



ceos
data

Retenční model

Retenční model CEOS Data pro věrnostní programy

Každý věrnostní program je nutné postupem času vylepšovat aby bylo možné nejen oslovovat stávající zákazníky k dalším nákupům, ale zároveň reaktivovat ty zákazníky, kteří už nenakupují nebo nakupují méně. Všechny tyto vzorce nákupního chování lze predikovat na základě stávajících dat o zákazníkovi a frekvenci nákupů. Vstupem informací jsou agregované ekonomické a socio-demografické údaje vztažené ke klientovi, výstupem pak pravděpodobnost odchodu nebo změny chování jednotlivce. Produkt retenčního modelu umí vylepšit stávající aplikace tak, aby se daly maximalizovat očekávané přínosy pro váš business. Na predikci nákupního modelu pak lze připravovat cílené marketingové kampaně.

Při identifikaci zákazníků e-shopu a jeho chování ve věrnostním programu vyplývají vzorce chování, jako jsou: rušení svého členství, vymazání ze systému, nepokračování v nákupech a setrvávání v členství, změněné nákupní chování a další. V retenčním modelu je tedy nutné identifikovat zákazníka a připravit segmentaci.

www.ceosdata.com

V retenčním modelu využíváme kombinaci dvou metod, které přinášejí rozdílné benefity.

První metodou je použití Machine Learning (dále ML) vycházející z segmentace klientů, tj jejich zařazení do reprezentativních skupin. Model je dále obohacen o různorodé atributy, které ovlivňují jeho přesnost. Machine Learning model je běžně používaným analytickým postupem pro identifikaci odcházejících zákazníků.

Druhá metoda je založena na využití tzv. intervalu spolehlivosti (dále IS) a na rozdíl od metody první zkoumá klienta jako suverénní identitu. Tento model je vhodný pro zákazníky s vyšší frekvencí nákupů a umožňuje včasnější detekci změny chování. Identifikuje změnu chování či odchod klienta na základě jeho vlastní nákupní historie. Vstupem jsou data nákupů klienta a jejich obraty. V toku času se mu od posledního nákupu postupně zvyšuje pravděpodobnost dalšího nákupu (plní se mu pomyslný progress bar k dalšímu nákupu), a pokud v mezním čase nenakoupí (překračuje povolenou toleranci), konstatujeme, že u něj došlo ke změně chování.

Díky kombinaci obou metod dokážeme

- » včasně reagovat i na rychlé změny chování (nečekáme pevně dané období 3 měsíců)
- » zároveň využít celou šířku ML modelování pro predikci odchodu v situacích, kde odchod není na první pohled jednoduše detekovatelný běžným pohledem na data

Nečastěji řešené změny chování připadají na:

- » identifikace odcházejícího zákazníka
- » identifikace zákazníka, který podezřele mění své chování oproti minulosti

Obě změny jsou řešitelné podobnými postupy, ale zásadně se liší v definici odezvy. Definice odezvy je přitom klíčová jak z hlediska business využití, tak i způsobu vyhodnocení přesnosti.

Predikci změny chování klientů řešíme dvěma odlišnými metodami, které se dobře doplňují

- » IS model vychází z nákupní historie každého klienta; určuje pravděpodobnost, že klient k určitému datu nakoupí
- » ML model vychází z agregovaných dat; určuje pravděpodobnost, že klient v jeho specifickém časovém okně nenakoupí

Implementace retenčního modelu a doporučení k jeho praktickému využití

Proces implementace modelů se provádí v následujících fázích:

- » příprava dat
- » tvorba odezvy a příprava vhodných proměnných
- » vytvoření modelů a validace
- » automatizace tvorby dat
- » aplikace modelů a export pravděpodobností

Hlavní informace poskytované modely:

- » klienti, kteří již „spí“, nebo změnili chování
- » klienti se blíží k časovému bodu, kdy by měli nejpozději nakoupit
- » klienti, kteří ještě „nespí“, ale brzy „usnou“

Návrhy dalšího rozvoje základního modelu (předpokládané rozšiřující prediktory v budoucnu) v závislosti na dalších sebraných datech o nákupech a nákupním chování. Z modelu jsou zároveň patrné faktory a jejich statistická významnost, tj. jak ovlivňují chování zákazníka (v rámci segmentu).

Model může být snadno obohacen o novou proměnnou, na které půjde ověřit, zda má vliv na chování zákazníka. K ověření dojde buď prostřednictvím p-hodnoty, nebo srovnáním přesnosti referenčního modelu bez dané proměnné. V budoucnu je možné rozšířit výstup o vliv proměnných na konkrétního člověka.

Předpokládaná a garantovaná úspěšnost modelu

Náš odhad úspěšnosti vychází z předběžné analýzy na vzorku dat, na kterém jsme se pokusili validovat oba výše zmíněné přístupy (resp. modely). Validací modelu rozumíme vyhodnocení, zda vytvořený matematický model prokazuje shody s reálnými daty.

Souhrn úspěšnosti:

- » dosažený záchyt až 95 % ohrožených klientů
- » úspěšnost záchytu je řízena volitelným parametrem
- » vysvětlení, proč není vhodné cílit pouze na nejvyšší míru záchytu - vzhledem k ceně retence
- » ukázka validace ML i IS modelu

Validace modelu ML

Validace tohoto modelu je přímočará. Známe predikované i skutečné hodnoty a lze tedy jednoznačně změřit přesnost. Samotná přesnost je dle našeho názoru však jen orientačním kritériem pro posouzení kvality modelu. Vzhledem k tomu, že nepredikujeme jasný odchod, ale změny chování, snažíme se primárně vyhledávat zákazníky, u nichž dochází k snížení frekvence nákupů – tzv. "usnutí". Model se tedy snaží vyhledávat tyto spáče.

V rámci experimentální analýzy, kterou jsme provedli na vzorku dat, vykazoval ML model senzitivitu až 95%, čili zachytil 95% všech spáčů

Trénovací vzorek obsahoval 10tis. náhodně vybraných klientů nakupujících v období 05/2017 až 12/2019. Validace byla provedena na datech prodloužených do 10/2020.

Validace modelu IS

Validace tohoto modelu je komplikovaná. Vzhledem k tomu, že zde nemůžeme použít jasně ověřitelnou odezvu, můžeme dosáhnout různé přesnosti podle toho, jak odezvu definujeme.

Předpokládaná a garantovaná úspěšnost modelu nabízí:

- » vysvětlení úspěšnosti (95%) dosahované naším modelem
- » informace o důležitosti prahových hodnot pro stanovení úrovně modelu