

**Dr. sc. Brano Markić**

Izvanredni profesor na Ekonomskom fakultetu Sveučilišta u Mostaru i Fakultetu za turizam i vanjsku trgovinu u Dubrovniku

## **POSLOVNA INTELIGENCIJA I POSLOVNE OBRADE PODATAKA**

UDK/UDC: 681.3(075)

JEL klasifikacija/JEL classification: C81, L86

Pregledni rad/Review

Primljeno/Received: 9. siječnja 2001./January 9, 2001

Prihvaćeno za tisak/Accepted for publishing: 22. veljače 2001./February 22, 2001

### **Sažetak**

*U radu se analizira poslovnu inteligenciju u kontekstu poslovnih obrada. Poslovna inteligencija je skup procesa, aktivnosti i postupaka pronalaženja i generiranja korisnih i uporabljivih informacija iz podataka pohranjenih u bazama i skladištima podataka. S informatičkog stajališta je temeljna zadaća poslovne inteligencije davanje odgovora na analitičke upite o poslovanju: «Zašto je prodaja na tržišnom segmentu A smanjena u odnosu na prodaju iz istog razdoblja prošle godine?», «Koje proizvode kupci najčešće kupuju istodobno s kupovinom proizvoda P<sub>1</sub> i P<sub>2</sub>?», «Zašto je prosječna razina zaliha povećana u odnosu na prodaju tržišnom segmentu C?», «Koliki promet očekujemo za svibanj sljedeće godine za proizvod P<sub>1</sub> na tržišnom segmentu A?» i sl. Softver koji omogućava pronalaženje odgovora na ovakve upite predstavlja najvažniji dio sustava poslovne inteligencije i može se podijeliti na tri temeljne kategorije: Upite i izvješća (Query and Report), sustave za potporu odlučivanju (Decision Support Systems) i izvršne informacijske sustave (Executive Information Systems). U radu se prate i analiziraju područja usko povezana s poslovnom inteligencijom: otkrivanje znanja u bazama podataka, alati za traganje po podacima, OLAP alati te skladišta podataka kao osnovicu podataka u analizi i generiranju znanja.*

**Ključne riječi:** *poslovna inteligencija, data mining, OLAP, skladišta podataka, knowledge discovery.*

## UVOD

Upiti i izvješća omogućavaju analizu primarnih (transakcijskih) ili deriviranih podataka i informacija pohranjenih u bazama i skladištima podataka. Za relacijske baze podataka ponajprije se koristimo upitima operacije selekcije kao jedne od dostupnih operacija relacijske algebre. Podaci se mogu pretraživati, sortirati, pregledavati i prikazivati na različite načine, ovisno o postavljenim zahtjevima i interesu korisnika. Upiti i izvješća omogućuju uvid u podatke, redak po redak relacijskih tablica, izvršavanje jednostavnijih matematičkih i logičkih operacija uporabom odgovarajućih funkcija (npr. Average, Sum, Max, Min, Stdev i sl.).

Sustavi za potporu u odlučivanju se grade i koriste u sofisticiranim oblicima analize i sofisticiranim upitima. Odgovori na takve upite nisu ni jednostavni ni trivijalni. Zato sustavi za potporu u odlučivanju koriste različite matematičke i statističke modele i funkcije koji omogućavaju analizu prošlih i predviđanje i prognoziranje budućih rezultata. Najvišu razinu aplikacija poslovne inteligencije predstavljaju izvršni informacijski sustavi. Oni po svojoj definiciji pokazuju korisnicima manji broj ključnih indikatora poslovnih rezultata i uspjeha.

Softver za poslovnu inteligenciju se veoma često oslanja na skladišta podataka i skupa se s njima koristi. Odgovor na poslovne upite često iziskuje podatke i informacije iz više izvora koji se u pravilu ne nalaze isključivo unutar organizacijskog sustava.

Skladišta podataka su predmetno orijentirani, nepromjenljivi i samo za čitanje pohranjeni podaci. Ovo se odnosi na cjelinu tvrtke i nastaje u procesima filtriranja, pročišćavanja i ekstrakcije podataka iz više baza. Terminom "data mart" označavamo skladište podataka za neku od poslovnih funkcija.

Poslovna inteligencija i skladišta podataka predstavljaju moćne tehnologije u rješavanju složenih poslovnih problema. "International Data Corporation" analizirala je benefite od implementacije 62 skladišta podataka za velike korporacije. Rezultati istraživanja su pokazali da prosječan povrat ulaganja za tri godine iznosi 401% i da jedna četvrtina korporacija ima povrat veći od 600%. Softver za poslovnu inteligenciju omogućava smanjivanje prosječne veličine zaliha, što smanjuje ukupne troškove i povećava financijski rezultat korporacije.

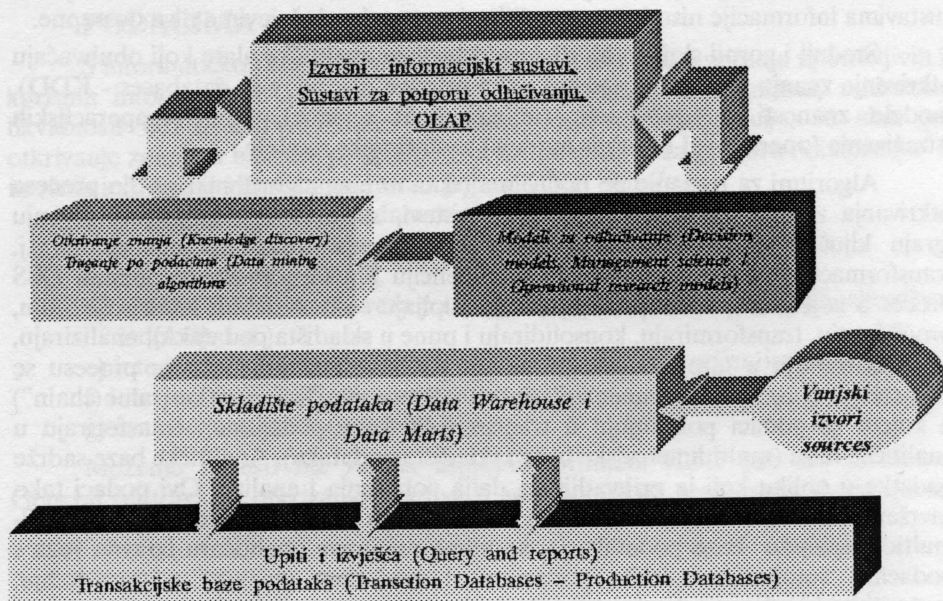
## 1. POSLOVNA INTELIGENCIJA

Većina organizacija, pojedinaca, tvrtki i institucija akumulirala je ogromne količine transakcijskih (primarnih) podataka u operativnim bazama. Cilj zapravo nije ogromna količina pohranjenih podataka nego informacije koje se mogu izvući iz njih. Želi se učiti iz podataka koji kao činitelji determiniraju prodaju i način, kako najbolje zadovoljiti kupce, kako najučinkovitije alocirati i kombinirati raspoložive resurse i minimizirati gubitke.

Podaci često predstavljaju šumu iz koje nije moguće jednostavno izvoditi smislene zaključke. Zato znanstvenici tragaju za odgovorom u kreiraju nove tehnologije za potporu izvođenju ispravnih i valjanih zaključaka na temelju

"ogromnih" količina podataka. Poslovna inteligencija je upravo stoga za velik broj organizacija neophodna u procesu pronalazjenja potrebitih informacija iz gomile raspoloživih podataka. Pohranjeni podaci se mjere u gigabajtima, terabajtima ili čak u pedabajtima. Takvi primarni podaci o transakcijama su jasan pokazatelj stanja, u organizaciji. Oni ne daju jasan odgovor zašto se nešto događa. Njima se ne dijagnosticiraju procesi i promjene, ne interpretiraju rezultati, ne klasificiraju podaci, ne određuju klasteri, ne modeliraju međuovisnosti podataka, niti se oni sumiraju, niti se otkrivaju promjene i odstupanja u odnosu na postavljene ciljeve, ne određuju korelacije između podataka, ne generiraju asocijativna pravila ili korelacije između slogova u bazama. Zadaća poslovne inteligencije je uporabom odgovarajućih softverskih alata dati odgovore na spomenute postavljene zadaće.

U sustavu poslovne inteligencije možemo izdvojiti tri podskupa softverskih alata (slika 1): Upiti i izvješća (Query and Report), sustavi za potporu odlučivanju (Decision Support Systems) i izvršni informacijski sustavi (Executive Information Systems).



Slika 1.: Softverski alati u sustavu poslovne inteligencije

Na slici je predstavljena veza između različitih softverskih alata u sustavu poslovne inteligencije. Na najnižoj razini su alati za ad hoc upite i izvješća. Oni korisniku daju detaljan uvid u proizvode, kupce, dobavljače i druge objekte unutar transakcijskih baza podataka. Posljednja dostignuća i uspjehe u području informacijskih sustava predstavlja OLAP (On Line Analytical Processing Tools) tehnologija. Njezin cilj je poboljšati odlučivanje na različitim organizacijskim i upravljačkim razinama. OLAP integrira upravljačke informacijske sustave (Management Information Systems – MIS), DSS i EIS i uvodi multidimenzijski pogled na podatke. Funkcionalne veze između informacijskih sustava i alata za odlučivanje jasno prikazuje slika 1.

Rad Coda ima za rezultat jasno razlikovanje analitičkih baza podataka od transakcijskih baza [1]. Razliku ćemo pokazati na primjeru aplikacije u obradi troškova (njihovo evidentiranje i analiza). Transakcijske baze podataka sadrže podatke kao što su: vrsta troška, datum, mjesto troška, odgovorna osoba i proračunsko polje koje predstavlja ukupan iznos troškova. U relacijskom modelu podataka ovi podaci su predstavljeni u obliku relacijskih tablica. Korisnik (menadžer) treba podatke u obliku koji se razlikuje od oblika u relacijskoj bazi podataka. Menadžer treba informaciju o ukupnim troškovima po mjesecu, relativnu promjenu troškova u tekućem u odnosu na isto razdoblje prethodne godine, strukturu troškova, prosječne mjesečne troškove. U tradicionalnim informacijskim sustavima informacije nisu izravno vodljive i u mnogim slučajevima nisu dostupne.

Srednji i gornji slojevi na slici predstavljaju analitičke alate koji obuhvaćaju otkrivanje znanja u bazama podataka (knowledge discovery in databases - KDD), modele znanosti o upravljanju (management science), modele operacijskih istraživanja (operational research models) te statističke modele.

Algoritmi za traganje po podacima (data mining algorithms) su dio procesa otkrivanja znanja. Izvršni informacijski sustavi, sustavi za potporu odlučivanju igraju ključnu ulogu u pretraživanju, pronalaženju i prikazivanju znanja tj. transformaciji podataka u poslovnu inteligenciju [2]. Ovo je dobro poznati DSS proces u kojem se podaci prikupljaju (prikupljaju se iz različitih izvora, filtriraju, pročišćavaju, transformiraju, konsolidiraju i pune u skladišta podataka), analiziraju, i transformiraju u informacije i pretvaraju zatim u znanje. U ovom procesu se uspostavlja dobro poznati lanac vrijednosti informacija ("information value chain") u kojem se podaci pohranjuju u transakcijske baze podataka i transferiraju u analitičke baze (multidimenzijske baze i skladišta podataka). Analitičke baze sadrže podatke u obliku koji je prihvatljiv za dalja pobljšanja i analizu. Ovi podaci tako završavaju kao informacije. Ključnu i vitalnu ulogu u ovom procesu ima OLAP-multidimenzijske baze podataka jer se njime osigurava analitičaru izravna veza s podacima (multidimenzijskim podacima) i modelima (znanost o upravljanju, operacijska istraživanja ili statistički modeli).

Često se postavlja pitanje u kakvom odnosu stoje OLAP i alati za traženje podataka (data minig)? Smatramo da su to različiti ali komplementarni alati. OLAP je dio šireg skupa alata za potporu odlučivanju. Tradicionalni alati za upite i izvješća, kao dio softverskih alata za poslovnu inteligenciju opisuju što je pohranjeno u bazi podataka. OLAP ide dalje. On koristi za pronalaženje odgovora i dodatno pitanje «Zašto su određene hipoteze istinite»?

Korisnik formira hipotezu o odnosima unutar baze ili skladišta podataka i provjerava je serijom upita o podacima. Npr. analitičar želi utvrditi čimbenike koji determiniraju ne vraćanje kredita odobrenog komitentima neke poslovne banke. On može pretpostaviti da su korisnici kredita s manjim dohotkom kreditno rizični i analizirati uporabom OLAP alata bazu podataka kako bi dokazao istinitost ili odbacio hipotezu. Drugim riječima OLAP - analitičar generira seriju hipotetičkih pravilnosti i relacija a upite prema bazi podataka koristi da dokaže pravilnosti ili ih odbaci.

Ako se broj varijabli koje analiziramo povećava na stotine ili tisuće onda je gotovo nemoguće pronaći dobru hipotezu i OLAP-om analizirati bazu podataka. OLAP pokazuje ograničenja. Ne možemo ga djelotvorno uporabiti jer teško pronalazimo hipotezu o odnosima između polja unutar relacijske tablice. Alati za traganje po podacima (data mining) upravo se koriste podacima u bazi podataka za otkrivanje pravilnosti i veza (patterna), zakonitosti umjesto da provjerava istinitost hipoteza. Analitičar koji želi identificirati rizične faktore za neotplaćivanje zajma (kredita) koristi data mining alate. Oni otkrivaju korisnike s visokim dugom i malim dohotkom ali se analizom mogu otkriti i novi faktori o kojima analitičar nije razmišljao. Upravo ta činjenica čini OLAP i data mining alate komplementarnim.

## 2. POSLOVNA INTELIGENCIJA I OTKRIVANJE ZNANJA U PODACIMA

S informatičkog stajališta najznačajniji dio softvera za kreiranje uporabljivih i korisnih informacija o kupcima, prodaji, proizvodima, dobavljačima, zalihama, likvidnosti itd. a koji nazivamo sustavom poslovne inteligencije, su alati za otkrivanje znanja u bazama podataka. Oni su dio sustava za potporu odlučivanju<sup>1</sup> i izvršnih informacijskih sustava. Danas pojam KDD<sup>2</sup> uključuje:

- a) traganje po podacima (data mining),
- b) integraciju sustava temeljenih na znanju i statističkih metoda,
- c) metodu strojnog učenja,
- d) metodu za otkrivanje veza među podacima (asocijativna pravila, algoritmi indukcije i sl.).
- e) integraciju sustava temeljenih na znanju i objektno orijentiranih sustava,
- f) interpretaciju otkrivenog znanja,
- g) evaluaciju otkrivenog znanja...

Možemo navesti niz veoma uspješnih primjera KDD kao što je sustav General Motorsa za otkrivanje kvarova u automobilskim nesrećama. Razvijaju se

- 
- 1 Vidjeli smo da su sustavi za potporu odlučivanju jedna od kategorija poslovne inteligencije zajedno s izvršnim informacijskim sustavima i sustavom upita i izvješća.
  - 2 Pojam KDD često se povezuje i s drugim nazivima kao što su arheologija podataka (data archeology), žetva informacija (information harvesting), obrada patterna (pattern processing), traganje (svrdlanje) kroz podatke (data mining), ekstarkcija znanja (data extraction)...

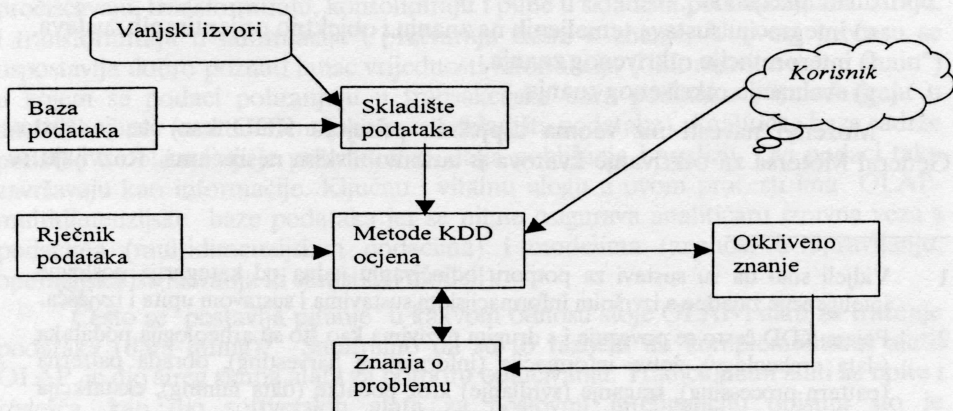
ekspertni sustavi za dijagnosticiranje. Drugi primjeri uspješne primjene su u bankarstvu glede odobravanja zajmova i kredita.

Discipline, izravno povezane s KDD-om, su baze podataka, ekspertni sustavi, agenti za pretraživanje, statistika, strojno učenje, skladišta podataka, OLAP alati.

Otkrivanje znanja je proces ekstrakcije informacija iz podataka koji su uporabljivi, korisni, implicitni i prethodno nepoznati. Informacije su po ispravnosti, zakonitosti ili jednostavno patternu razumljive korisniku. Pattern se definira kao iskaz (E) u jeziku (L) o odnosima između podskupa podataka ( $S_p$ ) i njegovog skupa S. Iskaz (E) nije potpuno siguran i uvijek se povezuje s određenim stupnjem neizvjesnosti i nesigurnosti pojave. Stupanj izvjesnosti ćemo označiti «c». Pattern se mora izraziti jasnim i jednostavnim jezikom, najčešće u obliku Ako...Onda.

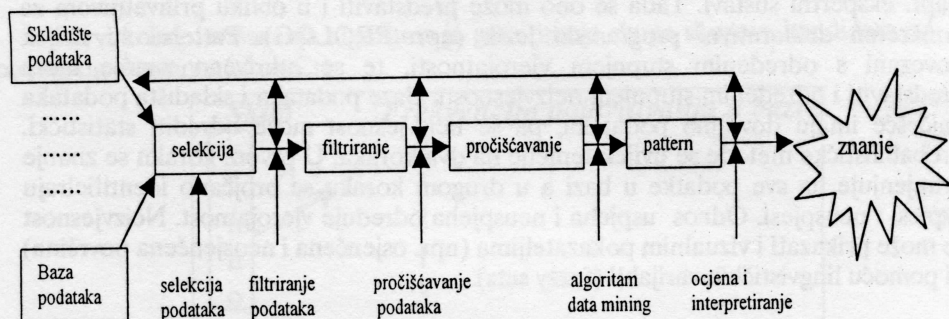
Npr. Ako je likvidnost «nezadovoljavajuća», onda je uzrok nenaplaćenog potraživanja od kupaca ( $c=0.8$ ). Ovi patterni mogu biti ulazi u bazu znanja ekspertnog sustava. Otkriveno znanje predstavlja tako izlaz iz programa koji pretražuje, analizira, ispituje podatke u bazama ili skladištima podataka. U ovim postupcima analize patterni se generiraju. Znanje, predstavljeno u obliku patterna, ovisi o kvaliteti analiziranih podataka i mora sadržavati dostatnu razinu izvjesnosti definiranu od strane korisnika. Bez dostizanja razine izvjesnosti definirane od strane korisnika pattern ne predstavlja znanje.. Pattern osim izvjesnosti, mora sadržavati i neka dodatna svojstva kao što su: netrivialnost, što znači da pattern nije jednostavno izvodljiv iz podataka, mora predstavljati novost za korisnika pa mora do tada korisniku biti nepoznat. Nov, koristan i primjenljiv - pattern iskazuje znanje mora biti primjenljiv i pomoći korisniku u rješavanju zadaća.

Proces otkrivanja znanja se onda temelji na algoritmima određene složenosti počevši od matematičkih i statističkih do heurističkih. Okvir procesa otkrivanja znanja čine: baza podataka, skladište podataka, rječnik podataka, domenska znanja o problemu, različite metode pretraživanja i otkrivanja pravilnosti podataka te sam korisnik. Interakcija ovih elemenata je prikazana na sljedećoj slici:



Slika 2.: Okolina KDD kao dijela poslovne inteligencije

Podaci u bazama podataka su dinamični. Mijenjaju se u vremenu za razliku od skladišta podataka koje predstavlja vremenski ovisne i nepromjenljive podatke, predmetno orijentirane. U tom smislu otkrivanje znanja možemo opisati nizom međusobno povezanih i ovisnih aktivnosti. One predstavljaju istodobno i najvažnije postupke u sustavu poslovne inteligencije. Prva je selekcija podataka a zatim je slijede njihovo filtriranje i pročišćavanje. Rezultat selekcije su ciljni podaci koje moramo dodatno transformirati kako bismo mogli primijeniti algoritme traganja po podacima. Takav sekvencijalan niz aktivnosti s povratnim vezama prikazan je na slici 3.



Slika 3.: Aktivnosti procesa otkrivanja znanja u bazama podataka kao dio sustava poslovne inteligencije

Prvu aktivnost nazivamo selekcija (izbor) podataka. Rezultat su ciljni podaci koje selektiramo iz skladišta podataka ili baze podataka. Podaci su najčešće u obliku tablica relacijskog modela. U selektiranju se koncentriramo na podskup podataka iz baze podataka na temelju kojeg će se izvršiti KDD. Naravno da ovo nije jednostavna aktivnost. Ona pretpostavlja prethodno istražene ciljeve korisnika i rječnika baze podataka. Filtriranje, zatim pročišćavanje podataka imaju za cilj ukloniti potencijalnu inkonzistenciju podataka na koje će se primijeniti algoritam traganja po podacima (data mining). Izbor algoritma ovisit će o postavljenom zadatku. Postavljeni zadatak je najčešće klasifikacija, modeliranje zavisnosti, otkrivanje promjene ili odstupanja u odnosu na postavljeni cilj, određivanje klastera. Klasifikacija označava pripadnost podataka (slogova) nekoj od unaprijed određenih klasa. Modeliranje zavisnosti označava opisivanje međuovisnosti podataka. Ovisnost postoji ako se može prognozirati promjena jednog podatka u zavisnosti od promjene drugog podatka ili formalno  $A \rightarrow B$ . Promjene ili odstupanja se mjere uvijek u odnosu na neku ciljnu veličinu. Određivanje klastera je postupak formiranja klasa na temelju podataka. Razlika između klasifikacije i klasteriranja je u tome što kod klasteriranja klase nisu unaprijed poznate. Izbor metoda otkrivanja patterna te modela i parametra je kritičan korak. Algoritam mora biti u skladu s korisnikovim ciljevima. Patterni su izlazi iz algoritma ili modela. Oni najčešće opisuju relacije između više atributa (polja) i odnose se na isti slog ili na grupe

slogova i njihove međusobne odnose. Tako mogu pokazivati trendove ili identificirati klastere kojima slogovi pripadaju. Patterni mogu biti kvalitativni ili kvantitativni. Kvalitativni opisuju logičke operacije između polja, a kvantitativni matematičke jednadžbe između polja (atributa). Interpretacija i evaluacija patterna označava tumačenje otkrivenih patterna korisniku, generiranje njegova prikaza u skladu sa zahtjevima korisnika. U procesu generiranja patterna često se moramo vraćati na prethodne korake, tako da se pojedini patterni mogu zanemarivati i eliminirati kao nebitni. Otkriveno znanje je ovisno o namjeni. Čovjek znanje najbolje razumije u obliku prirodnog jezika, grafike, slika ili formalne logike. Znanje dobiveno u ovom procesu se može koristiti i u drugom okruženju kao što su napr. ekspertni sustavi. Tada se ono može predstaviti i u obliku prihvatljivom za konkretan deklarativni programski jezik (npr. PROLOG). Patterni su uvijek povezani s određenim stupnjem vjerojatnosti, te se otkriveno znanje mora predstaviti i određenim stupnjem neizvjesnosti. Baze podataka i skladišta podataka najčešće imaju dovoljno podataka, pa se neizvjesnost može odrediti statistički. Probabilističke metode se uvijek temelje na dva koraka. U prvom koraku se znanje primjenjuje na sve podatke u bazi a u drugom koraku se brojčano identificiraju uspjesi i neuspjesi. Odnos uspjeha i neuspjeha određuje vjerojatnost. Neizvjesnost se može prikazati i vizualnim pokazateljima (npr. osjenčena i neosjenčena površina) ili pomoću lingvističkih varijabli (fuzzy sets).

## 2.1 Alati za traganje po podacima

Algoritmi za traganje po podacima (data minig algoritmi) su konačni nizovi koraka koji pronalaze patterne iz podataka. Ti algoritmi identificiraju pattern i daju njegov opis. Najčešće su to algoritmi koji nisu samo specifični za data mining već preuzimaju i adaptiraju podatke iz drugih disciplina (npr. statistike). Patterni se mogu predstavljati i opisivati uporabom oblika Ako... Onda, stabilima odlučivanja, neuronskim mrežama, genetičkim algoritmima, statističkim metodama, linearnim i nelinearnim modelima. Danas ne postoji univerzalno dobar i prihvatljiv algoritam za traganje po podacima. Za ilustraciju algoritama data mininga predstaviti ćemo algoritam «apriori».

### 2.1.1 Algoritam apriori

Svrha algoritma apriori je pronalaženje asocijativnih pravila. To su pravila oblika  $A \rightarrow B$ . Prevedeno na relacijske baze podataka možemo reći da algoritam apriori izražava tvrdnju prema kojoj slog koji sadrži skup elemenata A sadrži i skup elemenata B uz uvjet da je njihov presjek prazan skup  $A \cap B = \emptyset$ . Za asocijativna pravila se vezuju faktori povjerenja i potpore. Faktor povjerenja  $c$  predstavlja postotak slogova koji sadrže A sadrže i B. Semantičko značenje faktora povjerenja je jačina implikacije  $A \rightarrow B$ . Faktor potpore «s» označava postotak slogova u bazi koja sadrži  $A \cup B$ . Pravila s visokim stupnjem povjerenja i potpore nazivaju se jaka asocijativna pravila. Algoritam Apriori se koristi za otkrivanje jakih skupova elemenata. Jaki skupovi elemenata su skupovi asocijativnih elemenata sa stupnjem potpore «s» iznad neke unaprijed zadane vrijednosti. Drugo važno pojmovno značenje ovog algoritma je nadovezivanje skupova koje označavamo operatorom  $\oplus$ . Za skupove skupova od jednog elementa  $S_1 \oplus S_2 = \{k, l, m\} \oplus \{k, l, m\} = \{kl, km, lm\}$ .



Bit algoritma Apriori prikazat ćemo na primjeru. Zadana je sljedeća baza podataka B:

Baza podataka

<i>Pk</i>	<i>A<sub>1</sub></i>	<i>A<sub>2</sub></i>	<i>A<sub>3</sub></i>	<i>A<sub>4</sub></i>	
01	k	m	n		Traži se faktor potpore $s = 0.5$ (50%).
02	l	m	o		
03	k	l	m	o	
04	l	o			

Prvi korak izvršavanja algoritma je generiranje skupa skupova kandidata od po jednog elementa (skup  $C_1$ ).

<i>Skup</i>	<i>Frekvencija pojavljivanja u bazi</i>
k	2
l	3
m	3
n	1
o	3

$C_1$  →

Drugi korak predstavlja generiranje skupa jakih skupova od po jednog elementa (skup  $L_1$ ).

<i>Skup</i>	<i>Frekvencija</i>	<i>Skup</i>
k	2	a
l	3	l
m	3	m
n	1	
o	3	o

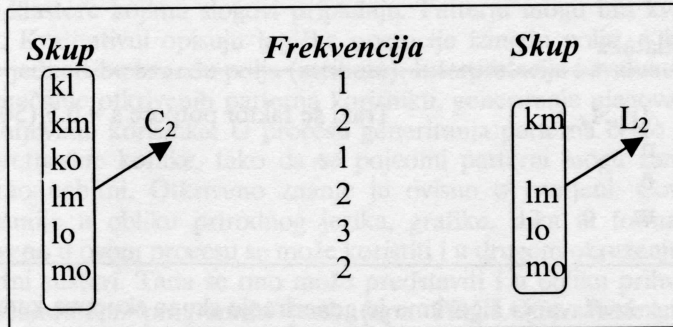
$C_1$  →       $L_1$  →

Treći korak je pravljenje skupa skupova kandidata od po dva elementa pomoću  $L_1 \oplus L_1$  (skup  $C_2$ ).

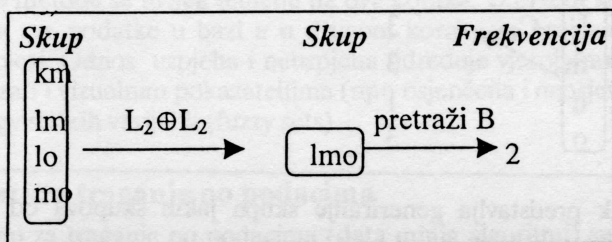
<i>Skup</i>	<i>Skup</i>	<i>Frekvencija</i>
k	kl	1
l	km	2
m	ko	1
o	lm	2
	lo	3
	mo	2

$L_1 \oplus L_1$  →      pretraži D →

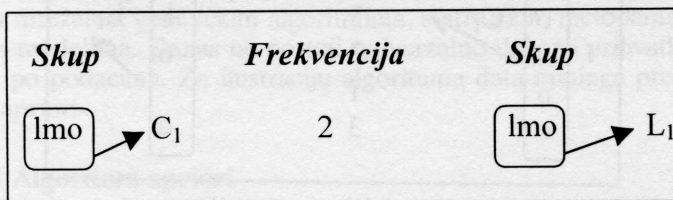
Korak 4 je generiranje skupa jakih skupova od po dva elementa (skup  $L_2$ ):



Peti korak je generiranje skupa skupova kandidata od po tri elementa pomoću  $L_2 \oplus L_2$  (skup  $C_3$ )



Šesti korak je generiranje skupa jakih skupova od po tri elementa (skup  $L_3$ )



Ukratko, algoritam apriori se temelji na formiranju jakih skupova tako da se njihov broj elemenata povećava za jedan. Zatim se pretražuje baza i filtriraju samo elementi koji zadovoljavaju kriterij potpore «s». U sljedećem koraku se skupovi nadovezuju, baza ponovo skenira, filtrira i formiraju se jaki skupovi. Skeniranje baze podataka moguće je smanjiti. Ali to nije predmet našeg istraživanja.

Postoje i drugi algoritmi koji imaju istu namjenu i zadatak kao i algoritam apriori. Njihov cilj je pronalaženje i određivanje asocijativnih pravila. Jedan od takvih algoritama je i DHP koji se u odnosu na apriori algoritam razlikuje samo po korištenju tzv. hash tablice. Ova tablica smanjuje veličinu skupa  $C_k$  a posebno veličinu  $C_2$ .

## ZAKLJUČAK

Ključni čimbenik koji određuje uspješnost primjene metoda poslovne inteligencije u organizaciji je kreiranje kvalitetnih i dostupnih podataka u kojima će se primjenjivati algoritmi za traganje po podacima. Skladišta podataka postaju tako neophodan uvjet razvitka poslovne inteligencije. Izlaz iz takvih algoritama su paterni u obliku iskaza koji sadrže odnose između podskupa podataka i njegovog skupa. Alati za traganje po podacima su dio šireg koncepta otkrivanja znanja u bazama i skladištima podataka. Poslovna inteligencija u informatičkom smislu pretpostavlja primjenu tri skupa alata: upita i izvješća (Query and Report), sustava za potporu odlučivanju (Decision Support Systems) i izvršnih informacijskih sustava (Executive Information Systems).

Poslovna inteligencija se promatra kao sustav koji integrira različitu softversku potporu koja osigurava rješenja iznad zbroja njihovih pojedinačnih mogućnosti. Zato su cilj i potreba razvijanja koncepta poslovne inteligencije u potpori poslovanju u potpunosti jasni.

## LITERATURA

- B. Markić, V. Bevanda, D. Tomić, The cube data model and business data processing, Varaždin, IIS, 2000.
- E.F-Codd.S.B Codd, C.T Salley, Providing On-Line Analytical Processing to User-Analysts: An IT Mandate, E.F. Codd and Associates, 1993.
- N.-S Koutsoukis, G. Mitra and C.Lucas, Adapting on-line analytical processing for decision modelling: the interaction of information and decision technologies, Decision support systems DSS, Vol. 26 (1999), Number 1. North Holland.
- R. Kimball, The data warehouse toolkit, Willey, 1996.
- Web site: <http://www.DataPerceptions.co.uk>

**Brano Markić, Ph. D.**

Associate Professor

Faculty of Economics, University of Mostar, Faculty of Tourism and Foreign Trade, Dubrovnik

**BUSINESS INTELLIGENCE AND BUSINESS DATA PROCESSING**

*This paper examines business intelligence in the context of business data processing. Business intelligence comprises a group of processes, activities, and projects through which useful and usable information is searched for in and generated from databases. From the information point of view, the main task of business intelligence is to give answers to analytical operation-related questions, such as: "Why are sales in market segment A reduced compared with the sales in the same period last year? What products do buyers most often purchase together with P<sub>1</sub> and P<sub>2</sub> products? Why is the average level of supplies increased in relation to the sales in market segment C? What turnover of the P<sub>1</sub> product is expected next year in May in market segment A," etc.*

*The software facilitating answer to these questions is the most crucial part of the business intelligence system, which can be divided into three categories: Queries and Reports, Decision-Making Support Systems, and Executive Information Systems. The paper examines the areas closely related to business intelligence, such as: discovery (retrieval) of knowledge from databases, search tools, OLAP tools as well as data storage, functioning as databases for analyses and generation of knowledge.*

**Key words:** *business intelligence, data mining, OLAP, data base, knowledge discovery (retrieval).*

**JEL classification:** C81, L86