

Creación de líneas de producto a contratar de acuerdo a perfiles de cliente

Presenta: Haile Jacobo Meneses Moreno

Introducción

Contamos con datos provenientes de una campaña de marketing directo de una institución bancaria de Portugal. Esta campaña está basada en llamadas telefónicas que tienen como objetivo que los clientes contraten un depósito a plazo fijo.

Estás llamadas con diversa duración y frecuencia obtenían como resultado en estos datos que el depósito es contratado o no (si/no), los resultados basados en frecuencia y duración obtienen diversos resultados según el momento de la temporada.

Objetivos del Proyecto

¿Cuáles son los objetivos del negocio?

- Aumentar la cantidad de clientes que contraten un depósito estableciendo un tipo de instrumento a ofertar basado en el perfil del cliente, buscando evitar un sesgo que limite la cantidad de clientes que puedan contratar servicios con la entidad.

2. ¿Qué decisiones o procesos específicos se quieren mejorar o automatizar con ML?

Con los datos que se tienen actualmente se pueden identificar diversos perfiles a los que pueden ofrecerse instrumentos que estén orientados a ellos, buscando con ello no excluir a la mayor cantidad posible de clientes que puedan tener una oportunidad de contratar algún producto con la entidad.

3. ¿Se podría resolver el problema de manera no automatizada?

Si, aunque implica realizar una tarea compleja, con una cantidad de variables que puedan hacer difícil la identificación de los perfiles y que además puede ocupar un periodo de trabajo muy extenso.

Metodología Propuesta

Segmentar a los clientes para encontrar grupos homogéneos en los datos me hace decantarme por un algoritmo como K-means Clustering; el objetivo es agrupar a los clientes según su perfil para identificar a qué grupo pertenece a cada cliente, permitiendo ofrecerles productos personalizados y con la idea de obtener mayores ratios de contratación a la entidad que puede tomar la decisión de crear estos productos.

La idea básica es que el algoritmo intentará agrupar a los clientes en K grupos (clusters), donde cada cliente dentro de un grupo es similar entre sí y diferente de los otros grupos.

Después de aplicar K-Means, cada cliente pertenece a uno de los clusters, y podemos interpretar estos grupos para saber qué tipo de perfil de cliente representa cada cluster.

Una vez que tenemos los clientes agrupados en perfiles, la entidad financiera podría ofrecer productos específicos a cada grupo.

Para evaluar el rendimiento de este algoritmo se tiene que utilizar una métrica distinta a los modelos supervisados, ya que en el clustering no se tienen etiquetas de verdad.

En este caso utilizaría Silhouette Score o Índice de Silueta, que es una de las métricas más utilizadas para evaluar la calidad de un clustering. Mide qué tan cerca están los puntos dentro del mismo cluster y qué tan alejados están de los otros clusters. Toma un valor entre -1 y 1, donde:

- Valor cercano a 1: Los puntos están bien agrupados y claramente separados de los otros clusters.
- Valor cercano a 0: Los puntos están en el borde de un cluster o mal agrupados.
- Valor negativo: Los puntos probablemente están asignados al cluster incorrecto.

Datos Disponibles

Dentro del conjunto de datos relacionados con el perfil financiero disponibles para trabajar con este algoritmo se tienen:

- Edad
- Balance de cuenta bancaria
- Crédito impagado
- Hipoteca (Sí/No)
- Préstamos personales (Sí/No)

Métrica de Éxito del Proyecto

Independientemente de tener en consideración la métrica técnica del algoritmo, en este caso el Índice de Silueta, el verdadero éxito de la implementación de este algoritmo se mide por el impacto que los clusters tienen en el negocio.

En este caso, se busca aumentar la cantidad de clientes que contraten un depósito y ofrecer productos basados en perfiles adecuados. Una de las métricas de negocio que podría utilizar sería la Tasa de Conversión por Cluster:

Una vez que se segmente a los clientes en diferentes clusters, se puede medir la tasa de conversión dentro de cada grupo:

Tasa de conversión = Número de clientes que contratan un producto / Total de clientes en el cluster.

Un éxito sería observar una tasa de conversión más alta después de implementar las recomendaciones basadas en los clusters, en comparación con las campañas anteriores.

Por ejemplo, si antes se tenía una tasa de conversión del 10% y después del clustering se aumenta al 15-20% en ciertos clusters, eso sería un claro indicativo de éxito.

Responsabilidades Éticas y Sociales

La principal meta de implementar un proyecto como el que se presenta, radica en evitar el sesgo y ofrecer productos a más clientes que pueden ser excluidos si se implementa un modelo que excluya por diversas razones que tradicionalmente los puede hacer inviables a contratar un producto financiero (edad, estado civil, antecedentes bancarios, etc.). Pero adicionalmente se tendrían en cuenta las siguientes consideraciones:

- Evitar el sesgo y la discriminación mediante la revisión de los datos y el monitoreo de los clusters generados.
- Ser transparente sobre cómo se toman las decisiones y por qué se asignan productos a determinados grupos.
- Respetar la privacidad y los derechos de los clientes, manejando los datos de manera responsable y con su consentimiento.
- Fomentar la inclusión financiera en lugar de la exclusión, ofreciendo productos a un espectro más amplio de clientes.
- Garantizar la supervisión humana y la posibilidad de que los clientes apelen decisiones automatizadas.