

Małgorzata Śniegocka-Łusiewicz

Analiza koszykowa w badaniach marketingowych

Słowa kluczowe: *analiza koszykowa, data mining, tabele kontyngencji, reguły asocjacji, analiza sekwencji, analiza skojarzeń*

Abstrakt: Analiza koszykowa, pozwala na analizę zawartości koszyka klienta. Analiza koszykowa polega na rozpoznaniu reguł, którymi kierują się klienci przy zapelnianiu „koszyka”, zwyczajów danego klienta, prawidłowości w korzystaniu z usług danego typu, badaniu, jakie produkty kupowane są razem lub w określonej sekwencji. Rozpoznanie tych reguł może mieć ogromny wpływ na zwiększenie wartości klienta.

Analiza koszykowa pozwala na zdobycie wiedzy o zwyczajach i upodobaniach klienta, dzięki czemu jest przydatna przy planowaniu kampanii promocyjnych, weryfikacji efektywności i skuteczności tychże kampanii, skutecznym oferowaniu dodatkowych usług i optymalizacji pakietów usług, taryf i opłat, trafnej rekomendacji zakupów kolejnego produktu, zwiększaniu sprzedaży oraz do przeciwdziałania rezygnacji klientów z usług.

Wprowadzenie

Dlaczego idąc do sklepu po jeden produkt wracamy z siatką pełną zakupów? Czy jest to tylko nasza spontaniczna decyzja włożyć coś jeszcze do koszyka? Jednym z podstawowych zadań marketingu jest skierowanie reklamy do odpowiedniej grupy odbiorców, uświadomienie potrzeby zakupu produktu potencjalnym klientom, nakłonienie do nabycia produktu i utrzymanie jako stałego klienta. Jeżeli zasugerujemy klientowi zakup kolejnego produktu, to zakupi on go tylko w przypadku „trafienia” w gust, potrzebę, aktualny nastrój. Ale jak trafić? Tym właśnie zagadnieniem zajmuje się analiza koszykowa. Celem artykułu jest przybliżenie zagadnienia analizy koszykowej i pokrewnych, i wyłożenie jej ogromnej przydatności w marketingu

praktycznym. Współczesnym i ciągłym zjawiskiem w gospodarce jest potrzeba zwiększenia przychodów firmy przy jak najmniejszym zwiększeniu wydatków, można to osiągnąć poprzez zwiększenie: liczby klientów, wartości klienta i wydłużenie czasu pozostawania klientem¹.

Analizę koszykową, w sposób uproszczony, możemy nazwać analizę wartości koszyka klienta. Przychodząc do sklepu (np. supermarketu, sklepu internetowego, punktu usługowego) ma on do wyboru wiele różnego typu produktów. Do swojego „koszyka” klient może włożyć to, co chce, w jakiej kolejności chce i ile chce. Analiza koszykowa polega na rozpoznaniu reguł, którymi kierują się klienci przy zapełnianiu „koszyka”, zwyczajów danego klienta, prawidłowości w korzystaniu z usług danego typu, badaniu, jakie produkty kupowane są razem lub w określonej sekwencji. Rozpoznanie tych reguł może mieć ogromny wpływ na zwiększenie wartości klienta².

Metody analizy koszykowej

Analiza koszykowa należy do zagadnień *Data Mining*³. Eksploracja danych (*Data Mining*) jest „analizą (często ogromnych) zbiorów danych obserwa-

cyjnych w celu znalezienia nieoczekiwanych związków i podsumowania danych w oryginalny sposób, tak aby były zarówno zrozumiałe, jak i przydatne dla ich właściciela”⁴. Innymi słowy „eksploracja danych jest procesem odkrywania znaczących nowych powiązań, wzorców i trendów przez przeszukiwanie dużych ilości danych zgromadzonych w skarbnicach danych, przy wykorzystaniu metod rozpoznawania wzorców, jak również metod statystycznych i matematycznych”⁵.

Podsumowanie zrozumiałe i przydatne musi nieść z sobą możliwość zrozumienia i uporządkowania danego zbioru danych, wykrycia praw i reguł nim kierujących. Reguły takie nazywamy **regułami asocjacji**⁶. Są to zasady, prawa, którymi kierują się klienci biorący produkty do „koszyka”. Reguły te dotyczą nabywania łącznego dwóch lub większej liczby produktów. Część

¹ J. Wachnicki, P. Komornicki, *Data mining dla marketera*, „Modern Marketing” 6/2001.

² Zobacz również: K. Jajuga, M. Walesiak, *Taksonomia 15. Klasyfikacja i analiza danych – teoria i zastosowania*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2008, s. 527–534.

³ W polskiej literaturze używa się określenia eksploracja danych, wybieranie danych, wyszukiwanie danych.

⁴ D. Hand, H. Mannila, P. Smyth, *Eksploracja danych*, Wyd. Naukowo-Techniczne, Warszawa 2005, s. 35.

⁵ The Gartner Group, www.gartner.com.

⁶ D. T. Larose, *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2006, s. 185.

reguł jest łatwa do wychwycenia i dość oczywista (np. oprócz chleba kupowane jest masło), ale najważniejsze są te reguły z pozoru nielogiczne i nieoczywiste, których znajomość pozwoli na skuteczne sterowanie sprzedażą.

Kolejną metodą analizy koszykowej jest tabela, **tablica kontyngencji** (inaczej tablica wielodzielcza, tablica krzyżowa). „Jest to tabela dwuwymiarowa, w której wiersze przedstawiają kategorie jednej zmiennej, a kolumny kategorie drugiej zmiennej”⁷. Przykład tablicy kontyngencji przedstawia tabela 1.

Tabela 1. Przykład tabeli kontyngencji – liczebności 4 grup produktowych w diecie 5 grup wiekowych uczniów i studentów

Grupa wiekowa	Przedszkolak	Uczeń szkoły podstawowej	Uczeń gimnazjum	Uczeń szkoły ponadgimnazjalnej	Student	Suma
Grupa produktów						
Owoce	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	$n_{1.}$
Nabiał	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}	$n_{2.}$
Mięso	n_{31}	n_{32}	n_{33}	n_{34}	n_{35}	$n_{3.}$
Słodycze	n_{41}	n_{42}	n_{43}	n_{44}	n_{45}	$n_{4.}$
Suma	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	$n_{.5}$	N

Źródło: opracowanie własne.

Do powyższej analizy możemy wykorzystać m.in. takie narzędzia, jak⁸:

– **test niezależności χ^2** – w celu weryfikacji hipotezy o niezależności zmiennych X i Y ;

gdzie: n_{ij} – liczba jednostek, które mają jednocześnie wariant x_i zmiennej X i y_j zmiennej Y , liczebności teoretyczne z tablicy niezależności, które spełniają warunek niezależności zmiennych losowych X i Y , $i = 1, 2, \dots, k$; $j = 1, 2, \dots, r$;

– **współczynnik Czuprowa** służący do mierzenia siły zależności między zmiennymi X i Y (budowany na podstawie statystyki χ^2),

gdzie: χ^2 – wartość statystyki (1), n – liczba elementów próby, k – liczba wierszy tablicy niezależności, r – liczba kolumn tablicy niezależności;

– **współczynnik Φ Yule’a** mierzy siłę związku pomiędzy dwiema zmiennymi mierzonymi na skalach niemetrycznych nominalnych,

gdzie: χ^2 – wartość statystyki (1), n – liczba elementów próby.

⁷ M. Lasek, *Data Mining: zastosowania w analizach i ocenach klientów bankowych*, Biblioteka Menedżera i Bankowca, Warszawa 2002, s. 17.

⁸ M. Piłatowska, *Repetitorium ze statystyki*, Wydawnictwo Naukowe PWN, Warszawa 2006.

Inne przydatne narzędzia do analizy to m.in.: współczynnik korelacji C-Pearsona oraz współczynnik V Cramera.

Analiza sekwencji⁹ polega na badaniu kolejności dokonywanych zakupów, kolejności wkładania produktów do „koszyka”. Pozwala ona na uwzględnienie powiązań zdarzeń w czasie. Nie bada ona tylko zwykłej kolejności, ale również siłę i kierunek zależności pomiędzy odpowiednimi produktami. Może się zdarzyć, że zakup jednego produktu automatycznie oznacza brak zakupu innego – produkty się wykluczają (np. jeżeli klient kupi masło, to nie kupi już margaryny, albo jeżeli kupi podpaski, to nie kupi wkładek higienicznych) lub zdecydowanie zmniejszają prawdopodobieństwo zakupu (raz na trzy zakupione masła klient kupuje margarynę, ponieważ jeden z członków rodziny nie jada masła). Analiza sekwencji pozwala między innymi na odpowiednie rozplanowanie produktów w sklepie na podstawie kolejności robienia zakupów tak, bo zachęcić klienta do jak największych zakupów i by uświadomić mu wiele potrzeb, o których istnieniu nie wiedział.

Badając analizę koszykową, nie sposób nie poruszyć zagadnienia **analizy skojarzeń**. „Termin **analiza skojarzeń** pojawia się wtedy, gdy techniki wydobywania z danych, sekwencyjnych czy niesekwencyjnych reguł asocjacji stosowane, są by zapanować nad tym, co właściwie wiemy, czym dysponujemy, ale co ukryte jest w skomplikowanych strukturach bazy danych. Zachowania pojedynczych klientów podlegają opisowi w kategoriach **koszyka zakupów** i kolejnych **transakcji**, gdy zakłada on konto w banku, wymienia informacje z innym, konkretnym klientem i wykonuje różne inne, temu podobne działania. [...]”

Technologia zastosowana do tego typu baz danych pozwala szybko wydobyć na jaw ważne struktury i skojarzenia dotyczące klientów i ich działań¹⁰. Zagadnienie to ma ogromne znaczenie przy analizach cech niemierzalnych. Pozwoli na zwiększenie efektywności marketingu bezpośredniego, ponieważ będziemy kierowali odpowiednią ofertę do właściwego klienta. Cechą charakterystyczną analizy sekwencji jest potrzeba podejścia nie maszynowego, a bardziej inteligentnego, nie czysto matematycznego, a bardziej „ludzkiego”, by zrozumieć zachowanie klienta.

Do powyższych analiz możemy wykorzystać m.in.:

– **korelację rang Spearmana** (inaczej korelacja rangowa Spearmana) – służy do obliczania zależności dla cech jakościowych, które mają charakter cech porządkowych. „Miernik ten służy do mierzenia korelacji rang dwóch zmiennych. Nie jest konieczne dysponowanie wartościami cechy X i cechy

⁹ Zobacz również: P. Lula, G. Paliwoda-Pękosz, R. Tadeusiewicz, *Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2007, s. 123.

¹⁰ <http://www.statsoft.pl/products/sal.htm> 27.07.2008.

Y, wystarczy znać ich rangi (pozycje), co często zdarza się w analizach zjawisk społecznych i ekonomicznych. Zaletą tej miary jest fakt, że może być ona stosowana do mierzenia siły korelacji zmiennych mierzalnych jak i porządkowych¹¹,

gdzie: $d_i = a_i - b_i$, a_i , b_i – rangi odpowiadających sobie wartości cech x_i i y_i , $i = 1, 2, \dots, n$,

– współczynnik korelacji rang Kendalla¹², który jest unormowaną miarą korelacji,

przy czym:

$N(N-1)/2$ – maksymalna nota, gdy wszystkie rangi są rosnące,

R – całkowita nota, jaką otrzymujemy na podstawie konkretnego uporządkowania,

k_i – liczba rang y_i , dla których jednocześnie $j > i$ oraz $y_j > y_i$.

Obszary wykorzystania analizy koszykowej

Analiza koszykowa pozwala odpowiedzieć na szereg pytań dotyczących klienta, np. jakie produkty kupowane są najczęściej razem, z jakim prawdopodobieństwem klienci po zakupie produktu A zakupią produkt B, zagadnienie zakupów w ramach różnego typu promocji i programów lojalnościowych. Programy lojalnościowe i akcje promocyjne (np. poprzez rabaty, karty kredytowe danego sklepu; zbieranie punktów w systemie) pozwalają na przypisanie klientowi wszystkich produktów, jakie zakupił (łącznie z ilością, częstotliwością zakupów), pozwalają na przewidzenie, kiedy znów zjawi się w sklepie i jakich produktów będzie poszukiwał. Przekazuje do systemu informacje o współkupowanych produktach, natomiast w przypadku usług można również zaobserwować sekwencje tychże usług. Pamiętajmy również o tym, że przed odbiorem karty kredytowej sklepu, karty do programu lojalnościowego wypełniamy wniosek, który odpowiednio zredagowany da analitykowi pełną wiedzę na temat cech demograficznych danego klienta. Pozwoli to nie tylko na przygotowanie charakterystyki klienta, który zjawia się w sklepie pierwszy raz, ale na takie dostosowanie oferty, by został stałym klientem sklepu, ale również pozwoli na opracowanie oferty standardowej dla osoby o odpowiednich cechach demograficznych.

Analiza koszykowa może być wykorzystana m.in. do planowania kampanii promocyjnych, weryfikacji efektywności i skuteczności tychże kampanii,

¹¹ W. Starzyńska, *Statystyka praktyczna*, Wydawnictwo Naukowe PWN SA, Warszawa 2000, s. 326.

¹² J. Steczkowski, A. Zeliaś, *Statystyczne metody analizy cech jakościowych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1981, s. 169.

skutecznego oferowania dodatkowych usług i optymalizacji pakietów usług, taryf i opłat (np. banki, firmy ubezpieczeniowe, telekomunikacyjne); trafnej rekomendacji zakupów kolejnego produktu (np. sklepy internetowe), zwiększenia sprzedaży (np. supermarkety i oferty specjalne) oraz do przeciwdziałania rezygnacji klientów z usług.

Analiza koszykowa pozwala rozpoznać wiele zagadnień związanych z odpowiednim planowaniem sprzedaży i promocji w sklepach typu hipermarket i nie tylko. Jednym z podstawowych zagadnień jest problem, jakie produkty kupowane są najczęściej razem? W jakich okolicznościach zewnętrznych? Badanie stopnia prawdopodobieństwa, z jakim klienci, którzy kupili produkt A, kupią następnie produkt B. Analizując paragony można wyznaczyć produkty, które są ze sobą w jakimś stopniu skorelowane. Podstawą zwiększenia sprzedaży jest nie tylko odkrycie par oraz grup tych produktów, ale również stopnia skorelowania, czyli stopnia prawdopodobieństwa, z jakim klient kupujący produkt A i B sięgnie również po produkt C. Znając takie zależności, można wpłynąć na sprzedaż samym tylko układem produktów w sklepie. Przy wysokim stopniu skorelowania może wystarczyć ustawienie produktu C pomiędzy produktami A i B. Obecność produktu, który inni klienci kupowali często jako towarzyszący do produktu A i B, po pierwsze, uświadomi klientowi, iż chciałby on taki produkt zakupić, ale również dostarczy od razu usprawiedliwienia w postaci czynnika ekonomicznego, tzn. „skoro już tu jestem, to kupię, nie wydam drugi raz na benzynę/bilet”. Dodatkowym bodźcem ekonomicznym może być odpowiednio zastosowana promocja.

Najważniejsze w dobrze przygotowanej promocji jest rozpoznanie produktu bazowego – tego – po który klient przyszedł, nawet jeżeli nie jest tego świadomy, zjawiając się w sklepie. Znając produkt bazowy (lub produkty bazowe), należy odpowiednio zapromować produkt skorelowany. Istnieje możliwość zastosowania strategii cenowej, która zachęci klienta do zakupu „jeszcze tylko tego”, np. poprzez zmniejszenie ceny albo dodawanie większej ilości produktu w tej samej cenie. Innym sposobem jest bardzo atrakcyjne zaprezentowanie produktu.

Zupełnie odmienną strategią marketingową jest skupienie się na produkcie bazowym. Znając produkt skorelowany, można tworzyć odpowiednie zestawy lub promocje polegające na dodawaniu produktu skorelowanego do produktu bazowego jako gratis. Może to spowodować, że pomimo wyższej ceny lub słabszej marki produkt będzie wybierany przed produktami konkurencyjnymi.

Niezwykle przydatną funkcją, jaką spełnia wyliczony wskaźnik asocjacji, jest odpowiednia propozycja dla klienta, który robi zakupy w sklepie internetowym. Poniżej znajduje się przykład strony z jednego ze sklepów internetowych. Po wybraniu odpowiedniego produktu u dołu strony znajdujemy

informacje o tym, jakie produkty zakupili klienci, którzy wybrali szukany przez nas produkt. Ma to na celu głównie uświadomienie nowej potrzeby u klienta (np. zakupienie pozostałych elementów zestawu), który zdecydował się lub jest skłonny do zakupu jednego z produktów. Zwłaszcza w sklepach internetowych w takich sytuacjach działa czynnik psychologiczny i ekonomiczny, tzn. klient jest bardziej skłonny do zakupu, ponieważ, po pierwsze, uświadamia sobie po raz pierwszy, że oprócz szukanego produktu ma potrzebę kupić inny produkt, drugą kwestią jest podejście ekonomiczne – płaci się za przesyłkę tylko raz. Oddzielną sprawą jest wyliczenie pozornie nieistniejących zależności. O ile można przewidzieć, że osoba kupująca doniczkę najprawdopodobniej zakupi pasujący do niej podstawek, o tyle już zależności wynikające z mody albo z typowych cech charakteru są bardzo trudne do odgadnięcia. Możemy posłużyć się wymyślonym na potrzeby tej pracy przykładem. Kupujące przez Internet sukienki balowe kobiety kupują również satynowe zestawy pościeli. Produkty te pozornie nie mają ze sobą nic wspólnego, jednakże po analizie okazuje się, że sukienki kupowane są na ślub koleżanki/kolegi, a zestaw pościeli jest prezentem dla młodej pary. Właśnie takie niezauważalne na pierwszy rzut oka zależności są kluczem do zwiększenia sprzedaży.

Zależności te można wyliczyć na podstawie przeszłych zdarzeń i rozszerzyć na pozostałych nabywców. Sprawnie działający mechanizm wyznaczający produkty kupowane razem z poszukiwanym produktem może wielokrotnie popyt w samym sklepie internetowym (porównaj rysunek 1).

Rysunek. Prezentacja innych produktów, które wybierali klienci kupujący książkę *Ścigany*

Ścigany [brak opinii]

Autor: **Simon Kerrick**
 Region: **Kraj: polski**
 Opinie: **Miejska**
 Data wydania: **2008-07-25**

Cena rynkowa: **27,90 zł** Cena empik.com: **25,99 zł**
 Czas dostawy: **wysyłamy w 24 godziny**

Kup używane: Brak produktów używanych.

O produkcie: [Recenzje >](#) [Szczegółowe dane >](#)

Jeśli chcesz sprzedać taki produkt, kliknij:
[Sprzedaj taki produkt >](#)

Dodaj do schowka >
Powiadom znajomego >

Opis książki: Ścigany
 Ciepły wiosenny powiew! Harlan Coben napisał: "Kerrick pisze z nogą przyciśniętą do pedału gazu. Trzymaj się mocno". Jeden telefon od przyjaciela powoduje, że spokojnie popołudnie Toma Mertona, sprzyjający opłogi omawiana, przyładowego ojca rodziny, zamienia się w szur realistyczny koszmarny. Jack Casey, prawnik, z którym nie miał kontaktu od lat, przeryszy biega go o pomoc. Zanim rozmowa zostanie przerwana, Tom słyszy w słuchawce, jak niezmiernie przestawiona mordują jego przyjaciela. Słyszysz też coś jeszcze, co zmusza go do ucieczki.

Klienci, którzy kupili ten produkt, wybierali również:

Książka Kompleks winy David Morrell	Książka Dowody zbrodni Philip Margolin	Książka W głębi lasu Harlan Coben	Książka Kiedy nęka tama Jonathan Kellerman	Książka Królestwo piekła Sven Hassel	Książka Znak Jason Pinter

Źródło: <http://www.empik.com/scigany-ksiazka,prod2270218,p> [03.10.2008].

Zakończenie

Zwiększenie przychodów firmy przy jak najmniejszym zwiększeniu wydatków można osiągnąć poprzez zwiększenie: liczby klientów, wartości klienta i wydłużenie czasu pozostawania klientem. Aby osiągnąć powyższy cel, niezbędna jest wiedza o kliencie, taka jak: jakie produkty kupuje on najczęściej razem, z jakim prawdopodobieństwem klient po zakupie produktu A zakupi produkt B, co jest skłonny zakupić w ramach promocji i programów lojalnościowych, w jakiej kolejności kupuje produkty. Analiza koszykowa pozwala na zdobycie takiej wiedzy i dzięki temu jest przydatna przy planowania kampanii promocyjnych, weryfikacji efektywności i skuteczności tychże kampanii, skutecznym oferowaniu dodatkowych usług i optymalizacji pakietów usług, taryf i opłat, trafnej rekomendacji zakupów kolejnego produktu, zwiększaniu sprzedaży oraz do przeciwdziałania rezygnacji klientów z usług.

Następnym etapem studiów nad analizą koszykową będzie zastosowanie przedstawionych metod w badaniach empirycznych i wykazanie ich przydatności w analizie koszykowej.

Literatura

- Hand D., Mannila H., Smyth P., *Eksploracja danych*, Wydawnictwo Naukowo-Techniczne, Warszawa 2005.
- Jajuga K., Walesiak M., *Taksonomia 15. Klasyfikacja i analiza danych – teoria i zastosowania*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2008.
- Kurzawa I., Wysocki F., *Wykorzystanie analizy koszykowej do identyfikacji zachowań konsumpcyjnych gospodarstw domowych w Polsce*, [w:] *Taksonomia 15. Klasyfikacja i analiza danych – teoria i zastosowania*, pod red. K. Jajugi, M. Walesiaka, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2008.
- Larose D. T., *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Lasek M., *Data Mining: zastosowania w analizach i ocenach klientów bankowych*, Biblioteka Menedżera i Bankowca, Warszawa 2002.
- Lula P., Paliwoda-Pękosz G., Tadeusiewicz R., *Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2007.
- Piłatowska M., *Repetytorium ze statystyki*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Starzyńska W., *Statystyka praktyczna*, Wydawnictwo Naukowe PWN SA, Warszawa 2000.
- Steczkowski J., Zeliś A., *Statystyczne metody analizy cech jakościowych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa 1981.
- Wachnicki J., Komornicki P., *Data mining dla marketera*, „Modern Marketing” 6/2001.
www.gartner.com
www.statsoft.pl

Market Basket Analysis in Marketing Research

Summary

Market basket analysis is, to put it simply, analysis of the contents of a customer's shopping trolley. Market Basket Analysis is based on the identification of rules that customers apply while filling their trolleys, their habits, proper use of services, and research regarding products purchased together or in a defined order. Recognition of the aforementioned rules may have a great influence on customer's value increase. Basket analysis may be applied in order to plan promotional campaigns, verify their effectiveness, offer additional services and optimize service packages, tariffs and fees (e.g. for banks, insurance companies, telephone operators, etc.); receive accurate recommendation concerning other products to buy (e.g. online shops), increase sales (e.g. supermarkets, special offers) and prevent the companies providing services from losing their customers.

