

Predicción SLA Natura



INTEGRANTES DEL GRUPO: LEANDRO M. ABRAHAM

EMPRESA: URBANO EXPRESS S.A.

FECHA DE PRESENTACIÓN: 12/02/2022

• DESCRIPCIÓN DEL CASO DE NEGOCIO

Trabajo para una empresa del rubro logístico (URBANO EXPRESS S.A.) y la idea es poder desarrollar un modelo de IA que pueda predecir el cumplimiento o no del SLA de las piezas del cliente Natura. Dicho SLA es una medida de tiempo que se asigna a cada pieza teniendo en cuenta variables como código postal, zona peligrosa, etc. Si la pieza es entregada dentro de los plazos establecidos, cumplirá con el SLA.

• OBJETIVOS DEL MODELO

Se plantea la necesidad de crear un algoritmo de IA que pueda predecir cuales piezas del shipper Natura no cumplirán el SLA, a fin de poder tomar acciones para revertir esa situación y poner el foco en el cumplimiento de dichas piezas.

• DESCRIPCIÓN DE LOS DATOS

Para este proyecto se utilizó una BD con alrededor de 80.000 registros. Cada registro corresponde al envío de un paquete del cliente Natura. Las variables que posee son:

cli_codigo: Tipo Categórico - Funciona como PK ya que es el código de la pieza.

sucursal: Tipo Categórico - Es la sucursal de Urbano encargada de realizar la entrega del paquete.

shipper: Tipo Categórico - Es el cliente, ya que Natura tiene varias divisiones (Natura E - Commerce, reventa, etc.)

zona_peligrosa: Tipo Categórico - Informa si la zona de envío es peligrosa o no.

servicio: Tipo Categórico - Informa el tipo de servicio contratado para ese paquete.

región: Tipo Categórica - Informa la región correspondiente al envío del paquete.

sla: Tipo Binaria - Informa si la pieza cumplió o no con el SLA. Es mi variable predictiva.

Cabe aclarar que la implementación se hace tomando los datos de una tabla desde el motor de BD PostgreSQL. Una vez aplicado el modelo, la idea es pasar esos datos a otra tabla en PostgreSQL donde se informen las predicciones de cada envío que este pendiente de entrega y con esa tabla armar un tablero con el visualizador de Power BI, que es el utilizado por nuestro equipo de BI.

• **HALLAZGOS ENCONTRADOS POR EL EDA**

A través del EDA se pudo determinar que el Data Set con el que contamos para este proyecto esta desbalanceado. Existen 68774 piezas que Cumplen el SLA (86%) y 10479 que no cumplen el SLA (14%). Además, las sucursales que más envíos reciben de este shipper son LLO (Logística Liviana Oeste) y LLS (Logística Liviana Sur)

• **ALGORITMO ELEGIDO**

Se probaron los siguientes algoritmos (todos incluidos en el código del proyecto final):

- ***Árbol de Decisión***

- ***Regresión Logística***

- ***KNN***

- ***Random Forest***

- ***XGBoost***

- ***XGBoost + Oversampling*** (Debido a la problematica del desbalance en la variable predictiva)

Si bien se han realizado pruebas de todos estos algoritmos, el algoritmo seleccionado para este proyecto fue XGBoost junto a un proceso de Oversampling, debido a que fue el que predijo con mayor precisión las piezas que no cumplirían el SLA, que es en definitiva lo que más interesa predecir en este proyecto de IA.

- **MÉTRICAS DE DESEMPEÑO DEL MODELO**

Las métricas del modelo seleccionado son:

Precision: 0.5539951291073492

Recall: 0.5515327759602568

Matriz de confusión:

[13581 7051]

[11455 9178]

CROC: 0.5601860679576701

- **FUTURAS LINEAS**

Los pasos a seguir una vez implementado el modelo es ir verificando el correcto funcionamiento del mismo. En un principio, las metricas que arroja el modelo no son las mejores. Se han probado varios tipos de modelos y se han mejorado las métricas a traves de diversas técnicas pero la problematica, probablemente sea la carencia de variables que ayuden a mejorar la performance del modelo. Se deben crear nuevas variables como puede ser, por ejemplo, una variable binaria que informe si el envío esta a cargo de un chofer que en el lapso de 7 días, tiene un indice alto de incumplimiento de SLA. Entre otras opciones posibles.

• CONCLUSIONES

La implementación de un modelo de IA que ayude a mejorar la performance de los envíos se presenta como una ventaja competitiva más que interesante para el negocio logístico. La posibilidad de poder hacer un seguimiento de los envíos que tienen alta probabilidad de no cumplir en tiempo y forma, sin duda aumentaría considerablemente la productividad y la calidad del servicio logístico. Es por ello que el uso de herramientas predictivas nos permiten mejorar considerablemente nuestra labor y son claves en la transformación digital de todas las industrias.

CODER