

Doc. dr. Mirjana Pejić Bach

RUDARENJE PODATAKA U BANKARSTVU

DATA MINING IN BANKING

SAŽETAK: Cilj rada je pokazati da se primjenom tehnika rudarenja već postojeći podaci u bankama mogu koristiti za donošenje kvalitetnijih odluka, čime se poboljšava poslovni rezultat banaka. Analizom mogućih primjena rudarenja podataka u bankarstvu pokazalo se da se ovom metodom mogu pronaći vrijedne informacije koje pomažu poboljšanju poslovnog rezultata banke. U anketi koja je provedena u hrvatskim bankama pokazalo se da 48% banaka koristi rudarenje podacima, a 9% banaka planira koristiti tehnologiju u idućoj godini. U najvećem broju banaka metoda se koristi za analizu i procjenu rizičnosti.

KLJUČNE RIJEČI: rudarenje podataka, bankarstvo, poslovna inteligencija.

ABSTRACT: Goal of the paper is to demonstrate that data mining techniques could use existing data in banks for increasing quality of decision making, which in turn improves business performance of banks. Analysis of possible applications of data mining in banking has demonstrated that this method helps with finding useful information for improving business performance of the bank. Survey research showed that 48% of Croatian banks use data mining, and 9% of the banks plan to use technology in the next year. In most of the banks data mining is used for risk analysis and estimation.

KEY WORDS: data mining, banking, business intelligence

1. UVOD

Rudarenje podataka je nova metodologija kojom se otkrivaju vrijedni podaci u bazama podataka poduzeća. Metoda se naziva rudarenje podataka, jer se u velikim količinama podataka traže informacije koje „vrijede zlata“. U ovom radu prikazat ćemo metodologiju i najvažnije primjene rudarenja podataka u bankarstvu.

Rudarenje podataka može biti vrlo korisno u bankarstvu. Brojni su primjeri, a navest ćemo tipične. Banka može na temelju karakteristika klijenata predvidjeti koji klijenti bi mogli koristiti određene usluge, te usmjeriti marketinšku kampanju samo na taj uski segment čime se smanjuju troškovi i povećava lojalnost klijenata. S druge strane banka može na temelju podataka iz prošlosti odrediti karakteristike klijenata za koje je vjerojatno da neće moći vratiti kredit te na taj način smanjiti rizik svog poslovanja. Sigurno je da banka ima već razvijene metode za odabir klijenata koji bi mogli koristiti dodatne proizvode kao i metode za detekciju potencijalnih loših dužnika. Novost u rudarenju podataka je u tome što na temelju podataka iz prošlosti i sofisticiranih metoda radi to isto s većim uspjehom.

Cilj rada je pokazati da se primjenom tehnika rudarenja već postojeći podaci u bankama mogu koristiti za donošenje kvalitetnijih odluka čime se poboljšava poslovni rezultat banaka. Hipoteza rada zasnovana je na primjeni tehnike rudarenja podataka čime se može povećati kvaliteta poslovnih odluka i poboljšati poslovni rezultat banke.

2. ŠTO JE RUDARENJE PODACIMA?

Banke dnevno bilježe goleme količine podataka. Za svakog klijenta vode se podaci o računima, transakcijama po svakom računu, kreditnim obvezama te demografski podaci.

Ovi se podaci bilježe u transakcijske baze podataka koje su nužne za tekuće poslovanje. Transakcijske baze podataka općenito obavljaju tri opće funkcije: vođenje evidencije o obavljenim poslovnim događajima, generiranje dokumenata potrebnih u poslovanju i izvještavanje o stanju poslovnog procesa (Parker *et.al.*, 2004.).

Najbolje se uči iz iskustva, stoga su prije petnaest godina isprva znanstvenici, a kasnije i menadžeri shvatili da su transakcijske baze podataka bogat izvor znanja kojim se može poboljšati poslovanje poduzeća (Piatetski, 1991.). Postalo je jasno da poduzeća imaju mnogo podataka, a malo informacija i vrlo malo znanja o mnogim aspektima svoga poslovanja.

Međutim, transakcijske baze podataka su goleme¹. Zamislimo da menadžment banke želi utvrditi karakteristike klijenata koji su u prošlosti bili nelikvidni. Takvu informaciju može tražiti od informatičara u poduzeću koji, osim svoga redovnog posla, trebaju potrošiti

¹ Veličina transakcijskih baza podataka mjeri se u terabajtovima (TB). Jedan terabajt sastoji se od 1,024 gigabajta (GB). Ako uzmemo u obzir da jedan gigabajt u prosjeku sadrži 64,782 stranica word dokumenta (LexisNexis, 2004), onda jedan terabajt ima 66.336,768 stranica. Teško je uopće i shvatiti koliko je to podataka, a kamoli koristiti ih u poslovanju bez pomoći informacijske tehnologije.

dosta vremena da naprave traženi izvještaj. Kad je izvještaj na stolu menadžera, možda je već prekasno za donošenje odluka.

Metoda kojom se može povećati uspješnost korištenja transakcijskih baza podataka za poboljšanje poslovanja poduzeća naziva se rudarenje podacima. Mnogo je definicija rudarenja podataka, a istaknut ćemo sljedeće:

Rudarenje podacima je traženje vrijednih informacija u velikim količinama podataka.

Rudarenje podacima je istraživanje i analiza velikih količina podataka pomoću automatskih ili poluautomatskih metoda s ciljem otkrivanja smislenih pravilnosti.

Do prije nekoliko godina metoda se razvijala prvenstveno u znanstvenim krugovima. Tek je nedavno zaživjela u poduzećima, kada je postalo jasno da je korištenje rudarenja podataka neizbježno za stjecanje komparativne prednosti poduzeća. U hrvatskim velikim poduzećima korištenje rudarenja podataka doživjelo je pravi bum. Na konferencijama posvećenim poslovnoj inteligenciji i upravljanju znanjem predstavljaju se projekti hrvatskih vodećih poduzeća koji kao sastavni dio imaju rudarenje podataka.

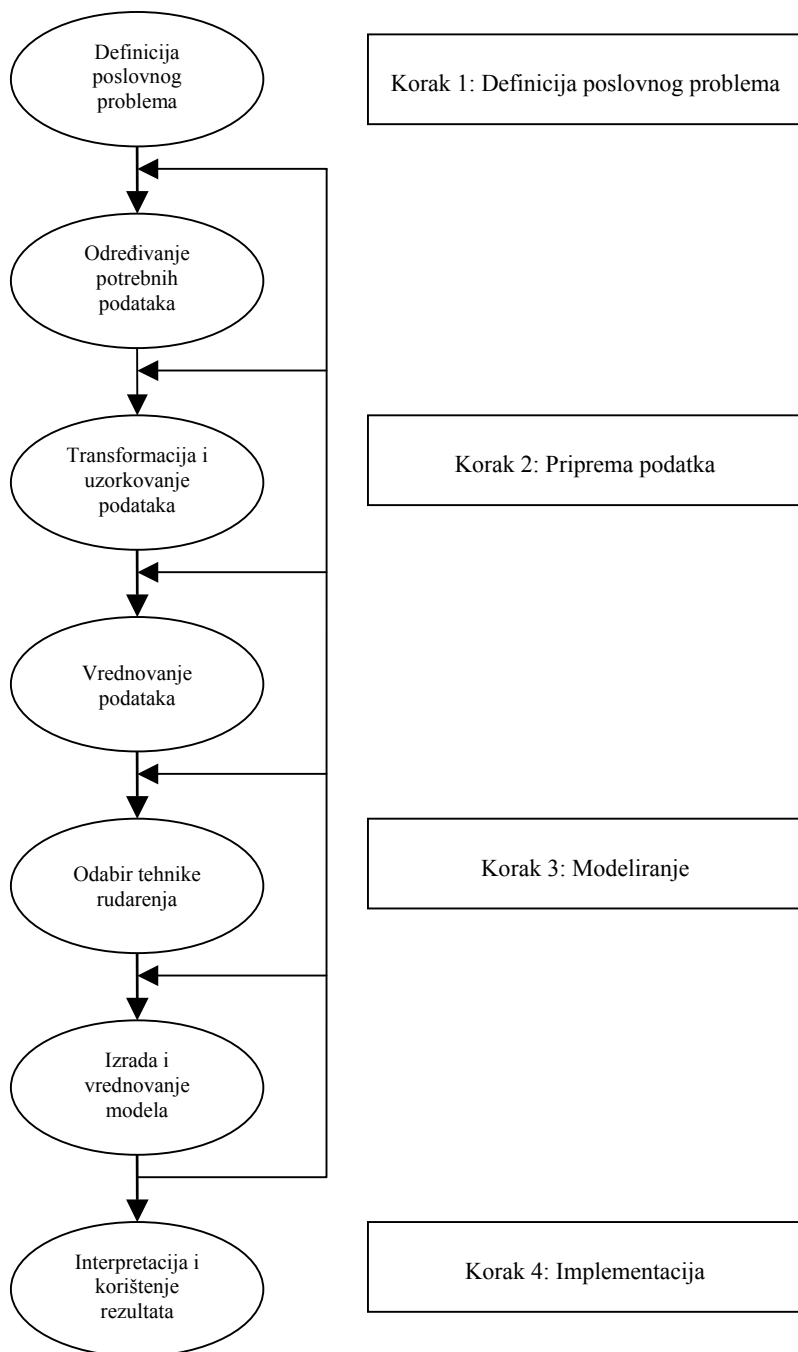
Hrvatske banke slijede ovaj trend. U anketi o korištenju naprednih informacijskih tehnologija u bankarstvu, provedenoj u rujnu 2004. godine, pokazalo se da 48% banaka koristi rudarenje podacima, a 9% banaka planira koristiti tehnologiju u idućoj godini. U najvećem broju banaka metoda se koristi za analizu i procjenu rizičnosti klijenata (Pejić Bach et.al., 2004.).

3. PROCES RUDARENJA PODACIMA

Važno je istaknuti da je rudarenje podataka više umjetnost nego znanost. Ne postoji recept za uspješno rudarenje podataka koje će sigurno rezultirati pronalaženjem vrijednih informacija. Međutim, vjerojatnost uspjeha će se povećati ako se slijede koraci procesa rudarenja podacima (Baragojin *et.al.*, 2001.), koji su prikazani na slici 1.

U prvom koraku se definira poslovni problem. Drugi korak je priprema podataka koja uključuje određivanje potrebnih podataka, transformaciju i uzorkovanje, te vrednovanje podataka. Modeliranje je treći korak, a obuhvaća odabir metode rudarenja te izradu i vrednovanje modela. Četvrti korak odnosi se na implementaciju koja obuhvaća interpretaciju i korištenje rezultata.

Proces rudarenja podataka je iterativan, što znači da se je u svakom trenutku moguće vratiti na neki od prethodnih koraka. Na primjer, u procesu odabira tehnike rudarenja možemo shvatiti da nismo dobro odabrali potrebne podatke, stoga se možemo vratiti na drugi korak i početi sve ispočetka. Ovakav „skok unatrag“ više će biti pravilo nego izuzetak, jer je u rudarenju podataka najvažnije dobro definirati problem te dobro odabrati i pripremiti podatke, što je teško „od prve“ napraviti kako treba. S druge strane, tijekom procesa rudarenja podataka povećava se naše znanje o poslovnom problemu i podacima, te je takva „revidirana“ definicija poslovnog problema često bolja od originalne. U nastavku ćemo detaljno opisati korake rudarenja podataka.



Slika 1. Proces rudarenja podataka

3.1. Definicija poslovnog problema

Rudarenje podataka je nova metodologija i često je u poduzećima prisutan stav da u bazama podataka postoje vrijedni podaci koje treba otkriti, pri čemu se ne definira koji bi to podaci bili niti za koju svrhu.

Ako se vratimo analogiji kopanja zlata koja je spomenuta na početku rada, ovaj pokušaj možemo usporediti s traženjem zlata na bilo kojem mjestu. Vjerojatnost pronalaženja zlata je znatno veća ako koristimo suvremeni alat. Međutim, moramo znati na kojem mjestu treba kopati.

Prvi korak u procesu rudarenja podataka je definicija poslovnog problema te izražavanje tog problema u obliku pitanja na koja se na završetku procesa može odgovoriti. Najbolji pristup definiranju poslovnog problema je analiziranje područja gdje je rudarenje podataka već uspješno korišteno. Nakon što se dobro upoznamo sa uspješnim primjenama rudarenja podataka, možemo odabrati područje koje je najkritičnije za naše poduzeće.

Na primjer, cilj poslovanja može biti povećanje prodaje životnog osiguranja. Moguće je napraviti model kojim se predviđa koji će klijenti banke kupiti životno osiguranje. Banka svoje marketinške napore može usmjeriti samo na te klijente, čime se smanjuju troškovi, ali se i povećava učinkovitost marketinške kampanje. U sljedećem poglavlju opisat ćemo detaljno tipične primjene rudarenja podataka u bankarstvu.

U ovom koraku određuje se koje će osobe sudjelovati u projektu rudarenja podacima. To će tipično biti specijalist za rudarenje podacima, informatičar koji dobro poznaje baze i skladišta podataka banke, te bankarski stručnjak koji je dobro upoznat s potencijalnom primjenom. Važno je da se na čelu tima nalazi ključna osoba iz menadžmenta koja ne mora direktno raditi na projektu, ali ga treba podržati te pomoći u rješavanju eventualnih teškoća. Još uvijek u hrvatskim bankama postoji otpor prema primjeni novih tehnologija, premda se stanje u posljednje vrijeme popravlja, što je pokazalo i ovo istraživanje. Uloga osobe iz menadžmenta je upravo u tome da pomogne u prevladavanju ovih otpora.

3.2. Priprema podataka

Priprema podataka obuhvaća određivanje potrebnih podataka, transformaciju i uzorkovanje te vrednovanje podataka. Ova faza je vremenski najzahtjevnija, a obuhvaća 60-90% vremena potrebnog za rudarenje podataka (Pyle, 1999.). Podaci za rudarenje mogu biti pohranjeni u različitim oblicima, od kojih su najčešći relacijske baze podataka ili skladišta podataka. Može se raditi o operacionim sustavima, kao što su POS, bankomati, telefonski razgovori, web serveri i slično. Također se mogu koristiti podaci prikupljeni istraživanjem tržišta ili eksterni izvori podataka – npr. Hrvatski registar obveza po kreditima (HROK), koji sadrži podatke o kreditnom opterećenju građana. Specijalist za rudarenje podacima, informatičar, te bankarski stručnjak zajedno određuju koji podaci će biti potrebni za izradu modela.

ODREĐIVANJE POTREBNIH PODATAKA

Drugi korak u rudarenju podataka je određivanje podataka koji će se koristiti za izradu modela. Podaci koji se tipično koriste za rudarenje podataka pohranjeni su u obliku transakcijske baze podataka te baze klijenata. Transakcijska baza bilježi podatke za svaku

transakciju, pri čemu je njezin tipičan sadržaj sljedeći: šifra klijenta, broj računa, vrsta, iznos i datum transakcije. Baza podataka o klijentima tipično sadrži šifru klijenta, šifru kućanstva, broj računa, ime i prezime kupca, adresu, telefon, demografske podatke, proizvode i usluge, dosadašnje ponude i segmentaciju.

U ovom se koraku također određuje koje varijable će se izbaciti iz analize te koja će varijabla biti ciljna ili zavisna. Na primjer, ako se radi o analizi kreditnog rizika, ciljna varijabla će biti ona koja opisuje je li klijent vratio kredit ili ne. Konačan rezultat određivanja potrebnih podataka je popis varijabli koje će se koristiti u izradi modela.

TRANSFORMACIJA PODATAKA

U ovom se koraku varijable iz dostupnih baza podataka transformiraju u oblik pogodan za rudarenje podataka. Podaci moraju biti u tabelarnom obliku pri čemu se u stupcima trebaju nalaziti varijable (obilježja), a u recima opažanja. Svaki redak mora opisivati podatak značajan za poduzeće (npr. kupca, proizvod).

Često se podaci iz transakcijske baze podataka moraju sumirati da bi bili korisni, pri čemu se dosta koristi ukupan i prosječan mjesečni iznos transakcija po svim računima klijenta.

Na temelju dostupnih varijabli iz baza podataka računaju se varijable koje određuje bankovni stručnjak. Primjeri takvih varijabli su razlike između trenutnog stanja tekućeg računa i odobrenog minusa te vrijeme koje je proteklo između dana otvaranja računa i dana prve transakcije.

UZORKOVANJE PODATAKA

U transakcijskim bazama podataka i bazama klijenata nalaze se velike količine podataka. Za izradu modela nije potrebno toliko podataka, stoga se koristi uzorkovanje podataka kako bi se odabrala manja količina podataka potrebnih za model.

Ovdje se često postavlja pitanje: Koliko je podataka dovoljno? Nema jednoznačnog odgovora jer broj potrebnih podataka ovisi o algoritmu. Za izradu stabla odlučivanja dovoljno je imati dvije do tri tisuće podataka, ali za treniranje neuronskih mreža potrebno je mnogo više.

Podaci za uzorak izabiru se najčešće slučajnim izborom. Često se dogodi da je udio događaja koji se analizira u uzorku vrlo mali. Na primjer, ako želimo napraviti model koji će prognozirati vjerojatnost da će klijent kupiti neki proizvod, potrebna nam je baza podataka sa sličnim podacima iz prošlosti. U takvoj bazi od, na primjer 100.000 klijenata, može biti samo 4.000 klijenata koji su kupili proizvod. Na temelju njihovih karakteristika napraviti će se model. Za izradu modela nije nam potrebno 100.000 podataka već mnogo manje – na primjer 10.000. Međutim, ako se podaci izabiru slučajnim izborom, broj od 4.000 klijenata koji su kupili proizvod u uzorku će biti mnogo manji – oko 40. Preporučuje se stoga da se u takvim slučajevima u uzorak od 10.000 uzme svih 4.000 klijenata koji su kupili proizvod, a ostalih 6.000 klijenata odabire se slučajnim izborom. Pokazalo se da takav pristup daje pouzdanije rezultate (Scott et. al., 1986.).

Pošto je izabran uzorak za izradu modela, potrebno ga je dodatno podijeliti na dva dijela – dio podataka za izradu modela i dio podataka za testiranje modela. Takav pristup tipičan je za rudarenje podataka jer se na taj način provjerava efikasnost modela na podacima koji nisu korišteni za njegovu izradu.

VREDNOVANJE PODATAKA

Kod vrednovanja podataka potrebno je analizirati postojanje netipičnih vrijednosti i prljavih podataka.

Netipične vrijednosti javljaju se u svakoj bazi podataka, a radi se na primjer o klijentima s izrazito visokim ili izrazito niskim primanjima. Potrebno je odlučiti što napraviti s netipičnim vrijednostima. Moguće je napraviti analizu i s tim podacima, izbaciti podatke o klijentima koji imaju netipične vrijednosti, izbaciti iz analize varijablu koja ima mnogo netipične vrijednosti ili netipične vrijednosti zamijeniti s nekom vrijednosti – npr. minimumom, maksimumom ili prosjekom. Vrijednosti varijable mogu se također podijeliti u razrede (npr. niska, srednja, visoka primanja).

Prljavi podaci odnose se na nepostojeće vrijednosti, nejasne definicije podataka i netočne vrijednosti. Vrijedi pravilo da su svi podaci prljavi podaci, odnosno u svakoj bazi moguće je naći neku od ovih pojava. Nepostojeće vrijednosti česte su, a obično se radi o stanju da za klijenta nemamo neke demografske podatke – npr. dob. Potrebno je vidjeti može li se dob izračunati na temelju nekih drugih varijabli- npr. jedinstvenog matičnog broja građana. Ako to nije moguće, postupak s nepostojećim vrijednostima isti je kao i postupak za netipične vrijednosti. Nejasne definicije podataka česte su kod „prelaska“ podataka iz jedne baze podataka u drugu. U tom slučaju varijabla Cijena 1 u jednoj bazi može označivati cijenu kave, a u drugoj cijenu kapučina. Netočne vrijednosti najčešće su posljedica pogrešnog unosa podataka u računalo.

3.3. Modeliranje

Na početku procesa modeliranja često se izrađuje analiza profila klijenata, pri čemu se analiziraju odabrane karakteristike klijenata – npr. spol, dob, zanimanje, primanja i slično.

U procesu rudarenja podacima koriste se sljedeće metode: statistika, baze i skladišta podataka, umjetna inteligencija i vizualizacija. Postoji dosta dvojbe o tome što je zapravo rudarenje podataka, ali važno je znati da se metode rudarenja podataka mogu podijeliti u tri kategorije: otkrivanje, klasifikacija i predviđanje (Berry et. al., 2000.).

Metode otkrivanja odnose se na mnoge postupke koje traže pravilnost u podacima bez prethodnog znanja o njihovom obliku. Brojne su metode kojima se otkrivaju pravilnosti u podacima, a neke od njih su metode segmentiranja (engl. cluster analysis) i asocijativna pravila.

Metode za klasifikaciju varijabli koriste se za predviđanje kategorije – npr. hoće li klijent vratiti kredit ili ne. Za klasifikaciju se često koriste stablo odlučivanja, logit regresija te neuralne mreže.

Metode za predviđanje vrijednosti varijabli koriste se za predviđanje numeričkih vrijednosti – npr. iznosa koliko će novca klijent godišnje kupiti na temelju njegove dobi, zanimanja i dosadašnje potrošnje. Za predviđanje se također koriste neuronske mreže, ali i linearna regresija te metode vremenskih serija.

Odabir metode ovisi o cilju analize. Na primjer, ako je cilj analize otkrivanje novih segmenata na tržištu, koristit će se neka od metoda otkrivanja. Najčešće korištene metode rudarenja podataka su metode klasifikacije, a najrjeđe se koriste metode predviđanja vrijednosti.

U ovoj se fazi rijetko odlučuje samo za jednu metodu već se obično provjerava više metoda. Tek nakon njihove usporedbe na konkretnom uzorku podataka odabire se metoda rudarenja podataka.

Primjena ovih metoda bila bi teško moguća bez softvera za rudarenje podataka. Na tržištu postoje brojni besplatni i komercijalni softveri. Obje skupine softvera mogu sadržavati više metoda ili biti specijalizirane samo za jednu. Iscrpan popis softvera dostupan je na web mjestu www.kdnuggets.com

Nakon primjene metoda njihovi rezultati se vrednuju. Podsjetimo se da su podaci koji su izabrani u uzorak, podijeljeni na dva dijela - dio podataka za izradu modela i dio podataka za testiranje modela. Podaci za testiranje modela koriste se za vrednovanje metode, jer se na taj način provjerava efikasnost modela na podacima koji nisu korišteni za njegovu izradu.

3.4. Implementacija

Implementacija rezultata odnosi se na interpretaciju rezultata i njihovo korištenje. U ovoj fazi ključna je uloga bankarskog stručnjaka koji na temelju specifičnih bankarskih znanja može interpretirati rezultate. Budući da bankarski stručnjaci obično nisu i stručnjaci za rudarenje podataka, važno je da rezultati modela budu u jednostavnom obliku za interpretaciju – npr. u obliku grafikona ili pravila.

Korištenje rezultata ovisi o njihovom predstavljanju i integraciji u svakodnevno poslovanje. Što su rezultati bolje predstavljeni, to će se više koristiti. Dobro je također ako se modeli rudarenja podataka implementiraju u informacijski sustav poduzeća. Na primjer, model za predviđanje odlaska klijenata konkurenciji trebao bi se integrirati u bazu podataka klijenata u obliku varijable koja sadrži vjerojatnost odlaska klijenta. Isto tako, model za prodaju dodatnih proizvoda trebao bi se integrirati u bazu klijenata u obliku niza varijabli koje prikazuju proizvode koje bi kupac mogao vjerojatno kupiti.

4. PRIMJENE RUDARENJA PODACIMA U BANKARSTVU

4.1. Rizik

Model rizika tipičan je upravo za banke i osiguravajuća društva. Za banke je jako važno da ne daju kredit osobi za koju je vjerojatno da ga neće moći vratiti. Tada se koriste modeli rizika za predviđanje vjerojatnosti hoće li ili neće dužnik moći vratiti kredit. Ovakvi modeli mogu se koristiti za klasične kredite koji imaju neki oblik osiguranja (jamac, hipoteka, založno pravo), ali i za neosigurane kredite (revolving kreditne kartice, minusi po tekućem računu). Osiguravajuća društva imaju rizik da će klijent iskoristiti osiguranje, na primjer, zbog ozljede. Poznato nam je da vodeće hrvatske banke i osiguravajuća društva koriste u svom poslovanju modele predviđanja rizika.

Rizik od prijevare također je važan za banke i osiguravajuća društva. Kod krađe kreditnih kartica, banke preuzimaju na sebe dio štete. Modeli koji na temelju ponašanja kupaca

brzo detektiraju krađe kreditnih kartica smanjuju gubitak banke. Na primjer, rudarenjem podataka pokazalo se da broj velikih transakcija u vrlo kratkom vremenu raste nakon krađe kreditne kartice. Moguće je kod svake transakcije u sustavu autorizacije usporediti prosječan broj i vrijednost transakcija po danu te na taj način brže reagirati ako postoji sumnja u krađu.

Osiguravajuća društva imaju rizik da će, na primjer, klijent pokušati iskoristiti svotu osiguranja od imovine nakon podmetnutog požara. Moguće je izraditi model koji olakšava detekciju potencijalnih pokušaja prijevare te model koji predviđa koji klijenti će tek u budućnosti pokušati prevariti osiguravajuće društvo.

Case study: HFC Bank otkriva zlouporabe kreditnih kartica

HFC Bank izdaje kreditne kartice koje koristi preko 3 milijuna britanskih građana. U prosjeku mjesečno obrađuju 9 milijuna transakcija, od kojih je 2.500 zlouporaba kreditnih kartica. Otkrivanje takvih transakcija je izrazito teško, prevaranti kopiraju informacije s magnetskih vrpca kartica i zatim pomoću njih pokušavaju kupiti proizvode. Obično se radi o nizu brzih kupovina proizvoda visoke vrijednosti poput nakita ili elektronike.

Rudarenje podataka zainteresiralo je menadžment banke koji je angažirao tim od 5 analitičara. Korišteno je 60 varijabli za identificiranje zlouporaba. Pokazalo se da su neka prodajna mjesta izrazito pogodna za zlouporabe (npr. kupnja nakita u inozemstvu), stoga je banka svim transakcijama na takvim prodajnim mjestima dala veću pozornost.

U samo prvih devet mjeseci rezultati operacije bili su spektakularni. Iako je u 2003. godini predviđen rast prijevara u industriji kreditnih kartica za 20%, HFC Bank uspjela je „izbjeći“ negativan trend te smanjiti broj štetnih transakcija za 55% uz mjesečnu uštedu 220.000 USD. Ovi rezultati dvostruko su bolji od ostalih banaka. Prosječan gubitak u kartičnoj industriji nastao kao rezultat prijevara je 0,27% ukupnog prihoda, a gubitak HFC Bank je 0,10%.

Izvor: www.sas.com/success/index.html

4.2. Prodaja dodatnih proizvoda postojećim klijentima

Modeli prodaje dodatnih proizvoda postojećim kupcima određuju vjerojatnost da će klijent banke kupiti dodatni proizvod. Cilj ovakve analize nije samo povećati broj klijenata koji će odgovoriti na ponudu, čime se smanjuju troškovi. Ponudom odabranih proizvoda odabranim klijentima također se povećava kvaliteta odnosa s klijentima. Na taj način raste profitabilnost poslovanja, jer je trošak prodaje drugih proizvoda postojećim klijentima mnogo niži od privlačenja novih kupaca uz istodobno povećanje lojalnosti postojećih klijenata. Vodeće hrvatske banke u svom poslovanju već duže vrijeme koriste ovakve modele (Vrančić, 2001.).

Case Study: First National Bank povećava efikasnost marketinga

First National Bank jedna je od najjačih banaka u Africi s 3.2 milijuna klijenata. U posljednje vrijeme, s dolaskom stranih banaka, konkurencija se zaoštrila. Menadžment banke shvatio je da ne iskorištava dovoljno podatke koje ima o svojim klijentima. To se posebno očitovalo u loše koordiniranom marketingu zbog čega su jedni klijenti dobivali previše ponuda, a drugi premalo. Efikasnost marketinških napora uopće se nije mjerila.

Menadžment banke odlučio je koristiti rudarenje podataka s ciljem povećanja postojećih prihoda prodaje dodatnih proizvoda te dizajnirati nove proizvode prilagođene posebnih potrebama klijenata i tako stvoriti nove izvore prihoda.

U prvom koraku dizajniran je sustav koji prikuplja i pročišćava podatke o klijentima čime se sprječava višestruko slanje ponuda na isto ime kroz više različitih kanala (pošta, telefon, mail). Podaci iz sustava analizirani su i korišteni za dizajn ponuda prilagođenih individualnim klijentima na temelju njihovih dosadašnjih transakcija i usluga koje koriste. Klijenti su odlično reagirali na individualno prilagođene ponude koje su prihvatili u 9% slučajeva, što je odličan rezultat.

Najznačajnije se pokazalo otkriće da najprofitabilniji klijenti (gornjih 5%) nemaju značajan broj ključnih proizvoda banke. Dio najboljih klijenata do tada bio je dakle zapostavljen zbog lošeg informacijskog sustava. Menadžment banke računao je da će se investicija u rudarenje podataka isplatiti tek za 10 mjeseci, ali su već prve marketinške kampanje vratile ulaganje uz profit od 3000%.

Izvor: <http://www.sas.com/success/fnb.html>

4.3. Zadržavanje postojećih klijenata

Odlazak klijenata konkurenciji problem je mnogih djelatnosti. Zbog zasićenosti tržišta jedine mogućnosti rasta poduzeća svode se na preotimanje klijenata od konkurencije ili prodaju drugih proizvoda postojećim klijentima. Klijenti često prelaze konkurenciji zbog pogodnosti koje im se nude. Kartične kompanije u svijetu već godinama vode rat kamatama kako bi privukle klijente. Niske kamate vrijede u početnom razdoblju (npr. godina dana), a kartične kompanije nadaju se da će klijenti nastaviti koristiti karticu i nakon isteka pogodnosti. Međutim, u svijetu se pokazalo da dio klijenata vješto koristi niske kamate kod nekoliko kartičnih kompanija. Rudarenjem podataka izrađuju se modeli kojima se predviđa vjerojatnost da će klijent, nakon što se kamate podignu na normalnu razinu, prijeći konkurenciji ili će smanjiti potrošnju.

Case study: Dreyfus Corporation predviđa odlazak klijenata konkurenciji

Zbog zasićenosti tržišta investicijski fondovi postali su vrlo agresivni u privlačenju klijenata konkurencije. Agresivno se natječu u pridobivanju novih klijenata, dok u isto vrijeme ulažu velika sredstva kako bi zadržali postojeće. Mnogi fondovi koriste agresivne mjere prevencije, kao što je naplata visokih provizija klijentima koji povlače svoja sredstva prije utvrđenog roka ili poticajnim mjerama poput naplate manjih provizija lojalnim klijentima.

Dreyfus Corporation koristi rudarenje podataka za predviđanje vjerojatnosti o prelasku pojedinih klijenata konkurenciji. Na temelju demografskih podataka te podataka o postojećoj kreditnoj aktivnosti, potrošačkim navikama te transakcijama dizajniran je model za detekciju klijenata koji razmišljaju o odlasku.

Model reagira na znakove upozorenja, kao što su nagli porast ili pad broja kontakata s klijentom te povećan broj transakcija između fondova. Sustav Dreyfus Corporation predviđa potencijalni odlazak klijenta tri do šest mjeseci prije nego što se to doista dogodi s točnošću od 80 do 85%. S potencijalno problematičnim klijentom se telefonski razgovara, te mu se nude posebne pogodnosti prilagođene njegovoj dobi, primanjima, portfelju i nedavnim investicijskim aktivnostima.

Izvor: <http://www.sas.com/success/dreyfus.html>

4.4. Segmentacija

Glavni resurs banke su njezini klijenti. Tek na temelju poznavanja njihovih karakteristika, preferencija i specifičnih potreba banka im može prilagoditi ponudu svojih usluga. Pri tome treba uzeti u obzir da se karakteristike klijenata mijenjaju svakih nekoliko godina. Mlada osoba koja tek studira vjerojatno koristi samo jedan ili tek nekoliko bankarskih proizvoda. Ista osoba koristit će drugačije proizvode nakon što se vjenča, a drugačije nakon umirovljenja.

Velike količine podataka o klijentima banke bi trebale koristiti za analizu karakteristika klijenata te na temelju njih formirati segmente kojima se mogu posebno prilagoditi usluge. Banke već mnogo godina koriste tradicionalne segmentacije sektora stanovništva i poduzeća. Međutim, takva tradicionalna segmentacija često može „zamagliti“ stvarno stanje. Korištenjem rudarenja podataka mogu se pronaći segmenti koji su do sada bili zanemareni te im se mogu ponuditi specijalno prilagođeni proizvodi čime se povećava profitabilnost poslovanja, a ne zagovara se nužno napuštanje tradicionalne segmentacije.

4.5. Životna vrijednost klijenta

Životna vrijednost klijenta je očekivana vrijednost zarade od pojedinog klijenta kroz određeno razdoblje. Na primjer, banci je interesantno privući što više studentske populacije od koje će velik dio postati profitabilni klijenti. Trenutna zarada od usluga studentima može

biti mala, ali ako se stvori dobar odnos s klijentom, u budućnosti će se ostvariti velika korist. Tek diplomiranom studentu trebat će kredit za auto, stan, tekući račun, kreditne kartice, mirovinsko i životno osiguranje... Zbog visokog obrazovanja očekuje se da će takva osoba imati i natprosječna primanja te će si moći priuštiti sve te proizvode. Rudarenjem podataka izrađuju se modeli kojima se predviđa životna vrijednost klijenta kako bi bankarski službenici mogli posvetiti više pažnje klijentima koji nisu trenutno profitabilni, ali bi to mogli postati u budućnosti.

4.6. Odaziv

Model odaziva ima za cilj predvidjeti koji će kupci pozitivno odgovoriti na ponudu za kupovinu proizvoda ili usluge, pri čemu se obično radi o direktnom marketingu. Ponuda može biti odaslana različitim kanalima – poštom, telefonom, Internetom. Cilj ovog modela može biti privlačenje novih kupaca, ali i starih kupaca s kojima duže vrijeme ne poslujemo. Takve će kupce često biti lakše pridobiti na kupovinu proizvoda od novih kupaca.

4.7. Aktivacija

Modeli aktivacije predviđaju vjerojatnost da će klijent kojega smo pridobili, postati profitabilan. Na primjer, klijent često sklopi ugovor za životno osiguranje nakon čega ne uplaćuje premiju. Isto se događa s kreditnim karticama u bankarskom sektoru. Klijenti koji dobiju kreditnu karticu, postaju profitabilni tek kada počnu plaćati karticom. Dio klijenata to nikada ne učini te se modelom aktivacije može predvidjeti koji će to biti klijenti. Takvim se klijentima mogu ponuditi dodatne pogodnosti da bi ih se potaklo na aktivaciju ili se može odustati od poslovanja s njima.

4.8. Racionalizacija poslovanja

Rudarenjem podataka moguće je racionalizirati poslovanje na mnogo načina čime se ostvaruju znatne uštede. Opisat ćemo nekoliko primjera. Punjenje bankomata veliki je izazov za banke u organizacijskom i logističkom smislu. Tehnički maksimalna količina novca koju bankomat može držati vrlo je velika, ali iz ekonomske perspektive nema smisla puniti sve bankomate maksimalnim iznosom jer je prosječan dnevni promet na bankomatima mnogo manji. Također, novac na bankomatima ne donosi nikakvu kamatu. Rudarenje podataka može se koristiti u svrhu izrade sustava za optimizaciju upravljanja gotovinom koji bi predviđao kada i koliko novca treba isporučiti na koje mjesto te bi pri tome uzimao u obzir tjedne, mjesečne i godišnje oscilacije. Rudarenje podataka može se koristiti za izradu modela koji daje preporuku što treba napraviti kad klijent kasni obročnim plaćanjem kredita ili premije osiguranja – čekati, telefonirati, slati opomenu ili pokrenuti ovrhu. Nekoliko banaka koristilo je rudarenje podataka u organizaciji zaštitarske službe banke. Modelom su analizirane karakteristike poslovnica koje su opljačkane u posljednje vrijeme, stoga je u poslovnicama sličnih karakteristika organizirano jače osiguranje.

5. ZAKLJUČAK

Cilj rada je pokazati da se primjenom rudarenja podataka može povećati kvaliteta poslovnih odluka, čime se poboljšava poslovni rezultat banke. Prikazan je proces rudarenja podataka u bankarstvu, koji se sastoji od nekoliko faza: definicije poslovnog problema, pripreme podataka, modeliranja i primjene modela. Deks top istraživanjem identificirane su najčešće primjene rudarenja podataka u bankarstvu, te je prikazana jedna primjena rudarenja podataka na cross sellingu bankarskih usluga sektoru poduzeća. Anketno istraživanje primjene naprednih informatičkih tehnologija u bankarstvu pokazalo je da hrvatske banke slijede ovaj trend te da se 48% banaka koristi rudarenjem podataka, a 9% banaka planira koristiti tehnologiju u tijeku iduće godine.

LITERATURA

1. Baragoin, C., Andersen, C.M., Bayerl, S., Bent, G., Lee, J., Schommer, C. (2001). Mining Your Own Business in Banking Using DB2 Intelligent Miner for Data. Dostupno na: <http://www.redbooks.ibm.com/>
2. Berry, M.J.A., Linoff, G.S. (2000). Mastering Data Mining. Wiley, Chichester.
3. LexisNexis (2004). How many pages in a gigabyte? Dostupno na: http://www.lexisnexis.com/applieddiscovery/lawlibrary/whitePapers/ADI_FS_PagesInAGigabyte.pdf
4. Parker, C., Case, T. (1998). Management Information Systems: Strategy and Action. McGraw Hill, New York.
5. Pejić Bach, M., Bosilj Vukšić, V., Ćurko, K., Požgaj, Ž., Vlahović, N. (2004). Informatički aspekti primjene upravljanja znanjem i poslovnim procesima u bankarstvu. Interni projekt Ekonomskog fakulteta u Zagrebu.
6. Piatetski-Shapiro, G., Frawley, W. (1991). Knowledge Discovery in Databases. MIT Press.
7. Privredni Vjesnik (2004). Leasing u Hrvatskoj. Privredni Vjesnik, prosinac, 2004.
8. Pyle, D. (1999). Data Preparation for Data Mining. Morgan Kaufmann, San Francisco.
9. Scott, A.J., C. J. Wild. (1986). Fitting Logistic Models under Case-control or Choice Based Sampling. Journal of the Royal Statistical Society Series B 48(2),170-82.
10. Vrančić, I. (2001). Metode raspoznavanja uzoraka za analizu poslovno-financijskih podataka. Magistarski rad. Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva.