141/2004

Raport Badawczy

Research Report

RB/31/2004

Lingwistyczne podsumowania baz danych jako element wspomagania zarządzania

M. Gola

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 8373578

fax: (+48) (22) 8372772

Kierownik Pracowni zgłaszający pracę:
Prof. dr hab. inż. Janusz Kacprzyk

Warszawa 2004

Lingwistyczne podsumowania baz danych jako element wspomagania zarządzania

Magdalena Gola
Pracownia Systemów Inteligentnych
Instytut Badań Systemowych PAN

Wprowadzenie

Sprawne i efektywne zarządzanie jest obecne jednym z najważniejszych elementów składających się na sukces firmy. Nie tylko przedsiębiorstwa produkcyjne, ale również organizacje non-profit poszukują skutecznych metod zarządzania. Wprowadzane są na rynek systemy wspomagające zarządzanie w oświacie, kulturze, służbie zdrowia. Wyższe uczelnie coraz częściej oferują nie tylko studia na klasycznych kierunkach menadżerskich takich jak studia MBA, ale również rozszerzają swoją ofertę o różnorakie podyplomowe studia kształcące kadrę menedżerska muzeów, placówek oświaty, teatrów. Zwiększony „popyt" na właściwe i sprawne zarządzanie jest bezpośrednim bodźcem, który uaktywnia poszukiwania coraz to nowych i skuteczniejszych systemów wspomagania zarządzania.

Podstawą właściwego zarządzania, na każdym szczeblu, jest informacja. To ona pozwala wprawnemu menażerowi właściwie ocenić sytuacje firmy i podjąć takie decyzje, które zaowocują najlepszymi wynikami w przyszłości. Informacja jest, więc obecnie najbardziej poszukiwanym „towarem" na rynku, niejednokrotnie towarem bezcennym.

Aby wyjść naprzeciw oczekiwaniom decydentów, dokonuje się w firmach bieżącej archiwizacji danych dotyczących kondycji firmy jak również wybranych parametrów cechujących stan gospodarki w danym momencie czasu oraz poczynań konkurencji. Posiadanie jednak gigabajtów danych nie jest receptą na sukces. Trudno, bowiem oddzielić dane istotne w danym momencie od tych, których znaczenie jest marginalne. Sam manager nie zawsze jest świadom, jakich informacji potrzebuje na danym etapie podejmowania decyzji, a przeszukiwanie olbrzymich baz i hurtowni danych rosnących z dnia na dzień wydaje się iście syzyfową pracą.

Z pomocą przychodzą w takim momencie systemy wspomagające, których zadanie polega nie tylko na systematycznym akumulowaniu potrzebnych danych, ale przede wszystkim na jasnym i klarownym przedstawieniu informacji i trendów zawartych w bazach danych. Zapotrzebowanie na wyżej wymienione systemy sprawia, że poszukiwane są coraz to doskonalsze sposoby odkrywania informacji zawartej w terabajtach danych, a przy tym nie wymagające olbrzymich nakładów obliczeniowych (ze względu na wielkość baz danych jest to nietrywialne wymaganie).

Jednym ze sposobów prezentacji informacji zawartej w zbiorach danych są oparte na teorii zbiorów rozmytych podsumowania lingwistyczne. Wychodzą one naprzeciw

wymogom jasnego i komunikatywnego przekazu informacji zawartej w zebranych danych.

Czym są podsumowania lingwistyczne?

Podsumowania lingwistyczne, to oparte na zbiorach rozmytych, wyrażone w formie słownej tendencje i trendy „zaobserwowane" w zbiorze danych. Jeśli na przykład mamy do czynienia z relacyjną bazą danych w agencji reklamowej możemy otrzymać następujące podsumowanie: „Większość pracowników jest młoda". Jest to dość trywialne podsumowanie, do którego otrzymania nie potrzeba złożonych systemów wspomagania zarządzania, pozwala ono jednak lepiej zrozumieć, czym są podsumowania lingwistyczne.

Przy pomocy podsumowań lingwistycznych możemy poszukiwać informacji dodając na wstępie pewne założenia, na przykład, że poszukujemy informacji związanych z grupą młodych pracowników firmy (generowanie podsumowań z dodatkowymi założeniami wymaga mniejszego nakładu obliczeń), lub nie dodając żadnych dodatkowych założeń (wtedy oczywiście obliczenia są bardziej złożone, a w przypadku dużych baz danych praktycznie niemożliwe do przeprowadzenia).

Odkryte informacje mogą pozwolić na lepsze zarządzanie firmą a tym samym na zmniejszenie kosztów możliwe, że przy równoczesnym zwiększeniu zysków. Nie należy jednak upatrywać w podsumowaniach lingwistycznych panaceum na wszystkie problemy firmy lub niezawodnego środka do osiągnięcia sukcesów. Pomagają one odnaleźć wiedzę, którą w pewnych przypadkach trudno byłoby wyekstrahować z bazy

danych tak, aby w jasny i czytelny sposób mogła posłużyć do podjęcia właściwych decyzji. Czasami może ona być bardzo użyteczna a czasami potwierdzi tylko to, o czym już dawno wiedzieliśmy.

Generator podsumowań lingwistycznych najlepiej włączyć w system wspomagania zarządzania, który jest wykorzystywany w danej firmie. Ważne jest, aby był to system kompleksowy (nie zaś oddzielny dla każdego z działów), pozwalający czerpać z ogółu informacji udostępnianych przez poszczególne działy.

Zbiory rozmyte – rys historyczny

Teoria zbiorów rozmytych została wprowadzona przez profesora Zadeha w 1965 roku [1]. Polega ona na zastąpieniu klasycznej funkcji charakterystycznej (używanej w konwencjonalnej teorii mnogości), która przyjmując jedynie dwie wartości 0 lub 1 opisywała przynależność elementu do zbioru (gdy wartość funkcji wynosi 1 dla danego elementu – element ten należy do zbioru, natomiast gdy wartość funkcji wynosi 0 dla danego elementu - element nie należy do zbioru), tzw. funkcją przynależności, mogącą przyjmować wartości z całego przedziału [0,1]. Dzięki temu możliwe jest uchwycenie „częściowej" przynależności do zbioru.

Definicja 1

Zbiór rozmyty A w przestrzeni rozważań X, jest określony jako para

$$A=\{(\mu_A(x),x)\}$$

Gdzie $\mu_A(x)$: $X \rightarrow [0,1]$ jest funkcją przynależności zbioru rozmytego A, a $\mu_A(x)$ jest stopniem przynależności elementu $x \in X$.

Dzięki wprowadzeniu pojęcia zbioru rozmytego, możliwe jest naśladowanie języka naturalnego, w którym np. poza pojęciem „ciepło i zimno" istnieją np. pojęcia: „raczej ciepło", „raczej zimno", „bardzo ciepło". Tak, więc teoria zbiorów rozmytych powstała z potrzeby opisania słabo definiowalnych pojęć z języka naturalnego, dla których klasyczny aparat matematyczny okazał się zbyt „ubogi".

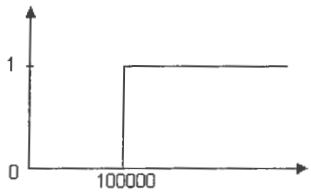
Przykład:

Zbiór klasyczny

bogaty = {x | stan posiadania (x) ≥ 100 000 zł}

Funkcja charakterystyczna zbioru „bogaty"

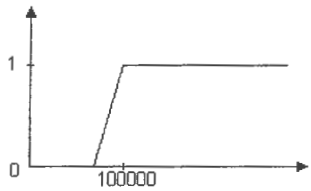
$$\mu_x = \begin{cases} 0 & s\tan posiadania\ (x) < 100000 \\ 1 & s\tan posiadania\ (x) \geq 100000 \end{cases}$$



Wykres funkcji charakterystycznej μ_x zbioru „bogaty"

Tak, więc za pomocą tego zbioru możemy określić tylko dwie możliwości: ktoś albo „jest bogaty", (czyli jego stan posiadania jest większy bądź równy 1000000) albo „nie jest bogaty" (posiada mniej niż 1000000). Łatwo zauważyć, że nie wyczerpujemy tym sposobem opisu wszystkich możliwości. Co z osoba, której stan konta to 999999 zł? Według powyższej definicji nie jest ona bogata, choć jej stan posiadania różni się tylko o 1 zł od osoby bogatej, podobnie jak nie jest bogata osoba, której stan posiadania to 50zł. Trudno jednak powiedzieć, że stan posiadania wyżej wymienionych osób jest sobie bliższy niż pomiędzy osobą o stanie konta 999999zł oraz 1000000zł.
Język naturalny wyróżnia wiele stanów pośrednich, takich jak na przykład: „raczej bogaty", „raczej niebogaty", „zdecydowanie niebogaty", itp. Aby móc je przedstawić musimy porzucić klasyczną teorię zbiorów i rozważyć zastosowanie w tym przypadku zbioru rozmytego.

Zbiór rozmyty



Wykres funkcji charakterystycznej μ_x zbioru rozmytego „bogaty"

Jak łatwo zauważyć funkcja charakterystyczna tego zbioru przyjmuje nie tylko wartości $\{0,1\}$ ale również wszystkie wartości pośrednie. Możemy, zatem przyjąć, że gdy $\mu_x \in$

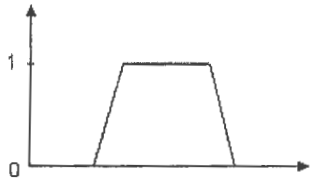
[0.7,0.9] możemy mówić, że x jest „średnio bogaty", a gdy $\mu_x \in [0.2,0.4]$ to x jest „raczej niebogaty".

Powyższy trywialny przykład pozwala zilustrować, w jaki sposób zbiory rozmyte rozszerzają możliwość uchwycenia nie do końca precyzyjnych określeń języka naturalnego.

Oczywiście istnieje wiele innych możliwych funkcji charakterystycznych dla zbiorów rozmytych, które pozwalają nam na właściwe dopasowanie do warunków, które chcemy opisać przy pomocy danego zbioru rozmytego.

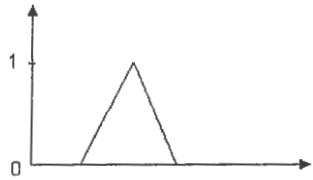
I tak funkcja może być:

- Funkcją „trapezową":



która może opisywać zbiór „w środku tygodnia".

- Funkcją „trójkątną":



która może opisywać zbiór „w okolicach środy".

- Funkcją „okresową":

która może opisywać zbiór „na początku miesiąca".

- Itp.

Od czasu wprowadzenia teorii zbiorów rozmytych jest ona nieustannie rozwijana i „zaszczepiana" na różnych polach i tak Bellman i Zadeh [2] wprowadzili ogólne podejście do podejmowania decyzji w warunkach rozmytości (tzn., gdy cele i

ograniczenia są rozmyte), Zadeh [4] zaproponował użycie zbiorów rozmytych w analizie systemów itd.

Dziś możemy spotkać się z „rozmytością" praktycznie na każdym kroku. Kupując aparat fotograficzny, pralkę czy wsiadając do metra mamy pośredni kontakt z teorią zbiorów rozmytych. Wprowadzenie, zatem teorii zbiorów rozmytych do procesów ekstrakcji informacji jest, więc w pewnym sensie naturalnym krokiem na drodze „ekspansji" zbiorów rozmytych.

Teoria zbiorów rozmytych

Konieczne jest wprowadzenie jeszcze kilku dodatkowych definicji stanowiących fundamenty teorii zbiorów rozmytych.

Oznaczenia

W przypadku zbiorów rozmytych w skończonej przestrzeni rozważań X, będziemy używać notacji:

$$A = \{(\mu_A(x), x)\} = \{\mu_A(x)/x\} = \mu_A(x_1)/x_1 + \ldots + \mu_A(x_n)/x_n,$$

gdzie $\mu_A(x_1)/x_1$ oznacza singleton rozmyty.

Definicja 2

Zbiór rozmyty A w X nazywamy zbiorem pustym, wtedy i tylko wtedy, gdy

$$\mu_A(x) = 0$$

dla każdego $x \in X$.

Definicja 3

Zbiór A jest podzbiorem zbioru B, co oznaczamy $A \subseteq B$, *wtedy i tylko wtedy, gdy*

$$\mu_A(x) \leq \mu_B(x)$$

dla każdego $x \in X$.

Definicja 4

Zbiór rozmyty A nazywamy normalnym, wtedy i tylko wtedy, gdy

$$\max_{x \in X} \mu_A(x) = 1.$$

W przeciwnym wypadku zbiór rozmyty nazywamy subnormalnym.

Definicja 5

Nośnikiem zbioru rozmytego A, nazywamy zbiór

$$suppA = \{x \in X : \mu_A(x) > 0\}.$$

Definicja 6

Licznością nierozmytą zbioru rozmytego $A = \mu_A(x_1)/x_1 + \ldots + \mu_A(x_n)/x_n$, *nazywamy wielkość*

$$\sum Count(A) = \sum_{i=1}^{n} \mu_A(x_i)$$

Podstawowe operacje na zbiorach rozmytych

Dopełnienie *zbioru rozmytego* ($\neg A$)

$$\mu_{\neg A}(x) = 1 - \mu_A(x)$$

Przecięcie *dwóch zbiorów rozmytych A i B*

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$$

Suma *dwóch zbiorów rozmytych A i B*

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$$

Definicja 7

Funkcję $t : [0,1] \times [0,1] \rightarrow [0,1]$ *nazywamy t-normą, jeżeli dla każdego* $a, b, c \in [0,1]$:

1. $t(a,1) = a$

2. $a \leq b \Rightarrow t(a,c) \leq t(b,c)$

3. $t(a,b) = t(b,a)$

4. $t[a,t(b,c)] = t[t(a,b),c]$

Przykładem t-normy jest funkcja minimum oraz iloczyn algebraiczny.

Definicja 8

Funkcję $s:[0,1]\times[0,1] \rightarrow [0,1]$ *nazywamy s-normą, jeżeli dla każdego* $a,b,c \in [0,1]$:

1. $s(a,0) = 0$

2. $a \leq b \Rightarrow s(a,c) \leq s(b,c)$

3. $s(a,b) = s(b,a)$

4. $s[a,s(b,c)] = s[s(a,b),c]$

Przykładem s-normy jest funkcja maksimum oraz iloczyn probabilistyczny ($s(a,b) = a + b - ab$).

Definicja 9

Relacją rozmyta R między dwoma zbiorami nierozmytymi X={x} i Y={y} nazywamy zbiór rozmyty określony na iloczynie kartezjańskim $X \times Y$, *tzn.*

$$R = \{(\mu_R(x,y),(x,y))\} = \{\mu_R(x,y),(x,y)\},$$

dla każdej pary $(x,y) \in X \times Y$.

Wprowadzenie relacji rozmytej pozwala nam wyrazić nieprecyzyjne zależności pomiędzy elementami, (w jakim stopni elementy są ze sobą w relacji), podczas gdy relacje nierozmyte pozwalały jedynie określić czy elementy są czy nie są sobą w relacji.

Definicja 10

Złożeniem typu s-t dwóch relacji rozmytych R w $X \times Y$ *oraz S w* $Y \times Z$ *nazywamy następującą relacje rozmytą* $R \circ_{s-t} S$ *w* $X \times Z$

$$\mu_{R \circ_{s-t} S}(x, z) = s_{y \in Y}[\mu_R(x, y)\, t\, \mu_S(y, z)],$$

dla każdego $x \in X$, $z \in Z$.

Zmienne lingwistyczne

Zmienne lingwistyczne to zmienne, które nie przyjmują zwykłych wartości numerycznych, ale wartości lingwistyczne, które są utożsamiane z pewnymi zbiorami rozmytymi. I tak zmienną lingwistyczną może być „wzrost", która może przyjmować wartości: „wysoki", „niski", „średni", czy też bardziej złożone np. „średnio wysoki" itd. Wartości lingwistyczne są utożsamiane z pewnymi zbiorami rozmytymi.

Z prostych wartości lingwistycznych można utworzyć bardziej złożone wyrażenia lingwistyczne stosując negację („nie niski"), spójniki („nie niski i nie wysoki") oraz

modyfikatory („mniej więcej średni"), które również utożsamia się semantycznie ze zbiorami rozmytymi stosując odpowiednie reguły semantyczne.

Logika rozmyta

Definicja 11

Logika rozmyta to podstawa wnioskowania opartego na niejednoznacznych i nieprecyzyjnych stwierdzeniach, w którym stosuje się metody teorii zbiorów rozmytych (Zadeh).

Załóżmy, że istnieje predykat:

$$u \ is \ P$$

np. „inflacja jest wysoka gdzie „u" jest zmienną a P jest terminem nieprecyzyjnym utożsamianym ze zbiorem rozmytym.

Konieczne jest również ustalenie stopnia prawdy stwierdzenia *P*, który oznaczamy $\tau(P)$ i obliczamy w następujący sposób:

$$\tau(u \ is \ P) = \mu_P(u).$$

Rozmyte kwantyfikatory lingwistyczne

W przeciwieństwie do logiki klasyczne, w której używamy dwóch kwantyfikatorów tzn. „istnieje" i „dla każdego" w logice rozmytej mam do czynienie z całą gamą kwantyfikatorów np. „większość", „prawie żaden", „kilka". Dzięki temu możliwe jest ujęcie logiczne takich zdań języka naturalnego jak: „większość Polaków posiada samochód".

Niech dane będzie zdanie z kwantyfikatorem logicznym: „większość samochodów jest srebrna", które można ogólnie zapisać jako:

$$Qy's\ are\ F$$

gdzie Q jest kwantyfikatorem rozmytym „większość", $Y=\{y\}$ jest zbiorem obiektów np. „samochody", a F jest właściwością (w tym przypadku „srebrna"). Podobnie jak w poprzednim przypadku konieczne jest również obliczenie stopnia prawdy zdania z kwantyfikatorem (wartość stopnia prawdy będzie oczywiście z przedziału [0,1]), który jest określany przy pomocy nierozmytych liczb kardynalnych (patrz *Definicja 6*).

W pierwszym kroku obliczamy wartość r:

$$r = \frac{\sum Count(F)}{\sum Count(Y)} = \frac{1}{p}\sum_{i=1}^{p}\mu_F(y_i)$$

a następnie wartość stopnia prawdy

$$\tau(Qy's\ are\ F) = \mu_Q(r).$$

Operatory uporządkowanej średniej ważonej (OWA)

Definicja 12

Operator OWA (ang. ordered weighted averaging) o wymiarze n to odwzorowanie

$$F:[0,1]\rightarrow[0,1]$$

o wektorze wag

$$W=[w_1,\ldots,w_n]^T$$

takim, że:

- $w_i \in [0,1]$, *dla każdego* $i=1,\ldots n$,
- $\sum_{i=1}^{n} w_i = 1$ *oraz*

$$F(a_1,\ldots a_n)=W^T B=\sum_{j=1}^{n} w_j b_j$$

gdzie b_j *jest j-tym elementami wektora* B, *powstałego ze zbioru* $\{a_1,\ldots,a_n\}$ *przez uporządkowanie nierosnące jego elementów (* $b_i \geq b_j$ *dla* $j>i$ *).*

.

Operatory OWA jest ogólnym operatorem agregacji, który w zależności o doboru wektora wag umożliwia agregację z użyciem żądanego operatora agregacji. I tak, jeżeli wektor wag W ma postać

$$W=[1,0,\ldots,0],$$

wówczas otrzymujemy agregację typu maksimum, która odpowiada kwantyfikatorowi egzystencjalnemu. Drugim krańcowym przypadkiem jest operator o wektorze wag

$$W = [0,\ldots,0,1]$$

który odpowiada kwantyfikatorowi szczegółowemu. Wszystkie pośrednie wektory wag pozwalają otrzymać pośrednie typy operatorów agregacji np. „większość".

Operatory OWA posiadają następujące własności:

- są przemienne
- są monotoniczne
- są idempotentne
- oraz

$$\max_{i=1,\ldots,n} a_i \geq F(a_1,\ldots,a_n) \geq \min_{i=1,\ldots,n} a_i$$

Istnieją procedury określania wektora wag dla zadanego z góry rozmytego klasyfikatora lingwistycznego.

Podsumowania lingwistyczne

Yager [5] jako pierwszy wprowadził podsumowania lingwistyczne, których celem jest odnajdywanie charakterystycznych cech analizowanych danych i przedstawianie ich w postaci opisowej (za pomocą języka naturalnego). Tu z pomocą przychodzą nam opisane już wcześniej zdania z kwantyfikatorem logicznym.

Istnieją dwie możliwości:

- podsumowania są tworzone jako odpowiedz na zapytanie użytkownika a następnie następuje weryfikacja ich poprawności i dokładności w oparciu o

zbiór danych (obliczany jest stopień prawdy dla danego podsumowania na podstawie zbioru danych),

- poszukiwanie podsumowań odbywa się automatycznie na danym zbiorze obiektów (dla każdego podsumowania obliczany jest stopień prawdy).

Oczywiście drugie podejście stwarza więcej możliwości odkrycia nietrywialnych trendów i zależności w zbiorze danych, wymaga również dużo większego nakładu obliczeń, co powoduje, że w praktyce większą rolę odgrywa poszukiwanie podsumowań na podstawie zapytań.

Mając bazę danych firmy reklamowej możemy sformułować zapytanie o: „zlecenia w okresie wakacyjnym". Istotnym jest w takim przypadku ustalenie minimalnego stopnia prawdy, który będzie dla nas do zaakceptowania dla otrzymanych podsumowań (unikniemy w ten sposób generowania podsumowań nieistotnych, dla których wsparcie w bazie jest niewielkie). Jednym z podsumowań, które możemy otrzymać jako odpowiedź na wyżej postawione zapytanie jest: „niewiele zleceń jest w okresie wakacyjnym" lub też „ około jedna trzecia zleceń w okresie wakacyjnym pochodzi od stałych klientów".

Warto również zastanowić się czy oprócz „stopnia prawdy" podsumowania nie wprowadzić dodatkowych miar jakości podsumowania. Może się, bowiem okazać, że właściwsze będą inne miary lub ich kombinacje.

Podsumowania a zarządzanie

Bazy danych stanowią źródło wiedzy nie tylko dla kierownictwa, lecz także dla wszystkich osób odpowiedzialnych za sprawne funkcjonowanie firmy. Tak, więc analizując lingwistyczne podsumowania bazy danych należy pamiętać, że powinny być one dostosowane swoim zakresem do osób na różnych stanowiskach.

Baza danych jest prawie chroniona i nie każdy pracownik firmy ma dostęp do wszystkich jej zasobów. Dane dotyczące dochodu ze sprzedaży nie są udostępniane kierownikowi magazynu itp. Tak, więc baza danych zawierająca moduł do generowania podsumowań lingwistycznych musi umożliwiać administratorowi personalizację interfejsów poszczególnych użytkowników, tak, aby dany użytkownik miał dostęp tylko do tych podsumowań, które są oparte na danych jemu dostępnych. Wymaga to oczywiście większego nakładu pracy, pozwalając jednak na właściwą ochronę strategicznych danych firmy. Tworzenie systemu opartego na podsumowaniach lingwistycznych wymaga, więc współdziałania z jego przyszłymi użytkownikami oraz z kierownictwem danej firmy, dla której dany system jest przygotowywany.

Wiedza zawarta w danych może znacznie usprawnić funkcjonowanie firmy, nie wystarczy, bowiem tylko „mieć" kompletną bazę danych, lecz prawdziwą sztuką pozostaje maksymalne wykorzystanie informacji zawartych w bazie. Użyteczne może okazać się również użycie podsumowań lingwistycznych do danych „mieszanych", tzn. pochodzących zarówno z bazy jak i z innych źródeł. W tym przypadku Internet staje się dla nas drugą potężną bazą danych. W wyniku takiej kompilacji danych możemy

otrzymać wynik: „większość transakcji dokonywana jest w ciepłe dni" (co mogłoby sugerować zwiększanie podaży w okresie letnim), czy też „niski kurs euro zwiększa zamówienia"(co pozwala uzależnić stan magazynu od aktualnej koniunktury).

Istotnym problemem pozostaje nadal odnalezienie tych danych spoza bazy danych, które mogą być w istotny sposób powiązane z danymi z bazy. Trudno założyć na początku, jakie dodatkowe dane będą nam potrzebne i użyteczne przy generacji podsumowań. Poszukiwanie zależności pomiędzy wysokością sprzedaży a temperatura i opadami nie jest trywialne i możemy dopiero po pewnym czasie próbować je odszukać. Nie mamy, zatem na bieżąco gromadzonych danych pogodowych dla interesującego nas rejonu i zatem aby spróbować wygenerować takie podsumowania musimy odnaleźć bazy danych pogodowych z okresu, jaki chcemy rozpatrywać, a takowe bazy są prowadzone tylko dla niektórych miejsc na kuli ziemskiej. Optymalnym rozwiązaniem w takim wypadku byłaby wcześniejsze założenie o trendach, jakie chcemy zaobserwować i bieżące gromadzenie dodatkowych informacji w naszej bazie (np. dane o temperaturach w danym rejonie, które mogą być na bieżąco pobierane ze stron serwisów pogodowych).

Oczywiście strategicznym działem w każdej firmie jest dział sprzedaży i najważniejsze wydają się być te podsumowania lingwistyczne, które odnoszą się do wyników sprzedaży. Jednak należy pamiętać, że sukces tego działu zależy od sprawnego funkcjonowania wszystkich działów firmy, dlatego też podsumowania lingwistyczne powinny wspomagać każdy z działów, nie naruszając równocześnie poufności strategicznych danych zawartych w bazie.

FIRMA

Spróbujmy teraz przyjrzeć się poszczególnym działom firmy, ich zadaniom, potrzebom i oczekiwaniom.

Jak już wspomnieliśmy wcześniej, działem najważniejszym w firmie jest dział sprzedaży. To on bezpośrednio generuje zyski, które wypracowuje cała firma. Sukces firmy zależy jednak nie tylko od działu sprzedaży, ale również od każdego z działów firmy, od ich sprawności jak również od kosztów ich funkcjonowania. Oczywistym jest, że brak odpowiedniego gospodarowania stanem magazynów może spowodować znaczne zwiększenie kosztów jego utrzymania, a tym samym wpłynie na wyniki firmy. Podsumowania lingwistyczne powinny wspomagać wszystkich managerów poszczególnych działów, których zadaniem jest sprawne zarządzanie powierzonymi im działami.

Dział sprzedaży

Dział ten dostarcza do bazy informacje o zmianach wielkości sprzedaży, z wyszczególnieniem zamówień dokonywanych przez poszczególnych kontrahentów w poszczególnych okresach. W bazie mogą znaleźć się również plany dotyczące sprzedaży na kolejne okresy ogólnie i z podziałem na konkretnych odbiorców. Większość z tych informacji ma znaczenie strategiczne dla firmy, dlatego też podsumowania oparte na ich podstawie nie mogą być ogólnodostępne, podobnie jak

same dane. „Wyciek" tych informacji mógłby, bowiem poważnie zaszkodzić interesom firmy.

Dzięki zgromadzonym danym możliwe jest poszukiwanie trendów pomiędzy ceną i wielkością sprzedaży, oraz wpływ polityki rabatowej na tą wielkość. Możemy poszukiwać podsumowań mówiących czy dane spływające z poszczególnych działów nie wykazują korelacji pomiędzy wysokością sprzedaży. Sprzedaż może również zależeć od koniunktury w gospodarce, działań konkurencji, pory roku, itd. Trudno wymienić wszystkie możliwe zależność, jakie mogą się pojawić, ponieważ te najbardziej nieprzewidywalne trendy pozwolą nam na najskuteczniejsze działania mające na celu zwiększenie sprzedaży.

Dostrzeżenie trendów w zmianach sprzedaży pozwala na dostosowanie produkcji do przewidywanych potrzeb, zminimalizowania ilości produktów zalegających w magazynach, zaplanowania akcji promocyjnych, które pozwolą na zwiększenie sprzedaży w okresach zastoju, korektę tych akcji promocyjnych, które nie przyniosły zamierzonego efektu

Dział marketingu

Aktywność tego działu jest ściśle powiązana z działem sprzedaży (w niektórych firmach jest to wręcz część działu sprzedaży, choć logicznym posunięciem wydaje się jednak wyodrębnienie działu marketingu jako samodzielnego działu w firmie ze względu na jego specyfikę).

Dział marketingu dostarcza informacje o przeprowadzonych akcjach promocyjnych, ich zakresie i grupie docelowej, do której były adresowane, wynikach badań marketingowych, itp., których analiza i przełożenie na wyniki sprzedaży jest niezwykle istotna. Należy się zastanowić czy do bazy nie powinny być dołączone informacje o kampaniach reklamowych bezpośrednich konkurentów, ponieważ to również może mieć wpływ na fluktuacje sprzedaży.

Dział marketingu będzie czerpał informacje z działu sprzedaży, aby móc oceniać odbiór i skuteczność przeprowadzonych akcji reklamowych (w wypadku tego działu nie wskazanym jest utajnianie danych dotyczących sprzedaży, ponieważ są one podstawą analiz dla tego działu). Oprócz danych dotyczących sprzedaży dział marketingu powinien rozważyć możliwość włączenia do swoich analiz danych dotyczących sytuacji w gospodarce i innych wskaźników makroekonomicznych. Od kreatywności tego działu zależy, bowiem jego efektywność.

Dział produkcji

Struktura i zadania tego działu uzależnione są od specyfiki firmy. Trudno porównywać działy produkcji firmy specjalizującej się w przetwórstwie owocowo-warzywnym i firmy komputerowej. Niemniej jednak istnieją pewne wspólne cechy działów produkcji nawet dla tak odmiennych firm.

Działalność działu produkcji jest ściśle powiązana z działem sprzedaży, który dostarcza informacji o wielkości produkcji, na jaką jest aktualnie zbyt, oraz z działem

zaopatrzenia, który dostarcza potrzebnych substratów. Tak, więc dynamika sprzedaży wpływa na wielkość produkcji natomiast dział zaopatrzenia warunkuje możliwości wytwórcze działu (oczywiście w pewnych określonych granicach).

Wszelkie analizy dokonane przez dział sprzedaży oraz dział marketingu pozwalają na dokładniejsze planowanie poziomu produkcji, natomiast wygenerowanie podsumowania lingwistyczne dotyczące procesów produkcyjnych (ich generacją zajmuje się manager działu produkcji) stają się narzędziem do optymalizacji tych procesów, a w konsekwencji do obniżania kosztów produkcji.

Wykorzystanie podsumowań lingwistycznych w dziale produkcji powinno odbywać się na dwóch płaszczyznach:

- jedna to płaszczyzna optymalizacji wielkości produkcji oparta na wskazówkach otrzymanych przez działy sprzedaży oraz marketingu
- natomiast druga to płaszczyzna optymalizacji wewnętrznych procesów działu, oparta na wewnętrznych analizach otrzymywanych podsumowań.

Otwartą kwestą pozostaje, które z danych, które może dostarczyć dział produkcji powinny być dołączone do bazy danych. Najprostszym rozwiązaniem byłoby dołączenie wszelkich możliwych do uzyskania danych, (np. o zastojach spowodowanych awariami, kosztach napraw i przeglądów parku maszynowego, zwolnieniach lekarskich poszczególnych pracowników itd.). Istnieje jednak ryzyko nadmiernego przepełnienia bazy danych, co będzie miało swoje odzwierciedlenie w złożoności obliczeń podczas generowania podsumowań lingwistycznych.

Dział zaopatrzenia

Dział ten współpracuje ściśle z działem produkcji i pośrednio z działem sprzedaży. Jego główne koszty są związane z magazynowaniem surowców. Dane dostarczane do bazy dotyczą dostawców, terminowości zamówionych dostaw, czasu dostaw, czasu składowania w magazynie, itd. Dane te nie są szczególnie chronione, tak wiec mogą być ogólnodostępne.

W przypadku działań mających na celu optymalizację kosztów dane gromadzone przez dział zaopatrzenia mogą okazać się bardzo pomocne.

Dział transportu

W przypadku działu transportu istnieją dwie możliwości:

- firma posiada własną flotę transportową, co uniezależnia ją od pośredników i pozwala na optymalizowanie całego łańcucha dostaw,
- firma nie posiada własnej floty i korzysta z usług przewoźników dostępnych na rynku.

Możliwa jest również sytuacja pośrednia, w której tylko część transportu realizowana jest przez własny park maszynowy, częściowo zaś firma korzysta z usług innych przewoźników.

W pierwszym przypadku na podstawie danych działu sprzedaży i marketingu, generowane podsumowania są szansą na optymalizację kosztów utrzymania parku maszynowego jak również kosztów transportu. W przypadku drugim nie mamy tak dużych możliwości, jeżeli chodzi o optymalizację transportu, choć oczywiście pewnym zakresie optymalizacja jest możliwa (choćby ze względu na terminowość i cennik dostawców).

Dział kadr

Dział kadr jest z założenia działem pomocniczym, który wspiera pozostałe działy firmy. Trudno jednoznacznie stwierdzić, w jaki sposób podsumowania lingwistyczne mogą wpłynąć na usprawnienie tego działu. Z pewnością dla kadry zarządzającej firmą istotne mogą być trendy zaobserwowane w wysokości płac dla poszczególnych grup pracowników (nie mówimy oczywiście o trywialnych podsumowaniach: „menadżerowie zarabiają dużo"), w terminach i długości zwolnień lekarskich, w terminach i długości urlopów itd.

Zarząd

Podsumowania lingwistyczne mogą stanowić również źródło informacji dla zarządu. Opis trendów i tendencji odkrytych w bazie za pomocą języka naturalnego pozwala na szybszą ocenę sytuacji niż w przypadku, gdy kadra zarządzająca otrzymuje tylko liczby. Należy wziąć pod uwagę fakt, że dane liczbowe nie są równie czytelne dla wszystkich osób pełniących kierownicze stanowiska. Możliwość operowania podsumowaniami

lingwistycznymi pozwala na większe zrozumienie sytuacji, w jakiej znajduje się przedsiębiorstwo i głównych czynników, jakie mają wpływ na jego wyniki.

Bibliografia

[1] Zadeh L.A. - Fuzzy sets. *Information and Control 8.* 1965

[2] Bellman R.E. Zadeh L.A.- *Decision making in a fuzzy environment. Management Science17.* 1970

[3] Kacprzyk J. - Wieloetapowe sterowanie rozmyte. *Wydawnictwa Naukowo-Techniczne.* 2001

[4] Zadeh L.A. - Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics* SMC-2. 1973

[5] Yager R.R. – A new approach to the summarization of data. *Information Sciences.* 1982

[6] Zadeh L.A. – Fuzzy logic and approximate reasoning. *Synthese 30.* 1975

[7] Strykowski P. – Lingwistyczne podsumowania baz danych z użyciem logiki rozmytej i algorytmów genetycznych. 1998