

Ejercicio 2

Rubén Cruz García

En un conjunto de 100 tiendas físicas de retail tenemos instalada nuestra tecnología de digital signage y audience analytics. La instalación, en cada una de las tiendas, consta de 10 pantallas ubicadas en distintas zonas de la tienda (3 escaparates, entrada, cola caja, entrada probadores, zona mujeres, zona hombres, zona niños, zona complementos) de las cuales se controla la emisión de la totalidad del contenido mediante nuestro sistema, midiéndose por completo. Cada una de estas pantallas está acompañada de una cámara frontal que permite medir las visualizaciones de contenido y características de las personas que realizan dichas visualizaciones. Además, la instalación contiene cámaras cenitales de otra tecnología en el techo de la tienda que permiten medir, usuario a usuario y con un ID para cada uno, las visitas y su movimiento por el espacio correspondiente a toda la superficie de cada una de las tiendas.

Los metadatos de las BBDD que disponemos de las distintas fuentes de información son estos:

- Contenido emitido por las pantallas y cámaras frontales ubicadas en las pantallas:

Variable	Descripción
<i>Store ID</i>	Identificador único de tienda
<i>Display ID</i>	Identificador único de pantalla
<i>Content ID</i>	Identificador único de contenido emitido
<i>Content tags</i>	Metainformación conceptual del contenido
<i>Start ts</i>	Time stamp del inicio de impresión del contenido
<i>End ts</i>	Time stamp del final de impresión del contenido
<i>Age</i>	Rango de edad estimado por la cámara frontal ubicada encima de la pantalla(<15, 15-25, 26-45, 46-60, >60)
<i>Gender</i>	Género estimado por la cámara frontal ubicada encima de la pantalla (masculino o femenino)
<i>Expression</i>	Expresión detectada estimado por la cámara frontal ubicada encima de la pantalla (sonrisa, neutra, seria o disgusto)
<i>Views</i>	Número de miradas detectadas a esa emisión concreta de contenido en esa pantalla concreta realizadas por personas con esas características de edad, género y expresión

- Cámaras cenitales:

Variable	Descripción
<i>Store ID</i>	Identificador único de tienda
<i>Zone ID</i>	Identificador único de zona de la tienda (podemos definir cuantas queramos y de la forma que queramos en el momento de la instalación del sistema)
<i>User ID</i>	Identificador único del usuario que visita la tienda. Este identificador se mantiene a través de las distintas zonas, desde que el usuario entra por la puerta hasta que sale de la tienda
<i>Entrance ts</i>	Time stamp del momento en el que el usuario entra a una zona concreta
<i>Exit ts</i>	Time stamp del momento en el que el usuario sale a una zona concreta
<i>Gender</i>	Género estimado estimado por la tecnología de las cámaras cenitales (masculino o femenino)

Supón que en el equipo de Data Science vamos a hacer un ejercicio de brainstorming de posibles soluciones avanzadas (estadística, Machine Learning, Inteligencia Artificial, etc.) que podríamos añadir a nuestro producto dados estos datos y otros que podamos obtener de fuentes o APIS públicas.

¿Qué propuestas harías tú? Explícalas a nivel conceptual y nombra la técnica o técnicas estadísticas que utilizarías en cada caso

Propuesta:

Nuestro objetivo es adaptar y predecir el contenido de las pantallas, para que éstas sean más efectivas y repercutan positivamente en las tiendas, generando más compras. Por tanto, de forma similar al *Ejercicio 1*, tenemos que comenzar con un análisis exploratorio de datos para entender su estructura, y así poder encontrar patrones, anomalías, tendencias y relaciones entre ellos. Por ejemplo: quizás gente de más edad tienda a ir a comprar los días de diario por la mañana, mientras que los jóvenes lo harán más por la tarde y los fines de semana.

Con las cámaras cenitales y el User ID habría que estimar qué pantallas ha podido visualizar cada usuario (si ha visualizado más de una y cómo se ha movido por la tienda), así como mejorar la localización de las pantallas si se obtiene que alguna de ellas está en un lugar poco idóneo (pocas visualizaciones o tiempo medio de visualización demasiado bajo). Por eso la Zone ID de las cámaras cenitales deberían determinarse a la hora de la instalación según donde se localicen las pantallas (10 zonas) para poder establecer un enlace entre ambas fuentes de información. Con la identificación de grupos de personas se podría generar otra variable predictora. Además si se detecta que un cliente se mueve hacia otra pantalla y hemos obtenido un tiempo de visualización elevado, habría que sincronizar la segunda pantalla a modo de continuación.

El ejercicio planteado bien podría tratarse de un problema de clasificación, en el que dados unos *features* queremos predecir el tipo de contenido que se reproduce en la pantalla, dentro de una selección disponible (en este caso el *Content ID* sería el *label* o *target*).

Antes de generar un modelo, podemos crear alguna variable adicional calculando clusters con K-Means, y quizás podamos agrupar varias variables. Por ejemplo: chicas de 15 a 25 años que sonríen u hombres de 46 a 60 años con gesto neutro, o clusters basados en la hora del día y el tiempo de visualización.

Para generar un modelo de predicción del contenido de acuerdo a la audiencia empezaría con probando varios modelos de ensemble como un Random Forest Classifier o un Gradient Boosting (donde cada árbol de decisión intenta corregir los fallos del anterior), ya que son algoritmos que no necesitan escalado ni procesamiento previo de los datos, y que además nos dan una idea de la importancia relativa de cada una de las variables predictoras, que podemos utilizar posteriormente para generar modelos más eficientes.

Lo primero es dividir los datos en Train y Test, empezando con una proporción de 70-30. Para una mejor selección de los hiper-parámetros empezaría con un Randomized Search, que nos daría una estimación de hacia dónde deberíamos establecer o *tunear* los parámetros. Posteriormente haría un Grid Search con Cross-Validation (k-fold de 5 particiones) para distintas opciones de parámetros más ajustadas gracias al Randomized Search.

Para evaluar los modelos utilizaría distintas métricas: recall, precisión (este es especialmente importante, ya que nos informa de los falsos positivos, que queremos evitar para no reproducir anuncios de ropa infantil a gente adulta, por ejemplo), F1-score,

especificidad, confusion matrix, ROC curve y AUC. Estas métricas pueden estar muy ajustadas entre los distintos modelos, por lo que sería necesario aplicar un test de Student o de Fisher, para verificar diferencias significativas.

Otro modelo sería entrenar una red neuronal dando como input las variables más importantes del Random Forest, y cambiando el peso (*weight*) de cada una de ellas hasta obtener la mejor combinación. En este caso habría que normalizar los datos.

Por último, podríamos plantearnos diseñar un modelo de Reinforcement Learning que recompensase al modelo si se lleva a cabo una compra. Este, sin embargo, necesitaría inicialmente datos simulados debido a las fases iniciales de prueba y error del modelo, que dificultan llevarlo a la práctica directamente.

En cuanto a añadir fuentes de información externa a las ya medidas por las cámaras, el contenido de las pantallas podría estar conectado a un servicio meteorológico o a APIs de geolocalización. Esta nueva información sería añadida como variables predictoras en los modelos explicados anteriormente. Por ejemplo, si es un día lluvioso o frío, reproducir vídeos que anuncien ropa cálida. Asimismo, la localización de la tienda puede ser un dato clave: si se encuentra cerca de la playa, los anuncios deberán ser diferentes a si la tienda está en el barrio de Gracia. Ciertamente, el barrio de cada ciudad o la ciudad misma puede determinar información útil a la hora de elegir el contenido: se podría hacer un estudio del precio de la vivienda de ese barrio, de los medios de transporte para llegar a esa localización o de la presencia de comercios del mismo sector. Además, se podría coordinar la reproducción en las pantallas con eventos programados, como la salida de una película en el cine o la llegada de un avión procedente de un lugar determinado (para anuncios en el aeropuerto, por ejemplo). Por último, sería indispensable tener información del stock para no promocionar productos no disponibles. De esta forma, se publicitarían productos sin vender desde hace tiempo en momentos específicos.