

Entrega 2 - Proyecto

Entregan: Julian Bolaños, Giovanni Mosquera, Juan Rodríguez y Juan Zorrilla

Modelamiento

Para avanzar con el desarrollo de un portafolio de soluciones de analítica para LÍMPIK, se pasará a la siguiente etapa, el modelamiento de nuestra propuesta para solucionar problemas/preguntas de interés del negocio.

Técnicas de Modelamiento

A continuación se listan nuevamente los objetivos de analítica:

1. Elaborar un modelo de regresión para predecir cuántas compras realizará un cliente por medio del catálogo, página web y tienda física.
2. Construir un modelo de clasificación de clientes acorde a los artículos que adquiere para generar perfiles de compradores
3. Elaborar un modelo de clasificación para identificar los usuarios que podrían aceptar una campaña de marketing/oferta a futuro, en base al número de campañas aceptadas y diferentes características del comprador.
4. Elaborar un modelo de regresión para predecir el dinero gastado por un cliente del comercio, en base al número de niños y jóvenes que viven en su hogar.

Para dar cumplimiento a cada uno de los objetivos, se plantea la elaboración de diferentes modelos de predicción que nos faciliten llegar a una respuesta correcta a los objetivos de negocio. Por cada objetivo de analítica se presentan el o los posibles modelos a emplear:

1. Regresión Lineal o Polinómica
2. K-Means
3. Regresión Logística (Gio)
4. Regresión Lineal o Polinómica

Protocolos de Evaluación para Modelos

Para los modelos de regresión se plantea el uso de las siguiente métricas: MSE, RMSE, Coeficiente R^2 y la gráfica de residuos.

Para el modelo de K-Means, se usarán estimadores como el método de codo, silueta y el índice Calinski-Harabasz, para encontrar la mejor definición de clusters.

En el último modelo, para la clasificación se emplearán la exactitud, precisión, TPR, FPR, puntaje f1 y la matriz de confusión.

Construcción de Modelos

Revisar los modelos creados en la carpeta de código, junto con su respectivo ajuste de hiper parámetros.

Evaluación

Resultados y Métricas

Revisar en el apartado de “Resultados” y “Métricas” de cada modelo en la carpeta de código.

Análisis de Resultados y Métricas

A continuación se listan los resultados obtenidos y las métricas relacionados a estos:

- Se obtuvo un perfil de comprador el cual es casi seguro que no aceptaría una campaña
- Se obtuvo un perfil de comprador que posiblemente acepte campañas
- Se obtuvo que la cantidad de jóvenes y niños por hogar, no es un factor determinante en el consumo
- Tenemos un error cuadrático demasiado alto para nuestra predicción de consumo de acuerdo a la cantidad de jóvenes y niños por hogar.
- Al intentar realizar predicciones de ventas con respecto a la tienda física no serán tan exactos como al realizar predicciones por catálogo
- Por el bajo ajuste de los modelos al intentar predecir la cantidad de compras que un cliente realizará por cualquier medio, no se realizará una predicción efectiva.
- Definimos 3 perfiles de compradores diferentes, cada uno con sus respectivas características

Acciones a tomar

En base a los resultados y métricas vistas dentro de los scripts de cada modelo, podemos empezar identificar que es necesario realizar ajustes con los modelos de regresión lineal presentados. Para dar cumplimiento al primer objetivo, se plantea alternar nuestro modelo de regresión lineal por un árbol de regresión, al igual que para el 4to objetivo se buscará realizar un análisis por medio de K-Means, ya que los resultados de la regresión no fueron óptimos.

Despliegue

Plan de Despliegue

Se plantea para la etapa de despliegue hacer uso de una página web simple, en donde se puedan consumir como recursos los 4 modelos generados. Para el uso de los modelos será necesario que quien interactúe con el programa, ingrese todos los datos de un comprador.