

Projecte de ML per a incrementar la despesa anual de clients a la botiga

Introducció

Es tracta d'una botiga exclusiva especialitzada en la confecció i venda de roba a mida. Es destaca per oferir consultories altament personalitzades. Els clients visiten la botiga per rebre assessorament directe d'estilistes experts que ajuden a crear peces úniques que s'ajustin perfectament a les seves preferències i mesures. Després de la sessió de consultoria, els clients poden fer comandes de roba mitjançant una aplicació mòbil o del lloc web de l'empresa.

Objectius del Projecte

1. Quins són els objectius del negoci? Augmentar les vendes avaluant on concentrar els esforços: millorant l'experiència dels clients al lloc web o l'aplicació mòbil.
2. Quines decisions o processos específics voleu millorar o automatitzar amb ML? Es cerca optimitzar les decisions empresarials relacionades amb l'experiència del client i automatitzar la predicció de la despesa anual de cada client.
3. Es podria resoldre el problema de manera no automatitzada? Si bé la predicció de la despesa anual de cada client es pot fer utilitzant fulls de càlcul o programari estadístics, l'ús de machine learning permetrà automatitzar aquest procés mitjançant entrenaments programats que incorporin ràpidament les transaccions diàries sense necessitat d'intervenció manual.

Metodologia Proposta

4. Quin és l'algorisme de Machine Learning més adequat per resoldre aquest problema? Com justifica l'elecció d'aquest algorisme? Que mètriques d'avaluació s'utilitzaran per a mesurar el rendiment del model?

Un bon punt de partida seria utilitzar algorismes supervisats, ja que coneixem la variable objectiu que volem explicar, que seria la despesa anual del client i, a més a més, partim de dades etiquetades, ja que la una de les dades disponibles és la informació del client, que és informació rellevant sobre la variable o grup d'estudi.

D'aquesta manera, considero que un bon model seria la regressió lineal múltiple, ja que tenim la variable dependent, que seria la despesa per client, que es podria explicar analitzant les altres dades disponibles, que serien les variables independents: informació del client (entenent que

demogràfiques, socials, d'interessos, etc.), la despesa prèvia anual, el temps de connexió i l'estat de membre.

D'altra banda, hem de considerar que per avaluar un model de regressió s'hauran d'utilitzar alguna de les mètriques següents: MAE, MSE, RMSE i R2. L'elecció de la mètrica dependria del model real, però en aquest cas, i pels avantatges que ens proporciona la mètrica, podríem utilitzar el MAE o error absolut mitjà; així, aquesta mètrica ens permetria avaluar la magnitud de l'error i els valors atípics (clients que compren molt o no ho fan gairebé mai) no influirien tant per a prendre decisions.

Dades Disponibles

5. Quines dades estan disponibles per abordar aquest problema? La botiga compta amb un conjunt de dades actualitzat que inclou informació identificativa de cada client, la suma anual que ha gastat a la botiga, el temps dedicat a interactuar tant al lloc web com a l'aplicació mòbil, i l'estat de membres.

Mètrica d'Èxit

6. Quina és la mètrica d'èxit per a aquest projecte? Augment en la Despesa Anual Mitjana per Client. Aquesta mètrica reflectiria directament l'efectivitat del model en millorar les decisions de l'empresa.

Responsabilitats Ètiques i Socials

7. Quines responsabilitats ètiques i socials és important tenir en compte?

En primer lloc destacaria la responsabilitat de respectar la privacitat de l'usuari i no incorporar al model dades que són irrellevants per aconseguir l'objectiu d'incrementar la despesa a l'aplicació o web. A més a més, s'hauria d'evitar demanar que l'usuari incorpori dades especialment sensibles com l'origen ètnic o racial, dades sobre la salut, dades relatives a l'orientació o identitat sexual, etc.

Per tal de garantir un bon ús del machine learning, també s'hauria de proporcionar informació a l'usuari i demás persones interessades que els permetés comprendre el model i com aquest ha extret els resultats; és a dir, garantir la traçabilitat de les conclusions a través d'un exercici de transparència.

Per últim, és especialment important determinar qui és la persona física o jurídica responsable del tractament de les dades que es recullen a l'aplicació o web i garantir que les interessades puguin fer valdre els seus drets davant d'aquest responsable.

