

Andrzej Kluza

Katedra Ekonometrii i Informatyki SGGW

Zastosowanie sztucznej inteligencji do wyboru produktów bankowych z użyciem bazy przypadków

Wstęp

Świadczenie usług w zakresie doradztwa finansowego stanowi na rozwiniętych rynkach jedną z wyspecjalizowanych dziedzin bankowości inwestycyjnej. Wielość produktów bankowych oferowanych na rynku jest przyczyną obszernej podaży informacji na ich temat oraz publikowania dużej liczby reklam. Produkty bankowe znalazły w sieci Internet bardzo dogodnie w formie medium informacyjne. Portale takie jak bankier.pl, money.pl, biznes.onet.pl, expander.pl, pb.pl (Puls Biznesu), www.wabank.pl informują o kontaktach, udzielanych kredytach i innych produktach, często wraz z obszernym wsparciem przyszłego klienta za pomocą informatycznego systemu doradczego. Również bezpośredni twórcy produktów bankowych i inwestycyjnych oferują szczegółowe wyjaśnienia dotyczące własnych produktów, wraz z porównaniami i kalkulacjami opłacalności, np. cu.com.pl (Commercial Union), mbank.com.pl. Usługi te dają możliwość znajdowania ofert według ustalonych przez odwiedzającego portal kryteriów. Pokrywa się to z tendencjami coraz większego informatycznego wspierania całego cyklu sprzedaży i wykorzystania sieci internetowej dla celów komercyjnych [8].

Jednym z nowoczesnych sposobów implementacji funkcji doradczych w środowisku systemu informacyjnego jest wykorzystanie metody wnioskowania opartego na bazie przypadków (Case-Based Reasoning, CBR) [3, 6]. Metoda ta należy do grupy metod sztucznej inteligencji. W odróżnieniu od metod grupujących wiedzę w bazach reguł, w metodzie wnioskowania opartego o bazę przypadków, wiedza jest podana *explicite* w zbiorze przypadków. Pracujący na takiej zasadzie system ma możliwość informowania rozwiązującego problem o tym, jak osiągnąć określony cel na podstawie doświadczeń z przeszłości.

Zastosowania wnioskowania opartego na bazie przypadków można podzielić za Kolodner [6] na:

- analityczne – zadania klasyfikacji obiektów, diagnozy i doradztwa; zadania prognozowania przebiegów czasowych, zadania oceny ryzyka i kosztów,
- syntetyczne – zadania projektowania i tworzenia konfiguracji na podstawie wyboru komponentów i planowanie przedsięwzięć,
- organizacyjne – organizacja procesów zarządzania oraz zadania z obszaru zarządzania wiedzą.

Na stronach internetowych AI-KBG Uniwersytetu Kaiserslautern [2] oraz firmy Kaidara Software [5] opisane są liczne przykłady wykorzystania systemów doradztwa opartego na metodzie CBR w zakresie:

- wsparcia i doradztwa produktów i rozwiązań sieciowych firmy 3COM,
- polepszenia wyników w naprawach samochodów – PSA Citroen Peugeot,
- ulepszonej analizy błędów produktów półprzewodnikowych – National Semiconductor,
- ulepszonej analizy błędów chipowych kart telekomunikacyjnych – MET Ericsson,
- zmniejszenia czasu naziemnego samolotów pasażerskich przez przyspieszenie wykrywania usterek w silnikach – CFM International/Snecma Services.

Z innych firm korzystających w swej działalności z doradztwa w technologii CBR można wymienić: Ruhrgas, Cisco, NEC Computers, SNCF, Legrand, Dailmer-Chrysler, General Motors, by podać tylko bardziej znane.

Handel elektroniczny (Electronic Commerce, e-commerce) jest coraz częstszym i jednym z najbardziej obiecujących obszarów zastosowań metody CBR [8, 4]. Przykładami zastosowania metod wnioskowania opartego na bazie przypadków w dziedzinie handlu elektronicznego mogą być firmy: Phillips (przeszukiwanie bazy produktów sprzętu AGD), Analog Devices (wyszukiwanie w bazie komponentów elektronicznych). Warto również wspomnieć o systemie inteligentnego wsparcia dla aplikacji handlu elektronicznego WEBSELL opartym na technologii CBR, opracowanym na Uniwersytecie Kaiserslautern [3].

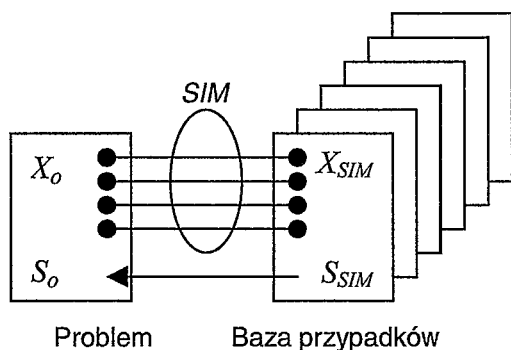
Przedstawiony w pracy system wnioskowania opartego na bazie przypadków przeznaczony dla doradztwa w zakresie produktów bankowych stanowi pierwszy znany autorowi przykład wykorzystania metody CBR w dziedzinie bankowości. Zadaniem aplikacji stworzonej przez autora jest wybór z rynkowej oferty bankowych rachunków oszczędnościowo-rozliczeniowych propozycji najbardziej pasujących do preferencji potencjalnego klienta, chcącego założyć konto w banku. Na podstawie tego przykładu udało się również przedstawić podstawowe cechy metody wnioskowania opartego na bazie przypadków.

Zastosowana metoda

Wnioskowanie oparte na bazie przypadków posługuje się przeszukiwaniem tej bazy. Warunkami koniecznymi zastosowania metody są: niezmiennosc reguł dotyczących przypadków, możliwość powtarzania się zjawisk oraz ciągłość części wartości cech obiektów. Przypadkiem nazywamy zgromadzoną wiedzę związaną z pewnym zaistniałym kiedyś doświadczeniem oraz opis postępowania, które może stać się wskazówką dla rozwiązującego problem, jak może on osiągnąć swój cel [6].

Rysunek 1 przedstawia sposób działania metody wnioskowania opartego na bazie przypadków. Zgromadzone dane o przeszłych doświadczeniach, w postaci pliku kart porównywane są z pojedynczą kartą umieszczoną po stronie lewej, oznaczającą doświadczenie bieżące, inaczej nazywane problemem. Proces porównywania nazywany jest zapytaniem (inaczej kwerendą).

Do wyznaczania zapisów najbardziej podobnych do bieżącego zagadnienia wykorzystywana jest tzw. funkcja podobieństwa, którą można traktować jako odwrotność odległości opartej na pewnej metryce [3, 6, 7]. Metoda dopuszcza zarówno minimalizację funkcji odległości, jak i maksymalizację funkcji podobieństwa. Po znalezieniu podobnego zapisu z przeszłości poznajemy również rozwiązanie problemu, które było w przeszłości pomyślnie zastosowane. Ostatnim krokiem jest próba zaadaptowania dawnego rozwiązania S_{SIM} do rozwiązania problemu bieżącego S_o , co na rysunku oznaczono strzałką skierowaną w lewą stronę.



Rysunek 1

Schemat działania metody wnioskowania opartego na bazie przypadków. X_o – wartości cech przypadku zawarte w zapytaniu, X_{SIM} – wartości cech najbardziej podobnego przypadku, SIM – funkcja podobieństwa globalnego, S_{SIM} – rozwiązanie pochodzące od najbardziej podobnego przypadku, S_o – rozwiązanie problemu otrzymane przez zaadaptowanie S_{SIM}

Źródło: opracowanie własne na podstawie [6].

Wektor wartości cech zapytania przedstawia się następująco:

$$X_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,n}),$$

gdzie: n – liczba wszystkich cech produktu,

$x_{0,j}$ – pojedyncza pożądana wartość j -tej cechy, $j = 1, 2, \dots, n$.

W bazie k przypadków $x_{i,j}$ oznacza j -tą cechę i -tego produktu z bazy. Funkcja podobieństwa lokalnego j -tej cechy $x_{0,j}$ zapytania i j -tej cechy $x_{i,j}$ i -tego produktu jest funkcją określoną na wartościach [3, 6, 7]:

$$p_j(x_{i,j}, x_{0,j}, \Delta x_{\max,j}), \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, n.$$

$\Delta x_{\max,j} = \max_{1 \leq i, l \leq k} (x_{i,j}, x_{l,j})$, oznacza maksymalną różnicę między dowolnymi

dwoma wartościami j -tej cechy produktu przy uwzględnieniu tylko tych przypadków, które są rozpatrywane w kwerendzie.

Wartość funkcji podobieństwa lokalnego określa, jak bardzo podobna jest dana cecha podana w zapytaniu w stosunku do odpowiadającej jej cechy dla i -tego produktu X_i .

Podobieństwo globalne $SIM(X_i)$ mówi o sumarycznej wartości podobieństwa problemu X_0 do i -tego przypadku X_i z bazy przypadków. Podobieństwo globalne może być liczone względem tylko niektórych, wybranych cech przypadku. O tym, które cechy będą wybrane do obliczania podobieństwa globalnego mówi tzw. reguła kompletności przypadku [1]. Podobieństwo globalne zdefiniujemy, jako ważoną sumę podobieństw lokalnych:

$$SIM(X_i) = \frac{\sum_{j=1}^n m_j w_j p_j(x_{i,j}, x_{0,j}, \Delta x_{\max,j})}{\sum_{j=1}^n m_j w_j}$$

gdzie: w_j – przyznana wartość wagi określająca stopień preferencji danej cechy produktu przez klienta, $w_j \in (0, +\infty)$. Współczynnik $m_j \in \{0, 1\}$ wynika z reguły kompletności przypadku. Wartość $m_j = 1$ określa, że przy obliczaniu funkcji podobieństwa lokalnego dla danego zapytania X_0 należy uwzględnić j -tą cechę przypadków z bazy. Zarówno funkcje podobieństwa lokalnego, jak i funkcja podobieństwa globalnego przybierają wartości z przedziału $<0, 1>$. Jak widać, funkcja $SIM(X_i)$ jest adaptacją tzw. metryki miejskiej [7].

Zarówno funkcje podobieństwa lokalnego, jak i globalnego można tworzyć z wykorzystaniem funkcji nieliniowych. Daje to możliwość zaimplementowania szczegółowej wiedzy eksperckiej dotyczącej zależności obserwowanych w obszarze konkretnej cechy produktu.

Dane empiryczne

Materiał empiryczny, na którym przeprowadzono obliczenia, pochodzi z portalu internetowego bankier.pl [9] z dnia 8.06.2003 r. Obejmuje on 142 bankowych rachunków rozliczeniowo-oszczędnościowych, w tym 106 przeznaczonych dla osób fizycznych, a 36 dla klientów instytucjonalnych. Dane dotyczą produktów 27 banków działających na terenie Polski.

Tabela 1 przedstawia nazwy cech rachunków bankowych oraz przedziały wartości wyliczonych cech. W tabeli zostały wskazane zastosowane funkcje podobieństwa lokalnego osobno dla każdej cechy produktu.

Tabela 1

Cechy rachunków bankowych i funkcje podobieństwa lokalnego. (m.w. - wartość regularnych miesięcznych wpływów na konto).

Nazwa cechy	Wartości cechy	Funkcja podobieństwa lokalnego
Nazwa banku	27 nazw	A
Nazwa handlowa rachunku	142 nazwy	A
Czy rachunek jest przeznaczony dla firmy	Tak/Nie	A
Stopa oprocentowania [%]	0–5,25	B
Minimalne wymagane wpływy [PLN]	0–10 000	B
Oплата za prowadzenie [PLN]	0–140	B
Koszt przelewu [PLN]	0–7	B
Koszt zlecenia stałego [PLN]	0–7	B
Kwota debetu [PLN] *	0–300 000	B
Kwota debetu w m.w. *	10	B
Oprocentowanie debetu [%]	0–35	B
Maksymalna kwota kredytu [PLN] **	0–1200 000	B
Maksymalna kwota kredytu w m.w. **	10	B
Oprocentowanie kredytu [%]	0–20,1	B
Dostępność przez telefon	Tak/Nie	A
Dostępność przez Internet	Tak/Nie	A

Źródło: opracowanie własne na podstawie [9].

Funkcje podobieństwa lokalnego zdefiniowano w następujący sposób [3, 6, 7]:

$$A. p_j = \begin{cases} 0, & \text{dla } x_{o,j} = x_{i,j} \\ 1, & \text{dla } x_{o,j} \neq x_{i,j} \end{cases}$$

$$B. p_j = 1 - \frac{|x_{o,j} - x_{i,j}|}{\Delta x_{\max,j}}$$

Reguły kompletności, mówiące o braniu pod uwagę w obliczeniach j -tej cechy rachunku, są zdefiniowane w sposób następujący:

$m_j = 1$, gdy $x_{o,j}$ podano i gdy $x_{i,j}$ ma wartość ustaloną,

$m_j = 0$, gdy $x_{o,j}$ nie podano \vee gdy $x_{i,j}$ ma wartość nieustaloną.

Reguły kompletności dla cech oznaczonych w tabeli 1 przez * i ** mogą dodatkowo być zastosowane zamiennie w parze, jeśli klient zadeklaruje wartość miesięcznych wpływów na rachunek.

Wybór produktów bankowych

Początek pracy z opisywanym systemem doradczym obejmuje przedstawienie potencjalnemu klientowi cech, jakie mają rachunki bankowe i zapoznanie go z zakresem zmienności wartości poszczególnych cech w populacji przypadków. Następnie należy dowiedzieć się, wobec których cech klient ma preferencje i jakie wartości preferowanych cech jest skłonny zaakceptować. Po wprowadzeniu preferencyjnych wartości cech produktu bankowego klient powinien obdarzyć je własnymi subiektywnymi wagami, określającymi stopień ważności poszczególnych własnych preferencji. Wagi mogą być dowolnymi liczbami rzeczywistymi dodatnimi.

Uruchomienie procedury wyszukiwania najbardziej podobnego produktu rachunku bankowego daje w rezultacie listę produktów, którą można posortować malejąco według wartości funkcji podobieństwa globalnego. Lista rachunków tworzona jest zgodnie z preferencją – osobno dla rachunków osobistych i dla klienta instytucjonalnego. Pierwszy na liście jest produkt najbardziej podobny do podanego profilu preferencji.

Adaptacja rozwiązania rozpoczyna się wskazaniem potencjalnemu klientowi banku listy kont bankowych otrzymanej z obliczeń. Już w tym stadium może on ewentualnie dokonać ostatecznego wyboru produktu. Jeśli okaże się, że istotne cechy pierwszego na liście rachunku nie mają akceptowanych wartości, można zastanowić się, czy następne produkty z listy w zadowalający sposób spełniają kryteria pytającego. Klient może również rozważyć dopasowanie swoich preferencji do podanego rozwiązania, być może bardzo korzystnego z innych powodów. Kolejnymi sposobami na dojście do satysfakcjonującego rozwiązania są zmiana preferencji lub wag w zapytaniu i wyszukiwanie produktu po ponownym uruchomieniu kwerendy. Adaptacja rozwiązania może

również polegać na próbie odzyskania w rozwiązaniu kont o funkcjonalności podobnej do opisanej w kwerendzie.

Przykłady przypadków dokonywania wyboru produktu bankowego zamieszczono w tabeli 2. Dla każdego przykładu podano wartości preferencji i wag oraz pierwszą propozycję konkretnego rachunku, najbardziej podobną do zapytania.

Tabela 2

Przykłady wyszukiwania produktu rachunku bankowego na podstawie ustalonych cech preferencyjnych i nadanych wag. Ust. – ustalana indywidualnie. Cechy zawarte w kwerendzie wytłuszczono. [-] – cecha nie dotyczy danego rachunku. Pusta komórka tabeli – brak preferencji dla tej cechy.

Numer przykładu	1			2			3		
Najbardziej podobna propozycja									
Preferencje zawarte w kwerendzie									
Wartości wag ustalone w kwerendzie									
Nazwa cechy									
Nazwa banku							3	N1	N2
Rachunek jest przeznaczony dla firmy		Nie	Nie		Nie	Nie		Tak	Tak
Stopa oprocentowania	5	3%	4,5%			4,6%	3	3%	2,5%
Minimalne wymagane wpływy [PLN]	3	1000	300	5	800	1500			0
Opłata za prowadzenie [PLN]			7	10	0	0			0
Koszt przelewu [PLN]			7	10	0	0	10	0	0
Koszt zlecenia stałego [PLN]			0	10	0	0			0
Kwota debetu [PLN]	5	3000	3000			0	5	100000	-
Kwota debetu w m.w.			-			0			-
Oprocentowanie debetu			16,5			0			-
Maksymalna kwota kredytu [PLN]			-			5000			Ust.
Maksymalna kwota kredytu w m.w.			-			-			-
Oprocentowanie kredytu			-	5	12%	12,5%	10	0	5,95%
Dostępność przez telefon			Tak			Tak	3	Tak	Tak
Dostępność przez Internet			Tak			Tak	3	Tak	Tak

Źródło: badania własne na podstawie [9].

Przykład 1.

Osoba fizyczna chce założyć rachunek, zależy jej na oprocentowaniu rządu 3% i uzyskaniu debetu do 3000 PLN. Mniej ważnym warunkiem jest minimalna wpłata, którą deklaruje w wysokości 1000 PLN. Po wykonaniu zapytania okazuje się, że znajdujemy rachunek o korzystniejszych cechach: wyższej stopie oprocentowania, mniejszych minimalnych wpłatach i żądanej kwocie debetu.

Przykład 2.

Przyszły klient szuka rachunku osobistego. Najbardziej istotne dla niego są zerowe stawki opłaty za prowadzenie rachunku, zerowe koszty przelewów i zleceń stałych. Tylko w połowie tak istotne są minimalne wymagane wpływy, które klient deklaruje na 800 PLN i oprocentowanie kredytu rządu 12%. Na liście propozycji pierwsze miejsce zajmuje rachunek spełniający większość założeń, jednak, aby go założyć, klient musi zadeklarować wpływy 1500 PLN. Jeżeli ta propozycja wyda się pytającemu korzystna, może on zdecydować się na podane warunki, zmieniając nieco swe preferencje.

Przykład 3.

Dla założenia rachunku firmy określono następujące preferencje: jak najmniejsze oprocentowanie kredytu oraz zerowe koszty przelewów. Mniej istotna była wartość debetu [100 000 PLN]. Za najmniej ważne uznano: dostęp przez telefon, dostęp przez Internet, 3% stopę oprocentowania rachunku, oraz nazwę banku, kojarzącą się z rzetelnością usług. Wynik zapytania podany w tabeli wskazuje na najbardziej podobny produkt. Widać jednak, że brak debetu nie odpowiada preferencjom zapytania. Tu jednak można zacząć negocjacje – zamiast debetu przyszły klient prawdopodobnie będzie skłonny do przyjęcia możliwości zaciągania kredytu o ustalonej indywidualnie wysokości, co może pełnić podobną debetowi funkcję. Tak więc, nie dopasowujemy tu wartości danej cechy rachunku, ale próbujemy zaadaptować do pożądanego funkcji inną cechę rachunku.

Wnioski

Przedstawione przykłady pokazują, że funkcja podobieństwa lokalnego przy zastosowaniu wartości skrajnych cechy produktu redukuje się do operacji sortowania w danym wymiarze. Gdy jednak w zapytaniu wybierzemy wartość z wnętrza przedziału zmienności, metoda wykazuje cechy rozwiązania problemu optymalizacyjnego. Przy nieliniowych funkcjach podobieństwa można otrzymać narzędzie optymalizacji nieliniowej.

Wiedza ekspercka o zależności pomiędzy wartościami cech produktów bankowych może być wykorzystana przy: a) ustalaniu funkcji podobieństwa, b) ustalaniu preferencji klienta w trakcie tworzenia zapytania, c) adaptowaniu rezultatów kwerendy.

Opisany w pracy system bazuje na wiedzy potencjalnego klienta banku co do własnych preferencji. Gdy jednak klient taki nie jest zorientowany, jakie są jego potrzeby w zakresie cech konta, warto dokonać przeliczeń na podstawie danych o kliencie i jego działalności. Aplikacja przeliczająca te dane może stanowić dodatkowy element systemu, wykorzystywany w procesie przygotowania kwerendy.

Metoda wnioskowania opartego na bazie przypadków, jako metoda sztucznej inteligencji mająca szeroki krąg zastosowań, ma wiele korzystnych cech, [6]:

- a) może być stosowana bez konieczności dogłębnego rozumienia czy opisanie praw rządzących zjawiskami, których realizacje stanowią zgromadzone przypadki,
- b) nie ma konieczności formułowania hipotez w celu przeprowadzania obliczeń, w odróżnieniu od testów statystycznych,
- c) nie wymaga posiadania wiedzy zapisanej w postaci reguł o tym, jak rozwiązać problem, w odróżnieniu od systemów sztucznej inteligencji opartych na regułach wiedzy,
- d) daje możliwość przetwarzania danych innych niż liczby i odczytania przesłanek, na podstawie których otrzymano rezultat, czym różni się od sztucznych sieci neuronowych.

Mimo wielu zalet, metoda wnioskowania oparta na bazie przypadków wymaga ustalenia funkcji podobieństwa lokalnego i globalnego oraz istnienia wiarygodnej bazy przypadków, na podstawie której zachodzi wnioskowanie.

Literatura

1. BERGMAN R., WILKE W., VOLLRATH I., WESS S., 1996: Integrating General Knowledge with Object-Oriented Case Representation and Reasoning. 4th German Workshop: Case-Based Reasoning – System Development and Evaluation.
2. ARTIFICIAL INTELLIGENCE – KNOWLEDGE BASED GROUP, 2003: Case-Based Reasoning homepage. University of Kaiserslautern, <http://www.cbr-web.org/>
3. Case-Based Reasoning Technology, From Foundations to Applications, 1998: ed. Mario Lenz, Brigitte Bartsch, Hans-Dieter Burkhard, Stefan Wess. Lecture Notes in Computer Science, Springer, Heidelberg.
4. DONNER M., ROTH-BERGHOFER T., 1999: Architectures for Integrating CBR-Systems with Databases for E-Commerce. Proceedings of the 7th German Workshop on Case-Based Reasoning, Wurzburg, Germany.

5. KAJDARA SOFTWARE, 2003: web page <http://www.acknosoft.com/>
6. KOLODNER J., 1993: Case-Based Reasoning. Morgan Kaufman Publishers, Inc., San Francisco, CA.
7. KUKUŁA K., 2000: Metoda unitaryzacji zerowej. Biblioteka Ekonometryczna, Wydawnictwo Naukowe PWN, Warszawa.
8. LENZ M., 1999: Experiences from Deploying CBR Applications in Electronic Commerce. Proceedings of the 7th German Workshop on Case-Based Reasoning, Würzburg, Germany.
9. BANKIER.PL 2003, Zestawienie kont bankowych z dnia 08.06.2003 <http://www.bankier.pl/fo/konta/narzedzia/zestawienie/>

The Case-Based Use of Artificial Intelligence in Choosing Banking Products

Abstract

Banking product extension is a source of widespread information and advertisement, especially fluently supported by Internet media. Case-Based Reasoning (CBR) is an Artificial Intelligence method being used as extension tool, with e-commerce as fast growing application field. Article presents the methodology of applying similarity measures of CBR method to extract bank account products most similar to presented customer preferences. Local similarity measures as well as global similarity measure are described for each account attribute. Computer application based on this method was created within project and applied to search banking products. Real world data of 142 Polish banks' accounts were used in computations. Examples of search for an account are presented to show searching results and different case adaptation techniques.