

**¿Cuál es el algoritmo de Machine Learning más adecuado para resolver este problema?
¿Cómo justifica la elección de ese algoritmo? ¿Qué métricas de evaluación se utilizarán
para medir el rendimiento del modelo?**

El algoritmo que considero utilizar es el de Regresión Lineal Múltiple (desde ahora RLM); al buscarse predecir un valor numérico y basándonos en diversos factores (tiempo de interacción de la web como de la app, estado de miembro, la suma anual gastada) un modelo de regresión basado en este algoritmo parece adecuado.

Aunque una ampliación de los datos podría hacer factible un algoritmo como la regresión de árboles de decisión, el uso de la RLM nos permite analizar como las variables independientes (tiempo en web/app, estado miembro, etc.) influyen de forma conjunta en una variable dependiente (el gasto anual).

Si se asume que existe una relación lineal entre estas variables predictoras y el gasto anual la RLM puede modelar esta relación de manera efectiva. Los coeficientes de la regresión nos indican la importancia relativa de cada variable predictora en la predicción del gasto anual, permitiéndonos comprender qué factores tienen mayor influencia en el comportamiento de los clientes tanto de la web como de la app.

Por ejemplo, mediante este algoritmo se podría saber si después de una mejora en la navegación de la página web y en la app se puede monitorizar la profundidad de las páginas vistas, el tiempo en el sitio, la tasa de rebote y su comportamiento de compra; esto también puede formar parte del proceso de poder predecir su gasto anual como del esfuerzo de la mejora de las herramientas web de la tienda.

Métricas para evaluar el rendimiento del modelo

Dentro de las métricas que podrían usarse en este caso, se pueden incluir:

Métricas centradas en el Usuario:

- **Tasa de Rebote:** Porcentaje de usuarios que abandonan el sitio o la app después de ver una página. Una tasa de rebote baja sugiera un compromiso (Engagement) con la marca.
- **Tiempo en el sitio/app:** Mide el tiempo promedio que los usuarios pasan en la plataforma. Tiempo más prolongado sugiere mayor interacción.
- **Profundidad de las páginas vistas:** Indica el número promedio de páginas que un usuario visita por sesión. Mayor profundidad sugiere una navegación más exploratoria.
- **Tasa de Conversión:** Mide el porcentaje de usuarios que realizan la acción deseada. Un aumento en la tasa indica una mejora en la experiencia de compra.

Métricas centradas en el Negocio:

- **Tasa de retención de clientes:** Mide el porcentaje de clientes que continúan usando la herramienta evaluada a lo largo del tiempo.
- **Ingresos por usuario:** Mide el valor monetario promedio que genera cada usuario.

Métricas Específicas para la Experiencia Digital

- **Tiempo de carga de la página:** Un tiempo de carga rápido mejora la experiencia del usuario.

- **Tasa de clics en llamadas a la acción (CTA):** Mide la efectividad de los elementos diseñados para impulsar acciones específicas.
- **Facilidad de uso:** Se puede medir a través de encuestas o tests de usabilidad, y se evalúa qué tan fácil es para los usuarios navegar y encontrar lo que buscan.

¿Qué responsabilidades éticas y sociales es importante tener en cuenta?

Al desarrollar y utilizar modelos de Machine Learning para predecir el gasto de los clientes y evaluar la eficiencia en el uso de las herramientas como la página web de la tienda o su app, es fundamental adoptar una perspectiva ética y social.

Esto implica respetar la privacidad de los datos, garantizar la transparencia, evitar la discriminación y utilizar los datos de manera responsable. Al hacerlo, se puede construir una relación de confianza con los clientes y contribuir a un desarrollo tecnológico más justo y equitativo.

Las consideraciones que pueden tenerse en cuenta para cumplir con ello pueden incluir, entre otros:

Privacidad de Datos

- **Consentimiento informado:** Se debe de obtener el consentimiento explícito de los clientes y ser informados sobre cómo se utilizarán sus datos y qué medidas se toman para protegerlos.
- **Minimización de datos:** Solo se deben recopilar y utilizar los datos estrictamente necesarios para alcanzar los objetivos del proyecto.
- **Seguridad de los datos:** Implementar medidas de seguridad robustas para proteger los datos de accesos no autorizados, pérdidas o alteraciones.
- **Anonimización y seudonimización:** En la medida de lo posible, los datos personales deben ser anonimizados o seudonimizados para proteger la identidad de los clientes.

Transparencia y Explicabilidad:

- **Transparencia en el uso de datos:** Los clientes deben ser informados de cómo se utilizan sus datos para personalizar la experiencia de compra.
- **Explicabilidad de los modelos:** Los modelos de ML pueden ser complejos, es importante poder explicar de manera sencilla cómo se han llegado a las predicciones.

No discriminación:

- **Evitar sesgos:** Los modelos de ML pueden perpetuar sesgos presentes en los datos de entrenamiento. Es fundamental analizar los datos para identificar y mitigar estos sesgos.
- **Trato equitativo:** Los modelos deben tratar a todos los clientes de manera justa y equitativa, independientemente de su género, raza, edad u otras características.

Uso responsable de los datos:

- **Evitar la manipulación:** Los datos no deben utilizarse para manipular o engañar a los clientes.
- **Cumplimiento de la normativa:** Asegurarse de cumplir con la normativa vigente en materia de protección de datos (en el caso de España el RGPD).