

Data mining a retence zákazníků

ve finančních institucích

Lenka Posch

Známa poučka říká, že důležitější než získávat nové klienty je udržet si ty současné. Nejen proto, že spokojený zákazník je pro každou firmu tou nejlepší reklamou. Také z hlediska nákladů je částka spojená se získáváním nových zákazníků zpravidla podstatně vyšší než náklady na udržení klientů stávajících. To obzvláště platí u finančních institucí, kde je velká konkurence a nabídka produktů a služeb je napříč společnostmi většinou srovnatelná.



Zákazníci v sektoru finančních společností mají tendenci setrvat u zvolené instituce poměrně dlouho. Proto může retence zákazníků, kteří mají zřejmý úmysl, podstatně ovlivnit zisk společnosti. Pro některé společnosti totiž může pět procent zákazníků odrazených od odchodu znamenat nárůst zisku o dvacet pět až sto procent. Klient není pouze zdrojem jednorázového příjmu, ale přináší finance dlouhodobě.

Jak na to?

Úplně na začátku je potřeba spočítat, jestli se vyplatí zabývat se myšlenkami na aktivitu pro udržení stávajících zákazníků. Jako podklad pro toto rozhodování slouží velikost výnosů, o které společnost přichází kvůli odcházejícím klientům. Konkrétně pak jde o tu část výnosů, kterou představují klienti, u nichž je rozhodnutí odejít odvrátitelné.

Prvním krokem je vždy identifikace skupiny zákazníků, u nichž je nejvyšší pravděpodobnost, že přestanou využívat naše služby či produkty. Vžitý je pro tento typ analýz pojem churn analysis. Podaří-li se zjistit důvod a pravděpodobný čas odchodu zákazníka a najdeme-li levné a efektivní řešení, které ho přiměje zůstat u naší společnosti, je vyhráno. Tato základní myšlenka je jednoduchá, ale její realizace je možná jen v případě, že máme k dispozici přesné modely, s nimiž můžeme pracovat. Dále je třeba si uvědomit, že kromě dobrých klientů existují i zákazníci, kteří přinášejí více problémů než zisku, a že není v zájmu společnosti snažit se udržet si tento typ klientů. Churn analysis je proto zaměřena pouze na bonitní klienty uvažující o odchodu.

Moderní software pro tento účel poskytuje řadu sofistikovaných nástrojů, bez kterých se při obrovských objemech dat, jež jsou dnes denní rutinou, neobejdeme. Nezdírkou je výhodné zahrnout všechny dostupné informace, tedy nejen data uložená v přehledných tabulkách, ale i nestrukturované texty, jež v minulosti neměly valné uplatnění. Churn modely jsou tak zpřesněny a vylepšeny s použitím textminingových algoritmů. V ideálním případě model předpoví tendenci zákazníka odejít dříve, než si ji zákazník sám uvědomí. Na základě historických dat klientů hledáme proto rizikové faktory typické pro odchod klienta, které jsou přítomny již dva nebo tři měsíce před jeho odchodem. Po nasazení takového modelu identifikujeme klienty s úmyslem odejít a zároveň máme ještě dostatek času k cílení nabídky a k přesvědčení zákazníka, aby u naší společnosti setrval. Používané modely využívají dataminingové techniky, jako jsou rozhodovací stromy, asociační pravidla nebo

pokročilé regresní modely. V praxi se ukazuje, že klíčová není volba algoritmu, nýbrž pečlivý výběr použitých proměnných.

Jakmile je definována ohrožená skupina klientů, je třeba ji blíže prozkoumat a porozumět, proč je v ohrožení. Zjistíme například, že všichni nespokojení zákazníci používají jeden konkrétní produkt. V takovém případě je potřeba nejen přesvědčit zákazníky, ale také daný produkt vylepšit. V dalším kroku je proto třeba provést segmentaci potenciálně nespokojených klientů, podle důvodu jejich nespokojenosti a preferencí – na každého klienta bude fungovat jiná taktika. Na každou z těchto cílových podskupin je následně aplikována vhodná retenční kampaň či aktivita. Někdy z modelů a segmentace získáme výborné podklady pro oporu rozhodování o budoucí politice a strategiích společnosti, v praxi bohužel ne vždy dochází ke změně, ačkoli data říkají, že je namístě.

Asociační pravidla

Jeden z dataminingových přístupů, který se využívá pro účely churn analýz, jsou asociační pravidla. Tato metoda je založena na velmi jednoduchém principu. U konkrétních zákazníků sledujeme současný výskyt dvou či více událostí, a vyvozujeme závěry o tom, jaký vliv má výskyt těchto konkrétních událostí na současný výskyt událostí jiných. V případě churn analýzy zkoumáme na základě historických údajů, co jsou asi pravděpodobné příčiny klientova rozhodnutí odejít od naší společnosti v horizontu tří měsíců před jeho odchodem. Celá analýza je založena na relativních četnostech výskytu určitých kombinací sledovaných událostí v historických datech. První ze sledovaných veličin pro nalezená WQ pravidla je označována jako support a udává procentuální část všech klientů, u nichž se vyskytuje jak množina událostí A (příčiny), tak množina událostí B (následky). Čím větší je support, tím lépe. Support je jakási základna všech našich zákazníků, u nichž je dané asociační pravidlo aktivní. Spolu s hodnotou support se uvádí ještě tzv. spolehlivost, veličina odpovídající svou povahou podmíněné pravděpodobnosti, která vztahuje počet zákazníků, u nichž nastaly všechny události v množinách A a B zároveň, k počtu zákazníků, u nichž nastaly pouze uvažované příčinné události A (tj. počet klientů, kde A má za následek B, vůči všem klientům, u nichž došlo k událostem z množiny A bez ohledu na možné důsledky). Opět, čím větší je spolehlivost uvažovaného asociačního pravidla, tím lépe. Současně ale chceme i rozumnou hodnotu support pro toto pravidlo.

Prostřednictvím dataminingového nástroje vyfiltrujeme pouze takové následky (nebo příčiny), které pro nás jsou zajímavé, a asociační pravidla, která je obsahují, podrobíme detailnější analýze. Jako směrodatná bereme v úvahu pravidla, která mají jak vysokou hodnotu support, tak spolehlivost (pojem vysoké hodnoty je relativní, neboť konkrétní hranice vždy závisí na úloze, kterou řešíme). Výstupem jsou při použití asociačních pravidel tedy jednak seznamy nalezených pravidel, která mají support a spolehlivost odpovídající kritériím stanoveným analytikem (příčiny jsou označovány jako „body“, následky jako „head“, viz tabulka), ale také grafy pravidel, které poskytují přehled o hodnotě support (velikost uzlových bodů) a spolehlivosti (jejich barevná škála) pro možné dvojice příčin a následků.

Je zde patrné, že pokud mají rodinní příslušníci zákazníka účet u firmy D, zákazník velmi často přejde k této společnosti. Lze předpokládat, že firma D propaguje nabídku, která zvýhodňuje klienty v rámci celých rodin. Další nejproblémovější oblastí je podle grafu i skupina zákazníků, kteří využívají produkt P2.

Kde všude lze asociační pravidla využít?

Jak jsme ukázali v předchozím odstavci, asociační pravidla přináší cenné informace o nejčastějších možných příčinách odchodu klientů ke konkurenci. Dále můžeme tuto metodu využít i pro nastavení vhodných aktivit ve skupinách ohrožených klientů pro jejich odrazení od přechodu ke konkurenci. Poté, co proběhne segmentace klientů podle věku, preferencí a pravděpodobných příčin odchodu, lze opět na základě historických dat nebo testovacího vzorku klientů stanovit, které aktivity jsou pro dané klienty nejúspěšnější a nejčastěji vedou k setrvání zákazníka u naší společnosti po zvolené časové období (například jeden rok).

Výhodou asociačních pravidel je jejich snadná interpretace a s ní související bezproblémová obhajoba závěrů před managementem společnosti. Kromě retence zákazníku je tato metoda také hojně využívána při analýze nákupního koše. ■



Autorka je odbornou konzultantkou a analytičkou firmy Statsoft.