



WSPOMAGANIE KAMPANII SPRZEDAŻY KRZYŻOWEJ (CROSS-SELLING) NA PRZYKŁADZIE OFERTY BANKU

Grzegorz Migut, StatSoft Polska Sp. z o.o.

Wstęp

W ciągu ostatnich lat bardzo dużą popularność zdobyły sobie strategie marketingowe oparte na filozofii marketingu relacyjnego i związanej z nim metodyki CRM. Dla przedsiębiorstw stało się oczywiste, że w działaniach marketingowych ważne jest nie tylko zdobywanie nowych klientów. Równie duże znaczenie ma budowa indywidualnych trwałych relacji z aktualnymi klientami i maksymalizacja korzyści płynących z tej relacji.

Budowanie indywidualnych, trwałych kontaktów ma na celu zwiększenie wartości klienta. Choć termin „wartościowy klient” ma zwykle dla każdej firmy odmienne znaczenie, jednak niezmiennymi składowymi tego pojęcia są: lojalność i dochodowość klienta. Wpływ tych dwóch czynników na wartość klienta wydaje się być oczywisty. Wartościowym klientem jest osoba, która przynosi nam dochód i dodatkowo jej relacja z nami jest trwała.

Oba kluczowe aspekty związane z wartością klienta mogą być optymalizowane dzięki kampaniom sprzedaży krzyżowej. Sprzedaż krzyżowa polega na oferowaniu aktualnym klientom pewnych dodatkowych usług i produktów związanych z wcześniejszymi zakupami. Możemy na przykład zaproponować osobie kupującej odtwarzacz DVD zakup stolika pod ten odtwarzacz lub posiadaczowi rachunku skorzystanie z kredytu. Pewną odmianą sprzedaży krzyżowej jest sprzedaż rozszerzona (*up-selling*). Polega ona na proponowaniu klientowi produktu lub usługi, z jakiej już korzysta lub zamierza korzystać, lecz o większej funkcjonalności (i zwykle wyższej cenie). Typowymi przykładami takiej sprzedaży są propozycje zmiany taryfy ze standardowej na wyższą, zmiana karty kredytowej ze zwykłej na złotą, czy też propozycja zakupu samochodu z pełnym wyposażeniem zamiast z wyposażeniem standardowym.

Akcje sprzedaży krzyżowej i rozszerzonej możemy podzielić na dwie grupy: na sprzedaż bezpośrednią i niebezpośrednią. Sprzedaż bezpośrednia występuje w momencie podejmowania przez klienta decyzji o zakupie danego produktu, np. klient jest w sklepie i pragnie kupić telefon komórkowy, a sprzedawca zachęca go do zakupu modelu lepszego niż planował pierwotnie (*up-selling*) lub zestawu głośnomówiącego lub etui (*cross-selling*). W tym przypadku kluczową rolę odgrywa wiedza sprzedawcy na temat oferowanych produktów oraz jego umiejętności sprzedażowe.



Nieco inna sytuacja występuje, gdy klient dokonał zakupu w przeszłości, naszym celem jest natomiast przygotowanie dla niego oferty na dodatkowe produkty, które z największym prawdopodobieństwem byłby skłonny kupić. Oferta ta może dotrzeć do niego wieloma kanałami komunikacji: tradycyjną pocztą, pocztą elektroniczną lub zostać przedstawiona podczas rozmowy telefonicznej.

Korzyści, zagrożenia oraz wymagania

Prowadzenie kampanii sprzedaży krzyżowej może przynieść wiele korzyści, zarówno sprzedającemu, jak i jego klientom, niestety niesie także ze sobą szereg zagrożeń.

Jeśli chodzi o korzyści z tego typu kampanii, to jest nią na pewno zwiększenie dochodowości aktualnych klientów. Jeśli w odpowiednim czasie dotrze do klienta dobrze przygotowana oferta, istnieje duża szansa, że klient z niej skorzysta. Korzystał będzie tym samym z szerszego zestawu produktów lub z produktów droższych. Liczba produktów i usług, z jakich dany klient jednocześnie korzysta, ma także ogromny wpływ na jego lojalność. Jeśli przykładowy klient posiada rachunek w jednym banku, lokatę w drugim, a posługuje się kartą kredytową wystawioną przez trzeci, tak naprawdę nie identyfikuje się z żadnym z nich. Jeśli natomiast w pełnym zakresie korzysta z usług jednego banku istnieje szansa, że do pewnego stopnia się z nim identyfikuje i z kolejnych usług bankowych skorzysta właśnie za jego pośrednictwem. Nie bez znaczenia jest również fakt, że korzystanie z danej usługi znacznie utrudnia zerwanie kontaktów z dostawcą. Jest tak na przykład w przypadku klienta banku, który skorzystał z kredytu hipotecznego. Odejście klienta posiadającego taki kredyt, choć oczywiście możliwe, wiąże się jednak z wieloma uciążliwymi formalnościami, które zwykle skutecznie zniechęcają przed podejmowaniem takich działań.

Inną zaletą kampanii sprzedaży krzyżowej jest fakt, że są one znacznie tańsze od kampanii związanych z pozyskiwaniem nowych klientów. W przypadku nowych klientów koszty ich pozyskania są zwykle duże, ponieważ wiążą się z koniecznością przeprowadzenia kosztownych kampanii reklamowych oraz oferowaniem nowym klientom specjalnych warunków w początkowym okresie korzystania z usług. Obsługa stałych klientów jest zdecydowanie mniej kosztowna. Największy koszt związany z prowadzeniem kampanii sprzedaży krzyżowej związany jest z odpowiednim wytypowaniem osób skłonnych kupić dany produkt. By móc wykonać tego rodzaju typowanie, konieczny jest najczęściej zakup specjalnego oprogramowania, służącego do budowy modeli *data mining* oraz zarządzania kampaniami marketingowymi. Koszt ten jednak zwraca się niekiedy już po przeprowadzeniu jednej kampanii.

Sprzedaż krzyżowa przynosi również korzyści klientom. Pierwsza korzyść wiąże się z kompleksowością świadczonej im usługi. Dla klientów może być dużo bardziej wygodne korzystanie z usług jednej instytucji, ponieważ w jednym miejscu mogą załatwić szereg kwestii, oszczędzając przy tym czas. Ponieważ są stałymi klientami, mogą również liczyć na specjalne traktowanie, np. korzystać z różnego rodzaju zniżek i rabatów.



Oczywiście sprzedaż krzyżowa niesie za sobą również szereg niebezpieczeństw. Podczas planowania tego typu kampanii należy pamiętać, że nie mogą one polegać na wysłaniu wszystkich przewidzianych ofert do każdego z klientów. Akcja przeprowadzona w ten sposób przyniosłaby raczej skutek odwrotny do zamierzonego, powodując spadek zainteresowania klientów ofertami od nas, cofnięcie zgody na przysyłanie informacji reklamowych, a w niektórych przypadkach wręcz rezygnację z naszych usług. Dlatego też kluczowym zadaniem planujących kampanię sprzedaży krzyżowej jest zapewnienie, by oferty dotarły jedynie do osób potencjalnie nimi zainteresowanych oraz by liczba ofert docierających do klienta nie była zbyt duża, najlepiej tylko jedna w danej kampanii.

Przed planowaniem tego typu kampanii należy wziąć pod uwagę różnego rodzaju ograniczenia. Należy pamiętać, że prowadzenie sprzedaży krzyżowej ma sens jedynie wtedy, gdy nasza oferta zawiera szeroki zestaw różnych produktów, komplementarnych względem siebie, jeśli planujemy kampanię *cross-sellingową*, lub względem siebie substytucyjnych, jeśli planujemy kampanię *up-sellingową*. Jeśli nasza oferta jest mało zróżnicowana, wtedy stosowność tego typu kampanii jest ograniczona.

Należy także pamiętać, że główna zaleta kampanii sprzedaży krzyżowej polega na maksymalizacji korzyści z już posiadanych klientów. Cel ten jest szczególnie ważny w sytuacji, gdy rynek, na którym działamy, jest już nasycony i trudno jest zdobyć nowych klientów. Na rynkach młodych, gdzie zdobycie klienta jest stosunkowo proste, zasadność prowadzenia tego typu kampanii jest mniej oczywista.

Poza ograniczeniami wynikającymi z posiadanej oferty i poziomu nasycenia rynku, występują również pewne zagrożenia związane z przyjętym modelem organizacyjnym sprzedaży i związaną z tym zagadnieniem dostępnością informacji. Jeśli każda grupa produktów jest obsługiwana przez inny dział sprzedaży, a każdy z tych działów przechowuje informacje o swoich klientach w odrębnych repozytoriach, wtedy stosowność kampanii sprzedaży krzyżowej ze względu na niekompletność informacji ograniczyć się może do pojedynczych działów. Oczywiście trudno oczekiwać od takich kampanii zadowalającej skuteczności.

Przeprowadzenie skutecznej kampanii sprzedaży krzyżowej wymaga więc integracji danych z tych źródeł, tak by informacje o klientach były możliwie jak najbardziej kompletne. Operacja ta, poza oczywistymi wyzwaniem technologicznymi, niesie również wyzwania organizacyjne związane z obawami sprzedawców przed zabraniami im „ich klientów”, o których informacja stanie się dostępna dla osób z innych działów.

Rola technik analizy danych w *cross sellingu*

Niezależnie od typu sprzedaży krzyżowej kluczowym czynnikiem decydującym o jej powodzeniu jest znajomość oczekiwań i zwyczajów klientów. Bardzo pomocne w poznawaniu klientów są narzędzia służące do zgłębiania danych (*data mining*). Wykorzystanie tych technik jest naturalne z powodu z reguły sporej ilości danych zgromadzonych na temat każdego z klientów – dane demograficzne klienta, historia jego transakcji. Dzięki



analizie tych danych możemy odkryć ukryte, nieznanne wcześniej zależności oraz zidentyfikować reguły zachowań klientów niemożliwe do wykrycia w inny sposób.

Pierwszym rodzajem analizy, jaki może być przydatny podczas akcji sprzedaży krzyżowej, jest analiza koszykowa oraz analiza sekwencji. Analiza koszykowa umożliwia odkrycie w danych dotyczących zakupów pewnych wzorców i reguł mających kształt wyrażeń logicznych typu: *JEŻELI klient kupił produkty A i B, TO kupił również produkt C*. Analiza sekwencji tworzy podobne reguły, z tym że uwzględnia również aspekt czasowy w dokonywaniu zakupów. W tym przypadku zakupy produktów A, B i C nie muszą być dokonywane jednocześnie, lecz mogą zachodzić w pewnych odstępach czasu. Znajomość tego typu reguł daje nam wiedzę, jakie dodatkowe produkty mogą być skłonni kupić klienci i co im zaoferować. W przypadku reguły przedstawionej powyżej naturalnym jest zaproponowanie wszystkim klientom posiadającym produkty A i B produktu C, ponieważ był on często kupowany w połączeniu z nimi. Reguły te możemy w naturalny sposób wykorzystać zarówno podczas bezpośredniej, jak i pośredniej sprzedaży krzyżowej. Bardzo często takie reguły zaimplementowane są w mechanizmach sklepów internetowych, które sugerują dalsze zakupy na podstawie wybranego produktu i transakcji dokonanych wcześniej. W tradycyjnych sklepach (innych niż samoobsługowe) reguły te mogą być stosowane bezpośrednio przez sprzedających.

Drugi rodzaj analiz polega na wykorzystaniu technik zgłębiania danych do budowania modeli predykcyjnych służących do typowania klientów, do których warto skierować określoną ofertę. Celem ich budowy jest określenie, jaki produkt powinien zostać zaoferowany konkretnemu klientowi i jaki jest najlepszy kanał komunikacji z klientem. Modele te stosowane są podczas sprzedaży niebezpośredniej w sytuacji, gdy klient dokonał zakupu w przeszłości i chcemy wskazać dla niego najbardziej odpowiedni kolejny produkt. Na przykład planujemy zaoferować klientom kilka produktów bankowych i chcemy wytypować osoby skłonne odpowiedzieć na poszczególne oferty. Ogólna zasada budowy tego typu modeli polega na przygotowaniu skoringów określających szanse na przyjęcie przez klienta danej oferty. Skoringi te budowane są na podstawie zachowań innych klientów w przeszłości

Budowa modeli skoringowych

Budowa modeli skoringowych służących do przewidywania skłonności klientów do zakupów zostanie zaprezentowana na przykładzie pewnego banku amerykańskiego¹. Bank ten, by zwiększyć dochodowość swoich klientów, zaplanował akcję marketingową, która miała zachęcić ich do skorzystania z różnych produktów bankowych. Kampania ta miała być realizowana na podstawie wyników uzyskanych dzięki modelom *data mining*. Ponieważ była to pierwsza kampania wspomagana technikami analizy danych, bank zdecydował się na projekt o ograniczonym zasięgu, który miał sprawdzić skuteczność proponowanego rozwiązania na przykładzie jednego produktu. W dalszej perspektywie

¹ Przykład oparty został na przykładach przedstawianych w pozycjach [1] i [2].



planowano wykorzystać metody *data mining* do wspierania kampanii odnoszącej się do większej liczby produktów.

Pierwsze zadanie, jakie należało rozwiązać, polegało na określeniu grupy docelowej, spośród której będą typowane osoby, do których trafi dana oferta. Zdecydowano, że kampania dotyczy będzie jedynie posiadaczy kont osobistych, nie będzie natomiast kierowana do posiadaczy kont komercyjnych.

Kolejna kwestia była związana z wytypowaniem produktów, jakie powinny być oferowane podczas kampanii, zarówno na jej testowym etapie, jak i podczas akcji docelowej. Ponieważ w ofercie banku znajdowało się kilkaset różnych produktów, oczywistym stało się, że nie można będzie przeprowadzić kampanii uwzględniającej w sposób szczegółowy każdego z nich. Konieczne stało się więc przygotowanie bardziej ogólnych kategorii, zawierających w sobie szczegółowe rodzaje produktów. Na przykład kilkanaście rodzajów lokat terminowych różniących się między sobą szczegółami uogólniono do jednej kategorii - Lokata Terminowa. Operację przygotowania ogólnych kategorii przeprowadzono oczywiście w porozumieniu z pracownikami banku, zgodnie z ich wytycznymi.

Zmiana poziomu szczegółowości produktów bankowych spowodowała wydzielenie 45 ogólnych grup produktów. Spośród tych produktów 25 mogło być oferowane klientom indywidualnym i one właśnie stanowiły cel planowanej kampanii. Akcja przeprowadzona miała być w ten sposób, że modele *data mining* miały wskazywać, z jakiej kategorii produktów dany klient byłby skłonny skorzystać, natomiast o szczegółowym rodzaju oferowanego produktu z tej kategorii miały decydować względy biznesowe. Jako produkt, na podstawie którego miano przetestować skuteczność proponowanego rozwiązania, wybrano konto maklerskie. Tego typu usługa cieszyła się małym powodzeniem wśród klientów banku, korzystało z niej mniej niż 5% wszystkich klientów.

Zarówno model testowy, jak i pozostałe modele miały być budowane na podstawie danych historycznych opisujących klientów banku. Dane te zawierały informacje demograficzne dotyczące klientów, informacje o produktach, z jakich korzystają oraz generowanym obrocie. Tak więc w przypadku konta maklerskiego informacja o tym, czy dana osoba z niego korzysta czy też nie, była zmienną objaśnianą, natomiast zmienne informujące o korzystaniu z pozostałych produktów oraz dane demograficzne były zmiennymi objaśniającymi.

Ogólna zasada działania tych modeli polega na wskazywaniu osób skłonnych do zakupu danego produktu na podstawie ich podobieństwa do osób, które już z tego produktu korzystają. Jeśli klient jest podobny (pod względem cech demograficznych, zachowania) do klientów, którzy kupili dany produkt w przeszłości, to istnieje duża szansa, że on również skorzysta z tej oferty i być może warto skierować do niego ofertę. Podczas budowy modelu należy uwzględnić fakt, iż zakup danego produktu może diametralnie zmienić rejestrowane zachowanie klientów. Na przykład po założeniu lokaty terminowej klienci mogą przechowywać znacznie mniejsze kwoty na rachunku bieżącym niż przed jego założeniem. Analizując dane takich osób, możemy stwierdzić, że większość osób posiadających lokatę ma niewielkie saldo na rachunku i błędnie skierować ofertę skorzystania z lokaty do podobnych osób. By uniknąć tego typu błędów, podczas analizy



należy brać pod uwagę informacje o klientach, którzy zakupili dany produkt, ale z okresu poprzedzającego jego zakup – jeśli dany klient założył lokatę pół roku temu, to nie analizujemy jego aktualnych danych, lecz te sprzed jej założenia.

Analizując częstość korzystania z poszczególnych produktów bankowych, można zauważyć, że różne produkty cieszą się różnym zainteresowaniem. Różnice w popularności poszczególnych produktów mogą wynikać z naturalnych zjawisk i zależności występujących wśród klientów. Liczba zleceń stałych lub poleceń zapłaty może być wyższa niż kredytów hipotecznych (rachunki musimy płacić wszyscy, natomiast nie wszyscy mamy niespłacone mieszkania). Należy jednak pamiętać, że pewne zależności mogą być wynikiem strategicznych działań banku podejmowanych w przeszłości. Niewielka liczba kredytów hipotecznych może wcale nie wynikać z małego potencjału rynku, ale może z faktu, iż bank unikał wcześniej udzielania tego typu kredytów. Jeśli w przeszłości kierowaliśmy ofertę na dany produkt jedynie do mężczyzn w średnim wieku, to musimy pamiętać, że model zbudowany na podstawie takich informacji również będzie typował wyłącznie takie osoby.

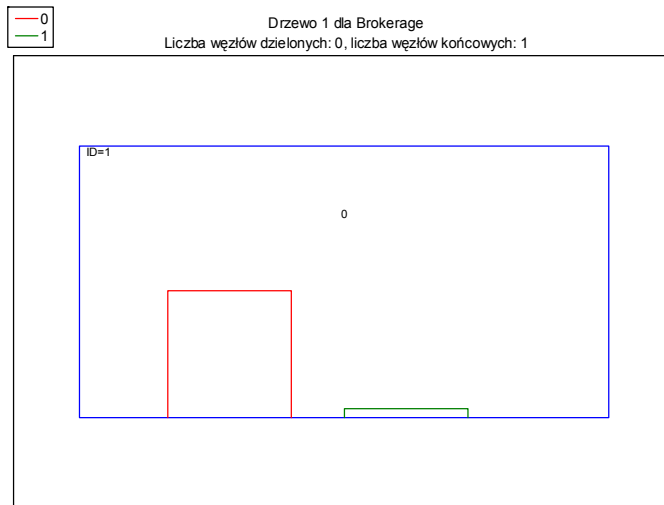
Budowa modelu prognozującego skłonność klientów do skorzystania z konta maklerskiego zostanie zaprezentowana w środowisku *STATISTICA Data Miner*. Podstawą tej analizy będzie zbiór danych zawierający 55 zmiennych objaśniających. Informacje demograficzne o klientach zapisane zostały w 11 cechach, natomiast pozostałe 44 zmienne informowały o transakcjach dokonanych przez tych klientów. Zmienną objaśnianą była zmienna *Brokerage*, informująca, czy dany klient posiada konto maklerskie (zmienna przyjmowała wartość 1) czy też nie (wartość 0). Analizowany zbiór zawierał blisko 25000 przypadków. Zbiór ten zawierał oczyszczone dane, przygotowane do zasadniczego etapu modelowania.

Za pomocą polecenia *Statystyka – Statystyki podstawowe i tabele – Tabele licznosci* zbadamy rozkład zmiennej *Brokerage*.

Analizując rozkład zmiennej *Brokerage*, możemy zauważyć, że jest on zgodny ze strukturą korzystania z konta maklerskiego przez klientów banku. Aż 95% przypadków przyjmuje wartość 0, a jedynie 5% przyjmuje wartość 1. Taki rozkład wartości zmiennej zależnej może znacznie utrudnić proces budowy modelu i spowodować, że zbudowane modele, mimo iż będą miały bardzo mały odsetek błędnych klasyfikacji, to jednak będą bezużyteczne w praktyce.

Dla zobrazowania tego problemu zbudujemy model predykcyjny na danych o zastanym pierwotnie rozkładzie zmiennej *Brokerage*. Do budowy modelu wykorzystamy moduł *Ogólne modele drzew klasyfikacyjnych i regresyjnych*. Zmienną *Brokerage* wskazujemy jako zmienną zależną, pozostałe cechy, w zależności od skali, na jakiej są przedstawione, wskazujemy jako jakościowe lub ilościowe zmienne niezależne. Przed uruchomieniem analizy na zakładce *Walidacja* wybieramy zmienną *Test*, informującą, czy dany przypadek ma znaleźć się w zbiorze uczącym czy testowym. Następnie uruchamiamy proces analizy.

Zgodnie z przewidywaniami algorytm nie dokonał żadnego podziału, cały zbiór został potraktowany jako jednorodny. Zbudowane drzewo składa się tylko z jednego węzła, nie ma żadnych węzłów potomnych.

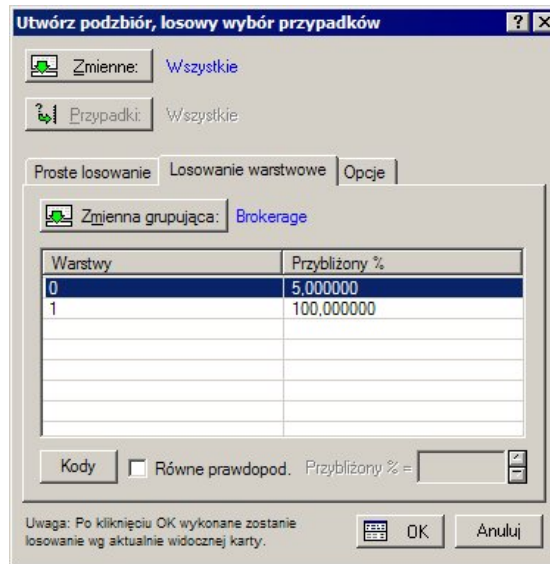


Zbudowany model ma co prawda bardzo dobrą jakość; jeśli chodzi o sumaryczny procent poprawnych odpowiedzi w 95% przypadkach podał prawidłową odpowiedź, co jest wynikiem praktycznie nieosiągalnym w rzeczywistości. Choć model ten bardzo dobrze odwzorował rzeczywistą zależność, że „prawie nikt nie korzysta z konta maklerskiego”, to jednak jest bezużyteczny w praktyce. Na jego podstawie nie jesteśmy w stanie wskazać, komu zaoferować konto maklerskie.

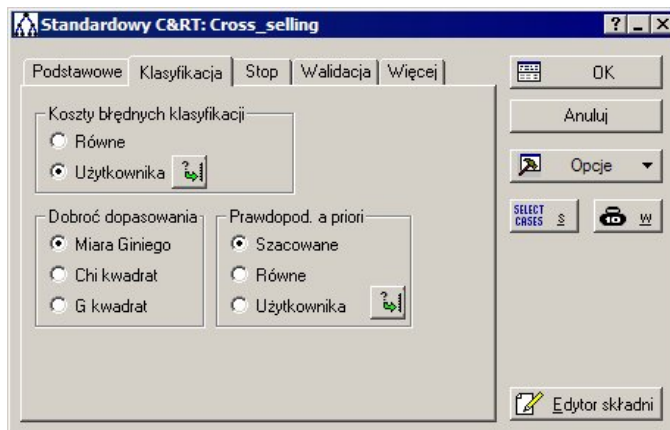
Takie zachowanie modelu można wytłumaczyć jego dążeniem do minimalizacji ogólnego poziomu błędów. Przy tak nierównej proporcji klas modelowi po prostu „nie opłaca się ryzykować” i wskazywać, że dany klient ma konto maklerskie. Jeśli wskaże, że wszyscy go nie mają, to osiągnie poziom błędów 5%, co jest wynikiem prawie idealnym.

By móc zbudować model, który będzie przydatny w praktyce, konieczna jest bądź modyfikacja zbioru danych, bądź zmiana parametrów algorytmu uczącego. Jedną z możliwości polega na zastosowaniu techniki *oversamplingu*, pozwalającej na stworzenie zbioru o proporcjach poszczególnych odpowiedzi wskazanych przez analityka.

Losowanie wykonujemy za pomocą polecenia *Dane-Podzbiór...* Następnie na zakładce *Losowanie Warstwowe* określamy proporcje poszczególnych klas. Losowanie przeprowadzamy w ten sposób, by do nowego zbioru danych trafiły wszystkie przypadki reprezentujące klientów, którzy skorzystali z konta maklerskiego, i 5% klientów, którzy z niego nie skorzystali. Tego typu operacja zrównoważy proporcje pomiędzy klasami i pozwoli zbudować model, który chociaż przypuszczalnie będzie cechował się większym błędem klasyfikacji, to najprawdopodobniej będzie bardziej użyteczny w praktyce.



Drugą możliwością, która zostanie wykorzystana w niniejszym przykładzie, jest modyfikacja kosztów błędnych klasyfikacji. W przypadku wybranej przez nas metody drzew klasyfikacyjnych parametr ten możemy określić na zakładce *Klasyfikacja* okna definiowania analizy.



Po wybraniu opcji *Użytkownika* w obszarze *Koszty błędnych klasyfikacji* możemy zmienić koszty pomyłek (które domyślnie są sobie równe) w taki sposób, aby pomyłki odnoszące się do klientów posiadających konto maklerskie były bardziej kosztowne. Przez to algorytm będzie miał mniejszą skłonność do przypisywania ich do grupy klientów, którzy tego konta nie posiadają. Po naciśnięciu przycisku w wyświetlonym oknie zmieniamy domyślny koszt błędnej klasyfikacji osób posiadających konto maklerskie z 1 do 15.



prawdopodobieństwo, że skorzystają z konta maklerskiego. Niestety, uzyskanej w ten sposób informacji nie będzie można wykorzystać podczas planowanej kampanii, ponieważ każdy z klientów korzystających z usługi *Private Banking* jest traktowany indywidualnie i nie są do niego wysyłane masowe oferty. Reguły klasyfikujące osoby, które nie korzystają z usługi *Private Banking* są bardziej skomplikowane. Cechami różnicującymi pozostałych klientów są między innymi: segment demograficzny, fakt posiadania lokaty oraz wielkość obrotów na rachunku.

Analizując wyniki modelu, możemy zauważyć, że ogólny błąd klasyfikacji jest większy od modelu początkowego i wynosi około 30%. Jednak wartość informacyjna tego modelu jest nieporównywalnie większa, ponieważ błąd rozkłada się bardziej równomiernie pomiędzy obiema klasami. Na podstawie tego modelu możemy podjąć już pewne kroki związane z optymalizacją wysyłania ofert

Oferowanie kilku produktów równocześnie

Przeprowadzanie kampanii polegającej na oferowaniu klientom tylko jednego produktu jest mało skomplikowane z organizacyjnego punktu widzenia – polega po prostu na zbudowaniu modelu wskazującego skłonność do zakupu tego produktu i wysłaniu ofert do osób wskazanych przez model.

Sytuacja staje się nieco trudniejsza w przypadku, gdy podczas kampanii rozważamy wysłanie do klientów oferty na produkt wybrany z grupy produktów. Jeśli chodzi o samą budowę modeli prognostycznych, to przebiega ona w sposób analogiczny do przykładu przedstawionego powyżej. Dla każdego produktu, jaki bierzemy pod uwagę podczas danej kampanii, budujemy po prostu osobny model wskazujący, czy klient jest skłonny kupić ten produkt czy też nie. Następnie na podstawie wyników wszystkich zbudowanych modeli wybieramy, jaki produkt będzie najbardziej odpowiedni dla danego klienta.

Wybór odpowiedniego produktu jest trywialny w przypadku, gdy tylko jeden ze zbudowanych modeli wskazuje, że klient skłonny jest skorzystać z oferowanego produktu. Sytuacja jest nieco bardziej skomplikowana w przypadku, gdy taką skłonność sugeruje więcej niż jeden model lub żaden z nich (jeśli zakładamy, że każdy klient ma otrzymać ofertę).

W takiej sytuacji, aby wybrać najodpowiedniejszą opcję, konieczne jest porównanie ze sobą odpowiedzi poszczególnych modeli. By móc to zrobić, powinniśmy zdefiniować pewną miarę, na podstawie której będziemy mogli porównywać wyniki dla poszczególnych produktów. W [1] przedstawiono trzy cechy, jakimi powinna się charakteryzować taka miara:

- ♦ ocena skłonności do zakupu w każdym modelu musi mieć taki sam zakres wartości,
- ♦ jeśli dana osoba już korzysta z danego produktu, powinna otrzymywać ocenę 0,
- ♦ miara powinna brać pod uwagę względną częstość korzystania z danego produktu wśród klientów.



Biorąc pod uwagę wymienione powyżej punkty, naturalną miarą pozwalającą porównywać ze sobą modele jest miara oparta na prawdopodobieństwie, że dana osoba będzie skłonna kupić dany produkt. Zaletą tej miary jest to, że ma stały zakres wartości (od 0 do 1), co spełnia pierwszy z powyższych wymogów. By spełnić kolejny wymóg, konieczna jest jedynie zamiana prawdopodobieństwa na wartość 0, jeśli dana osoba już korzysta z danego produktu. Jeśli chodzi o uwzględnienie częstości korzystania z danego produktu, to warunek ten jest spełniony w sposób naturalny w przypadku zachowania podczas budowy modelu pierwotnego rozkładu zmiennej zależnej. Tak więc, jeśli dany produkt jest stosunkowo mało popularny, to również oceny prawdopodobieństwa zakupu tego produktu nie będą wysokie. W przypadku gdy przed zbudowaniem modelu stosujemy technikę *oversamplingu*, zaburzającą naturalne proporcje osób, które korzystają i nie korzystają z danego produktu, konieczne jest przemnożenie uzyskanych prawdopodobieństw przez prawdopodobieństwa *a priori*, że dany klient będzie skłonny kupić dany produkt.

Przed wykonaniem powyższej operacji warto jeszcze zadać sobie pytanie, na ile proporcje korzystania z danego produktu występujące wśród naszych klientów są naturalne i rzeczywiście odpowiadają potencjałowi rynku, a nie są wynikiem przyjętej przez nas w przeszłości polityki.

W przypadku zbudowanego przez nas modelu prawdopodobieństwa poszczególnych klasyfikacji otrzymujemy, wybierając przycisk *Przewidywanie* na karcie *Przypadki* znajdującej się w oknie wyników analizy.

Dane: Przewidywane 1					
Przewidywane 1 (Cross_selling)					
Zmienna zależna: Brokerage					
Opcje: Jakościowa zależna, Drzewo numer 1, Próba testowa					
Obszew. wartość	Przewid. wartość	Prawdopodobieństwo 0	Prawdopodobieństwo 1	Końcowy węzeł	
5	0	0	0,982540	0,017460	7
11	0	0	0,994307	0,005693	21
13	0	0	0,996656	0,003344	47
14	0	0	0,996656	0,003344	47
16	0	0	0,996656	0,003344	47
35	0	0	0,996656	0,003344	47
36	0	0	0,996656	0,003344	47
40	0	0	0,996656	0,003344	47
42	0	0	0,996656	0,003344	47
2045	0	0	0,996656	0,003344	47
2048	0	0	0,996656	0,003344	47
2053	0	0	0,996656	0,003344	47
2055	0	0	0,996656	0,003344	47
2057	0	0	0,996656	0,003344	47
2059	1	1	0,000000	1,000000	3
2062	0	0	0,996590	0,003410	19
2068	0	0	0,996656	0,003344	47
2070	1	1	0,077000	0,923000	5

Oferty sprzedaży krzyżowej warto jest optymalizować pod kątem efektu finansowego, jaki mają przynieść. Samo prawdopodobieństwo przyjęcia oferty nie mówi nam, jaka jest oczekiwana korzyść finansowa wynikająca z jej przyjęcia. By uwzględnić tę kwestię



w typowaniu oferowanych produktów, należy każde prawdopodobieństwo zakupu danego produktu pomnożyć przez średni zysk, jaki bank uzyskuje dla tego produktu. Możliwe jest również stosowanie bardziej zaawansowanej optymalizacji, uwzględniającej na przykład oczekiwany dochód odpowiedni dla segmentu, w którym znajduje się dany klient.

Literatura

1. Berry M., Gordon L., *Mastering Data Mining. The Art and Science of Customer Relationship Management*, John Wiley & Sons, Inc, New York 2000.
2. Berson A., Smith S., Thearling K., *Building Data Mining Applications for CRM*, McGraw Hill, New York 2000.