



ANALIZA DANYCH I CRM - PRZEGLĄD

Grzegorz Migut

StatSoft Polska Sp. z o.o.

Wstęp

Działania organizacji w konkurencyjnym otoczeniu są uwarunkowane pojawiającymi się wokół różnorodnymi ograniczeniami i wymuszeniami. Umiejętność radzenia sobie z oddziałującymi na organizację czynnikami zewnętrznymi decyduje o jej przetrwaniu bądź upadku. Za jeden z najistotniejszych czynników, które obecnie kształtują obraz rynku, należy uznać ogromne zróżnicowanie produktów oraz „szum informacyjny” wywołany niespotykanym wcześniej natężeniem aktywności reklamowej. Dodatkowo można także zaobserwować wzrost przeciętnego rozporządzalnego dochodu. Czynniki te powodują znaczące obniżenie się poziomu lojalności klientów oraz występowanie efektów częstego przesuwania preferencji. Kolejnym czynnikiem decydującym o obliczu dzisiejszego rynku jest rewolucja technologiczna. Jej owocami są nowe, elektroniczne kanały dystrybucji i komunikacji, które umożliwiły bezpośredni kontakt z oddalonym klientem, a co za tym idzie spowodowały globalizację konkurencji [1].

Czynniki te, połączone często z niską jakością obsługi klienta, spowodowały, że firmy, pragnąc przetrwać w obecnych warunkach, zmuszone zostały do zmian swoich strategii marketingowych. Odpowiedzią na nowe warunki stał się marketing relacyjny, który umieszcza potrzeby klienta w centrum zainteresowania organizacji. Nowe podejście zakłada konieczność budowania indywidualnych, trwałych kontaktów pomiędzy dostawcą a klientem. Komunikacja i dystrybucja powinna mieć charakter spersonalizowany, indywidualny. Kontakty z klientem traktuje się nie jako zbiór odrębnych epizodów, lecz jako proces, na który można oddziaływać.

Budowanie trwałych relacji z klientem powoduje wiele wymiernych korzyści wynikających z następujących przyczyn:

- ◆ stali klienci przynoszą firmie zdecydowanie większe korzyści niż klienci incydentalni,
- ◆ bardziej kosztowne są kampanie mające na celu pozyskanie nowych klientów, niż te, których celem jest zatrzymanie obecnych,
- ◆ wiele więcej kosztuje odzyskanie klienta, który odszedł do konkurencji, niż sprawienie by był on usatysfakcjonowany, a przez to mniej zainteresowany ofertą konkurencji,
- ◆ łatwiej jest sprzedać nowy produkt osobie, która jest już naszym klientem, niż osobie zupełnie nieznannej.



W przypadku firm mających olbrzymią ilość klientów nawiązywanie indywidualnych relacji z pojedynczym klientem jest praktycznie niemożliwe. Możliwa jest jednak uważna interpretacja rutynowo gromadzonych informacji o klientach i podejmowanie na tej podstawie takich działań, by klienci odnosili wrażenie, że są indywidualnie traktowani, że rozumiemy ich potrzeby. Jeśli firma potrafi pokazać klientom, że ich rozumie i może zaspokoić ich potrzeby, to jest bardziej prawdopodobne, że kontakty będą trwały nadal [6].

CRM - podstawowe pojęcia

Odpowiedzią na potrzebę zarządzania relacjami z klientami stała się metodyka CRM, będąca praktyczną realizacją filozofii marketingu relacyjnego. Często CRM jest traktowany jako pewien rodzaj systemu informatycznego, co znacznie zawęża znaczenie tego terminu. W rzeczywistości CRM jest pewną strategią działania, a systemy informatyczne są jedynie narzędziami wspierającymi tę strategię. Samo wdrożenie takiego systemu nie wystarczy, by poprawnie zarządzać relacjami z klientem, konieczne są jeszcze radykalne zmiany w organizacji i kulturze pracy.

By móc skutecznie zarządzać relacjami z klientem konieczne jest sprawne gromadzenie, przechowywanie i analiza napływających do organizacji informacji, a także prawidłowe wykorzystanie wyników tych analiz. To z kolei jest możliwe dopiero po wprowadzeniu odpowiednich standardów i strategii działania.

W CRM możemy wyróżnić trzy podstawowe poziomy (podsystemy): operacyjny, analityczny oraz interakcyjny.

CRM operacyjny ma na celu zbieranie danych transakcyjnych i danych o klientach, ich opinii o produktach, sprzedawcach i komunikacji. Zadaniem operacyjnego CRM jest również wsparcie telemarketingu. Do tego celu służą teleinformacyjne systemy automatycznego rozdzielania rozmów przychodzących (*Automatic Call Distribution*) czy też interaktywnej obsługi głosowej (*Interactive Voice Response*).

Zadaniem **CRM analitycznego** jest przetwarzanie i analiza danych, data mining w celu planowania marketingowego, segmentacji i strategii instrumentalnych.

CRM interakcyjny ma na celu kształtowanie bezpośrednich kontaktów z klientem i bieżące reagowanie, za pomocą różnorodnych kanałów dystrybucji i komunikacji, na pojawiające się wymagania.

CRM analityczny

Podsystemem, którego działanie ma największy wpływ na podejmowanie strategicznych decyzji odnośnie różnego rodzaju akcji i kampanii marketingowych, jest CRM analityczny. Zadaniem tego podsystemu jest dostarczenie informacji niezbędnych do podejmowania efektywnych działań skierowanych na klienta. Informacje te są często efektem bardzo skomplikowanych analiz data mining, tak więc informacja docierająca do osób



korzystających z tego podsystemu jest z reguły wysoce przetworzona. By móc uzyskać poprawne, wcześniej nieznane oraz pełne informacje, a także by móc mądrze je wykorzystać, konieczne jest:

- ◆ zdobycie i integracja zarówno wewnętrznych, jak i zewnętrznych informacji oraz zapewnienie ich kompletności w skali całej organizacji – **aspekt danych**,
- ◆ wydobycie z przygotowanych danych ukrytej w nich wiedzy oraz przedstawienie wyników w sposób umożliwiający ich praktyczne wykorzystanie – **aspekt analizy**,
- ◆ zapewnienie odpowiedniej kultury organizacyjnej umożliwiającej korzystanie z wyników analiz – **aspekt organizacji**.

Aspekt danych – stosowanie tego typu systemu ma sens jedynie w przypadku, gdy jesteśmy w stanie zapewnić odpowiednią jakość danych – ich dokładność, zgodność, kompletność itp. W organizacji powinna obowiązywać wysoka kultura gromadzenia i przechowywania danych. Wszystkie informacje, jakie docierają do organizacji, powinny być traktowane jako potencjalne źródło użytecznych informacji, mogących pomóc w zrozumieniu klientów, i rejestrowane w systemie. Systemem, który może ułatwić gromadzenie tego typu informacji, jest CRM operacyjny.

Z punktu widzenia zapotrzebowania organizacji na dane do analizy kluczową rolę odgrywa wdrożenie bazy danych służącej celom analitycznym, tak zwanej korporacyjnej hurtowni danych. Hurtownia danych integruje różnego rodzaju dane mogące pochodzić z CRM operacyjnego, z różnego rodzaju systemów transakcyjnych, księgowych, produkcyjnych itp. Innymi słowy, hurtownia integruje dane pochodzące z całej organizacji oraz jej otoczenia i umożliwia ich wykorzystanie do celów analitycznych. Tego typu baza danych musi spełniać szereg wymagań, z których najważniejsze to:

- ◆ zawartość bazy nie zmienia się w trakcie analizy – aktualizacja następuje w ściśle określonym czasie;
- ◆ jednym z wymiarów danych jest czas;
- ◆ struktura danych umożliwia generowanie wielowymiarowych zestawień i podsumowań;
- ◆ baza działa w trybie do odczytu – standardowy użytkownik nie ma możliwości jej modyfikacji;
- ◆ hurtownia jest odizolowana od innych systemów bazodanowych w przedsiębiorstwie.

Bardzo ważne jest, by utworzona baza analityczna przechowywała dane atomowe, niezagregowane, ponieważ tylko na podstawie takich danych możemy przeprowadzać dogłębne analizy statystyczne i data mining.

Aspekt analizy – zgromadzone przez nas dane zawierają w sobie wiele ukrytych informacji, niedostępnych przy pobieżnej analizie. Ukryte zależności, trendy i wzorce, innymi słowy ukryta wiedza, która może w znacznym stopniu zoptymalizować nasze wysiłki marketingowe, może zostać odkryta dzięki zastosowaniu odpowiednich narzędzi do analizy danych. Narzędzia analityczne tego typu noszą wspólną nazwę narzędzi do zgłębiania danych (*data mining*).



Tego typu narzędzia powinny:

- ◆ udostępniać wszystkie metody analityczne potrzebne podczas analizy;
- ◆ umożliwiać przeprowadzenie pełnego cyklu analizy danych, począwszy od preprocessingu, czyli wstępnej analizy, przez zasadniczą część analizy, aż po generowanie raportów;
- ◆ umożliwiać łatwe stosowanie zbudowanych modeli;
- ◆ oferować intuicyjny, prosty w obsłudze interfejs;
- ◆ umożliwiać szerokie dostosowywanie systemu oraz budowanie własnych procedur analitycznych.

Aspekt organizacyjny – nawet najlepiej rozwiązany problem gromadzenia i analizy danych nie da efektu, jeśli równocześnie z nimi nie nastąpią zmiany organizacyjne w przedsiębiorstwie. Dobrze skonstruowany model zbudowany na danych o dobrej jakości nie będzie przydatny, jeśli jego wyniki nie będą dostępne dla osób, które ich potrzebują. Dlatego też w parze ze zmianami technologicznymi, muszą iść zmiany w kulturze organizacyjnej. Tak naprawdę kultura CRM powinna obejmować całą firmę. Do osiągnięcia sukcesu konieczna jest współpraca różnych części i działów przedsiębiorstwa, które dostarczają informacji potrzebnych do lepszego zrozumienia naszych klientów.

Wartość klienta

Kluczowym pojęciem występującym w CRM jest wartość klienta. W literaturze można znaleźć następującą definicję [6]:

CRM jest strategią mającą na celu zoptymalizowanie całościowej wartości klientów.

Oczywiście trudno jest zdefiniować jednoznacznie pojęcie wartości klienta. Istotnymi czynnikami wpływającymi na jego wartość są na pewno dochody, jakie przynosi organizacji, a także jego lojalność oraz skłonność do polecenia nas znajomym. W zależności od podejścia i organizacji te definicje mogą się jednak znacznie różnić. Na pewno jednak wartości klienta nie można rozstrzygać, patrząc jedynie przez pryzmat kwoty, jaką u nas wydał, równie istotne są korzyści, jakie potencjalnie dany klient może przynieść w przyszłości.

Oczywiście sposób jego obliczania może być indywidualną kwestią danej organizacji, możemy obliczać go za pomocą technik data mining (również popularny wskaźnik wartości życiowej klienta *LTV-Life-time value*) lub za pomocą innych metod, niemniej jednak umiejętność obliczania tego typu wskaźnika pozwala na racjonalne planowanie i ocenę różnego rodzaju kampanii.

Znajomość wartości klienta jest bardzo istotną miarą pozwalającą ocenić skuteczność planowanych akcji marketingowych. Bez niego nie jesteśmy w stanie stwierdzić, czy planowana przez nas akcja przyniesie wymierny efekt.



Wykorzystując wskaźnik określający wartość klienta, przy pomocy odpowiednio zaplanowanego badania rynkowego, możemy oszacować końcowy efekt ekonomiczny planowanej kampanii. Tego typu badania mogą polegać na przykład na wyodrębnieniu spośród naszych klientów dwóch grup: testowej, do której wysłamy ofertę, oraz kontrolnej niebiorącej udziału w kampanii, która posłuży nam do oceny wyników przeprowadzonego badania.

Jeśli nie jesteśmy w stanie określić wartości klienta, korzyści z analiz data mining będą w dużym stopniu ograniczone. W takim przypadku możemy co prawda określić charakterystyki klientów lub określić, jakie grupy klientów będą zachowywać się podobnie w danej sytuacji. Jeśli jednak pragniemy zoptymalizować kampanię mającą określić, którzy z naszych klientów są skłonni odpowiedzieć na naszą ofertę, to bez prognozy opartej na wartości klienta nie będziemy w stanie stwierdzić, czy tego typu akcja w ogóle ma sens ekonomiczny. Musimy mieć pewną miarę, która decydowałaby o sensie podejmowania zaplanowanych działań.

Bez tego typu miary istnieje ryzyko, że za sukces uznamy kampanię, dzięki której udało się powstrzymać od odejścia znaczną liczbę klientów, którzy jednak nigdy nie przynosili nam zysku i nie roszą nadziei na poprawę dochodowości.

Możemy też stwierdzić, że kampania mającą na celu pozyskanie nowych klientów zakończyła się porażką, ponieważ jej koszty były większe od uzyskanego dochodu. Nie bierzemy jednak pod uwagę faktu, że klienci Ci, jeśli tylko byli zadowoleni z naszej usługi, mogą wydać w przyszłości kolejne sumy, które na pewno zrekompensują poniesione wydatki.

Rodzaje zadań analitycznych

Metody statystyczne i data mining mogą być ogromnie użyteczne w doskonaleniu relacji z klientem. Do najczęstszych obszarów ich zastosowań należy zaliczyć:

- ◆ segmentację,
- ◆ analizy związane z cyklem życia klienta,
- ◆ analizy dotyczące sekwencji zakupów oraz podobieństwa i powiązań między produktami,
- ◆ analizę satysfakcji klientów.

Segmentacja

Analiza *data mining* może pomóc zidentyfikować grupy podobnych do siebie klientów, innymi słowy dokonać ich segmentacji. Tego typu analiza może być przydatna do optymalizowania naszych działań i bardziej indywidualnego traktowania naszych klientów, ponadto może być wstępem do dalszych analiz. Dalsze analizy przeprowadzamy już jedynie na wyodrębnionych segmentach, dzięki czemu możemy uzyskać bardziej



precyzyjne modele i uzyskać bardziej szczegółowe informacje. Możemy wyróżnić kilka rodzajów segmentacji. Pierwszy z nich to segmentacja opisowa, do której zaliczyć można:

- ◆ **segmentację demograficzną** - jest przeprowadzana na podstawie danych takich jak: dochód klienta, wiek, płeć, wykształcenie, stan cywilny, ilość osób w rodzinie, status mieszkaniowy, typ mieszkania, grupa etniczna, wyznaniowa itp.,
- ◆ **segmentację behawioralną** - jest przeprowadzana w oparciu o dane reprezentujące zachowanie klientów. W przypadku sklepu mogą to być informacje na temat częstości zakupów, ilości oraz rodzaju zakupionych produktów,
- ◆ **segmentację pod względem motywacji** - opiera się na zmiennych opisujących przyczyny, z powodu których klient dokonał zakupu. Dane tego typu zwykle uzyskiwane są na podstawie badań (np. ankietowych), gdyż z reguły nie jesteśmy w stanie odczytać motywacji klientów na podstawie danych opisujących ich zachowanie.

Drugi rodzaj segmentacji to segmentacja predykcyjna. Jest ona użyteczna do tego, by zrozumieć na przykład, jakie zmienne odróżniają dobrych klientów od złych. By móc przeprowadzić tego typu analizę, powinniśmy na początku ustalić zmienną określającą „dobrego” klienta. Dla tej zmiennej określamy, które z pozostałych zmiennych mają na nią największy wpływ. Oczywiście należy pamiętać, że segmentacja klientów ma sens jedynie wtedy, gdy zakładamy, że nasze zachowanie będzie różne w odniesieniu do różnych segmentów.

Cykl życia klienta

Wartościowy klient rzadko jest klientem statycznym. Relacja z nim ewoluuje i zmienia się w czasie. Zrozumienie tej relacji jest krytyczne dla zarządzania kontaktami i podejmowania odpowiednich działań. Różne poziomy w relacjach z klientami można opisać za pomocą modelu cyklu życia klienta. Zakłada on, że niektórzy z potencjalnych klientów, dokonując zakupu, zmieniają się w naszych nowych klientów, ci zaś, jeśli przez pewien czas korzystają z naszych usług, mogą być postrzegani jako nasi stali klienci. Następnie pewna ich część może zrezygnować, stając się naszymi byłymi klientami. Budowanie modeli predykcyjnych związanych z cyklem życia klienta jest najbardziej typowym przykładem wykorzystywania metod data mining w zarządzaniu relacjami z klientem. Oczywiście w zależności od etapu cyklu zmieniają się typy analiz, jakie można przeprowadzić.

Potencjalni klienci

Zdobywanie nowych klientów jest dla wielu firm głównym źródłem wzrostu sprzedaży. Skuteczność kampanii mającej na celu zdobycie nowych klientów jest więc kluczowa dla funkcjonowania organizacji. Tego typu akcje kierowane są do wybranej wcześniej grupy **potencjalnych klientów**, czyli osób, które mogą potencjalnie być zainteresowane naszymi produktami bądź usługami, lecz nie są naszymi klientami. Bardzo ważnym zadaniem stojącym przed pracownikami działu marketingu jest wybranie odpowiednich osób, do których zostanie skierowana przygotowana oferta. Bardzo pomocne w wyborze segmentu



potencjalnych klientów są narzędzia data mining. Sprawdzają się one zwłaszcza w sytuacji, gdy dysponujemy olbrzymią bazą potencjalnych klientów, z których musimy wskazać jedynie tych najbardziej odpowiednich (ze względu na koszty kampanii nie planujemy wysyłać oferty do wszystkich).

Budowanie modeli data mining służących do selekcji osób, do których warto skierować ofertę, jest dość trudna, z powodu charakteru danych wykorzystywanych w analizie. Z reguły są to dane zakupione z zewnętrznej instytucji, zawierające najczęściej jedynie informacje demograficzne. Na podstawie tego typu informacji trudno budować skuteczne modele prognozujące zachowanie, stąd też zwykle cechują się one znacznym błędem klasyfikacji. Nie zmienia to jednak faktu, że mogą być one ogromnie pomocne w ograniczaniu kosztów kampanii. Zwykle kampanie tego typu cechują się znikomym odsetkiem (mniej niż 1%) osób, które na nie odpowiadają. Celem modelu nie jest tak naprawdę dokładne wskazanie tych osób, lecz pewne zawężenie grupy, w której się znajdują. Jeśli model wskaże, że warto wysłać ofertę do połowy naszych potencjalnych klientów, wśród których będą wszyscy rzeczywiście skłonni na nią odpowiedzieć, to może to w dużym stopniu ograniczyć koszty kampanii, nie zmniejszając jej skuteczności.

Aktualni klienci

Kolejna grupa (faza cyklu życia klienta) to osoby, które odpowiedziały na naszą ofertę i są nią zainteresowane, ewentualnie **dokonały zakupu po raz pierwszy**. Każda z tych osób jest inna. Nie wszyscy będą wartościowymi klientami, lojalnymi i kupującymi za duże kwoty. Duża część z nich nie będzie przynosiła zysków (czasami może wręcz generować straty). Modele budowane dla tej grupy klientów mają na celu wskazanie, czy nowy klient będzie wysoce dochodowy, przeciętny czy też będzie on przynosił niski dochód (stratę). Skorzystanie z wyników analizy umożliwia odpowiednie skanalizowanie działań marketingowych i skupienie się jedynie na osobach mogących potencjalnie być dobrymi klientami. W tym wypadku ilość informacji na temat klientów, jaką dysponujemy, jest o wiele bogatsza; po pierwsze dysponujemy pełnym i aktualnym zestawem informacji demograficznych, dodatkowo bardzo dobrym predyktorem dalszego zachowania klienta jest jego pierwszy zakup, niejako inicjujący kontakty z firmą.

Gdy pomiędzy dostawcą a klientem została nawiązana relacja, kolejnym krokiem jest maksymalizacja korzyści z niej wynikającej. Jedną z możliwości osiągnięcia tego celu jest stosowanie wobec klientów tak zwanej sprzedaży krzyżowej (*cross-selling*). Polega ona na oferowaniu klientom pewnych dodatkowych usług i produktów związanych z wcześniejszymi zakupami (np. posiadacz rachunku w banku decyduje się na korzystanie z karty kredytowej). Można również starać się, by nasi klienci zaczęli korzystać z poszerzonych wersji zakupionych wcześniej produktów i usług (*up-selling*). Przykładem takiej oferty może być propozycja przejścia na wyższą taryfę użytkownikom korzystającym z taryfy standardowej.

Prawidłowo przeprowadzone kampanie mające na celu zwiększenie jednego lub obu typów sprzedaży przynoszą z reguły prawie same korzyści zarówno sprzedającym, jak i kupującym. Dzięki nim firmy osiągają zdecydowanie lepsze wyniki finansowe, ograniczają



ryzyko podejmowanych decyzji oraz zapewniają sobie lojalność klientów. Klienci natomiast oszczędzają czas i pieniądze, mają lepszy dostęp do informacji i mogą korzystać ze specjalnych rabatów przyznawanych stałym klientom. Dodatkowo koszt dotarcia do aktualnych klientów jest znacznie mniejszy niż koszt kampanii skierowanej do całego rynku [5].

Umiejętne zaplanowanie i przeprowadzenie tego typu kampanii wymaga wiedzy odnośnie zachowania i zwyczajów klientów w zakresie dokonywanych przez nich zakupów. Na przykład klient kupujący pieluszki może być zainteresowany nabyciem innych produktów do pielęgnacji niemowlaka. Ten przykład jest stosunkowo prosty do przewidzenia i tak naprawdę każdy może określić szereg tego typu zależności. Prawdziwa wartość wynika z odkrycia pewnych zależności i wzorców, które są o wiele bardziej subtelne i trudne do wychwycenia. Często jedynie wnikliwa analiza danych może pomóc w odnalezieniu tego typu wzorców.

Metody data mining mogą być bardzo pomocne we wspieraniu tego typu akcji. Zbudowane modele służą do prognozowania, którzy klienci byliby skłonni zakupić dany produkt. W większości forma analizy niewiele się różni od analizy stosowanej dla potencjalnych klientów określającej, jaki produkt im zaoferować (za wyjątkiem faktu, iż w takim przypadku dysponujemy szeregiem informacji pochodzących z wcześniejszych kontaktów z klientem). Każda możliwość transakcji sprzedaży krzyżowej jest traktowana jako oferowanie klientowi jednego produktu. Każdemu produktowi odpowiada jeden model prognostyczny. Kluczowe jest następnie zoptymalizowanie oferty w stosunku do wszystkich klientów tak, by oferta lub oferty, jakie otrzyma klient, przyniosły najwięcej korzyści zarówno sprzedającemu, jak i kupującemu. Podczas budowy tego typu modeli ważne jest takie ich skonstruowanie, by prognozy były sporządzane nie dla stanu obecnego, lecz z pewnym wyprzedzeniem (np. 2 miesiące). Tego typu wyprzedzenie jest konieczne, by oferta zdążyła dotrzeć do klienta jeszcze przed jego ewentualnym kontaktem z konkurencją.

Klienci zagrożeni odejściem oraz byli klienci

Na wartość klienta wpływa nie tylko kwestia ilości dokonywanych przez niego zakupów. Drugim równie istotnym czynnikiem jest trwałość tej relacji. Klient regularnie korzystający z naszych usług jest bardzo cenny dla organizacji, dlatego zapewnienie jego lojalności jest jednym z kluczowych zadań, jakie stoją przed osobami zajmującymi się zarządzaniem relacjami z klientami. Odejścia (*churn*) klientów stanowią poważny problem dla wielu firm, zwłaszcza tam, gdzie z jednej strony występuje wysoki koszt pozyskania nowych klientów, a z drugiej strony klient bardzo łatwo może odejść do konkurencji, ponosząc przy tym niewielkie lub żadne koszty. W sytuacji gdy rynek nasycy się produktem lub usługą i jest coraz mniej potencjalnych klientów, znaczenie utrzymania klienta staje się coraz większe.

Chociaż klienci rezygnują z naszych usług z wielu przyczyn, tak naprawdę wstępują dwa różne typy odejść. Pierwszy z nich to odejścia dobrowolne. Odnoszą się do stałych klientów, którzy po prostu nie chcą być naszymi klientami. Sytuacja taka może być



spowodowana wieloma różnymi czynnikami: zmiana miejsca zamieszkania, zmiana stylu życia czy zainteresowań lub też lepsza oferta konkurencji. Z drugiej strony istnieje druga grupa rezygnacji, gdzie inicjatywa stoi po stronie dostawcy. To on ucina kontakty z klientem, w związku z niewywiązywaniem się przez klienta z zawartej umowy (np. odłączenia abonentów niepłacących za rozmowy).

Oczywiście obydwie grupy wymagają uważnej analizy, niemniej jednak cel analiz dla obu grup znacznie się różni. Zasadniczym pytaniem, jakie należy postawić sobie, zajmując się kwestią dobrowolnych odejść brzmi: *Kiedy klient zrezygnuje z naszych usług?* Znajac odpowiedź na to pytanie, możemy podjąć pewne kroki zapobiegawcze, ograniczające skalę odejść. W przypadku drugiej grupy cel analizy jest inny (nie ma sensu zabiegać o klienta, który nie płaci). Dla tego typu odejść pytanie brzmi: *Czy istnieje ryzyko, że klient przestanie wywiązywać się z warunków zawartej umowy?* Odpowiedź na to pytanie może umożliwić szybką interwencję i zapobiec generowaniu strat przez klienta.

Metody analizy danych, a w szczególności data mining stosowane są w zapobieganiu migracji klientów w rozmaity sposób. W szczególności techniki statystyczne wykorzystuje się do badania i monitorowania satysfakcji klienta, wykrywania przyczyn decydujących o zadowoleniu klienta i jego lojalności, monitorowania zmian stopnia satysfakcji klientów.

Najpopularniejsze rozwiązanie, które bezpośrednio odnosi się do sformułowanych powyżej pytań, polega na znalezieniu tych klientów, dla których prawdopodobieństwo odejścia jest największe. Tego typu analizy przeprowadzamy w oparciu o dane dotyczące naszych obecnych i byłych klientów za pomocą technik ukierunkowanego data mining. Podobnie jak w przypadku modeli dla sprzedaży krzyżowej, również tutaj prognozy buduje się z wyprzedzeniem czasowym, by dostawca miał czas podjąć pewne kroki zapobiegawcze. Drugim zadaniem jest identyfikacja przyczyn potencjalnego odejścia. Wiedza ta jest konieczna, by podjąć w stosunku do zagrożonych klientów optymalne kroki zapobiegawcze (odejścia dobrowolne) bądź optymalizujące działania dotyczące potencjalnie niewypłacalnych klientów. W przypadku drugiej części zadania do analizy wykorzystujemy najczęściej metody nieukierunkowanego data mining.

Ostatni etap cyklu życia klienta to osoby, które zrezygnowały z naszych usług. W tym przypadku także bardzo pomocne mogą być analizy data mining wskazujące klientów, do których warto ponowić ofertę. Tego typu modele budowane są przy pomocy metod ukierunkowanego data mining.

Analizy dotyczące sekwencji zakupów oraz podobieństwa i powiązań między produktami

Kolejne zadanie wykonywane dzięki technikom data mining polega na określeniu powiązań między poszczególnymi produktami, a dokładniej na próbie odpowiedzi na pytanie: *Jakie produkty są zwykle kupowane razem lub w pewnych odstępach czasu oraz w jakiej kolejności?* Odpowiedzią na to pytanie jest zwykle zestaw pewnych reguł zbudowany na podstawie historycznych danych (transakcyjnych), zawierających zapis



dokonywanych zakupów. Znajomość takich reguł może być bardzo przydatna w różnego rodzaju działaniach skierowanych na klienta. Wiedza na temat kupowanych produktów umożliwia także uzyskanie informacji odnośnie naszych klientów oraz motywów, jakimi się kierują, dokonując określonych zakupów.

Najpopularniejsze zastosowanie tego typu reguł polega na wykorzystywaniu zawartej w nich wiedzy do planowania optymalnego układu towarów w sklepie. Jeśli wykryta reguła wskazuje, że dwa (lub kilka) produkty kupowane są zwykle razem, wtedy można je umieścić na sąsiadujących półkach, by ułatwić klientom zakupy (lub przeciwnie na odległych w nadziei, że klient, mijając kolejne półki, zdecyduje się na zakup dodatkowych towarów). Na podstawie otrzymanych reguł możemy także określić, które towary powinny być szczególnie eksponowane, oraz które warto zaoferować po promocyjnej cenie.

Bardzo często wyniki takich analiz stanowią wstęp do dalszych dociekań. Tego typu analizy poprzedzają często opisywane powyżej kampanie sprzedaży krzyżowej. Znajomość reguł postępowania klientów pozwala nam określić, jaki następny produkt najprawdopodobniej zakupi dany klient, i na tej podstawie podejmować działania odnośnie kierowania ofert.

Znajomość tego typu reguł jest również bardzo przydatna w handlu internetowym. Podczas zakupów w sklepie internetowym wszystkie transakcje zapisywane są w bazie danych, a klient, który ich dokonywał, jest jednoznacznie identyfikowany w systemie, tak więc przechowywana jest cała historia jego zakupów. Na podstawie tych danych generowane są wzorce, dzięki którym program z dużym prawdopodobieństwem „wie”, jakimi produktami dany klient może być zainteresowany, i już na wstępie może zaproponować odwiedzającemu określony asortyment.

Generowanie reguł jest możliwe dzięki zastosowaniu metod nieukierunkowanego data mining: analizy asocjacji (*association rules*) oraz analizy sekwencji i połączeń (*sequence and link analysis*). Oczywiście nie wszystkie odnalezione reguły muszą być użyteczne. Część z nich na pewno będzie przedstawiała zależności znane analizującym, część reguł może być niemożliwa do interpretacji. Jednak pozostałe mogą znacznie poszerzyć wiedzę na temat badanego zjawiska, sugerując określone działania lub wskazując obszary dalszej analizy.

Bardzo przydatne w zrozumieniu badanego zjawiska są narzędzia wizualizacji, dzięki którym możemy analizować całą siatkę połączeń pomiędzy poszczególnymi obiektami, co umożliwia głębsze zrozumienie wzajemnych zależności i relacji.

Pomiar satysfakcji klientów

Satysfakcja klientów określana jest jako poziom zaspokojenia potrzeb klienta przez całkowity produkt, jaki nabył u dostawcy. Pod pojęciem produktu całkowitego rozumiemy również obsługę klienta oraz wszystkie usługi późniejsze (przeeglądy serwisowe itp.). Pomiar satysfakcji klienta jest kluczowy w zrozumieniu rzeczywistych jego potrzeb i oczekiwań. Dzięki niemu jesteśmy w stanie powiedzieć, co jest naszą mocną stroną,



a jakie aspekty naszej działalności powinniśmy zmienić. Satysfakcja klienta bezpośrednio wpływa na jego stosunek do nas. Im bardziej nasz produkt zaspokaja potrzeby klienta, tym chętniej powtarza zakupy, natomiast przykre incydenty zwykle skłaniają go do zmiany dostawcy. Dlatego też kluczową sprawą w zarządzaniu relacjami z klientem jest pomiar, a następnie monitorowanie poziomu satysfakcji klientów.

W celu mierzenia i monitorowania satysfakcji wśród klientów (rzeczywistych i potencjalnych) przeprowadza się odpowiednie badania marketingowe i na tej podstawie określa się przyczyny ich konkretnych postaw i zachowań. Badania marketingowe można podzielić na kilka typów, które stosuje się w zależności od wiedzy na temat badanego zjawiska, jaką na wstępie dysponuje badający. Typy te możemy traktować jako etapy pewnego ciągłego procesu. Możemy wyróżnić badania:

- ◆ eksploracyjne lub rozpoznawcze, w których główny akcent spoczywa na odkryciu idei danego zagadnienia, podziale problemu na mniejsze części, sformułowaniu hipotez dotyczących zachowania klientów,
- ◆ opisowe - zajmujące się określeniem częstotliwości występowania jakiegoś zjawiska lub znalezieniem związków pomiędzy dwoma zjawiskami,
- ◆ przyczynowe - polegające na określaniu zależności przyczynowo-skutkowych, często w oparciu o eksperyment.

Podczas planowania badań oraz analizy ich wyników niezbędne jest wykorzystanie technik statystycznych. Głównymi zadaniami, jakie stoją przed analitykiem, jest: wyznaczenie odpowiedniej liczebności (reprezentatywności) próby, na której ma być przeprowadzane badanie, określenie skali, na której będzie odbywał się pomiar satysfakcji, opracowanie odpowiednich pytań, a następnie analiza wyników, podczas której obok całej gamy testów statystycznych można używać innych, często bardzo zaawansowanych metod, jak np. karty kontrolne, analiza czynnikowa czy głównych składowych.

Bardzo ważne jest, by pomiary satysfakcji klientów nie były przedsięwzięciem jednorazowym. Aby określić efekty naszych wysiłków, bądź w porę wychwycić niekorzystne zjawiska, konieczne jest regularne monitorowanie poziomu satysfakcji.

Metody data mining w CRM analitycznym

Wszystkie przedstawione powyżej rodzaje analiz w praktyce wykonuje się za pomocą technik *data mining*. Pod tą nazwą kryje się w rzeczywistości szereg różnego rodzaju metod analitycznych. Ich zastosowanie zwykle się różni, dlatego też wprowadzono szereg klasyfikacji porządkujących te metody. Najbardziej odpowiedni z punktu widzenia przedstawionych zastosowań wydaje się podział na ukierunkowany *data mining* (uczenie z nauczycielem) oraz nieukierunkowany *data mining* (uczenie bez nauczyciela).

Metody ukierunkowanego *data mining* są wykorzystywane w przypadkach, gdy interesujący nas podział został już zaobserwowany i zapisany w próbie uczącej. Wykorzystując te dane, budujemy model, który ma na celu przewidzieć ten podział dla nowych danych.



Mamy na przykład zbiór zawierający pewne dane o klientach oraz informację, czy dany klient odpowiedział na naszą specjalną promocję (mamy więc podział na klientów, którzy odpowiedzieli, i tych, którzy nie odpowiedzieli). Na podstawie zgromadzonych danych budujemy model, którego celem jest postawienie prognozy, kto z nowej grupy klientów (nieuczestniczących w starej promocji) byłby skłonny odpowiedzieć na podobną akcję promocyjną. Tego typu modele mogą znacznie obniżyć koszty naszych działań, nie obniżając naszej skuteczności, ponieważ nasze wysiłki kierujemy jedynie do osób, wobec których istnieje szansa, że odpowiedzą na naszą ofertę. Techniki ukierunkowanego *data mining* stosujemy w segmentacji predykcyjnej, jak również we wszystkich analizach odnoszących się do cyklu życia klienta. Do metod ukierunkowanego *data mining* zalicza się m.in.: drzewa decyzyjne, sieci neuronowe, *MARSplines*, *Naive Bayes* oraz *Support Vector Machine*.

W sytuacji nieukierunkowanego *data mining* sytuacja jest inna. Mamy pewne dane o kliencie, które jednak nie zawierają żadnego podziału narzuconego z góry. Naszym celem jest wyjaśnienie struktury tych danych, na przykład określenie skupisk czy zależności niewidocznych na pierwszy rzut oka. Możemy przykładowo dysponować bazą danych klientów zawierającą różnego typu zmienne, które potencjalnie mogą mieć duży wpływ na zachowanie klientów. Celem analizy może być odnalezienie segmentów rynku, czyli grup osób podobnych do siebie, bądź odnalezienie związków pomiędzy kupowanymi towarami. Do technik nieukierunkowanego *data mining* zaliczyć można m.in.: analizę skupień, analizę koszykową (asocjacji), analizę sekwencji i połączeń, sieci neuronowe oraz analizę czynnikową i składowych głównych.

Literatura

1. *Analiza satysfakcji i lojalności klienta*, Materiały Seminaryjne StatSoft, Kraków 2003.
2. Berry M., Gordon L., *Data Mining Techniques for Marketing, Sales and Customer Support*, John Wiley & Sons, Inc, New York 1998.
3. Berry M., Gordon L., *Mastering Data Mining. The Art and Science of Customer Relationship Management*, John Wiley & Sons, Inc, New York 2000.
4. Berson A., Smith S., Thearling K., *Building Data Mining Applications for CRM*, McGraw Hill, New York 2000.
5. Szczepaniec M., *Sprzedaż krzyżowa (cross-selling) produktów bankowych*, Bank i kredyt, 2003.
6. Todman C., *Projektowanie hurtowni danych. Zarządzanie relacjami z klientami (CRM)*, Wydawnictwo Naukowo-Techniczne, Warszawa 2003.