\} \sum_{i=1}^n [W_i] Y_A > \sum_{i=1}^n [C_k^P]$$

$$\{\text{Para el individuo}\} \left[\frac{C_k^P}{W_i} \right] < Y_A$$

Con la participación del intermediario, se asume que el costo de la información es repartido por todos los participantes que se encuentran asociados al intermediario " $C_k^I = C_k^P \left(\frac{1}{n}\right)$ ", debido a que el intermediario tomara todos los costos para producir dicha información, de tal modo que en la economía se tomara solamente una vez el costo, al emplear el intermediario $\left(\frac{1}{n}\right)$; Evidentemente, los montos invertidos por el intermediario son mayores, a causa de la asociación de los distintos montos individuales. ($W_i > [W_i^I = nW_i]$), de tal modo, la relación entre el costo y el monto invertido $\frac{C_k^I}{nW_i}$ será cada vez menor, por lo que el beneficio esperado " Y_A " podrá ser menor para compensar dichos costos y aceptar la inversión de los recursos. Adicionalmente el intermediario requiere de pagar una gratificación por la captación de sus recursos:

(Ecuación 3) con efecto intermediación

sea, $R_i := \text{Retribución al cliente}$;

$$\{\text{Para la economía}\} \sum_{i=1}^n [W_i] Y_A > C_k^P \left(\frac{1}{n}\right) + \sum_{i=1}^n R_i$$

$$\left[\frac{nW_i}{nR_i + C_k^I} \right] > Y_A^{-1}$$

$$\{\text{Para el Intermediario}\} \left[\frac{R_i}{W_i} \right] + \left[\frac{C_k^I}{nW_i} \right] < Y_A$$

En ningún punto, cliente no conoce el beneficio obtenido por el intermediario: Y_A ó $(R_i - Y_A)$, y no existe incentivo del intermediario por revelar dicha diferencia, por lo que no existe la transferencia plena de la información entre el agente informado (intermediario) al agente desinformado (cliente). Sin embargo, puede existir incentivo de elevar la retribución ante la situación de competencia dentro de la economía, en donde el intermediario deberá de elevar la tasa de interés pasiva para poder elevar su captación de recursos del público en general.

Al comparar las condiciones individuales de la ecuación 2 y 3:

(Ecuación 4) Condición de la intermediación

$$\left[\frac{R_i}{W_i} \right] + \left[\frac{C_k^I}{W_i^I} \right] \leq \left[\frac{C_k^P}{W_i} \right] \Leftrightarrow [R_i + C_k^I] < C_k^P < [Y_A * W_i]$$

$$W_i^I (Y_A - (R_i + C_k^I)) := \text{Ganancia por intermediacion}$$

Condiciones obtenidas:

- $[(C_k^P - Y_A) \leq R_i]$ Si el costo de la información de cada individuo es ínfimo, no se requiere de emplear los servicios del intermediario, en estos casos, las transacciones son realizadas por los mismos individuos.
- $[(R_i + C_k^I) > Y_A]$ Los costos de intermediación son superiores a la posible recompensa por el manejo del activo "A". Por lo tanto, no hay incentivo del intermediario de participar en el riesgo cuando el costo de la información es elevado.
- $\left[\frac{C_k^I}{W_i^I} < Y_A \right]$ Los montos de inversión y costos de la información son un factor limitante a la colocación de los activos. Mayores montos de inversión "diluyen" los costos y hacen más atractivas las recompensas "YA", de esta forma, existe mayor flexibilidad en la colocación de los activos. (Campbell, Tim S. & William A. Kracaw, 1980)

- [$(R_i + C_K^I + [Y_A * W_i^I]) = C_k^P$] Existe una recuperación de la eficiencia económica debido a que los repetitivos costos de la información de cada individuo no son efectuados; parte de estos costos que debieron ser tomados de manera privada, son convertidos en una retribución al emplear el servicio del intermediario. Cuanto menor sea el costo del intermediario, mayor será la recuperación de la eficiencia económica. Aun cuando la asimetría de la información exista. (Levine & King, 1993).
- [$(C_K^I = C_k^P) \approx 0$] Si la información individual fuese perfecta, existiría redundancia en el empleo de intermediarios.

Ante las distintas características del mercado real que son capturados por los instrumentos financieros; Los costos de información, o la información requerida asociada a cada polo y línea de negocio del banco es diferente.

En la bibliografía revisada sobre teoría de la intermediación, se menciona ocasionalmente: la información creada posee economías de escala, ante el efecto de ser repartido entre "n" número de clientes; pero no se menciona explícitamente que la información no posee economías de alcance, solamente en las corrientes de la teoría evolutiva (Nelson & Winter, 1982) y de la teoría basada en los activos (Grant, 2001), en las cuales se hace mención de la especificidad de la información y del conocimiento; de esta forma, la información obtenida solamente es útil en determinadas condiciones y en determinados lapsos de tiempo.

Por ejemplo: el adquirir información sobre los mercados de algodón, no aporta ninguna señal sobre las actividades de la empresa dedicada a la producción de pan, como la empresa Bimbo; pero dicha información puede aportar alguna señal sobre empresas dentro del ramo textil al pertenecer a la misma cadena productiva. Así mismo, se requerirá de conocer a los agentes individualmente, ya que su participación dependerá de la organización individual que las empresas

posean; dicha estructura no será siempre la misma y cambiará ocasionalmente ante los cambios de gerencia. Por lo que la información se deprecia rápidamente ante los cambios del ambiente y así mismo, la información es específica a determinadas cadenas productivas.

El banco al participar en un sinnúmero de mercados distintos, subvierte en algunos aspectos el beneficio creado en la ecuación 4. Debido a que acumula distintas tareas que requieren la extracción de la información específica y de la actualización constante de esta, por lo que deberá de monitorear continuamente los distintos sectores y ramos de la economía; incrementando los costos de la información.

Al integrar dentro de la organización distintas actividades, la organización interna del intermediario se vuelve tan compleja como la economía misma, de tal forma que al participar en un mayor número de mercados los costos de transacción internos se vuelven mayores y así mismo el riesgo organizacional propio del intermediario (Holland, 2010); pero al mismo tiempo la participación en las distintas cadenas productivas, le da al banco la habilidad de transferir tanto los riesgos como los costos de la información de los distintos mercados, permitiéndole diversificar las inversiones y promediar los riesgos de los distintos sectores de la economía. (Duffie, 2008)

Esta peculiaridad fomenta la necesidad de convertir al intermediario en un ente "místico", el cual requiere de conocer profundamente las características del ambiente en el cual se desarrollan los recursos que son asignados bajo su jurisdicción, requiriendo crear la suficiente información con un ínfimo costo para el máximo aprovechamiento de dichos recursos; Beneficiándose oportunamente de las valorizaciones de los activos y minimizando las posibles pérdidas económicas implicadas por la compleja organización, es decir: maximizar la relación "Ganancia/Riesgo" mediante la práctica de la Administración de Riesgos: actividad

la cual implica tomar decisiones asertivas o esquemas de acciones con el propósito de identificar riesgos; transferir, evitar o moderar los riesgos absorbidos y seleccionar los métodos requeridos para el monitoreo de los riesgos (King, 2014).

1.4 La naturaleza de los Riesgos y su Administración

Los intermediarios, en el polo de la banca comercial desempeñan operaciones de fondeo y operaciones relacionados a la asignación de créditos, denominadas como operaciones pasivas y activas respectivamente. Estas poseen diversas características tanto en los plazos de recuperación, como en los flujos y montos recuperados; en dichas actividades se asumen diversos riesgos que se encuentran entre sí: (Heffernan, 2005)

- **Riesgo de Crédito:** Probabilidad de pérdida de valor del monto otorgado como crédito; el cual puede perderse en su totalidad o recuperarse parcialmente.
- **Riesgo de Liquidez:** incapacidad del banco de dotar liquidez, a causa de fondos líquidos insuficientes, debido al fondeo, recursos comprometidos en otras actividades.
- **Riesgo de liquidación** (Settlement Risk): también conocido como "Herstatt risk"; el cual se refiere a la probabilidad de incumplimiento o demora de los pagos
- **Riesgos de Mercado:** Probabilidad de cambio en el valor de los activos a causa de la interacción del mercado, Ej. Tasa de interés, inflación, tipo de cambio, cambios en el precio del mercado.
- **Riesgos de apalancamiento:** Probabilidad en la cual las reservas de capital o activos líquidos, no sean suficientes para cubrir el financiamiento obtenido.
- **Riesgos operacionales** (internos)
- **Riesgos Soberanos, políticos y legales** (del ambiente)

Los riesgos poseen diversas características, tal como la diferente probabilidad de su materialización, las cuales se encuentran amoldadas por la naturaleza de los factores que las originan (conocidos habitualmente como factor de riesgo "Risk Hazard"); A causa de los diversos orígenes, crean impactos desiguales en frecuencias discontinuas, aparentemente aleatorias; aun cuando se crea improbable las repercusiones de un evento funesto, sean pérdidas mínimas o totales, estas tienen siempre la posibilidad de ocurrir pese a la pequeña probabilidad existente, tal como lo expresó Edward A. Murphy Jr.; por lo que siempre es requerido tener planes de contingencia ante las consecuencias de los riesgos, aun cuando no se conozca la posible fuente del riesgo.

En la banca comercial, bajo el sentido de intermediación, funciona como el puente que permite la conciliación entre los superávits y los déficits de sus clientes; los cuales poseen la peculiaridad de ser tanto clientes corporativos como personas físicas o firmas pequeñas. Los cuales difieren en su comportamiento individual y en los planes elaborados para el empleo de los recursos adquiridos por medio del crédito. De forma analógica a la observación de los asteroides y el riesgo que poseen de impactar en el planeta tierra: las entidades de mayor tamaño (como clientes institucionales) son ampliamente conocidos, evaluados y usualmente tienden a poseer un menor riesgo; en contradicción con las entidades más pequeñas (como personas físicas y PyMEs), las cuales no son conocidos en su totalidad, son difícilmente monitoreados y poseen una mayor probabilidad de impacto.

Para ejercer las funciones de monitoreo e intermediación, los bancos deben de adquirir la suficiente información, necesitando de la suficiente capacidad tecnológica para recolectarla, analizarla e implementarla para la toma diaria de decisiones y para la dirección en el largo plazo de la firma. Permitiéndole al banco actuar ante cualquier anomalía que pueda provenir de un amplio grupo de

individuos que poseen conductas específicas, que se rigen bajo su propia estructura (productiva o familiar); ante la variabilidad de dichas facultades, los agentes se comportan de distinto modo aun cuando las condiciones del ambiente son similares, de tal modo que asumen distintos niveles de riesgos involucrados en dichas decisiones (riesgo idiosincrático); los cuales pueden comportarse de diversas formas para la resolución de un mismo objetivo. (Rae, 1834)

De acuerdo con Heffernan, S. (2005), los bancos poseen al menos 5 mecanismos distintos con las cuales le permita minimizar los riesgos asociados al Riesgo de crédito:

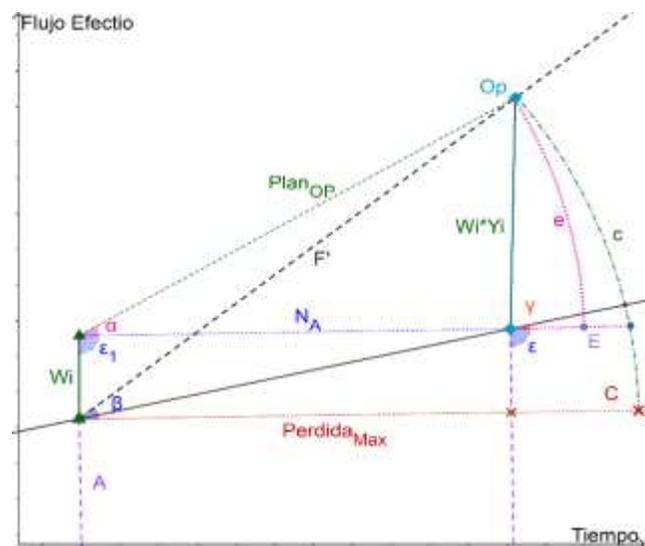


Ilustración 2.A Gráfica que representa la interacción entre los riesgos; Elaboración propia.

"default", ante la probabilidad de incumplimiento de pago el precio se es determinado para mantener un determinado valor esperado de la inversión, de tal modo: a mayor riesgo, mayor será su precio. (Kao, 2000)

Fijación del

Precio: Bajo la perspectiva del "pricing" por medio del modelo estructural, son evaluadas determinadas características, tales como: el valor de la compañía, la capacidad de generar ganancias, los flujos de capital y la estructura del capital de esta; dichas características permiten calcular la probabilidad o el riesgo de

En el caso de los créditos, el precio se entiende como la tasa de interés que se cobrará. La singularidad del modelo estructural radica que, en el caso de los agentes con una mayor necesidad del financiamiento, son aquellos cuya estructura

organizacional posee debilidades en su sustentación financiera, por lo que se les ofrece una tasa de interés mayor ante el mayor riesgo de crédito que poseen.

Bajo el sentido de St. Thomas Aquinas (1225) el "precio justo" (α) puede ser difícilmente localizado y/o expresado matemáticamente; El precio justo puede ser acordado en una transacción donde los agentes posean las mismas condiciones, es decir: que ambos agentes deberán de poseer el conocimiento suficiente entre ellos mismos, poseyendo el conocimiento sobre los costos de producción, el posible beneficio obtenido (β) de emplear el bien y de vender el bien; y así mismo de las condiciones del ambiente (mercado). Al tener todos estos factores en cuenta, no existe ningún tipo de oportunismo o riesgo moral por parte de ninguna de las partes en un entorno de información perfecta, y en un mercado de bienes homogéneos. De esta forma se deduce que no existe más precio justo que el que es acordado socialmente (precio de mercado).

Los créditos poseen una naturaleza no-homogénea, esto, debido a que los recursos colocados por el intermediario, son empleados por el agente receptor para el financiamiento de diversos planes quienes poseen características heterogéneas; Para que pueda ser aplicada la definición del precio justo, el intermediario deberá de conocer a la persona justa (indicada) la cual estará dispuesta pagar un precio (precio de reserva) ya que encajaría ante las expectativas que el agente posee de sí mismo, en determinado tiempo " T_0 " (Ilustración 1)⁵, de tal modo, que conocerá

⁵ Para los clientes empresariales, su capacidad productiva y de liquides se mide bajo el flujo de efectivo que posee; por otro lado, las personas físicas, su expectativa estaría dada por la capacidad de acumulación de activos: $F = W - (C_f + \Delta dep + i)$; la capacidad de acumulación de activos depende de su flujo actual de efectivo, mediante fuentes de ingresos del individuo menos el gasto en bienes perecederos (Ej. Alimentos y servicios), la depreciación de sus activos fijos (bienes duraderos, títulos financieros) y el interés de los préstamos anteriores.

el valor esperado de emplear los recursos en sus planes de producción o consumo en el tiempo "T_i".

Tasa de Interés := α

La tasa de interés (α), ante la asunción del precio justo, deberá de alinearse perfectamente los posibles beneficios del financiamiento del plan del agente T_i (F'),

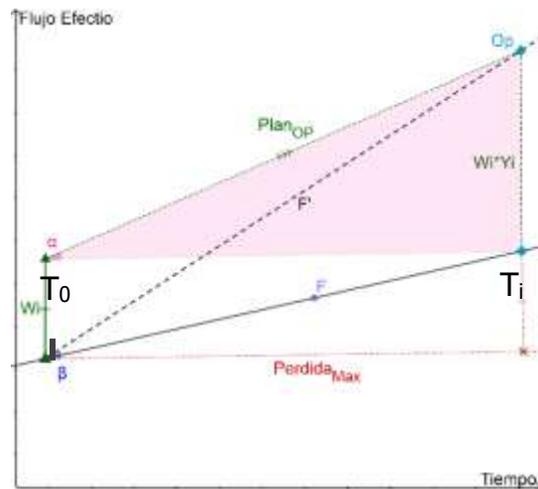


Ilustración 2.B Gráfica que muestra el beneficio esperado del crédito y por lo tanto su tasa de interés α ; Elaboración propia.

de este modo: $Intereses = F'_{t_i} - F_{t_i}$; Lo que reduce teóricamente a "0" la probabilidad de incumplimiento, si el agente a quien se le otorga el crédito es lo suficientemente perspicaz para seguir la línea de su plan óptimo (Plan OP).

Riesgo Idiosincrático := β

$$= R^M + R^L$$

$$\beta = \gamma + \alpha$$

Cuando el intermediario asigna los recursos a un determinado plan de producción, los recursos se encuentran fuera del control del intermediario, de tal modo, los recursos son expuestos a la conducta del agente quien adquiere el financiamiento; surgiendo el riesgo idiosincrático.

Al ser el cliente quien administra los recursos asignados, existen un sin-número de posibles trayectorias o decisiones que puedan ser tomadas para su aprovechamiento de los recursos; Dichas posibles rutas se encuentra ilustradas como la curva "C", en donde las posibles trayectorias pueden generar atrasos en el pago del crédito al no aprovechar o perder el valor de los recursos, a tal riesgo de incumplimiento bajo la administración del agente se define como el "riesgo de

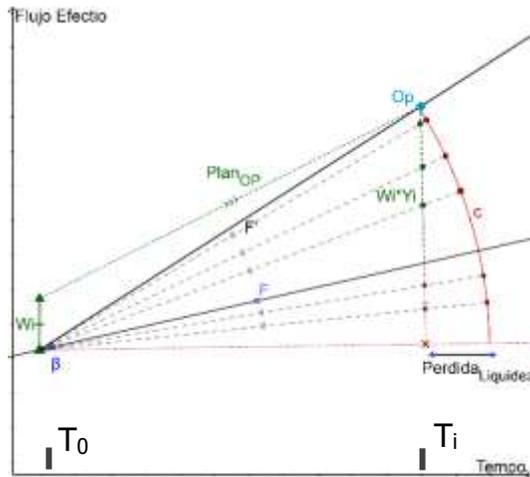


Ilustración 2.C Gráfica que muestra los posibles caminos que el cliente puede tomar; Elaboración propia.

liquidación" (R^L); las posibilidades en la trayectoria, también está compuesto, por los riesgos existentes en mercado (R^M) de los cuales ni el agente ni el intermediario tienen control.

Estos factores ilustrados con el ángulo β , hace referencia a los posibles escenarios finales en los que pueda finalizar el plan del cliente, en la búsqueda de alcanzar el plan óptimo (Op).

Tal como indicaba Stiglitz (1981), la tasa de interés puede afectar en el comportamiento del individuo ante el aprovechamiento de los recursos que se le son proporcionados. Por lo que ante a mayores tasas de interés, o en este caso un sobreprecio en base al precio justo, el agente debería de alcanzar el aprovechamiento que no puede llegar con su plan óptimo. El sobreprecio del crédito dará como consecuencia demoras en el pago acordado del crédito o absorciones del colateral el cual es expuesto por el consumidor (A) para acceder al crédito, quien deberá de sacrificarlo para liquidar el pago en " T_{i+1} ". De tal modo, que la asignación de un precio elevado se convertiría el banco en una entidad

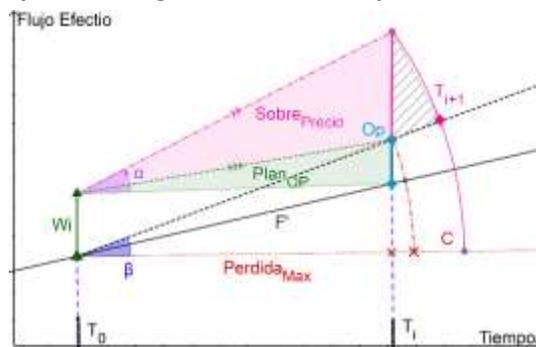


Ilustración 2.D Gráfica que muestra un sobreprecio del crédito, inalcanzable con el plan OP; Elaboración propia.

parasitaria, lo cual limita la operatividad del agente que solicita el crédito, ya que pierde valor económico debido al costo del crédito y el aumento de la dificultad para gestionar los recursos de manera efectiva. Ante tales consecuencias, crea al mismo tiempo un impacto financiero en el agente y en la economía en general: al afectar a los

agentes relacionados con el cliente dentro de la cadena productiva, creando un efecto sistémico que ocasionaría un aumento en el riesgo de liquidación de otros clientes por las alteraciones en el sistema y por lo tanto en el riesgo del mercado. (Baldan, 2012)

Límites de Crédito y Uso de Garantías, Son dos mecanismos por el cual, el banco limita la colocación de recursos; Siendo la fijación de las garantías, un posible costo hundido para el cliente, quien desee tomar un comportamiento oportunista después de adquirir el financiamiento; Al incrementar el costo de participar en el mercado financiero, la calidad del cliente aumenta, al ser los más fiables quienes aceptaran apostar la garantía más elevada; mientras que aquellos oportunistas tendrán menos incentivo a participar ya que el costo hundido reducirá el beneficio de su comportamiento oportunista.

Diversificación, mediante la administración del portafolio de créditos, el banco participa en diferentes cadenas de producción, con tal de reducir las pérdidas ocasionadas por algún sector específico, para ello los factores de riesgo de las distintas cadenas productivas no deben de encontrarse correlacionados; Con la diversificación, el banco es capaz de aceptar mayores riesgos al promediar tales riesgo con aquellos clientes que poseen un nivel menor.

Dentro de la misma ilustración 2, se hace referencia al riesgo sistémico mediante el ángulo " ε " en el cual se desea expresar que siempre existe la posibilidad de que algo pueda salir mal, pese a la correcta medida y protección ante el riesgo idiosincrático. De este modo siempre existirá, aun cuando sea pequeña, la probabilidad de la ocurrencia de algún siniestro (conocido como Cisne Negro), por lo que el banco y cualquier empresa siempre tendrá cierto nivel de riesgo epistemológico; Ante la existencia de dichos riesgos que no pueden ser evitados o diluidos, estos pueden ser compartidos o trasladados mediante la creación de productos financieros, tales como los derivados del crédito. Con el uso

de dichos instrumentos financieros los bancos poseen la capacidad de transferir los riesgos que se asumen al impregnarlos en algún producto financiero específico.

El empleo de instrumentos derivados del crédito, le permite al intermediario estabilizar la exposición en el portafolio de inversión que posee, al diluir y diversificar el riesgo de crédito con los inversionistas. Sin embargo, el intermediario ante la opción de poder delegar los riesgos que implican las operaciones financieras, crea al mismo tiempo, ineficiencia en la colocación de los recursos; debido a que es reducida la necesidad del monitoreo, ante la disminución en las expectativas de pérdidas económicas a causa de la selección adversa (Duffie, 2008) y ante las disminuciones en las responsabilidades que el intermediario deberá de asumir ante la materialización del riesgo. Ante la alteración en las actividades del intermediario los recursos pueden ser colocados en inversiones de bajo rendimiento (créditos de baja calidad) sin la precaución del riesgo ni del rendimiento del plan de inversión, reduciendo la calidad de los proyectos y el beneficio económico que trae consigo la creación del proyecto de inversión (Morrison, 2001) (Stiglitz & Weiss, 1981).

Ante la inexperiencia y la poca información que el cliente posee, ante la complejidad del instrumento y ante el beneficio de liquidar el capital asignado a proyectos de alto riesgo; los bancos poseen los suficientes incentivos para transferir la mayor cantidad posible de los riesgos que han adquirido; Ej. En el año 2008, los bancos delegaron los riesgos que provenían del incumplimiento del pago de las hipotecas, al emplear un instrumento financiero denominado como "CDOs" (Collateralized Debt Obligation), los cuales brindaron liquidez a los créditos perdidos; Financiado las deudas del banco con recursos de terceros y esparciendo los riesgos del crédito al público en general. La transferencia de la información de dicho instrumento y de la situación del riesgo, fue mínima; por lo que los agentes poco precavidos, desconocían el riesgo de las operaciones financieras que el título poseía, lo cual generaría grandes pérdidas a los distintos agentes del sistema

económico que participaban en la inversión, lo cual formaría la crisis de 2008 por sus efectos sistémicos.

El poseer poca atención, poca habilidad cognitiva o poca capacidad de adquirir información sobre las posibles repercusiones o eventos funestos; puede en algún momento determinado, repercutir tanto en la principal entidad involucrada en la toma del riesgo como en el grupo primario que lo rodea, (o grafos que se relacionan con la entidad) (Newman, Barabasi, & Watts, 2006)⁶ los cuales son agentes que interactúan frecuentemente con la entidad y sus acciones están relacionados directamente a él (como acreedores, deudores, socios o trabajadores en el caso de las personas morales). Dentro del sistema, cada nodo posee sus propias características individuales, vínculos y grupos primarios; El nodo portador del riesgo, puede afectar al sistema dependiendo de la magnitud en la relación o dependencia que posee con los demás nodos. Las interacciones y los riesgos pueden repercutir de manera simultánea y directa a los nodos más cercanos (causalidad Newtoniana) afectando el comportamiento de los demás agentes interconectados de manera indirecta, siendo estos los nodos vinculados al grupo primario de la entidad observada, por las interacciones que el grupo primario posee con otros nodos fuera del grupo; De este modo, el comportamiento adverso es “contagando” en todo el sistema, afectando eventualmente a todo sistema económico por medio del encadenamiento de los

Ejemplo de Sistema

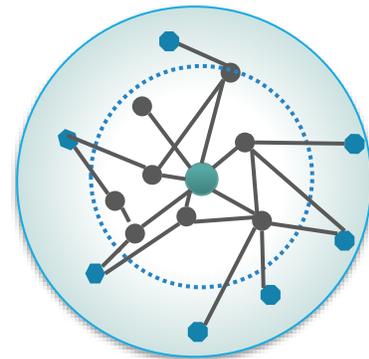


Ilustración 1 Ejemplo de Sistema sencillo, Elaboración Propia. El clúster, dentro del círculo punteado es el grupo primario de la entidad observada.

⁶ Dentro de la rama de la computación, los grafos, son un conjunto de nodos que se entrelazan de manera binaria (conectado o desconectado), es empleado en la teoría de grafos para ilustrar los actores participantes dentro de una red social (por ejemplo: redes sociales, flujos comerciales, flujos de información, organizaciones complejas, hechos históricos); “Structure & Dynamic Networks” (2006) Mark Newman.

vínculos del sistema; tal efecto es conocido habitualmente en las ramas de la economía como "externalidades".

Tras la probabilidad de contagio (definido como "riesgo sistémico"), ha sido este un tema de gran interés, debido a la ocurrencia de eventos que han requerido cambios dentro de forma de pensar y actuar ante la presencia constante de riesgos individuales, que se han transportado al sistema económico a causa del comportamiento oportunista de los agentes, ante la posibilidad de maximizar sus propias ganancias privadas en el muy corto plazo, sin importar las consecuencias sociales del largo plazo; tal como la formación de burbujas especulativas (Allen & Gale, 2002), o el contagio ocasionado por la transferencia del riesgo (Baur & Jooseens, 2005)

Tanto en el caso de la crisis de Bankhaus Herstatt en Alemania 1974 y de Franklin National Bank (Basel Committee, 2004). En ambos casos, las acciones individuales y oportunistas de los intermediarios, ocasionaron repercusiones que transformarían la estructura institucional y los mecanismos de la supervisión que se realizaban en aquella época. Por lo cual fue creado el comité de Basilea, para recuperar la confianza y la estabilidad financiera de los agentes afectados, incorporando normas y medidas de regulación; con tal de evitar repetir errores similares. Las regulaciones y mecanismos deben de ser lo suficientemente eficaces para suministrar la suficiente información sobre los participantes, reduciendo los costos de información que el cliente asume ante la búsqueda de un intermediario y en la selección de sus productos financieros; con la finalidad de que pueda valorarizar sus recursos financieros. Así mismo, la participación de las entidades reguladoras reduce los costos de transacción generados, al mantener el funcionamiento del sistema de intercambio económico, homologando los procedimientos de las instituciones financieras, lo cual disminuye la incertidumbre que los clientes asumen al emplear los servicios financieros.

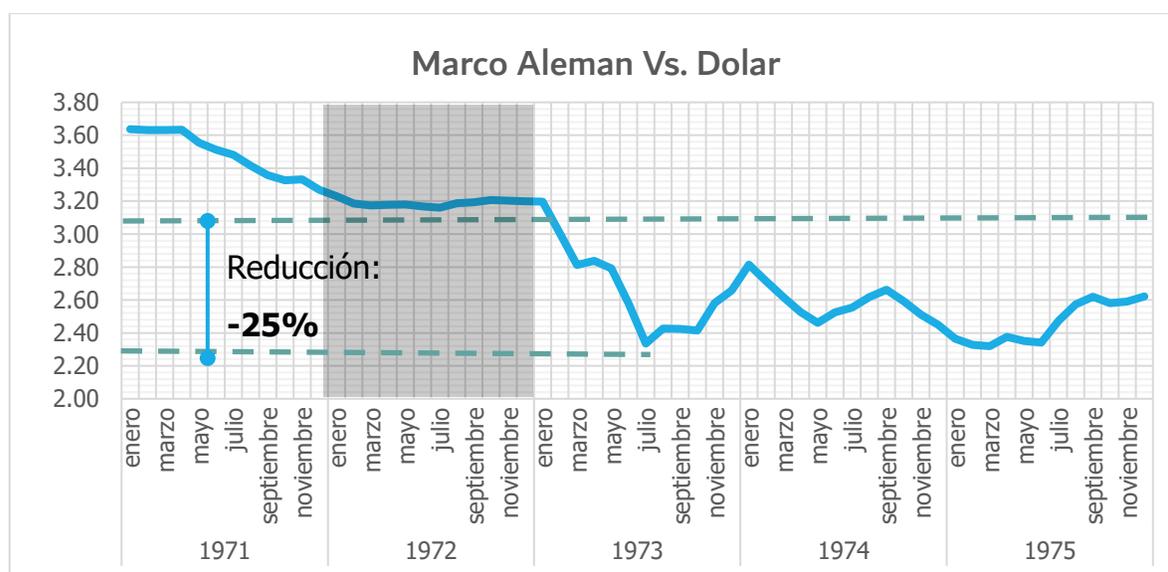
En el caso concreto de Bankhaus Herstatt, tanto la mala praxis de las inversiones ejecutadas (riesgos operacionales) como los riesgos de mercado, afectaron las operaciones internas; Originando pérdidas en los depósitos locales del banco Herstatt, desencadenando el efecto del contagio en el sistema financiero dentro y fuera de Alemania por el incumplimiento de sus obligaciones. Herstatt fue un banco pequeño (poseído en su mayoría por Hans Gerling) siendo el banco posicionado en número 35 en Alemania por su tamaño, pero famosa por poseer altos volúmenes en la inversión en divisas. El departamento encargado de invertir en divisas (especialmente dólares), liderado por Dany Dattel; incurrió en un gran aumento de las inversiones en divisas durante 1972, ante la abolición de la paridad oro-dólar establecida por Bretton Woods.

Al finalizar la paridad fija, inicio la volatilidad del dólar al mismo tiempo que inicio el ánimo por lucrar con dichas disparidades en el precio, así, el banco Herstatt había colocado 53% de sus activos en la inversión de divisas (y también el 28% de sus pasivos) (Mourlon-Druol, 2015); las inversiones con alto nivel de riesgo se manejaban sin conocimiento del público, se afirma que ni siquiera el dueño Iwan Herstatt poseía conocimiento sobre tal decisión.

El riesgo aumentaba progresivamente, y fueron expuestos 137 millones de marcos en las inversiones. Las autoridades no pudieron vaticinar las consecuencias de la excesiva exposición del banco, debido a la escasa información que el banco ofrecía sobre las inversiones que se llevaban a cabo. No existieron acciones o medidas de contingencia tras aquel fatídico desenlace. En el momento en que las inversiones habían perdido su valor; El banco perdió la liquidez, así como la confianza y los ahorros del público en general (además de ser clausurado posteriormente); las deudas que debieron de ser cubiertas por el banco, fueron parcialmente rescatadas con los recursos de las autoridades, la demora en el pago de las deudas del banco Herstatt, crearon al mismo tiempo pérdida de liquidez en

otras entidades relacionadas, como los bancos: William & Glyns, Chase Manhattan, Union Bank de Suecia, First National Bank of Milwaukee, entre otros.

Gráfica 1. Paridad Marco Alemán y Dólar durante el suceso de Herssat



Elaboración propia, Fuente: Federal Reserve Bank of St. Louis, Germany / U.S. Foreign Exchange Rate.

Los eventos descritos originaron cambios en la organización bancaria y la información que debe de ser reportada a las autoridades. Requiriendo de un área especializada, la cual le permitiera producir información privada, para su uso interno. Creando información crucial que le permitiese medir, monitorear y modelar los riesgos existentes en los distintos niveles de su propia operación.

El analizar los riesgos se vuelve una actividad compleja, a causa de la naturaleza compleja de los riesgos, de la información disponible y la imposibilidad de observar el futuro con el 100% de asertividad. Ante la diversa información que debe de ser procesada e interpretada, en un complejo contexto en donde cada individuo interactúa en un tablero dinámico con una infinidad de jugadores e intenciones desconocidas; Donde cada nodo dentro del sistema puede ser tanto un portador del riesgo como el objetivo de este. Los riesgos se encuentran esparcidos

en apariencia amorfa y oculta dentro de los diferentes dominios de las actividades empresariales: Producto, Proyecto o Proceso; y así mismo poseen diversos orígenes: Internos o externos. (Pandian, 2007)

Para poder prevenir los riesgos, se requiere tener previo conocimiento de estos, creando sistemas de alerta temprana ("Early Warning System"), permitiendo detectar de forma anticipada el advenimiento del riesgo, creando información de las diversas posibilidades del origen del riesgo; Analizando individualmente cada firma, cliente, proceso interno, ambiente macroeconómico, político... Necesitando de herramientas lógicas (como un diagrama de flujo) y de procesos de identificación de riesgo; Analizando la frecuencia de eventos pasados, observando la distribución de su comportamiento, analizando sus atributos y evaluando sus posibles impactos. Procurando siempre minimizar el daño posible: eliminando, aislando o evitando las causas que provocan efectos adversos, siendo estas causas el efecto de otra causa dentro del sistema (Thommas Aquinas, "Summa Theologica", 1225)⁷.

El riesgo es la probabilidad de la materialización de un costo, el cual está sujeto a la frecuencia con la que se ejerce la actividad a la cual es inherente. El temor, o la sobrevaloración del riesgo, influye negativamente en la ejecución de los planes, proyectos o transacciones que los agentes económicos realizan día a día. Por el otro lado, la osadía, la creencia en la cual el riesgo no existe o la ignorancia sobre los diversos riesgos, conlleva repetidamente a la materialización de las pérdidas no consideradas por el agente; ya que el agente no posee ninguna medida para su preservación ante los cambios en el ambiente, haciendo imposible

⁷ Axioma 1: Siempre existe una causa: $\exists x \exists y C(x, y)$, el evento "x" causa a "y"

Axioma 2: Nada es causa de sí mismo: $\forall x \neg C(x, x)$, el evento "x" ocasiona un efecto asimétrico a "x"

Axioma 3: Si no existe una primera causa, no habrá una causa final: $\forall y [\neg \exists z FC(z, y) \rightarrow \neg \exists x C(x, y)]$
Se requiere de la existencia de la primera causa "z", para que exista una causa "x" que cause "y"

la supervivencia de la entidad que no puede evaluar o contemplar el impacto de los riesgos tomados.

Para poder minimizar el riesgo, se requiere de un proceso de aprendizaje que permita corregir la conducta del individuo, con la finalidad de obtener el mayor beneficio en un ambiente adverso, En uno de los experimentos realizados por Antonio Damásio & Antoine Bechara (2005), los pacientes con lesiones cerebrales en la corteza ventromedial tienden a perder mayores cantidades de dinero en un juego de apuestas conocido como "Iowa gambling task", debido a que su comportamiento no es actualizado, aun cuando se posee el conocimiento del riesgo (es percibido por el paciente, pero no es tomado en cuenta para la decisión, debido a la lesión que no permite evaluar o apreciar los efectos adversos de tomar el riesgo).

La competitividad, los costos operativos y la velocidad de las decisiones; son algunos factores que determinan la eficiencia en la que un intermediario opera. Viene bien que los bancos posean los suficientes recursos para sus operaciones, sin embargo, los intermediarios deben de poseer la suficiente capacidad de administrar sus recursos óptimamente, descubriendo las mejores oportunidades de valorización de los activos y teniendo mecanismos que le permitan actuar con diligencia. De tal modo, se requiere de grandes fuentes de información, así como la capacidad cognitiva que le permita analizar la información, para obtener conocimiento de esta y emplearla a su favor.

Las nuevas fuentes de la información que las entidades (como los intermediarios financieros) podrían incluir, pueden provenir de diversos medios electrónicos tales como las redes sociales, así como de la interacción del usuario con chatbots, mensajes o incluso grabaciones de llamadas telefónicas; de este modo, se posee un mayor volumen de información sobre el individuo "A", la cual

le permita reconocer si este posee una conducta riesgosa o es una persona que atiende sus obligaciones, por ejemplo.

En el momento en que las actividades humanas evolucionan de "offline" a actividades primordialmente digitales; las acciones que realizan los individuos se encuentran registradas dentro de la red, generando grandes bases de datos que poseen un gran potencial para crear conocimiento con un alto grado de efectividad, que le permitan al agente forjar respuestas precisas a tiempo real.

La peculiaridad de esta información es su excesivo volumen, del cual se puede tener un pequeño grupo de datos que poseen gran información y así mismo un gran volumen de registros que pueden poseer poco significado. Por lo que se considera a esta información como "Bad Data", la cual requiere de intrincados procesos analíticos, coloquialmente referidos como "alquimia digital" (transformar "datos basura" en "oro"), requiriendo de un entorno informático con la suficiente inteligencia para procesar, mantener y transmitir la información que se necesita. El mal manejo de esta información puede ocasionar tanto la pérdida de esta (por filtros ineficientes) como el obtener respuestas poco asertivas y extremadamente lentas.

Capítulo II

2 La Información del Big Data

*“Nuestros sentidos nos permiten percibir sólo una pequeña porción del mundo exterior.”
- Nikola Tesla (1856-1943)*

La información, puede ser definida como un mecanismo de transmisión de señales, que poseen como finalidad la interacción entre dos o más agentes, el empleo de la información, se encuentra definido por la organización del sistema económico en el que participan los agentes quienes poseen objetivos individuales; los mensajes transmitidos por los agentes dentro del sistema económico, son elaborados bajo cierta codificación, la cual el mismo mensaje puede ser interpretado de distintas maneras. (Hayek, 1945)

La información contenida en los mensajes, se encuentra fragmentada (o sesgada), haciéndose poco clara ante el ruido que pueda existir en el canal de transmisión y así mismo por la capacidad del canal para transferir dicho mensaje información. (Stone, 2015) Los agentes poseen características heterogenias entre sí, y así mismo tienen objetivos que son incompatibles entre sí; debido a los distintos contextos sociales en los que se desenvuelven, las necesidades que poseen y la interpretación de la información (conocimiento). (Wells, 1981) La capacidad de los agentes para interpretaciones el mensaje y reacciones ante esta (Entropía) crea ajustes en su conducta, que no siempre son los deseados por el emisor del mensaje. Debido a esto, las capacidades del canal, el ruido que pueda tener este y los objetivos individuales son factores que limitan la capacidad de un sistema de introducir y esparcir información a los participantes, lo que genera incertidumbre en el sistema.

A causa de la existencia de la incertidumbre epistemológica, se producen eventualidades no esperadas, a causa de la existencia de factores no conocidos;

ocurridos ante la acción humana⁸. Tales eventos, brinda mayor información sobre los elementos del sistema, debido a que posibilita la reflexión de eventualidades no ocurridas u observadas previamente, incurriendo a la búsqueda de posibles acciones o procedimientos (rutinas) no considerados previamente. Este efecto de retroalimentación es definido por Hayek, como el *negative feedback*, por el cual, los agentes pueden generar ajustes en su comportamiento debido a las constantes modificaciones en las expectativas, ante la observación de sus propios fracasos son generados nuevas cepas de conocimiento en los centros de información (el mercado).

La necesidad de la información; su recolección y transmisión es crucial tanto para la organización de los sistemas económicos como para la elaboración de planes individuales de los agentes. Requiriendo de diversos canales para la transferencia de la información, dichos canales poseen sus propias características en cuanto a: costos, eficiencia y ruido. ante las diferentes características de la información, el agente recolecta conocimientos a través de los distintos medios: impresos, estadísticos, culturales (imitación social), verbales o electrónicos.

El desarrollo tecnológico en los medios de comunicación, produce cambios en la organización y en mecanismos por los cuales los agentes interactúan; Modificando la estructura de los acuerdos sociales y económicos, al implementar innovaciones tecnológicas, las cuales reducen las fricciones, los costos de transacción y de transporte; transforma la estructura interna del sistema y de los agentes: como las interacciones entre personas (P-2-P), la forma en como las

⁸“La teoría de la información, desarrollada por Claude Shannon nos dice que un evento inesperado nos informa sobre el mundo a nuestro alrededor. Podemos de la misma forma concluir que solamente no habría necesidad de ajuste si nuestro conocimiento fuese completo y cierto.” (Ribeiro, 2004)

personas interactúan para la adquisición de servicios (P-2-B) y en como las empresas comunican información entre si (B-2-B) o para sí mismas.

El modo en como la información es acumulada y difundida evoluciona con el paso del tiempo, al aumentar el número de participantes del sistema, el flujo de la información incrementa, generando grandes cúmulos de información que pueden dejarse de lado por la incapacidad de procesar dichos flujos; requiriendo de mejorar los medios de almacenamiento de la información y el desarrollo de canales eficientes para la transmisión de esta. Ante las fricciones del mercado y los sesgos de información, se genera sobre aglomeración en la información que debe de ser comunicada dentro del sistema para su funcionamiento; haciendo de una tarea habitual la elección de la información, el cómo almacenarla, analizarla, catalogarla y emplearla para la toma de decisiones individuales.

2.1 Definición del Big Data

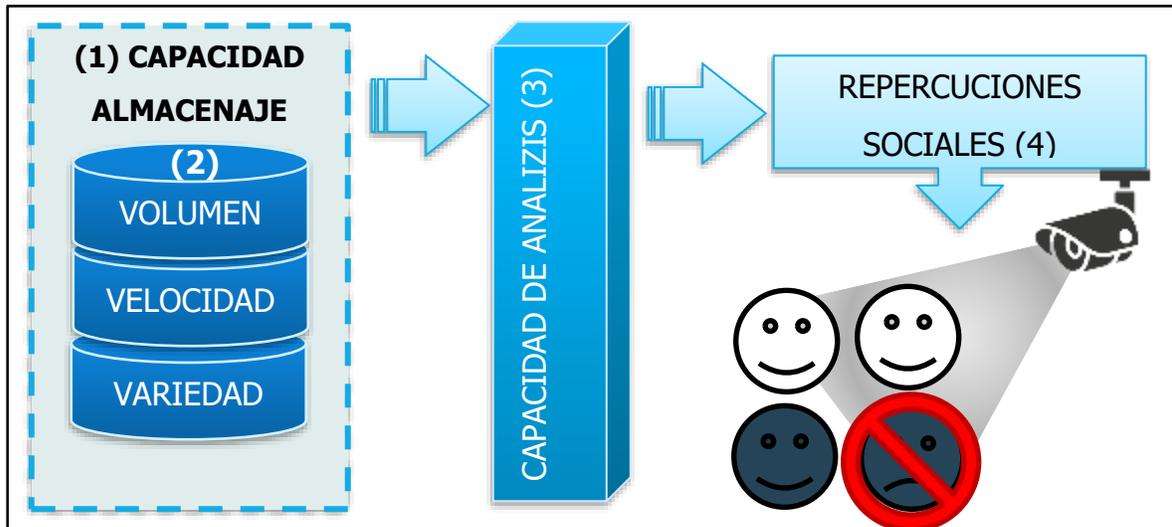
“Big Data” es un término ambiguo que es empleado en distintos ámbitos de la sociedad, incluyendo las distintas ramas científicas que se han desarrollado rápidamente y su interés por el tema (tales como en ramas de la biología, la genética, la astronomía, medicina, administración, topografía y entre otras), las cuales incluyen desarrollos tecnológicos bajo distintos enfoques o metas, y al mismo tiempo emplean el mismo termino para describir distintas atributos que son resaltados por el empleo que se le da en las distintas investigaciones de las distintas ramas, además de ser un término popular y ser empleado indiscriminadamente incluso en formas coloquiales con ninguna preocupación de su definición o de los procesos involucrados.

La definición del Big Data puede ser catalogada en cuatro grupos de definiciones (De Mauro, Greco, & Michele, 2015); de las cuales difieren en cierta medida por el valor literario que se le da al termino, pero todos se encuentran compuestos bajo el mismo tronco central: la información.

Los grupos comunes que emplean el termino son:

- 1) **Basados en las capacidades:** empleando el termino para hacer referencia a las limitaciones actuales que hay entre el flujo de información que es transmitida y las capacidades o requerimientos para el almacenamiento de dichos flujos.
- 2) **Basados en los atributos de la información:** por el cual se definió inicialmente como las tres "V"s (Volumen, Velocidad y Variedad), a la cual se le han incluido posteriormente nuevas "V"s como la del "Valor". Siendo esta una de las definiciones más aceptadas debido a las características descritas por el "volumen" que hace referencia al termino "big" de "Big Data". (Oracle Corporation, 2019)
- 3) **Basados en las necesidades tecnológicas:** a lo que se refiere al desarrollo de las capacidades tecnológicas que permiten manipular, generar y analizar información. Bajo este tópico, se ha entremezclado con definiciones como "Data mining" o "Machine Learning", tales definiciones hacen referencia a los métodos y mecanismos (algorithms) que se han desarrollado.
- 4) **Basado en el fenómeno social:** los cuales definen el termino, como un fenómeno social, basados en los impactos sociales del desarrollo de la inteligencia de las computadoras y los cambios sociales de la "no privacidad" y de la cultura informática.

Ilustración 3: Características descritas al emplear el término "Big Data", elaboración propia.



Tras la crisis de lo que se conoce como "punto COM", durante la década de los 2000, a lo que se podría marcar como el albor de la "era del Big data"; Siendo desarrollado un nuevo modelo, en el cual los agentes se les permite interactuar y crear propio contenido dentro de un entorno digital, evolucionando a lo que se le conoce como "web 2.0". En donde los usuarios participantes se les da la virtud de "inteligencia colectiva" convirtiendo a la red, en un cerebro global. (O'Reilly, 2005)

Esta virtud, es nada menos, que la capacidad de los agentes de registrar por sí mismos, sus conocimientos que poseen y las acciones que habitualmente realizan. Tal fenómeno social (Definición 4), da como resultado, la capacidad de los individuos de crear su propio entorno que puede ser observado públicamente. Modificando el esquema en como los agentes interactúan y en como la información cotidiana es almacenada y obtenida.

Algunos autores, dentro de las distintas corrientes sobre el origen y definición del Big data, describen que el comienzo del fenómeno en el aumento del flujo de la información, inicio incluso previo de la aparición del "internet": En 1944,

Freemont Rider describe en su publicación "The Scholar and the Future of the Research Library": En la biblioteca de la Universidad de Wesleyan; Ante el incremento en el número de investigadores, se percibió un incremento exponencial de los artículos académicos que han sido impresos, de tal manera que proyecta que en pocos años, para almacenar todas las nuevas publicaciones escritas se necesitaría de una inmensa biblioteca y un gran número de empleados para su administración (definición 1). En su momento, dicho fenómeno fue denominado como "Explosión de la información"; de acuerdo a la definición 2, la información almacenada en la biblioteca y el aumento del flujo a esta, puede ser considerada como "Big Data" a causa del volumen y variedad de los tópicos investigados.

Bajo el esquema de la "Web 2.0"; El aumento de la información disponible es característico de la era digital y la denominada era del "Big Data"; Distintas variedades y mecanismos que se han desarrollado para registrar la información del mundo real, generan a su paso, un conjunto de información tan complejo y diverso, ya que la información proviene de distintos dispositivos y sensores que poseen su propio lenguaje.

Sin embargo, a pesar de que exista almacenada la información en la biblioteca, su mera existencia no implica que sea usada para el incremento constante en el "stock de conocimientos" (Axioma 2, St. Thomas Aquinas). en el caso registrado por Rider, se requeriría del suficiente número de académicos y la suficiente organización de los ejemplares para poder hacer uso de las publicaciones y continuar su análisis o hacer discusión de estos, con la finalidad de crear nuevo conocimiento a partir de información que la biblioteca posee (definición 3).

Hoy en día, existen plataformas que permiten la búsqueda de contenido académico tal como Google Scholar, Pubmed, Elsevier, entre otros. Lo cual permite filtrar entre el stock, las publicaciones que son requeridas por el investigador para la elaboración de sus proyectos; Es decir, aun cuando el flujo de la información sea

exageradamente grande no necesariamente implica la creación de un conocimiento igualmente grande. Por ejemplo: si existen 204 millones de usuarios de Spotify, y cada uno de ellos deseara oír música al mismo tiempo, cada usuario emplea 2.5mb por minuto al reproducir la música en el dispositivo, para analizar dichos datos en tiempo real se requeriría de una banda que soporte 408 Terabytes por minuto; lo cual es un enorme flujo de información (redundante y repetitiva, es decir Bad Data), del cual el análisis integral de los 408 Terabytes sería totalmente fútil e innecesariamente complejo. Para esto, el análisis del comportamiento de los usuarios se simplifica al filtrar y analizar solamente la información relevante.

Para obtener conocimiento útil, para los servicios de la empresa y mejorar la satisfacción del consumidor; deberá de poseer la compañía información relevante que le permita analizar conductas de interés, tal como: la información, las características sobre los usuarios y las características de las canciones que son reproducidas; brindando a la compañía la capacidad de analizar tanto el comportamiento de los usuarios, como el brindar opciones basados en la estructura de la música que habitualmente escuchan, enlazando la información proveniente de dos fuentes distintas (la meta data de la canción y las estadísticas del usuario). La capacidad del análisis se puede expandir, al incluir variedad en la información que se obtiene del consumidor, por ejemplo: al incluir información proveniente de un dispositivo de Geo-Localización (GPS), con el cual es posible generar sugerencias dinámicas que se adapten al entorno en el que el usuario se encuentra, tal como la escuela, la biblioteca, el parque, etc. (Greenhalgh, Hazzard , McGrath, & Benford, 2016)

El volumen de los dispositivos conectados a la red a ido en gran incremento, que, en el año 2011 se hace entrega de la nueva versión de direcciones "Internet protocol" (IP) conocida como IPv6, la cual ayudaría a cubrir las direcciones únicas que su predecesor el IPv4 no podría cubrir, la cual se encontraba limitado a 4.3

Billones de direcciones únicas, por los dígitos que la componen. Lo cual implica, un incremento en los nodos que conforman la internet, los cuales son una fuente de información que transita por el sistema; se estima que para el 2020 exista un flujo de información de 40 trillones de Gigabytes (Gantz & Reinsel, 2012)

Algunos autores que definen al Big Data (BD) bajo las características que posee la información, usualmente incluyen en las "V"s, la "V" de valor, esto, debido a que su análisis posee potencialmente un alto valor. Sin embargo, no puede ser posible conocer si la información que se transmite posee algún valor hasta ser analizada; como en el ejemplo descrito, no toda la información posee un alto valor, a consecuencia de su redundancia y por el costo de analizar repetitivamente inmensas cantidades de registros, de tal forma que "big" no es necesariamente bueno, es decir, el alto flujo de información no necesariamente posee una alta calidad. De tal forma que el producto de la información de baja calidad es conocimiento de pésima calidad e inservible (a lo que se le conoce coloquialmente como "GIGO", Garbage in – Garbage out).

En cuanto a las características de la información, la temporalidad es un rasgo importante, debido a que describe sus limitaciones al momento de emplear la información para la creación de conocimiento; por lo cual, la velocidad con la que la información es adquirida y procesada, es un elemento clave para la maximización del beneficio obtenido de la información privilegiada:

$$\pi_K = K_P - K_C$$

$$\pi_K(t) = \frac{K}{(1+f)^t} - K_C$$

$$\Rightarrow \lim_{t \rightarrow \infty} \left\{ \frac{K}{(1+f)^t} \right\} = \psi - K_C \Rightarrow [(\psi + K_P) = K_C]$$

$$\therefore \pi_K(\infty) = \psi - 0$$

Es esta forma, el beneficio del conocimiento (π_K), es creado por la diferencia entre el conocimiento privado (KP) y el conocimiento común (KC). Cuyo beneficio permanece a lo largo del tiempo, mientras es descontado bajo un factor de propagación y mimetización del conocimiento (f), es decir: una vez publicado el conocimiento privado, esta se propagará a través de los demás agentes, haciendo que la información privada sea en algún punto en el tiempo información pública, de tal forma que el rendimiento de la información privada se vuelve 0. Al mismo tiempo existen factores que limitan la propagación del conocimiento dentro del sistema (ψ) tal como patentes, limitación cognitiva de los agentes rivales o la privacidad (Ej. secretos empresariales); Manteniendo el beneficio mientras esta barrera exista.

$$\Delta K = \log(X) \text{ s. a. } (C_k^P = C(X) + C(S))$$

Siendo la maximización del beneficio el objetivo principal de los agentes, la maximización de la " π_K ", el cual depende de la minimización del costo y del tiempo empleado en el análisis de la información; ante el constante incremento en el flujo de la información, los factores limitantes de la maximización son los costos de adquirir información adecuada " $C(X)$ " y las habilidades que el agente emplea para su cifrado " $C(S)$ ".

Dichas limitantes de producir conocimiento, son los factores que determinan la brecha entre el stock de información (la "data"), y el conocimiento que es obtenido del ambiente, empleado para toma de decisiones. De tal modo, se han desarrollado métodos automatizados que permiten reducir dichos costos, estos metodos son empleados para procesar y analizar, grandes cúmulos de información; tales como el "data mining" o el "machine learning", estos métodos permiten extraer información de los registros más complejos, con el fin de poder incluir dentro de los modelos estadísticos, información proveniente de registros no-estructurados (como imágenes y audio);

De tal forma se puede definir el termino de BIG DATA, de la siguiente forma:

***Def. Big Data,** Es el término que describe la información que es extraída a través de diversas fuentes, la cuales no poseen una estructura homogénea; para este grupo de información, existe algún método con el cual la información puede ser manipulada, analizada o visualizada, con el fin de obtener algún “conocimiento” útil para la toma de decisiones, de tal modo que el método empleado sea tan veloz como el flujo de la información. Por lo tanto, al no poseer algún método para la extracción del conocimiento, los registros recolectados serán solo un depósito de información. (data warehouse)*

Del BD la cualidad más importante, al criterio del autor de esta tesis, es la característica de la diversidad de las fuentes de información; la cual debería de ser considerada más importante que la característica que hace referencia al tamaño, stock o el número de registros de una base de datos. Debido que, para poder entender la complejidad de un fenómeno, es preciso obtener todos los fragmentos que lo componen, comprendiendo la naturaleza de todos los elementos que conforman a la estructura observada.

2.2 Características de la información.

Tras la masificación de los dispositivos electrónicos y de su conectividad con el entorno digital; las cámaras, micrófonos y mensajes poseen mayor cobertura, al registrar un mayor número de hechos o fenómenos que acontecen en la actualidad, este tipo de información es catalogada como “no-estructurada”; tal información, permite complementar la información que puede ser observada dentro de una base de datos estructurada con variables que pueden ser difícilmente mesurables y comparables; requiriendo de las habilidades cognitivas humanas para ser complementadas o tomadas en cuenta. (Ej. Describir las características físicas de la estructura de un automóvil provenientes de una foto).

Se estima que para el 2025, el 80% de la información mundial será no-estructurada; es decir, el 80% de la información estará almacenada en archivos de texto, fotografías y audio, que crece continuamente. Por lo cual se requiere de la suficiente capacidad técnica y humana para ser analizada, como en el ejemplo de Rider (1944); Ejemplo de ello, se estimaba que en año 2003, el 30 o 40 por ciento del tiempo laboral de los trabajadores administrativos se centra en la administración de los archivos con información no estructurada. (Blumberg & Atre, 2003) Siendo esta importante para la organización sin embargo su procesamiento requiere de un costo importante al requerir ser analizada manualmente.

Tras el cambio al "Web 2.0", las compañías no son las únicas capaces de crear y difundir información en la red, la transición a este esquema 2.0 ha brindado las herramientas a los usuarios para crear su propia información al emplear la diversidad de dispositivos que se encuentran conectados a la internet; los cuales, permiten realizar diversas actividades que poseen la capacidad de ser transformados en registros (o datos).

Con una amplia gama de dispositivos y servicios en-línea: los usuarios forjan un alto volumen de información sobre su actividad diaria, actualizándola por medio de diversos canales, creando un vasto flujo de información disponible:

Los dispositivos conocidos como IoT (Internet de las cosas) permiten medir numéricamente la actividad humana, animal o de algún objeto; genera un flujo de registros en alta frecuencia que se actualiza en tiempo real (o con baja latencia), en todo momento que la entidad permanezca activa. Capturando información de alguna característica específica del objeto analizado, tal como: ritmo cardíaco, localización, velocidad de un objeto, calor corporal, movimientos, transacciones, entre otros. Usualmente los dispositivos IoT son empleados por el público en general, para interconectar a la red elementos físicos (como las luces) que pueden ser manipulados por medio de otros dispositivos.

La información generada por los dispositivos IoT, es conocida como "machine data", denominado de esta forma, por ser el medio por el cual las maquinas se comunican entre sí, creando registros parcialmente legibles por el ser humano, pero totalmente legible por el computador para poder realizar sus funciones; para la lectura de su información se requiere de un dispositivo que traduzca dicha información al usuario que lo emplea.

Al mismo tiempo, mediante la interacción en el Internet, los usuarios generan un rastro de información que es heterogénea en términos cualitativos y cuantitativos, el comportamiento de los usuarios es almacenado en distintos registros, tales como: historial de navegación web, "micro-blogging" o posts en redes sociales, mensajes de textos, e-mails, formularios y meta data del dispositivo empleado y demás... Dicha información es almacenada en lenguajes de bajo nivel (código binario) dentro de las maquinas, la cual es traducida y expresada en diversas formas de archivos de texto o multimedia, que pueden ser comprendidos por el ser humano, pero su contenido no puede ser interpretado por una máquina; denominando este tipo de información como "Natural Language".

De esta forma, la información que puede ser recolectada en el entorno informático, posee las siguientes denominaciones según sus características (como se muestra en la ilustración 4)

Información estructurada: Información que ha sido procesada, es almacenada bajo una estructura predefinida y un orden homologado. su estructura está especificada por campos diseñados para la captura de datos (registros) sobre elementos específicos que conforman a la entidad observada (Ej. Elementos de un CD: Autor, # canciones, Portada, Descripción, Genero...).

La forma de almacenar este tipo de información es mediante el uso de bases de datos SQL (Structured Query Language), tal como bases de datos u hojas de

cálculo, las cuales permiten generar consultas organizadas, mediante el empleo de tablas que poseen elementos clave que les permite relacionar a las entidades observadas por los atributos que poseen (Ej. cliente, cuenta, # de compra).

Información no estructurada: la forma más común en la que la información es almacenada en la web 2.0; podría ser definida como “datos en estado natural”, ya que se requiere de la interpretación humana para dar estructura al archivo al obtener de ella información que permita etiquetarla posteriormente al ser interpretada, por ejemplo: otorgar una calificación a algún ensayo o tarea escolar, determinar el contenido de una fotografía o comprender alguna orden dentro de un archivo de audio.

Debido a su naturaleza dinámica y heterogénea, los datos no poseen un patrón común, pero pueden ser catalogados bajo el mismo campo semántico; por ejemplo: los datos provenientes dentro de un e-mail no poseen siempre un mismo asunto o fin, por lo que el contenido del mensaje no posee un patrón predefinido. Debido a estas características heterogéneas, su almacenamiento se aloja en bases de datos denominadas como “NoSQL”.

El esquema de la información no-estructurada, es ampliamente flexible, a la cual se le da determinada estructura dependiendo del tipo de información que se desea capturar o transferir:

- **Basados en Texto;** modelo altamente flexible que permite almacenar gran número de registros sin emplear “columnas” predefinidas, por lo que pueden ser añadidas o emitidas al crear un nuevo registro (por ejemplo: documentos XML y JSON).
- **Basados en Elementos Binarios;** también conocidos como “BLOB” (Binary Large Object), que permite almacenar archivos o información en forma de una cadena binaria en la base de datos, que puede ser traducida posteriormente

por la máquina, lo cual permite almacenar archivos multimedia (por ejemplo: Archivos multimedia que son enviados por medio de la aplicación Whats App)

- **Basado en Nodos;** Almacena la información de la entidad y las relaciones con que posee con otras entidades, lo cual permite crear mapas de nodos y observar sus interacciones.

Ilustración 4, Categorías de la información.



Fuente: Elaboración propia.

El incorporar nuevos métodos automatizados para el análisis de la información no estructurada; da como resultado, la ampliación en el abanico de opciones de las fuentes de información que pueden ser empleadas en él análisis; dotando al agente la capacidad de procesar mayor cantidad de información en lapsos de tiempo más cortos, determinando las fuentes (o visualizando la información) que puede ser de utilidad e incluso automatizar el proceso del análisis y obtener bajo la cognición del ordenador, el conocimiento necesario para la toma de sus decisiones.

Al dotar de los ordenadores de una capacidad cognitiva (la capacidad de decidir en base la información que se le es suministrada) y la capacidad de analizar el Lenguaje Natural, son elementos altamente novedosos, debido que su desarrollo y las posibles aplicaciones "inteligentes" crea nuevas formas dinámicas de producir y consumir información.

Tales atributos dotan a las computadoras la capacidad de poder analizar el contexto en cual la información se desenvuelve, permitiéndoles analizar la información que va más allá del "número", como las imágenes, videos o audio. Siendo capaces de describir y clasificar los elementos que componen la información en su "estado natural", Identificando los mensajes que se encuentran contenidos en los diversos canales de baja latencia, tal como la conducta de un individuo (Ej. expresiones faciales humanas, estrés en la voz, propiedades físicas del lugar y demás).

Algunos de estos avances, por ejemplo, aquellos llevados a cabo en la Universidad de Adelaide, Australia (Dick & Brooks, 2003): Quienes han desarrollado algoritmos capaces de analizar los videos que son producidos por las cámaras de seguridad pública, con el objetivo de crear diversas aplicaciones derivadas del análisis visual de la conducta humana; Tal como el comportamiento del tráfico automotriz, búsqueda de objetos (o rostros) específicos o la identificación de conductas sospechosas.

También, el desarrollo de algoritmos, da la posibilidad de procesar documentos de texto (como documentos en formato PDF) de elaborar análisis fundamental al instante mediante el empleo del "Text Mining"; empleando algunas formas básicas de este tipo de análisis, esto es capaz de reconocer los sentimientos de las palabras (Ej. opinión general del público en Twitter) o también, puede crear nubes con las palabras importantes de diversas fuentes de noticias y artículos, haciendo énfasis en aquellas que están siendo buscadas, como se muestra en la ilustración 5. De tal modo, que se puede facilitar la previsualización de la información contenida en el documento, permitiendo decidir al usuario si usara o no, la información contenida en algún documento.

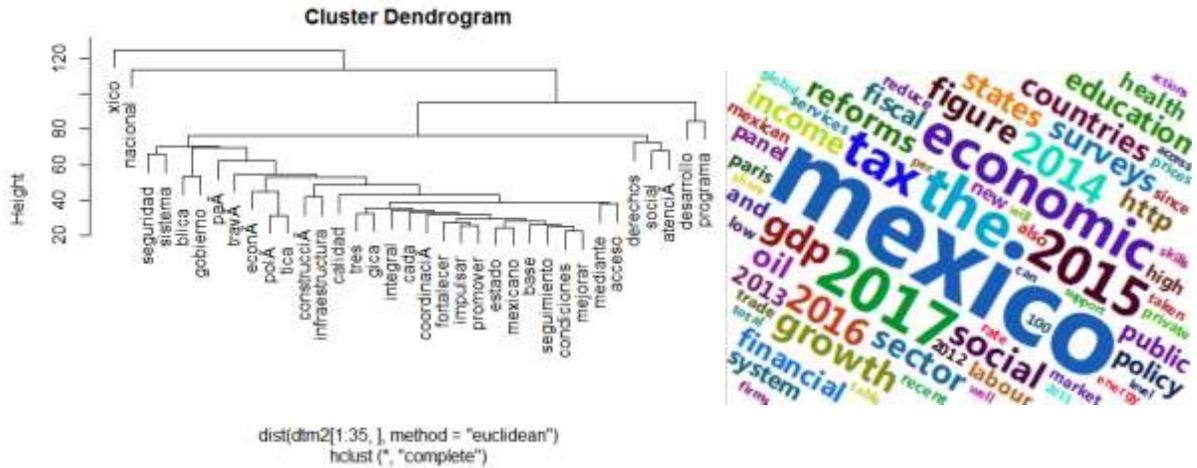


Ilustración 5 Dendrograma de palabras y Nube de Palabras, ejemplos básicos del text mining, Elaboración Propia; Fuente: 4to informe de gobierno (2015-2016) y OCDE Economic Surveys Mexico (2017); El dendograma muestra las palabras mas importantes (frecuencia) del texto y sus relaciones (palabras comunmente cercanas a ellas); la nube de palabras solo muestra la frecuencia de las palabras.

Con estos métodos, se le ha incorporado la capacidad de que el usuario pueda interactuar con la máquina, permitiéndole obtener respuestas automáticas de acuerdo a la información que se tiene y la información que se solicita mediante el uso del lenguaje natural; adelantos como el ChatGPT, permite al usuario obtener información instantánea sin recurrir a diversos costos, tal como el encontrar a la persona adecuada que posee la información o tener que revisar todo el documento para obtener la pieza adecuada de información (o algún consejo personalizado). (George, 2023)

Como se ha descrito, la variedad de la información es una cualidad del big data, con el modelo web 2.0, permite a cualquier agente generar sus propios fragmentos de información bajo su propia voluntad. Por lo que, los agentes pueden publicar información sin requerir de una entidad central que regule o verifique el flujo de información; de tal modo la calidad de la información puede ser cuestionable, al tratarse de fuentes tan diversas no es posible verificar la veracidad de cada una de las fuentes que produce la información ni de los estándares

empleados en la elaboración de la meta data o las etiquetas con las cuales es almacenada la información, dificultando su búsqueda, recopilación y organización.

En una forma metafórica: El Big Data, es una biblioteca extensa de libros de toda clase, tamaño y era. Cuya información recopilada en ella puede ser monumental, por lo que no le es factible para un individuo leer y “aprender” de cada uno de estos libros; El potencial descriptivo de la información se ve limitada por las limitaciones físicas de la búsqueda del conocimiento, en este caso: el tiempo. Por lo consiguiente, la interrogante importante sigue siendo: ¿Que conocimientos deseo adquirir de ella? y ¿Para qué?

2.3 El Beneficio Del Análisis Del Big Data.

Dentro del mercado de los seguros (EIOPA, 2019), siendo este el ejemplo más ilustrativo de las consecuencias de la información asimétrica, tal como el Riesgo Moral. Algunas empresas dentro de este mercado, han desarrollado algunas iniciativas basadas en Big Data para mejorar los servicios ofrecidos. Al emplear nuevas fuentes de información, proveniente de la IoT, de la actividad en línea del cliente (actividad web, perfil social y laboral) y de la información interna, como: registros en el uso de los medios de comunicación con la aseguradora (Ej. Call-center, aplicaciones o página web), mejorara la calidad de su negocio en diversos aspectos:

- **Desarrollo de productos ajustados a la exposición al riesgo:** conocidos como seguros “basados al uso”; por medio del uso de dispositivos IoT, tal como en el caso de la compañía AIG (SXKM, Seguro por kilómetro), se desarrolla un servicio de seguro que se adapta a las necesidades del cliente y adicionalmente brinda servicios requeridos por el uso del vehículo (como asistencia vial y servicios de mantenimiento). Otorgando la suficiente información para recompensar los

buenos hábitos (conducta de bajo riesgo) y ponderar correctamente a los conductores con alta probabilidad de siniestralidad.

- **Venta de productos:** Por medio del empleo de análisis detallado del cliente, se puede determinar el perfil o el segmento del mercado al que pertenece; con tal de ofrecer los servicios necesarios de acuerdo con sus hábitos, ofrecer asesoría personalizada y distribuir los productos, empleando los canales precisos y preferidos por el cliente. Que adicionalmente, se busca reducir la fricción creada en los procesos de contratación y de los servicios otorgados.

- **Detección de fraudes:** empleando la información proveniente de sensores biométricos, se puede legitimar que el cliente es quien realiza los procesos de reclamación de la suma asegurada; se emplea algoritmos especializados para evaluar la autenticidad del siniestro y evaluar los daños ocasionados, al analizar las fotografías tomadas del siniestro, evaluándolas con algoritmos especializados.

- **Fijación del precio:** por medio de procesos analíticos sofisticados, se puede micro-segmentar en distintos segmentos, ofreciendo propuestas de seguro a nivel personal, basados en las características económicas, sociales y la conducta del cliente; para ello es empleado la información individual de cada cliente, con la finalidad de mantener su fidelidad dentro del proceso de renovación de la póliza, en el cual usualmente el cliente observa los precios de la competencia. Sin embargo, ante las regulaciones en la protección de información privada y las políticas de no discriminación (ej. sexo, raza o religión), puede dificultar las evaluaciones individuales, a lo cual puede ser un freno moral al momento de otorgar un "pricing" individualizado. (Xu, Jian, Wang, Yuan, & Ren, 2014)

Los beneficios prometidos en el análisis y aplicación del conocimiento obtenidos de los grandes cúmulos de información: son la minimización de las fricciones en los procesos económicos, la predicción de los siniestros

(materialización del riesgo) y en la disminución del riesgo moral. Ante las nuevas capacidades analíticas se abren nuevas formas de organización y nuevas herramientas en las distintas áreas, por ejemplo:

- **Planeación Urbana:** Simulaciones ambientales y optimización en el empleo del suelo, ante las necesidades de los individuos que viven en las áreas adyacentes.
- **Ciencias Sociales computacionales:** Estudio y comprensión del individuo dentro de su propio ambiente natural, en el cual habitualmente se emplea el microblogging, al ser la plataforma por la cual los individuos expresan sus reacciones.
- **Inteligencia en Transportación:** Optimización en la logística, mediante la creación de los recorridos más cortos o con menor consumo de combustible y tiempo.
- **Modelación del ambiente:** Comprensión del ambiente, bajo la simulación histórica; tomando en cuenta gran cantidad de variables físicas, con la finalidad para comprobar el estado final observable, con las hipótesis establecidas previamente.
- **Materiales Inteligentes (Biotecnología):** Desarrollo de materiales sintéticos, al simular los posibles resultados al combinar distintos materiales para la fijación del polímero con las características deseadas en resistencia, peso o densidad.
- **Seguridad Pública, y Web:** Por medio de algoritmos automatizados (Inteligencia Artificial), es posible monitorear de tiempo real todos los eventos de los cuales se posea información, como en el ejemplo de los análisis de las

imágenes captadas por las cámaras de seguridad, de la universidad de Adelaide (2003).

- **Bussines Intelligence:** Generación de conocimiento sobre los procesos internos, para generar mejores decisiones ante las circunstancias del ambiente económico y los problemas internos.

La necesidad de la información es crucial; para cada agente económico, la información y la creación de su propio conocimiento influye en la toma cotidiana de decisiones, las cuales determinan resultados eminentes y las consecuencias del largo plazo. (Rae, 1834) El agente al poseer información completa en el mismo instante en que fue divulgada; le permite ejercer una decisión oportuna y precisa, permitiendo maximizar el beneficio obtenido al emplear sus recursos de la mejor manera posible (eficientemente) (Grant, 2001); el conocimiento pleno permite aprovechar cada eventualidad u oferta existente en el mercado: Ante el ambiente de competencia, la creación de su conocimiento permite obtener una ventaja competitiva en el mercado (Hayek, 1945), haciendo del agente el más "creativo"⁹ y adaptable ante situaciones de desequilibrio.

Los agentes, con la mera interacción con el ambiente; genera pequeñas modificaciones en el mercado, al mismo tiempo que obtiene fragmentos de información del mismo mercado, interpretando las señales de los demás agentes y el análisis de los resultados obtenidos posterior a la interacción¹⁰. Para que la asimetría de la información desapareciera, sería necesario de un mecanismo efectivo, que permita recibir y transmitir toda la información del mercado al sistema

⁹ Entiéndase "Creatividad", como la capacidad del agente de poder idear "n" soluciones distintas ante un problema específico, sujeto al conocimiento previo que posee.

¹⁰ El conocimiento social, corresponde a la mimetización del comportamiento de otros participantes que el agente observa, el cual es transferido automáticamente a medida que el agente interactúa o permanece en un entorno específico; "How Social Perception Can Automatically Influence Behavior" Ferguson & Bargh (2004)

económico; de esta manera, cada uno de los agentes conocería exactamente la sucesión de acciones que deberán de realizar en el sistema económico, de tal modo, cada una de las acciones realizadas no interferiría con los planes de otros agentes participantes. (Hayek, 1948)

En el sistema económico existen elementos que no permiten alcanzar la eficiencia de la información: Ante la premisa de que los recursos son finitos, yace la imposibilidad de interferir con los demás participantes cuyo fin u objetivo es el mismo, por lo que existe el incentivo de preservar la información privada; ante la existencia de la división de trabajo, cada uno de los participantes de la economía posee solamente un fragmento de la información que obtiene de acuerdo a su entorno especializado, de tal manera que existe la posibilidad de que un agente oculte información a los demás agentes, dando como resultado la posibilidad de riesgo moral (tal como el dilema del Agente - Principal); así mismo, el agente tendrá incentivo a aumentar el stock de información privada ante la insistente búsqueda por descubrir lo desconocido (nuevas fuentes de los recursos, nuevos clientes, proveedores, modelos y métodos productivos). Con ello se generan continuos desequilibrios y modificaciones en el mercado, que alteran continuamente las expectativas y los planes de los agentes.¹¹

En una economía de recursos finitos, de división de trabajo y de la percepción de la existencia de lo desconocido; existirá dentro del mercado constantes desequilibrios económicos, por los cuales, los agentes necesitaran adaptar continuamente su comportamiento, requiriendo de buscar reiteradamente

¹¹ Las acciones conocidas como "societales", son las acciones del agente que nacen de su instinto; influenciado bajo la información social que se encuentra segmentada y deja de lado la racionalidad creada por la información que el agente a recolectado. De tal modo, la acción "societal" genera constantemente desvíos en las expectativas de otros agentes; modificando el comportamiento de los agentes dentro del sistema cuando una "nueva conducta" es descubierta y mimetizada por los demás agentes.

el conocimiento necesario que le permita al agente minimizar sus costos y maximizar su beneficio al emplear óptimamente sus recursos. Requiriendo de esta forma, acciones (que poseen costos) para revelar sus preferencias (lo que el agente oferta) y acciones para revelar el "tipo" de los agentes contrarios con los que se planea realizar la negociación (sea proveedor o cliente).

El factor determinante para alcanzar la eficiencia, se encuentra en la estructura por la cual la información viaja; la información es empleada para la comprensión del mercado, por ejemplo: ¿Cuál será la demanda del producto "A"?, ¿Cuál es la utilidad del bien para el consumidor? O ¿Cuál será el precio que otorgue la máxima ganancia?; Dichas preguntas pueden ser contestadas por medio de un análisis de mercado (marketing), requiriendo de información sobre el mercado meta, obtenida de eventos ocurridos en el pasado (Ej. Volumen de ventas del año pasado) y de la interacción directa con los consumidores (como un grupo de enfoque). Permitiendo perfeccionar los elementos que conforman la mercancía producida con las características deseadas por el cliente, en base a las expectativas descubiertas de los agentes observados. (Pandian, 2007)

El Big Data, tienen la cualidad de poseer información en gran abundancia sobre los individuos que participan dentro de la red, permitiendo conocer a profundidad a cada uno de los individuos sin tener la necesidad del trato presencial. (Ravi & Kamaruddin, 2017) La información puede ser analizada mediante el empleo de algoritmos específicos, los cuales permitan filtrar la información, separarla y conjugarla; con la finalidad de obtener de ella, un conjunto de conocimientos sobre las características que describen al agente y su tipo; permitiendo conocer cuál puede ser la reacción futura de este, prediciendo los posibles planes de acciones "racionales", de acuerdo con la información que se posee de sus acciones previas,

descubriendo el patrón de las acciones que el agente repite sistemáticamente bajo el principio de “placer-dolor”. (Foxall, 1993)¹²

Los mercados que poseen un gran número de participantes son proclives a caer en estados de información asimétrica sobre las características de algún bien, servicio o persona. Ante la singularidad de cada individuo, su capacidad cognitiva y la búsqueda de maximizar su utilidad a expensas de los demás participantes; propician en el sistema económico, un ambiente propenso a la selección adversa. Por lo que es necesario para el banco, el análisis exhaustivo de sus clientes, con la finalidad de colocar los recursos de manera óptima: al minimizar las repercusiones por exponerse a los riesgos de la contraparte.

La información recolectada puede ser analizada de las siguientes 3 formas distintas: (Ravi & Kamaruddin, 2017)

- **Análisis Descriptivo:** Mediante la información recolectada de los eventos pasados y mediante la evaluación estadística se puede determinar los factores que influyen en el comportamiento de un objeto de estudio; comprendiendo las relaciones entre los factores y sus efectos.
- **Análisis Predictivo:** mediante el análisis descriptivo, pueden ser intuidas las futuras reacciones, distinguiendo los patrones que le permitan inferir estadísticamente las expectativas sobre la conducta de las entidades observadas.
- **Análisis Prescriptivo:** Permite conocer el impacto que las entidades pueden experimentar, al simular distintos escenarios posibles a partir de un

¹² De acuerdo con el modelo del comportamiento del de Freud (la teoría psicoanalítica): como un método de selección natural, los agentes repiten constantemente las acciones que le otorga mayor bienestar personal, y evita el comportamiento que les provoca algún daño o dolor; cuyas sensaciones son originadas de las consecuencias de sus acciones, como el miedo o el deseo, contribuyen en el mecanismo de toma de decisiones de los agentes. “Consumer Behaviour as an Evolutionary Process” Gordon R. Foxall (1993)

punto de referencia; permitiendo deducir las consecuencias esperadas, por medio del conocimiento recolectado del análisis descriptivo y predictivo.

A través del procedimiento llamado "data mining", pueden ser creados algoritmos que poseen modelos estadísticos automatizados, que cuentan con la suficiente capacidad para poder ser capaz de comprender y descubrir patrones sobre el comportamiento del objeto analizado con constantes flujos de información; Permitiendo a las empresas comprender un panorama que difícilmente puede ser observado a simple vista, debido a la gran complejidad de su naturaleza y la imposibilidad de analizar globalmente la información. (Computing Community Consortium, 2012) Por ejemplo, el panorama mundial de la economía, puede ser difícilmente analizada: ya que deberán de ser analizados los distintos eslabones de una cadena productiva y la información local necesaria de cada uno de los eslabones.

***Def. Data Mining**, o "Minería de datos", es la creación de algoritmos que permiten limpiar, ordenar y analizar la información que se posee; permitiendo identificar patrones, crear pronósticos y simular posibles escenarios mediante los métodos científicos de la estadística.*

Algunos bancos han aprovechado los beneficios del "Data Mining", lo cual ha permitido mejorar distintos aspectos en la operatividad del banco; tal como la detección de fraudes y una mejoría en la calidad de la relación con el cliente (Marr, 2016):

El Banco Real de Escocia (RBS), como parte del plan para ganar terreno en la industria de la banca comercial se optó por crear iniciativas de "personología", creando conocimiento de los clientes a nivel personal. Pretendiendo comprender las necesidades de cada uno de ellos y de la utilidad que concedía los créditos personales que eran otorgados; Ideando una forma de generar sugerencias personales y mejorar el trato individual en las negociaciones de los distintos productos ofrecidos.

Entre 1970 y 1980, afirma Nelissen (Analista de RBS): los bancos han perdido la relación y la confianza de sus clientes, concentrándose solamente en la fijación de metas en ventas (y de ganancias), creando productos de los cuales no se poseía ninguna certeza sobre su utilidad o beneficio que el cliente obtenía de ellos. De esta forma la estrategia de competitividad de los bancos se basaba en las ofertas y en el precio, más que en las necesidades reales del cliente.

La iniciativa de Nelissen, fue desmontada por los clientes quienes no demandaban dichos servicios, siendo este considerado como un costo innecesario del que no se obtenían grandes ventajas, más que recibir una felicitación en el día de sus cumpleaños; Esto, debido a que se experimentaron un aumento en el precio por el servicio adicional, debido a que los procesos de análisis se realizaban continuamente, lo cual resultaba altamente costoso al realizar dicho proceso manualmente; tal costo, se trasladó al precio final del servicio, afectando la satisfacción del cliente.

Tras la evolución de las capacidades tecnológicas, en el 2015 (7 años después de la crisis financiera del 2008), se optó por retomar el concepto propuesto por Nelissen, empleando la información contenida en los datos personales de sus clientes y en el historial de sus transacciones, determinando la utilidad individual de los servicios y productos ofrecidos; perfeccionándolos en base al conocimiento obtenido por el "análisis profundo del cliente" (proceso conocido Customer Analytics). Analizando la información mediante el empleo software tipo CRM (Customer Relationship Management) desarrollado por Pegasystems; mejorando la administración bancaria, sus servicios y contribuyendo la adaptación de los servicios de ventanilla, a la "banca digital" multicanal.

Mediante el Data Mining, RBS, inicia un proceso el cual le permite subcategorizar los nichos de mercado a niveles individuales: aunque los clientes posean tarjetas de crédito con características similares en el mismo banco, cada

persona hace uso de ella bajo su propio gusto y criterio. Determinando mediante el CRM: ¿Cuál es el propósito del producto?, comprendiendo el comportamiento del cliente, centrándose en sus necesidades y el empleo del financiamiento. Perfeccionando los servicios relacionados que son ofrecidos por el banco, tal como el “cross-selling”, ajustando tales promociones a las necesidades de cada segmento del mercado y a la individualidad de cada cliente.

Los beneficios esenciales del CRM en cualquier perspectiva empresarial, se traslada a todas las fases que se tienen en la relación cliente-empresa: la adquisición del cliente, su retención, la construcción de una relación profunda con este y el aumento del volumen de las transacciones:

Cuadro 2. Aplicación de Análisis Big Data en Customer Analytics¹³ (algunos ejemplos)

| Análisis de Sentimientos | Customer 360 | Segmentación de los clientes | Siguiente Mejor Oferta. | Omnicanal |
|---|---|---|---|---|
| <ul style="list-style-type: none"> • Aceptación de los productos. • Identificación de clientes influyentes. • Análisis de retroalimentación. | <ul style="list-style-type: none"> • Personología del cliente. • Comprender la relación entre el cliente-producto. • Detección oportuna de abandono. | <ul style="list-style-type: none"> • Marketing focalizado • Pricing óptimo de los servicios | <ul style="list-style-type: none"> • Asociación de productos. (Cross-sell) • Programas de Lealtad | <ul style="list-style-type: none"> • Comunicación mediante canales deseados por el cliente. • Optimizar la masificación de la información. • Reducir los costos de comunicación. |

| | |
|------------------------|---|
| Risk Management | Análisis predictivo de exposición |
| | Detección de cuentas de alto riesgo |
| | Detección de fraudes y procesos de lavado de dinero |
| | Customer Life Time value |
| | Sensibilidad del cliente con respecto al precio (elasticidad) |
| | VaR estimado del cliente |

En otro caso, en el Banco Fifth third (FTB), el cual reside y opera en la localidad de Ohio, EUA. Mediante el empleo de las herramientas fabricadas por IBM

¹³ Basado en la información de “Big Data on Banking for Marketers”, Evry (2013) & “Big Data Alchemy”, Capgemini consulting (2013)

(2018): las cuales consisten en algoritmos creados a partir de bibliotecas en los lenguajes computacionales como "R" y "Python", entre otras estructuras como "Keras" y "Tensorflow", siendo los algoritmos diseñados y almacenados en libretas elaboradas en "jupyter". Lo cual, fue una total sorpresa para la expectativa de la compañía, ya que la innovación tecnológica se encontraba en programas de código abierto que su uso se encuentra abierto a todo el público.

La principal aflicción de FTB, fue el crear un entorno seguro para manejar información sensible, permitiéndole almacenar la información y los resultados en ambientes seguros bajo sus propias políticas de seguridad. Al mismo tiempo, el FTB, buscaba crear dicha estructura sin sacrificar la capacidad de analizar gran cantidad de información al menor costo posible, con la mayor efectividad y eficiencia viable. (Al incrementar los procesos de seguridad, se incrementan los costos de transacción, debido a las fricciones generadas ante la inclusión de los procesos de seguridad; en el caso de la incorporación del software de IBM, se consigue mejorar los procesos de análisis sin incrementar en gran medida los costos por los procesos de seguridad)

Con la finalidad de maximizar el beneficio obtenido de la información; se mantiene el objetivo de producir los resultados en el menor tiempo posible y amentando su capacidad de exploración. El FTB emplea para el análisis un clúster de servidores bajo la estructura "Hadoop" que es escalable horizontalmente, la cual permite almacenar una gran cantidad de información en distintos nodos (servidores) y administrarlos como si fueran uno solo. Permitiendo aumentar la velocidad del análisis a medida que el número de nodos aumenta, este proceso usualmente llamado "Map Reduce", reparte la información al número de nodos disponibles para que cada uno realice su trabajo individualmente y entregue sus resultados al nodo principal el cual compila dichos resultados obtenidos.

Los nodos de este modelo trabajan paralelamente, empleando modelos estadísticos tradicionales, así como los modelos basados en algoritmos tipo "machine learning", tal como el "Deep Learning" (Sistemas de Redes Neuronales "NNet"). Con estas herramientas se crearon modelos capaces de predecir los volúmenes de los depósitos, al virtualizar múltiples escenarios con distintas tasas de interés. Esto, con la finalidad de buscar predecir el volumen futuro de las cuentas de los clientes, permitiéndole crear estrategias agresivas para su atracción e incrementos o disminuciones en los depósitos; aumentando la certeza sobre la liquidez que el banco poseerá en el futuro y permitiéndole mantener a los clientes que ya posee.

Para poder relacionar los distintos fragmentos obtenidos de fuentes heterogéneas de información que componen el "Big Data", se requiere de la estandarización de la información producida dentro de las distintas industrias, lo cual permitiría producir modelos dentro de toda la cadena productiva, desde la creación del insumo hasta el consumo final del producto:

Empresas como Alibaba (Zhou, Chen, & Li, 2018), una empresa basada en plataformas digitales; puede obtener información de la cadena productiva al poseer plataformas virtuales que cubren las distintas necesidades en las distintas fases de la cadena, así como también canales para la venta de artículos (e-commerce). Alibaba, desarrolla actividades financieras, atendiendo a grupos poco habituales de la banca tradicional, tal como los productores agrícolas y las pequeñas empresas.

Alibaba por medio de "Ant-finance", posee un modelo financiero que se ajusta a los ciclos productivos y a las necesidades de los trabajadores rurales a nivel microempresa (crédito) al mismo tiempo que a nivel personal (préstamo); Sin necesidad de una tarjeta de crédito, los clientes de Rural Taobao (o T-Mart, parte del grupo Alibaba) son analizados en tiempo real en base a la información obtenida a sus compras, a su capacidad de cumplimiento, características sociales y

conexiones sociales; empleando a la organización Zhima Credit (Subsidiaria de Alibaba) determina con cero intervención humana el nivel de riesgo, dándole al cliente un score instantáneo el cual es actualizado constantemente, con el cual Ant-Finance mantiene un nivel bajo de morosidad, financiando a un público considerado de muy alto riesgo.

2.4 Herramientas para el Análisis

Para poder adquirir los resultados de los casos descritos previamente; se requiere de un proceso que se encuentra sujeto a limitaciones técnicas y tecnológicas del entorno. Pese a que existe software abierto al público en general, los usuarios deben de poseer para su empleo el suficiente conocimiento para desarrollar sus propios modelos, para un Data Scientist o Científico de la información, debe de poseer habilidades computacionales, estadísticas y el conocimiento sobre el tema del cual realizara el análisis, con el fin de poder interpretar los resultados obtenidos.

Def 5. Data Scientist, Extensión evolutiva de un estadista, siendo capaz de manejar grandes cantidades de información. Requiriendo habilidades computacionales para la creación de algoritmos que sean eficientes: sin redundancias, mínimo uso de recurso y menor tiempo de respuesta. (Cielen, Meysman, & Ali, 2016)

Adicionalmente a las fuentes de información alternas (como los datos obtenidos del Iot) se requiere de una metodología y de algoritmos específicos que permitan elaborar modelos estadísticos, con la finalidad de encontrar de ella patrones que no son visibles a simple vista o que no se habían contemplado previo a la investigación. Lo cual es distinto al obtenido de los procesos estadísticos descriptivos, en donde dichos se centran en comprender las variables y sus relaciones; determinando los factores de causalidad y probando su significancia mediante hipótesis establecidas previo a la investigación; los cuales son empleados en el "Data mining", para determinar la relevancia de usar cierta variable para la generación de predicciones.

A diferencia de la estadística tradicional, el "Data mining", es un proceso automatizado que explora, organiza, y compila la información; Buscando en ella patrones, que permite generar predicciones acertadas incluso al conjugar variables con poco poder de predicción. (Capgemini Consulting, 2013), con lo cual en ocasiones imposibilita comprender los resultados obtenidos, siendo este efecto conocido como "black box". Aunque, es muy común encontrar frases flameantes como "machine learning", o "Deep learning" dentro del análisis de Big Data, estas es una pequeña parte del proceso que se requiere para la alquimia de la información (transformación de datos en oportunidades).

Con la finalidad de explorar las capacidades y limitaciones técnicas del análisis del Big data; Se examinará la metodología de la compañía IBM (2012), denominada como CRISP-DM (Cross Industry Standard Process for Data Mining), el cual como su nombre lo indica, es una metodología estándar para realizar data mining; el cual es un método reiterativo, donde sus procesos deben de realizar de manera cíclica, con la finalidad de brindar resultados satisfactorios.

Cuadro 3: Método CRISP-DM. Elaboración propia; basado en IBM (2012)

| | |
|------------------------|---|
| Bussines Understanding | <ul style="list-style-type: none"> •Conocer previamente las problemáticas de la empresa •Elegir las razones para elaborar un proyecto •Conocer las capacidades de la empresa |
| Data Understanding | <ul style="list-style-type: none"> •Organización de la información disponible •Exploración de la información •Conocer las relaciones de la información •Determinar la Calidad de la información |
| Data Preparation | <ul style="list-style-type: none"> •Limpiar y seleccionar la información •Integrar la información, (crear nuevas variables) |
| Modeling | <ul style="list-style-type: none"> •Selección de modelo a emplear |
| Evaluación | <ul style="list-style-type: none"> •Interpretación de los resultados obtenidos •Comprobación de la confiabilidad de los resultados del modelo. |
| Toma de Decisión | <ul style="list-style-type: none"> •Toma de decisiones en base a los resultados obtenidos. |

2.4.1 “Business understanding” o Características Empresariales.

El primer paso, es el determinar los objetivos del proceso, hacer las preguntas correctas y determinar qué clase de respuestas se están buscando. Aunque podría ser entretenido emplear variables al azar; los resultados obtenidos resultarían ser incomprensibles o poseer poco beneficio al obtener algún significado que no es de utilidad para la organización. Aunque pareciera que el Data Mining (DM) ofrece resultados “mágicos”, se requiere de la previa apreciación del problema, para establecer la estructura correcta que permita recolectar toda la información relevante a la problemática, de acuerdo a los costos que el agente está dispuesto a pagar por la infraestructura de la información y así mismo, el costo de la toma de decisiones.

Por Ejemplo: Si se requieren conocer la probabilidad de fallas del equipo de producción; con tal de llevar a cabo mantenimiento preventivo, la empresa debe de requerir los recursos suficientes para realizar el mantenimiento con la frecuencia indicada. De lo contrario, si solo se conocieran los resultados obtenidos del DM, solo se conocerán las características, las frecuencias de los fallos; pero no se podría hacer nada al respecto. De esta forma: Aunque se conozca la solución del problema, se deben de colocar recursos necesarios para llevar a cabo la solución; es decir, se debe de trasladar el conocimiento a la toma de decisiones.

Algunos beneficios esperados y objetivos buscados al emplear Data Mining en las empresas son las siguientes: Una Inmejorable toma de decisiones, perfeccionamiento en la colaboración interna, minimización en los costos o residuos, optimizar la productividad y la calidad de los recursos humanos; ser capaz de reaccionar velozmente a los cambios del mercado, mejorar la atención al cliente y su retención. (Capgemini Consulting, 2017)

La principal dificultad de incorporar a las empresas DM, es el costo de la infraestructura necesaria: servidores, softwares, procesos de recolección de información y personal calificado; así también, las complicaciones que se encuentran en la estructura de la transmisión de información dentro de la empresa, De acuerdo con la encuesta realizada por Capgemini en el 2014; solamente el 27% de los ejecutivos encuestados (de distintos sectores: Financiero, manufactura, farmacéutica, ventas, etc.) reportaron como exitoso la implementación del DM; Indican que los principales desafíos para su implementación es: el integrar la información en un espacio donde pueda ser empleada en su totalidad, 75% de los encuestados de Camgemini, poseen la información esparcida en las distintas áreas de la organización; por lo que el reto de implementar el análisis Big data, implica el edificar una estructura administrativa basada en la calidad de la información, con una visión empresarial "pro-data" y "pro-tecnologica"; por lo cual el cambio de la cultura organizacional es un problema para las empresas tradicionales, ya que la organización y la colaboración debe de romper las barreras entre los "pro-data" y de los que mantienen la privacidad de la información.

De acuerdo con los estudios realizados por Dresner Advisory Services (Columbus, 2017), de las industrias que adoptan las tecnologías de Big Data, son principalmente las industrias de la telecomunicación y del sector financiero; siendo la industria educativa, la que menos interés a tenido en tal tecnología.

El sector financiero, de acuerdo con la encuesta, busca emplear el análisis de Big Data, para la administración y optimización de la información interna, mantenimiento predictivo y análisis de los clientes; Desarrollando herramientas e iniciativas para crear Reportes, Dashboards y visualización de la información (reducir los costos del monitoreo); así como herramientas de "Auto-servicio", las cuales permiten al usuario (o consumidor) emplear servicios especializados a través de plataformas, tal como cotizadores, calculadoras o asistentes virtuales "Chat-bot"

con la finalidad de facilitar la contratación del producto o servicio. (es decir, reducir los costos de transacción del cliente)

Algunas instituciones financieras como el “Тинькофф Банк” (banco Tinkoff) de Rusia (Informatica, 2018). El empleo del Big Data ha favorecido al banco para ser capaz de automatizar sus procesos internos y los procesos con los cuales interactúa con el ambiente (procesos de contratación, búsqueda de clientes y proveedores); Mejorando su competitividad al reducir los costos por unidad de venta y minimizando las fricciones creadas por la gestión administrativa. Ofreciendo servicios altamente personalizados y enfocados al cliente, al emplear un ecosistema de plataformas virtuales o “apps” en diversos dispositivos que cuentan con acceso remoto.

Con ello, el Тинькофф Банк, prácticamente invita a cualquier persona que poseyera un nombre y un teléfono, a usar los servicios ofrecidos por el banco; reduciendo la barrera a la entrada del cliente por la fricción generada por el costo (en tiempo) de la tramitología de identificación y contratación; Y así mismo, el banco se protege constantemente de los clientes dañinos mediante un sistema avanzado de monitoreo de riesgos, el cual no entorpecerá o no generará alguna fricción por los procesos de vigilancia. De esta manera el banco diseñó una potente infraestructura que posee dos centros básicos en su funcionamiento: La automatización de los procesos y el monitoreo constante de los riesgos. (Ivanova, 2018)

Por las dificultades del ambiente en el que se encuentra el banco Tinkoff, la adopción de la banca digital es complicada: Siendo Rusia un lugar, en donde un gran flujo de migrantes mantiene a gran parte de la población sin algún tipo de registro e identificación oficial; donde la infraestructura física de telecomunicaciones no garantiza el acceso a todo momento ni la accesibilidad universal al internet. Y adicionalmente, los aspectos culturales de los clientes, los

cuales se encuentran acostumbrados al uso del papel (cheques, tarjetas de plástico y contratos físicos) para llevar a cabo sus transacciones.

Por lo que el banco Tinkoff, siendo un banco totalmente online cuyo core-bussines se centra en los procesos automatizados y la personalización de sus productos; la empresa emplea la tecnología disponible en los dispositivos para realizar procesos identificación por medio de mecanismos biométricos (huella digital, reconocimiento facial y de voz) al emplear las herramientas convencionales de los teléfonos celulares (como la cámara y micrófono) y así mismo se ha creado iniciativas para poder crear una versión de SIM Card de teléfono, como una identificación que pueda servir para la realización de transacciones y traspasos financieros entre clientes (P-2-P).

El banco posee también, un catálogo flexible de productos el cual le permite crear un amplio ecosistema virtual centrado en las necesidades financieras de sus clientes, quien tiene la opción de añadir o eliminar módulos a su gusto, de tal forma que los servicios ofrecidos por el banco se adhieran íntegramente al estilo de vida del cliente; convirtiendo los servicios del banco, parte sustancial y trascendental de las transacciones financieras personales de los clientes y de su modo de vida; Con el uso de DM y de la inteligencia artificial, le permite a la aplicación generar toda clase de información relevante (sugerencias o promociones en locales cercanos), en base a la hora del día, a la ubicación del cliente y la configuración de su ecosistema. (Тинькофф Банк, 2019)

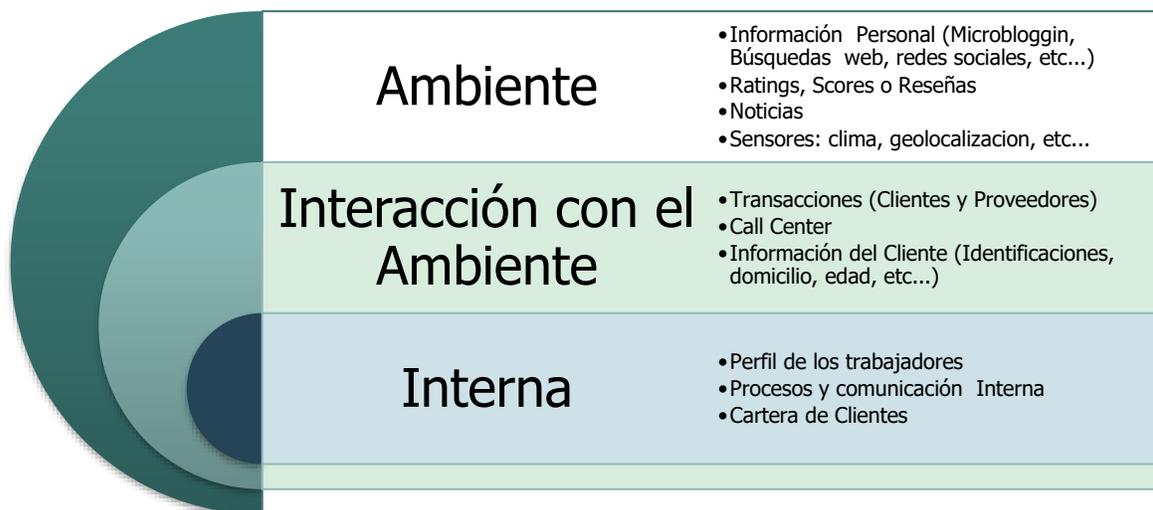
2.4.2 “Data understanding” o Entendimiento de la información.

Dentro del proceso de DM, parte fundamental es la información y la calidad que esta tenga de esta, ya que se trata de la materia prima del proceso, para ello se debe de poseer algún mecanismo con la finalidad de producir e integrar toda la información relevante que puede ser recolectada por los diversos canales que

puedan existir; Siendo necesario comprender el tipo de información que se produce y posee almacenada internamente, y la que puede ser adquirida por medio de terceros.

Como se había mencionado, uno de los atractivos principales del Big Data, es el emplear fuentes alternativas de información para llevar a cabo análisis complejos. De la cual se puede ser categorizada en tres niveles en base a la entidad observada: la información interna, la que se genera de la interacción con el ambiente y la información propia del ambiente.

Ilustración 6 Ejemplos de Información Alternativa (Niveles del sistema)



Fuente: Elaboración propia.

Parte de los beneficios del DM, no solamente se encuentra en los pronósticos que puedan ser obtenidos a partir de la información extravagante o alternativa, tal como "selfies de los clientes" o el "historial de películas vistas en Netflix"; si no que, también es factible el desarrollar proyectos de DM incluyendo toda la información que la empresa posee de sí misma, tal como la información financiera, la información sobre sus trabajadores, el flujo de la comunicación interna, etc. Al implementar una infraestructura de información de baja latencia (actualizada en tiempo real) que permita entrelazar y emplear la información disponible; permite

obtener una visión amplia sobre las operaciones internas, la cual puede ser analizada en tiempo real; permitiendo detectar ciertas anomalías en los procesos internos de la empresa, tal como cuellos de botellas que implican costos operativos y/o generación de residuos:

- Bajo la iniciativa que fue desarrollada en el 2003, el proyecto ORION (On-Road Integrated Optimization and Navigation) de la compañía de paquetería UPS; Empleando la información recolectada, de una gran cantidad de sensores que se instalarían en los vehículos de logística, los cuales transmiten señales de posicionamiento global (GPS); Le permite analizar el desempeño de sus conductores y genera las mejores rutas posibles. Por lo que, en el 2013, la compañía UPS, disminuyó el costo de combustible en 8 galones, alrededor de \$50 millones de dólares en él mismo año, 2013. (Capgemini Consulting, 2014) Al mismo tiempo que se reducían los tiempos de entrega.

En lo que respecta a la información recopilada sobre las interacciones que la organización lleva a cabo con el entorno; es decir, las negociaciones e interacciones que la organización realiza con los demás agentes que habitan en el mismo sistema para llevar a cabo sus operaciones: Los clientes, los proveedores y los competidores.

Algunos casos, en el que se emplea esta clase de información:

La banca en línea, multicanal; posee información del cliente por la interacción que posee este mismo con la empresa: navegación web e información del dispositivo móvil (del portal o aplicación de la misma empresa), los cajeros, grabaciones de servicio Call-Center, expediente y características socioeconómicas del cliente. Mediante los procedimientos DM, es posible recopilar la información que se encuentra dispersa en los distintos canales de contacto con los clientes,

creando expedientes que pueden ser analizados, para la aplicación de ofertas personales. (O'Brien, 2017)

Un ejemplo de su aplicación, es el caso del banco Bank of America (Fletcher, 2014) el cual desarrolla el programa BankAmeriDeals, el cual recompensa al cliente con promociones Cash-back, dichas recompensas están basadas a la calidad del cliente, recompensando su lealtad. Con la finalidad de no generar anuncios no deseados (o SPAM), se generan recomendaciones personalizadas en base a las preferencias del cliente y su entorno geográfico (la cercanía a los establecimientos).

Finalmente, el monitorear las características del ambiente también puede ser empleado dentro del análisis DM; este tipo de información, le permite al agente mejorar sus capacidades adaptativas, debido a que permita al agente ajustar sus conductas rápidamente de acuerdo a las condiciones del contexto económico

- La plataforma en línea Airbnb (Airbnb, 2018); ofrece dentro de sus funciones una herramienta llamada "Smart Pricing", con la cual la plataforma le sugiere al anfitrión un precio por los servicios que ofrece; para ello estima mediante modelos DM (un modelo Boost Gradient Machine) curvas de demanda, con las cual le permite generar un pricing dinámico, el cual se ajusta a las características cualitativas del anfitrión, la estacionalidad y la demanda (la probabilidad de registro o contratación).

Para crear el modelo de probabilidad de registro; se toman como datos de la oferta del anfitrión (tipo de cuarto, capacidad de personas, reseñas y su historial) de la temporalidad (días del calendario y época del año) y las características de la competencia (número de anfitriones en la locación visitas y numero de búsquedas relacionadas).

Tabla 2; Cuadro Resumen, de los casos mencionados. Elaboración propia.

| Empresa | Tipo de Datos | Resultados Obtenidos con DM | Decisión Tomada |
|-----------------|-----------------------------|--|---|
| UPS | Interna | Medir con gran precisión, la eficiencia del traslado de los paquetes. | Generar rutas optimas, que permitan disminuir el costo en combustible y del tiempo de entrega. |
| BANK OF AMERICA | Interacción con el ambiente | Recopilación de la información esparcida del cliente y crear expedientes individuales. | Ofrecer recompensas personalizadas a los clientes, con tal de mantener una estrecha y longeva relación con los clientes meritorios. |
| AIRBNB | Ambiente | Crear modelos para estimar la posible demanda de los servicios ofrecidos por los anfitriones | Sugiere precios, para facilitar la gestión del anfitrión y ser una empresa atractiva para sus proveedores y clientes; en base a las características del ambiente. |

Con el uso de la información alternativa (o no-tradicional), permite al sector financiero, generar modelos con mayor precisión al incorporar y completar tales modelos con nuevas fuentes de información, lo cual le permite conocer a profundidad las características sobre la entidad observada; empleando la información alternativa es posible cubrir la población habitualmente excluida por la escasa información que se tiene del cliente, tal como las pequeñas y medianas empresas (PyMEs), migrantes, indocumentados, estudiantes y gente que no posee un score crediticio. Algunas fuentes de información alternativa pueden ser las siguientes: (Oliver Wayman, 2017)

- Consumo de Servicios (Gas, Agua, Electricidad, Telefonía móvil e internet)
- Valor de los activos que posee y compras realizadas (“e-commerce”)
- Información financiera (Ingresos, gastos balance promedio)
- Patrones en la interacción social (Red social)
- Test psicométricos
- Información Geo-espacial

Al emplear este tipo de información, se pueden crear alguna especie de “score” que no necesariamente prediga a la perfección la probabilidad de incumplimiento, sino que simplemente ayude a categorizar y separar los distintos clientes por el tipo riesgo que porta; permitiendo medir aquellas entidades que no poseen un historial crediticio y eliminar algunas barreras de la inserción financiera.

2.4.3 Visualización de la información.

Los avances de las capacidades analíticas de las computadoras y la automatización de la integración de la información, no servirían de gran cosa si no se poseyera un sistema que permita comunicar de forma legible la información contenida. Parte importante del proceso de DM radica en la visualización de la información, con el fin de poder interpretar y conocer las características de la información, con la finalidad de poder emplear las distintas variables dentro del modelo, así como la detección de posibles anomalías dentro de la base de datos.

Se puede entender como “Visualización de los datos”, como la representación gráfica, en donde se emplea un conjunto de elementos visuales, los cuales son creados computacionalmente al emplear una serie de datos que puede ser representado en un plano gráfico; Los principales objetivos de visualizar son:

- Comprender las relaciones entre las variables y la comprensión de ideas abstractas
- Identificación, exploración de la información
- Monitoreo de eventos y operaciones
- Comunicación de ideas mediante el empleo de elementos estéticos.
- Entretener y contar historias

Al incorporar distintas fuentes de información, con variables de distintas características; la visualización de los elementos, se convierte prácticamente en un proceso artístico, debido a que su uso implica el uso de colores, formas y tamaños;

con la finalidad de presentar un sin número de datos en un plano limitado de espacio, en donde se deben de ser acomodados distintos tipos de variables:

- **Ordinal Categórica;** es una variable categórica que posee un valor ordinal lógico o numérico (frio, tibio, caliente); generalmente esta denotada por poseer valores numéricos enteros.
- **Nominal Categórica;** describe características de la información, usualmente está contenida en formato "character", las cuales no poseen un orden específico; Ej. Rojo, Azul, Canino, Etc...
- **Cuantitativa continua;** constituido por valores reales, Se describe información, como cantidades, distancias, temperatura y demás.
- **Cuantitativa Discreta;** constituido por valores naturales, describe información que posee características contables, como las cantidades en número de cualquier objeto.

Al poseer un limitado espacio de presentación, los algoritmos diseñados para la visualización de la información, usualmente poseen las características de ser interactivas y actualizadas en tiempo real. Al poseer una imagen dinámica, en esta pueden ser incluidos ajustes para mostrar la información que el usuario desea visualizar.

En la ilustración 7, Productos como "GAPMINDER", permite crear visualizaciones dinámicas (que permite observar su evolución a través del tiempo, lo cual deja una estela de puntos), permitiendo observar la relación de hasta de 4 variables al emplear los ejes, colores y tamaño (adicionalmente se podría incluir el empleo de formas para incluir una quinta variable categórica, por ejemplo: el continente al que pertenece cada país.); la plataforma permite escoger y cargar dinámicamente un conjunto de variables disponibles de cada país.

Ilustración 7, Relación tener el crédito doméstico, el ingreso, la población y la formación de capital (en dólares constantes) y su evolución a través del tiempo



Elaboración propia, Empleando GAPMINDER, con datos del Banco Mundial;

Interpretando la ilustración 7, se puede observar que a medida que existe dentro de la economía una mayor cantidad de recursos, es posible crear un ambiente propicio para otorgar créditos “sanos” y coadyuvar a la formación del capital. Existen tres curiosos casos: China, Japón y EUA, cuya pendiente a lo largo del tiempo es más inclinada al cenit. Lo cual al simplemente observar la gráfica se genera la siguiente duda: ¿Qué factores han desarrollado tan abruptamente el mercado de crédito en los mencionados países?

Dentro de la visualización se incluyen otros países peculiares, estadísticamente conocidos como Out-liers, sin embargo, su existencia no implica la autoeliminación dentro de un modelo, siendo que estos son posibles casos de estudio ya que se encuentran fuera del comportamiento normal de los demás

países. Estos Out-liers son Arabia Saudita y Bruselas, quienes poseen un alto nivel de ingresos, pero poseen un mercado de crédito muy pequeño.

Ante los volúmenes de la información, se requiere de herramientas más sofisticadas, diseñadas para poder procesar e integrar la información necesaria para la visualización; para la creación de estos algoritmos, existen programas de código abierto tales como R, Python o Weka. Con este tipo de programas es posible desarrollar herramientas las cuales no poseen ningún obstáculo en cuanto a costo del software; sino que su uso solamente se limita a las capacidades técnicas del usuario.

Cuadro 4, Ejemplos de software de código abierto para el análisis y visualización de la información. Elaboración Propia



Weka "Waikato Environment for Knowledge Analysis"

- Proyecto que fue desarrollado en la universidad pública de Waikato, ubicada en Nueva Zelanda, forma parte de una iniciativa pública de innovación, con la finalidad de brindar al sector privado, herramientas e innovaciones que ayude al desarrollo del país.
- Weka, posee una colección de algoritmos precargados, escritos en lenguaje "Java" para realizar "machine learning"; por medio de una interfaz amigable al usuario, permite analizar grandes cuerpos de información sin ninguna complicación técnica-computacional, ya que el programa está dedicado principalmente a usuarios industriales.



Python

- Python, fue desarrollado originalmente en 1980 por Guido Van Rossum, como un hobby personal, al intentar crear un lenguaje intérprete alternativo, orientada a objetos tal como Java, ABC, C++ o Ruby.
- Dentro de sus principales características se encuentra en la versatilidad y flexibilidad de sus aplicaciones; entre sus distintos usos, pueden ser desarrollados en el lenguaje Python: Aplicaciones web, computación numérica y científica, herramientas GUI (interfaz gráfica de usuario), desarrollo de videojuegos y evidentemente: Algoritmos de Data Mining y Machine Learning.



R

- Lenguaje denominado como "R", es parte del proyecto del sistema operativo GNU (se pronuncia ñu)
- "R", es un programa desarrollado por laboratorios Bell, diseñado específicamente para la estadística y el manejo de la información numérica. Provee gran cantidad de herramientas para el análisis estadística lineal y no lineal; Incluyendo, también, bibliotecas de "Regular Expresion" que permite diseñar algoritmos para el análisis de textos (Text Mining)

2.4.4 Análisis y Modelación de la información

Debido a la existencia de la asimetría de la información, los intermediarios financieros como los bancos, emplean el uso de "Score Cards" o "Scoring", para generar las expectativas sobre los clientes existentes o los prospectos, vaciando toda la información que se posee del cliente en uno o varios índices que le permitan al intermediario identificar el perfil del cliente; permitiéndole administrar los riesgos de contrapartida de acuerdo a las características "esperadas" del cliente: creando políticas de rechazo, límites en los créditos y condiciones que permitan disminuir los riesgos tomados por la compañía. (Bessis, 2002).

La variable "Scoring", es una variable arbitraria, ya que se encuentra determinada por factores subjetivos de acuerdo a las políticas y aceptación de riesgo del intermediario; tal como el establecer lo que es "bueno" o "malo", se describe bajo una variable categórica la consecuencia observable (o estado final) de una entidad, dicha consecuencia se encuentra sujeta a un conjunto de causas no observables por él intermediario, por ejemplo: el banco puede poseer un registro de sus clientes "buenos" y "malos", de los cuales puede observar solamente el cumplimiento de los pagos y de los rendimientos obtenidos por cada cliente; pero no es posible determinar las causas específicas de su cumplimiento o morosidad.

En 1941, David Durand desarrolla académicamente un método para poder cuantificar la credibilidad de los clientes, en base a sus características principales. Tal como sus ingresos, sus activos (como cuentas bancarias, tipo de propiedad de su hogar), de su estabilidad laboral y de su ocupación; analiza también las características del crédito otorgado: Monto otorgado, periodo, frecuencia del pago, uso de colateral y tasa de interés.

Bajo el empleo del modelo de análisis lineal discriminante (LDA), se determinan los factores principales, con la finalidad de poder clasificar y separar estadísticamente las entidades observadas a partir del promedio sobre las características observadas de cada clase. Sin embargo, el modelo LDA no posee la suficiente eficiencia al clasificar los factores, probablemente a causa de que las variables observadas, no describen las cualidades personales que los clientes poseen.

Con el uso del Big Data, el empleo de información alternativa; es posible complementar el "Scoring", ya que, al incorporar nuevos datos y características, es posible describir cualitativamente a los clientes, tal como: su comportamiento, su atención al riesgo, administración de sus recursos. Además de poder realizar validaciones cruzadas y detectar comportamientos congruentes con la información que el cliente revela en la solicitud del crédito. (Jagtiani & Lemieux, 2018).

Al incorporar diferentes y heterogéneos tipos de datos, se requieren de distintos tipos de modelos y métodos que mejor se adapten a las características de la información recolectada, con el fin de determinar efectivamente las variables o componentes que poseen mayor capacidad de generalización que permitan describir correctamente los datos que no pueden ser observados (como las características del cliente).

Con algoritmos denominados como "Machine Learning", los modelos son ajustados automáticamente a las características del continuo flujo de información que es vertida, creando modelos que mejoran de acuerdo con su experiencia, es decir, mejoran a medida que procesan cada vez más información, y se ajusta ante cambios en los patrones que contiene la información.

Def 6. Algoritmo, es la sucesión de ordenes lógicas que es empleado por el ordenador para completar alguna actividad designada por un usuario.

Dentro de los algoritmos Machine Learning, se encuentra el modelo de "Sistema de Redes Neuronales" (Neural Network o NNet). Este procedimiento implica un nivel de aprendizaje profundo conocido como "Deep Learning", a través del procedimiento conocido como "Backpropagation", con el cual el algoritmo, de forma automática ajusta el peso o el coeficiente de los perceptrones, mejorando su efectividad al evaluar constantemente los nodos que componen el modelo, corrigiendo sus errores y ponderando los nodos con mejor capacidad descriptiva.

Arthur Samuel (1959), pionero en programación de Inteligencia Artificial; Hizo popular el concepto "Machine Learning", al exponer sus avances en la programación de un jugador virtual, el cual sería capaz de realizar el mejor movimiento sin la necesidad de programar manualmente la mejor respuesta para cada evento presentado en el tablero. Mediante un proceso iterativo de "Ensayo y Error", el programa almacena los resultados obtenidos de cada acción, evitando los resultados negativos y repitiendo los positivos, para ello es necesario especificar al programa los premios y los castigos de sus acciones; a este esquema de aprendizaje, usualmente se le conoce como Inteligencia Artificial (IA), a través de esta inteligencia, es posible que el programa modifique su comportamiento mediante la observación de sus resultados y las condiciones del escenario. (De tal modo que se le es otorgada al computador, cierta capacidad cognitiva freudiana)

La idea principal de Arthur Samuel, era el poder crear un jugador virtual capaz de anticipar los movimientos del jugador humano y evaluar los resultados de los posibles escenarios al realizar determinados movimientos, para ello, el programa emplea la grabación de sesiones en juegos de prueba, conocidos como book games, llamando al proceso de aprendizaje como book learning. El programa al analizar los resultados de estas rutinas de entrenamiento, era capaz de implementar tal conocimiento aprendido contra algún oponente humano. De tal modo, que mientras más jugaba más aprendía (machine learning).

En los procesos de modelación y “aprendizaje” en el DM, los modelos poseen dos tipos de clasificación, de acuerdo a las características del escenario en donde se desarrollan y el tipo de problemática que se desea resolver (Talabis, McPherson, Miyamoto, & Martin, 2015):

- **Supervised Machine Learning**, Implica la creación de un modelo del cual se poseen los datos de la variable objetivo o variable a predecir. La información que es obtenida es dividida en dos partes: una muestra la cual es empleada para calcular los coeficientes del modelo (denominada como muestra “train”) y otra muestra con la que son generados los outputs (denominada como muestra “test”) esto, con la finalidad de evaluar la efectividad del modelo en base a los errores obtenidos, al comparar las observaciones estimados por el modelo con los datos reales de la muestra “test”.
 - Las técnicas de este tipo de ML son modelos de características predictivas:
 - **Modelos de Regresión**, el cual es empleado para predecir variables cuantitativas o numéricas, tal como: la Regresión Lineal.
 - **Modelos de Clasificación**, empleado para poder predecir variables de clasificación y binomiales, tal como: Regresión Logística, LDA, “K-más cercano”, Arboles de decisión, Sistema de Redes Neuronales y Support Vector Machines (SVM)
- **Supervised Machine Learning**, Implica la creación de un modelo, el cual sea capaz de encontrar una estructura o patrones dentro de la información al poseer información sin una variable objetivo previamente definida; de esta forma el modelo clasifica y agrupa los registros por medio de sus elementos comunes. De tal modo que, al poseer parámetros de clasificación

tras obtener los resultados del modelo, es posible clasificar los nuevos registros en base a las características determinadas por el modelo.

- Las técnicas de este tipo de ML son:

- **Modelos de Clustering**, permite la agrupación de los datos, en distintas clasificaciones y grupos con variables categóricas no-ordinarias (por ejemplo, grupo "A","B","C", etc...). Tal como los modelos: "Clusterización K-Media", "Mean Shift Clustering".

Al poseer gran cantidad de observaciones y de variables disponibles, se pueden crear un modelo que posea dentro de este, más de un modelo que sea capaz de realizar predicciones o clasificaciones por medio del consenso, de los resultados obtenidos de los distintos modelos. A este tipo de modelos se les conoce como "modelos en conjunto"; los cuales generan submuestras de la muestra "train", con la finalidad de crear "n" modelos denominados como "weak learners", los cuales mediante un proceso de ajuste (por moda, media o media ponderada), se obtiene resultados habitualmente más robustos que el emplear un solo modelo descriptivo. Tal como los modelos Bootstrap, el AdaBoost o el bosque de árbol de decisiones.

2.5 Dificultades con el uso del Big Data

Pese a los beneficios que pueden ser obtenidos al emplear el análisis del Big Data o Data Mining; al emplear este tipo de información se hallan distintas dificultades, tales como las complicaciones en su uso y la calidad de la información; así mismo, existen posibles características oportunistas emplear los resultados obtenidos y creación de riesgos debido a la custodia de la información sensible de los clientes.

De la información, tanto la calidad como el costo y la accesibilidad son factores que obstaculizan el empleo del Big Data; al poseer un gran número de

fuentes, la información obtenida puede estar contaminada, ser reiterativa, poco confiable o simplemente erróneamente registrada. (Cai & Zhu, 2015) Al poseer una gran cantidad de registros, los errores contenidos en la base de datos probablemente pasen de desapercibidos o sean difícilmente encontrados, generando modelos con resultados erróneos, que no permiten captar la realidad.

Entre una de las críticas en la que se abunda sobre el uso del DM, es el empleo de los datos provenientes de las plataformas de Redes sociales; (Computing Community Consortium, 2012) debido a su heterogeneidad de la información, es complicado integrar las distintas plataformas sociales y así mismo recolectar los registros indicados para la investigación¹⁴, siendo que en dichos ecosistemas virtuales se encuentra un inmenso ruido el cual es producido por cuentas falsas, clones de cuentas, re-posteos publicitarios y bots; las cuales generan señales erróneas en la estadística al generar información que no posee ninguna veracidad y no presenta ningún estándar de calidad.

La calidad de la información puede ser caracterizada bajo las siguientes dimensiones, de acuerdo con Cai & Zhu (2015):

- **Disponibilidad**, Se refiere a la barrera a la adquisición de la información, siendo limitada por poseer un costo elevado al integrar la información, así como su latencia y la secrecía de la información. La información al hacerse obsoleta en el tiempo, esta pierde su capacidad de generar valor al portador

¹⁴ Dentro de la experiencia al elaborar esta tesis, se intentó realizar un modelo basado en "Sentiment Analysis" con información de Twitter para observar e intentar mesurar las experiencias que los clientes poseían con sus bancos, el proceso no es muy complicado ya que las herramientas se encuentran disponibles en el internet. Sin embargo, la información recopilada por los algoritmos diseñados, no proporcionaron la data suficiente para evaluar dicho análisis; debido a que la información fue contaminada con innumerables tweets que no poseían ninguna trascendencia al tema observado ni a las entidades estudiadas. El problema fundamental del experimento se encontró en la simplicidad del algoritmo, ya que no poseían los filtros necesarios en la recopilación de los datos y la disponibilidad temporal de la información, ya que solamente se podían obtener datos de 07 días anteriores a la fecha.

al existir "n" competidores en mejores circunstancias en el análisis de la misma información.

- **Aplicable**, Se refiere a la certificación de la fuente de información, es decir, la información debe de ser capturada por fuentes e instrumentos que generen información verídica, tal como estadísticas generadas por instituciones de investigación o sensores perfectamente calibrados. así también, son requeridos estándares sobre la "información de la información" con el fin para poder interpretar los valores capturados dentro de los registros (es decir, la información debe de poseer meta-data que la describa correctamente).
- **Utilidad**, los datos recolectados debe de estar apegada a la realidad, representar las propiedades reales de la entidad observada, los datos deben de poseer una estructura homogénea y constante para poder ser comparada en distintas fases del tiempo.
- **Relevancia**, los datos deben de poseer la información importante al tema a investigar, lo cual permite capturar las peculiaridades del grupo objetivo u observado que desean ser estudiados; además debe de ser capaz de explicar las hipótesis que fueron alineadas con el objetivo de la información.
- **Presentación**, La información deberá de ser fácil de leer e interpretar, así mismo, debe de ser posible de comprobar la puntualidad y la medida empleada, al representar la realidad observada.

Pese a que usualmente se jacta, de que es posible emplear complejos y heterogéneos grupos de información con el Data mining; Usualmente la información del ambiente permanece privada por la utilidad que le brinda a su portador, y difícilmente es accesible. Y así mismo, por las características heterogéneas de la información, se poseen dificultades al integrar la información

que se encuentra en los distintos niveles del sistema (nodo, grupo primario y sistema o interna, interacción con el ambiente y el ambiente en sí).

La nomenclatura y la utilidad de la información, son pertenecientes a la entidad quien la genera, es decir: La entidad que genera la información posee un objetivo específico al producir su propia información (por ejemplo, una compañía comercial) que puede no alinearse a los objetivos perseguidos por el consumidor de la información (por ejemplo, un académico). De tal modo que, el desear observar las condiciones exactas en las que el ambiente se encuentra; la observación puede mantenerse en proceso intrincado aun con la tecnología del Big Data, por lo que la información puede no encontrarse completa y no tener la relevancia suficiente sobre el objeto de estudio.

Para poder observar plenamente el ambiente y sus peculiaridades, la información debería de ser totalmente publica o accesible, la estructura de la información requeriría de una estandarización universal en los formatos de los distintos instrumentos de medición que producen la información, de tal manera que pueda ser relacionada o referenciada a las distintas entidades observadas (por ejemplo relacionar los datos de: un habitante – una casa - un vehículo) con muestras recolectadas de distintas fuentes y creadas con en distintos instrumentos o dispositivos.

El acceso limitado al internet y las restricciones que pueden tener ciertos grupos sociales; es otra de las características que genera problemas en la representatividad de la población. De acuerdo con los datos del Banco Mundial, en el 2017 solamente el 50% de la población posee algún medio para acceder al internet. Gran parte de los sesgos se encuentra

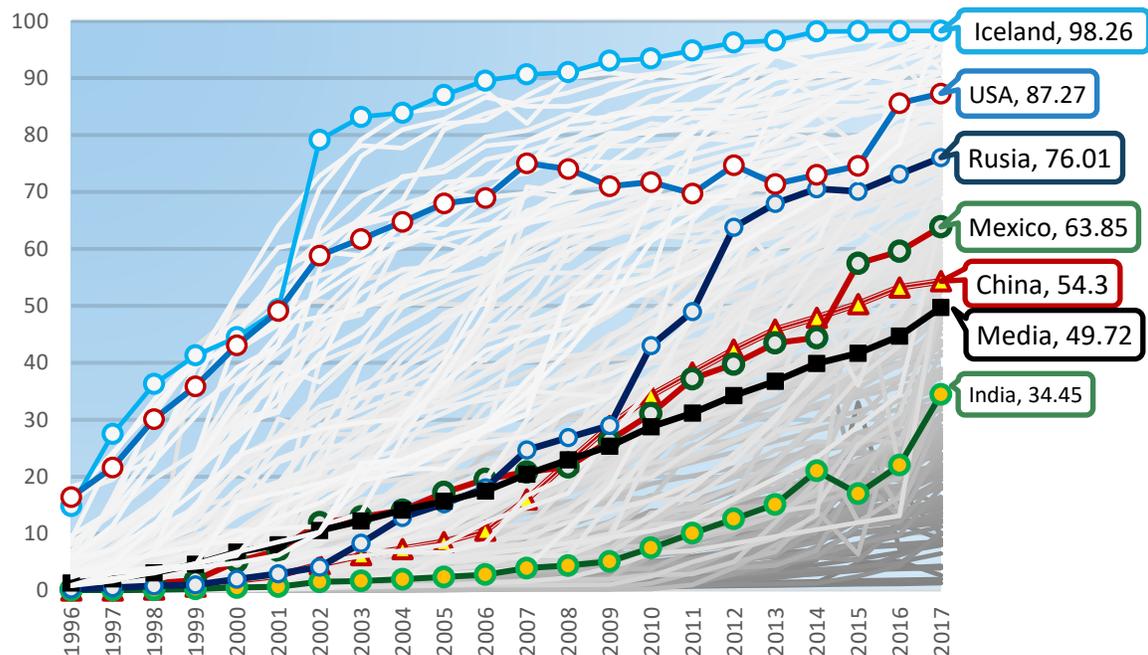
| <i>Uso de Redes Sociales en la población de USA (2018)</i> | |
|---|-----|
| <i>18 a 29 Años</i> | 88% |
| <i>30 a 49 Años</i> | 78% |
| <i>50 a 64</i> | 64% |
| <i>Mas de 65 Años</i> | 37% |

Tabla 3, Uso de Redes Sociales por grupos de Edad, Fuente: "Social Media Use in 2018" Pew Research Center

en países de baja infraestructura y recursos, tal como los países africanos como Somalia, Nigeria, Madagascar, Togo, Malawi, etc... en los cuales, solo un puñado de personas posee acceso al internet (3%-10% de la población)

En países desarrollados, como EUA el 87% de la población posee algún tipo de acceso al internet, De acuerdo con las encuestas realizadas por Pew Research Center, la masificación del internet se esparce incluso a las edades más avanzadas. En el 2012, el uso del internet y de las redes sociales, se encontraba limitado a los usuarios de entre 18 a 29 años. Por lo que se le asocia a este pequeño grupo de internautas, una muestra no representativa de la población total, ya que se excluye la población de edades más avanzadas; Así mismo, existen en el internet ciertos incentivos que puede hacer que un internauta modifique su conducta habitual lo cual generaría gran cantidad de ruido; por anuncios y recompensas en línea "Paid-to-do" que recompensan al internauta por realizar ciertas acciones.

Gráfica 2, Porcentaje de la población con acceso a internet (Todos los países, registrados en el banco mundial)



FUENTE: Elaboración Propia, con datos del Banco Mundial; Porcentaje de la población que usa internet, sea por dispositivos móviles, computadora, asistente personal o televisor

2.5.1 Mala praxis; Oportunismo e invasión a la privacidad

De la misma manera que existe el riesgo moral por el lado del cliente, por el lado del intermediario existe la misma característica de información asimétrica. Mediante el empleo de Scoring tipo "Black Box", el cliente no posee ninguna certeza de factores son empleadas en su score y de la misma manera no puede saber si la información es correcta o si se le está juzgando bajo un criterio "justo".

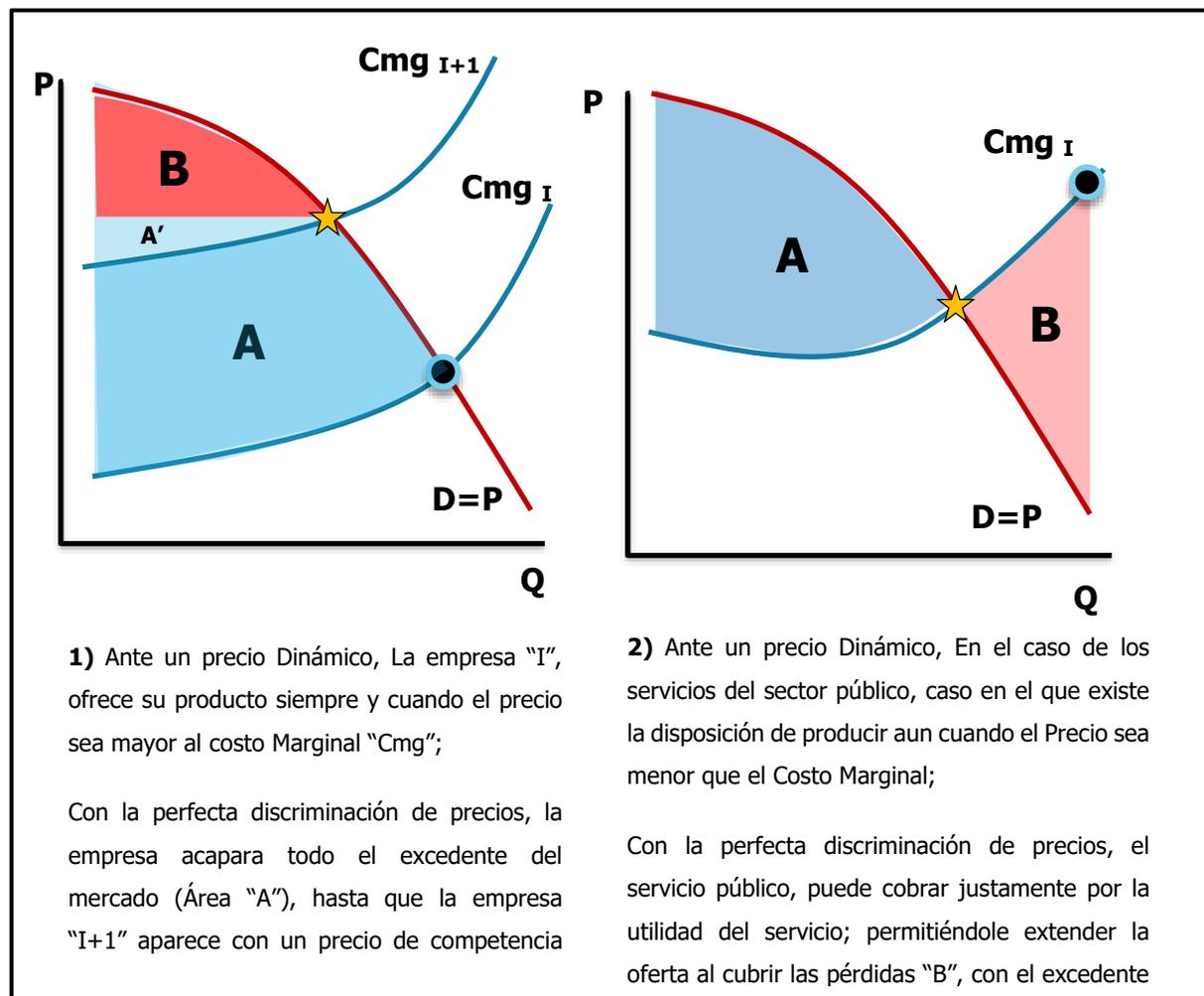
Pese a que existe un factor de aparente imparcialidad humana, tal como el uso de Inteligencia Artificial o modelos de Clusterign; existirá siempre un factor arbitrario, ya que el diseño y los estímulos de la IA, "aprenden" y adaptan su comportamiento basado en un esquema de recompensas, requiriendo de un medio de traducción o aprendizaje que involucre la ponderación humana, del cual puede basarse en la selección de las variables empleadas y de la interpretación de los resultados que se encuentran; siendo esto amoldados al desarrollar los objetivos del modelo. Siendo que la personalidad de la IA, no nace de la mera casualidad (Axioma 2 de St. Tomas de Aquinas), esta inteligencia se encuentra diseñado bajo los objetivos particulares al arquitecto del sistema informático.

Al asumir que la firma obtiene el conocimiento pleno sobre el cliente, mediante el empleo del Big data, este puede plantear un precio dinámico ajustado a cada uno de sus clientes (como en el caso de AIRBNB). De esta forma, ante el comportamiento oportunista el banco puede fijar un precio de venta que es ajustado automáticamente al precio de reserva de cada individuo, reduciendo a cero el excedente del consumidor y por lo tanto reduciendo el bienestar o capacidad adquisitiva de los clientes. (Reed Shiller, 2014)

Bajo la perspectiva, la utilidad de un bien (y por lo tanto el precio de reserva) está determinada por la expectativa que el consumidor tiene de consumir el bien "X": $E_t(X|\kappa) = U_{t_{i+1}}$; Ante la alteración de esta expectativa, la cual puede ser

causada por los estímulos generados por el ambiente (Bechara, 2005) generando la necesidad o urgencia de consumir el bien en el momento actual "t0"; tales estímulos pueden darse por condiciones pavlovianas inducidas por los anuncios, la acompañamiento, el clima, las tendencias sociales e incluso estados mentales o facultades mentales características del individuo (Ej. Lesiones, tal como en la regiones orbito prefrontal).

Figura 1, Comportamiento del mercado, con perfecta discriminación de precio.



FUENTE: Elaboración Propia.

Suponiendo, que la tecnología del análisis del Big Data, puede conocer perfectamente dichos factores, y por lo tanto conocer las vulnerabilidades en tiempo real; les permitiría a las firmas, obtener el suficiente conocimiento para determinar e identificar a consumidores vulnerables, maximizando los ingresos por medio de tácticas comerciales indebidas al ofrecer los productos sobrevalorados en los momentos en el que el cliente estaría dispuesto a pagar tal sobreprecio, tal como en los momentos de necesidad o urgencia (Hurley & Adebayo, 2017) o cuando un bien se encuentra en "tendencia" y por lo tanto en un precio elevado).

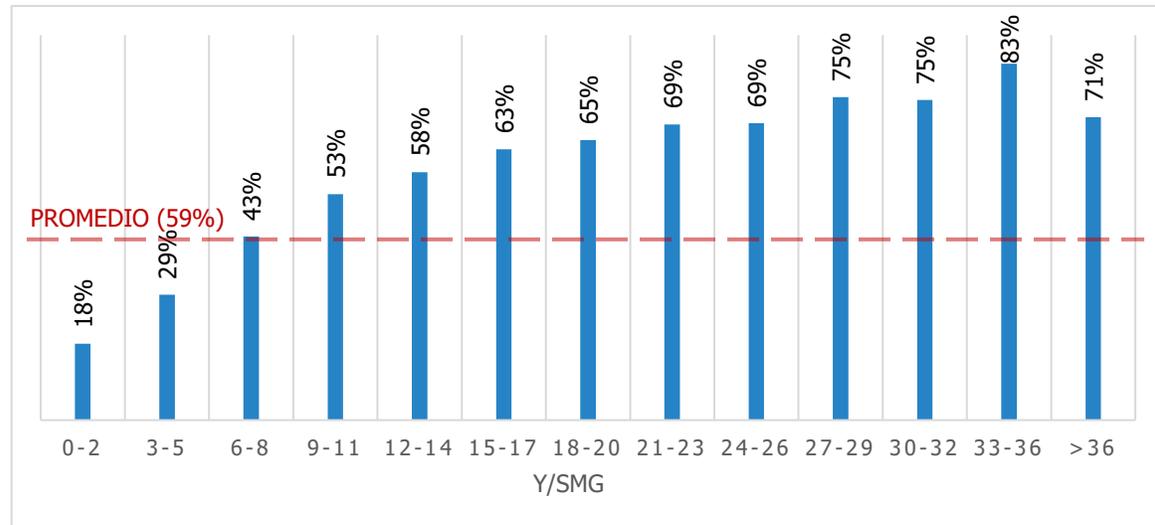
Por el lado amable, al asumir que el precio de reserva está sujeto solamente al nivel de ingresos que el consumidor percibe, y bajo la misma discriminación de precios: sería posible crear una estructura de precios que permitiera ampliar la oferta de los productos (por debajo de su costo marginal) o establecer un mecanismo de precios equitativo, al ejercer un precio dinámico en función al nivel de ingresos que posee cada consumidor.

Con la falta de regulación o transparencia en los mecanismos de discriminación, el uso del Big data en el sistema financiero presenta ser tanto una oportunidad para los grupos de consumidores de bajos ingresos, como una amenaza para el público en general, al ser objetivo de créditos accesibles con precios "injustos". En el año 2017, en México, existe un amplio grupo sin acceso a servicios financieros; entre los grupos de consumidores que poseen ingresos menores a 14 Salarios Mínimos Generales (SMG); solamente el 32% de este grupo posee acceso algún tipo de servicio financiero.

De la misma manera; se puede observar un pequeño sesgo dentro de la distribución de los ingresos y la actividad financiera que poseen los consumidores; en la gráfica 3, se puede observar que en valores absolutos el grupo de individuos entre 6 y 8 SMGs, son quienes realizan una mayor cantidad de movimientos financieros (actividad financiera), por motivo de pagos a tarjetas de crédito y

deuda, y el uso de préstamos y depósitos a cuentas bancarias. Así mismo, se puede observar que este grupo en específico posee la mayor masa de ahorro disponible en el País.

Gráfica 3, (%) POBLACION CON SERVICIOS FINANCIEROS.

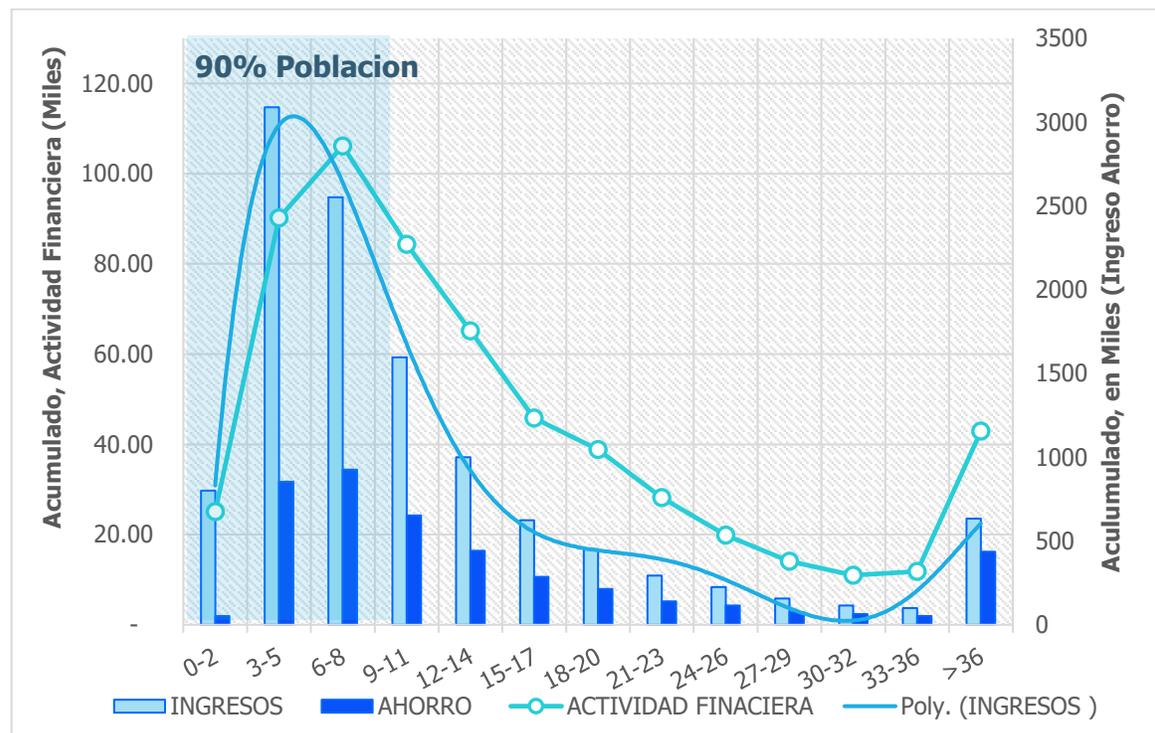


Elaboración Propia, Fuente: Encuesta Nacional de Ingresos y Gastos de los hogares INEGI 2017; agrupados por SMG (= \$7,952.40); el uso de servicios se determinó al contar la población que posee gastos financieros de deuda o pago a tarjeta, así como depósitos a cuentas de ahorro.

Dentro de la gráfica 4 se puede observar, la suma de los ingresos y ahorros de los grupos formados por la característica de su ingreso (SMG, Salario Medio Gneral); la mayor parte de la población se encuentra clusterizada en el grupo con ingresos de entre 3 y 5 SMG, siendo el grupo predominante en México, y así mismo, en este clúster se centra gran parte de los recursos empleados en la economía; dentro de los grupos de 0 a 5 SMG, el 64% de la muestra encuestada y de 6 a 8 SMG el 26.81%, (componiendo estos rangos el 90% de la población total de México). Para el agregado de estos grupos de 0 a 8 SMG; poseen en promedio un gasto en productos financieros (depósitos y pagos de préstamos) del 5% en Promedio por hogar, con respecto al nivel de su ingreso (Y/SMG). Los grupos de 20 SMGs, poseen en promedio 16%.

El sesgo en términos relativos de la actividad financiera de cada grupo, puede observarse gráficamente, al ver la disparidad entre las curvas del nivel de actividad financiera de cada grupo y la curva suavizada del nivel de ingresos denotado en la gráfica 4 como: "Poly. (Ingresos)".

Gráfica 4, INGRESOS, AHORRO Y ACTIVIDAD FINANCIERA POR AGRUPADOS EN SMG



Elaboración Propia, Fuente: Encuesta Nacional de Ingresos y Gastos de los hogares INEGI 2017.

Por medio de la innovación financiera, con el empleo del Big Data, debe de ser factible el diseñar productos especializados, para cada subgrupo de acuerdo a sus necesidades y las características de sus hábitos (tal como en el caso del banco Tinkoff). Con lo cual permitiera cubrir, los sesgos que existen en la inclusión financiera, al reducir los costos de transacción y aumentando la empleabilidad de los servicios financieros en la vida diaria de los clientes de distintos niveles

económicos. Pero su empleabilidad de dichos subgrupos dependerá de los beneficios que puedan aportar los productos y servicios que el banco ofrece.

Otra de las características negativas que se señalan con el uso del Big Data, es el enfoque en la privacidad tanto por el lado de la información obtenida, como del conocimiento producido; ya que son expuestos detalles sensibles de la vida personal del consumidor. (Xu, JIan, Wang, Yuan, & Ren, 2014) y de la misma manera, son cuestionados los métodos por los cuales la información es recolectada, como el uso del Web Mining o dispositivos IoT (como los instrumentos de medición de los teléfonos móviles, tal como el acceder al GPS), que permiten recolectar toda clase de información con la autorización del usuario; pero en ocasiones sin el consentimiento real del usuario, por el limitado conocimiento técnico que el consumidor puede poseer y por no conocer exactamente cuál es la información que está siendo compartida con las compañías.

Para que un individuo desee revelar información privada o sensible, este debe de poseer alguna especie de recompensa¹⁵ o pago por su exposición. En el caso del financiamiento, el consumidor o empresa posee un incentivo de revelar parte de su información con tal de poder obtener de ello una "buena puntuación" y, por ello, un financiamiento con menor costo. De la misma manera los individuos poseen incentivo a ocultar la información por su propia percepción negativa de revelar información que describa características poco favorecedoras, lo cual repercutiría a la puntuación y al costo del financiamiento.

La mera posesión de datos confidenciales sobre los agentes con los cuales la organización interactúa, representa un riesgo para la firma; Al poseer y

¹⁵ En muchos casos, los usuarios sacrifican su privacidad por obtener servicios electrónicos (o de plataforma) de forma gratuita; al permitir a las plataformas acceder libremente a su información, los usuarios pueden obtener el servicio sin tener que pagar monetariamente por este. Sean los casos como las plataformas de WatsApp, Facebook, servicios Google, Juegos Online, etc.

recolectar información sensible sobre los clientes, la firma se convierte en el portador de diversos riesgos, tanto riesgos internos (Ej. riesgos de reputación por el mal uso de la información) como riesgos externos (Ej. el riesgo de brechas en la ciberseguridad)

Ante la existencia de este tipo de riesgos, la empresa debe de implementar políticas de seguridad para el ocultamiento de la información para preservar la privacidad de la información. (tal como la generalización, supresión, segmentación, o perturbación de la información.) (DeYoung, 2018); lo cual implica la creación de fricciones por los procesos de seguridad, adicionalmente, se posee perturbaciones en las señales generadas con la información alterada por las distintas capas criptográficas que pueden volver ilegibles o ambiguos los datos almacenados. Dificultando la interconectividad de la información obtenida, por lo que la firma requeriría de capacidades aún mayores que van más allá del uso de los algoritmos y de la modelación estadística.

Dentro de las peculiaridades de usar todo tipo de información, los datos pueden incluir variables proxy que se encuentran dentro de las variables recolectadas, poseyendo estas características sujetas a la discriminación subjetiva de religión, sexo o raza; Los datos proxy de este tipo es ampliamente criticado, ya que se sugiere que su uso implica la discriminación racial, variables tales como: la edad, la altura, el color de piel y el código postal. Al usar este tipo de información dentro la creación de un "score", este score puede ser considerado injusto e incluso ilegal, por excluir a grupos pertenecientes a ciertas religiones o razas (Ej. Cómo a la gente de colores o a los musulmanes) (Hurley & Adebayo, 2017)

Por lo que se cuestiona la inclusión de dichas variables para su inclusión en los modelos de discriminación y pricing; Debido a que las mismas condiciones discriminatorias imposibilitan la justa participación de las distintas razas, manteniéndolas en el mismo ambiente socio-económico en la cual quedan

estancadas por la misma discriminación racial. Es decir, el mecanismo del modelo que está basado en datos pasados implica la preservación del sistema tal y como se encuentra; al predecir siempre un comportamiento que puede ser considerado adverso, de alguna raza en específico. Tales prejuicios, no permite desvíos en la conducta actual ya que no se les propicio un ambiente favorable para su evolución.

Evidentemente, la funcionalidad de la información está determinada por la finalidad de esta. Por ejemplo: Es bastante absurdo buscar determinar con la información de Twitter, el nivel de gasto público en infraestructura educativa. Sin embargo, tal vez sea posible medir con un sofisticado análisis la capacidad analítica del usuario (ergo nivel educativo o capacidad intelectual), al observar las características de cada post, al observar el empleo de la gramática, el uso del idioma, el valor literario que se le otorga (el lenguaje) y el tópico discutido.

Por lo que, el uso de la información podrá ser justa o injusta, de acuerdo al uso que se le de ha esta. Es probable que la incorporación de esta información pueda ser empleada para mejorar la sugerencia financiera de acuerdo a su raza y religión, ya que dichas características determinan ciertos hábitos y necesidades propias de su propia estructura orgánica (Dawkins, 2016). Por ejemplo, hombres jóvenes entre 24 a 32 años, se les podría limitar el monto del crédito, debido a la cantidad de testosterona que poseen, lo cual puede inducir en conductas altamente riesgosas, ya que bajo su estructura biológica la perspectiva del beneficio esperado de alguna acción será mayor, al mismo tiempo que es subestimando el impacto del riesgo; en comparación de una persona mayor, cuyo nivel de testosterona suele ser menor y por lo tanto la adversidad del riesgo será menor.

Capítulo II

3 Empleando el Big Data en el Sistema Bancario.

“La vida es inherentemente riesgosa, pero el único mayor riesgo a evadir a toda costa, es el riesgo de no hacer nada”
- Denis Waitley, autor de “Psychology of Winning” (1979)

El corazón de cualquier empresa se centra en la óptima selección de “Q” (cantidad) dentro de un determinado margen de ganancia, el cual se encuentra delimitado por los precios del mercado y los costos de producción, buscando siempre maximizar su beneficio al reducir los costos (maximizar sus ingresos). Ante las distintas condiciones del ambiente y ante las diferentes características de los productos ofrecidos a los clientes con características disimiles; los costos y los riesgos varían según la composición de estos factores: ambiente, consumidor y producto (Pandian, 2007).

Ante la existencia de los diversos riesgos y variaciones en el ambiente, los agentes poseen una utilidad esperada por producir el bien “X”, de la cual, la materialización de dicho beneficio dependerá de los procesos que se tomen para llegar a la liquidación de la mercancía, como se ha mencionado esta perspectiva se encuentra sujeto al conocimiento “k” que el agente posee sobre cada factor:

$$E_t(X|\kappa) = \prod_{i=t-\infty}^{t_0} \vartheta(x|\kappa)_{t_i} \delta^{t_i} - \prod_{i=t-\infty}^{t_0} C(x|\kappa)_{t_i} \delta^{t_i} - P(\text{Risk}|\kappa) * \epsilon$$

$$\left[\prod_{i=t-\infty}^{t_0} \vartheta(x|\kappa)_{t_i} \delta^{t_i} \right] := \text{El ingreso está determinado por las características de los consumidores (ingreso de los consumidores y demanda).}$$

$\left[\prod_{i=t-\infty}^{t_0} C(x|\kappa)_{t_i} \delta^{t_i} \right] :=$ Los costos, están determinados por los procesos productivos seleccionados por el agente, para producir el bien "X"

$[P(\text{Risk}|\kappa) * \epsilon] :=$ Componente de incertidumbre epistemológica, se encuentra sujeto plenamente a las condiciones del ambiente, por lo que su materialización, no depende de las acciones del agente.

Al emplear la tecnología del Big Data, el agente puede maximizar su conocimiento "k", y de ello le permite maximizar su utilidad: al conocer ampliamente a sus clientes y maximizar su ingreso, el identificar los procesos más eficientes y reducir sus costos; y finalmente, el conocer el ambiente en donde se desarrolla el agente, le permite reducir cualquier impacto o pérdida que le pueda ocasionar cualquier imprevisto.

De esta forma, para cada producto, existe un conjunto de riesgos que están relacionados a este; en el caso del crédito, existen distintos grupos de solicitantes que poseen distintas características y para cada grupo, el intermediario posee un nivel distinto de exposición al riesgo.

$$\pi = \sum_{i=1}^n \{P_i Q_i - C_i Q_i - [P(x_i)_{Risk} * Impacto] Q_i\}$$

- $[P(x_i)_{Risk} * Impacto] :=$ Exposición al riesgo $i = Ex_i$

$$\frac{\partial \pi}{\partial Q_i} = \sum_{i=1}^n \{P_i - C_i - Ex_i\} \Rightarrow P_i = C_i + Ex_i$$

"Pricing" de un bien al considerar los riesgos:

$$P_i - (C_i + Ex_i), \quad \{\forall i \exists P(x_i)_{Risk} \in \mathbb{R}(0 \leq P(x) \leq 1)\}$$

De esta manera el "Pricing justo" para el productor, es aquel precio que compensa: tanto el costo de producción como el factor a la exposición al riesgo (probabilidad del riesgo y su impacto de suceso)

$$Ex_i \begin{cases} \text{En caso de Siniestro: } Ex_i \neq 0 \\ \text{Sin materialización del riesgo: } Ex_i = 0 \end{cases}$$

Reflejándose en la utilidad marginal:

$$\partial\pi \begin{cases} \text{En caso de Siniestro: } P = (C_i - Ex_i) \Rightarrow \partial\pi = 0 \\ \text{Sin materialización del riesgo: } P \neq (C_i - Ex_i) \Rightarrow \partial\pi = Ex_i \end{cases}$$

$\therefore Ex_i$ es una ganancia extraordinaria,
que compensara la materialización del riesgo en algun momento

Para el intermediario financiero, el medir y administrar los riesgos de las distintas cadenas productivas; permiten crear nuevos flujos de crédito hacia las nuevas cadenas productivas con altos niveles riesgos (inversión de nuevas tecnologías productivas); (Schumpeter, 1911) de tal modo, la repercusión de un proyecto de un alto nivel de riesgo es compensado al promediar el riesgo con una mayor cantidad de proyectos con un menor nivel de riesgo.

Al suponer que los costos de producción son igual a "0", se puede denotar que el nivel del riesgo es el único factor limitante en la colocación del crédito. Sin embargo, dentro del proceso de interacción con los clientes, estos generan costos individuales por las fricciones generadas en la contratación y el intercambio de información con cada cliente. En el caso de los microcréditos, y a las aplicaciones a pequeña escala de ahorro (como de los segmentos de hogares que perciben ingresos de entre 3 y 5 Salarios mínimos Generales), los costos administrativos y el costo de la información (entrevistas con el cliente), son los factores que limitan la entrada a agentes con actividades financieras de pequeña escala. (Jagtiani & Lemieux, 2018).

En los procesos de intermediación, dentro de las actividades de la banca comercial. Los costos del banco pueden catalogarse de la siguiente manera (Bessis, 2002):

$$C = c_T + c_{op} + c_R$$

c_{inf} := *Costos de Transaccion (transaccion, adquisicion, distribucion, analisis)*

c_{op} := *Costos Operativos (mantenimiento instalaciones, personal de servicios)*

c_{op} := *Costos Retribucion (Tasa de Interes Pasiva)*

Para poderse relacionar el intermediario con los clientes, deben de ser generados diversos costos que son cruciales para la creación y mantenimiento de canales de comunicación con los clientes (tanto con inversionistas como con sujetos de crédito); denotados usualmente como "Costos de Transacción" (Williamson, 1989) dentro de estos se encuentran los procesos administrativos de búsqueda de nuevos clientes, la creación de los contratos, de su verificación y creación de medidas de seguridad; dichos procesos generan fricciones en la economía, ante los procesos requeridos que interrumpen o impiden el flujo de las negociaciones.

En un ambiente en donde se empleaban procesos mono-canal, se poseían barreras a la entrada de los clientes, por las fricciones generadas al contratar los servicios financieros forzosamente en las ventanillas físicas, de las cuales, su ubicación o lejanía determinan los costos del cliente por emplear los servicios financieros del banco. Con la aparición de las nuevas tecnologías relacionadas al análisis Big Data y a la infraestructura del internet, las cuales permiten crear y unificar los múltiples canales que posee la banca con el cliente (sucursales físicas, cajeros automáticos, servicios en línea o móvil.), dichas innovaciones permiten reducir los tiempos de respuesta (los costos de transacción), al mejorar la interacción que el cliente puede realizar con el conjunto de canales que el banco actualmente brinda, (O'Brien, 2017) permitiéndole al banco ofrecer productos

financieros de autoservicio y productos preautorizados que facilitan su uso y contratación.

De esta forma, los costos administrativos y las fricciones pueden ser reducidos, al masificar las aplicaciones de crédito y ahorro mediante el uso de dispositivos electrónicos como medio de contratación, siendo estos dispositivos móviles un canal continuo de comunicación; Mediante algoritmos Machine Learning, los documentos que son enviados por los clientes pueden ser analizados para comprobar su veracidad y la información contenida, sin la necesidad de la intervención humana (Ravi & Kamaruddin, 2017).

Al incorporar servicios Multi-Canal, la efectividad de los canales puede ser analizado, mejorando el tipo de información que es enviada al usuario, el cual puede ser alineados al tipo o perfil del consumidor para que este tome provecho de la información enviada, la cual deberá ser relevante a su estilo de vida (United States of America Patente nº 7,062,510 B1, 2006); mejorando la efectividad de las campañas de mercadeo, minimizando los costos por la adquisición y retención de los clientes; Al mismo tiempo que es preservada la reputación de la empresa y la relación que el cliente posee con el banco, al reducir los mensajes no deseados (o SPAM).

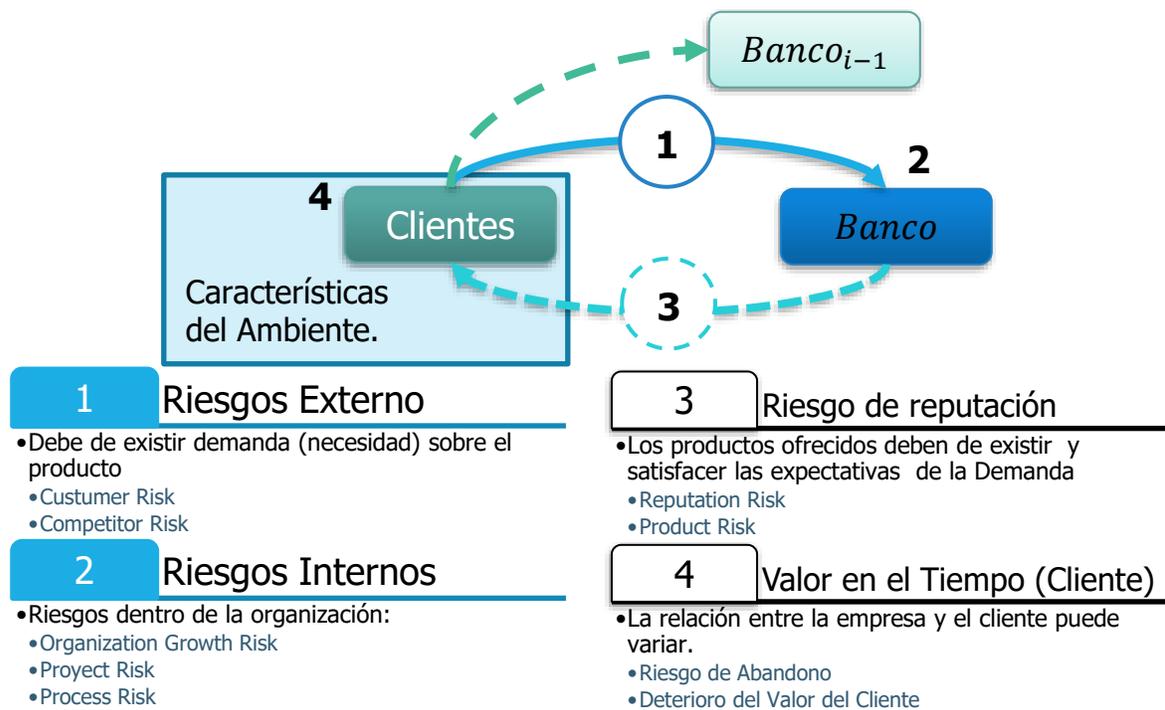
De la misma manera, que son reducidos los costos de transacción del cliente al emplear los servicios omni-canal que el banco pueda ofrecer; La cobertura de dichos servicios es extendida, lo cual permite al intermediario abarcar un área mayor con menor costo, como en el caso del banco Tinkoff, quien no posee sucursales físicas, por lo cual no requiere de la construcción de más sucursales para amplificar su área de servicio.

La organización entre el cliente y la empresa, puede ser fraccionado en 4 fases: adquisición, producción, distribución y retroalimentación, cada fase posee

riesgos inherentes por los procesos involucrados como se describe en la Figura 2 elaborado en base a (King, 2014) y (Bessis, 2002):

- Adquisición de Clientes (costos de adquisición)
- Servicio al cliente (costos de negociación)
- Relación y Retención del consumidor (Costos de distribución)

FIGURA 2. Riesgos en las relaciones Cliente-Empresa (en este caso: Banco)



3.1 Riesgo Externos, Existencia del Mercado

Determinado aquí como el axioma 2 de St. Thomas de Aquinas; nada es efecto de su propia causa. De tal forma la existencia de la oferta no existe por mera casualidad, por lo que esta debe de encontrarse sustentada en una necesidad definida. La banca comercial posee una singularidad en la demanda de su producto, debido a que el producto ofrecido posee una forma económicamente abstracta usualmente percibida como: dinero (el monto del crédito).

De esta forma, existe una causa material (dinero), la cual es transformada en una causa final (mercancías), de acuerdo a las necesidades de la causa eficiente (plan del cliente). Por transitividad: la causa primera (el otorgamiento de un crédito), está sujeto a la existencia de la necesidad del cliente (causa final). Por lo que el servicio ofrecido por el banco, debe de estar alineado al objetivo final que el cliente posee, con los montos obtenidos mediante el financiamiento de sus proyectos. Tal como la visión que tenía Nielsen del banco RBS.

Mediante el enfoque, de administración preventiva del riesgo, con las capacidades de procesamiento de las bases de datos; los intermediarios se convierten parte importante de tablero, al administrar los riesgos, al transformar o modificar el comportamiento de sus clientes, mediante la creación de esquemas y productos con los suficientes incentivos y utilidad (para el consumidor), de tal forma que permitan generar mejores scores a los clientes, al brindar servicios que incentiven o permitan el buen comportamiento. (King, 2014)

Mediante los algoritmos de clusterización (Unsupervised Learning), los clientes pueden ser agrupados o clusterizados, generando "n" cantidad de subgrupos; dicho proceso es habitualmente conocido como "microsegmentación". Con la suficiente información sobre los individuos y/o los subgrupos que poseen características similares: se pueden conocer sus hábitos, su comportamiento y sus necesidades particulares. Al analizar minuciosamente los comportamientos del cliente, pueden ser diseñados ofertas especializadas y personalizadas para cada microsegmento, induciendo sus comportamientos a la baja exposición al riesgo (Rajan, 2019): creando especiales ofertas tales como, microcréditos para la adquisición de productos específicos, basadas en los intereses y necesidades del cliente, por ejemplo estos pueden ser montos y plazos que permitan cubrir el financiamiento de: un nuevo videojuego, libros de texto de uso escolar, equipo de ejercicio, pequeños lotes de insumos básicos o nueva maquinaria requerida.

Con el uso de la información tradicional, los clientes pueden ser agrupados por nivel de ingreso, código postal y profesión u ocupación (Arya, Eckel, & Colin, 2013); al incorporar la información alternativa, las agrupaciones pueden realizarse con mayor detalle, lo cual permite crear mayor conocimiento sobre los grupos específicos, permitiendo crear expectativas precisas sobre las conductas de los nuevos clientes que son categorizados con el uso de la clusterización.

Al poseer una mayor cantidad de detalle sobre los clientes y su información personal, es posible crear segmentos pequeños que absorban las características comunes y sea posible agrupar los riesgos con la menor desviación sobre la media del cluster; siendo que estos consumidores con características similares y necesidades semejantes, percibirán una mayor satisfacción al obtener un producto diseñado específicamente para el subgrupo, que al mismo tiempo tendrán menor desviación sobre la media del subgrupo.

Los procesos de clusterización, pueden realizarse bajo distintos enfoques; siendo posible realizar segmentaciones en base tanto de las necesidades como del valor que el cliente le da a la firma. Siendo posible realizar distintos modelos de agrupación con los distintos lotes de variables (Evry, 2014); permitiéndole al banco crear análisis detallados de cada grupo, comprendiendo ampliamente el tipo de servicio y/o producto que el cliente necesita, determinando así los motivos por los cuales el cliente abandona al banco, con lo cual también puede ser calculada la probabilidad de recuperación del crédito; Poseyendo una expectativa de beneficio para cada grupo, con el cual puede determinar el nivel óptimo de los costos destinados para cada segmento:

$$CLV = \sum_{i=0}^n \frac{R_i}{(1+d)^i}$$



Ilustración 8, Factores que determinan el CVT

comunicación; por lo que este modelo posee 4 factores importantes para medir el valor de la relación con el cliente: lealtad (expectativa de vida), ingresos, costos y tasa de descuento.

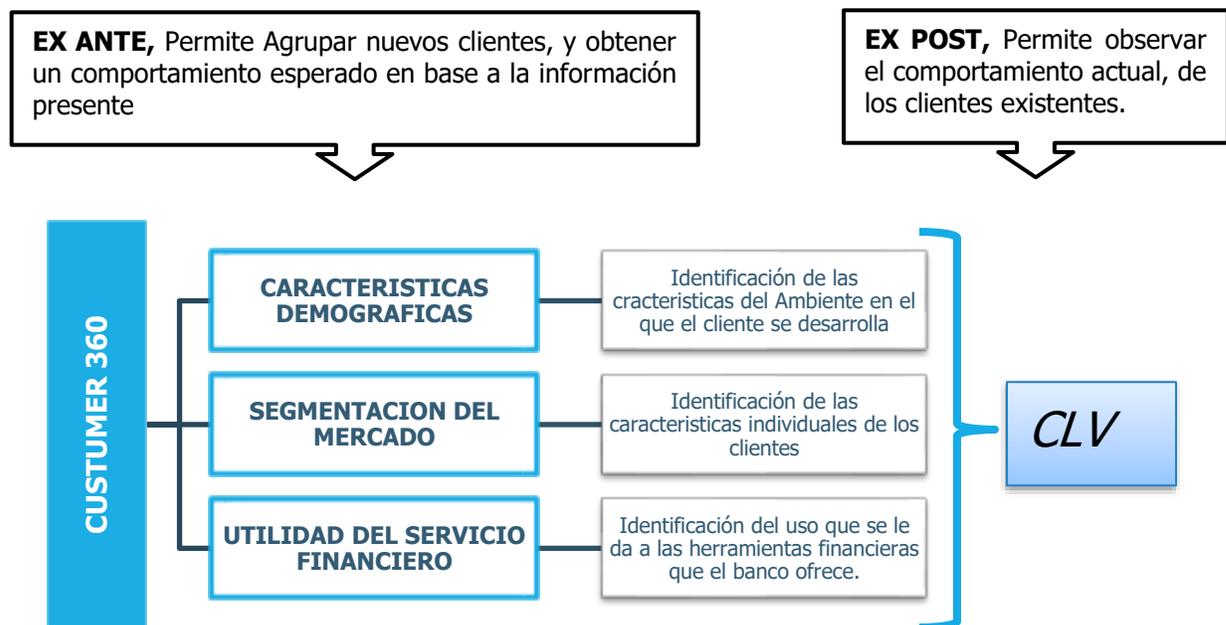
Para tener una aproximación del CLV, se emplea el modelo "RFM" (lo Resiente, la frecuencia y lo Monetario) (Sohrabi & Khanlari, 2007), Bajo este modelo, se asume que es posible evaluar los patrones del comportamiento futuro, en base al patrón de la actividad actual y anterior del cliente. Al utilizar variables tales como: la frecuencia con la que el cliente usa los servicios bancarios, los montos de su operación, el tiempo de vida del cliente y el tiempo de su última adquisición; En base al RFM, refleja la relación que posee el banco con el cliente en función de sus actividades cotidianas.

Debido a que existen un sin número de factores que pueden determinar la actividad del cliente con el banco, el "CLV" es un valor ampliamente abstracto que no explica la conducta o los motivos reales del cliente (Blattberg, Malthouse, & Neslin, 2009); ya que su valor puede estar inducido tanto por las características

Bajo el enfoque "Customer Lifetime Value" (CLV), se generan estimaciones sobre el valor presente de los ingresos obtenidos por la actividad del cliente durante su expectativa de vida, siendo este el valor presente de todos los flujos futuros de efectivo (Pfeifer, Haskins, & Conroy, 2004), lo cual permite observar y diferenciar los clientes "valiosos" y que poseen una relación más estrecha con el banco. Como herramienta de discriminación, el CLV ayuda a determinar el nivel de aceptable de los costos de adquisición y del costo de mantener el canal de

del banco (campañas de marketing, herramientas o servicios, productos, reputación, etc...), como por el estado del ambiente (como la estabilidad económica, temporadas, seguridad, competencia, etc...), y las características del individuo (Su nivel de ingresos y su comportamiento habitual). Para poder comprender los elementos que mejoran la calidad del cliente, es preciso segmentar los clientes en base a las características del mercado y su personalidad.

Figura 3, Distintos enfoques, de Clusterización.



Con el uso de variables alternativas, es posible tener mayor flexibilidad en las variables que puede ser empleadas y obtener distintos grupos de segmentación bajo distintos enfoques, mejorando el análisis de los grupos de clientes que pueden ser comparados entre sí; entre estas nuevas variables, se encuentra la "huella digital" que el cliente deja al momento de realizar sus compras o utilizar las distintas herramientas en línea.

Analizando de esta forma, el tipo de uso o utilidad que el cliente le da a los servicios financieros ofrecidos y a las herramientas que el banco pone a su

disposición. Ayudando a la comprensión del "CLV" al analizar los distintos aspectos del servicio y comparando los clientes que pertenecen al mismo grupo, obteniendo medidas relativas que ayudan a detectar anomalías o inconvenientes en los servicios ofrecidos y la efectividad de los costos asociados a la adquisición (C_d) y mantenimiento del cliente. (C_i) (Khajvand & Tarokh, 2011)

Algunos autores emplean el FRM y sus variaciones como método de segmentación, sin embargo, la posibilidad de utilizarse como medio de análisis se encuentra limitado solamente a los clientes existentes quienes ya poseen algún historial o antecedente dentro de la firma; por lo que, para obtener mayor beneficio de esta herramienta, se podría considerar este método como la evaluación del subgrupo en un enfoque ex-post (posterior a su aplicación).

Al generar clasificaciones y agrupaciones en base a variables que pueden ser observadas ex-ante del análisis. Se pueden analizar los valores CLV o el Score promedio sobre los clusters, de esta manera el valor medio CVL de cada subgrupo, será el valor esperado de los nuevos clientes que sean agrupados en determinado conjunto; De esta forma, se posee una mejor expectativa de su comportamiento futuro (riesgo) y de la rentabilidad esperada en base a la información que se poseen de los clientes existentes del grupo. Obteniendo de esto, un parámetro de discriminación en base a la utilidad esperada a lo largo de la vida de cada cliente, lo que permitirá elegir una estrategia adecuada de comunicación y servicio de acuerdo a sus características, ajustando dinámicamente los costos empleados del servicio (PLV), en base a la evolución que se posea en a la relación con el subgrupo:

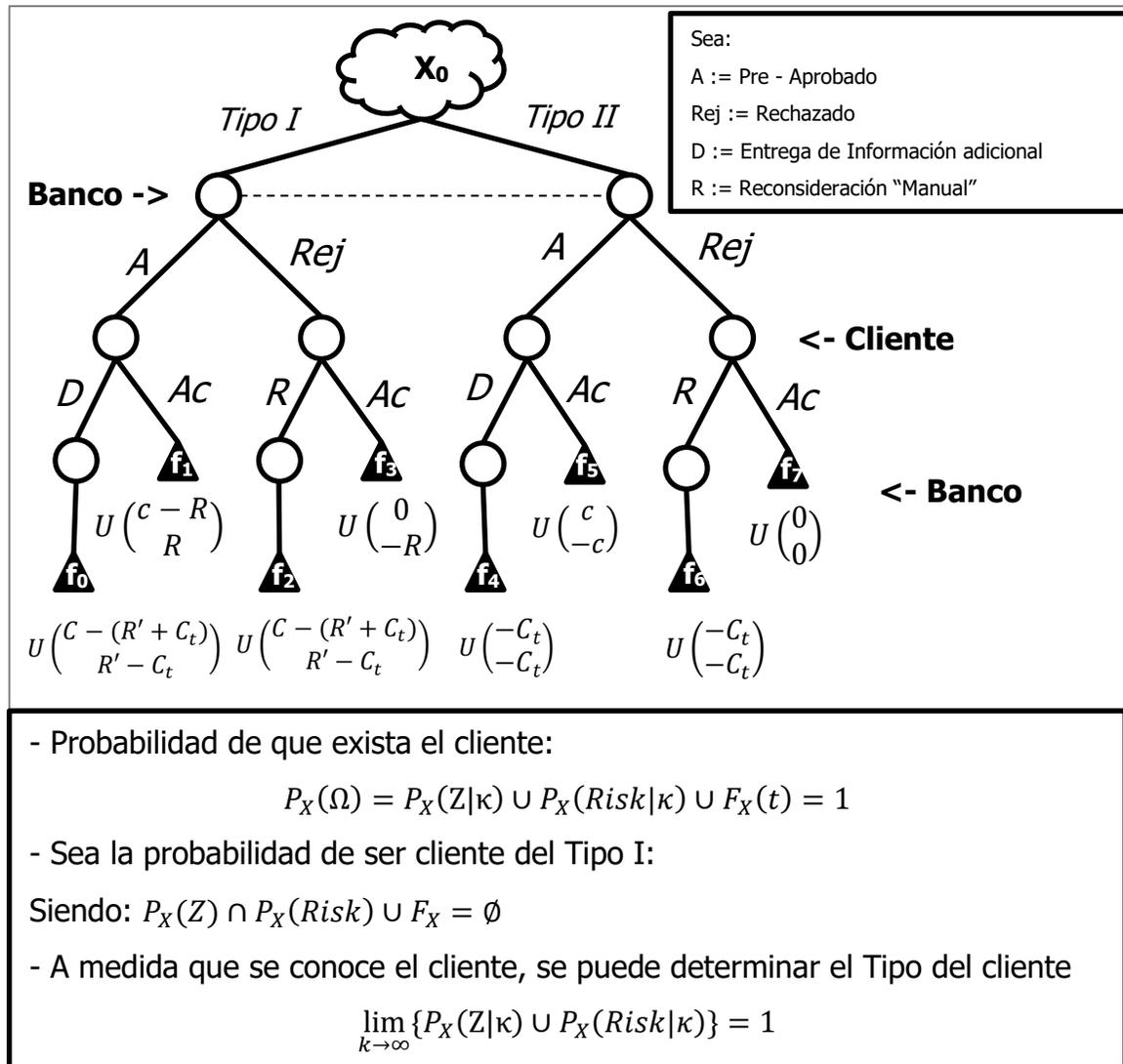
$$CLV > PLV \text{ (Proyect Life Value)}$$

$$\sum_{i=0}^n \frac{R_i}{(1+d)^i} > \sum_{i=0}^n \frac{C_i}{(1+d)^i} - C_d$$

$$\sum_{i=0}^n \frac{R_i - C_i}{(1+d)^i} - C_d > 0$$

Al incorporar servicios multicanal y autoservicio, se puede mejorar drásticamente la interacción que el cliente posee con el intermediario, con lo cual se mejora la experiencia del cliente al brindar un sentido de pertenencia; permitiendo obtener el "feedback" y monitoreo sobre el servicio y los productos financieros que han sido adquiridos. (MBAMA, 2018)

Figura 4, Teoría de Juegos, esquema de autoservicio (caso del banco)



Elaboración propia

Para poder incorporar la colocación de créditos, por tipo de venta de "autoservicio" a través de aplicaciones móviles; debe de existir el suficiente incentivo o certeza, para ello, se requiere del adecuado nivel de información sobre el aplicante, para que el banco determine el tipo de cliente que solicita el crédito, esta información puede ser obtenida tanto de los medios tradicionales como los alternos (como la huella digital, el microblogging, etc.)

Con el uso de los canales de "autoservicio", el banco mejora la calidad de servicio en distintos aspectos, con los cuales la satisfacción de los clientes aumenta: mejorando la accesibilidad de sus productos y mejorando los servicios brindados al personalizarlos a las características del cliente (Sindwani & Goel, 2015); Al mejorar los procesos de contratación de los servicios financieros, los costos de distribución asumidos por el banco y los costos de transacción que son asumidos por el cliente, son reducidos significativamente, en comparación con los costos del servicio en ventanilla.

De esta forma, el banco mejora la penetración del mercado al exponerse a los riesgos ligados a la banca de auto servicio ¹⁶. Por lo que deben de fijarse esquemas o mecanismos que minimicen tales riesgos (como el ilustrado en la figura 4), el banco puede establecer un esquema el cual incentive al consumidor a revelar toda la información necesaria para su administración, comprobando la identidad del cliente y minimizando el impacto por el comportamiento oportunista que el cliente pueda tener; Para incentivar un comportamiento favorable, el banco puede limitar los servicios provistos al cliente de autoservicio "nodo terminal f1" ($R >$

¹⁶ Del comic de Peter Steiner (1993), se dice que: "por el internet, nadie sabe que en realidad eres un perro" Ya que se encuentra en la actividad en línea un pozo de riesgo moral por el uso de avatares y demás símbolos que cubren la identidad real del individuo. De tal modo, que no se puede conocer de forma contundente quien se encuentra detrás de la pantalla y cuales sean sus intenciones. Por lo que la operatividad en línea de las empresas, deben de poseer mecanismos de seguridad que limitan la operatividad de estas, es decir, mecanismos que generan fricciones tal como la generación de tokens de seguridad.

$R' \& C > c$); y puede mejorar progresivamente el precio a medida de que se posea mayor información del cliente, por medio de las interacciones que se tengan con este.

Dadas las características del ambiente, los jugadores (los agentes) pueden ser de dos tipos: clientes que retribuyan por los créditos otorgados (Tipo I) y clientes oportunistas que no poseen intenciones de pagar el crédito (Tipo II); La determinación de un tipo o de otro se encuentra designado por el tipo de señales que se hallen en la información que se posee del prospecto ($PX(Z|\kappa) \cup PX(Risk|\kappa) \cup FX(t)$), en la información que se tiene de los clientes ya existentes ($P(Z)$) y de la reputación que el banco posee ($P(d)$), (el tipo de banco que el cliente percibe, para facilitar el análisis, solo se considera un "tipo" de banco).

$$P(Z|k') = \frac{P(Z|k) * P(Z)}{P(d)}$$

Al no conocer el tipo de cliente quien realiza la aplicación al crédito, se determina la utilidad esperada de la pre-aceptación:

$$U(A; Ac, Ac) = P(Z|k') * (R) + P(Risk|k') \cup F_X * (-c)$$

$$U(Rej; Ac, Ac) = P(Z|k') * (-R) + P(Risk|k') \cup F_X * (0)$$

$$U(A; Ac, Ac) = U(Rej; Ac, Ac)$$

$$P(Z|k') * (R) + P(Risk|k') \cup F_X * (-c) = P(Z|k') * (-R)$$

$$P(Z|k') * R = P(Risk|k') \cup F_X * c$$

Siendo que cuando la utilidad esperada es equivalente: $U(A; Ac, Ac) = U(Rej; Ac, Ac)$, el banco es indiferente de brindar o no el "auto-servicio". De esta forma, la condición necesaria para que el banco posea el incentivo suficiente para

la colocación del crédito es: La utilidad esperada debe de ser mayor a la exposición del riesgo.

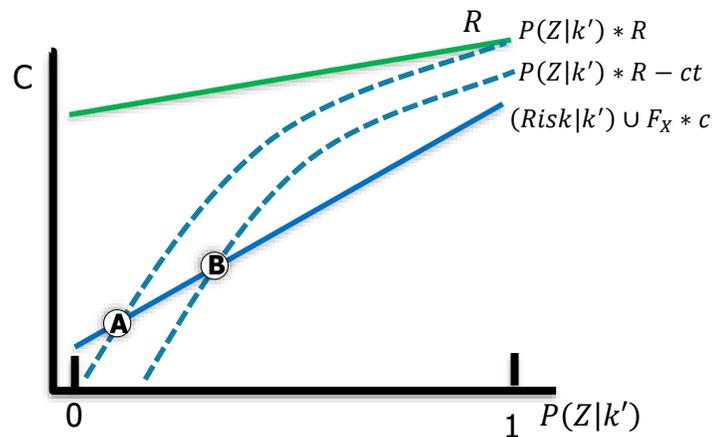
$$P(Z|k') * R > (Risk|k') \cup F_X * c$$

$$\Rightarrow [P(Z|k') * R] > [P(Z|k') * R - ct]$$

$$U(A; Ac, Ac) > U(A; D, D) > U(Rej; Ac, Ac)$$

De esta forma, el banco limita la colocación del crédito en base al conocimiento que se tiene del consumidor y en base a los costos que el banco posee en los trámites relacionados a la solicitud del crédito.

Figura 5, Riesgos aceptados por el banco



Tal como se Gráfica en la figura 5; Al reducir los costos de transacción y los costos administrativos el punto de quiebre, en donde existe el incentivo suficiente para que el banco asigne cierta cantidad de recursos al cliente "X", de este se requiere de un nivel menor de información ($P(Z|k')$); de esta forma, se puede deducir teóricamente: al reducir los costos de transacción al mínimo (cero, en este caso), así como los costos de distribución (sean los costos fijos de las ventanillas y cajeros), la flexibilidad de los recursos es mayor, permitiendo al banco ofrecer microcréditos, que pueden estar ajustados a las necesidades de las pequeñas empresas y clientes con bajo nivel de ingresos (Jagtiani & Lemieux, 2018); Los cuales, pueden ser aprobados con la información disponible en el "smartphone" del consumidor. A medida que el cliente interactúa con la compañía, este puede generar un mayor nivel información por el historial que va creando, permitiendo

mejorar su reputación o "score" y calcular el CLV personal a lo largo de su relación con el banco, que permitirá incrementar el límite del crédito al reducir la exposición del banco.

Los servicios del e-banking o los servicios bancarios de autoservicio a través del internet, son canales de alta exposición de riesgos debido a la posible actividad oportunista del cliente. Sin embargo, la tecnología ayuda a crear mecanismos para identificar al cliente sin necesidad del trato presencial; mediante el uso los mecanismos e-KYC (Identificación digital biométricos del cliente) se analizan por medio de los instrumentos de medición del smartphone (como la cámara o la grabadora) los rasgos característicos del cliente tal como las características biométricas: el reconocimiento facial, de voz, de ubicación e incluso tipográfico; lo cual permite validar la legitimidad del cliente y al mismo tiempo poseen un menor costo de transacción. (Bank of Thailand, 28 August, 2018) Tal como en el caso del sistema WeBank, la cual emplea las cámaras de los dispositivos como herramienta de reconocimiento facial, con la finalidad de poder rastrear la identidad del cliente, empleando la información provista por el Ministerio de Seguridad Pública de China. (Kshetri, 2016).

Aun cuando existen herramientas o mecanismos que permiten reducir los riesgos de la banca digital, prevalece por parte del banco, el riesgo de canibalismo de los servicios ofrecidos por los cajeros y sucursales; los cuales poseen costos hundidos por la construcción de la infraestructura, los cuales poseen cierto grado de obsolescencia al incorporar ampliamente los servicios ofrecidos digitalmente. Bajo las costumbres actuales, aun es requerido la infraestructura necesaria para el empleo de los servicios bancarios, como depósitos y retiros de efectivo, siendo esta costumbre, un factor limitante para poder incorporar en su totalidad dicha clase de servicios en línea. De acuerdo con la encuesta realizada por Capgemini, a los bancos comerciales alrededor del mundo en el 2016: El 93% de los bancos están

dispuestos a mover sus operaciones a un entorno digital, aunque solo el 13% posee las capacidades de lograrlo por los costos que implica el cambio de la cultura y el canibalismo de sus activos fijos, claro que, para incorporarse al entorno digital, el 64% de los bancos encuestados, ven a las Fintech y a los startups como un medio para avanzar al entorno digital, al colaborar con estas o adquirirlas.

Al introducirse en los entornos digitales, los bancos permite reducir los costos en la comunicación con los clientes y así mismo poseen más herramientas para la recolección de datos, facilitando el análisis individual y siendo herramienta fundamental de las decisiones financieras que el cliente toma (Тинькофф Банк, 2019); Por lo cual algunos intermediarios financieros han expandido sus operaciones, ocupando el espacio de los intermediarios mercantiles y viceversa: Por ejemplo: Alibaba (Zhou, Chen, & Li, 2018) siendo una plataforma de e-commerce, expandió sus operaciones al nicho de los productos financieros, empleando la información que obtiene a partir de las compras de sus clientes dentro de la misma plataforma de comercio. Así, Alibaba opto por ofrecer préstamos destinados al consumo además de ofrecer servicios de ahorro. Determinando el monto del préstamo en base a las características de los productos que son vendidos en su plataforma, así como la hora de la compra que es efectuada la compra y el historial de artículos de lujo previamente adquiridos (no se especifica, pero probablemente estos artículos de lujo son considerados como una garantía ante el incumplimiento).

En el caso de los bancos que expanden sus operaciones a la distribución y venta: En Europa, la asociación entre bancos Deutsche Bank, HSBC, KBC, Natixis, Rabobank, Société Générale y UniCredit; crearon un consorcio fundado como Digital Trade Chain. (Groenfeldt, 2017) El cual creo una herramienta basada en plataforma llamada "We.Trade" dirigido especialmente para las Pequeñas y Medianas Empresas (PyMEs) con la finalidad de crear una red comercial diseñada

para establecer vínculos entre potenciales clientes y proveedores, encadenando los planes de producción por medio de dicha aplicación móvil.

Sus funciones permiten mejorar el flujo de efectivo de sus clientes al facilitar la creación de canales de comunicación y así mismo los bancos proporcionan apoyo en las capacidades de liquidez mediante la colocación del crédito; De esta forma, dicha herramienta aumenta la probabilidad de retorno del crédito y el volumen de los préstamos otorgados. Con el uso de la aplicación, es recolectada la información de los usuarios la cual es almacenada mediante el uso de la tecnología Blockchain (siendo esta una herramienta en seguridad), en el cual pueden ser almacenados los historiales de compras de los usuarios registrados; con lo cual se obtiene la información de forma segura y veraz sobre las cadenas productivas y las necesidades de sus clientes.

3.1.1 Análisis de Sentimientos (Customer Understanding)

Para cada decisión tomada individualmente existe un conjunto de información (estimulo) el cual induce las conductas del agente, la cual es procesada en base a las características neurobiológicas del cliente, forjando alguna particular personalidad al nodo el cual interactúa con su grupo primario y es retroalimentado por las consecuencias observadas de sus acciones; por lo que cada acción individual posee efectos dentro del grupo primario (Efecto Adaptación). Cada nodo del sistema posee acciones que se encuentra limitadas dentro de un marco institucional (moral-religiosa, política y legal), lo cual reduce la incertidumbre y delimita el efecto sistémico derivado de algún comportamiento adverso que pueda ser realizado en conjunto, por los nodos que pertenece al sistema; De tal forma las instituciones están diseñadas dentro del ideal humano, para minimizar los daños que puedan ser ocasionados por los efectos del caos. (North, 1991)

Tras la aparición de ciertas eventualidades, forjadas de la naturaleza caótica del sistema, ejercen dentro del sistema, fluctuaciones que desequilibran y modifican las condiciones del ambiente (entropía). Forzando a los agentes económicos a reestructurar tanto su comportamiento y como sus actitudes, como respuesta a la propagación de tal disrupción; tal comportamiento se le conoce "Negative Feedback System"; de tal forma el agente observa dichas perturbaciones, y experimenta al observar los resultados derivados de sus decisiones, modificando su comportamiento ante cambios de dichos resultados; ante un sistema económico dinámico, las conductas adaptativas permiten reestructurar el sistema, alcanzando un equilibrio temporal, hasta que una nueva perturbación aparezca (Hayek, 1945).

Tras los cambios irreversibles que son provocados por las perturbaciones, el sistema es reestructurado en base al conocimiento que es forjado por la experiencia de dicha perturbación. Modificando la estructura del sistema, con la finalidad de ser capaz de afrontar los impactos de eventos idénticos o similares. Dentro del ambiente de competencia (caótico), el sistema posee una energía dinámica y siempre fluctuante (entropía), describe Mahoney (1991), las fluctuaciones crean oportunidades en la reorganización efectiva para algunos subsistemas, pero no crea una metamorfosis total del sistema, extinguiendo los subsistemas que no son capaces de adaptarse. La capacidad de reorganización, está sujeta a las capacidades del nodo, de sus recursos y de sus interacciones con otros nodos del sistema (transferencia de la información).

Sintetizando, los individuos actúan ante la adversidad dependiendo de su ubicación dentro del sistema. Esto quiere decir: toman sus decisiones dentro de un entorno delimitado por el ambiente y la estructura socio-económica que lo rodea. Los métodos que son empleados por el agente para satisfacer sus necesidades, obedecen a la personalidad que el individuo posee, la cual se encuentra esta

moldeada por las recompensas o castigos (emociones), inducidos por las decisiones tomadas y experiencias de sus acciones (Freud, 1940); es decir, se encuentra basada en su conocimiento y la formación de hábitos que minimizan el costo de la toma de decisiones.

Para saciar sus necesidades, los agentes buscan dentro de los objetos (mercancías) características particulares, Karl Marx lo define a estas características como valor de uso; definiéndose que la misma mercancía posee diferentes valores de uso dependiendo del individuo que lo porta, dependiendo de la necesidad que lo motive y la utilidad esperada de su empleo. Por ejemplo: A una hoja de papel se le pueden otorgar un distinto valor de uso, sea este una herramienta una de estudio, de trabajo, para decorar o para satisfacer al ocio (al doblarlo y aventarlo). De tal forma, los consumidores poseen cierto interés y utilidad el cual se encuentra determinado por las características específicas del producto.

En los modelos clásicos de economía, la demanda está determinada por diversos factores: precio del bien, precio de bienes relacionados, ingreso, gustos y necesidades estacionales; determinando modelos de demanda mediante la variables observables y tangibles como lo son el precio y los ingresos, las cuales poseen relativa sencillez de ser observados. La variable "Gusto" o las emociones, es una variable altamente subjetiva y difícil de cuantificar, pero esta es la variable vital en las elecciones cotidianas de los individuos (o nodos).

Un nodo, posee la característica de ser consiente de sí mismo, posee la noción de su propia existencia, el cual se desarrolla en un ecosistema natural y social. Crea información de las experiencias de los eventos pasados (recuerdos); de cada acción posee recompensas (placer) y castigos (dolor) sociales o personales, físicos o emocionales. Los nodos modifican su comportamiento frecuentemente ante una continua y constante creación de información; creando

experiencias con diferentes impactos y peso en su memoria (o trauma) (Locke, 1960).

De esta forma el ser humano es racional, ya que es inducido a preservar su especie bajo un criterio de adaptación, saciando sus necesidades individuales y las necesidades sociales (necesidad de pertenecer a un club); Los individuos forjan su propio conocimiento en base a sus experiencias el intercambio de la información (Ej. Opiniones y reseñas) y por la mimetización de ambiente; La mimetización, dentro de un espacio cultural el cual es heredado genéticamente y aprendido mediante la observación de sus manifestaciones en el ambiente (fenotipo), tal transmisión de información, definidas como "Memes" por Dawkins (1982) dan forma al proceso evolutivo de la cultura al transferir el conocimiento, generando conjuntos de subsistemas heterogéneos, los cuales poseen distintas características y comportamientos.

La heterogeneidad de las culturas, crea un conjunto diverso de expectativas, poseyendo cada una de estas culturas un desarrollo cognitivo distinto, generando una interpretación distinta de la información que es recolectada y distintas respuestas ante las situaciones adversas. (Staats, 2003) A medida que el conocimiento del agente y su capacidad cognitiva sea mayor; su creatividad¹⁷ será mayor y será capaz de adaptarse al ambiente con mayor velocidad. (Rae, 1834) Ante el bajo conocimiento sobre las eventualidades (la ignorancia de los posibles desenlaces o incertidumbre), el individuo replicara el comportamiento encontrado en la estructura social a la cual pertenece (Ferguson & Bargh, 2004), de esta forma la personalidad es influencia de las opiniones de sus cercanos y de los actores

¹⁷ Las ideas y la creatividad de una persona están limitada al conocimiento que este posee sobre el mundo, no se pueden crear objetos nunca antes vistos (Ej. no es posible imaginar un planeta sin conocer previamente un círculo y posteriormente una esfera), las ideas "nuevas" no es más que la simple combinación de objetos previamente conocidos. (Platón, 400 A.C.)

importantes¹⁸ tal como: calificadoras, críticos, mentores, figuras públicas, personas de un gran lazo emocional o la opinión de la mayoría.

El individuo reacciona de diversas maneras ante la información recolectada: el comentario o recomendación posee un impacto diferente, dependiendo de la fuente de la que provenga y la eficiencia del canal. Ej. Ver un letrado "los mejores tacos del país" posee un valor de credibilidad menor que la recomendación de un amigo cercano. Cabe que mencionar que, no necesariamente una opinión o la difusión de la información diluye completamente el costo de la información ni la incertidumbre, debido a que el individuo se encontrara en la situación de arriesgar sus recursos en buscar su propio conocimiento ante la permanencia de la incertidumbre.

Las decisiones que son tomadas por los agentes, en función al conocimiento obtenido posee una última polaridad (Sigmund Freud, 1940), en donde el individuo real, expresado como "ego", se encuentra en la búsqueda del balance entre el "Id" y el mundo exterior; manteniendo una actitud ante las respuestas de los distintos estímulos (recompensas, castigos o peligros) de sus características

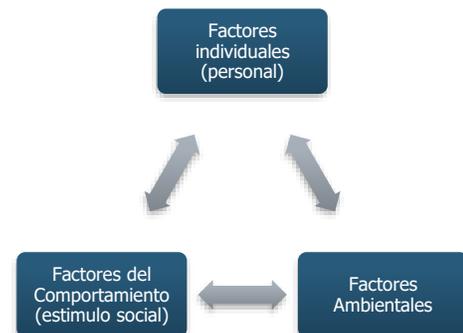


FIGURA 6, Determinación recíproca del comportamiento adaptado de (Albert Bandura, 1986)

individuales, biológicas e instintivas (el Id), cuya saciedad se encuentra delimitada por los estímulos del mundo exterior, siendo estos estímulos posibles respuestas de las acciones y decisiones tomadas dentro de un ambiente con determinadas

¹⁸ "Pero hay que saber disfrazarse bien y ser hábil en fingir y en disimular. Los hombres son tan simples y de tal manera obedecen a las necesidades del momento, que aquel que engaña encontrará siempre quien se deje engañar... Pues los hombres, en general, juzgan más con los ojos porque todos pueden ver, pero pocos tocar. y estos pocos no se atreven a oponerse a la opinión de la mayoría" ("El Príncipe", Nicolás Maquiavelo, 1532)

estructuras sociales. Es decir, que el comportamiento es inducido en parte por las características sociales del ambiente y por el papel social que el nodo desempeña dentro en él sistema.

Dentro de estas teorías sobre el comportamiento y de las decisiones individuales, se encuentran una serie de variables que solamente pueden ser observados mediante una compleja recopilación de información a través de herramientas de medición neurocientífica, tales como el fMRI (functional magnetic Resonance imagin), el EEG (electroencephalography), el GSR (Galvanic Skin Response). Así mismo, es probable que variables como los enlaces sentimentales entre los individuos sea difícilmente observable, pero el análisis de las plataformas de redes sociales puede dar un acercamiento a tal medida, de tal forma, el Text Mining puede ser un acercamiento al análisis de las características particulares de un individuo. Permitiendo observar mediante su actividad en línea, algunos rasgos de su comportamiento, las características cognitivas y las emociones que los individuos poseen.

Dentro de las nuevas herramientas de análisis, nacidos para el estudio del Big data, se encuentran las ramas del Análisis de Sentimientos, también conocido como "Opinión Mining"; la cual consiste en analizar información en formato de texto; para el análisis de este tipo de información se emplean algoritmos del tipo "Text mining" y de "Natural Language Procesing" (NLP). El objetivo de este campo, se centra en análisis de los hechos y opiniones de las entidades observadas, observando con ello las actitudes, emociones y sentimientos que una empresa, organización o persona física posee sobre un evento u objeto (Pawar, Jawale, & Kyatanavar, 2016). Siendo esta una variable que permite medir la satisfacción y las actitudes de los clientes, así también las características del mercado, la reputación de la firma y de la competencia.

El NLP, es un proceso que funciona como puente entre el lenguaje computacional y el lenguaje humano; con el cual la computadora es capaz de descifrar mediante el uso de la Inteligencia Artificial, los textos de formato no-estructurado, traduciéndolo a un formato estructurado el cual puede ser analizado mediante algoritmos Machine Learning. Siendo esta, una manera de capturar información y generar herramientas que permitan reducir significativamente los costos de analizar y explorar la información contenida en los textos; produciendo conocimiento con el menor esfuerzo posible al elaborar resúmenes sobre los documentos analizados e incluso de generar respuestas en base a preguntas formuladas en lenguaje natural. (Liddy, 2001)

Para ello se requiere de un proceso dedicado a la extracción y análisis de la información contenida en las palabras; ponderando el valor de cada palabra de las expresiones que son empleadas de acuerdo al léxico, cuyas reglas y estructura varía dependiendo al uso que se le da al mensaje, el cual puede ser usado tanto para la interacción como para la transmisión de la información, la cual puede encontrarse distintos canales; donde cada canal posee sus propias características condicionando la estructura de la información y la eficiencia en su transmisión, así como los costos de la transmisión.

El análisis de sentimientos permite a las empresas medir numéricamente, la información contenida en los distintos canales, analizando de esta forma las opiniones que el público posee sobre algún objeto en particular, las interacciones que comparte con los demás agentes y el feedback de sus experiencias; evaluando en gran volumen y al unisonó: las encuestas de opinión, los comentarios expresados en las redes sociales (sean: Twitter, Facebook, etc.), sitios de reseñas, blogs, chatbots, e-mails y diversos documentos electrónicos. Obteniendo con ello, un gran cúmulo de conocimiento que permite analizar detalladamente las características, los intereses y las emociones de cada individuo; además de poder

observar las condiciones del ambiente en el cual se desarrollan dichas experiencias, describiendo en gran medida, la personalidad y las actitudes de los sujetos analizados.

El análisis de la información que recorre la internet, tal como las noticias, blogs informativos y el intercambio en las redes, permiten obtener el panorama y las condiciones del ambiente; observando y analizando las distintas fuentes de información que los agentes consumen, siendo que estas fuentes de información son sumamente importantes en el comportamiento (la toma de decisiones) de los individuos. Por estos medios es transferido el conocimiento a los agentes, con un bajo costo, debido a que el agente adquiere dichos conocimientos sin la necesidad de la experimentación o la indagación. (Lak & Turetken, 2017)

Por su parte, las expresiones de los clientes tales como las opiniones, reseñas de los productos y servicios; revelan algunas características del cliente, tal como la forma de pensar del individuo, el conocimiento que posee, el provecho obtenido por el servicio ofrecido y de las prestaciones del producto; de esta forma, puede ser determinado los factores que explican la retención del cliente y anticipar su abandono y sus motivos (factores que ayudan a explicar el valor obtenido del CVT).

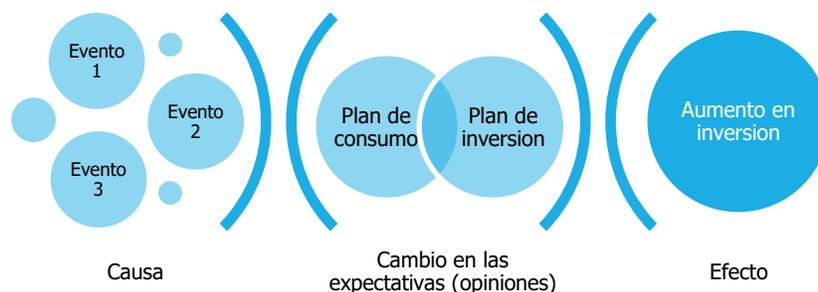
Mediante el empleo del Text Mining (Hambury & Nopp, 2015), se han podido analizar los reportes anuales de diversas empresas: catalogando estos reportes bajo el nivel de riesgo que se poseía, ante la evaluación de las actitudes individuales de las compañías ante el riesgo, etiquetándolas como: {Positivo, Negativo, Incierto}, al analizar los comentarios contenidas dentro de la carta del CEO de cada reporte. Se obtuvo una serie de datos numéricos (estructurada), con la cual podían observar numéricamente la evolución de las actitudes de las compañías en los distintos años. Esto, con la finalidad de poder observar la tendencia de las distintas organizaciones mediante el empleo de regresiones lineales simples.

De esta manera fue posible observar y poseer índices numéricos, que permitiesen medir los riesgos. Las cuales fueron analizadas individualmente, comparando las evaluaciones de actitudes y reacciones ex-ante de una toma de decisiones. Obteniendo con esta tecnología las expectativas sobre los planes de inversión desarrollados por las compañías y de los planes de consumo de los agentes. Al emplear este tipo de tecnología, algunos bancos han podido elaborar promociones “cross-selling” (ofertas de bienes complementarios de un artículo adquirido previamente) al localizar tanto las tendencias del mercado como la utilidad del bien adquirido (Perales-González, 2017)

Al Analizar las expectativas individuales de cada uno de los nodos, dan como resultado, un análisis de las causas previas a los efectos derivados de las decisiones tomadas por el agente. Ejemplo: el aumentar la inversión de la compañía es el efecto, que fue causado por la oportunidad de negocios, el cual es efecto de un cambio en la tendencia general del mercado a causa de la aparición de ciertos eventos. De esta manera, se puede obtener una herramienta que es capaz de analizar los posibles comportamientos futuros y medir la prudencia de los agentes al tomar tales decisiones.

Poseyendo de esta manera, una alerta temprana sobre los riesgos idiosincráticos y de los riesgos sistémicos generados por el ambiente y la entropía de la colectividad de las decisiones individuales.

Ilustración 9. Sucesión de causas-efectos del ejemplo mencionado. Elaboración propia.



Mediante la acumulación del conocimiento de las personalidades de los nodos y de los eventos del ambiente (Ej. Macroeconómicos), se puede determinar las actitudes que los individuos pueden tomar ante la aparición de ciertos eventos;, con lo cual puede determinarse la probabilidad de una elección adversa que podrá incurrir el nodo observado que posee determinadas características; tal como las actitudes que se toman en un ambiente altamente nocivo o las que se toman ante la presencia de un exceso de confianza ante los resultados positivos que fueron obtenidos previamente bajo determinadas condiciones del ambiente, que ya obsoletas.

**Tabla 4. Investigaciones, con aplicación del análisis de sentimientos.
Elaboración propia.**

| Título | Autor | Objetivo | Año |
|--|--|--|----------------------------------|
| "A Sentiment-Aware Model for Predicting Sales Performance Using Blogs" | <ul style="list-style-type: none"> • Yang Liu • Xiangji Huang • Aijun An • Xiaohui Yu | Crear modelos en base de algoritmos que obtengan la suficiente información para predecir el nivel de ventas. Empleando los sentimientos recuperados de los blogs. | 2007; Toronto Canadá |
| "Tweets Sentiment and Their Impact On Stock Market Movements" | <ul style="list-style-type: none"> • Johan Bollen • Huina Mao • Xiao-Jun Zeng | Buscar la relación entre el ánimo de los comentarios en Twitter y el impacto en el precio de acciones. Existiendo el poder de predicción de dichas emociones, su implementación no sustituye las herramientas existentes, las complementa. | 2010; Reino Unido, Manchester |
| "Identifying and Following Expert Investors in Stock Microblogs" | <ul style="list-style-type: none"> • Roy Bar-Haim • Elad Dinur • Ronen Feldman • Moshe Fresko • Guy Goldstein | Similar al trabajo de Bollem; se busca dentro del Twitter opiniones sobre el mercado de valores, en este trabajo, se buscan solamente los usuarios con experiencia e influencia. | 2011; Jerusalén, Israel |

En estos trabajos, se exploran las capacidades del Text Mining para el análisis del ambiente, empleando el Análisis de Sentimientos para predecir el comportamiento futuro de los agentes en base a la información contenida en la red; Es analizando en función de los comentarios y reseñas (Ej. de las películas) y su capacidad de determinar el nivel de ingreso esperados, Así también, es analizado las opiniones y perspectivas que el público posee sobre el mercado bursátil, permitiendo obtener

indicios y modelos que mejoren las capacidades de predecir el comportamiento futuro, al analizar la fuente de dichos movimientos (los inversionistas)

Los algoritmos que son empleados han ido evolucionando para superar los retos de analizar la información no estructurada y obtener de ella conocimiento útil. Su función se centra en el analizar los distintos niveles del texto (palabra, oración, párrafo). Uno de estos algoritmos básicos o simples es el Dictionary-based approach, (también conocido como Bag-of-words o "BoW"), los algoritmos del tipo "bolsa de palabras" o también conocidos como "enfoque basado en diccionario", del cual, su función radica en la ponderación de las palabras al otorgarles valores que son arbitrariamente otorgados, de esta forma es decidido arbitrariamente lo "bueno" y lo "malo".

Con dichos valores asignados, el algoritmo busca las palabras y transforma en número las palabras, lo cual, al otorgar cierta estructura, puede ser empleada para obtener el peso emocional positivo o negativo de las oraciones y poder observar su trayectoria en el tiempo.

Tabla 5. Tabla Resumen del análisis de sentimientos.

| Objeto de estudio | Información | Resultado |
|------------------------------|---|---|
| Cliente Institucional | <ul style="list-style-type: none"> • Reportes • Reseñas de sus clientes • Actividad en web (comunicación con el público) | Se puede analizar el comportamiento de la empresa, al comprender y dar seguimiento al desarrollo de sus planes de inversión. [Determinar el riesgo de proyecto de una empresa] |
| Cliente Físico | <ul style="list-style-type: none"> • Actividad en web del cliente específico • Encuestas de opinión • Mensajes o quejas | Elaboración de un perfil, en el cual se puedan deducir los comportamientos habituales y futuros de los planes de consumo del cliente. [Determinar el riesgo de incumplimiento] |
| Mercado | <ul style="list-style-type: none"> • Noticias • Actividad en general del publico | Seguimiento de eventos específicos (dentro de los grupos primarios), los cuales puedan repercutir a los clientes. [Determinar la existencia del mercado] [Identificar nuevas necesidades singulares u oportunidades de negocio] |
| Ventaja General | Se poseen siempre resultados dinámicos, y siempre actualizados; además de la capacidad de evaluar instantáneamente la información en forma de texto, como los informes. | |

3.1.2 “Scoring” en base a información alternativa

Los bancos emplean mecanismos que les permiten seleccionar a los mejores sujetos con la mejor capacidad y voluntad de pagar el crédito otorgado. Con el constante aumento de los aplicantes de créditos, son requeridas herramientas que reduzcan la fricción de las transacciones y minimicen los costos de analizar la información presentada por el cliente. Obteniendo con estas herramientas un puntaje (o score), el cual ayuda a la colocación de los recursos, buscando la recuperación de su valor y minimizar la pérdida por la selección adversa; Para ello se opta por emplear distintas herramientas estadísticas, tales como el Análisis Discriminante Lineal (LDA), que permite clasificar si el cliente es “bueno” o “malo”, en base a la información recolectada en los cuestionarios y entrevistas con este. (Bessis, 2002)

De esta forma, se posee un trade-off entre la información que es recolectada y la fricción creada por los procesos de recolección y análisis de esta, es decir, a medida que se tiene una mayor cantidad de información se obtiene mayor certeza (y por lo tanto la incertidumbre epistemológica es reducida) pero al mismo tiempo se poseen mayores costos en el análisis de la información y los costos derivados del tiempo dedicado a ello. Estos costos crean una barrera de entrada a los clientes, cuya decisión de adquirir “X” producto se reduce al poseer gran cantidad de requisitos (o costos) en su aplicación.

Para disuadir de alguna probabilidad del riesgo moral (de la veracidad y legalidad de la información) y reducir las fricciones del papeleo; se han creado iniciativas para crear identificaciones virtuales en base a blockchain; dicha identificación puede acarrear la información necesaria la cual es actualizada de forma automática al generar nuevos documentos asociados a la identidad del individuo. (Deloitte, s.f.). El cual puede ser empleado para reducir los costos en los procesos KYC (Know Your Customer). Aun así, se mantienen los riesgos de

contraparte y los costos de oportunidad (de rechazar potenciales clientes "buenos") al no conocer la capacidad y voluntad de pago del cliente. El "Credit Scoring", es para el banco una herramienta de discriminación, la cual reduce la probabilidad del riesgo, minimizando el impacto al evadir clientes que potencialmente ocasionarían pérdidas a la firma; Un scoring rígido o altamente adverso al riesgo, representa una barrera al acceso al financiamiento a agentes que son invisibles, como las Pymes e individuos sin historial crediticio.

En el 2015, de acuerdo con la encuesta de INEGI en México, el 90% de las micro empresas y el 73% de las pequeñas empresas no pueden acceder al financiamiento; debido a que no se poseía colateral o aval, no se tenía historial crediticio o forma de comprobar sus ingresos; En la Unión Europea, 300 instituciones de banca comercial afirmaron que 87% de las aplicaciones fueron rechazados debido a que no pudieron ser calificadas por la escasa información que se posee del cliente. (Henry & Morris, 2018)

Así también, las empresas que no poseen una calificación, deben asumir altas tasas de interés o recurrir a préstamos clandestinos de alto riesgo; Generando desequilibrios en la liquidez del deudor debido la elevada carga de los intereses y el corto beneficio que puede presentar el crédito. Causando pérdida de valor a los deudores afectados, disminuyendo de ello el flujo financiero que afecta al grupo primario que interactúa directamente a ellos (proveedores, trabajadores, compradores, etc.), contagiando entrópicamente a los distintos eslabones de la cadena productiva.

Sugerido por la organización "European MicroFinance" (Henry & Morris, 2018): La inclusión de los clientes que nunca han poseído un historial crediticio y que poseen muy poca información sobre sus actividades o situación financiera, consta en crear nuevos modelos de discriminación, los cuales posean "información alternativa". ya que, por sus características, la actividad de las pequeñas y micro

empresas se encuentra estrechamente ligado a las capacidades y hábitos del dueño, estas pueden ser evaluadas mediante las características del dueño y/o asociados involucrados.

Como se ha mencionado previamente, las tecnologías que permiten analizar gran cantidad de información (Big data), mejora las capacidades analíticas de la entidad o agente, permitiendo analizar fuentes de información que difícilmente pueden ser digeridas; Al poseer modelos que miden y cuantifican los riesgos en base a las características del comportamiento del consumidor, se poseen la cualidad de ser más asertivo en la colocación de los créditos, al observar de cerca el carácter y las consecuencias de sus acciones dentro del sistema. (Brett, 2014)

De esta forma, los intermediarios tienen la opción de realizar acciones en el sistema, el cual brinde incentivos, con tal de inducir positivamente el comportamiento del cliente; Por ejemplo: Una empresa de bicicletas que se encuentra endeudada por el banco y tiene problemas de liquides, el banco podría inducir la liquides de su mercancía mediante promociones (como Meses sin Intereses) y créditos al consumo, dotando a la empresa de ingresos suficientes para solventar sus deudas.

Al incorporar información alternativa, no solamente se analiza el estatus financiero de la entidad observada, sino también pueden ser observadas las características del grupo primario (nodos que interactúan directamente con la entidad observada) y los rasgos psicológicos que determinan el comportamiento del individuo. Siendo información útil, debido que la capacidad de pagar el crédito se encuentra determinada por factores por financieros y la voluntad de pagar está definida por factores psicológicos.

Mediante un experimento elaborado en la Universidad de Texas, Dallas. (Arya, Eckel, & Colin, 2013) Se recolecto información sobre el comportamiento;

evaluando la correlación entre el Score del crédito y los aspectos conductuales, como: el Ímpetu, la preferencia en el tiempo (o paciencia) y la honradez personal. En el experimento, estas medidas psicológicas dieron como resultado, positivas capacidades de discriminación en base al cuestionario psicométrico que fue elaborado, siendo los resultados comparados con scores obtenidos con métodos tradicionales.

Para poder aplicar los métodos de evaluación psicológica, deben de ser tomados en cuenta la información del mercado y del clúster en el cual el cliente pertenece, ya que, por las diferencias culturales, no permiten obtener generalizaciones absolutas en los modelos que emplean este tipo de información.

Las Startups como Kreditech, Lendo o СоцСкор (Socscor); emplean distintas herramientas móviles, para la recopilación de información de compras recientes en e-commerce (como en Amazon), las conexiones en red sociales (en el caso de Rusia Vkontakte o VK) extrayendo de estas la información del perfil, tal como: tiempo de perfil, ultima conexión, estado marital, edad, posts y contenido. (МАЦИУТИН, 2015) Obteniendo de estas distintas señales en los niveles de egresos (el cual debe de estar emparejado y ser coherente a su nivel de ingresos) y señales conductuales, al analizar y clasificar los "post" que son publicados por el agente, los que son seguidos y sus interacciones con otros agentes.

Entre otros ejemplos interesantes, se ha buscado realizar modelos de Scoring en base al "Huella Digital" (Berg, Burg, Gombovic, & Puri, 2018), es decir el registro de actividad en línea durante la aplicación o compra de un producto: tal como la ortografía en los formularios, tipo de correo usado, características del dispositivo usado para navegar, hora en la que se hizo la navegación, proveedor del internet, etc. Como se había descrito previamente, mundialmente, el 50% de la población pose algún canal de acceso al internet; de tal forma que el poseer un conducto por el cual hacer uso del internet puede ser considerada como una

variable proxy de bienestar y/o nivel de ingresos, Tal como otras variables, como el uso de correos institucionales, la velocidad de carga de la página web que pueden dar más señales a la capacidad adquisitiva del cliente.

Este tipo de variables provenientes de la información alternativa, por sí mismas, poseen una capacidad de descripción teóricamente cuestionable; Pero su utilidad no es mutuamente excluyente a los métodos tradicionales, ya que los resultados obtenidos pueden ser complementados con los métodos tradicionales.

4 Conclusiones

De acuerdo, a como se había planteado en el objetivo de esta tesis, se ha hecho una exploración de que los avances tecnológicos en las tecnologías de la información. Aunque no se posee una medida cuantitativa del fenómeno, ya que la información que las compañías comparten sobre sus procesos internos es limitada. No obstante, se puede intuir cualitativamente, que su implementación ha creado beneficios a las compañías quienes la han incorporado en su toma diaria de decisiones.

En los ejemplos aquí descritos, sugieren que dicha tecnología ha creado mejoras en como el intermediario financiero interactúa con los agentes, tanto los clientes como sus deudores. En los últimos años, se ha incrementado el uso de la banca digital, y a experiencia propia, los procesos del banco BBVA se ha vuelto un proceso mas amigable para la aplicación de microcréditos, y de productos financieros tal como la apertura de una cuenta de debito y de la contratación de una tarjeta de crédito de bajo volumen, denominada como "crea", ya que dicho producto está destinado para jóvenes y que permite "crear" un historial crediticio con la entidad, por medio de la interacción. Tales productos se han contratado sin la necesidad de acudir a la sucursal.

Así mismo, se sugiere que las nuevas posibilidades analíticas, permite aprovechar y crear oportunidades comerciales; tal como el modificar el entorno y la cultura social, tal como en el caso de banco Tinkoff en Rusia, el cual supero adversidades en el ambiente, para propiciar uno mejor con el empleo de sus invocaciones en cuanto a servicios, convirtiéndose en un banco importante de rusia, siendo frenado por los conflictos atañidos por su nacionalidad. Así mismo, de acuerdo a algunos comentarios en el portal "startpack.ru" (servicios en nube, para empresas) sus aplicaciones poseen en general calificaciones favorables, sin

embargo, algunos clientes mantienen desconfianza y una cultura conservadora en cuestión al uso de sus datos y de su información.

Tal cultura conservadora, limita la información que el banco puede obtener y así mismo, en cuanto a la teoría; limitaría las capacidades analíticas del intermediario para hacer frente a la administración de los riesgos. En resultado, el completo anonimato, daría como resultado el tratamiento homogéneo de cada cliente, por lo cual no sería una característica "justa", ya que los servicios y el precio no irían de acuerdo al cliente, es decir. El rendimiento que el intermediario deberá de aceptar será uno de mayor riesgo, ya que el elevado precio estaría promediado sujeto únicamente a las condiciones del ambiente mas no en las habilidades del agente de superar la adversidad o aprovechar las condiciones generales de dicho ambiente.

Tales culturas conservadoras, en cuanto a la privacidad, han prevalecido debido a cambios en los términos y condiciones de algunos productos, tales como el aviso de cambio de privacidad de la aplicación de "WhatsApp". Siendo la privacidad alguna especie de bien, que es valorizado por la sociedad; dificultando la observación de los agentes, quienes mantiene el empleo de mercados y/o herramientas anónimas, tal como el uso de buscadores alternativos tipo "duckduckgo" o navegadores como "tor", quienes aseguran el mantener la privacidad.

En algunos casos como en el de we.trade y el de Ali-baba (con ant-finance), la tecnología le permite maximizar el valor de sus activos, y a que dicha innovación, influye en la economía para crear un propicio mercado económico real. Permitted crear canales financieros y canales de comunicación: Al implementar tecnologías de herramientas automatizadas, los costos de crear y mantener canales de comunicación, es reducido; ya que, en el caso de we.trade, dichos canales son mantenidos por los mismos clientes quienes emplean el servicio y categorizan a

sus compañeros mediante el uso de ratings, que brinda a los demás agentes información útil para el manejo de sus propios riesgos. Es decir, confiar o no en el proveedor o cliente que se busca para realizar transacciones económicas.

En el caso de bankamerideals de Bank of America, a empleado la tecnología para mejorar su capacidad analítica, con lo cual le ha permitido obtener y acumular conocimiento sobre los clientes, realizando mejoras en los productos ofrecidos, mejorando las capacidades de personalización y adaptación a las necesidades del cliente. En este caso en particular, el bank of america, mejora las ofertas ofrecidas en "cashback", con lo cual permite ofrecer una selección delimitada por las conductas de los clientes. Que, al mismo tiempo, le permite al banco elegir efectivamente las compañías con las que se realizan convenios, para su inclusión en dichas ofertas, manteniendo conexiones "sanas" tanto con el cliente que emplea el servicio como con las empresas con las que se tienen dichas asociaciones estratégicas.

Bajo la visión del Banco Real de Escocia (RBS), el empleo de esta tecnología le ha ayudado a implementar la "personología" del servicio bancario. Lo cual le permite crear una relación más estrecha con el cliente, dando la oportunidad de no solo de discriminar o negar a los sujetos de crédito, sino que tiene la oportunidad de establecer esquemas de beneficios que permitan mejorar la conducta crediticia del cliente, dando la posibilidad de mejorar su score. Así mismo, al analizar y conocer plenamente las necesidades del cliente y sus características, los bancos pueden conocer en forma amplia sus necesidades y explicar las actitudes, con lo cual, el banco puede adaptar sus funciones, manteniendo una ventaja competitiva y manteniendo un servicio o la experiencia de los distintos clústeres de clientes.

Debido a que el "scoring" y los métodos internos de discriminación se mantienen como secretos empresariales, es prácticamente imposible describir cualitativamente los efectos de la aplicación de la tecnología, en estos aspectos.

Sin embargo, desde el aspecto teórico: las herramientas como el análisis de los textos, (Sentiment Analysis), los bancos pueden evaluar la como información no estructurada (como los textos, imágenes de video y audio). Dándole la capacidad de “leer” o de resumir inmediatamente esta clase de información; la cual es necesaria para poder realizar un análisis fundamental a profundidad, permitiendo mejorar los resultados de dichos análisis, al poder compilar y revisar en menor tiempo distintos fragmentos de la realidad.

Por lo que el campo de visión de los intermediarios es ampliado, permitiéndole cuantificar numéricamente información que tomaría décadas analizarla, como en el ejemplo de personología del RBS en los 70’s. Esta información ahora puede ser tomada y ser introducida en modelos estadísticos, transformando la información no-estructurada en información estructurada). los cuales pueden ser empleados para predecir el comportamiento de los agentes, obteniendo señales sobre los factores de riesgos ubicados en el ambiente y los posibles efectos que tendrá sobre los agentes afectados al analizar simultáneamente las opiniones de los agentes y las noticias que se encuentra circulando.

También, es posible incluir gran diversidad en la información al emplear aquella a la que se le denomina como “alternativa”; con las capacidades de los procesadores actuales, es quizás factible crear un modelo en conjunto, que permitan emplear toda la información disponible. Con lo cual se complementaría la información tradicional y ampliaría dichos métodos para obtener scores, de agentes que no poseen suficiente información financiera, mejorando las capacidades de clasificación, permitiendo penetrar mercados que no habían sido alcanzados con los métodos tradicionales. Evaluando las características conductuales y del ambiente, los cuales permiten mesurar las cualidades del cliente, que determinan la voluntad de retribuir los créditos y la capacidad de pago.

Pese a los beneficios que pueden ser obtenidos de la información alternativa, el poseer este tipo de información produce distintas complicaciones en su proceso; por lo cual se requieren de algoritmos sofisticados para poder lidiar con el ruido producido por las fuentes cuya veracidad no puede ser comprobada; adicionalmente, por el inmenso flujo de información, los errores en los registros afectan significativamente los modelos elaborados, afectando a las acciones basadas con los resultados obtenidos con errores, los cuales difícilmente pueden ser localizados por la magnitud de la información.

Por el proceso tipo "Black Box", en los procesos de selección de crédito y Scoring; los consumidores no poseen la certeza con la información con la que están siendo medidos sus parámetros. Ya que al brindar el cliente información privada, el intermediario posee una ventaja y control sobre el cliente. De tal forma que puede existir un comportamiento oportunista por parte del intermediario, siéndole posible localizar los clientes desesperados por recursos, quienes aceptaran tasas de interés muy elevadas por obtener recursos urgentemente. De tal forma que los precios dinámicos, pueden propiciar precios monopólicos, reduciendo el excedente del consumidor y absorbiendo los recursos de la economía ante la succión de estos de forma parasitaria.

De tal modo, se sugeriría la creación de regulaciones o métodos que permitieran conocer como los datos son empleados para la discriminación de precios y/o de los créditos otorgados; ya que como se ha comentado con la información de discriminación ética (sexo, raza o religión), un dato es empleado "justamente", dependiendo al uso que se le dé; es decir, cuando el cliente mantiene un beneficio esperado positivo al exponer el valor de la privacidad.

5 Referencias

- Acevedo Balcorta, J. A. (2012). *El sistema Bancario Mexicano*. México: Alzar.
- Airbnb. (2018). *Customized Regression Model for Airbnb Dynamic Pricing*. London, United Kingdom: KDD (Applied Data Science Track Paper).
- Allen, F., & Gale, D. (1998). *Innovations in Financial Services, Relationships and Risk Sharing*. Pennsylvania, USA: Whartoon School.
- Allen, F., & Santomero, A. (1998). *The theory of Financial intermediation*. Journal of Banking & Finance.
- Arrow, K. (1973). *Information and Economic Behavior*. USA, Cambridge: Harvard University.
- Arya, S., Eckel, C., & Colin, W. (2013). *Anatomy of the Credit Score*. Dallas, TX: MPRA.
- Baldan, C. &. (2012). Liquidity risk and interest rate risk on banks: are they related? *The IUP Journal of Financial Risk Management, Forthcoming*.
- Ballard K, K. B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage*, 143-50. doi:10.1016/j.neuroimage.2008.11.004
- Bechara, A. &. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and economic behavior*, 336-372.
- Berg, T., Burg, V., Gombovic, A., & Puri, M. (2018). *On the Rise of the FinTechs – Credit Scoring using Digital Footprints*. Berlin, Alemania: Federal Deposit Insurance Corporation.
- Bessis, J. (2002). *Risk Management in Banking*. West Sussex, England: John Wiley & Sons.

- Blattberg, R., Malthouse, E., & Neslin, S. (2009). Customer Lifetime Value: Empirical Generalizations and Some Conceptual Questions. *Interactive Marketing*, 157 - 168.
- Blumberg, R., & Atre, S. (2003). *The Problem With Unstructured data*. DM Review.
- Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*.
- Campbell, Tim S., & William A. Kracaw. (1980). Information Production, Market Signalling, and the Theory of Financial Intermediation. *The Journal of Finance*, JSTOR.
- Capgemini Consulting. (2013). *Big Data Alchemy: How can Banks Maximize the Value of their Customer Data?*
- Capgemini Consulting. (2014). *Cracking the Data Conundrum: How Successful Companies Make Big Data Operational*.
- Capgemini Consulting. (2014). Gearing Up for Digital Operations. *Digital Transformation Review*.
- Capgemini Consulting. (2017). *The Big Data Payoff: Turning Big Data into Business Value*.
- Chester, J., & Mierzwinski, E. (2014). *Big Data Means Big Opportunities and Big Challenges*. USA: Center for Digital Democracy - Federal Trade Commission.
- Cielen, D., Meysman, A., & Ali, M. (2016). *Introducing Data Science*. Shelter Island, NY: Manning.
- Coase, R. (1937). *The Nature of the Firm*.

- Columbus, L. (24 de December de 2017). *53% Of Companies Are Adopting Big Data Analytics*. Obtenido de Forbes: <https://www.forbes.com/sites/louiscolombus/2017/12/24/53-of-companies-are-adopting-big-data-analytics/#2a11fd1e39a1>
- Computing Community Consortium. (2012). *Challenges and Opportunities with Big Data*.
- CRISP-DM consortium. (2000). *CRISP-DM 1.0*. SPSS.
- Dawkins, R. (2016). *he extended phenotype: The long reach of the gene*. Oxford University Press.
- Day, J. J. (2007). The Nucleus Accumbens and Pavlovian Reward Learning. *The Neuroscientist*, 13(2), 148-159. doi:10.1177/1073858406295854
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). *A Formal Definition of Big Data Based on its Essential Features* (Vol. 65). Cassino, Italy: Library Review.
- De Mauro, A., Greco, M., & Michele, G. (2015). *What is Big Data? A Consensual Definition and a Review of key research topics*. AIP Publishing.
- Demsetz, H. (1969). *Information and Efficiency: Another Viewpoint*. Chicago, USA: The University of Chicago.
- Diamond, D. (1996). *Financial Intermediation as Delegated Monitoring*. Federal Reserve Bank of Richmond Economic Quarterly.
- Duffie, D. (2008). *Innovations in credit risk transfer: implications for financial stability*. Basel, Switzerland: Bank For International Settlements.
- EIOPA. (2019). *Big Data Analytics in Motor an Health Insurance*. Luxemburgo: Office of European Union.

- Eldering, C. (2006). *United States of America Patente nº 7,062,510 B1*.
- Escott Manga, L., Olmedo Fernandez, E., & del Pozo Garcia, E. (2002). *Optimizacion Dinamica*. España. Madrid: Universidad Complutense.
- Evry. (2013). *Big data in banking for marketers*.
- Ferguson, M., & Bargh, J. (2004). *How Social Perception Can Automatically Influence Behavior* (Vol. 8). TRENDS in Cognitive Sciences.
- Flage, R., & Aven, T. (2009). *Expressing and Communicating Uncertainty in Relation to Quantitative Risk Analysis*. University of Stavanger, Stavanger, Norway.
- Fletcher, H. (1 de February de 2014). *Bank of America's BankAmeriDeals retains customers and their trust*. Recuperado el 30 de November de 2019, de Target Marketing: <https://www.targetmarketingmag.com/article/bank-america-s-bankamerideals-retains-customers-their-trust/all/>
- Foxall,, G. (1993). *Consumer Behaviuor as am Evolutionary Process* (Vol. 27). Birmingham, UK: Emerald Insight.
- Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). *Time Discounting and Time Preference: A Critical Review*. Journal of Economic Literature, Vol.XL.
- Freud, S. (1940). An Outline of Psycho-Analysis. *The International Journal of Psychoanalysis*, 27-84.
- Gantz, J., & Reinsel, D. (2012). *The Digital Universe In 2020; Big Data, Bigger Digital Shadows*. EMC Corporation.
- George, A. S. (2023). A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal*,, 9-23.

- Greenhalgh, C., Hazzard, A., McGrath, S., & Benford, S. (2016). *GeoTracks: Adaptive Music for Everyday Journeys*. Nottingham, UK: The University of Nottingham.
- Grossman, S., & Stiglitz, J. (1976). *Information and Competitive Price Systems* (Vol. 66). *The American Economic Review*.
- Hambury, A., & Nopp, C. (2015). *Detecting Risks in the Banking System by Sentiment Analysis*. Lisbon, Portugal.: Association for Computational Linguistics.
- Hayek, F. A. (1945). The Use of Knowledge in Society. *American Economic Review*.
- Hayek, F. A. (1948). *Individualism and Economic Order*. Chicago, Illinois: The University of Chicago Press.
- Heffernan, S. (2005). *Modern Banking*. West Sussex, England: John Wiley & Sons.
- Helton, J. C. (1996). *Guest editorial: treatment of aleatory and epistemic uncertainty in performance assessments for complex systems*. Northern Ireland: Reliability Engineering and System Safety, Elsevier Science Limited.
- Holland, J. (2010). *Banks, Knowledge and Crisis - a case of knowledge and learning failure* (Vol. 18). (J. o. Compliance, Ed.) Glasgow, UK: University of Glasgow.
- Hurley, M., & Adebayo, J. (2017). *Credit Scoring in the Era of Big Data*. New Haven, Connecticut, USA: Yale, Journal of Law and Technology.
- Informatica. (2018). *Tinkoff Tames the Big Data Bear, Innovates at Lower Cost with Informatica*.

- Jagtiani, J., & Lemieux, C. (2018). *Do Fintech Lenders Penetrate Areas That Are Underserved by Traditional Banks?* Philadelphia, USA: Federal Reserve Bank of Philadelphia.
- John P. O'Doherty, P. D. (2003). Temporal Difference Models and Reward-Related Learning in the Human Brain. *Neuron*, Pages 329-337,. doi:doi.org/10.1016/S0896-6273(03)00169-7
- Kao, D.-L. (2000). *Estimating and Pricing Credit Risk: An Overview* (Vol. 56). (F. A. Journal, Ed.) CFA Institute.
- Kaplan, S., & Garrick, B. (1981). *On the Quantitative Definition of Risk*. Society of Risk Analysis.
- Khajvand, M., & Tarokh, M. J. (2011). *Analyzing Customer Segmentation Based on Customer Value Components*. Theran, Iran: Toosi University of Technology, Journal of Industrial Engineering.
- King, B. (2014). *Breaking Banks; The Innovators, Rogues, and Strategists Rebooting Banking*. Solaris South Tower, Singapore: John Wiley & Sons.
- Kshetri, N. (2016). *Big Data's Role in Expanding Access to Financial Services in China*. International Journal of Information Management.
- Laeven, L. L. (2015). Financial innovation and endogenous growth. *Journal of Financial Intermediation*, 24(2), 1-24.
- Lak, P., & Turetken, O. (2017). The Impact of Sentiment Analysis Output on Decision Outcomes: An Empirical Evaluation. *Transactions on Human-Computer Interaction*, 1-22.
- Levine, R. (1997). Financial Development and Economic Growth: Views and Agenda. *Journal of economic literature*, 35(2), 688-726.

- Levine, R., & King, R. (1993). *Finance, Entrepreneurship and Growth: Theory and evidence*. Journal of Monetary Economics.
- Liddy, E. D. (2001). *Natural Language Processing*. Encyclopedia of Library and Information Science.
- Marr, B. (2016). *Big data in practice: how 45 successful companies used big data analytics to deliver extraordinary results*. John Wiley & Sons.
- MBAMA, C. (2018). *Digital banking, customer experience and bank financial performance : UK customers' perceptions*. Emerald, Journal of Bank Marketing.
- Merton, R. C. (1989). On the application of the continuous-time theory of finance to financial intermediation and insurance. *The Geneva Papers on Risk and Insurance-Issues and Practice*(14), 225-261.
- Morrison, A. (2001). *Credit Derivatives, Disintermediation and Investment Decisions*. Oxford, UK.
- Mudzingiri, C., Muteba Mwamba, J. W., & Nicolaas Keyser, J. (2018). *Incentivized Time Preferences, Level of Education in a Household and Financial Literacy; Laboratory Evidence*. South Africa: Journal of Economics and Behavioral Studies.
- Nelson, R. R., & Winter, S. G. (1982). *An Evolution Theory of Economic Change*. USA, Cambridge: The Belknap Press of Harvard University Press.
- O'Brien, E. (2017). *Omni Channel Business Strategies in Banking*. Mercator Advisory Group .
- Oliver Wayman. (2017). *Alternative Data and the Unbanked*.

- O'Reilly, T. (30 de September de 2005). *What is Web 2.0*. Obtenido de O'Reilly Media: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?page=1>
- Pandian, C. (2007). *Applied software risk management : a guide for software project managers*. Boca Raton, Florida: Auerbach.
- Pawar, A. B., Jawale, M. A., & Kyatanavar, D. N. (2016). *Fundamentals of Sentiment Analysis: Concepts and Methodology*. Switzerland: Springer International Publishing.
- Pfeifer, P., Haskins, M., & Conroy, R. (2004). CUSTOMER LIFETIME VALUE, CUSTOMER PROFITABILITY, AND THE TREATMENT OF ACQUISITION SPENDING. *Journal of Managerial Issues*.
- Rajan, H. (. (19 de Febrero de 2019). *Big-Data Driven Cross-Selling for Retail Lenders*. Obtenido de Medium: <https://medium.com/finbox/big-data-driven-cross-selling-for-retail-lenders-27f00c065e05>
- Ravi, V., & Kamaruddin, S. (2017). *Big Data Analytics Enabled Smart Financial Services: Opportunities and Challenges*. Hyderabad, India: Springer.
- Reed Shiller, B. (2014). *First-Degree Price Discrimination Using Big Data*. USA, MA: Brandeis University.
- Ribeiro, F. C. (2004). *El Pensamiento de Hayek Y La Teoria De La Informacion*. Argentina: Instituto Universitario ESEADE .
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal*, 3.

- Scholtens, B. &. (2003). The theory of financial intermediation: an essay on what it does (not) explain. *SUERF - The European Money and Finance Forum, Vienna*.
- Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience volume, 199–207*.
- Sindwani, R., & Goel, D. (2015). The Impact of Technology Based Self Service Banking Dimensions On Customer Satisfaction. *International Journal of Business Information Systems Strategies*.
- Sohrabi, B., & Khanlari, A. (2007). *Customer Lifetime Value (CLV) Measurement Based on RFM Model*. Theran, Iran: Springer, Iranian Accounting & Auditing Review.
- Stiglitz, J. E. (1981). Credit rationing in markets with imperfect information. *The American economic review, 393-410*.
- Talabis, M. R. (2015). Analytics Defined. Information . *Information Security Analytics, 1-12*.
- Xu, L., Jian, C., Wang, J., Yuan, J., & Ren, Y. (2014). *Information Security in Big Data: Privacy and Data Mining*. China: IEEE Access.
- Zhou, Q., Chen, X., & Li, S. (2018). *Innovative Financial Approach for Agricultural Sustainability: A Case Study of Alibaba*. Shangai, China: MDPI.
- Иванова, И. (2018). Современные Методы Прогнозирования Финансовых Результатов Деятельности Предприятия. Еп Н. К. Гуап, *Международная Студенческая, Сборник Докладов, Часть 4* (págs. 19-23). San petersburgo, Russia: Министерство Образования И Науки Российской Федерации.

Масютин, А. (2015). *Кредитный Скоринг На Основе Данных Социальных Сетей (Credit Scoring, based on social media)*. Russian Federation, Moscow: Business Informatics, No. 3 (33).

Тинькофф Банк. (09 de December de 2019). Tinkoff super-app boosts digital leadership in European fintech space. Moscow, Russia. Obtenido de Тинькофф Банк: <https://www.tinkoff.ru/eng/media/news/09122019-tinkoff-launches-first-russian-super-app-eng/>