

Modelos predictivos de Customer Experience – CX

eTb

Reporte



Julio 2024

emərgia

Objetivo	3
Descripción del Objetivo	3
Descripción de datos usados	3
Datos de interacciones previas	3
Ejemplo de Datos de Interacciones Previas	4
Datos del cliente	4
Ejemplo de Datos del Cliente	5
Datos contextuales	5
Ejemplo de Datos Contextuales	6
Cantidad de datos históricos	7
Análisis Predictivo	7
Definición del Problema y Recolección de Datos	7
Preprocesamiento de Datos	8
Análisis Exploratorio de Datos	8
Modelado	8
Despliegue	8
Algoritmos Utilizados	8
Resultados del Modelo	9
Métricas de Rendimiento	9
Documentación de la Aplicación del Modelo	11
Estrategias de Negocio y Personalización de Acciones	11
Acciones Proactivas	11
Personalización de Acciones	11
Comunicación Dirigida	11
Optimización de la Experiencia del Cliente	11
Seguimiento y Ajuste Continuo	12
Impacto en el Negocio	12
Comunicación de Resultados	13

Objetivo

El objetivo de este proyecto fue desarrollar un modelo predictivo avanzado que permitiera anticipar el motivo de las llamadas de los clientes de la empresa, con el fin de reducir el tráfico de llamadas entrantes, mejorar la experiencia del cliente y optimizar la atención mediante una gestión efectiva que incremente la resolución en el primer contacto (FCR) y potencie las oportunidades de cross-selling.

Descripción del Objetivo

El principal propósito de este modelo predictivo es predecir y anticipar la experiencia del cliente utilizando datos históricos de interacciones previas con la empresa. Este enfoque permite identificar patrones y tendencias en las razones de las llamadas, lo que posibilita implementar estrategias proactivas para mejorar la eficiencia operativa y la satisfacción del cliente. A continuación, se detallan los principales objetivos específicos del modelo:

- **Reducción del Tráfico de Llamadas Entrantes:** Al predecir los motivos más comunes de las llamadas, la empresa puede tomar medidas preventivas, como mejorar la información en su sitio web o a través de otros canales de comunicación, para resolver las inquietudes de los clientes antes de que sientan la necesidad de llamar.
- **Mejora de la Experiencia del Cliente:** Al anticipar los problemas y necesidades de los clientes, la empresa puede ofrecer soluciones más rápidas y efectivas, aumentando la satisfacción y la lealtad del cliente.
- **Optimización de la Atención al Cliente (FCR):** Incrementar la tasa de resolución en el primer contacto al proporcionar a los agentes de atención al cliente información relevante y soluciones predefinidas para los problemas más comunes.
- **Potenciación de Oportunidades de Cross-Selling:** Identificar momentos clave durante las interacciones con los clientes para ofrecer productos o servicios adicionales que se alineen con sus necesidades y comportamientos.

Descripción de datos usados

Para construir el modelo predictivo que anticipe el motivo por el cual un cliente se comunicará la próxima vez (Explicación factura, Soporte técnico, PQR factura, Traslado), es esencial recolectar y preparar diversos tipos de datos. A continuación, se describe brevemente cada tipo de dato usado:

Datos de interacciones previas

En la construcción del modelo fue usada toda la información registrada de las comunicaciones pasadas entre los clientes y la empresa. Estos datos fueron fundamentales para entender el comportamiento histórico de los clientes y predecir futuros motivos de contacto. A continuación, se detallan los principales componentes de los datos de interacciones previas:

- **Fecha y hora de la comunicación:** El momento exacto en que se produjo cada interacción (fecha y hora). Esta variable permitió analizar patrones temporales, como

picos de contacto en ciertos días y horas, además relacionar eventos específicos con el aumento de interacciones.

- **Motivo de la comunicación:** La razón por la cual el cliente se puso en contacto con la empresa, lo cual permitió entender los problemas o consultas más frecuentes e identificar tendencias y áreas de mejora.
- **Duración de la llamada o interacción:** El tiempo que duró la interacción entre el cliente y la empresa. Esta variable permitió evaluar la complejidad de los problemas y la eficiencia del servicio al cliente. Las interacciones más de mayor duración pueden relacionarse a problemas más complejos o insatisfacción.
- **Resultado o resolución de la interacción:** El resultado final de la interacción, como si el problema fue resuelto, si se escaló a otro departamento, si quedó pendiente, etc. Esto permitió evaluar la efectividad de las respuestas y resoluciones dadas, como también estimar el impacto en la satisfacción del cliente.
- **Canal de Comunicación:** El medio a través del cual se realizó la interacción (e.g., teléfono, correo electrónico, chat en línea, redes sociales). Esta variable permitió identificar los canales más utilizados y su efectividad para manejar diferentes tipos de consultas.

Ejemplo de Datos de Interacciones Previas

Fecha y Hora	Motivo	Duración (min)	Resultado	Canal
1/07/2023 10:15:00	Explicación factura	12	Resuelto	Teléfono
2/07/2023 14:30:00	Soporte técnico	25	Escalado a nivel 2	Chat
3/07/2023 9:45:00	PQR factura	15	Resuelto parcialmente	Correo
4/07/2023 11:20:00	Traslado	30	Pendiente	Teléfono

Datos del cliente

Se refieren a la información específica sobre cada cliente que puede influir en su comportamiento y en las razones por las cuales se comunican con la empresa. Estos datos ayudan a personalizar las interacciones y a construir modelos predictivos más precisos. A continuación, se describen los principales tipos de datos del cliente necesarios:

- **Datos demográficos:** Información básica sobre el cliente, como edad, género, ubicación, estado civil, nivel educativo, etc. Esta variable ayudó a segmentar a los clientes y entender patrones de comportamiento en diferentes grupos demográficos.
- **Historial de compras o servicios contratados:** Registro de todos los productos o servicios adquiridos por el cliente, incluyendo fechas de compra, cantidades, precios y frecuencia de compra. Esta variable proporcionó una visión de los intereses y necesidades del cliente, así como de su relación a largo plazo con la empresa.
- **Datos de facturación:** Información sobre las facturas emitidas al cliente, incluyendo montos, fechas de emisión, fechas de vencimiento y pagos realizados. Permitió identificar posibles problemas o consultas relacionadas con la facturación, como errores, retrasos en los pagos, o solicitudes de aclaración. Permitió identificar posibles problemas o consultas relacionadas con la facturación, como errores, retrasos en los pagos, o solicitudes de aclaración.
- **Datos de uso del producto o servicio:** Información sobre cómo el cliente utiliza los productos o servicios, como la frecuencia de uso, características más utilizadas, problemas reportados, etc. Esta variable ayudó a detectar problemas técnicos y a identificar oportunidades para mejorar la experiencia del cliente.
- **Encuestas de satisfacción y feedback:** Esta variable detalla opiniones y niveles de satisfacción del cliente obtenidos a través de encuestas, comentarios, reseñas y otros medios de retroalimentación, proporcionando insights sobre la percepción del cliente respecto a los productos o servicios, y ayuda a identificar áreas de mejora.

Ejemplo de Datos del Cliente

Cliente ID	Edad	Género	Ubicación	Historial de Compras	Monto Total Facturado	Último Pago	Frecuencia de Uso	Satisfacción (1-5)
1	35	F	Bogotá	Producto A, Producto B, Servicio C	\$150000	2023-06-15	Diario	4
2	28	M	Cali	Producto D, Servicio E	\$75000	2023-06-10	Semanal	3

Datos contextuales

Estos datos se refieren a la información externa que puede influir en el comportamiento y las decisiones de los clientes. Estos datos no están directamente relacionados con las interacciones previas o las características individuales de los clientes, pero proporcionan un

contexto adicional que puede afectar el motivo por el cual un cliente se comunica con la empresa. A continuación, se describen los principales tipos de datos contextuales necesarios:

- **Promociones actuales:** Información sobre las ofertas, descuentos, campañas de marketing y promociones especiales vigentes. Las promociones pueden provocar un aumento en las consultas y las interacciones con los clientes, así como influir en los motivos de comunicación.
- **Cambios en políticas o servicios:** Modificaciones recientes en las políticas de la empresa, términos y condiciones, tarifas, o servicios ofrecidos, las cuales pueden generar dudas, quejas o solicitudes de aclaración por parte de los clientes.
- **Eventos especiales:** Eventos importantes como lanzamientos de nuevos productos, actualizaciones de software, cambios en la administración, fusiones o adquisiciones, etc. Estos eventos pueden desencadenar un aumento en las interacciones y consultas por parte de los clientes, buscando más información o asistencia relacionada.
- **Factores Económicos y Sociales:** Información sobre el contexto económico y social que podría afectar a los clientes, como recesiones, pandemias, desastres naturales, etc. Estos factores pueden influir en el comportamiento de los clientes y en el tipo de consultas que realizan.
- **Tendencias del Mercado:** Información sobre las tendencias actuales en el mercado, la industria y la competencia. Las tendencias del mercado pueden afectar las expectativas y demandas de los clientes, así como el volumen y tipo de interacciones.

Ejemplo de Datos Contextuales

Fecha	Promoción Actual	Cambio en Política	Evento Especial	Factores Económicos/Sociales	Tendencia del Mercado
1/07/2023 0:00:00	Descuento del 20% en Producto X	Nueva política de devolución	Lanzamiento de Producto Y	Aumento del costo de vida	Creciente demanda de servicios online
15/07/2023 0:00:00	2x1 en Servicio Z	Cambio en tarifas de suscripción	Actualización de software V2	Recesión económica	Competencia ofreciendo precios más bajos

Cantidad de datos históricos

La cantidad de datos históricos necesarios para el entrenamiento de un modelo predictivo puede variar dependiendo del tipo de problema, la frecuencia de las interacciones y la estacionalidad de los datos. Sin embargo, para la construcción del modelo entrenado fueron usados datos históricos de un año, ya que esto permite capturar cualquier estacionalidad o patrones anuales en el comportamiento de los clientes. Es importante tener en cuenta que para la construcción de este tipo de modelos, puede ser óptimo usar un histórico de 2 a 3 años, ya que proporciona una cantidad suficiente de datos para entrenar el modelo y capturar tendencias y patrones a largo plazo. Además, los datos fueron sometidos a un análisis de calidad de los mismos con el fin de garantizar fueran precisos, completos, consistentes, relevantes y actuales para asegurar que las predicciones fueran fiables y útiles para la toma de decisiones.

Adicionalmente, se detallan algunas consideraciones adicionales para la construcción de este tipo de modelos:

- **Frecuencia de Interacciones:** Si las interacciones con los clientes son muy frecuentes (por ejemplo, diarias o semanales), un año de datos podría ser suficiente. Si son menos frecuentes (mensuales o trimestrales), puede ser necesario un período más largo.
- **Cambios recientes en Políticas o productos:** Si ha habido cambios significativos en políticas, productos o servicios recientemente, es importante incluir suficiente historial para que el modelo pueda capturar el impacto de estos cambios.
- **Estacionalidad:** Asegúrate de cubrir al menos un ciclo completo de cualquier estacionalidad relevante. Por ejemplo, si las interacciones aumentan durante ciertos meses del año, asegúrate de incluir estos períodos en el historial.

Análisis Predictivo

El análisis predictivo llevado a cabo con el modelo predictivo sigue una metodología estructurada para entender y prever las expectativas, preferencias y comportamientos futuros de los clientes. La metodología utilizada se detalla a continuación:

Definición del Problema y Recolección de Datos

- **Entendimiento de Objetivos de Negocio y KPIs:** Se trabaja en estrecha colaboración con los stakeholders de la empresa para definir claramente los objetivos del modelo y los indicadores clave de rendimiento (KPIs) que se quieren mejorar.
- **Datos Históricos:** Se recopilan datos históricos de chat, correo y teléfono de las interacciones de clientes, incluyendo datos sociodemográficos, detalles de productos y servicios utilizados, y registros de interacciones anteriores.

Preprocesamiento de Datos

- **Limpieza de Datos:** Se realizan procesos de limpieza para eliminar valores atípicos, manejar datos faltantes y corregir errores en los datos.
- **Creación de Características Relevantes:** Se generan nuevas características a partir de los datos existentes para mejorar la capacidad predictiva del modelo.

Análisis Exploratorio de Datos

- **Distribución de los Datos:** Se analiza la distribución de las variables clave para identificar posibles sesgos y características importantes.
- **Identificación de Relaciones y Patrones:** Se utilizan técnicas de análisis exploratorio para descubrir relaciones entre diferentes variables y patrones en el comportamiento de los clientes.
- **Detección de Tendencias Potenciales:** Se identifican patrones y tendencias que pueden revelar áreas de mejora en la experiencia del cliente, así como eventos recurrentes que afectan la satisfacción del cliente.

Modelado

- **Algoritmo de Aprendizaje Automático:** Se entrena el modelo utilizando el algoritmo de XGBoost, ajustando los parámetros para optimizar su rendimiento.
- **Validación Cruzada K-fold:** Se emplea validación cruzada para evaluar el rendimiento del modelo y garantizar su robustez.
- **Evaluación del Desempeño:** Se utilizan métricas como el accuracy score, precisión y recall para evaluar la precisión y efectividad del modelo.
- **Interpretabilidad del Modelo:** Se analiza la importancia de las características para entender qué factores son más influyentes en las predicciones.

Despliegue

- **Entorno de Producción:** El modelo se despliega en el entorno de producción de la empresa.
- **Monitoreo del Desempeño del Modelo:** Se monitorea continuamente el desempeño del modelo en relación con los KPIs definidos, asegurando que siga siendo efectivo con el tiempo.
- **Reentrenamiento Periódico:** El modelo es re-entrenado de manera periódica con nuevos datos para mantener su precisión y relevancia.

Algoritmos Utilizados

El modelo predictivo implementado para la empresa utiliza el algoritmo de "XGBoost" (Extreme Gradient Boosting), una técnica de machine learning basada en árboles de decisión que ha demostrado ser altamente eficaz para tareas de predicción debido a su robustez y capacidad para manejar grandes conjuntos de datos con múltiples características. A continuación, se describen los componentes clave y beneficios de este algoritmo:

- **Ensamblado de Árboles:** XGBoost combina múltiples árboles de decisión para mejorar la precisión de las predicciones. Cada nuevo árbol corrige los errores

cometidos por los árboles anteriores, lo que resulta en un modelo más preciso y robusto.

- **Regularización:** XGBoost incluye técnicas de regularización que ayudan a prevenir el sobreajuste, mejorando la generalización del modelo a nuevos datos.
- **Eficiencia Computacional:** Este algoritmo es conocido por su eficiencia en términos de tiempo de entrenamiento y uso de memoria, lo que permite manejar grandes volúmenes de datos de manera efectiva.
- **Importancia de Características:** XGBoost ofrece métricas que indican la importancia de cada característica en la predicción, lo cual es útil para interpretar los resultados y entender qué factores influyen más en las llamadas de los clientes.

Resultados del Modelo

El modelo predictivo de Customer Experience (CX) es un modelo multiclase que se entrena utilizando un enfoque binario. Para ello, se toma cada una de las posibles clases (explicación factura) y se entrena el modelo para distinguir entre una clase específica y el resto de las clases. Este proceso se repite para cada una de las clases, lo que permite al modelo aprender a identificar cada clase individualmente frente a todas las demás. Para realizar las predicciones con este tipo de modelo, se utiliza un solo modelo con cuatro posibles salidas. Durante la fase de predicción, el modelo evalúa la probabilidad de cada clase por separado y selecciona la clase con la mayor probabilidad como la predicción final. Este enfoque permite al modelo manejar múltiples categorías de manera efectiva, garantizando que las predicciones sean precisas y que se identifiquen correctamente las necesidades y problemas de los clientes.

Métricas de Rendimiento

Precisión, Recall: Son métricas clave para evaluar el rendimiento de modelos de clasificación en Machine Learning. Estas métricas ofrecen información específica sobre la calidad de las predicciones del modelo, especialmente en problemas donde las clases están desbalanceadas. Los datos presentados en la Tabla 1. hacen referencia a las métricas por clase de interés (Explicación factura, Soporte técnico, PQR factura y Traslado).

#	Clase	XG Boost 79.80 %		
		Precisión [%]	Recall [%]	F1 - Score [%]
1	Explicación Factura	80	86	83
2	Soporte Técnico	80	89	85
3	PQR Factura	40	89	56
4	Traslado	42	88	57

Tabla 1. Reporte de las métricas de clasificación de cada una de las clases.

Matriz de Confusión: La matriz de confusión proporciona una representación visual de las predicciones del modelo frente a los valores reales, permitiendo identificar no solo la cantidad de predicciones correctas sino también los errores específicos. A continuación se muestran las matrices de confusión relacionadas a cada problema de clasificación además de su respectiva Área Bajo la Curva (AUC) que también es usada como métrica para evaluar el rendimiento de un modelo de clasificación.

Tabla 2. Reporte de matrices de confusión por cada una de las clases.

Clase Explicación Factura	Clase Soporte Técnico
<p>Heatmap showing the confusion matrix for 'Clase Explicación Factura'. The y-axis is 'True label' (0, 1) and the x-axis is 'Predicted label' (0, 1). The diagonal elements (0,0) and (1,1) are light orange, indicating high true positives. The off-diagonal elements (0,1) and (1,0) are dark purple, indicating low false positives and false negatives. A color scale on the right ranges from 0 (dark purple) to 100 (light orange). The overall accuracy is ACC = 86.17 %.</p>	<p>Heatmap showing the confusion matrix for 'Clase Soporte Técnico'. The y-axis is 'True label' (0, 1) and the x-axis is 'Predicted label' (0, 1). The diagonal elements (0,0) and (1,1) are light orange, indicating high true positives. The off-diagonal elements (0,1) and (1,0) are dark purple, indicating low false positives and false negatives. A color scale on the right ranges from 0 (dark purple) to 100 (light orange). The overall accuracy is ACC = 87.51 %.</p>
Clase PQR Factura	Clase Traslado
<p>Heatmap showing the confusion matrix for 'Clase PQR Factura'. The y-axis is 'True label' (0, 1) and the x-axis is 'Predicted label' (0, 1). The diagonal elements (0,0) and (1,1) are light orange, indicating high true positives. The off-diagonal elements (0,1) and (1,0) are dark purple, indicating low false positives and false negatives. A color scale on the right ranges from 0 (dark purple) to 100 (light orange). The overall accuracy is ACC = 71.87 %.</p>	<p>Heatmap showing the confusion matrix for 'Clase Traslado'. The y-axis is 'True label' (0, 1) and the x-axis is 'Predicted label' (0, 1). The diagonal elements (0,0) and (1,1) are light orange, indicating high true positives. The off-diagonal elements (0,1) and (1,0) are dark purple, indicating low false positives and false negatives. A color scale on the right ranges from 0 (dark purple) to 100 (light orange). The overall accuracy is ACC = 73.65 %.</p>

Documentación de la Aplicación del Modelo

Estrategias de Negocio y Personalización de Acciones

El uso del modelo predictivo permitió implementar acciones de negocio específicas y estrategias de retención basadas en los insights generados. A continuación, se detallan las estrategias y acciones concretas:

Acciones Proactivas

- **Proporcionar Información Anticipada:** Basado en las predicciones del modelo, la empresa puede enviar información relevante a los clientes antes de que ellos tengan que llamar. Por ejemplo, si el modelo predice un aumento en las consultas sobre facturación, la empresa puede enviar correos electrónicos con explicaciones detalladas y soluciones a problemas comunes relacionados con las facturas.
- **Automatización de Respuestas:** Implementar respuestas automáticas en el IVR (Interactive Voice Response) que aborden las preocupaciones más probables, disminuyendo la necesidad de intervención humana.

Personalización de Acciones

- **Ofertas y Promociones Personalizadas:** Utilizando los datos del modelo, la empresa puede crear y enviar ofertas personalizadas que se alineen con los intereses y comportamientos pasados de los clientes. Por ejemplo, ofrecer descuentos en servicios adicionales a clientes que frecuentemente llaman para preguntar sobre esos servicios.
- **Segmentación de Clientes para Atención Prioritaria:** Identificar a los clientes de alto valor y priorizar su atención para mejorar su satisfacción. Esto puede incluir asignar a estos clientes a agentes más experimentados o ofrecerles servicios exclusivos.

Comunicación Dirigida

- **Campañas de Comunicación Proactiva:** Desarrollar campañas de comunicación dirigidas a segmentos específicos de clientes en función de las predicciones. Por ejemplo, enviar guías detalladas y videos explicativos sobre la factura a clientes que frecuentemente solicitan explicaciones de facturación.
- **Mensajes Contextualizados:** Crear mensajes que aborden directamente los problemas o intereses predichos para cada cliente, mejorando la relevancia y efectividad de la comunicación.

Optimización de la Experiencia del Cliente

- **Mejorar Contenidos de Autoayuda:** Ampliar y mejorar la base de conocimientos y las secciones de preguntas frecuentes en el sitio web de la empresa, enfocándose en las consultas más comunes predichas por el modelo.

- **Capacitación Específica para Agentes:** Capacitar a los agentes de atención al cliente en las áreas identificadas como más problemáticas o frecuentes, asegurando que estén mejor preparados para resolver consultas de manera rápida y efectiva.

Seguimiento y Ajuste Continuo

- **Monitoreo de Impacto en Tiempo Real:** Utilizar dashboards y cuadros de mando en tiempo real para monitorear el impacto de las acciones implementadas, permitiendo ajustes rápidos basados en los resultados observados.
- **Encuestas de Satisfacción Post-Interacción:** Implementar encuestas de satisfacción después de cada interacción para obtener feedback directo de los clientes sobre la efectividad de las soluciones ofrecidas y ajustar las estrategias en consecuencia.
- **Análisis de Feedback y Ajustes Iterativos:** Recoger y analizar feedback de los clientes para identificar áreas de mejora y realizar ajustes iterativos en las estrategias de retención.

Estas acciones específicas permiten a la empresa no solo predecir las necesidades de sus clientes sino también abordar proactivamente sus preocupaciones, mejorando significativamente la experiencia del cliente y aumentando las tasas de retención.

Impacto en el Negocio

- **Indicadores Clave de Rendimiento (KPIs):** Muestra cómo el modelo ha mejorado los KPIs relevantes (e.g., reducción en el tiempo de respuesta, aumento en la satisfacción del cliente).

poner KPIs impactados

-

Impacto en el Negocio:

- Reducción del tiempo de respuesta en un 15%.
- Aumento de la satisfacción del cliente en un 10% según encuestas post-interacción.

Visualización:

- Enlace al dashboard interactivo: [Enlace]
- Gráficos y tablas: (Incluir visualizaciones relevantes aquí)

Conclusión:

- Resumen de los resultados obtenidos.

- Próximos pasos y recomendaciones para futuras mejoras del modelo.

- **Casos de Uso:** Proporciona ejemplos concretos de cómo el modelo ha ayudado a resolver problemas específicos o a mejorar procesos.

Gráficos con tres meses pre- pos intervención Kpis escogidos

Comunicación de Resultados

a. Informes y Presentaciones

- **Informes Detallados:** Elabora informes escritos que incluyan toda la información anterior, con explicaciones claras y detalladas.
- **Presentaciones:** Prepara presentaciones para las partes interesadas, destacando los aspectos más importantes y los impactos clave del modelo.

b. Revisión y Retroalimentación

- **Revisión Periódica:** Programa reuniones periódicas para revisar el rendimiento del modelo y discutir posibles mejoras.
- **Retroalimentación de Usuarios:** Recopila y analiza la retroalimentación de los usuarios que interactúan con el modelo para identificar áreas de mejora.

POner un governance (ver el de data quality)