

BUDOWA MODELI DLA POTRZEB ANALITYCZNYCH SYSTEMÓW CRM

*Mariusz Żytniewski
zyto@sulu.ae.katowice.pl
Akademia Ekonomiczna im. Karola Adamieckiego w Katowicach*

Streszczenie

Celem niniejszego referatu jest przedstawienie analiz stosowanych w systemach CRM, etapów budowy modelu analitycznego oraz określenie miejsca modeli analitycznych w procesie poprawy oddziaływania na klienta. Autor przedstawia metodę budowy modeli analitycznych pod nazwą CRISP-DM oraz jej umiejscowienie w zaproponowanym przez firmę Terradata modelu oddziaływania na klienta.

Słowa kluczowe: CRM, analityczny CRM, CRISP-DW, data mining.

Wstęp

Analityczne systemy CRM podlegają stałemu rozwojowi i wykorzystują coraz większe możliwości, jakie dają im analizy związane z odkrywaniem wiedzy w danych. Zastosowanie systemów data mining, czyli metod eksploracji danych w systemach CRM wiąże się z coraz szerszą integracją tych systemów w organizacji. Rozwój systemów analitycznych w obszarze wspomagania relacji z klientem prowadzi do zastosowania najnowszych technik analitycznych dla potrzeb tychże systemów. W niniejszym referacie przedstawiona zostanie koncepcja budowy modelu analitycznego wykorzystywana między innymi przez firmę SPSS. CRISP-DM jest podzieloną na etapy koncepcją budowy modeli analitycznych wykorzystywanych tak przez systemy CRM jak i data mining. Zanim jednak zajmiemy procesem budowy modeli prześledźmy, jakie systemy wykorzystywane są w analitycznych CRM.

Zastosowanie analitycznych systemów CRM

Systemy analityczne CRM za zadanie mają analizować dane dotyczące całej organizacji dla celów poprawy relacji z klientami. Pozwalają one dzięki swojemu potencjałowi przewidywać odpowiedzi klientów na reklamę firmy, zdefiniować, co jest przyczyną niskiego zadowolenia klientów oraz określać grupy klientów, którzy z dużym prawdopodobieństwem mogą opuścić firmę. Dzięki nim można także identyfikować klientów o podobnych cechach, oceniać wartość

klienta w czasie oraz tworzyć grupy produktów, które są kupowane w określonym czasie i w określonej kolejności.

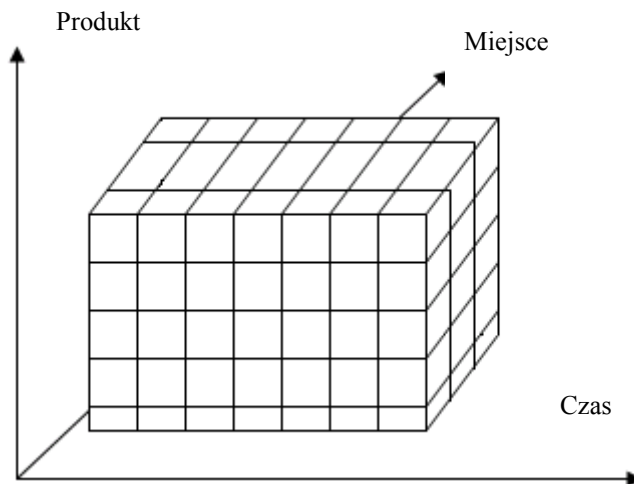
Podsumowując przedstawione tutaj zadania analitycznego CRM można wyszczególnić takie cele analiz jak [HnC**B**02], [JwPk**0**2], [Gr98]:

- Określanie podobieństwa i powiązań między produktami;
- Współ kupowanie produktów i usług np. analizy asocjacyjne;
- Analizowanie zależności czasowych np. sekwencje zakupów;
- Określanie wartości klienta (w tym LTV);
- Określanie zachowań klientów np. migracja klientów;
- Lojalność i satysfakcja klientów;
- Skuteczność akcji promocyjnych;
- Segmentacja klientów;
- Profilowanie klientów;
- Określanie ryzyka np. ryzyko kredytowe i ryzyko nadużyć.
- Analizy linii trendów

Systemy CRM, a dokładniej analityczne systemy CRM swój potencjał opierają na analizach związanych z drążeniem danych. Dzięki rozwojowi platformy sprzętowej, systemy te posiadają wystarczającą moc obliczeniową dla drążenia danych długookresowych, pochodzących z całej organizacji. Oprócz odkrywania wiedzy systemy tego typu stosują także różne metody raportowania i wielowymiarowych analiz danych OLAP.

OLAP (OnLine Analytic Processing)

Systemy OLAP pozwalają na wielowymiarowe analizy danych. Za ich pomocą możliwe staje się przeprowadzenie analiz wskazujących np. na wielkości sprzedaży produktów, na różnych terenach w pewnych przedziałach czasu. Za pomocą tego typu systemów istnieje możliwość tworzenia wielowymiarowych kostek reprezentujących analizowane dane (Rysunek 1).



Rysunek 1 Kostka OLAP

Źródło: Opracowania własne

Tego typu systemy w głównej mierze stosowane do tworzenia zapytań i generowanie raportów na podstawie danych dotyczących analizowanego zagadnienia. Systemy OLAP pozwalają na zagłębianie się w danych tzw. Drill-down, czyli coraz większe uszczegółowianie badanego zjawiska.

Data mining

Systemy odkrywania wiedzy nastawione są na odnajdywanie zależności w dużych ilościach danych. Stosowane analizy pozwalają już nie tylko analizować przeszłość jak miało to miejsce w przypadku OLAP, ale także wspomagają procesy predykcji.

Zastosowanie modeli predykcyjnych w konkretnym przypadku zmusza twórców modelu do zastanowienia się czy za jego pomocą będziemy starali się sklasyfikować dany podmiot np. czy dana osoba odpowie na naszą ofertę czy też nie, czy też może potrzebują oni odpowiedzi, jaki procent klientów zrezygnuje z naszych usług po wprowadzeniu zmian.

Do analiz stosowanych w tych systemach zaliczymy np.[Ss02]:

- Analizy klasyfikacyjne
- Drzewa decyzyjne
- Analizy asocjacji i sekwencji

- Sieci neuronowe

Najnowsze systemy stosujące powyższe analizy opierają się na interfejsie graficznym pozwalającym na przygotowywanie zaawansowanych modeli. Budowa modelu jednak jest procesem trudnym i wymagającym określenia pewnych etapów niezbędnych dla prawidłowego przygotowania modelu.

Budowa modeli na przykładzie CRISP-DM

Jedną z metod budowy modeli analitycznych jest CRISP-DW (Cross Industry Standard Process for Data mining).

Model referencyjny CRISP-DM składa się z 6 kolejnych kroków. [JkOI02], [SPSS03], [CKRSW00]:

1. Określenie celów biznesowych
2. Zrozumienie danych
3. Przygotowanie danych do modelowania
4. Budowa modelu
5. Ewolucja modelu
6. Rozpowszechnienie modelu i rezultatów analiz

1. Określenie celów biznesowych

Jest to pierwszy etap niezwiązany bezpośrednio z analizami, ale przygotowujący analityka do ich przeprowadzenia. W etapie tym określany jest cel budowy modelu. Bardzo ważne jest, aby osoby definiujące model zwracały uwagę, na to, iż model ten ma mieć zastosowanie biznesowe, dlatego też, jeżeli nie są pracownikami organizacji budowę jego powinny konsultować z osobami, które będą go w przyszłości eksploatowały. Określenie celów biznesowych ma za zadanie umożliwić także przygotowanie kryteriów oceny badanego modelu. Inne kryteria oceny będzie miał model, którego zadaniem będzie określenie grupy klientów, którzy potencjalnie mogą odejść do konkurencji, a inne kryteria będzie miał model przewidujący rezultaty kampanii marketingowej.

2. Zrozumienie danych

- **Budowa marketingowej bazy danych**

Dane zastosowane w modelu powinny być jak najbardziej aktualnymi danymi. Dodatkowo decydent powinien posiadać wiedzę merytoryczną na ich temat dotyczącą znaczenia tych danych dla organizacji. Problemy z danymi pojawiają się już na poziomie ich przesyłania między systemami. Dane w organizacji mo-

gą być różnie opisane i przechowywane w systemach słabo zintegrowanych, co dodatkowo utrudnia tworzenie efektywnych modeli. Systemy wewnątrz organizacji operować mogą na różnych bazach związanych z działalnością operacyjną oraz hurtowniami danych. Zebranie danych w spójną całość dla potrzeb analizowanego zjawiska często zajmuje od 50% do nawet 90% czasu (według danych SPSS).

- **Eksploracja danych**

Przyjrzenie się analizowanemu danym jest procesem żmudnym, ale bardzo potrzebnym. Prawidłowe rozpoznanie badanych danych podnosi skuteczność analiz jak i pozwala wyeliminować dane niepotrzebne jeszcze przed rozpoczęciem procesu analitycznego. Pomagają tutaj takie proste działania jak generowanie sum, średnich, odchyleń.

3. Przygotowanie danych do modelowania

Jest to ostatni etap związany z przygotowaniem danych. Podzielić go można na cztery podetapy:

- **Wybór zmiennych**

Jest to ważny etap związany z wydajnością procesu drążenia danych jak i jego „siłą” predykcji. Zbyt wielka ilość zmiennych w modelu doprowadzić może do problemów z wydajnością procesu analitycznego. Innym problemem może być wybór zmiennych, które zamiast poprawiać możliwości predykcyjne będą osłabiały wyniki.

- **Przygotowanie predykatów**

Czasami w procesie analitycznym lepszym i bardziej czytelnym jest zastosowanie predykatu będącego kompilacją kilku innych zmiennych. Przykładem może być zastosowanie w modelu wskaźnika zamiast podstawowych danych.

- **Wybór danych lub zbioru danych**

Wykorzystanie wszystkich danych wchodzących w skład zbudowanej bazy nie jest konieczne. Zastosowanie próby losowej dla potrzeb budowy modelu jest korzystniejsze, gdyż nie powoduje tak dużego obciążenia systemu mającego te dane przetwarzać. Dzięki zastosowaniu mniejszych grup danych możliwe staje się modelowanie większej grupy modeli relatywnie w tym samym czasie, co zbudowanie mniejszej ilości modeli, ale z większej grupy danych

- **Dostosowanie zmiennych do algorytmu**

Określenie parametrów budowanego modelu związane jest z wynikami, jakie otrzymamy. Określenie zbyt rygorystycznych wartości spowodować może, iż model nie odnajdzie żadnych zależności. Z drugiej strony określenie parametrów na zbyt niskim poziomie powoduje często problemy z odnalezieniem tych najbardziej cennych wyników w wygenerowanej odpowiedzi.

4. Budowa modelu drażenia danych

Jest procesem automatycznym związanym z zastosowaniem przygotowanych danych. Proces ten polega na nadzorowaniu działań systemu. Wyniki, które wskazuje system mogą posłużyć nam albo do przyjęcia modelu albo do powrotu do wcześniejszych kroków.

5. Ocena otrzymanych wyników

Jedną z miar skuteczności naszego modelu jest jego dokładność. Określa ona jak bardzo wyniki naszej analizy przystają do rzeczywistych danych. Innymi metodami wykazania prawidłowości działania modelu są dzwignia, osiągnięty zysk lub wskaźnik ROI.

Aby określić wartość modelu należy zadać sobie następujące pytania [JkOI02]:

- Czy model spełnia postawione przed nim cele biznesowe?
- Czy jakieś ważne cele biznesowe nie zostały zasygnalizowane?
- Czy zbudowany model ma sens z punktu widzenia eksperta?
- Jakie korzyści uzyskamy dzięki modelowi?
- Jak zmienią się wyniki analiz w przypadku zmiany procesów biznesowych, na których bazują rezultaty modelu?

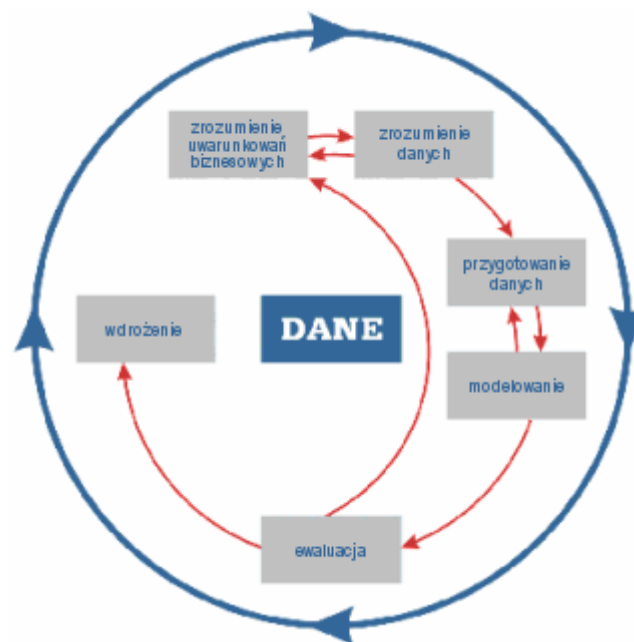
Odpowiedzenie sobie na powyższe pytania pozwala określić jak bardzo model odpowiada naszym oczekiwaniom.

6. Integracja drażenia danych z rozwiązaniami CRM

Analitik powinien być przygotowany na reakcje, jakie przyniesie wprowadzenie modelu do użytku. W całym toku życia modelu powinien on być monitorowany pod kątem nieprawidłowego zastosowania oraz zmian kryteriów doboru modelu do analiz. Raz stworzony model powinien ewoluować wraz ze zmianami wewnątrz organizacji.

Przedstawiany tutaj proces budowy nie zapewnia gwarancji stworzenia prawidłowego modelu analitycznego, ale wspomaga proces tworzenia, dzięki podziałowi kolejnych zadań na etapy.

Graficzny model procesu tworzenia modeli przedstawia rysunek 2.



Rysunek 2. Model CRISP-DM

Źródło: Materiały SPSS

Koncepcja CRISP-DM budowy modelu analitycznego jest początkiem procesu działania systemu CRM. Kolejne etapy oddziaływania na klienta odnaleźć można w modelu zaproponowanych przez firmę Terradata.

Jako podstawowe etapy pracy organizacji dla potrzeb CRM można tutaj wyróżnić [Terra02]:

- **Analizy**
 - analiza klientów
 - określanie docelowego klienta
 - szacowanie odpowiedzi
 - analiza przypadków

- **Modelowanie**
 - Budowa modelu
 - Wyniki

- **Personalizacja**

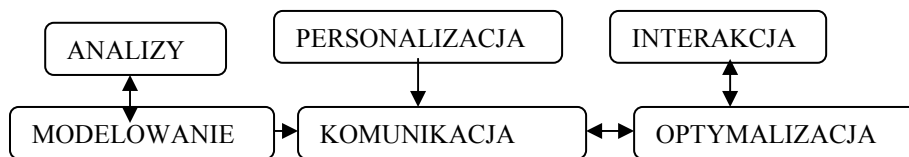
- generowanie personalizowanych ofert do klientów
- budowa zasad personalizacji

- **Komunikacja**
- planowanie dróg komunikacji
- definiowanie zasad działania oraz wyzwalaczy

- **Optymalizacja**
- określanie priorytetów i limitów w kanałach komunikacji
- określanie częstotliwości i jakości komunikacji

- **Interakcja**
- personalizacja czasu rzeczywistego
- zarządzanie oddziaływaniem na klienta

Graficzne powiązania między poszczególnymi etapami przedstawia rysunek 3



Rysunek 3. Model CRM Terradata

Źródło: Opracowania własne na podstawie „Teradata CRM Response Modeling”

Jak widać sam proces tworzenia modelu analitycznego, określanego tutaj przez pierwsze dwa kroki, jest jedynie wstępem do dalszych procesów wewnątrz organizacji. Jednak jako początkowy element łańcucha, jest najważniejszy. Popętnienie błędów w tym etapie spowoduje, iż niezależnie od powodzenia dalszych kroków cały proces oddziaływania na klientów zakończy się niepowodzeniem.

Podsumowanie

Zastosowanie nowoczesnych systemów informatycznych w organizacji wiąże się z coraz większym upowszechnieniem wiedzy i nowymi możliwościami jej odnajdywania. Przedstawione tutaj informacje związane z zastosowaniem analitycznych systemów CRM oraz etapami budowy modeli dla potrzeb tych systemów wskazują ich dużą rolę w procesie funkcjonowania przedsiębiorstwa. Każde działania w obrębie organizacji można podzielić na etapy. Tak też jest w przypadku systemu oddziaływania na klienta. Określenie błędów już na początku pracy powoduje zmniejszenie prawdopodobieństwa niepowodzenia całego procesu. Z drugiej strony nie wykryte błędy w początkowej fazie powielane są w dalszych etapach.

Nie należy zapominać, iż zastosowanie odpowiednich metod analitycznych w procesie analizowania danych wiąże się jednoznacznie z wiedzą osób analizujących dane. Jak w przypadku systemów ekspertowych wiedza eksperta była niezbędna dla zbudowania bazy wiedzy, tak tutaj jest on niezbędny dla określenia czy wiedza odkryta przez system jest ważna z punktu widzenia organizacji. Wiedza analityka pozwala mu także na odpowiednie przygotowanie modelu analizującego dane. To właśnie do tego typu pracowników kierowana jest koncepcja CRISP-DM

Nowością na rynku jest oprogramowanie przeznaczone do budowy baz modeli dla potrzeb systemów analitycznych. Koncepcja takiego systemów przedstawiona między innymi [SsMz03] pokazuje nowe możliwości systemów przeznaczonych już nie tylko dla wysoko wykwalifikowanych pracowników działu analiz, ale także dla pracowników innych komórek organizacji.

Literatura:

- [JkOI02] James King; Orin Linden “Data mining isn't a 'cookbook' activity” National Underwriter; Erlanger; Sep 30, 2002
- [SPSS03] „Budowanie korzystnych relacji klientami” SPSS Ploska
- [HnC02] Hamid R Nemati, Christopher D Barko “Enhancing enterprise decisions through organizational data mining” The Journal of Computer Information Systems, Summer 2002.
- [JwPk02] Janusz Wachnicki, Piotr Komornicki „Modern Marketing” numer 2/2002
- [Gr98] Groth, R. “Data Mining: A Hands-on Approach for Business Professionals”. Prentice Hall, 1998
- [SsMz03] S. Stanek, M.Zytniewski “Concept of the Self-educating CRM System with a Base of Models” ISDSS 2003
- [CKRSW00] Pete Chapman (NCR), Julian Clinton (SPSS), Randy Kerber (NCR), Thomas Khabaza (SPSS), Thomas Reinartz (Daimler-

Chrysler), Colin Shearer (SPSS) and Rüdiger Wirth (Daimler-Chrysler) "CRISP-DM 1.0 Step-by-step data mining guide" SPSS 2000

[Ss02] Sanjay Saxena "CRM Analytics – KEY to ROI on CRM" www.realmarket.com 2002

[Terra02] "Teradata CRM Response Modeling" www.teradatalibrary.com

Abstract

The main goal of this article is to show analysis using in analytics CRM, stages of constructing the base of models and describe model's place in customer relationship management process. Author shows CRISP-CM as the method of building analytical models and place for it in proposed by Terradata "Response Modeling".