

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 a l'aplicació mòbil

2. Quines decisions o processos específics volem 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?

- Completar pels estudiants

L'objectiu del projecte és predir la despesa anual de cada client que és una variable continua. A més, en el conjunt de dades disposem de diverses característiques predictives: suma anual de despesa, temps d'interacció, ..., que ens ajudaran a predir la nostra variable dependent (despesa anual). Per tant, tenim un cas **d'aprenentatge supervisat de tipus regressió**.

Tenim diferents algorismes que podem utilitzar per aquest cas: regressió lineal, Random Forest regressor, XGBoost, etc.

En un primer pas farem servir la **regressió lineal** que és l'algoritme més senzill i que pot ser el més adient si la relació entre les variables predictives i la variable dependent és lineal. Tanmateix, si la relació no fos lineal o més complexa podem provar els altres algorismes que són més sofisticats. Per últim, per avaluar el model podem fer servir **MSE (Mean Squared Error)**, que mesura l'error mitjà al quadrat entre les prediccions del model i els valors reals, ens interessa que sigui el més baix possible.

Una altra mètrica que podem utilitzar és **R^2** que ens diu quina part de la variació total dels valors observats pot ser explicada pel model. Aquest serà millor quant més s'apropi a 1.

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?*

- Completar pels estudiants

En el nostre projecte hauríem de tenir en compte els següents aspectes:

- **Transparència**, haurem de proporcionar informació clara i accessible als clients de com s'utilitzen les seves dades i desenvolupar models d'IA que permetin la seva comprensió tant per part dels clients com dels treballadors.
- **Justícia** per evitar discriminació i biaixos en el conjunt de dades. Si el model està esbiaixat pot ser que tractem a determinats segments dels nostres clients de manera diferent i injusta i, per exemple, que no els tinguem en compte en certes promocions o que la seva satisfacció no sigui positiva.

- **Privacitat**, com treballem amb les dades dels clients ens hem d'assegurar que les dades estiguin protegides per garantir la seva privacitat, per exemple anonimitzant les dades. Per aconseguir-ho hem d'obtenir el consentiment explícit dels clients per poder fer-les servir.
- **Responsabilitat**, l'empresa ha d'indicar quins són els responsables en cas d'errors o un ús inadequat de les dades i que puguin generar queixes dels clients. Per exemple, el responsable podria ser el desenvolupador del model o el propietari de la botiga.
- **Impacte social**, l'automatització de la predicció de la despesa anual per part dels clients podria fer que alguns treballadors, que feien abans aquesta tasca amb Excel, no es necessitin en un futur. És un dels reptes ètics la forma de fer servir l'IA sense desplaçar els treballadors..