

UNIVERSIDAD PANAMERICANA

FACULTAD DE INGENIERÍA

Con estudios incorporados a la
Secretaría de Educación Pública

**“USO DE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO
PARA LA CARACTERIZACIÓN Y MEJORA DE
EXPERIENCIAS DE SERVICIO”**

T É S I S

**QUE PARA OBTENER EL TÍTULO DE
MAESTRÍA EN DIRECCIÓN DE OPERACIONES**

**P R E S E N T A
GUILLERMO E. ESQUIVEL JAIMES**

**DIRECTOR DE TÉSIS:
DRA. MARÍA DE LOURDES GUADALUPE MARTÍNEZ
VILLASEÑOR**

Agradecimientos

A mi madre, padre y hermana por el apoyo y oportunidades que me brindaron.

A la Dra. Lourdes por su paciencia, tiempo, conocimientos y buenas discusiones.

A la Dra. Karina y el Dr. Hiram por su atención al detalle y rigor científico.

A mis amigos y seres queridos; Rodrigo, Javier, Juan Carlos y Samuel por impulsarme y apoyarme.

ABSTRACT

Una profunda desconexión entre las áreas de estrategia de soporte y el área directiva de una empresa produjeron una desalineación en el uso de los recursos de la compañía. Los esfuerzos y recursos para la mejora del servicio se estaban utilizando en dimensiones operativas las cuales no producían mejoras tangibles en la experiencia de servicio. Si bien aprendizaje máquina es comúnmente usado para curar experiencias de servicio hacia usuarios de plataformas tecnológicas altamente maleables, rara vez es utilizado para segmentar y analizar una experiencia de servicio física para hacer ajustes en tiempo real sobre los procesos de un centro de atención, como atender a un cliente primero, o tener agentes especializados en problemas dados. En este trabajo utilizamos un modelo predictivo de *random forest* para entender la relación entre una calificación de satisfacción de un servicio y las variables operacionales utilizadas para cuantificar el sistema. El poder predictivo de cada una de las variables se utiliza como un punto de referencia para modelar las necesidades de los clientes. De esta forma se modela un mapa que visualiza las diferencias entre las prioridades para el negocio y las necesidades para el cliente.

I - ÍNDICE

| | |
|-----------------------------|---|
| Abstracto | 1 |
| i - Índice | 1 |
| 1 - Introducción | 4 |
| 2 - Objetivos e hipótesis | 6 |
| 2.1 - Objetivo general | 6 |
| 2.2 - Objetivos específicos | 6 |
| 2.3 - Hipótesis | 6 |
| 3 - Limitaciones | 7 |
| 4 - Marco teórico | 7 |

| | |
|---|----|
| | 2 |
| 4.1 - Cuantificación de sistemas de servicio | 7 |
| 4.2 - Economía de experiencias | 9 |
| 4.3 - Experiencias en el ciclo de vida del cliente | 10 |
| 4.4 - Servicio al cliente y experiencias de servicio | 11 |
| 4.5 - Técnicas de evaluación para ecosistemas de servicio | 12 |
| 4.6 - Índice de satisfacción del cliente (CSAT) | 12 |
| 4.7 - Net Promoter's Score (NPS) | 13 |
| 4.8 - Otros indicadores de servicio | 15 |
| 4.9 - Experiencias de servicio idiosincráticas | 16 |
| 4.10 - Economías de la información | 17 |
| 4.11 - Información como un servicio | 18 |
| 4.12 - Tratamiento de datos | 20 |
| 4.13 - Muestreo balanceado | 21 |
| 4.14 - Tokenización | 21 |
| 4.15 - Dummy variables | 22 |
| 4.16 - Gestión de desempeño | 22 |
| 4.17 - Aprendizaje automático | 23 |
| 4.18 - Técnicas de selección de variables | 24 |
| 4.19 - Validación de modelo | 26 |
| 4.20 - Grados de libertad | 26 |
| 4.21 - Sistemas de toma de decisión | 26 |
| 4.22 - Uso de técnicas de estadísticas avanzadas para la optimización de experiencias. | 27 |
| 5 - Caso práctico; implementación de técnicas de aprendizaje automático para la mejora de la experiencia de servicio. | 27 |
| 5.1 - Situación actual | 28 |
| 5.1.a Problemas estructurales en la dirección. | 29 |
| 5.1.b Iniciativas anteriores | 31 |
| 5.2 - Estudio de centros de soporte en persona | 32 |
| 5.3 - Descripción de la metodología | 35 |
| 6 - Implementación de la metodología | 37 |
| 6.1 - Mapeo de los puntos de lectura | 37 |
| 6.2 - Agrupación las variables según su relación con aspectos del negocio | 38 |
| 6.3 - Segmentar las variables en variables primarias y secundarias: | 39 |
| 6.4 - Identificar la variable de salida del sistema | 40 |
| 6.5 - Hacer una priorización de los esfuerzos actuales | 41 |

| | |
|--|----|
| 6.6 - Análisis exploratorio | 43 |
| 6.6.a Base de datos y tratamiento | 43 |
| 6.6.b Demográficos | 44 |
| 6.6.c Distribuciones | 44 |
| 6.6.d Estudio correlacional | 48 |
| 6.7 - Optimización de variables a través de la iteración de un modelo simple | 52 |
| 6.7.a - Validación y balanceo del set de datos | 53 |
| 6.7.b - Integración de variables secundarias | 55 |
| 6.7.c - Simplificación de variables de salida | 57 |
| 6.7.d - Configuración y selección final de las variables | 58 |
| 6.8 - Selección de modelos | 58 |
| 7.8.a - Modelo final | 60 |
| 6.9 - Evaluar el poder predictivo de cada variable: | 60 |
| 6.10 - Sensibilización de las variables con expertos en el sistema: | 61 |
| 6.11 - Contraste contra las prioridades del negocio | 64 |
| 6.12 Utilizar las zonas de desenfoque para desarrollar mejoras | 66 |
| 7 - Implementación de las recomendaciones | 66 |
| 8 - Discusión | 71 |
| 9- Conclusiones | 73 |
| 10 - Recomendaciones para trabajos futuros | 74 |
| Referencias | 76 |

1 - INTRODUCCIÓN

Las experiencias son un diferenciador crucial ante los competidores. Conforme migramos hacia una economía de servicios, la percepción de calidad de un cliente se vuelve un factor crítico para la retención del mismo (Chandok, 2014). La hospitalidad, servicio al cliente y entrega de servicio son ahora diferenciadores y puntos críticos en la selección de un cliente ante un mercado competido (Collier, 2018).

Estos nuevos diferenciadores suelen ser procesos con alto involucramiento humano. Consiguientemente, debido a su naturaleza humana, son retadores en su estudio, observación, cuantificación y mejora. La gestión de procesos humanos suele tener sesgos personales de los individuos que diseñan los procesos y de la misma manera la cuantificación es subjetiva cuando la salida es una percepción humana (Weech-Maldonado, 2008).

A pesar de las dificultades que conlleva mejorar la percepción humana, si esta es vista como un conjunto agregado, es mucho más sencillo entender la perspectiva general y englobarla en una problemática específica, modelable y cuantificable. Se presume que, al acotar la problemática a un sistema específico, es posible encontrar una solución sistemática y llevarla a escala (F. Frei, 2018).

Adicionalmente, la evolución de los sistemas de gestión de datos nos ha permitido capturar y almacenar grandes volúmenes de datos. Los grandes volúmenes de datos capturados pueden ser transformados para modelar el sistema experiencial de forma concreta y con menos sesgos (Hartung 2018).

En este trabajo analizaremos las fallas en la priorización de una empresa de servicios y su área de soporte a través de su vasta recolección de datos.

Se presume que la estrategia actual dista diametralmente de las necesidades de los clientes. Pues la implementación de las estrategias no tiene efecto positivo en las calificaciones de satisfacción de los clientes. A pesar de invertir tiempo y esfuerzo, la compañía estudiada no tenía agencia en la mejora de los

indicadores de satisfacción. Cahill sugiere en un estudio similar que la causa de los indicadores de satisfacción estancados puede ser la falta de alineación entre los esfuerzos de la empresa y las expectativas de los clientes (Cahill, 1997).

Pretendemos encontrar una solución que haga uso efectivo de los bancos de datos de la empresa y a través de un modelo inspirado en aprendizaje automático podamos modelar de forma casi automática los factores que se ven reflejados en la satisfacción de los clientes. Lo anterior para dar luz sobre la brecha de prioridades anteriormente explicada.

De manera general se busca encontrar las dimensiones que mayor impacto tienen en la evaluación de la experiencia del cliente en un ecosistema de servicio por medio de técnicas de aprendizaje automático, con el fin de crear estrategias para la mejora del servicio.

Lo anterior se logra a través de una combinación de un análisis exploratorio completo del sistema de soporte donde se definen los componentes principales y una serie de selección de modelos que logra caracterizar las variables críticas para los clientes.

La metodología para la cuantificación de ecosistemas de soporte posteriormente es encapsulada con el objetivo de facilitar la toma de decisiones para su mejora. La metodología propuesta es fácilmente adaptable a cualquier tipo de ecosistema de servicio donde se capturen variables operativas y se evalúen con una encuesta de salida al terminar el servicio recibido.

También se propone un tratamiento de datos para eliminar los factores humanos que sesgan la información en conjunto con el uso de modelos de aprendizaje automático para eliminar el factor subjetivo de nuestros paradigmas de lo que creemos que es apreciado por los clientes (Farris, 2010).

La metodología propuesta fue puesta a prueba pilotando por seis meses las decisiones tomadas con el modelo producido por la misma con resultados positivos. El modelo producido con la metodología propuesta logró dar visibilidad sobre la brecha de prioridades e informar el diseño de los procesos de mejora que elevaron significativamente la satisfacción del cliente.

A pesar que la metodología fue implementada de manera particular, no hay evidencia que demuestre que no pueda ser implementada de una manera generalizada a cualquier sistema clásico de soporte, donde la captura de información se haga de una manera similar. Este trabajo propone incentivar el desarrollo de

técnicas cuantitativas para ramas tradicionales cualitativas pues las tendencias en gestión de la información apuntan a que la toma de decisiones debe ser cada vez más soportada en datos duros.

2 - OBJETIVOS E HIPÓTESIS

2.1 - Objetivo general

1. Encontrar las dimensiones que mayor impacto tienen en la evaluación de la experiencia del cliente en un ecosistema de soporte al cliente, por medio de técnicas de aprendizaje automático con el fin de crear estrategias para la mejora del servicio.

2.2 - Objetivos específicos

1. Establecer una metodología de evaluación de ecosistemas de servicio y definir sus componentes con el fin de entender los estados iniciales del sistema a mejorar
2. Realizar un análisis exploratorio de los factores individuales para entender las propiedades y descriptores estadísticos de cada uno de ellos.
3. Hacer un análisis comparativo de técnicas para encontrar el modelo descriptivo del ecosistema.
4. Desarrollar una matriz de sensibilidad con los distintos factores para encontrar los componentes principales de la percepción de servicio dada con el fin de identificar los elementos críticos para implementar en la estrategia de mejoramiento.

2.3 - Hipótesis

Las técnicas de aprendizaje automático pueden ser usadas para descomponer un ecosistema de servicio en sus factores primarios de los que depende la percepción de la experiencia para definir la estrategia de mejora.

3 - Limitaciones

Es importante notar que este estudio no pretende definir una solución única para la mejora del servicio al cliente, ni tampoco trata de encapsular recomendaciones universales para mejorar los indicadores de satisfacción de cualquier centro de soporte pues este estudio asume medularmente que cada ecosistema de soporte debe de responder a necesidades únicas del cliente. Este ejercicio busca lograr una manera sistemática de mejorar la toma de decisiones para ecosistemas de servicio al cliente, particularmente en servicios de atención y activación en persona (sitios físicos o de retail).

Es también importante notar que el estudio se realizó en un mercado particular (Ciudad de México), y no necesariamente las conclusiones y recomendaciones para el mercado se traducen de manera directa a distintos mercados o servicios distintos. Justamente se busca poder llevar a escala el modelo de toma de decisiones para poder entregar recomendaciones individuales y únicas según el mercado y tipo de servicio.

4 - MARCO TEÓRICO

4.1 - Cuantificación de sistemas de servicio

Cuando un producto o un servicio dada la cotidianidad de su uso se convierte en un *commodity* se vuelve fácilmente reemplazable por la siguiente opción más cercana en el mercado. Los diferenciadores clásicos como la disponibilidad o el precio se vuelven factores de accionables débiles para la retención de los clientes (Allen, 2019).

Se ha demostrado mediante estudios del comportamiento del cliente la correlación entre indicadores de experiencia de servicio y potencial de recompra que la experiencia de uso y el servicio al cliente se tornan como mecanismos de lealtad y retención a los mismos. La retención es deseable ya que es de 5 a 20 veces más costoso adquirir a un nuevo usuario que retener un usuario existente (Frei 2012), (Sabbeh 2018).

Utilizaremos de forma indiferente el término “cliente” o “usuario” pues permite la flexibilidad de referirnos a un producto o un servicio.

El soporte se entiende como la gama de servicios provistos para asistir a los clientes a hacer el uso correcto de un producto o servicio. Estos servicios incluyen: acciones de planeación, instalación, entrenamiento, solución de problemas, mantenimiento, actualización y deshecho de un producto (businessdictionary s.f.).

Se debe estudiar el ecosistema de servicio de un proveedor de un servicio o producto. Por definición, el servicio al cliente se entiende como el proceso del aseguramiento de la satisfacción de un producto o servicio; regularmente en la etapa de transacción del cliente, o posteriormente. El *servicio al cliente* puede ser llevado a cabo en persona, por teléfono, sistemas de autoservicio u otros métodos (Investopedia s.f.).

De manera particular profundizaremos en los canales en persona de puntos de venta o *centros de atención*; por los mismos entendemos aquellas subunidades o departamentos envueltos directamente en la generación de ganancias que proveen servicios de soporte a los clientes o usuarios de un servicio o producto (Investopedia s.f.).

Uno de los principales objetivos como centro de atención es proveer al cliente de una experiencia sin fricciones y resolver la mayor cantidad de problemas del cliente; para esto se requiere cierta visibilidad e introspección que no puede ser obtenida únicamente estudiando métricas operacionales individualmente.

Hay varios indicadores de negocio o *KPIs* (Indicador Clave de Rendimiento, por sus siglas en inglés), utilizados horizontalmente en las industrias para bosquejar la calidad de servicio entregada por los distintos canales de soporte mencionados anteriormente (en persona, teléfono, auto-servicio, etc) (Allen, 2019).

Profundizaremos en cada uno de ellos más adelante para evaluar sus cualidades y carencias; por lo pronto nos enfocaremos en el “*Customer Satisfaction Score*” (*CSAT*, por sus siglas en inglés). Este se captura con una encuesta rápida de salida después de una interacción de servicio (Kim, 2017).

El principio y direccionalidad del *CSAT* radica en tener una medida estándar para la calidad del servicio en una escala de Likert de 5 puntos siendo 1 para un servicio insatisfactorio y 5 para un servicio excepcional. El indicador es meramente perceptivo, por lo cual se busca una muestra lo suficientemente amplia para bosquejar la percepción común de un grupo de clientes (Kim, 2017).

Llevaremos a profundidad el estudio de estos indicadores de servicio y su impacto con la retención de usuarios. Particularmente se busca entender las principales diferencias en las capturas del *CSAT* en centros de atención con procesos estandarizados. El estudio se acota a una región específica, pues existen rasgos culturales que alterarán la forma de evaluación por parte de los usuarios (Weech-Maldonado 2008).

Estos mismos rasgos culturales que pueden ser geográficos, idiosincráticos o sociales afectan de manera lateral la experiencia de servicio y las dinámicas internas. A través de una metodología basada en *aprendizaje automático* buscamos aislar estos efectos (Weech-Maldonado 2008).

4.2 - Economía de experiencias

En la progresión del valor económico, a través del tiempo, primero nos encontramos con los consumibles de extracción o *commodities*, los cuales son procesados para transformarse en *bienes* que a su vez son utilizados para realizar *servicios*; y la diferenciación y elevación de estos servicios los transforman en *experiencias*. Las economías de las experiencias explican el porqué del cual podemos pagar más de 4 veces por una cena en un restaurante de lujo, contra lo que se pagaría en un restaurante de cadena; y 10 veces más contra lo que representaría cocinar la propia comida (Gilmore 1998).

La figura 1 muestra cómo el valor económico aumenta al pasar de la extracción de materias primas hacia la economía de las experiencias.

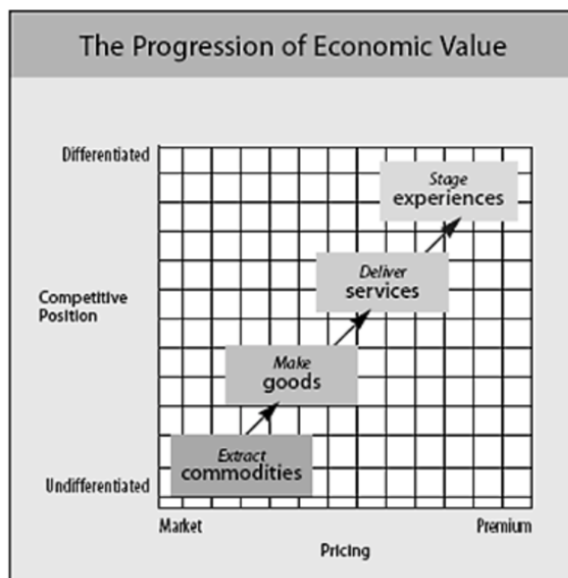


Fig 1. Progresión del valor económico (Gilmore 1998)

Una experiencia se construye a través de impresiones que el cliente puede llevarse consigo post consumo; estos eventos “memorables” son interpretados únicamente por el usuario. En la economía de las experiencias podríamos utilizar un servicio como el conductor hacia una experiencia (Chandok & Gupta, 2014).

La creación de las experiencias trata de acentuar lo que debe de ser memorable y corregir todo aquello que no aporte a la vivencia y expectativas del cliente. Para ello es fundamental el conocimiento pleno del negocio y los puntos a ser apalancados (Chandok & Gupta, 2014) (Collier et al., 2018).

Los aspectos principales de la experiencia son la absorción - inmersión y el grado de participación, y según el enfoque dado a la vivencia, el tipo de experiencia que se forma, podemos hablar de experiencias estéticas dónde se conjugan la participación pasiva y la inmersión o experiencias educacionales en las cuales hay absorción y participación activa. Lo interesante viene al uso de cada una de estas variaciones y el trato o retención del cliente (Collier et al., 2018).

También se cree que las experiencias están limitadas a las industrias del entretenimiento, como parques temáticos y restaurantes; pero ha probado la experiencia que un factor de éxito de negocios como Starbucks, Virgin Mobile, Southwest Airlines es la experiencia que entregan (Collier et al., 2018).

El enfoque de estas compañías ha sido ir más allá del pensamiento *commodity* dónde las empresas se limitan a su función primaria; las empresas que mencionamos tienen un profundo conocimiento de sus clientes y de su producto y convierten cada punto de contacto con sus clientes en una experiencia personal (Collier et al., 2018).

4.3 - Experiencias en el ciclo de vida del cliente

En la actualidad, pocas empresas logran dar el nivel de servicio que es esperado por sus clientes. Dentro del ciclo de vida de un cliente existen cuatro puntos de inflexión básicos donde la experiencia de usuario impacta el futuro éxito de la relación con el cliente; estas son: la fase de identificación o adquisición. En

esta fase se valora cada prospecto de cliente y su probabilidad de consumo del producto o de servicio (Gupta, N. 2014).

Fase de atracción o *lead*: se invita al cliente a comenzar el ciclo de vida del producto; se busca identificar grupos con rasgos similares para la segmentación y direccionamiento de información (Lejeune 2001).

Fase de retención: la retención de un cliente consta en todas las acciones tomadas por una organización para alargar el ciclo de consumo o de uso de un servicio, el objetivo es evitar el abandono; el abandono califica cuando un cliente cambia de proveedor de servicio por uno más competitivo o con mayor calidad o nivel de servicio, se busca reducir la razón abandonó pues adquirir un cliente nuevo al requerir inversión es mucho más costoso que mantenerlo según la industria (Lejeune 2001).

Por último, está la fase de desarrollo, donde se busca maximizar el consumo del cliente para generar mayores ganancias por usuario. En esta etapa se desarrolla la abogacía de marca y se define la futura lealtad del usuario (Lejeune 2001).

Es importante poder identificar la etapa en la que se encuentra cada uno de los usuarios para poder segmentar y caracterizar la experiencia de servicio a las expectativas puntuales que tiene en el momento dado. Según la fase del ciclo de vida es posible implementar distintos niveles de servicio, es decir, que un cliente en la etapa de adquisición requiere un tratamiento distinto a un cliente en etapa de desarrollo pues el valor promedio de ambos es distinto y por ende el retorno que pueda darse en cada uno se ve afectado, por lo que se invierte en ellos (Lejeune 2001).

4.4 - Servicio al cliente y experiencias de servicio

Dentro del ámbito del servicio al cliente existen pocas metodologías para la evaluación de una experiencia de servicio. Dos de los más comunes y aceptados son el “*Customer Satisfaction Score*” (CSAT score) y el “*Word of Mouth*” (WoM), ambos están fundamentados bajo la premisa de que si un servicio resulta bueno será recomendado orgánicamente dentro de una red de clientes (Collier 2018).

A pesar de ser metodologías que si bien logran dar dirección sobre la calidad (o preferencia) del servicio, fallan en demostrar qué elemento del proceso fue el más relevante dentro de la experiencia holística, quitando un poco de visibilidad sobre en qué aspectos se deben o no focalizar los esfuerzos (Collier 2018).

Por otra parte, se tienen los estudios dentro de la curación de experiencias por sí mismas. Grandes empresas dedicadas a servicios han desarrollado metodologías para incrementar la satisfacción del cliente. Analizando y mapeando todo el sendero que atraviesa el cliente, se busca que el cliente sea reconocido y sea entendido en su individualidad, resultando en experiencias de servicio únicas o *idiosincráticas*, confeccionadas explícitamente para el cliente. Para esto es necesario conocer a la perfección los puntos de dolor del proceso para el visitante (Collier 2018).

4.5 - Técnicas de evaluación para ecosistemas de servicio

En los ecosistemas de servicio la percepción es igual a la realidad, y gran parte de las veces se confunde la percepción de la experiencia de servicio con la calidad del proceso. Si el servicio resulta bueno en general el cliente suele tener mayor tolerancia a los defectos, por lo que una técnica evaluativa de percepción no suele reflejar la escena completa de la experiencia, por lo cual una técnica comúnmente usada es el análisis de sentimiento hacia una marca, usando la abogacía de los usuarios como referencia del proceso o para encontrar defectos notados por los mismos (Allen, 2019) (O'Connor, 2018).

Una de las industrias con mayor avance en la evaluación de servicios es la industria hotelera. Ellos han expuesto que es imperativo conocer los factores de éxito del servicio y hacer modelos independientes para cada situación que les permita encontrar sus deficiencias, en lugar de tener una métrica agregada donde se pierda visibilidad (O'Connor, 2018) (Sadhu, 2020).

4.6 - Índice de satisfacción del cliente (CSAT)

El Customer Satisfaction Score o (CSAT, por sus siglas en inglés), es una herramienta con la cual se puede cuantificar la medida en donde un producto o un servicio alcanzaron o superaron las expectativas percibidas de un cliente o usuario (Farris 2010).

La metodología de recolección puede tener múltiples variaciones. Usualmente este indicador suele ser obtenido a través de una encuesta de salida ya sea por un medio digital, cómo un correo electrónico, llamadas telefónicas, mensajes de texto o medios físicos tan simples como una plantilla (Farris 2010).

Comúnmente se define la encuesta de satisfacción con la pregunta: “¿Cómo calificaría el producto/servicio recibido?” bajo una escala de Likert de 5 puntos. La escala de Likert permite cuantificar el medir el grado de afinidad hacia una conducta o referencia dada (Likert 1932).

Tradicionalmente se utilizaba una escala numérica entre uno y cinco; actualmente las nuevas tendencias en medición de satisfacción proponen utilizar una escala en emoticones o *emojis*, la cual permite traducir un cuantificador numérico a un grado psicoemocional y alineación entre los aspectos perceptivos de la encuesta (Alismail 2018) (Cha 2007). Similar a la escala numérica se asignan emoticones de una cara molesta hasta una cara contenta, pasando por estados transitorios y neutros.

A cada *emoji* se le asigna un equivalente numérico, paralelo a la escala de Likert (1-5). El equivalente numérico de los resultados auto reportados por cada usuario o cliente es agregado ya sea con un promedio o con la media para entender la evaluación completa del producto o servicio (Likert 1932) (Vieira, 2016).

Existen algunos riesgos conocidos al utilizar el CSAT score como método evaluativo. Puede tener polarización de respuestas, es decir, los encuestados ignoran la escala y responden únicamente con el valor más bajo y más alto de la encuesta; esto puede comprobarse si la distribución de las respuestas es bimodal (Farris 2010).

A pesar de tener algunas limitaciones y de la complejidad de recolectar una muestra significativa, este indicador suele ser muy versátil y popular dentro de la industria de la experiencia del cliente pues permite hacer un sondeo rápido al cliente (Farris 2010).

4.7 - Net Promoters Score (NPS)

El Net Promoter Score o mayormente conocido como *NPS*, por sus siglas en inglés, (Reichheld 2003) es una métrica popular para medir la lealtad y entendida bajo la óptica de la experiencia del usuario. Fue popularizada por Frederick Reichheld en un artículo escrito para Harvard Business Review (HBR) en el año 2003 y desarrollada por la firma de consultoría Bain & Company (Markey, 2011).

La metodología de captura del NPS consiste en una encuesta hecha posteriormente al consumo o uso de un producto o servicio; la encuesta consiste de una pregunta; “¿Qué tan probable es que recomiendes [producto X] a un familiar o a un amigo?” en una escala del 0 al 10 donde 9 y 10 son considerados promotores; del 0 al 6, detractores; y del 7 al 8, neutrales.

El *NPS* se calcula cómo:

$$NPS = (\% \text{ Promotores} - \% \text{ Detractores}) * 100$$

La métrica tiene varias delicadezas y consideraciones; una de ellas son los factores de retención del cliente. Si un cliente está retenido por prácticas monopólicas, difícilmente se podrá medir su lealtad real a través de un NPS (Keiningham 2008).

Otro aspecto sobre el cual se sostiene la validez del *NPS*, es el aspecto conductual que asume que un usuario que recomendaría el consumo de un producto de igual manera lo compraría nuevamente. Se debe tener cuidado al considerar que un producto es consumible múltiples veces. En tales situaciones no existe una correlación entre el crecimiento y la salida de un *Net Promoter Score* (Keiningham 2008).

Existen opositores a la certeza que el *NPS* puede proveer sobre el crecimiento que una compañía puede dar. En una réplica del estudio realizado por Reichheld (*London School of Economics y Listening Co.*), se replicó una R^2 de 0.70 en el crecimiento contra NPS, y a pesar de que otros indicadores lograron un mejor puntaje como el ACSI (American Customer Satisfaction Index), del 0.76, los resultados varían diametralmente por industria. Se entiende por esto que el factor de lealtad no es la única variable característica del crecimiento de un negocio (Keiningham 2008).

Otros detractores de la métrica apuntan a que, si bien la métrica no es un buen predictor del crecimiento, es un excelente descriptivo del pasado (Mecredy 2017). Es decir, que las capturas de la intención de recomendación están íntimamente ligadas a la experiencia pasada más que a las intenciones a futuro del cliente, y a pesar de que la hipótesis es sensible, el estudio conducido por Mecredy no alcanza relevancia estadística (Mecredy 2017).

A pesar de eso el indicador sigue siendo usado ampliamente en múltiples industrias y si bien es cierto que modelar todas las variables de negocio con un indicador único puede ser una tarea poco confiable,

podemos entenderlo como un indicador de lealtad y a su vez, de retención, buscando la misma como variable a optimizar.

4.8 - Otros indicadores de servicio

Según la industria, el ecosistema y canales de servicio que provee a sus clientes pueden implementar distintas metodologías para la evaluación de la calidad de la experiencia. Discutiremos brevemente algunas de las más populares para entender el alcance de cada una de ellas y evaluar su inclusión en un modelo holístico de experiencia (Price & Jaffe, 2008) (Disney Institute, The & Kinni, 2011).

Volumen de soporte y quejas: se observa el volumen neto de solicitudes de soporte a través del tiempo, con la intención de buscar el impacto de nuevas acciones de negocio contra la satisfacción del cliente. Se puede medir con frecuencia diaria, semanal o mensual según la rotación del negocio a observar. A pesar de que es un buen señalador de causas raíz, el volumen neto no es comparable entre negocios y tiende a perder claridad en negocios de crecimiento volátil (Price & Jaffe, 2008).

Resolución al primer contacto: *First Contact Resolution (FCR)*. Inclinado hacia las búsquedas para optimizar los procesos, el FCR mide el porcentaje de quejas resueltas en el primer contacto; es un indicador usado ampliamente en servicios de soporte subcontratados pues la repetición de un contacto se traduce en un costo. Falla bajo el supuesto de que todo contacto puede ser resuelto y queda al criterio de un agente de soporte, indicar que un contacto se completó, no evalúa la satisfacción del mismo (Dixon et al., 2013).

Tiempo de resolución promedio: se entiende como el tiempo transcurrido en horas, días o minutos desde el primer contacto de un cliente o usuario hasta el punto de solución o abandono del mismo. Se fundamenta bajo la premisa de que a menor tiempo de resolución mayor satisfacción y se ha descartado que esto sea verdad para todas las industrias. La gravedad de las incomodidades es un factor que no se contiene en este indicador, por lo cual el tiempo de resolución no necesariamente implica una solución satisfactoria (Dixon et al., 2013) (Allen, 2019).

Tiempos de espera: Es una métrica utilizada principalmente en servicios con un nivel de servicio bajo; o altos volúmenes de soporte. Recae en el tiempo de espera antes de ser atendido por un agente de soporte, principalmente en centros de atención o soporte telefónico. Este indicador es recomendable si los problemas son altamente solucionables y se busca más eficiencia operativa (Disney Institute, The & Kinni, 2011).

Recontactos: esta métrica busca entender el número de contactos que debe de hacer un cliente antes de resolver su problema. Se logra capturar aquí la robustez y qué tan acotados son los problemas que puede experimentar un cliente. Se recomienda utilizar cuando el enfoque es fortalecer los procesos de solución. Cuando se usa este indicador, la experiencia de servicio por lo general tiene muchos defectos (Dixon et al., 2013).

Retención: captura el número relativo de clientes que se van perdiendo a lo largo del tiempo. Se suelen hacer puntos de lectura según la volatilidad del negocio; si la probabilidad de abandono es alta se suelen hacer cortes con mayor frecuencia. La métrica funciona a medida que el soporte o atención es un detractor de abandono; si las causas de salida no son experiencias negativas, suele tener poco efecto.

4.9 - Experiencias de servicio idiosincráticas

Entendamos naturalmente la experiencia que percibe un cliente al consumir un producto o servicio cómo proceso lineal que reside en el modelo de des-comprobación (Oliver 1977). El modelo pone una situación sencilla donde un individuo compara una experiencia recibida contra sus expectativas de la misma y el paralelismo o diametralidad de esta comparación define la satisfacción o insatisfacción. Este principio, si bien es válido, al día de hoy ha variado pues las expectativas de los nuevos consumidores cada vez son más difíciles de alcanzar.

Por lo ya visto bajo los estudios del Net Promoter Score u otros indicadores de lealtad, existe una correlación fuerte entre el crecimiento de una empresa y la lealtad de sus clientes (Reichheld 2003). Por lo que las tendencias de negocio tratan de optimizar la lealtad y el mayor conductor de lealtad es la abogacía entre usuarios que puede entenderse a través del WoM (*Word of Mouth*). La mejor forma de incentivar la promoción de boca en boca es a través de una experiencia extraordinaria, pues la calibración de expectativas propuesta en el modelo de des-comprobación ya no es suficiente (Raassens 2017).

Se pretende a través del entendimiento intrínseco del cliente, poner la luz sobre los factores que resultan relevantes para el cliente y entregarlos de forma dirigida. A esta confección de experiencias individuales se les denomina “Experiencias de Servicio Idiosincráticas” o *ISE (Idiosyncratic Service Experiences)*, por sus siglas en inglés (Collier 2018).

Una experiencia de servicio idiosincrática no se basa en alcanzar las expectativas de un cliente, se consigue superando lo inesperado, una experiencia que resulta en compartirla con otras personas. En los estudios entregados por Ravindra Chitturi, se demostró que el factor de satisfacción y “deleite” de un cliente está correlacionado con la recompra, así como la tolerancia a futuras fallas.

Actualmente nos encontramos en la época de la economía de las experiencias y el valor está latente. Si se tiene la intención de retener a un usuario se le debe de pagar con la moneda de la experiencia (Gilmore 1998) (Dixon et al., 2013).

Además del grado de retención obtenida a través de una experiencia de servicio que sobrepasa lo esperado por un usuario, se consiguen clientes con mucha mayor resiliencia contra los defectos de los productos, mayor difusión y recomendación de las empresas y menor sensibilidad a los precios. Se entiende por este fenómeno el valor agregado de las experiencias de servicio. Llevado al tiempo un cliente leal es 6 veces más rentable que un cliente de una visita única (Frei 2012).

Generar este tipo de experiencias puede tener un costo muy elevado puesto que, para poder controlar y acotar cada visita a un nivel de servicio elevado, se requiere flexibilidad de personal capacitado, lo cual puede ser costoso. (Disney Institute, The & Kinni, 2011).

4.10 - Economías de la información

Conforme se redujeron los costos de los sistemas, las industrias comenzaron a desarrollar mecanismos para la captura y lectura de datos y así enfocar la toma de decisiones a eventos fundamentados con información. Al igual que la materia prima, al ser llevada a procesos, va adquiriendo valor; igual los datos, al recibir ciertos tratamientos pueden transformarse en información valiosa. Existe un marco sobre el cual se recomienda su tratamiento para el desarrollo de su valor (Lim 2018).

Los 9 factores de valor son: la fuente, la recolección, los datos, el análisis, información de la fuente de datos, entrega de información, información del usuario, valor en el uso de la información (Lim 2018).

Dado el crecimiento y desarrollo de las tecnologías de recolección de datos múltiples, las industrias registran de manera rutinaria grandes cantidades de datos; aun así pocas industrias logran explotar la cantidad de datos que tienen a su alcance (Lim 2018). La ventaja de la producción de información a través de su propia cadena de valor está en los posibles usos que se extienden desde encontrar nuevas oportunidades de negocio, diversificación de valor o retención y lealtad de los clientes (Lim 2018).

Según la literatura de Lim (2018), el valor de los datos únicamente se da si se le da uso específico a la información aplicada a un proceso y se debe de atender a los 8 factores de la cadena de valor:

- 1) Fuente de los datos
- 2) Recolección de los datos
- 3) Datos
- 4) Análisis de datos
- 5) Información de la fuente
- 6) Entrega de la información
- 7) Información del usuario
- 8) Valor en la información

Esta cadena es el tratamiento que debe de dársele a cualquier banco de datos para su interpretación y transformación en información accionable. Los puntos de origen de los datos suelen variar respecto a las industrias. Es común encontrar en las ramas de la tecnología continuos puntos de captura en los servicios dados; pueden ir desde el seguimiento de “clics” en una página web hasta una encuesta de salida post utilización del producto o consumo del servicio (Lim 2018).

Los resultados obtenidos de la cadena de valor suelen ser de naturaleza predictiva, prescriptiva o descriptiva del sistema sobre el cual se capturaron los datos. El diseño y la implementación de dichos modelos son independientes al proceso de construcción de la cadena de valor de la información (Lim 2018).

4.11 - Información como un servicio

Actualmente es relativamente sencillo y económico recolectar información de múltiples fuentes; esto es un producto secundario de la velocidad del desarrollo tecnológico y con ello la rápida evolución de los mercados. Este fenómeno ha logrado que la toma de decisiones tome aún más importancia (Saad, 2013).

Los servicios de analítica e información permiten, gracias a las infraestructuras de datos actuales, tomar decisiones en tiempo real y de manera eficiente; de manera reduccionista podemos englobar los usos de los datos en tres categorías principales para los sistemas de toma de decisión (Chandok, 2014).

Modelos descriptivos: Trata de entender los sistemas de estudio holísticamente a través de los datos capturados (Delen 2013).

Modelos predictivos: Buscan extraer y entender las tendencias a futuro del sistema para anticipar futuros impactos o desarrollos (Delen 2013).

Modelos prescriptivos: A través de matemática y algoritmos se pretende determinar cursos de acción alternativos sobre un sistema de objetivos dado; estos sistemas no dependen únicamente en los datos, también se apoyan de la sensibilidad de negocio (Delen 2013).

Los datos por si solos no pueden ser procesados de manera sencilla. Los sistemas de recolección de datos suelen tener varias impurezas que comprometen el uso. A los datos sin algún pre procesamiento se les conoce como datos crudos o *raw data*; a estos se les debe dar distintos tratamientos para poderlos utilizar correctamente y que se pueda asegurar la confiabilidad de los modelos entrenados con estos sets de datos (Saad, 2013).

Un set de datos o *dataset* es un conjunto de valores, usualmente cuantitativos cuando se trata de registros, números y cualitativos cuando se trata de cadenas de caracteres; usualmente son almacenados en matrices rectangulares donde las columnas representan las variables o dimensiones; y las filas, las observaciones. Todos los valores de una columna deben de ser medidos en las mismas unidades (Delen 2013) (Saad, 2013) (Wickham 2014).

La semántica utilizada entre una variable y una observación suele ser confusa. Una regla de oro, para poder diagnosticar el tipo de una variable, es que la relación entre las columnas o variables debe de ser más sencilla de explicar que la relación entre las observaciones; y a su vez, la agrupación entre observaciones debe de ser más sencilla que la agrupación entre variables (Wickham 2014).

Según Wickman (2014), el estándar para facilitar el procesamiento de los datos se puede definir en tres reglas a cumplir:

- 1) Cada variable forma una columna.
- 2) Cada observación forma una fila.
- 3) Cada unidad observacional forma una tabla.

Esta estructura de datos les facilita a los procesadores encontrar y manipular datos. Cualquier otra configuración tiende a requerir un mayor volumen de recursos computacionales y a invitar al error.

A pesar de que el orden de los valores no afecta el resultado es importante pues facilita la observación y un primer análisis visual a los datos crudos. El estándar sugiere comenzar por las variables fijas, usualmente referidas como “dimensiones”, seguidos por los datos variables o medidos (Wickham 2014).

Los sistemas de recolección de datos por las frecuencias y aleatoriedad que pueden tener los puntos de lectura, no suelen seguir estos estándares; los errores más comúnmente encontrados en los datos crudos suelen ser:

- 1) Los títulos de las columnas son valores no variables.
- 2) Múltiples variables son registradas en la misma columna.
- 3) Variables registradas tanto en las observaciones como en las columnas.
- 4) Distintos tipos de variables observacionales son capturadas en la misma columna.
- 5) Una misma unidad observacional es guardada en múltiples tablas.

Siguiendo estas normas es mucho más fácil hacer la agregación de los datos para su análisis. Considerando los volúmenes de registros que se generan por minuto, es de suma importancia apegarse en lo posible a las recomendaciones de Wickham (2014), descritas anteriormente, pues un error multiplicado por el volumen puede dañar toda la caracterización de los modelos.

4.12 - Tratamiento de datos

Uno de los pasos críticos para el análisis de datos es la fase de limpieza y tratamiento. Los sistemas de recolección de datos no siempre pueden ser perfectos y algunas veces llegan a generar errores en los datos.

Algunas de las contaminaciones más comunes en los datos pueden ser: valores faltantes, líneas corrompidas, valores anómalos o separadores recorridos (Badr 2019).

Existen varios tratamientos para “imputar” las bases de datos. Dependiendo de la cantidad de datos disponibles y la facilidad de obtención, estas técnicas pueden ser eliminar los registros erróneos o arreglarlos con tal de no perder datos. A este tipo de tratamientos se les llama *técnicas de imputación* (Badr 2019).

Algunos de los métodos de imputación más comunes pueden ser *fillings* con la media o la mediana de los datos, ceros o valores nulos, regresión estocástica o métodos de *aprendizaje máquina* como *k-nearest neighbors* (Badr 2019).

Si los datos son abundantes también es posible eliminar los registros dañados y eliminar los extremos de las mediciones. Los casos varían según la calidad y cantidad de los datos, pero es común eliminar 1% de los registros más extremos (Badr 2019).

4.13 - Muestreo balanceado

Cuando el muestreo de datos no es balanceado, es decir, que una clase es mucho más abundante que el resto es posible hacer muestreo balanceado. Este permite balancear el número de muestras sin necesariamente añadir nueva información al sistema (Brownlee 2019). Ayuda para la visualización pues aumenta las probabilidades de un evento de ser observado, manteniendo las propiedades estadísticas de los valores (Brownlee 2019).

La forma más común de hacerlo es duplicando los registros de las muestras minoritarias; o si los datos son abundantes, también se puede reducir el número de muestras de la clase mayoritaria. Estos tratamientos son muy útiles al trabajar con eventos raros (Brownlee 2019).

4.14 - Tokenización

La “tokenización” es una técnica de procesamiento de lenguaje natural que permite convertir largas cadenas de texto en elementos más simples, y a su vez en una secuencia de *tokens*. Estos pueden generarse a través de separadores léxicos como espacios, puntos o comas; es decir, convertir un párrafo de texto en palabras individuales para su procesamiento (Mattsson, 2010).

Pero también es común la generación de n-grams. Un n-gram es una secuencia de **n** palabras. Dentro del proceso de “tokenización” es común generar tokens con n-grams de **n** = 1,2,3. Esto permite encontrar frecuencias de términos comunes y estudiar su relación con otras variables (Mattsson, 2010).

Los tokens de n-grams dan visibilidad sobre términos usados frecuentemente juntos, los cuales se pueden agrupar como categorías nuevas o valores de entrada. Estudiar la frecuencia de términos permite obtener información adicional y los tokens pueden ser tratados como clases nuevas y ser usados en sistemas de clasificación (Mattsson, 2010).

4.15 - Dummy variables

Una forma común de traducir variables categóricas en variables accesibles para un modelo de predicción es a través de “*dummy encoding*” o “*one-hot encoding*”. El principio es sencillo: si el número de categorías de las variables no es muy extenso es posible codificar cada categoría a un elemento binario y añadirlas como variables de entrada nuevas. Es decir, que una variable con tres categorías se puede codificar en 3 nuevas variables binarias, donde 1 representa que el valor de la categoría está presente, y 0 dónde no (Bruce et al., 2020).

Es importante considerar el concepto de grados de libertad, pues al tener 3 categorías y traducirlas en 3 nuevas variables, una de las nuevas variables puede ser inferida con únicamente 2 de las variables. Esto convierte el sistema en un sistema con 2 grados de libertad. Utilizar las 3 variables codificadas puede generar errores de multicolinealidad en los algoritmos de regresión (Bruce et al., 2020).

4.16 - Gestión de desempeño

De la misma forma que el incremento en las tecnologías de recolección y procesamiento de datos han facilitado el uso de los mismos, existe un síntoma de confusión en la gestión y utilización de ellos. Muchas industrias se han perdido en la cantidad de indicadores a controlar (Johnston 2002).

El uso excesivo de KPIs (Key Performance Indicator) puede llevar a la pérdida de foco del objetivo de una empresa. En teoría de procesos se dice que si algo no es medido no es controlable, pero se busca un balance entre lo que es lo suficientemente importante para la misión de una empresa para ser medido. Para evitar la ceguera se recomienda que la selección de KPIs, con los cuales se mide un negocio, describan de manera precisa la misión y objetivo de la empresa (Elhadjamor & Ghannouchi, 2020).

Existe una diferencia fundamental entre un KPI y una métrica. Como anteriormente describimos, un KPI debe de englobar la función primordial de un negocio y debe de dar una vista clave de lo que sucede en un sistema en un tiempo determinado. A su vez, su enfoque debe de ser el de mejora más que el de control (Elhadjamor & Ghannouchi, 2020).

4.17 - Aprendizaje automático

Dentro del campo de *aprendizaje máquina y grandes datos*, se han terminado de definir los casos en los que se usan de las técnicas y el alcance de la misma, determinado por la frecuencia, granularidad, variedad y volumen. En las industrias ricas en datos se puede apalancar el uso de los mismos para entregarles valor para la toma de decisiones de negocio (Gandomi, 2015).

En la mayoría de los casos de usos de aprendizaje automático en negocios, encontramos aplicaciones de predicción o categorización, pero también se les puede dar uso para la exploración y comprensión de un fenómeno basado en el modelado del mismo (Gandomi, 2015).

Algunas principales técnicas de algoritmos de aprendizaje automático que encontramos son:

Tareas de regresión: buscan entender la relación de dos variables a través del error de los datos observados contra una curva caracterizada, es ampliamente utilizada como un predictor o selector binario en su versión logística, es popular debido a su fácil implementación (Bruce et al., 2018).

Árboles de decisión: por medio de un puntaje de “pertenencia”, un árbol de decisión logra categorizar un grupo de datos; su estructura es figurativamente parecida a la de un árbol donde cada nodo se refiere a una variable de decisión para la clasificación del vector. Es un algoritmo sumamente flexible y por esta flexibilidad fue popularizado para su uso en datos de clientes (Bruce et al., 2018).

Técnicas bayesianas: estos algoritmos buscan explicar la probabilidad de un evento dado el comportamiento pasado de un grupo de variables. El algoritmo de Naïve-Bayes busca la interdependencia total y explica el aporte individual y comportamiento futuro de cada una de las variables, según sus históricos (Bruce et al., 2018).

Máquinas de vectores de soporte: es una técnica de aprendizaje supervisado donde, a partir del entrenamiento en un grupo de datos etiquetados, se optimiza para encontrar los hiperplanos que logran separar de mejor manera las categorías. Se logra categorizar nuevos datos según su cercanía a los hiperplanos (Bruce et al., 2018).

Aprendizaje basado en experiencia: las entradas nuevas son clasificadas según los registros almacenados en memoria, a través de distancias euclidianas u otras. Las entidades pasadas emiten votos sobre la última lectura según su cercanía. K – Means es el método más popular de este tipo. Existe el reto que se debe de dar el número de vecindades al algoritmo, muchas veces desconocido (Bruce et al., 2018).

Aprendizaje ensamblado: Es la combinación de múltiples técnicas ensambladas para producir predicciones basadas en las predicciones de otras metodologías. Entre los algoritmos ensamblados más populares se encuentran los bosques aleatorios, los métodos de *boosting*, y gradiente descendente estocástico (Bruce et al., 2018).

Redes neuronales artificiales: Inspiradas en la organización neuronal biológica, las redes neuronales artificiales forman capas de neuronas interconectadas por pesos; la información se transfiere en forma de sumas ponderadas desde una capa de entrada, pasando por capas ocultas y un selector final de salida. Las redes neuronales son fácilmente entrenadas, resistentes a los errores. Llevadas a escala pueden formar redes de aprendizaje profundo, que a pesar de requerir una gran cantidad de recursos computacionales pueden adaptarse a sistemas sumamente complejos (Bruce et al., 2018).

Análisis de discriminante lineal: es un clasificador matemático que busca la combinación de predictores que sean capaces de diferenciar entre un objetivo, es decir, busca expresar la relación entre una variable dependiente y un grupo de variables independientes, con la diferencia de que la variable dependiente suele ser categórica. Se estima la clasificación a través de un modelo probabilístico (Bruce et al., 2018).

4.18 - Técnicas de selección de variables

Uno de los principales retos en la aplicación de técnicas de *aprendizaje máquina* es la selección de las variables que describe un sistema. Las variables, si bien dadas las capacidades de procesamiento de las computadoras y servidores, cada día se vuelven más accesibles, así también lo hacen las técnicas de obtención de datos (Sheikh, 2017).

Nos enfrentamos a una n-dimensionalidad de variables para caracterizar un modelo; y si bien es fácil utilizar todas las variables en un algoritmo, esto no necesariamente garantiza el éxito o la convergencia del mismo (Sheikh, 2017).

Existen métodos heurísticos aplicables bajo la sensibilidad del negocio. Para ellos basta tener un entendimiento claro del ecosistema a estudiar y hacer una selección manual de las variables que deberían tener una relación directa con la variable de salida que se intenta estudiar, a pesar de ser un excelente primer acercamiento. El sesgo de confirmación de las hipótesis puede llevar a omitir variables relevantes (Sheikh, 2017). También cierra la idea de incluir variables exógenas que puedan tener alguna relación interesante. Este método limita la precisión del modelo a la experiencia del diseñador (Sahar F. 2018).

Como explica Hartung (2018), en "Making Sense of *Big Data*", si se trabaja dentro de un sistema generoso con los datos y observaciones se pueden tomar medidas más precisas para la selección de variables. Principalmente aspiramos a dos características de los modelos: la primera, que sean buenos predictores, y la segunda, que sean de fácil entendimiento, por lo que un modelo de n variables se vuelve poco explicativo y a su vez poco accionable para el negocio (Hartung, 2018).

De las formas más populares de entender la relevancia de un grupo de variables es a través de un análisis de regresión; de esta forma se puede entender el comportamiento matemático de la descripción de la variable según los pesos asignados a la regresión lineal. Es importante entender que el comportamiento matemático puro del modelo puede no ser descriptivo del sistema a estudiar, por lo que el análisis de pesos es necesario, pero no suficiente (Bruce et al., 2020).

Uno de los aspectos a considerar en una regresión optimizada por mínimos cuadrados ordinarios (OLS) puede tender a cargar los coeficientes de la regresión a una sola variable. Un OLS, a pesar de su sencilla implementación, se ha demostrado que no es un buen predictor ni de fácil interpretación (Bruce et al., 2020).

Se puede utilizar una *regularización red elástica* que combina las bondades de las regularizaciones de Ridge, que penaliza los coeficientes grandes, según un factor lambda dado y la regularización LASSO que, al enfrentar predictores con correlaciones altas, entre ellos descarta alguno permitiendo el modelado con las variables más representativas del sistema (Sahar F. 2018).

4.19 - Validación de modelo

Un modelo debe de apegarse a lo observable en la práctica. Las abstracciones que se puedan hacer de un modelo dado no deberían contradecir lo que se observa en el sistema; cuando esto sucede es probable que se esté enfrentando el caso de una paradoja de Simpson, que explica efectos contrarios cuando los datos son observados de manera agregada (Hartung, 2018).

4.20 - Grados de libertad

En tópicos de manejo de datos, el concepto de grados de libertad aplica para las estadísticas calculadas en muestreos de datos. Se refiere al número de variables libres para variar. Por ejemplo, si conocemos la media de una muestra de 10 valores, y también conocemos el valor de 9 de los valores, podemos conocer el valor del décimo elemento; únicamente 9 son libres para variar. Lo anterior puede generar una redundancia en los datos (Bruce et al., 2020).

Algunos algoritmos de regresión pueden fallar si existen variables predictivas redundantes. Esto es particularmente común cuando se convierte una variable categórica en elementos binarios (dummies). Los errores en algoritmos de regresión causados por variables redundantes se llaman errores de multicolinealidad. Pueden ser corregidos fácilmente eliminando uno de los elementos y dejando $n-1$ grados de libertad (Bruce et al., 2020).

4.21 - Sistemas de toma de decisión

Estudios empíricos han demostrado una relación entre la satisfacción del cliente y el crecimiento del negocio ligado a la lealtad y la rentabilidad a largo plazo (Anderson 1994). Por lo cual en un ambiente competitivo y altamente enfocado al cliente se busca entender y mejorar la experiencia del cliente ampliamente. Los sistemas digitales nos han permitido recolectar grandes cantidades de información del cliente en tiempo real para así poder tomar decisiones oportunas y hacer intervenciones relevantes para mejorar la experiencia del cliente y entrenar a los representantes de la firma (Bockhorst 2017).

4.22 - Uso de técnicas de estadísticas avanzadas para la optimización de experiencias.

Se presume que la experiencia percibida y capturada por medio de encuestas, en la mayoría de los casos, es producto de la calidad de los procesos; este ciclo permite la mejora de la experiencia a través de la mejora de los procesos (Bockhorst 2017). Se han utilizado en industrias similares, incluyendo centros de atención telefónica, sistemas basados en técnicas de inteligencia como “*Random Forest*” para predecir, basado en información histórica y lecturas en tiempo real de los procesos, los resultados de las encuestas de satisfacción en tiempo real para implementar correcciones en la lógica de soporte (Sandhu, 2020).

5 - CASO PRÁCTICO: IMPLEMENTACIÓN DE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO PARA LA MEJORA DE LA EXPERIENCIA DE SERVICIO.

En este documento se analiza el caso de un proceso de atención al cliente y su intención de mejora, enfoque de decisiones y elementos para alinear las prioridades del negocio. Este es un proceso clásico de centro de atención al cliente en el cual el cliente acude a una localidad a recibir un servicio de soporte.

La empresa estudiada tiene operaciones en todo el territorio mexicano y se dedica al transporte, sin embargo, el horizonte de este trabajo se enfoca en el servicio de atención al cliente de los conductores que operan para dicha empresa. En los centros de atención, los conductores reciben distintos servicios de soporte que van desde activación y afiliación hasta soporte técnico sobre la aplicación que les permite conducir.

La estrategia para la mejora de los servicios dados en los centros de atención suele ser pobre, al existir distancia entre las áreas directivas y las necesidades directas de clientes, sumada a una brecha metodológica, las decisiones no suelen ser acertadas y la estrategia actual rara vez cubre las necesidades reales de los clientes que acuden a los centros de atención.

El estudio se llevó a cabo en Ciudad de México, México. Durante el año 2019, la implementación se hizo a mediados del 2019 con resultados observados a inicios del año 2020. Se recolectaron alrededor de 80,000 encuestas de satisfacción recolectadas a través de una notificación al celular de los conductores que visitaron los centros de atención.

Por fines de confidencialidad, un factor fue aplicado a los datos y los nombres de las entidades.

5.1 - Situación actual

En la empresa que estudiamos en este trabajo, las decisiones relacionadas a los procesos de atención al cliente son tomadas de forma poco informada, heurística, y anecdóticamente. A pesar de poder lograr un grado de satisfacción para los clientes, con un acercamiento más anecdótico que cuantitativo, es difícil tomar decisiones acertadas para la optimización de la experiencia del cliente. La toma de decisiones actualmente es iterativa y con base en la prueba y error. Sin embargo, este método además de ser poco efectivo en tiempo de su implementación, no permite aislar fácilmente los elementos reales que se mejoraron en el proceso. Otra consecuencia de la lentitud en la implementación es el costo de oportunidad de no poder reaccionar rápidamente a un mercado cambiante y a un cliente cuyas exigencias se transforman cada día más rápido.

Al momento del comienzo de este estudio, las métricas de satisfacción de algunos de los centros de atención al cliente se veían de la forma mostrada en la Tabla 1.

Tabla 1. Calificación de satisfacción de los centros de atención promedio 3 meses antes del piloto

| Locación | Media de la encuesta de salida a la pregunta: ¿En general, cómo calificaría el servicio recibido?" (1-5) |
|----------------------|--|
| Centro de atención 1 | 4.12 |
| Centro de atención 2 | 4.23 |
| Centro de atención 3 | 4.34 |

Nota: No se especifica en cuáles centros específicos se llevó a cabo el piloto a petición de la empresa.

Nuestra empresa, al momento de iniciar el estudio, tenía problemas para optimizar y mejorar los indicadores de calidad. A pesar de esfuerzos para mejorar el servicio de atención, como por ejemplo el desarrollo de herramientas para gestionar filas, los indicadores finales (Tabla 1) solían mantenerse sin cambios.

Estudiando el CSAT de más de 50 centros de atención a lo largo de 12 meses, pudimos ver una relación que existe con la pérdida de usuarios en la etapa tardía del ciclo de vida de producto. Esta correlación es la que explicaremos a profundidad en el análisis exploratorio para obtener una visión completa de las variables medibles de impacto de la experiencia del cliente.

El factor percibido y englobado en las encuestas de salida (levantamientos hechos por correo electrónico o notificación al celular al terminar una visita a un centro de atención) se dividió a su vez en múltiples capturas del proceso como tiempos de espera, tipo de problema y efectividad en la resolución. Profundizaremos más adelante en cada uno de estos factores que describen la vivencia que captura la encuesta de salida.

Al tener una brecha en la metodología de cuantificación y priorización, se pierde visibilidad en cómo transformar los datos en una estrategia sólida que responda a las ineficiencias de cada uno de los centros de atención. Este trabajo propone la existencia de una metodología para capturar dichas necesidades de forma cuantitativa, darles luz y comunicarlas efectivamente para transformarlas en una estrategia que permita la mejora de la experiencia.

5.1.a Problemas estructurales en la dirección.

En nuestra empresa, y de forma general, la toma de decisiones en áreas del negocio con alto involucramiento humano suele ser complejo. Los llamados factores caóticos agregan una capa de complejidad a la cuantificación de la eficiencia de procesos y calidad de servicio (Leavy, 2020). Por lo anterior se le suele poner poca atención a la mejora de los procesos enfocados al cliente, pues al ser sistemas con entradas y salidas humanas, los enfoques suelen ser más cualitativos que cuantitativos (Allen, 2019). Sin embargo, dada la abundancia en información que se captura hoy en día, es posible traducir las señales en elementos accionables para la toma de decisiones (Leavy, 2020).

Las áreas de soporte tradicionalmente son consideradas como un centro de costos, por lo cual, suelen pasar desapercibidas en la estrategia general de los negocios, como se muestra en la figura 2. Sin embargo, para nuestro caso particular, la experiencia de soporte se ha convertido en un diferenciador ante la competencia. El mercado del producto que se ofrece es sumamente competido y las eficiencias en precios y accesibilidad ya han sido agotadas, por lo que existe un nuevo paradigma relacionado a la importancia del soporte (Sandhu, 2020).

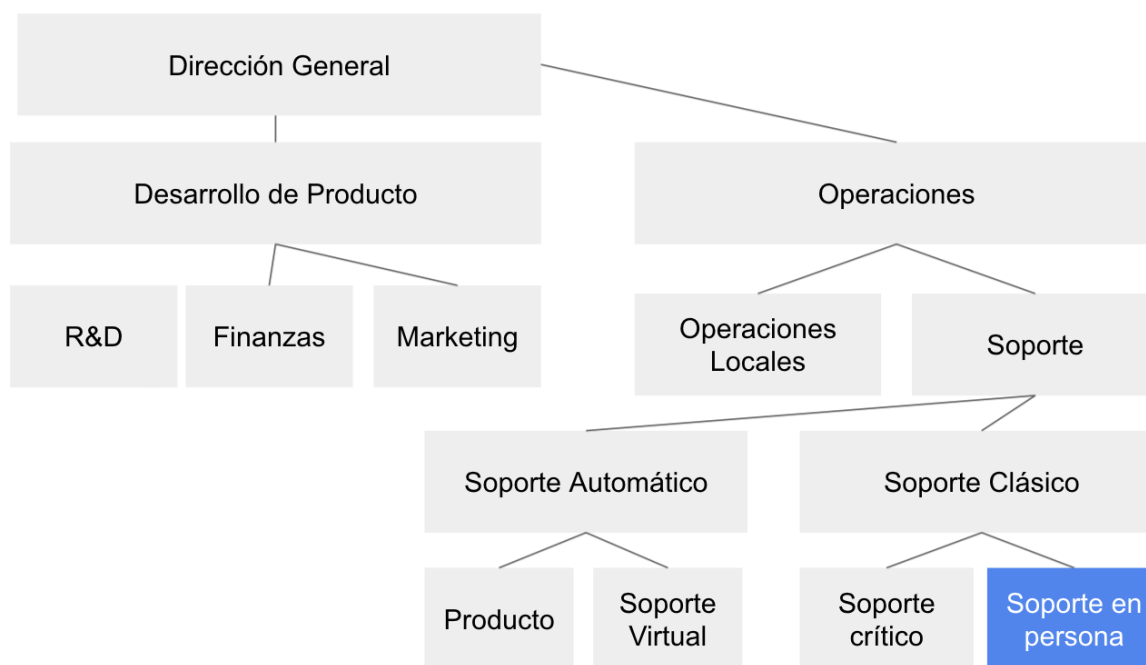


Fig 2. Organigrama (simplificado) de la empresa y posicionamiento del área referente al análisis

Sin embargo, existe un reto en poder encontrar la relación que existe entre las entradas y salidas de los procesos enfocados al soporte del cliente. Los datos generados por humanos, como las encuestas de salida, suelen tener una capa perceptiva que hace que el tratamiento de datos requiere procesos adicionales; los métodos clásicos de investigación de operaciones pueden capturar sesgos en las entradas y salidas y dirigir erróneamente una estrategia (Sandhu, 2020), (Goul & Corral, 2007).

Lo que buscamos fue caracterizar en un modelo de caja negra la experiencia que vive cada usuario con una métrica de servicio de salida y entender las dinámicas de cada una de las entradas para así tener una visión completa de un centro de atención (o cualquier experiencia que pueda mapear y medir en sus procesos).

Los sistemas actuales facilitan la captura y procesamiento de información, sin embargo, en la empresa estudiada este gran activo de información es raramente utilizado para diseñar de manera informada los procesos del negocio. Buscaremos a lo largo de este trabajo establecer una mejor administración de estos activos y la capitalización de los mismos para apegarnos a las tendencias actuales de los sistemas de toma de decisión (Goul & Corral, 2007).

La intención de esta investigación es agregar un nivel más objetivo a la forma de medir el desempeño de los centros de atención al cliente, a los agregadores clásicos de servicio para poder enfocar la estrategia general en las necesidades reales de los clientes.

Tenemos evidencia que las prioridades del negocio no necesariamente están correctamente calibradas pues a pesar de tener múltiples implementaciones de iniciativas, los resultados en los calificadores de servicio permanecen relativamente constantes.

Actualmente, la prioridad del área de estrategia está enfocada en los tipos de clientes, es decir, existe la hipótesis de que los clientes deben de ser tratados de manera distinta según su *lifetime value* o las probabilidades de retención del cliente.

5.1.b Iniciativas anteriores

A pesar de haber iniciado varias iniciativas como un programa de lealtad o filas prioritarias, para entregar un servicio diferenciado según el tipo de cliente, es difícil establecer la diferenciación. Los resultados

preliminares no muestran cambios significativos en los calificadores de servicio de los clientes a los cuales se les ha dado un servicio ajustado a sus rasgos de valor como cliente. Aunado a lo anterior, estas iniciativas son difíciles en su implementación y costosas en su ejecución, pues se requieren procesos únicos para un grupo de clientes pequeños, difícilmente escalable.

Al tener una visión completa de la experiencia que se da en el centro y la sensibilidad de los visitantes hacia todas las dimensiones, podrán entregarse diagnósticos direccionales y accionables que resulten en un impacto positivo a las mediciones de salida (Allen, 2019). Esto es, por ejemplo, si la saturación del centro resulta en un CSAT bajo, podrá evaluarse buscar un local más grande; si los tiempos de espera tienen impacto en sí, el problema es percibido como resuelto. Dada una interacción apresurada del agente, podrá evaluarse contratar más agentes; si la percepción de resolución cae bajo la asistencia de un solo agente, podrá evaluarse su desempeño de forma completa.

Así mismo podemos dar luz sobre las brechas entre las prioridades del negocio y las prioridades del cliente. Actualmente en la empresa estudiada, existe poca visibilidad entre las mismas, lo que lleva a una toma de decisiones poco acertadas y con poco a nulo impacto en la percepción de los clientes. De forma similar, podemos decir que la falta de visibilidad en las prioridades se traduce en inversiones poco rentables; por lo que encontrar un puente entre lo que es importante para el cliente y dónde el negocio invierte se puede traducir en grandes ahorros para la empresa (Thompson, 2021).

5.2 - Estudio de centros de soporte en persona

Estudiaremos de manera particular un ecosistema de servicio de atención al cliente, de una empresa de transporte, en el cual los clientes o conductores pueden recibir distintos tipos de servicios, desde afiliación a la plataforma, hasta soporte técnico de la aplicación con la cual los conductores reciben a los pasajeros.

El ecosistema de servicio es homólogo a una sucursal bancaria, donde pueden adquirirse nuevos productos que requieren un proceso de inscripción y puntos de soporte cotidiano como una consulta de cuenta, o una transacción.

Habitualmente, los centros de atención operan en redes con distintos locales para soportar el volumen de visitas de clientes. Actúan como puntos de venta accesibles para los servicios; los sistemas y procesos generalmente son estandarizados en miras de proveer la misma experiencia en cualquier centro de atención que se visite.

Estudiaremos los rasgos de distintos centros dentro de una misma ciudad, Ciudad de México. Generalizamos la metodología para diferentes geografías. A pesar de que los procesos pueden ser distintos entre cada una de las geografías, la metodología no busca evaluar la calidad de los procesos sino encontrar recomendaciones de los flancos débiles de la red en general.

La dinámica de un centro de atención de la empresa estudiada es la siguiente: el cliente es atendido a su entrada por un recepcionista que hace un primer diagnóstico del problema o entrega información básica del producto. El recepcionista asigna un turno con un pre-diagnóstico y el cliente es invitado a esperar a una sala. En nuestro caso, evaluaremos agentes generalistas con el entrenamiento suficiente para abordar cualquier posible problemática; por esto mismo el agente debe de ser ágil al dar soporte cómo al vender el producto.

De manera simplificada, cada interacción independiente de un cliente dentro del centro de soporte sucede como se muestra en la Fig. 3. Puede haber interacciones de seguimiento para el cliente, pero cada visita al centro es registrada como una entrada única, no ligada a visitas anteriores al centro u a otros centros. Cada visita es capturada por un sistema automático y agregado a una base de datos.

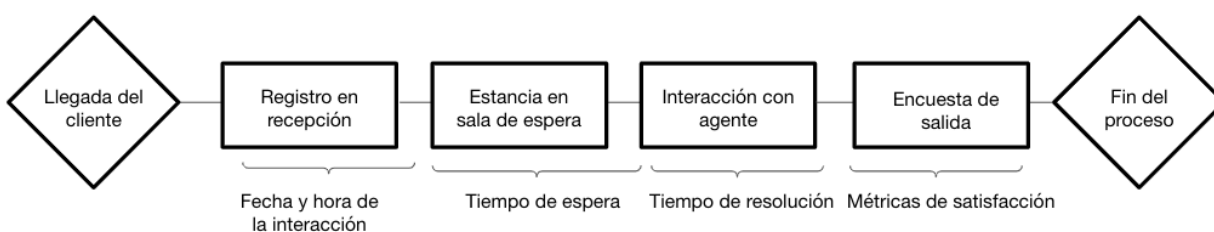


Fig 3. Proceso de soporte del centro de atención

Posteriormente a que el cliente es dirigido con un agente, el agente efectúa un segundo diagnóstico del problema o requerimiento (en caso de adquisición) del cliente y cataloga el servicio entregado según un árbol de diagnóstico.

El agente entrega al cliente la atención solicitada y se concluye la interacción entre el agente y el cliente. Posteriormente al evento, el cliente recibe una encuesta de salida donde se le solicita evaluar la atención entregada por el agente, señalar si el problema de su visita fue saciado y una aplicación de CSAT. En caso de no existir una solución al problema, el caso del cliente es escalado y la interacción concluida y capturada en las bases de datos. Para fines de este estudio, el escalamiento no es estudiarlas.

Todos los eventos mencionados anteriormente son registrados a través de una herramienta de gestión de servicios, es posible cuantificar y tener un mapeo exacto de cada una de las visitas de un cliente a un centro de atención.

El alcance de la metodología propuesta busca entender las mejoras al nivel del centro de atención, no a nivel de una interacción particular de un cliente. Trataremos de caracterizar el colectivo de las variables que definen la experiencia de usuario; para esto utilizaremos datos agregados en distintos órdenes. En las secciones siguientes profundizaremos en las técnicas de selección de variables. Para garantizar la captura fiel del sistema a estudiar utilizaremos desde demográficos de los clientes, hasta los rasgos individuales y capacidades de cada uno de los centros a estudiar.

Se estudiarán interacciones de seis meses ordenadas de manera aleatoria para evitar impactar el sistema por estacionalidades o eventos puntuales. Los eventos puntuales, como pueden ser fallas en el sistema, o cambios en la regulación, serán tratados como valores atípicos y serán descartados para el modelado del sistema.

Los centros que estudiaremos están ubicados en distintas áreas de la Ciudad de México y cada uno de ellos tiene distintas capacidades de agentes (número de agentes) y espacios, pero procesos idénticos. Como validación se llevará la metodología a nivel ciudad para entender el efecto colectivo de los centros.

Otro punto importante a considerar dentro de la selección de los centros son los tipos de problemas que pueden resolver los agentes de soporte o atención. Agregar una capa más de granularidad al estudio de los eventos puede tener resultados muy diversos, es decir, hay experiencias que por su naturaleza no son agradables, por ejemplo, un cobro duplicado, una cancelación de contrato o una violación de políticas. Se

evaluará considerar ambas opciones: una entrenando el modelo utilizando los tópicos de solución indiferentemente; y otra seleccionando únicamente las razones de visita que cubran el 80% de las interacciones y así obtener un modelo que logre comprender el sistema de manera extensa sin sobre ajustar a situaciones atípicas.

Utilizaremos para caracterizar el modelo únicamente aquellas interacciones donde la encuesta de salida asociada haya sido contestada para tener muestras calificadas. En Latinoamérica se espera un porcentaje de respuesta a encuestas enviadas del 15%, este número se considera orgánico cuando no hay un incentivo de por medio que promueva el contestar la encuesta (Harzing 2016). Nuestro sistema a estudiar tiene un rango en sus respuestas del 13%-20%, lo cual es un número considerable de respuestas, y una muestra representativa de los clientes.

Hemos observado que existe un incentivo secundario para llenar una encuesta de satisfacción que recae en saciar una necesidad psicológica ante una experiencia negativa; este es un supuesto asumido. Existe mayor probabilidad de tener una encuesta de una experiencia negativa donde el cliente decidió ser vocal a una experiencia excelente, pero es de esperarse que al solventar por estas experiencias negativas naturalmente las experiencias positivas mejorarán o si bien, no se verán afectadas.

5.3 - Descripción de la metodología

La metodología para identificar los puntos de mejora del área de negocio es la siguiente:

1. **Mapeo de los puntos de lectura:** De forma general, identificar todos los puntos de medición del sistema a analizar de manera exhaustiva. Los puntos de lectura deben de identificarse sin importar si es evidente que caracterizan la experiencia o no.
2. **Agrupar las variables según su relación con aspectos controlables del negocio:** Es decir, identificar qué variables pueden ser relevantes para evaluar un punto de esfuerzo del sistema. Por ejemplo, el número de agentes de soporte se ve directamente reflejado en los tiempos de espera.
3. **Segmentar las variables en variables primarias y secundarias:** Las variables primarias, siendo las variables más evidentes para la caracterización del sistema de negocio, y las variables secundarias, aspectos que suelen ser pasados por alto.

4. **Identificar la variable de salida del sistema:** La variable a optimizar debe de ser identificada y aislada. Ésta puede ser una encuesta de salida, o un indicador operativo a mejorar.
5. **Hacer una priorización de los esfuerzos actuales:** De manera sustentada, ordenar las prioridades del negocio e identificar su nivel de importancia relativo. El orden debe reflejar la visión estratégica. Es decir, ¿Qué tan importante es X factor para el negocio?
6. **Ejecutar un análisis exploratorio de las variables:** El objetivo es informar la selección del modelo y reducir el número de iteraciones necesarias para encontrar un modelo que pueda converger a la predicción de la variable de salida.
7. **Optimización de variables a través de la iteración de un modelo simple:** Comparando el desempeño de un modelo de fácil implementación como regresión lineal, logística o árbol se identificarán las variables de entrada y salida de forma iterativa.
 - a. **Validación y balanceo del set de datos** - utilizando las variables principales se busca encontrar sesgos o desequilibrios en el *dataset* original.
 - b. **Integración de variables secundarias** - este mismo modelo es ejecutado usando las variables secundarias (no sensibles al negocio), utilizando una técnica de *grid search*. Iterando el modelo, agregando y removiendo variables y observando su efecto en el desempeño del modelo se intenta buscar las variables que puedan ser agregadas a la selección final de variables.
 - c. **Simplificación de variables de salida** - utilizando nuevamente las variables principales, se busca encontrar una configuración óptima para los objetivos a predecir (granularidad, simplificación, agrupaciones, etc.)
8. **Probar distintos modelos con las configuraciones y variables seleccionadas:** De forma iterativa, usando distintos modelos de aprendizaje automático con la configuración y selección de variables, hasta encontrar el modelo que mejor prediga nuestro indicador a optimizar.
9. **Evaluar el poder predictivo de cada variable:** Ya elegido el modelo que mejor logre predecir el indicador a optimizar se debe hacer un análisis de sensibilidad, aislando cada una de las variables para entender cómo cambia el desempeño del modelo. Con ello, hacer un ranqueo de las variables según el cambio en el desempeño del modelo al quitarlas como variables de entrada.
10. **Sensibilización de las variables con expertos en el sistema:** Se utilizará el ranqueo del paso anterior de la metodología para comparar la importancia según el modelo con el conocimiento institucional del sistema.

- 11. Contraste contra las prioridades del negocio:** Se busca encontrar áreas donde las prioridades del negocio disten de la priorización del modelo. Estas áreas se identificarán como “áreas de desenfoco”
- 12. Utilizar las zonas de desenfoco para desarrollar mejoras:** Basado en las áreas donde las prioridades del negocio no tengan un nivel alto de acuerdo con el modelo, desarrollar iniciativas para alinearlas.

El modelo final pretende hacer una comparación objetiva entre lo que el modelo identifica como importante para el cliente y lo que la dirección identifica como clave para la mejora del indicador de servicio. Las áreas de desacuerdo entre el modelo y la dirección (área de negocio) muestran los puntos ciegos sobre los cuales se puede trabajar para mejorar el indicador objetivo.

6 - IMPLEMENTACIÓN DE LA METODOLOGÍA

6.1 Mapeo de los puntos de lectura

En el centro de soporte anteriormente mencionado, se capturan las siguientes variables para cada visita al centro. La visita al centro empieza a capturarse en el momento que el cliente visita al mostrador del centro de atención y termina cuando el cliente completa la encuesta de satisfacción al final de la interacción con el agente de soporte. La tabla 2 muestra puntos cómo son capturados:

Tabla 2: Variables del sistema

| Variable | Descripción | Tipo |
|--------------------|---|-----------|
| Estampa de llegada | Registro del momento en el cual el cliente llega al centro | TIMESTAMP |
| Estampa de llamado | Registro del momento en el cual el agente llama al cliente a ser atendido | TIMESTAMP |
| Estampa de cierre | Registro del momento en el cual el agente termina la interacción con el cliente | TIMESTAMP |

| | | |
|--|---|----------|
| Tipo de cliente | Tipo de cliente que visita entre las categorías: nuevo, existente, en riesgo. | CATEGORY |
| Tiempo de espera | Tiempo de espera en minutos desde ser atendido en recepción hasta ser llamado | FLOAT |
| Tiempo de resolución | Tiempo en minutos desde que la interacción comienza hasta que termina. | FLOAT |
| Tiempo promedio de resolución del centro | Tiempo promedio de resolución del centro en minutos | FLOAT |
| Tipo de visita | Categorización hecha por el agente, algunos de los tipos principales: activación, pagos, información, etc | CATEGORY |
| Primera visita | Agregado del tipo de visita, booleano si es la primera vez que el cliente visita el centro por el tipo de problema particular en los últimos 30 días. | BOOL |
| Tipo de visita nivel 3 | Categorización hecha por el agente dentro del árbol de decisión de problemas llevado al ancestro 3 | CATEGORY |
| Tipos únicos de problema | Agregado del tipo de visita nivel 3, es el número de problemas distintos de un centro o un cliente. | INT |
| Usuario de producto | Tipo de servicio contratado y modalidad | CATEGORY |
| Calificación al agente “Expert Helpfulness” | Respuesta numérica entre 1-5 sobre el atendimento del Agente | INT |
| CSAT | Respuesta numérica entre 1-5 sobre la experiencia general | |
| Calificación promedio del Agente de atención | Promedio de las calificaciones recibidas en una semana determinada para un agente | FLOAT |
| Demográficos del usuario-Edad | Edad del usuario | INT |
| Demográficos del usuario-Sexo | Sexo del usuario | CATEGORY |
| Volumen promedio de visitar al centro | Volumen de visitas ingresadas al centro por semana | FLOAT |
| Tiempo entre usuarios | Tiempo promedio del centro entre el atendimento de un usuario y el siguiente. | FLOAT |

| | | |
|----------------------------------|--|-------|
| Calificación interna del cliente | Calificación de 1 a 5 basada en un algoritmo interno | FLOAT |
|----------------------------------|--|-------|

6.2 Agrupación de las variables según su relación con aspectos del negocio

Inicialmente, cada variable debe de poder asociarse con un aspecto controlable del negocio. Este paso nos permite proveer recomendaciones accionables a la dirección de cómo mejorar cada aspecto operativo del negocio. Por ejemplo, mejorar las instalaciones, contratar más personal, o redoblar esfuerzos en entrenamiento y capacitación. La tabla 3 muestra las herramientas de control que pueden ser utilizadas para abordar cada una de las áreas operativas cuantificadas con las variables enlistadas.

Tabla 3: Variables mapeadas a las áreas del negocio

| Variable | Asociación Operativa | Herramienta de control |
|--|------------------------------------|-------------------------|
| Calificación al agente “Expert Helpfulness” | Calidad del agente | Capacitación |
| Calificación promedio del agente de atención | Calidad del agente | Capacitación |
| CSAT | Calidad del agente y procesos | Procesos |
| Tiempo promedio de resolución del centro | Capacidad de producción del centro | Eficiencia del personal |
| Tiempo de espera | Capacidad del centro | Eficiencia del personal |
| Estampa de llegada | Flujos de entrada | Eficiencia de colas |
| Estampa de llamado | Flujos de entrada | Eficiencia de colas |
| Volumen promedio de visitar al centro | Flujos de entrada | Eficiencia de colas |
| Estampa de cierre | Flujos de salida | Eficiencia del personal |
| Tipo de cliente | Segmentación | Especialización |
| Tipo de visita | Segmentación | Especialización |
| Tipo de visita nivel 3 | Segmentación | Especialización |
| Usuario de producto | Segmentación | Especialización |
| Demográficos del usuario-Edad | Segmentación | Especialización |
| Demográficos del usuario-Sexo | Segmentación | Especialización |
| Calificación interna del cliente | Segmentación | Especialización |
| Tiempo de resolución | Tiempo de producción | Capacitación |
| Tiempo entre usuarios | Tiempo de producción | Eficiencia del personal |
| Primera visita | Flujos de entrada | Procesos |

6.3 Segmentar las variables en variables primarias y secundarias:

Como siguiente paso de la metodología propuesta, haremos 2 grupos de variables: el primero, con la selección manual basada en sensibilidad de negocio; a estas variables se les referirá como “variables primarias”. La segunda selección extendida, que incluirá variables secundarias las cuales no estén sensiblemente relacionadas con la variable de salida, será el “grupo extendido”.

Variables primarias: variables elegidas bajo la luz del conocimiento del negocio buscando caracterizar al cliente y usar las métricas que naturalmente están impactando la experiencia:

- Tiempo de espera
- Tiempo de la interacción
- Tiempo total de la visita
- Calificación del agente
- Tipo de cliente
- Tipo de soporte/ sin tipo de soporte

Variables secundarias: variables no relacionadas sensiblemente con el negocio, es decir, que de primer entendimiento no tienen linealidad con lo percibido en la experiencia:

- Primera visita
- Problema recurrente
- Demográficos del usuario-edad
- Demográficos del usuario-género
- Calificación interna del cliente
- Día de la semana
- Tipo de producto
- Hora pico
- Usuarios esperando en la hora de visita

Algunas de las variables secundarias son agregadas o inferidas de los datos capturados, por ejemplo, para obtener la “primera visita”, los datos a nivel cliente son separados y se identifica si la visita ha sido primera en el último año. Similarmente para la “Hora Pico”, los datos a nivel centro de atención son agregados, y si una interacción sucede en la hora donde el centro recibe mayor volumen de visitas, la interacción es clasificada como en “Hora Pico”

De los grupos anteriores se podrá formar una selección final, posteriormente utilizando los mejores predictores de cada grupo y contrastarlo con el conocimiento de un experto en el sistema para validar la coherencia de las variables y darles un explicativo tangible de su relación real con el sistema.

Algunas de estas variables capturan naturalmente la experiencia, como los tiempos de espera o la calificación del agente atendido. En otra la relación no es tan natural, como podrían ser los demográficos del usuario o el número de visitas.

6.4 Identificar la variable de salida del sistema

En este caso de estudio, nuestra empresa busca optimizar el indicador de CSAT (*Customer Satisfaction Index*). Este es generado con las encuestas de salida. Al terminar una interacción, se envía a los clientes una encuesta de satisfacción a sus teléfonos. El 16% de las encuestas enviadas se responden, la encuesta tiene varias preguntas, entre ellas “¿En general, cómo calificaría el servicio?” en una escala de Likert de 5 puntos en emojis (mostrada en la figura 4). Cada emoji es traducido a un valor numérico discreto, desde 1 para una representación de una cara molesta a 5 para una cara satisfecha y sonriente.

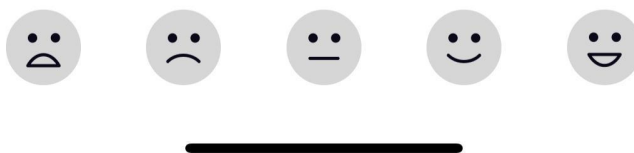


Fig 4. Escala de emojis utilizada para evaluar la experiencia de soporte.

La pregunta de la encuesta busca englobar la satisfacción generalizada del cliente. Por esta razón será utilizada como el indicador a optimizar en el modelo.

6.5 Hacer una priorización de los esfuerzos actuales

El siguiente paso en la metodología requiere identificar cómo se alinean los esfuerzos actuales del negocio para establecerlos en un nivel de prioridad y ver su relación con las métricas disponibles.

Esfuerzos y prioridades del área del negocio:

Cómo se menciona en la situación actual, el área de estrategia busca mejorar el indicador de servicio CSAT a través de la diferenciación del servicio. Buscar dar distintos niveles de servicio según el tipo de cliente que se presente al centro de atención.

Las variables en dónde se podría reflejar la segmentación del servicio son:

- Tipo de problema
- Calificación del cliente
- Demográficos
- Tipo de cliente

De manera secundaria la eficiencia en el servicio también es perseguida por el área de estrategia, pues se asocia a los costos generados por el soporte. La gestión busca sostener un nivel base de eficiencia, más no optimizarlo.

Las variables en donde se podría reflejar la eficiencia:

- Volumen de visitas
- Tiempos de espera
- Tiempos de resolución

Con un ejercicio de consenso con el área de estrategia se logró hacer un ranqueo de los esfuerzos-métricas de la dirección, representadas en el marco sugerido por Frei (Frei, 2012). La figura 5 muestra la alineación de las prioridades de la dirección. El fuerte enfoque en soporte segmentado en el tipo y demográficos del cliente los ponen en el tope de la jerarquía. Los tiempos totales y de espera son únicamente vistos como subproductos de la eficiencia del centro, dejándolos en segundo lugar. Al final se encuentran los elementos de desarrollo y entrenamiento de los agentes de soporte dónde actualmente no se destinan los esfuerzos (Frei, 2012).

Prioridades del negocio

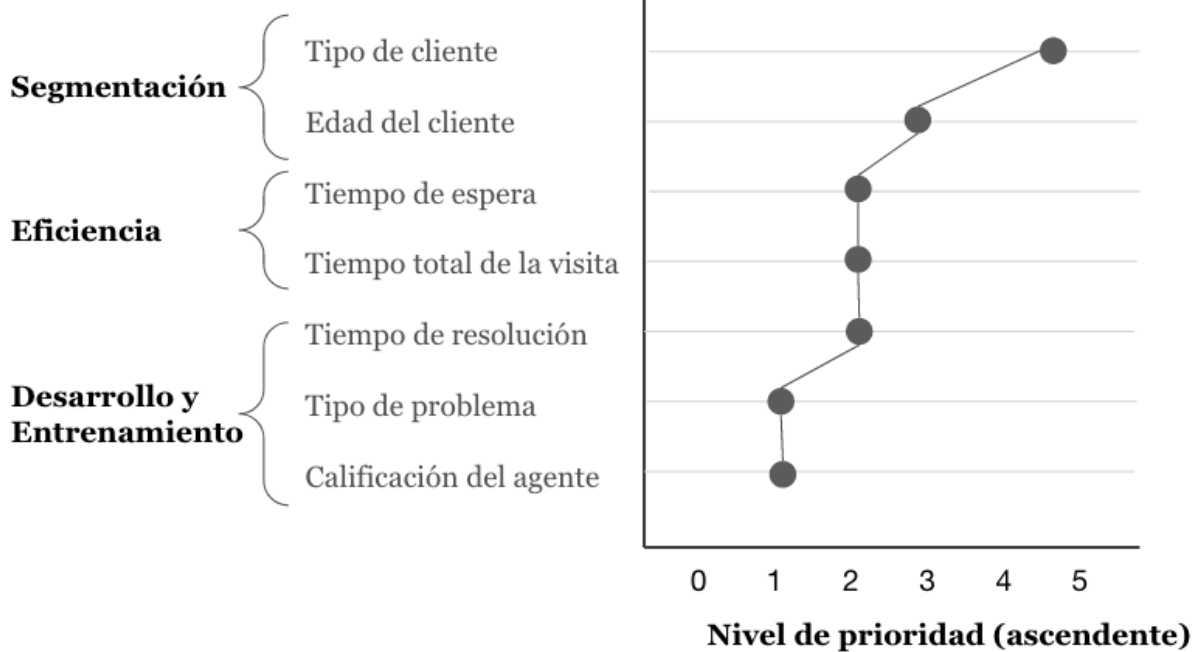


Fig 5. Mapeo prioridad-indicador

Como es mencionado en la sección de antecedentes, la dirección de nuestra compañía ha puesto la mayor inversión y esfuerzo en la creación de un soporte segmentado según el tipo de cliente, y sus datos demográficos. De forma menos acentuada pero igual presente, se busca siempre la eficiencia del centro de atención visto como un cargo operativo más que como un factor de la experiencia del cliente.

6.6 Análisis exploratorio

Este análisis exploratorio nos permite ver un poco más de cerca las descriptivas y estadísticas de las variables a observar, para darles los tratamientos correctos en el modelado. También nos permite tener una noción inicial sobre el estado actual de estado del sistema y cómo podemos mejorarlo.

6.6.a Base de datos y tratamiento

Comenzaremos con un análisis exploratorio clásico de datos sobre nuestro *dataset*, con los muestreos de experiencia de servicio.

Cada línea representa una visita a un centro de atención, con las mediciones de experiencia mencionadas en la sección anterior y su respectiva encuesta de satisfacción contestada.

Trabajaremos con una muestra de 121,415 interacciones de servicio, sucedidas en el año 2019 en 5 centros de servicio de la Ciudad de México, incluyendo la zona metropolitana, distribuidos de la siguiente manera: dos en la zona centro, uno en la zona sur, uno en el noreste y otro en el norte.

El 1% de los datos fueron removidos para evitar sesgos en los extremos superiores de los datos, por ejemplo, registros de tiempos de espera que no fueron cerrados.

La Tabla 4 muestra el número de encuestas capturadas para cada uno de los centros de soporte:

Tabla 4. Número de muestras

| Centro de Atención | Tamaño de la Muestra |
|--------------------|----------------------|
| Sitio Centro | 41,858 |
| Sitio Norte | 22,150 |
| Sitio Sur | 18,423 |
| Sitio Noreste | 11,180 |
| Sitio Aeropuerto | 20,512 |
| Total | 114,123 |

6.6.b Demográficos

En nuestra muestra de encuestas, el 95.6% de los clientes son hombres y el 4.4% restante, mujeres, con una edad mediana de 33 años.

6.6.c Distribuciones

Con el objetivo de tener un mejor entendimiento de los datos a trabajar, haremos un estudio simple de las distribuciones de cada una de las variables.

Tiempos de espera:

Los tiempos de espera se definen como el tiempo transcurrido entre el registro del cliente y el comienzo de la interacción con el representante.

Podemos observar en la figura 6 que la distribución de los tiempos de espera no es normal. En los cinco casos observados la distribución tiene un sesgo, es asimétrica con cola a la derecha. Este tipo de distribuciones es comúnmente observado en eventos no distribuidos, es decir, que existen horas pico donde los tiempos de espera tienden a ser mayores, pero el número de casos es menor. La mayoría de los clientes espera menos de 10 minutos, pero en horas pico el tiempo de espera puede superar los 50 minutos.

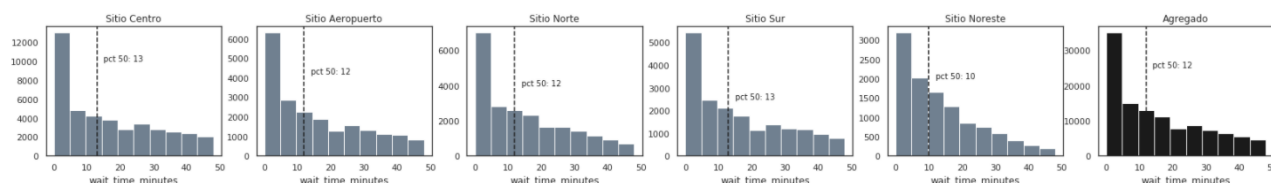


Fig. 6. Distribuciones de tiempos de espera

También podemos observar en la tabla 5, que la gran mayoría de los clientes (percentil 75) esperan menos de 30 minutos, y la mediana en los cinco casos se encuentra entre 10-13 minutos. En todos los centros tenemos números similares, que no necesariamente se alinean con el volumen de visitas y esto se debe a que los distintos centros tienen capacidades instaladas distintas.

Tabla 5. Distribución de los tiempos de espera en minutos

| Centro de Atención | Media | Mediana | Percentil 25 | Percentil 75 |
|--------------------|-------|---------|--------------|--------------|
| Sitio Centro | 16.25 | 13 | 2 | 27 |
| Sitio Norte | 14.85 | 12 | 3 | 24 |
| Sitio Sur | 15.96 | 13 | 3 | 27 |
| Sitio Noreste | 13.42 | 10 | 4 | 21 |
| Sitio Aeropuerto | 15.52 | 12 | 3 | 26 |
| Agregado | 15.52 | 12 | 3 | 26 |

Tiempo de interacción:

Por tiempo de interacción se entiende el tiempo transcurrido entre que el cliente comienza a ser atendido y el final de la interacción.

Al igual que los tiempos de espera, podemos observar en la figura 7 que las distribuciones de los tiempos de resolución tienen un sesgo de cola derecha.

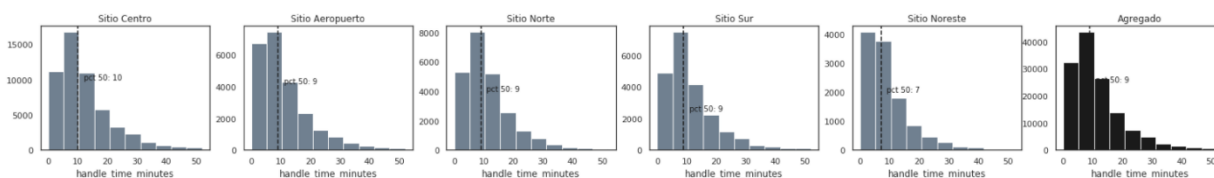


Fig 7. Distribuciones tiempo de resolución

Los agentes en los distintos centros de atención reciben el mismo entrenamiento y la lógica de soporte es la misma. Por lo anterior, los tiempos de resolución de los distintos centros debería de ser muy similar; en lo general lo es, pero hay dos centros que tienen tiempos distintos como podemos observar en la tabla 6.

Que los centros tengan tiempos iguales de resolución asume que la proporción y distribución de problemas en cada uno de los centros es la misma. Probaremos esto más adelante.

Tabla 6. Resumen tiempo de Resolución (minutos)

| Centro de Atención | Media | Mediana | Percentil 25 | Percentil 75 |
|--------------------|-------|---------|--------------|--------------|
| Sitio Centro | 12.38 | 10 | 6 | 16 |
| Sitio Norte | 11.5 | 9 | 6 | 15 |
| Sitio Sur | 11.52 | 9 | 6 | 16 |
| Sitio Noreste | 9.54 | 7 | 4 | 15 |
| Sitio Aeropuerto | 11.16 | 9 | 5 | 12 |
| Agregado | 11.62 | 9 | 6 | 15 |

Calificación de encuestas

Al terminar la interacción se envía a los clientes una encuesta de satisfacción a sus teléfonos. El 16% de las encuestas enviadas se responden; la encuesta tiene varias preguntas entre ellas “¿Cómo calificaría el servicio?” en una escala de Likert de 5 puntos en emojis. Cada emoji (figura 4) es traducido a un valor numérico discreto, desde 1 para una representación de una cara molesta a 5 para una cara satisfecha y sonriente.

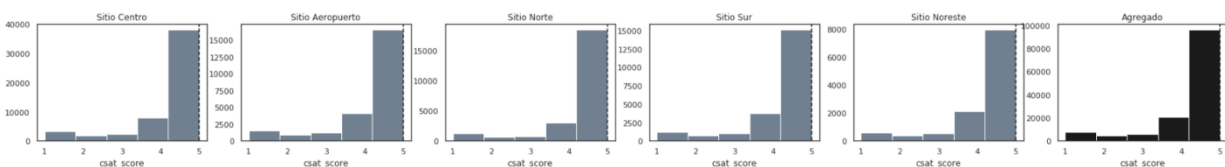


Fig 8. Distribuciones de encuestas de satisfacción

En la figura 8 podemos ver las distribuciones de los resultados de las encuestas. Nótese que el mayor volumen de respuestas se agrupa en 4 y 5, siendo 2,3 eventos raros y 1 ligeramente más común. Este es un efecto común en medición de satisfacción; las respuestas tienden a polarizarse formando distribuciones binomiales. Aunque no es el caso de nuestros datos, se puede observar el ejemplo.

Capacidad de los centros (clientes en espera)

Otro de los factores a considerar dentro de la experiencia del cliente en los centros de atención es el número de clientes en espera en un momento dado, es decir, la saturación del centro. Por lo general, los centros de atención suelen ser espaciosos, pero un espacio muy ocupado y tiempos de espera largos pueden generar una mala experiencia.

Para modelar esto, medimos cuántos clientes esperaban en el momento de la visita de otro cliente, es decir, si el cliente “n” visita el centro y en el momento hay 20 clientes esperando y 15 siendo atendidos, el cliente “n” percibirá una saturación de 35 clientes.

En la figura 9 podemos observar que algunos de los centros pueden ser altamente saturados con hasta 100 clientes esperando simultáneamente.

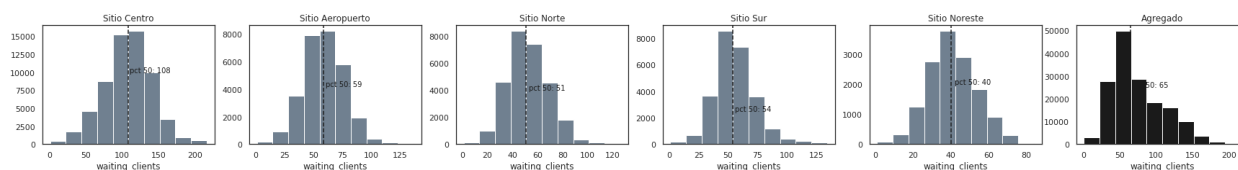


Fig 9. Distribuciones de número de clientes esperando simultáneamente

En la tabla 7 podemos notar de forma concreta el nivel de saturación de los centros. El sitio estándar tiene una capacidad de 80 personas. El Sitio Centro es considerado una central de soporte, su capacidad es de 350 personas.

Tabla 7. Agregados número de clientes esperando simultáneamente

| Centro de Atención | Mediana | Percentil 25 | Percentil 75 |
|--------------------|---------|--------------|--------------|
| Sitio Centro | 108 | 86 | 129 |
| Sitio Norte | 51 | 40 | 63 |
| Sitio Sur | 54 | 45 | 66 |
| Sitio Noreste | 40 | 32 | 49 |
| Sitio Aeropuerto | 59 | 47 | 71 |
| Agregado | 65 | 47 | 98 |

Conclusiones del análisis de distribuciones y muestreo:

- El número de muestras es sustancial y potencialmente adecuado para un modelo de aprendizaje automático pues las variables no están correlacionadas. El volumen de muestras es suficiente para tomar al menos 10,000 observaciones de cada centro. Las variables están etiquetadas y sin datos faltantes en cada una de las dimensiones.
- Como resultado de la estandarización operativa de los centros de atención (mismos recursos, mismos entrenamientos, capacidades similares), las distribuciones de las métricas de desempeño de los centros son homogéneas. Esto apunta a que se pueden utilizar observaciones de distintos centros para modelar la experiencia generalizada.
- La mayoría de las métricas requerirá tratamiento para normalizar las distribuciones, pues las dimensiones en distintas escalas pueden alterar el desempeño del modelo y el número de ocurrencias de cada uno de los eventos no es equivalente.
- Dos datos de CSAT tienen un sesgo importante. La mayoría de las respuestas se acumulan en el 5, se requerirá un muestreo sintético o ajustado para balancear el set de datos. Al ser eventos raros las calificaciones bajas es posible que los modelos hagan *overfitting* ante el gran volumen de respuestas con 5. Para resolver lo anterior, podría hacerse muestreo sintético para agregar muestras de los eventos raros o reducir intencionalmente las respuestas con 5 para ecualizar el set de datos. Dado el vasto volumen de observaciones, se optará por hacer submuestreo aleatorio.

6.6.d Estudio correlacional

En la siguiente sección buscaremos correlaciones entre las variables para evitar que los modelos fallen por multicolinealidad. Lo anterior es importante porque muchos algoritmos de regresión pueden fallar si la correlación en las métricas es alta (Bruce et al, 2018).

Tiempos de espera y encuesta de satisfacción

Una de las hipótesis comunes es que el tiempo de espera puede ser un factor definitivo sobre la percepción de calidad, para esto, agregamos los datos de las encuestas por día, así como los tiempos de espera promedio por día por centro de atención. Lo anterior se hizo como una técnica de tratamiento de datos para evitar calcular correlaciones sobre variables discretas como lo son las respuestas de las encuestas de satisfacción.

La figura 10 muestra la correlación entre los tiempos de espera y la calificación de la encuesta. El coeficiente de correlación (Pearson) es muy bajo, es de -0.17 , lo que indica que no existe evidencia fuerte de que la medida de cambio de tiempo de espera no influye por sí sola en la calificación del cliente.

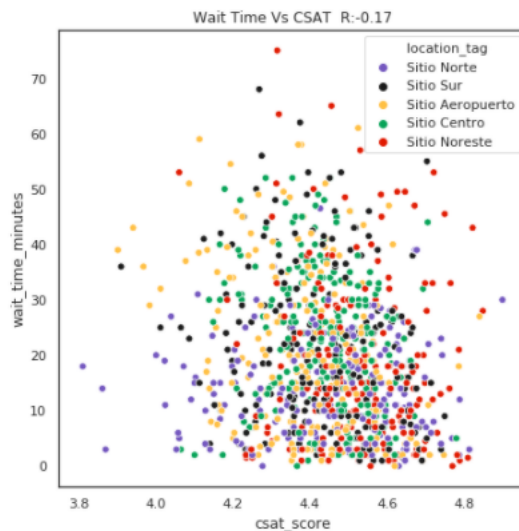


Fig 10. Correlación Tiempo de espera contra encuesta de satisfacción agregación por día

Tiempos de resolución y encuesta de satisfacción

Otro de los supuestos comunes en el servicio al cliente es que el tiempo de resolución es un factor clave para pronosticar la experiencia del soporte de manera general. Siguiendo la metodología anterior, logramos mostrar la relación que existe entre ambas variables.

La figura 11 muestra que la correlación entre las dos variables es virtualmente inexistente (-0.02 Pearson). Existen casos donde el tiempo de resolución es muy bajo y la experiencia pobre; y casos contrarios donde el tiempo de resolución es alto y la experiencia es satisfactoria.

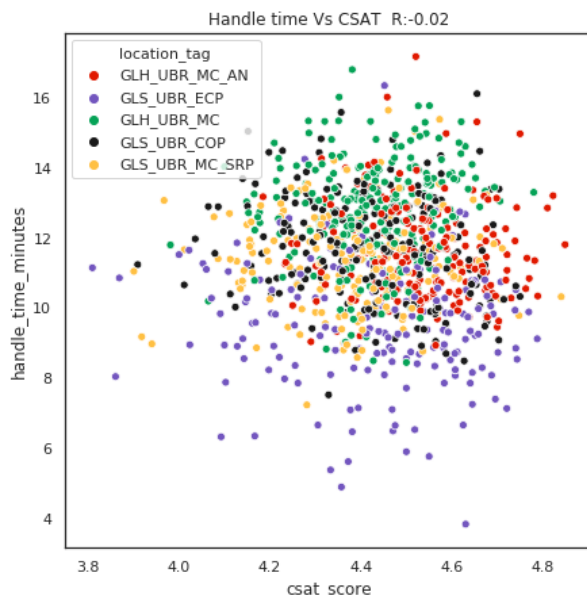


Fig 11. Correlación tiempo de resolución contra encuesta de satisfacción agregación por día

Calificación del agente de soporte y encuesta de satisfacción

Otra hipótesis se relaciona con la atención otorgada por el agente de soporte y la relación que existe con la experiencia general de servicio. Es común pensar que la calidad del soporte puede ser entendida por la calidez y eficiencia de los agentes.

Para explorar esta hipótesis estudiamos la relación existente entre la satisfacción general de la interacción de soporte y la calificación promedio de cada uno de los agentes.

Aunque la correlación no es fuerte (0.54) podemos empezar a ver un patrón en la figura 12; la calificación de los agentes sí juega una parte en la experiencia general percibida por el cliente. No podemos afirmar

causalidad a partir de la correlación, ni decir que este es el único factor definitorio de la calidad de la experiencia. Pero la correlación nos da una guía sobre las variables que serán buen predictor de la calidad.

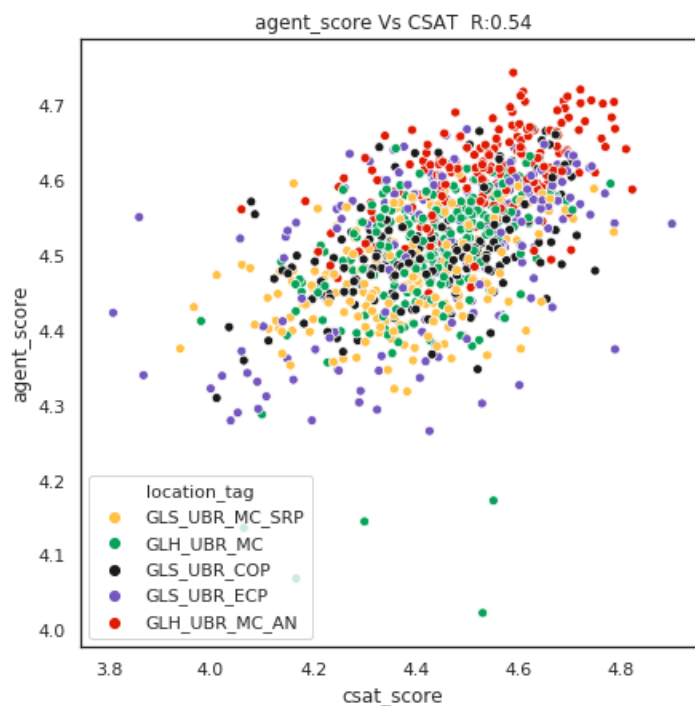


Fig 12. Correlación calificación del agente contra encuesta de satisfacción agregación por día

6.7 - Optimización de variables a través de la iteración de un modelo simple

Justificación de técnicas y configuraciones generalizadas

Antes de iniciar con las distintas iteraciones de modelos y variables, en esta breve sección se explica el racional de la selección de técnicas aplicadas a todos los modelos consiguientes.

Segmentación set de prueba y set de entrenamiento:

- Se utilizará una muestra de 25% de los datos para el set de prueba, y 75% de los datos para el set de entrenamiento.

Modelo simple a utilizar:

- La metodología no especifica un modelo particular, pero sugiere un modelo simple de fácil implementación, por lo que utilizaremos un modelo de regresión logística para que nos permita obtener el grado de pertenencia a cada una de las categorías. Una de las ventajas principales de utilizar un modelo de regresión logística es que obtendremos un resultado entre cero y uno que representa el grado de membresía a cada uno de los grupos.
- La estrategia será un *One vs Rest*, lo que quiere decir que trataremos cada categoría (posibles calificaciones 1-5) como una categoría booleana; después entraremos a un modelo de regresión logística para cada uno de ellos, individualmente. Así tendremos una función de membresía para cada posible respuesta, la categoría (posible respuesta) con un valor de membresía mayor será considerado la predicción.

Valores tokenizados:

- La variable de tipo de visita es una cadena (string) que identifica el problema diagnosticado por el agente de soporte.
- Tokenizar esta variable agrega 62 columnas al *dataset*, aumentando significativamente el costo computacional del mismo, sin una ganancia tangible en la interpretación final del modelo. Operativamente no sería viable diseñar la estrategia alrededor de un solo tipo de problema.

Evaluación de los modelos:

- Dado que los falsos negativos y falsos positivos no tienen un peso en particular, utilizaremos *Accuracy* para medir el desempeño del modelo. Este se define como la razón de valores correctamente categorizados contra el número de observaciones de la clase determinada.

6.7.a - Validación y balanceo del set de datos

Como pudimos ver en el análisis exploratorio, la distribución de respuestas en las encuestas de satisfacción (CSAT) no es homogénea. Las encuestas respondidas con un 5 sobrepasan los volúmenes de otras respuestas en una proporción de 1:10.

En un modelo de prueba inicial (regresión lineal múltiple one-vs-rest). Sin aplicar ningún tipo de técnica de balanceo de muestras, obtuvimos un modelo donde el 99% de las predicciones eran un 5 y a pesar de

tener un *accuracy score* de 71.7% el modelo es un caso ejemplar de *overfitting*, por lo cual necesitaremos balancear la muestra como se ve en la tabla 8.

Tabla 8. Tabla mostrando *overfitting* del modelo

| CSAT Score | % de predicciones |
|------------|-------------------|
| 1 | 0.30% |
| 2 | 0.00% |
| 3 | 0.00% |
| 4 | 0.00% |
| 5 | 99.70% |
| Agregado | 100.00% |

El set de datos se balanceó con submuestreo aleatorio utilizado 25,000 lecturas, 5,000 muestras aleatorias para cada tipo de respuesta (1-5) como se muestra en la tabla 9.

Tabla 9. Tabla resumen del muestreo

| CSAT Score | Muestras | Calif. Agente (prom.) | Tiempo de resolución (prom. Mins) | Tiempo de espera (prom. Mins) | Tiempo total de la visita (prom. Mins) | Tipos únicos de problemas |
|------------|----------|-----------------------|-----------------------------------|-------------------------------|--|---------------------------|
| 1 | 5000 | 1.06 | 12.91 | 27.44 | 40.34 | 41 |
| 2 | 5000 | 1.75 | 12.62 | 27.34 | 39.96 | 39 |
| 3 | 5000 | 2.29 | 12.48 | 26.82 | 39.41 | 41 |
| 4 | 5000 | 3.05 | 11.45 | 23.83 | 33.29 | 43 |
| 5 | 5000 | 4.06 | 11.67 | 22.26 | 33.93 | 38 |
| Agregado | 25000 | 2.44 | 12.24 | 25.54 | 37.78 | 55 |

Utilizando nuevamente las variables primarias (tiempo de espera, tiempo de la interacción, tiempo total de la visita, calificación del agente, tipo de cliente) como entrada, y las calificaciones de las encuestas como variables de salida. Se volvió a ejecutar la regresión logística one-vs-rest pero utilizando el *dataset* balanceado con submuestreo aleatorio con 20% de los datos para cada una de las posibles respuestas del 1 al 5.

La “exactitud” o *accuracy* del modelo con el *dataset* balanceado fue de 0.47, lo que significa que únicamente en el 47% de los casos del test *dataset* los casos fueron categorizados correctamente. Al estar sub-ajustado el modelo, posteriormente buscaremos aplicar tratamientos a los datos para resolver el problema. A pesar de que el modelo con el *dataset* completo tuvo un *accuracy* mayor (0.71), este no tiene *overfitting* (*accuracy* en test de entrenamiento 0.76) pues las predicciones no se concentran en una única clase; las clases en el set de entrada están balanceadas.

6.7.b - Integración de variables secundarias

Ahora integraremos variables tangenciales a la experiencia de servicio, es decir, que no son comúnmente relacionadas a la satisfacción. El objetivo es validar si alguna de estas variables tiene cualidades predictivas para ser añadidas a las variables primarias y así mejorar la exactitud del modelo.

Variabes secundarias:

Las variables secundarias pueden ser “variables de confusión” al no estar sensiblemente relacionadas con la experiencia del cliente, pero podrían influir indirectamente. Las variables son un subconjunto del *data set* utilizado anteriormente.

- **Primera visita:** Booleana, cierta si es la primera visita del cliente en los últimos 30 días.
- **Problema recurrente:** Booleana, cierta hay registros anteriores del cliente resolviendo el mismo problema en un centro de soporte.
- **Demográficos del usuario-edad:** Edad del cliente extraída del CURP de la licencia.
- **Demográficos del usuario-género:** Sexo del cliente extraído del CURP de la licencia.
- **Calificación interna del cliente:** Calificación del 1 al 5 del cliente.
- **Día de la semana:** Día de la semana en el cual sucedió la visita al centro de soporte.

- **Tipo de producto:** Producto al cual el cliente es asociado según su uso en la plataforma.
- **Hora pico:** Booleano, cierto si la visita ocurrió en la hora del día con mayor número de visitas, según la distribución por centro.
- **Usuarios esperando en la hora visita:** Número de clientes esperando simultáneamente a la hora de la visita del cliente.

Análisis de resultados:

La iteración con variables secundarias del modelo arrojó resultados de 23% de exactitud, casi 25 puntos porcentuales debajo de la iteración con variables primarias. Los resultados obtenidos no son extraños, es de esperarse que ante una selección de variables no lineales las cualidades predictivas de cada una de las variables son menores.

A pesar de ello, utilizaremos un método de optimización de *grid search* con las variables, así podremos identificar el valor predictivo de cada una de las variables al eliminarla de las variables de entrada y observar el cambio en el desempeño del modelo.

Las variables que encontremos utilizando este modelo las agregaremos al set de variables primarias en búsqueda de mejorar el desempeño del modelo.

Grid search diferencial:

Para encontrar las variables secundarias más importantes entrenaremos el modelo quitando una de las variables por vez y observaremos la diferencia en el desempeño. La tabla 10 muestra la variación del desempeño del modelo al eliminar individualmente cada una de las métricas.

Tabla 10. Incremento del desempeño de cada variable secundarios

| Métrica | Línea Base | Desempeño sin la Métrica | Delta (desempeño - línea base) |
|---------------------------------|------------|--------------------------|-----------------------------------|
| Primera visita | 23.3% | 22.7% | -0.60% |
| Problema recurrente | 23.3% | 23.4% | 0.03% |
| Producto | 23.3% | 23.3% | 0.00% |
| Demográficos del usuario-edad | 23.3% | 21.9% | -1.45% |
| Demográficos del usuario-género | 23.3% | 23.0% | 0.00% |

| | | | |
|--------------------------------------|--------|--------|--------|
| Calificación interna del cliente | 23.3% | 23.0% | -0.36% |
| Día de la semana | 23.3% | 23.3% | 0.00% |
| Hora pico | 23.3% | 23.8% | 0.48% |
| Usuarios esperando en la hora visita | 23.3% | 23.8% | 0.44% |
| Modelo ajustado | 23.30% | 23.60% | 0.28% |

Después de hacer el *grid search* podemos elegir qué variables conservar de las variables secundarias. Las tres variables que mostraron una mejora en el desempeño del modelo fueron: edad, indicador de primera visita y la calificación interna del cliente.

Añadiremos únicamente la variable de los demográficos al set final de variables pues las otras dos no muestran un cambio significativo.

6.7.c - Simplificación de variables de salida

Los resultados de las iteraciones anteriores y el análisis exploratorio muestran que los límites entre cada una de las calificaciones son difusos; una experiencia con 4 o 5 de calificación pueden ser virtualmente idénticas y el modelo no es capaz de identificarlas puntualmente.

Una técnica para corregir ante estas situaciones es simplificar las variables de salida en elementos más simples y sensibles al modelo.

Para nuestro caso particular, agrupamos (por frecuencias) las distintas calificaciones; agrupamos las muestras donde la calificación fue 4 o 5 en “buena experiencia”, y las muestras con 1, 2 o 3, en “mala experiencia”, basándonos en el objetivo de negocio de tener una mediana arriba de 4. De esta forma podremos facilitar la categorización de las muestras pues las buenas y malas experiencias son diametralmente distintas.

Para este paso de la metodología, se sugiere utilizar las primarias (tiempo de espera, tiempo de la interacción, tiempo total de la visita, calificación del agente, tipo de cliente) con un modelo simple, pues lo que se busca es entender el cambio del desempeño del modelo al simplificar las variables de salida.

Utilizaremos un modelo multiclase. En iteraciones siguientes explicaremos modelos de una clase, dado que la variable de salida fue simplificada.

Resultados:

Al simplificar las variables de salida, el modelo arrojó una exactitud de 79.1%, superando radicalmente el desempeño del modelo simple con variables principales y muestreo balanceado. Esto indica la eficacia de la simplificación de la variable de salida.

Para el modelo final utilizaremos la variable de salida binaria agrupada: “buena experiencia” si la calificación fue 4 o 5, y “mala experiencia” si la calificación fue 1, 2 o 3. Esto nos permitirá entenderlo como un modelo biclase y no cómo un modelo multiclase.

6.7.d - Configuración y selección final de las variables

Los tres subprocesos anteriores nos permiten diseñar nuestro set de datos de forma ordenada e informada. La utilización de modelos simples abre la posibilidad de iterar rápidamente el modelo sin invertir mucho tiempo en su implementación.

Las configuraciones finales, producto de los subprocesos de la metodología son:

- **Balanceo de muestra:** dado el desbalance de las calificaciones de las encuestas y los valores intermedios (2,3,4), siendo eventos raros, balancear el *dataset* es recomendable para poder producir el modelo.
- **Variables secundarias:** La variable secundaria de los demográficos del cliente será añadida a la selección final de variables, en conjunto con las variables primarias, pues demostró que puede tener poder descriptivo.
- **Simplificación de la variable de salida:** Al ser las encuestas sujetas a percepción, los límites entre las calificaciones son difusos a pesar de ser valores nominales. Por lo anterior se eligió agrupar los valores de las encuestas de satisfacción de forma binomial (buena experiencia / mala experiencia).

6.8 - Selección de modelos

El siguiente paso de la metodología consiste en utilizar la selección de variables especificada en el paso anterior con distintos modelos de aprendizaje máquina. En este caso, modelos de categorización hasta encontrar el modelo que mejor logre ajustarse a nuestro indicador a optimizar, CSAT.

El ajuste y calibración de los hiper parámetros de cada uno de los distintos modelos está fuera del alcance de este trabajo. Pero de manera general se utilizó un *grid search* y métodos de imputación y escalamiento para las variables de entrada.

Los modelos utilizados fueron:

- Regresión logística.
- *K-Nearest neighbors*
- Random forest
- Decision tree
- Support vector machine
- Multi-layer perceptron classifier

Los métodos utilizados fueron seleccionados según su complejidad de implementación y su adecuación al *dataset*.

Los resultados de cada uno de los modelos se muestran en la figura 13.

Precisión Individual por Modelo

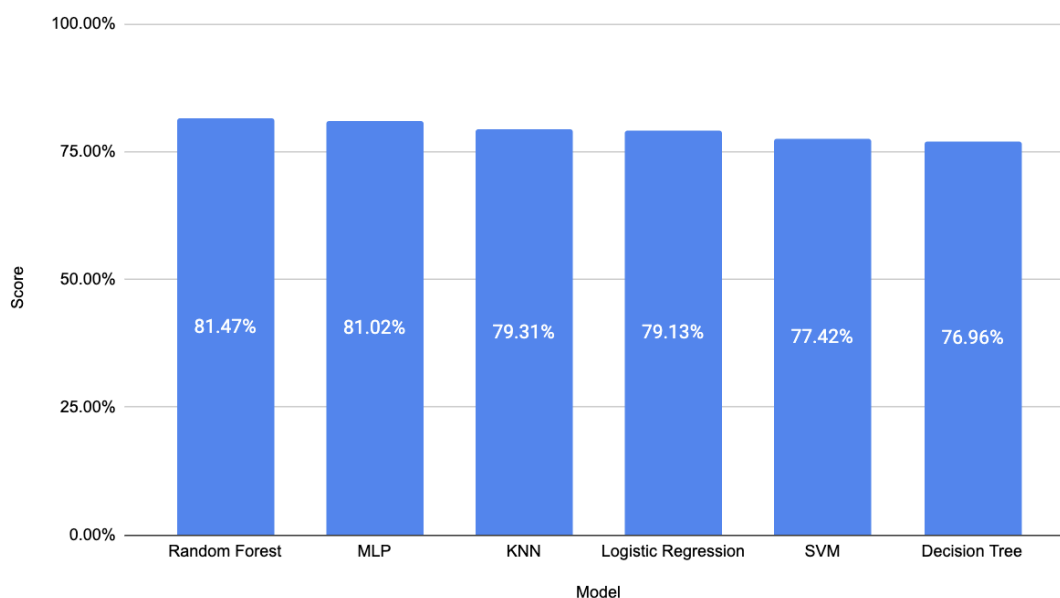


Fig 13. Desempeño individual de los modelos.

El modelo con mejores resultados fue el Random Forrest, con 100 estimadores (árboles) y una profundidad mínima de 3 ramas por árbol.

Como muestran los resultados, la mayoría de los modelos convergen en el rango de 77-81.5%, es importante recordar que con elementos humanos en las encuestas hay una porción de los datos que no puede ser explicada por un modelo con este nivel de complejidad (Farris, 2010).

Podría seguir buscándose una mejora en el desempeño del modelo con tratamientos en las variables y modelos embebidos, pero para el horizonte de este trabajo nos quedaremos con este modelo.

7.8.a - Modelo final

Después del análisis y evaluación de modelos, utilizaremos el algoritmo de Random Forest para obtener el modelo de la evaluación de la experiencia de servicio con las variables principales.

Utilizaremos también la versión binaria de la encuesta de satisfacción (buena experiencia/ mala experiencia) pues mostró mejores resultados y las conclusiones del modelo no pierden sensibilidad. La experiencia y las acciones a tomar después de los resultados pueden ser usadas de la misma manera.

Utilizaremos el modelo con 81% de precisión – al ser las salidas de las muestras elementos auto reportados por personas.

Modelo e hiper parámetros:

Random Forrest – profundidad mínima 3

```
'max_depth': 3,  
'min_samples_leaf': 4,  
'min_samples_split': 10,  
'n_estimators': 400
```

Los hiper parámetros ideales fueron encontrados a través de un *grid search* aleatorio de la paquetería *scikit learn*.

Variables:

- Tiempo de espera, Tiempo de la interacción, Tiempo total de la visita, Calificación del agente, Tipo de cliente, Edad del usuario

6.9 - Evaluar el poder predictivo de cada variable:

En esta sección utilizaremos el modelo final como base para hacer un ranking de las variables de servicio que modelan el servicio. La forma de evaluar cada una de nuestras variables es a través de su capacidad predictiva. El corolario es que, a mayor capacidad predictiva de una determinada variable, mayor la importancia que tiene en la definición de una buena contra una mala experiencia. Esta metodología es una adaptación del análisis iterativo de ingeniería de variables y puesta en práctica a través del marco de trabajo de este estudio (Sheikh, 2017).

Para evaluar cada variable, entrenaremos el mismo modelo bajo las mismas condiciones, removiendo una variable en cada iteración y observando el cambio en el desempeño del modelo. A mayor cambio mayor importancia. Una vez que todas las variables hayan sido sometidas a este proceso ordenaremos las variables según su cambio absoluto en el modelo.

Evaluando el cambio en el desempeño del modelo, aislando cada una de las variables finales, podemos observar que el cambio se relaciona con la capacidad predictiva y descriptiva de la experiencia de servicio para cada una de las variables.

Empezando con una línea base de 81.4%, y estudiando el cambio en el desempeño del modelo, se hizo un ranqueo de las dimensiones, resumido en la tabla 11.

Tabla 11. Ranking the variación del score del modelo

| Métrica | Línea Base | Desempeño sin la Métrica | Delta | Ranking |
|---------------------------|------------|--------------------------|---------|---------|
| Calificación del agente | 81.5% | 64.4% | -17.12% | 1 |
| Tiempo total de la visita | 81.5% | 71.3% | -10.17% | 2 |
| Tiempo de resolución | 81.5% | 75.3% | -6.13% | 3 |
| Tipo de problema | 81.5% | 76.3% | -5.15% | 4 |
| Tiempo de espera | 81.5% | 78.5% | -2.97% | 5 |
| Edad del cliente | 81.5% | 80.5% | -0.93% | 6 |
| Tipo de cliente | 81.5% | 80.7% | -0.82% | 7 |

6.10 - Sensibilización de las variables con expertos en el sistema:

Con la priorización producida por el modelo, en conjunto con expertos en el sistema a analizar, se busca racionalizar cada una de las variables para buscar discrepancias o puntos ciegos en la estrategia actual.

Se hizo un grupo de trabajo con gerentes de los centros de soporte, agentes y miembros del equipo de estrategia para buscar posibles relaciones entre las variables y lo observado en el campo.

Calificación del agente:

La variable con mayor peso en el modelo fue la calificación del agente, lo que quiere decir que parte de la definición de una buena o mala experiencia puede explicarse por la calificación del agente.

La calificación del agente se mide de manera similar con una pregunta en la encuesta de salida, al final de la interacción, para el modelo en que se utilizó el promedio de las calificaciones del agente de un mes al momento de completarse la interacción con el cliente.

La calificación del agente puede traducirse en elementos operativos, tal como los protocolos de comunicación con el cliente, el entrenamiento que recibe, su capacidad individual de diagnóstico y resolución de problemas, así como su uso y manejo de las herramientas.

Tiempo total de la visita:

La segunda variable con más peso en el modelo es el tiempo total de la visita. El tiempo total de la visita se define como la suma del tiempo de espera y el tiempo de resolución; razón por la cual la suma tiene implicaciones operativas distintas a sus elementos individuales. En algunas ocasiones el tiempo de espera puede ser muy corto pero el tiempo de resolución es largo, lo que puede igual percibirse como una mala experiencia. De manera similar, el tiempo de espera puede ser largo según la complejidad a resolver, del problema del cliente. Por esta razón, la experiencia percibida puede ser afectada de manera distinta por el tiempo total de la visita.

La experiencia de los centros puede ser mejorada al controlar de manera integral el tiempo total que los clientes invierten en el centro de atención y no únicamente en alguno de los componentes.

Tiempo de resolución:

Posteriormente en los pesos sigue la variable de tiempos de resolución. El tiempo de resolución es el tiempo que le toma al agente completar la interacción de soporte y dar seguimiento al problema del cliente. Los tiempos de resolución pueden variar, según la complejidad del problema y la lógica de soporte determinada para el problema. Empíricamente, un problema percibido como sencillo tiene la expectativa de ser resuelto rápidamente y existe mayor tolerancia al tiempo de resolución, cuando el problema es percibido como complejo.

Se puede mejorar la experiencia del cliente controlando y estandarizando la lógica de soporte para establecer tiempos máximos de resolución al tipo de problema. Otra forma de atender el problema puede ser manejando correctamente las expectativas del cliente durante el diagnóstico del problema y comunicar con mayor transparencia la complejidad para así expandir el horizonte esperado por el cliente.

Tipo de problema

Esta es la categoría dada por el agente tras el diagnóstico del problema. El tipo de problema da la pauta de la lógica de soporte a seguir, tras identificar el problema del cliente. De igual forma, el tipo de problema es un indicador clave de la complejidad del problema y también si hay una lógica de soporte disponible para el problema o no.

El tipo de problema para el cliente puede ser imperceptible. El agente se encarga de hacer la mayor parte del proceso de captura y diagnóstico. Se ha observado en mesas redondas con clientes que al no comunicarse con claridad puede generar una sensación de falta de control al cliente al no establecer correctamente la calibración de expectativas.

Tiempo de espera

El tiempo de espera es el tiempo transcurrido entre el registro del cliente en la recepción, en lo que es llamado por un agente de soporte para iniciar la interacción. En un contexto global, el tiempo de espera es un indicador clave de la satisfacción del cliente. En nuestro modelo fue uno de los indicadores débiles.

En nuestro contexto particular el tiempo de espera es prioritario. El 93% de los clientes son atendidos en menos de 20 minutos, lo cual es un estándar bastante exigente para la industria del servicio (Accenture, 2017). Por esta razón no es extraño que el indicador se posicione en un lugar bajo en el rango.

Edad del cliente

Los demográficos nos muestran cómo los clientes interactúan con los sistemas de soporte y nuestros productos en general. La edad del cliente fue tomada de los perfiles de cada uno de los clientes. Por cuestiones de privacidad se usan rangos de 5 años.

Los procesos de los centros de atención se diseñan para poder asistir a los clientes, sin importar sus grupos demográficos. Los agentes son capacitados para que pueden asistir de manera clara a cualquier grupo, incluyendo personas con discapacidad.

Dado que los procesos de soporte son agnósticos a los diversos grupos de clientes, no es extraño que el modelo se comporte de la misma manera.

Tipo de cliente

El indicador menos importante del modelo fue el tipo de cliente. El tipo de cliente se define en la etapa de consumo del cliente. Si es un cliente que apenas va a adquirir el servicio se considera un cliente nuevo. De manera similar, si el cliente ya tiene un historial existente y recibe servicio post venta, se cataloga como un cliente existente.

Este indicador fue el más débil para el desempeño del modelo. No hay una explicación más clara que la variedad de posibles problemas, para nuevos clientes, es similar a la variedad de problemas para clientes existentes, sin necesariamente ser un indicador fuerte o diferenciador en la experiencia de cada uno de estos tipos de clientes.

6.11 - Contraste contra las prioridades del negocio

Retomando el marco de alineaciones podemos visualizar, de mejor manera, la alineación del modelo y de las prioridades. En este esquema (*fig. 14*) tenemos en el eje *y* las cualidades del ecosistema de servicio; y en el eje *x*, tenemos el nivel de prioridad otorgado por el negocio y el cliente. En un modelo balanceado las prioridades para el cliente estarían perfectamente alineadas con las prioridades del negocio (Frei 2012).

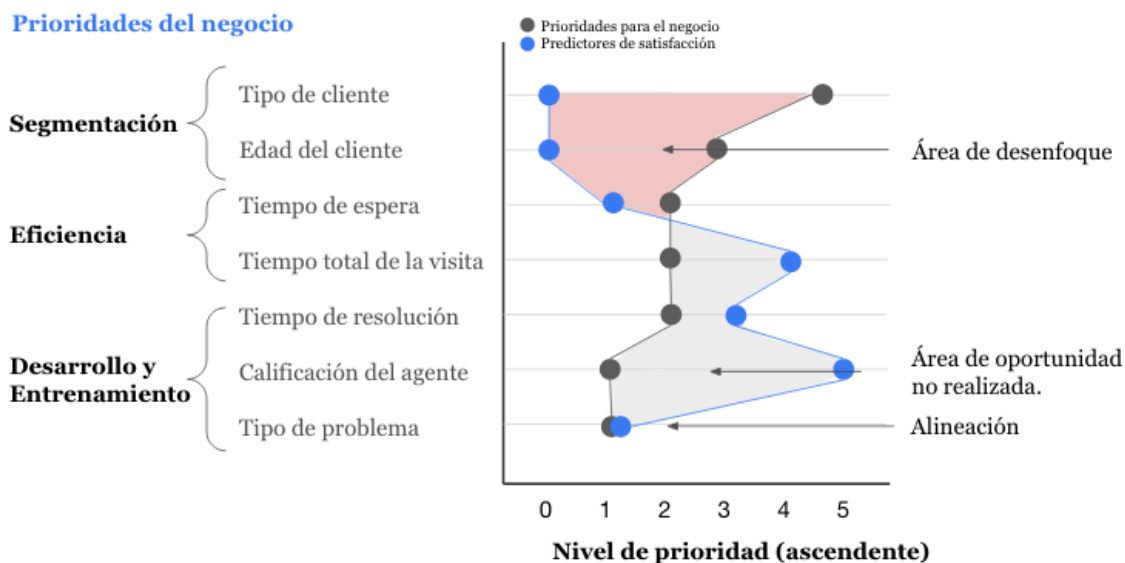


Fig 14. Marco de alineaciones generado a partir del modelo predictivo

Las brechas entre las prioridades nos permiten identificar fácilmente los cambios operativos necesarios para mejorar la experiencia percibida del cliente. Primeramente, en las áreas donde la prioridad del negocio es diametralmente distinta a las prioridades del cliente tenemos las “áreas de desenfoque”, sombreadas en rojo en la *fig. 14*.

Estas nos indican enfoques de la empresa. En nuestro caso de estudio, la empresa ha iniciado varios esfuerzos e iniciativas para entregar un soporte diferenciado según la etapa del ciclo del consumidor del cliente. La intención por parte de la empresa es entregar un servicio más completo y personalizado a clientes nuevos para así acelerar los ciclos de conversión.

Nuestra metodología nos apunta a que las iniciativas anteriores no mostraron resultados y que invertir más en iniciativas relacionadas al tipo de cliente no mejorará la percepción del nivel de servicio.

Por otra parte, tenemos las áreas donde las prioridades del cliente son mayores a las prioridades de la empresa. Estas áreas se denominan “áreas de oportunidad no realizada”. Estas áreas son puntos ciegos para la organización. En nuestro caso de estudio la calificación del agente es únicamente considerada como un factor de desempeño interno del colaborador, pero no como un descriptor de la experiencia de servicio.

Los resultados de nuestra metodología muestran que los clientes perciben el indicador de desempeño del agente como un factor de calidad en el servicio entregado. Una recomendación directa para nuestra organización será enfocar la contratación y entrenamiento de los agentes de soporte para optimizar sobre este indicador.

Otra observación naciente de nuestros resultados está ligada al tiempo total de espera. Actualmente la organización se enfoca en los tiempos de resolución mayoritariamente, pues el tiempo del cliente con el agente es la constricción del *throughput*. Los tiempos de espera individualmente no representan un uso significativo de recursos para el negocio y al ser un ecosistema de soporte, los clientes suelen tolerar más los tiempos de espera.

Nuestros resultados muestran que el cliente no únicamente tiene interés porque su problema sea resuelto rápidamente, sino el tiempo total que pasa en el centro de soporte. Un tiempo de espera corto pierde su valor si el tiempo de resolución es largo, y un tiempo de resolución corto no es apreciado de la misma

forma si fue antecedido por un tiempo de espera largo. La recomendación para la organización será optimizar los tiempos de espera y de atención de forma conjunta para mejorar la percepción del cliente.

Nuestra metodología nos permite identificar las áreas anteriores de manera efectiva y escalable para analizar rápidamente un sistema del negocio y cuantificar sus ángulos para así compararlos contra las prioridades de la administración del negocio. Así se pone a la luz sencillamente las oportunidades para mejorar y redireccionar los esfuerzos del negocio hacia un objetivo claro, medible y cuantificable.

6.12 Utilizar las zonas de desenfoco para desarrollar mejoras

El modelo nos permitió generar de manera programática, y basada en datos duros, el marco de alineaciones que saca a la luz las brechas que existen entre las prioridades de los equipos de operaciones y estrategia y las necesidades de los clientes. El marco muestra que las áreas de oportunidad no realizadas se encuentran principalmente en la calificación del agente, es decir, la calidad y nivel de servicio que el agente imparte independientemente del tipo de problema o tipo de soporte que da.

La segunda área de oportunidad no realizada estaba en el tiempo total de la visita, que equivale al tiempo de espera más el tiempo de resolución. La diferencia entre observar el tiempo total de la visita y sus partes individualmente está en que son dos procesos accionables distintos, los tiempos de espera se ligan a la capacidad en términos de empleados y los tiempos de resolución se ligan la lógica de soporte y la eficiencia de los agentes para diagnosticar el problema y dar el servicio.

De manera similar y como control, también observamos la prioridad inicial que era el servicio diferenciado. Donde a un grupo de clientes identificados como de alto valor se les daba un servicio prioritario, como poder agendar una cita, soporte telefónico y políticas más flexibles.

Las recomendaciones generadas tras el ejercicio con el área de operaciones y estrategia fueron las siguientes. Para la calificación del cliente, se propuso entender los factores que diferenciaban a los agentes mejores capacitados, y desarrollar entrenamientos basados en las cualidades. De igual manera el programa de incentivos para los agentes se enfocaría más en este indicador.

Para el tiempo total de la visita, se identificaron cuáles eran los subprocesos que alargaban el tiempo total de la visita entre ellos, la carga inicial de datos y el diagnóstico inicial que hace el agente. La propuesta fue utilizar el tiempo en la sala de espera para iniciar algunos de estos procesos.

Ambas propuestas fueron implementadas en pilotos, que cubriremos a más detalle en la siguiente sección.

7 - Implementación de las recomendaciones

Para comprobar si al enfocarse en las brechas identificadas por la metodología es posible mejorar la experiencia, y por lo tanto el indicador de satisfacción, se implementaron pilotos distintos enfocados en cada una de las dimensiones para aislar el efecto. Los pilotos buscan comprobar que las prioridades modeladas son mejores con respecto a las prioridades originales del área de estrategia del negocio.

Piloto A - Enfocado en calificación del agente

Para aislar el efecto de la capacidad del agente en la satisfacción final percibida del cliente, se corrió un piloto por 6 meses. Consistía en poner dos grupos de agentes; el primer grupo estaba compuesto por los agentes en el primer cuartil, en cuanto a su calificación promedio del reactivo para calificar la atención del agente en la encuesta de salida. El segundo grupo era la cola izquierda de la distribución de calificación del mismo indicador. La media de la calificación del grupo de agentes bien calificados era 4.88, y mediana de 4.87, con la distribución normal. Para el grupo de agentes con menor calificación, la media era 3.89, mediana 3.82, y la distribución, con un ligero sesgo.

En la medida de lo posible, el centro se mantuvo balanceado con 40% de los agentes en un polo de la distribución; otro 40% de los agentes en el polo opuesto, y se dejó un 20% abierto para no afectar la capacidad del centro y poder tener espacio para rotación de personal.

Los clientes se asignaban aleatoriamente a los agentes según su disponibilidad en un modelo de colas FIFO (*First-In, First-Out*). En el periodo de 6 meses se recolectaron 10,435 encuestas de los dos grupos de agentes polares.

Posteriormente, aplicamos nuestras técnicas de limpieza de encuestas para mejorar la muestra. Se removieron las encuestas contestadas en menos de 15 segundos (5% de las encuestas) y también se eliminaron las encuestas donde todos los reactivos eran iguales, todos los reactivos con 5 o todos los

reactivos con 1 (7% de las encuestas); la razón por la cual estas muestras se eliminan es para remover respuesta de baja calidad, apresuradas y poco objetivas que buscan mayormente penalizar a la empresa más que dar retroalimentación del servicio.

La muestra restante fue de aproximadamente 9,800 encuestas de los grupos polares, 57% del grupo de agentes con calificación positiva y 43% del grupo de agentes con calificación negativa.

Finalmente, las dos muestras fueron balanceadas con submuestreo aleatorio con 5,000 encuestas para cada uno de los grupos.

Los resultados se muestran en la tabla 12.

Calificación de soporte CSat del centro, piloto A:

Tabla 12. Resultados piloto A

| Grupo observado | Media | Intervalo de confianza (95%) | Razón de cambio contra pre-tratamiento |
|------------------------|-------|------------------------------|--|
| Centro pre-tratamiento | 4.23 | 4.13 - 4.35 | 0.00% |
| Agentes cola superior | 4.67 | 4.55 - 4.71 | +10% |
| Agentes cola inferior | 3.85 | 3.77- 4.1 | -8% |

Se realizó también un test de permutación con el p-value = 0.000034, lo cual indica que el modelo probabilístico es estadísticamente significativo.

Piloto B - Enfoque en el tiempo total de la visita.

Para comprobar el efecto de la segunda variable clave, el tiempo total de la visita, se diseñó otro piloto en un centro de atención distinto al centro del *piloto A*. Aleatoriamente a 30% de los clientes que lo visitaban se les ofrecía llenar un cuestionario que funcionaba como un pre-diagnóstico del problema y si se llenaba exitosamente, el cliente era llamado al inicio de la fila. Con este modelo el tiempo total de la visita se reducía hasta un 83%, según la hora del día y el tiempo de espera del centro.

Únicamente se consideraron 30% de los clientes para el grupo de tratamiento porque la capacidad del centro no permitía acelerar el tiempo de espera del 50%.

El modelo de prediagnóstico pretendía reducir el tiempo de atención al darle al agente la información necesaria para diagnosticar el problema, cargar los datos del cliente y dar una solución de forma rápida.

Mover a los clientes al inicio de la fila permitía acortar el tiempo de espera. Para este centro no se estableció un sistema FIFO, los agentes directamente podían tomar a los clientes del grupo registrado en espera. Por lo que el cliente no era consciente de haber tenido un servicio prioritario, pero su percepción de espera bajaba significativamente.

El tiempo promedio total de visita del centro era 43 minutos, cuando los clientes eran atendidos bajo el modelo del tratamiento el tiempo promedio era de 17 minutos.

El piloto B estuvo activo por 6 meses, en los cuales se recolectaron 12,554 encuestas. Se aplicaron los tratamientos necesarios para eliminar encuestas con sesgo o contestadas de forma apresurada. Las encuestas contestadas en menos de 15 segundos fueron eliminadas, al igual que las encuestas con reactivos polares y constantes (todos los reactivos 5 o todos los reactivos 1).

La muestra restante fue de 10,879 encuestas, 69% del grupo control y 31% para el grupo de tratamiento. Finalmente, la muestra final se balancea con 3,000 encuestas del grupo control y 3,000 para el grupo de tratamiento, aleatoriamente.

Los resultados se observan en la tabla 13.

Calificación de soporte CSat del centro el piloto B:

Tabla 13. Resultados piloto B

| Grupo observado | Media | Intervalo de confianza (95%) | Razón de cambio contra pre-tratamiento |
|------------------------|-------|------------------------------|--|
| Centro pre-tratamiento | 4.12 | 3.93 - 4.32 | 0% |
| Grupo tratamiento | 4.45 | 4.39 - 4.54 | +8% |
| Grupo control | 4.01 | 3.91 - 4.28 | -2% |

Se realizó también un test de permutación con el $p\text{-value} = 0.000088$ lo cual indica que el modelo probabilístico es estadísticamente significativo.

Es importante notar cómo hubo un pequeño cambio entre el grupo control y la media del centro pre tratamiento, los intervalos de confianza se intercalan, pero la hipótesis que lo explicaba era el efecto que tenía llevar a los clientes al inicio de la fila en tiempo de espera de los clientes del grupo control.

Piloto C - Control - Soporte personalizado.

Como control, se diseñó un piloto que probaba la prioridad inicial que respondía al servicio personalizado para clientes catalogados como de alto valor, basado en su *lifetime value*, una proyección del valor que le generarán a la compañía en los siguientes 3 años. El proyecto de servicio ya estaba implementado a nivel ciudad. La iniciativa se lanzó a inicios del 2019. El servicio prioritario consta de poder recibir servicios adicionales de soporte. Estos pueden ser: agendar una cita por teléfono, políticas más flexibles tal como reducir el rigor para aprobar un proceso de reembolso.

Para analizar este grupo se observaron los clientes con historial de soporte en un único centro de atención diferente al centro del Piloto A y el Piloto B. Lo anterior para evitar contaminación cruzada entre los experimentos.

La muestra de encuestas fue de 3,050 encuestas en total, y al aplicar los procedimientos de limpieza de encuesta la muestra final fue de 2,705; la muestra final se comparó con una muestra aleatoria de 2,705 encuestas de clientes no prioritarios en el mismo centro de atención.

Los resultados se observan en la tabla 14.

Calificación de soporte CSat del centro el piloto C:

Tabla 14. Resultados piloto C

| Grupo observado | Media | Intervalo de confianza (95%) | Razón de cambio contra el centro |
|-------------------|-------|------------------------------|----------------------------------|
| Centro (muestreo) | 4.34 | 4.02 - 4.50 | 0.00% |
| Prioritario | 4.31 | 3.97 - 4.47 | -0.60% |
| No prioritario | 4.35 | 4.02 - 4.57 | -0.20% |

Los resultados del Piloto C no son concluyentes; los intervalos de confianza están completamente superpuestos, por lo cual no se puede confirmar que los resultados no sean un problema de muestreo.

Se realizó también un test de permutación con el $p\text{-value} = 0.6$, lo cual indica que el modelo probabilístico no es estadísticamente significativo.

Dado que el Piloto C era una prioridad principal para el área de negocio, tuvimos más recursos para analizar, incluso de forma cualitativa, los resultados. En mesas redondas con los clientes anecdóticamente notamos cómo las expectativas de los clientes con servicio prioritario subían. Los servicios adicionales dados no satisfacen estas nuevas expectativas que se generan al comunicarles que recibirán un trato especial. Los datos no pueden comprobar esta hipótesis, pero es plausible que, al ofrecerles un servicio especial, los clientes respondan de manera más exigente las encuestas, como lo afirma la teoría de las expectativas de Oliver (Oliver, 1977).

8 - DISCUSIÓN

La figura 15 muestra la implementación de los pilotos basados en la escala de prioridades sugerida por el modelo creado a través de nuestra metodología, así como los resultados obtenidos en cada uno de ellos.

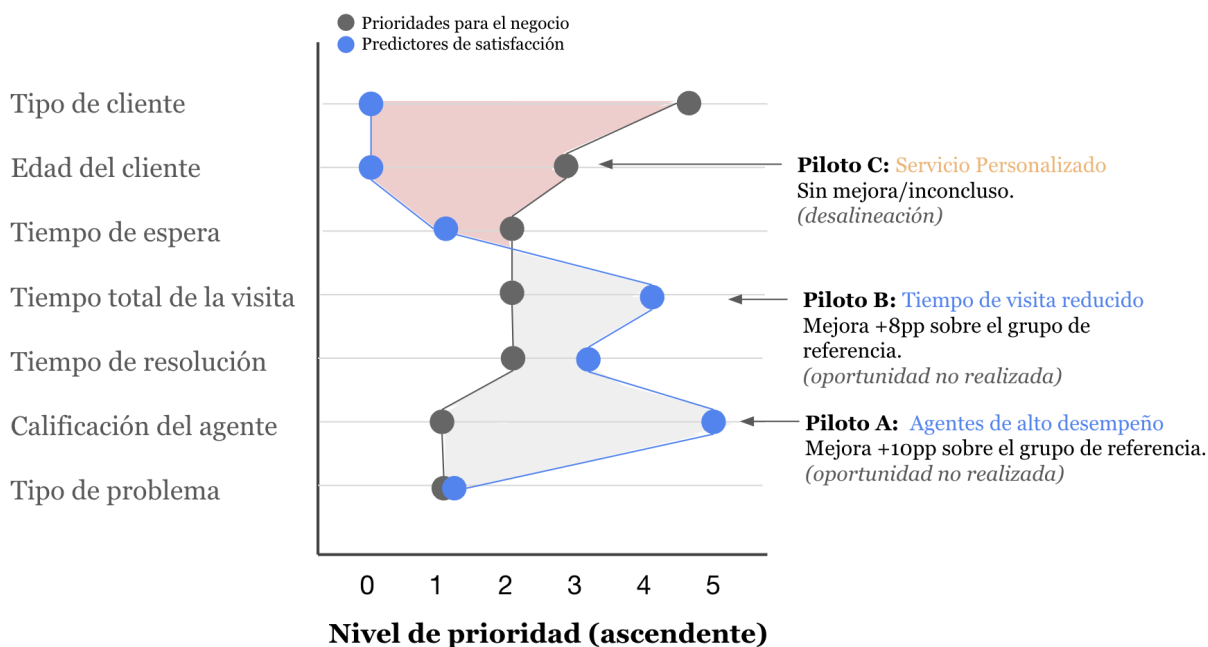


Fig 15. Alineación de los pilotos con las recomendaciones del modelo

1 - En este ejercicio práctico logramos sintetizar un sistema de servicio y modelar sus indicadores operacionales y salidas, o mejor dicho las encuestas de satisfacción. Los sistemas caóticos que incluyen entradas y salidas generadas por humanos pueden tener un nivel adicional de dificultad para cuantificar y observar.

2 - Observamos que puede existir una brecha entre los tópicos prioritarios para la empresa o las áreas de gestión, y lo que realmente aprecia o necesita el cliente. Dadas las estructuras organizacionales y operativas las áreas de gestión pueden percibirse lejanas al cliente o consumidor final, lo que genera desalineamientos entre los proyectos de mejora de una empresa y la percepción de calidad del cliente.

3 - La información es abundante y altamente capitalizable. El avance acelerado de los sistemas de gestión de información nos permite capturar grandes volúmenes de datos para modelar un sistema con amplio detalle. Sin embargo, el mismo exceso de información complica el uso de los mismos; es fácil perderse en la abundancia de la información porque es de suma importancia entender cada una de las variables del

sistema, su calidad y forma de recolección para poder entender los sesgos y vicios que pueden tener y darles el tratamiento específico antes de utilizarlas.

4 - Los sistemas de gestión de información suelen tener debilidades, por lo cual es necesario estudiar minuciosamente las variables de entrada y de salida de los sistemas y utilizarlas en conjunto, de forma agregada. Simplificar el sistema demasiado puede resultar en recomendaciones espurias.

5 - Se precisa hacer un ejercicio robusto de *feature engineering* antes de iniciar cualquier proceso de modelado. Este proceso no únicamente sirve para optimizar el costo computacional del sistema, también sirve para añadir una capa de sensibilidad de negocio más tangible al sistema. Empíricamente, una variable no ajustada a un indicador operativo, por buen predictor de una salida que sea, no suele poderse traducir a una recomendación accionable.

6 - En sistemas con salidas subjetivas (en nuestro caso, encuestas), la capacidad predictiva de una variable puede estar directamente relacionada con la percepción de satisfacción o el sesgo subjetivo de una encuesta de salida si el sistema está correctamente acotado a dimensiones tangibles para el cliente, cómo tiempos de espera, capacidad del agente de soporte, posibilidad de resolución del problema, etc.

7 - La selección del modelo es tan importante como la selección de variables. Es muy ambicioso buscar una solución única para todos los sistemas. Sin embargo, se puede hacer una selección precisa del modelo según las características de las variables de entrada y de salida mediante un proceso iterativo.

8 - Las necesidades del cliente no siempre se van a alinear con las prioridades del negocio. Desde el punto de vista del negocio se debe entender claramente el alcance del entendimiento del cliente y la diversidad de sistemas y clientes que pueden existir. Es más eficiente y accionable abordar la escala de prioridades de forma sistemática y cuantitativa, a pesar de que el paradigma de toma de decisiones de soporte sea tradicionalmente anecdótico y heurístico.

9 - CONCLUSIONES

En este trabajo se propuso una metodología para utilizar modelos de aprendizaje automático, como sistemas de toma de decisiones. Se estudia el caso de uso en un ecosistema de servicio al cliente, donde en la situación inicial existen sesgos en la toma de decisiones que llevan al desperdicio de recursos, esfuerzos y una implementación estratégica sin resultados.

La metodología delineada en este trabajo logra vislumbrar los aspectos operativos que tienen un impacto en la percepción del cliente sobre la calidad del servicio. De esta forma se logró contrastar las prioridades de la gerencia contra las salidas de un modelo predictivo que modela la experiencia de soporte.

A través de la combinación del análisis cuantitativo y acumen de negocio de expertos en el sistema se logró una evaluación robusta del ecosistema servicio. Es decir, se logró entender las relaciones que existen entre los datos que se capturan con los sistemas de gestión y la forma en la que se pueden traducir, a lo que el cliente percibe como calidad.

Con el análisis exploratorio del ecosistema de servicio se logra entender, de manera clara, el estado inicial de la experiencia y se identifican las levas que pueden ser utilizadas para la mejora del mismo, por ejemplo, tiempos de espera, expertise del agente de soporte, tiempos de resolución, etc.

También se abordan los distintos tratamientos que permiten utilizar los datos capturados por los sistemas de gestión para ser incluidos en el modelo del sistema de servicio. La metodología propuesta pone especial énfasis en la sensibilización de los indicadores para evitar los elementos caóticos naturales de los procesos humanos, especialmente en el manejo de información de encuestas donde es fácil tener muestras poco balanceadas o contaminadas.

Posteriormente exploramos distintas técnicas de aprendizaje automático para encontrar un modelo que pudiera predecir las respuestas de los clientes en una encuesta de satisfacción, según las dimensiones de servicio a los que fueron expuestos (tiempo de espera, tiempo de resolución, especialización del agente, etc). Se probaron seis modelos de categorización, todos con resultados similares.

El estudio del cambio en el desempeño de los modelos, al aislar cada una de las dimensiones de soporte, nos permitió entender el poder predictivo de cada una de las dimensiones de soporte y hacer una escala de

prioridades. Para así identificar las áreas de desenfoco entre las prioridades del equipo de gestión y lo que percibía el cliente.

Con la nueva escala de prioridades, se generaron recomendaciones para la mejora de la experiencia y se pusieron a prueba con pilotos, dando resultados satisfactorios con mejoras del 8% a 10% sobre el estado inicial (tabla 13 y tabla 12). Con ello se comprobó nuestra hipótesis de que las técnicas de aprendizaje automático pueden ser usadas para descomponer un ecosistema de servicio en sus factores primarios de los que depende la percepción de la experiencia para definir estrategia de mejora.

La misma metodología puede utilizarse para la mejora de sistemas de soporte similares, o incluso puntos de venta de *retail* y hospitalidad en múltiples industrias, puesto que las dimensiones estudiadas suelen ser un factor común en estos (tiempos de espera, resolución, toma de encuestas de satisfacción, etc). El ecosistema de soporte se estudia en este trabajo es un sistema tradicional de soporte en persona, fácilmente replicable para cualquier negocio que tenga localidades físicas con agentes de soporte o vendedores.

Este estudio se limita a factores operativos, en estudios de tiempos y eficiencia. Pero existen elementos de la experiencia aún más subjetivos que operan de forma inconsciente en la percepción del cliente que no fueron incluidos. Algunos ejemplos de estas dimensiones pueden ser: temperatura del centro de atención, dificultad para llegar al local, decoración y ambientación de las instalaciones, iluminación, disponibilidad de servicios (sanitarios, estacionamiento, etc). Estas dimensiones podrían ser incluidas en el modelo experiencial para robustecer las recomendaciones.

Este trabajo consolida una serie de técnicas y tendencias de gestión para abordar de la manera más objetiva y cualitativamente posible las mejoras en la experiencia de soporte.

En conclusión, pudimos encontrar una metodología que, a través un modelo basado en técnicas de inteligencia artificial, logra encontrar de manera sistemática las brechas entre las necesidades del cliente y las prioridades del negocio y sus variables operativas clave, para transformarlas en recomendaciones accionables y cuantificables.

10 - RECOMENDACIONES PARA TRABAJOS FUTUROS

Este proyecto abre varias oportunidades para profundizar y mejorar la metodología propuesta. Tanto en la optimización de selección de variables, como el tratamiento para encuestas y datos subjetivos, y la selección final del modelo.

En primera instancia, podría buscarse una manera más programática o automatizada para el análisis de las variables de entrada y su conexión con elementos sensibles para el cliente y el negocio. Es decir, encontrar de una forma escalable las variables iniciales que modelarán todo el sistema de soporte. Existen actualmente abundantes técnicas de *feature engineering* que podrían ser utilizadas para expeditar el proceso inicial de selección de variables primarias. Lograr esto lograría reducir las dependencias en el acumen del sistema que se quiere modelar.

El segundo punto a explorar es la selección del modelo de aprendizaje automático. En este trabajo, el modelo fue elegido de forma iterativa, de forma educada y acotada a los modelos que al mejor juicio podrían funcionar mejor, pero con base en el ensayo y error. Se propone también, como seguimiento a esta investigación, la existencia de una serie de principios que permita elegir el modelo de forma más precisa y directa, según el tipo de variables de entrada y las variables de salida. Con dichos principios, el proceso de generación de variables clave podría ser llevado a escala y producción más fácilmente.

Existe una buena oportunidad de mejorar las técnicas de limpieza de datos para encuestas. La rama de cuantificación de factores humanos es muy amplia y continuamente se perfeccionan las técnicas de recolección, pero esta evolución constante de las metodologías de captura, generan un retraso o poca estandarización en las técnicas de limpieza. Es posible aplicar las técnicas clásicas de limpieza de datos, pero las distribuciones no suelen ser normales, por lo que las técnicas tradicionales pueden tener consecuencias no previstas.

Finalmente, es necesario fomentar el desarrollo de sistemas para la toma de decisiones en áreas del negocio secundarias, cómo lo puede ser soporte o servicio al cliente. El conocimiento y desarrollo se enfoca mayoritariamente en áreas de retorno inmediato como marketing o ventas. Sin embargo, la mejora en dichas áreas puede toparse fácilmente y la falta de enfoque en las áreas secundarias genera disparidad en la evolución de la empresa. Estas disparidades, si no son atendidas, pueden rápidamente volverse

puntos débiles. Una organización con un enfoque completo, puede convertir sus áreas menos atendidas en fortalezas con el enfoque correcto.

REFERENCIAS

- Ahani, Ali, Rahim Nor Zairah Ab., and Mehrbakhsh Nilashi. 2017. "Forecasting Social CRM Adoption in SMEs: A Combined SEM-neural Network Method." *Computers in Human Behavior* 560-78.
- Akbaba, Atilla. 2006. "Measuring Service Quality in the Hotel Industry: A Study in a Business Hotel in Turkey." *International Journal of Hospitality Management* 170-92.
- Alismail, Sarah. 2018. "The Use of Emoji in Electronic User Experience Questionnaire: An Exploratory Case Study." *Hawaii International Conference on System Sciences*.
- n.d. *businessdictionary*. Accessed agosto 13, 2018.
<http://www.businessdictionary.com/definition/customer-support.html>.
- Allen, D. R. (2018). *Analysis of Customer Satisfaction Data*. New Age International Publisher.
- Bruce, A., & Gedeck, P. (2018). *Practical Statistics for Data Scientists*. O'Reilly.
- Cahill, D. J. (1997). What's Wrong with Service Providers? *Journal of Customer Service in Marketing & Management*, 3(2), 71–79.
- Cha, Young-Joo. 2007. "Identifying cross-cultural differences of emoticons." *Indiana University*.
- Chandok, M., & Gupta, N. (2014). Examining the attributes of customer experience and assessing the impact of customer experience on customer satisfaction: an empirical study of banking industry. *International Journal of Services Sciences*, 5(2), 154.
- Chitturi, Ravindra, Rajagopal Raghunathan, and Vijay Mahajan. 2008. "Delight by Design: The Role of Hedonic Versus Utilitarian Benefits." *Journal of Marketing* 48-63.
- Collier, Joel E., Donald C. Barnes, Alexandra K. Abney, and Mark J. Pelletier. 2018. "Idiosyncratic Service Experiences: When Customers Desire the Extraordinary in a Service Encounter." *Journal of Business Research* 150-61.
- Delen, Dursun, and Haluk Demirkan. 2013. "Data, Information and Analytics as Services." *Decision Support Systems* 359-63.
- Disney Institute, The, & Kinni, T. (2011). *Be Our Guest (Revised and Updated Edition): Perfecting the Art of Customer Service (A Disney Institute Book) (Revised, Updated ed.)*. Disney Editions.
- Elhadjamor, E. A., & Ghannouchi, S. A. (2020). AHP and SMART Criteria for KPI Selection in Higher Education. *International Journal of Knowledge-Based Organizations*, 12(2), 1–15.

- Farris, Paul W., Neil T. Bendle, Phillip E. 2010. *Marketing Metrics: The Definitive Guide to Measuring Marketing Performance*. Upper Saddle River, New Jersey: Pearson Education, Inc.
- Frei, Frances, and Anne Morriss. 2012. *Uncommon Service How to Win by Putting Customers at the Core of Your Business*. Illinois: HBR.
- Gandomi, Amir, and Murtaza Haider. 2015. "Beyond the Hype: Big Data Concepts, Methods, and Analytics." *International Journal of Information Management* 137-44.
- Gilmore, B. Joseph Pine II and James H. 1998. "Welcome to the Experience Economy." *Harvard Business Review*, July.
- Goul, M., & Corral, K. (2007). Enterprise model management and next generation decision support. *Decision Support Systems*, 43(3), 915–932.
- Hartung, T. (2018). Making Big Sense From Big Data. *Frontiers in Big Data*, 1.
- Harzing, Anne-Wil. 2016. "Response rate in international mail surveys: Results of a 22- country study." *International Business Review*,
- Hui Zou, Trevor Hastie. 2005. "Regularization and variable selection via the elastic net." *J. R. Statist. Soc* 301-320.
- n.d. Investopedia. Accesado en agosto 13, 2018.
<https://www.investopedia.com/terms/c/customer-service.asp>.
- Johnston, R., S. Brignall, and L. Fitzgerald. 2002. "Good Enough' Performance Measurement: A Trade-off between Activity and Action." *Journal of the Operational Research Society* 256-62.
- Kim, W. H. (2017). The Impact of Online Reviews on Customer Satisfaction: An Application of the American Customer Satisfaction Index (ACSI). *International Journal of Tourism Management and Sciences*, 32(5), 65–78.
- Keiningham, Timothy L., Lerzan Aksoy, Bruce Cooil, Tor Wallin Andreassen, and Luke Williams. 2008. "A Holistic Examination of Net Promoter." *Journal of Database Marketing & Customer Strategy Management* 79-90.
- Leavy, B. (2020). Frances Frei and Anne Morriss: The dynamics of empowering leader/follower relationships. *Strategy & Leadership*, 48(6), 27–33. <https://doi.org/10.1108/sl-09-2020-0125>
- Lejeune, Miguel. 2001. "Measuring the Impact of Data Mining on Churn Management." *Internet Research* 11 375-87.
- Likert, Rensis. 1932. "A Technique for the Measurement of Attitudes." *Archives of Psychology*. 1–55.
- Lim, Chiehyeon, Ki-Hun Kim, Min-Jun Kim, Jun-Yeon Heo, Kwang-Jae Kim, and Paul P. Maglio. 2018. "From Data to Value: A Nine-factor Framework for Data-based Value Creation in Information-intensive Services." *International Journal of Information Management* 121-35.

- Mattsson, U. T. (2010). A New Scalable Approach to Data Tokenization. SSRN Electronic Journal. Published.
- Markey Philip, Malcolm J. Wright, and Pamela Feetham. 2017. "Are Promoters Valuable Customers? An Application of the Net Promoter Scale to Predict Future Customer Spend." *Australasian Marketing Journal* 3-9.
- Merkey, Robert. 2011. "Introducing: The Net Promoter System" , Bain & Company.
<https://www.bain.com/insights/introducing-the-net-promoter-system-loyalty-insights>
- O'Connor, M. (2018). Measuring the unmeasurable. *Support for Learning*, 33(3), 255–266.
- Oliver, Richard L. 1977. "Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation." *Journal of Applied Psychology* 480-86.
- Price, B., & Jaffe, D. (2008). *The Best Service is No Service: How to Liberate Your Customers from Customer Service, Keep Them Happy, and Control Costs* (1st ed.). Jossey-Bass.
- Raassens, Néomie, and Hans Haans. 2017. "NPS and Online WOM." *Journal of Service Research* 322-34.
- Reichheld, Frederick. 2003. HBR.org. Accesado en Junio 2018.
<https://hbr.org/2003/12/the-one-number-you-need-to-grow>.
- Saad, A. (2013). Integration of Automated Decision Support Systems with Data Mining Abstract: A Client Perspective. *International Journal of Advanced Computer Science and Applications*, 4(2)
- Sandhu, D. V. (2020). ARTIFICIAL INTELLIGENCE AND ITS RELATION WITH MANAGEMENT & CUSTOMER SUPPORT SERVICES. *International Journal of Psychosocial Rehabilitation*.
- Sahar, F. 2018. "Machine-Learning Techniques for Customer Retention: A Comparative Study." *International Journal of Advanced Computer Science and Applications* 9,
- Sheikh, Y. (2017). Effective Feature Selection for Feature Possessing Group Structure. *International Journal Of Engineering And Computer Science*.
- Thompson, A. A. (2021). *Crafting & Executing Strategy: The Quest for Competitive Advantage: Concepts and Cases* (23rd ed.). McGraw-Hill Higher Education.
- Vieira, P. C. (2016). T-Test with Likert Scale Variables. SSRN Electronic Journal. Published.
- Wan, Yun, and Qigang Gao. 2015. "An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis." 2015 IEEE International Conference on Data Mining Workshop.
- Weech-Maldonado, Robert, Marc N. Elliott, Adetokunbo Oluwole, K. Cameron Schiller, and Ron D. Hays. 2008. "Survey Response Style and Differential Use of CAHPS Rating Scales by Hispanics." *Medical Care* 963-68. *Medical Care*.
- Wickham, Hadley. 2014. "Tidy Data." *Journal of Statistical Software*.

Yadegaridehkordi, Elaheh, Mehrbakhsh Nilashi, Mohd Hairul Nizam Bin Md Nasir, and Othman Ibrahim. 2018. "Predicting Determinants of Hotel Success and Development Using Structural Equation Modelling (SEM)-ANFIS Method." *ourism Management* 66 364-86.