



Sveučilište u Zagrebu
EKONOMSKI FAKULTET
Zagreb - Hrvatska

Trg J. F. Kennedyja 6
10000 Zagreb, Hrvatska
Telefon +385(0)1 238 3333
<http://www.efzg.hr/wps>
wps@efzg.hr

SERIJA ČLANAKA U NASTAJANJU

Članak broj 10-08

Sandro Juković

Segmentacijska analiza poslovnih klijenata banaka pomoću samo-organizirajućih mapa



SVEUČILIŠTE U
ZAGREBU



Segmentacijska analiza poslovnih klijenata banaka pomoću samo- organizirajućih mapa

Sandro Juković
sjukovic@gmail.com

Sve izneseno u ovom članku u nastajanju stav je autora i ne odražava nužno službena stajališta Ekonomskog fakulteta u Zagrebu. Članak nije podvrgnut uobičajenoj recenziji.

Članak je izrađen na temelju diplomskog rada pod nazivom
«Segmentacijska analiza poslovnih klijenata banaka pomoću samo-organizirajućih mapa»
izrađenim pod mentorstvom prof. dr. sc. Mirjane Pejić Bach.

Copyright 2010 by Sandro Juković

Sva prava pridržana.

Dijelove teksta dopušteno je citirati uz obavezno navođenje izvora.

Sažetak

Samo-organizirajuće mape (SOM) su dvoslojne umjetne neuronske mreže koje su inicijalno kreirane za rješavanje problema klaster analize, vizualizacije i apstrakcije podataka. Njihov najveći doprinos je u području vizualizacije više-dimenzionalnih podataka na dvo-dimenzionalnu mapu, koja odražava eventualne veze među ulaznim podacima. Cilj ovog rada je prezentirati teorijsku osnovu SOM-a i prikazati primjenu metode u svrhu segmentacije tržišta.

U radu je objašnjen algoritam SOM-Ward koji je implementiran u softveru Viscovery SOMine. Tada je u istom softveru provedena klaster analiza prema anketi poslovnih klijenata banaka. Nakon toga su prikazani i interpretirani rezultati te analize kao tri segmenta. Segmenti se razlikuju s obzirom na attribute trgovinskog poslovanja s inozemstvom (uvoz/izvoz), godišnje prihod, podrijetlo kapitala, stavove o odabiru kredita, planove zapošljavanja itd. Tako kreirani segmenti mogu biti korišteni za daljnje odlučivanje o poduzimanju marketinških aktivnosti.

Ključne riječi

samo-organizirajuće mape, SOM, neuronska mreža, klaster analiza, segmentacija tržišta, rudarenje podataka

JEL klasifikacija

C81

1. UVOD

U članku će biti prikazani rezultati segmentacijske analize poslovnih klijenata banaka u Republici Hrvatskoj pomoću metode samo-organizirajućih mapa (u daljnjem tekstu SOM). Neuronske mreže su model bioloških neuronskih mreža ljudskog živčanog sustava. To je zapravo, paralelan spoj međupovezanih (svaki sa svakim) neurona koji su organizirani u slojevima (Kumar i Ravi, 2007). SOM, za razliku od većine mreža koje imaju tri i više slojeva, imaju samo dva sloja (ulazni i izlazni sloj – mapu).

Glavno svojstvo umjetnih neuronskih mreža, kao i SOM-a, je svojstvo učenja iz podataka koji se unose u ulazni sloj mreže (Kohonen, 1995). Tako neke složene interakcije analiziranih podataka mogu biti modelirane bez potrebe za poznavanjem prirode pojave koju ti podaci opisuju. Drugim riječima, nije potrebno kreirati algoritam koji će te podatke analizirati, već neuronska mreža kroz trening nauči kako analizirati podatke.

Svojstvu učenja SOM-a treba dodati još dva svojstva, to su redukcija dimenzionalnosti analiziranih podataka i smještanje sličnih podataka na susjedne čvorove na mapi. Redukcija dimenzionalnosti omogućuje lakše razumijevanje međuodnosa atributa podataka, dok grupiranje sličnih podataka omogućuje klasteriranje podataka (Kiang i Fisher, 2008) Tako je moguće primijeniti ovu metodu u području klaster analize.

Klaster analiza (grupiranje podataka) predstavlja osnovnu metodu otkrivanja znanja iz baza podataka kojom je moguće izvoditi segmentiranje tržišta. Tehnike segmentiranja podataka spadaju u grupu tzv. „nenadziranih metoda“. Cilj ovih metoda je otkrivanje globalne strukture podataka i pri ovoj metodi nije definiran ciljni atribut pa ne postoji razlika među atributima (Lešković, 2007).

Segmentacija poslovnih klijenata banaka u RH će biti prikazana nakon određivanja kriterija segmentacije ovisno o ponašanju, demografskim, osobnim, operativnim, situacijskim i kriterijim vezanima za pristup nabavi (Lešković, 2007). Svrha provođenja ove segmentacije je u smanjenju troška marketinških aktivnosti u bankarskom sektoru.

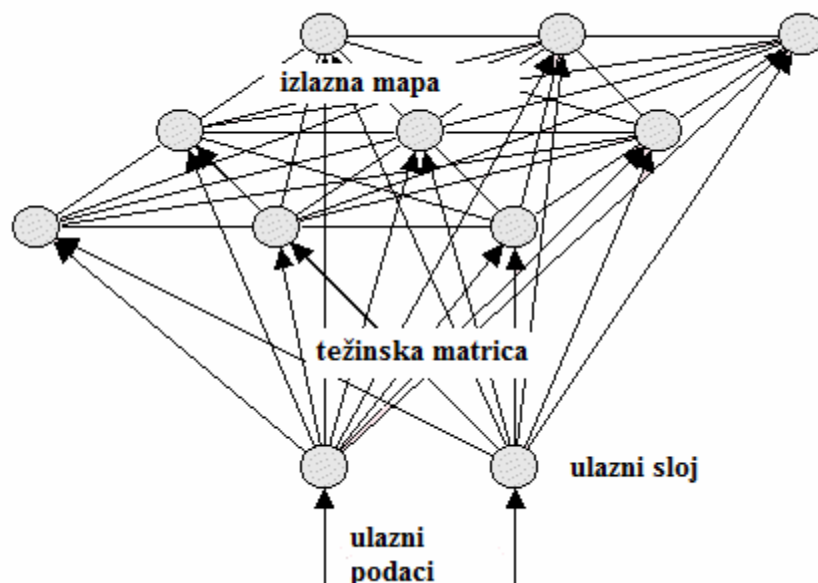
Rad je podijeljen u nekoliko dijelova. Prvi dio govori o teoretskim osnovama SOM metode, zatim slijedi dio o segmentaciji klijenata koji navodi odabrane kriterije segmentacije. Dalje su prikazani rezultati analize, a nakon toga interpretacija rezultata. Na kraju se daje napomena gdje se rezultati mogu koristiti te se predlaže način daljnjeg korištenja ove metode i ideje za daljnja istraživanja.

2. SOM METODA I SOM-WARD ALGORITAM

Samo-organizirajuće mape su utemeljene na konceptu neuronskih mreža i zapravo se mogu definirati kao dvoslojna neuronska mreža.

Neuronske mreže su model bioloških neuronskih mreža ljudskog živčanog sustava. To je zapravo, paralelan spoj međupovezanih (svaki sa svakim) neurona koji su organizirani u slojevima. Neuronske mreže se u praksi, između ostalog, koriste u financijsko uslužnoj djelatnosti (Kumar i Ravi, 2007). Također, treba napomenuti da svaka veza među neuronima ulazno-izlaznog sloja ima svoju pripadnu težinu (ponder) (Huysmans et al., 2006). To se može dobro vidjeti na slici 1 gdje je prikazan koncept samo-organizirajućih mapa.

Slika 1: Koncept modela samo-organizirajućih mapa



Izvor: Interactive Tutorial on Neural Networks (Jevtić, n. d.)

Pravilo ažuriranja težina pojedinih veza se zove *winner-take-all* pristup, to je temeljno pravilo kompetitivnog učenja koje nam govori da se čvoru koji je najbliži po vrijednostima atributa ulaznom vektoru ažurira težina veze. U konkretnom algoritmu koji se koristi, osim pobjedničkog čvora, ažuriraju se i njemu neposredno susjedni čvorovi (Lee et al., 2005).

Razni osjetilni podražaji kod ljudi su neurološki mapirani u mozgu tako da prostorni odnos među podražajima određenih dijelova mozga odgovara prostornom odnosu među neuronima organiziranim u dvodimenzionalnu mapu (Kohonen, 1995). Tako lokacija točaka na mapi pokazuje relativnu sličnost točaka u višedimenzionalnom prostoru, tj. što su dva sloga podataka sličnija, to će bliže biti prikazani na dvodimenzionalnoj mapi. U skladu s time, primarna funkcija SOM-a je prikazati višedimenzionalne podatke sa ulaza u mrežu, na dvodimenzionalnu mapu, pritom zadržavajući odnose među podacima (Kiang M. Y., 2001).

Da bi se razumjelo algoritam koji koristi softverski alat Viscovery SOMine, treba se napomenuti nekoliko činjenica o Ward metodi klasteriranja. Klasična metoda Ward klasteriranja spada u hijerarhijsko aglomeracijske algoritme koji imaju slijedeće karakteristike: počinju s klasteriranjem gdje svaka jedinica (čvor) tvori jedan klaster, a u svakom koraku klasteriranja se spajaju dva klastera s minimalnom udaljenosti (ovisno o mjeri udaljenosti definiranoj u algoritmu). Minimalna udaljenost se naziva nivo udaljenosti (engl. *distance niveau*) (Viscovery, 2009).

Dakle, svrha ovog postupka klasteriranja je minimizirati varijancu unutar pojedinih klastera, a maksimizirati varijance među klasterima kako bi se klasteriranje što kvalitetnije napravilo.

Smisao Ward klasteriranja se najlakše može razumjeti ako se promatra univarijantne podatke (samo jednu varijablu). Npr. ako imamo slijedeće vrijednosti neke pojave (2, 6, 5, 6, 2, 2, 2, 2, 0, 0, 0). Gubitak informacija koji bi bio rezultat tretiranja ovog skupa podataka kao jednog klastera sa srednjom vrijednošću od 2.5 bi bio jednak:

$$ESS_{\text{OneGroup}} = (2-2.5)^2 + (6-2.5)^2 + \dots + (0-2.5)^2 = 50.5$$

gdje je ESS suma kvadrata pogrešaka (engl. *Error Sum-of-Squares*) (XLMiner dokumentacija, 2010).

S druge strane, ako se ove vrijednosti klasteriraju u slijedeća četiri klastera:

{0,0,0}, {2,2,2,2}, {5}, {6,6}.

Tada će gubitak informacija izražen preko sume kvadrata pogrešaka iznositi:

$$ESS_{\text{OneGroup}} = ESS_{\text{Group1}} + ESS_{\text{Group2}} + ESS_{\text{Group3}} + ESS_{\text{Group4}} = 0.0$$

To znači da je klasteriranje podataka na četiri klastera u potpunosti minimiziralo varijance pojedinih klastera i tako u punom iznosu umanjilo gubitak informacija koji je postojao prije postupka klasteriranja.

2.1. SOM-Ward algoritam u Viscovery SOMine-u

Što se tiče ovog rada, softver pomoću kojeg se podaci analiziraju (Viscovery SOMine) ima implementirana tri algoritma za potrebe klasteriranja i vizualizacije podataka: SOM Single-linkage, SOM-Ward algoritam te Ward algoritam (Viscovery, 2009). Pošto je odlučeno u radu koristiti SOM-Ward algoritam, treba ga objasniti pobliže.

SOM-Ward algoritam je tzv. hibridni algoritam koji je nastao na temelju primjene paradigme *soft computinga*. Potonji termin je u upotrebu ušao u ranim 1990.-ima kako bi se odvojio od konvencionalne *hard computing* metode (Kumar i Ravi, 2007). Što se tiče *soft computing* paradigme, glavni razlog za hibridizaciju različitih inteligentnih tehnologija (metoda, algoritama) je u shvaćanju da su te metode prije komplementarne nego konkurentne u vidu karakteristika tehnologija kao što su efikasnost, tolerancija na pogreške/nepreciznosti i učenje iz primjera. Dalje, hibridni algoritmi, koji su proizašli iz ovakve filozofije, su minimizirali nedostatke i maksimizirali prednosti individualnih tehnologija korištenih u njihovoj konstrukciji (Kumar i Ravi, 2007). Jednostavnije rečeno, pomoću hibridizacije inteligentnih tehnologija postižu se sinergijski učinci u vidu efikasnosti, preciznosti i sposobnosti učenja softvera za analizu podataka.

U SOMine-u, matrica udaljenosti se inicijalizira na način da uzima u obzir broj slogova podataka koji odgovaraju čvorovima na mapi. Čvorovi s mnogo pripadajućih slogova podataka imaju veći ponder od čvorova koji imaju manji broj veza.

Mjera udaljenosti koja se koristi u algoritmu je modificirana Wardova udaljenost jer je vrlo vjerojatno da mapa sadrži „prazne“ čvorove (Viscovery, 2009):

Ako su r i s dva čvora između kojih se mjeri udaljenost, a n_r i n_s broj slogova podataka koji pripadaju čvorovima r i s , te \bar{X}_s i \bar{X}_r ulazni vektori. Tada je udaljenost d_{rs} definirana na slijedeći način:

$$d_{rs} := \begin{cases} 0 & \text{ako je } n_r = n_s = 0, \\ \frac{n_r \cdot n_s}{n_r + n_s} \cdot \|\bar{X}_r - \bar{X}_s\|^2 & \text{u ostalim slučajevima.} \end{cases}$$

Ova definicija osigurava da se tijekom prvih postupaka spajanja klastera spajaju oni čvorovi koji su prazni (npr. $n_r = 0$), sve dok ne ostanu samo klasteri koji sadrže slogove podataka ($n_r > 0$). Pošto u svakom koraku spajanja klastera postoji velik broj klastera kandidata, koristi se pravilo da se uzmu u obzir ona dva koji su najbliži na mapi.

U skladu s tim, za SOM-Ward mjera udaljenosti je prilagođena na slijedeći način (Viscovery, 2009):

$$d'_{rs} = \begin{cases} d_{rs} & \text{ako su klasteri } r \text{ i } s \text{ susjedi na SOM-u,} \\ \infty & \text{u ostalim slučajevima.} \end{cases}$$

Prema tome, SOM-Ward provjerava lokacije klastera na mapi, ako dva klastera nisu blizu na mapi, njihovo se spajanje čak ni ne razmatra.

Kako bi se odabralo najbolju kombinaciju klastera, SOMine izračunava indikatore kvalitete klastera (engl. *Cluster indicator*). Indikator se računa za svaki element ove hijerarhijske strukture. Indikator klastera se računa na slijedeći način: Serija Wardovih udaljenosti se izračuna za sve moguće brojeve klastera. Zatim se nivo udaljenosti normalizira eksponencijalnom funkcijom (Viscovery, 2009):

$$\mu(c) = d(c) \cdot c^\beta,$$

gdje je $d(c)$ Wardov nivo udaljenosti koji je korišten da se c klastera spoji u $c-1$ klastera, a β je koeficijent linearne regresije.

Odnos normaliziranih nivoa udaljenosti dva susjedna klastera tvori indikator klastera (Viscovery, 2009):

$$I'(c) = \frac{\mu(c)}{\mu(c+1)} - 1$$

Što je indikator klastera veći, to je proces klasteriranja kvalitetnije izvršen. Drugim riječima, razlike među pojedinim klasterima su veće. Ako je udaljenost ($d(c+1)$) među klasterima prije novog koraka velika, a nakon tog koraka ($d(c)$) mala, broj klastera koji je bio u prethodnom koraku se uzima kao optimalni broj klastera jer razlika među klasterima treba biti veća.

3. SEGMENTACIJA KLIJENATA

Segmentacija klijenata je podjela tržišta na homogene grupe potrošača uz osmišljavanje strategija i primjenu specifičnog marketinškog programa s namjerom zadovoljenja potreba odabranog segmenta.

U primjeni ovih marketinških strategija u financijskim institucijama to bi značilo da pružatelj financijske usluge pravi razlike među brojnim tržišnim segmentima, odabire jedan ili više njih i razvija usluge, marketinški miks, a kao njegov sastavni dio, i komunikacijski miks prilagođen svakom pojedinom segmentu. Ovakav način tržišnog ponašanja u teoriji je prepoznat kao ciljni marketing (Bevanda, 2008).

Više je načina segmentiranja tržišta, a najčešće rabljena metoda je socio-demografska segmentacija tržišta. Modeli neuronskih mreža omogućuju podjelu tržišta na određene homogene grupe ili segmente i kao takvi su se pokazali efikasnijima od linearnih metoda. Obje metode pružaju visok stupanj pouzdanosti (Dragičević, 2006).

Banke u Hrvatskoj koriste tradicionalne segmentacije sektora poduzeća, ali takva segmentacija često može zamagliti stvarno stanje. Stoga rudarenje podataka, tj. klaster analiza, može pronaći segmente koji su prije bili zanemareni. Tim segmentima se mogu ponuditi specijalno prilagođeni proizvodi čime banke mogu povećati profitabilnost poslovanja (Pejić-Bach, 2005).

To je glavna motivacija za provedbu segmentacije tržišta pomoću klaster analize. U daljnjem tekstu su navedeni kriteriji segmentacije poslovnih klijenata banaka.

3.1. Kriteriji segmentacije klijenata

Kriterije segmentacije se može podijeliti u dvije osnovne skupine prema kategoriji korisnika: poslovni i privatni korisnici. Poslovni korisnici (poduzeća i obrti) se nazivaju i tržište poslovne potrošnje, a predstavljaju B2B (engl. *business to business*) odnos među kompanijama. Privatni korisnici (građanstvo) predstavljaju B2C odnose (engl. *business to customer*), a smatramo ih tržištem krajnje potrošnje (Lešković, 2007).

Na poslovnom je tržištu kriterije segmentacije moguće razvrstati u nekoliko skupina.

3.1.1. Osobne karakteristike korisnika (ponašanje)

Čak i u B2B odnosima, u konačnici odluku donose pojedinci, ali koji su ograničeni politikom svog poduzeća. Podaci o značenju pojedinih uloga i osoba u procesu kupovine, kao i podaci o njihovim karakteristikama, izuzetno su skupi i teško dobavljivi. Postoji mogućnost prikupljanja takvih podataka od

strane zastupnika prodaje te skladištenja tih podataka u baze podataka. Možemo izdvojiti slijedeće osobne karakteristike korisnika u B2B odnosima (Lešković, 2007):

- a) sličnost osobe koja kod korisnika donosi odluku o kupovini i kompanije dobavljača proizvoda/usluga,
- b) odnos osobe koja odlučuje o kupnji prema riziku,
- c) naklonjenost osobe koja odlučuje o kupnji prema dobavljaču.

3.1.2. Demografski kriteriji segmentacije

Kako je demografija znanost o dinamici stanovništva, logično bi bilo zaključiti kako se ona zapravo odnosi na B2C dio poslovanja. U B2B odnosima, demografske varijable su veličina kompanije korisnika, potencijal korisnika (i potencijal gospodarske grane u kojoj on posluje) te lokacija korisnika (koja načelno implicira gospodarsku moć subjekta). U skladu s gore navedenim, možemo izdvojiti slijedeće varijable vezane za ovaj tip kriterija (Lešković, 2007):

- a) Veličina kompanije: utječe na segmentaciju jer velike kompanije uglavnom zahtijevaju prilagođena rješenja, koja ne moraju nužno biti profitabilna za poduzeće koje daje uslugu (prodaje proizvod).
- b) Poznavanje gospodarske grane: omogućuje bolje razumijevanje potreba korisnika i percepciju situacije kupovine. Npr. banka ima slične zahtjeve za telekomunikacijskim uslugama kao i osiguravajuće poduzeće, ali se oni kao potrošačke grupe znatno razlikuju po drugim faktorima pa se mora uzeti u obzir i segmentacija po djelatnosti.
- c) Lokacija korisnika: utječe na organizaciju prodajne sile. Posebno je važna kod organizacije prodajnih zastupnika kao kanala prodaje te određenja lokacije maloprodajnih mjesta (lokacija prodajnih partnera).

3.1.3. Operativni kriteriji segmentacije

Ovi kriteriji omogućuju precizniju segmentaciju u okviru pojedinih demografskih kategorija:

- a) Tehnologija proizvodnje: dugoročno determinira potrebe za potrošnim materijalima, uslugama i imovinom, stoga je moguće korisnika prilično detaljno segmentirati po ovom kriteriju.
- b) Potrošački status: predstavlja jedan od najjednostavnijih načina segmentacije koji se najčešće radi po proizvodima i po potrošačkoj marki. Naime, svi korisnici jednog proizvoda ili trgovačke marke uglavnom imaju neke zajedničke karakteristike. Poduzeća definiraju vlastiti pristup korisniku temeljem poznavanja ponašanja korisnika u kupnji konkurentskih proizvoda/usluga i analize slabih točki pristupa konkurencije. Tada preferiraju ono tržište, tj. korisnike čije karakteristike (slabosti i snage) bolje poznaju.
- c) Sposobnost korisnika: kriterij koji opisuje sposobnost korisnika u smislu potreba za većom ili manjom količinom proizvoda/usluga.

3.1.4. Kriteriji segmentacije vezani za pristup nabavi

Kriteriji segmentacije vezani za pristup nabavi mogu biti:

- a) Organizacija nabave: determinira veličinu i poslovanje nabave korisnika. Poduzeće se treba prilagoditi klijentu, a to uključuje i prilagodbu klijentovoj nabavi.
- b) Struktura moći: utjecaj pojedinih organizacijskih jedinica na odluku o kupovini.
- c) Nabavna politika: pravilo klijenta da radi s poznatim poduzećem, domaćim poduzećem, najjeftinijim poduzećem itd.
- d) Priroda postojećih odnosa: opisuje čvrstoću odnosa s pojedinim korisnicima.
- e) Kriteriji pri nabavi: opisuje bitne karakteristike klijenata u smislu njegova afiniteta prema cijeni, kvaliteti usluge i sl.

3.1.5. *Situacijski kriteriji segmentacije*

Ovi elementi nemaju stalni karakter i zahtijevaju dublje poznavanje klijenata (Lešković, 2007):

- a) Hitnost zahtjeva: kupovina može biti prva, ponovljena ili modificirana ponovljena. Neka se poduzeća mogu usredotočiti samo na hitne zahtjeve klijenata.
- b) Način primjene proizvoda/usluge: predstavlja značajnost istoga u ukupnom poslovnom procesu korisnika.
- c) Veličina prosječne narudžbe.
- d) Podaci o lojalnosti korisnika, odnosno da li korisnik jedino od nas nabavlja ovu vrstu usluge te da li od nas nabavlja naše ostale proizvode/usluge.

4. METODOLOGIJA I KARAKTERISTIKE UZORKA

Podaci za istraživanje su preuzeti iz ankete poslovnog sektora bankarskog tržišta koja je obuhvaćala 1200 hrvatskih poduzeća, neovisno o porijeklu vlasništva. Anketa je provedena u veljači 2003. godine.

Anketa je sadržavala pitanja o raznim aspektima povezanosti banaka i njihovih klijenata. Tako je prvi dio ankete o informiranosti poduzeća i poznavanju banaka nazvan „A-poznavanje banaka“. Drugi dio ankete vezan za ponašanje prilikom korištenja bankarskih usluga je nazvan „B-korištenje bankarskih usluga“. Slijedeći segment pitanja je nazvan „C-glavna banka i lojalnost glavnoj banci“. Nakon toga se ispitivao ugled banaka i percepcija o najboljoj banci („E-ugled banaka i najbolja banka“). Također je ispitana i navika poslovanja s bankama („F-navike poslovanja s bankama“), ponašanje prilikom zaduživanja i trošenja novca („G-kredit i kreditne kartice“) te odnos poduzeća prema tehnologiji („I-informatika“). Zadnji odjeljak s pitanjima daje demografske informacije o poduzećima kao što su: osnovna djelatnost poduzeća, broj zaposlenih, veličina poduzeća, ukupni godišnji prihod itd („Dm-demografija“).

Treba napomenuti da svaki od navedenih dijelova ankete sadrži pitanja koja većinom daju odgovore u obliku kategorijalnih podataka, o pojedinom pitanju ovisi koliko će biti ponuđenih kategorija. Većinom su to skale (Likertova skala) od 0 – 5 s pripadnim objašnjenjima pojedinog broja. Ove skale su temelj za analizu podataka pomoću metode samo-organizirajućih mapa.

Određeni problem sa skupom podataka je činjenica da osim podataka o uvozu/izvozu te broju zaposlenih, nema numeričkih podataka. Taj nedostatak u određenoj mjeri smanjuje atraktivnost i preglednost virtualnog prikaza, ali odnos među pojedinim varijablama se i dalje kvalitetno može prikazati. Pitanja se bila strukturirana na načina da daju kategorijalne podatke, tj. inače numeričke vrijednosti su samim procesom anketiranja obrađene i svrstane u kategorije radi pojednostavljenja procesa segmentacije.

4.1. Prikupljanje i odabir podataka značajnih za SOM analizu

Podaci za SOM analizu su odabrani uzevši u obzir kriterije segmentacije navedene u poglavlju „4.4 Kriteriji segmentacije klijenata“.

Podaci o osobinama korisnika

Što se tiče prvog kriterija segmentacije (Osobine korisnika), odabrane su varijable (podaci) koji su u skladu s „c) naklonjenost osobe koja odlučuje o kupnji prema dobavljaču“. Točnije, to su slijedeće varijable:

- b0 – važno kod koje banke se koriste usluge,
- e2 – najbolja banka u Hrvatskoj.

Demografski podaci

Podatke o demografskim karakteristikama ispitanih poduzeća nije trebalo puno tražiti, jer su u anketi označeni pod dijelom „Demografija – Dm“. Zbog velikog broja demografskih varijabli (26 varijabli) te

zahtjeva za što manjim brojem ulaznih varijabli, radi lakšeg razumijevanja izlaznih rezultata algoritma, odabrano je samo 13 varijabli.

Što se tiče dijela „a) veličina kompanije“, uzete se u obzir varijable:

- Dm6 – ukupni broj zaposlenih u poduzeću,
- Dm16 – rast/pad prihoda u odnosu na prethodnu godinu,
- Dm16a – iznos ukupnog prihoda u prethodnoj godini,
- Dm21 – planovi s brojem zaposlenih.

Podaci odabrani za prikazivanje karakteristika poduzeća ovisno o njenoj pripadnoj gospodarskoj grani su bili sljedeći:

- Dm8 – osnovna djelatnost poduzeća,
- Dm13 – vrsta vašeg poduzeća,
- Dm14 – postotak uvoza u godišnjem prometu,
- Dm15 – postotak izvoza u godišnjem prometu,
- Dm18 – oblik vlasništva nad poduzećem,
- Dm19 – podrijetlo kapitala,
- Dm22 – financijska situacija,
- Dm24_1 – način isplate plaća.

Zadnji podaci koji su od demografskih podataka odabrani su vezani za lokacijske i povijesne čimbenike, u skladu sa zadnjim demografskim kriterijem segmentacije navedenom u poglavlju „4.4.3 Demografski kriteriji segmentacije“:

- Dm7 – sjedište poduzeća,
- Dm9 – godina osnutka poduzeća.

Operativni podaci

Ovo su podaci u skladu s ranije spomenutim operativnim kriterijima segmentacije, nakon duljeg razdoblja razmatranja varijabli u anketi, za ovaj dio su odabrani podaci koji se odnose na status potrošača, poglavito podaci vezani za percepciju o lakoći dobivanja kredita:

- g1_1, g1_2, g1_3 – 3 najvažnija razloga koja utječu na izbor banke za dizanje kredita,
- g2 – najvažniji razlog koji utječe na izbor banke za dizanje kredita,
- k1_1 – kreditne kartice koje koristi poduzeće,
- k2_1 – kreditne kartice koje poduzeće planira koristiti.

Podaci o pristupu nabavi bankarskih usluga/proizvoda

Ovi podaci su odabrani u skladu s poglavljem „4.4.4 Kriteriji segmentacije vezani za pristup nabavi“. S napomenom da su odabrani podaci koji se mogu svrstati u skupinu kriterija „e) Kriteriji pri nabavi“. Kao što je već rečeno, oni opisuju bitne karakteristike klijenata u smislu njegova afiniteta prema cijeni, kvaliteti usluge i sl.:

- c3b_1, c3b_2, c3b_3 – tri najvažnije karakteristike/obilježja banke (npr. sigurnost i stabilnost, visina provizija i naknada, održavanje obećanja),
- e3b_1, e3b_2, e3b_3 – tri najvažnije karakteristike/obilježja usluga u banci (npr. široki spektar usluga, odnos kvalitete i cijene, brzo rješavanje zahtjeva).

Situacijski podaci

Varijable koje su odabrane kako bi predstavile poduzeća sa stajališta situacijskih kriterija su sukladno poglavlju 4.4.5 vezani za „d) lojalnost korisnika“:

- c1 – glavna banka (informacija o banci kod koje se koristi najviše usluga),
- c2 – stupanj zadovoljstva glavnom bankom.

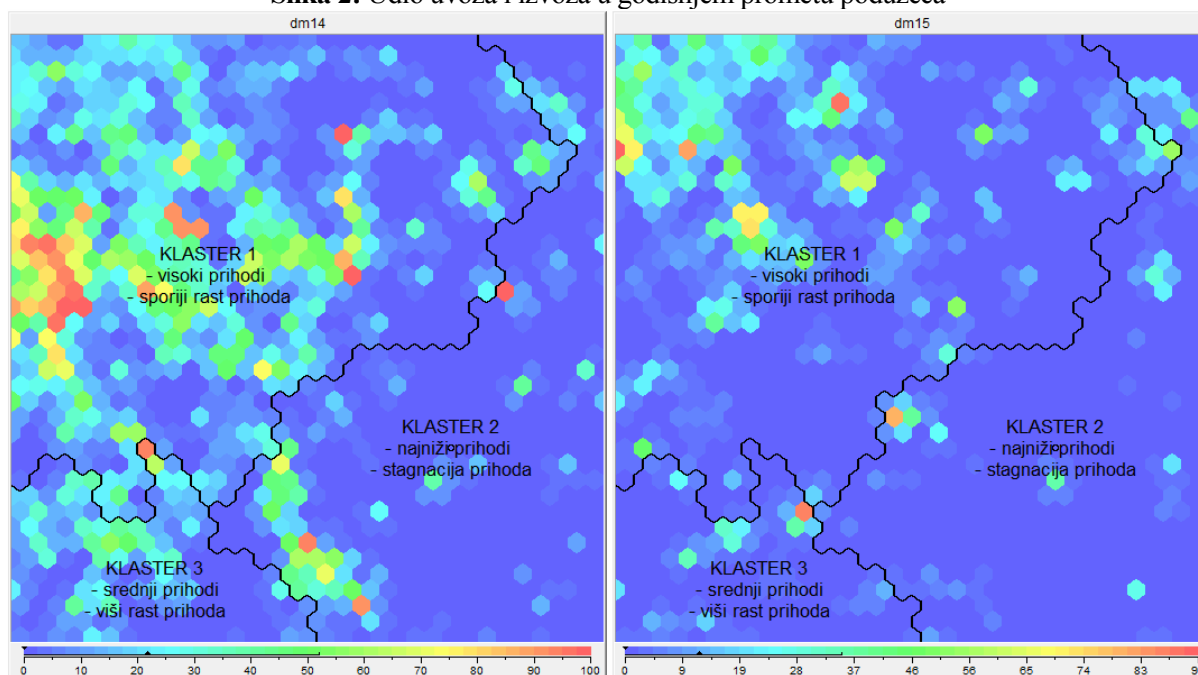
5. REZULTATI ISTRAŽIVANJA

5.1. Prikaz rezultata klasteriranja

Rezultirajuće mape, osim rezultata klasteriranja, prikazuju i redosljed kojim će biti navedene razlike među klasterima u interpretaciji rezultata klasteriranja. Vrijednosti na mapi su veće ako je čvor crvene boje, a manje kako se boja približava plavoj boji, što je vidljivo i na skali ispod svake slike.

Vizualni prikaz udjela uvoza i izvoza (Dm14, Dm15) uvelike ističe prirodu poslovanja hrvatskih poduzeća, koja se više okrenuta uvozu nego izvozu. Što je na mapama (slika 2) prikazano s većim udjelom crvene boje na lijevom dijelu slike. Zaključak je da su poduzeća unutar klastera 1 najaktivnija u trgovinskom poslovanju.

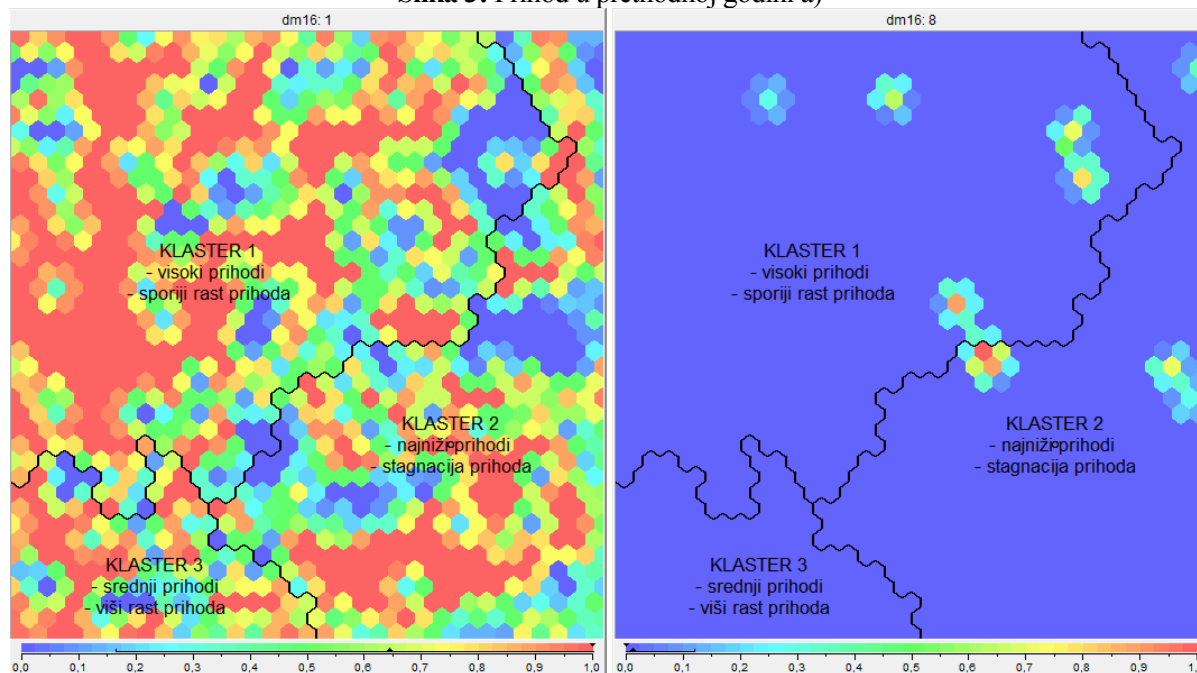
Slika 2: Udio uvoza i izvoza u godišnjem prometu poduzeća



Izvor: izrada autora

Prikaz razlika među klasterima s obzirom na ukupni prihod prethodne godine (Dm16) se može vidjeti na slici 3. Lijevo se nalazi mapa koja prikazuje razdiobu binarnih vrijednosti za vrijednost prihoda u prethodnoj godini do 3,5 milijuna kuna. Desno je prikazana mapa za vrijednosti prihoda od 75 do 300 milijuna kuna. Dobro se vidi da su to velika poduzeća sa sporijim rastom prihoda (utvrđeno detaljnom analizom podataka).

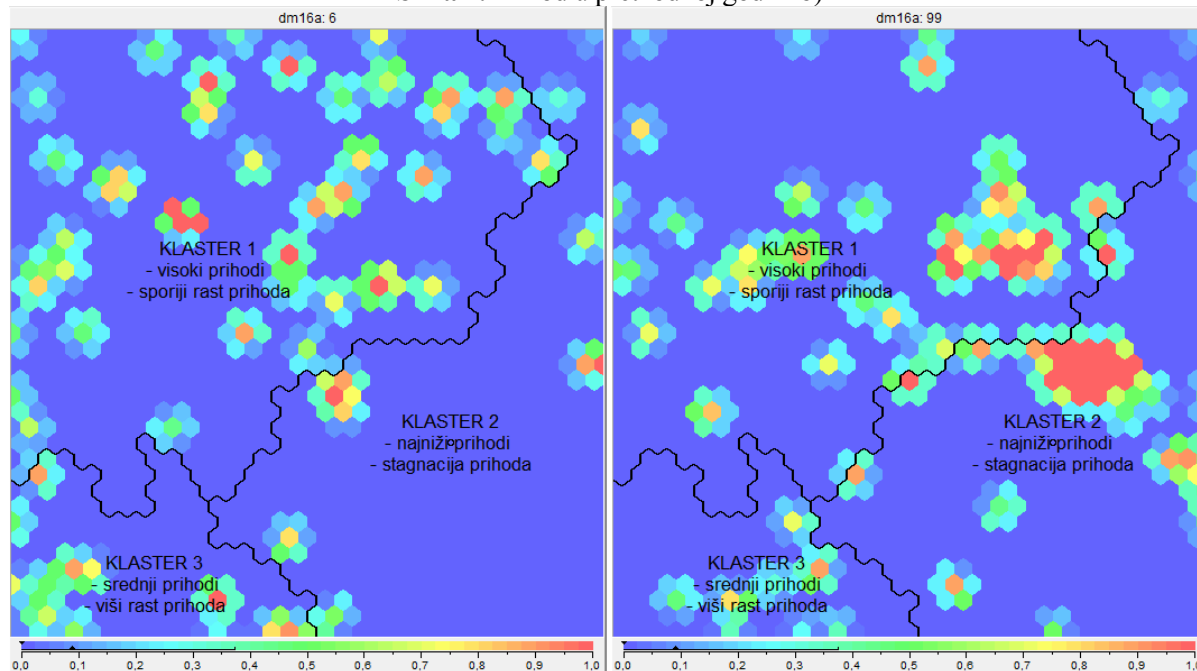
Slika 3: Prihod u prethodnoj godini a)



Izvor: izrada autora

Na slici 4 se vide vrijednosti prihoda od 33 do 56 milijuna kuna (Dm16a:6). Na desnoj strani se vidi koja poduzeća su zatajila informaciju o prihodu prethodne godine (Dm16a:99). Tako je prihod namjerno prešutjelo 9% poduzeća iz klastera 1 i 9,8% u klasteru 2.

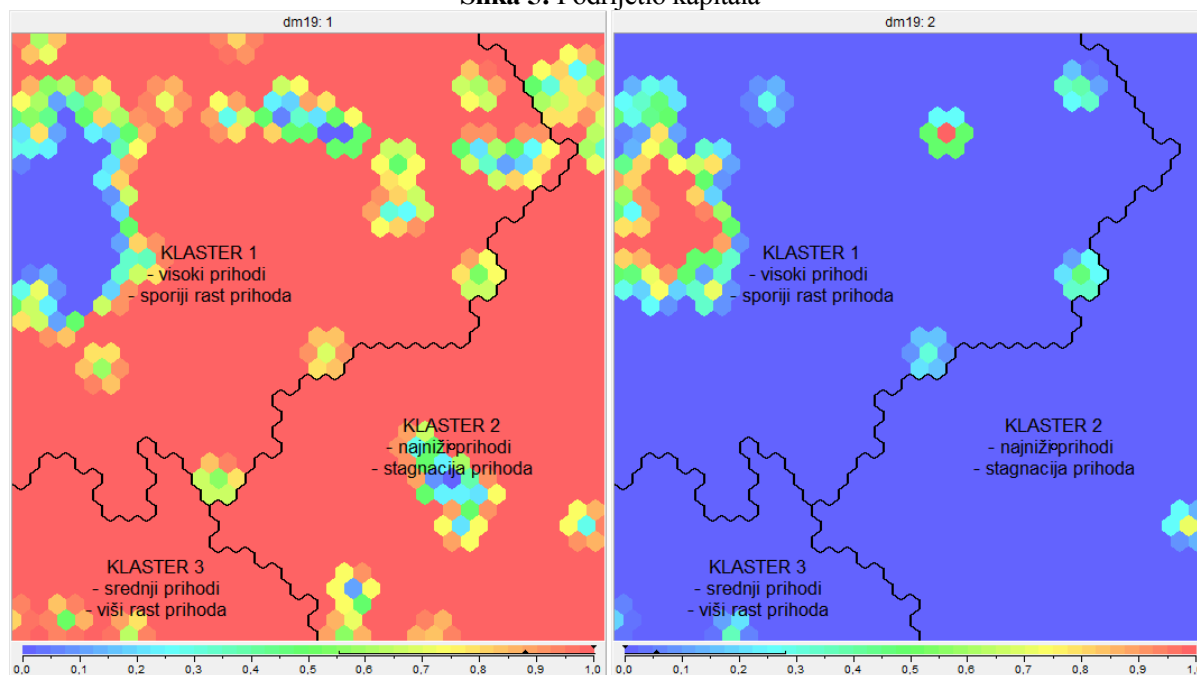
Slika 4: Prihod u prethodnoj godini b)



Izvor: izrada autora

Vrlo korisna ilustracija mogućnosti samo-organizirajućih mapa se vidi na slici 5. Na lijevoj mapi se vide domaća poduzeća (crveno obojana), a na desnoj strani su crvenom bojom označena strana poduzeća. Ako bi se ove dvije mape preklapile jedna preko druge, rezultirajuća mapa bi bila gotovo cijela crvena. Jedino bi plave boje bili čvorovi koji pripadaju poduzećima s mješoviti podrijetlom kapitala.

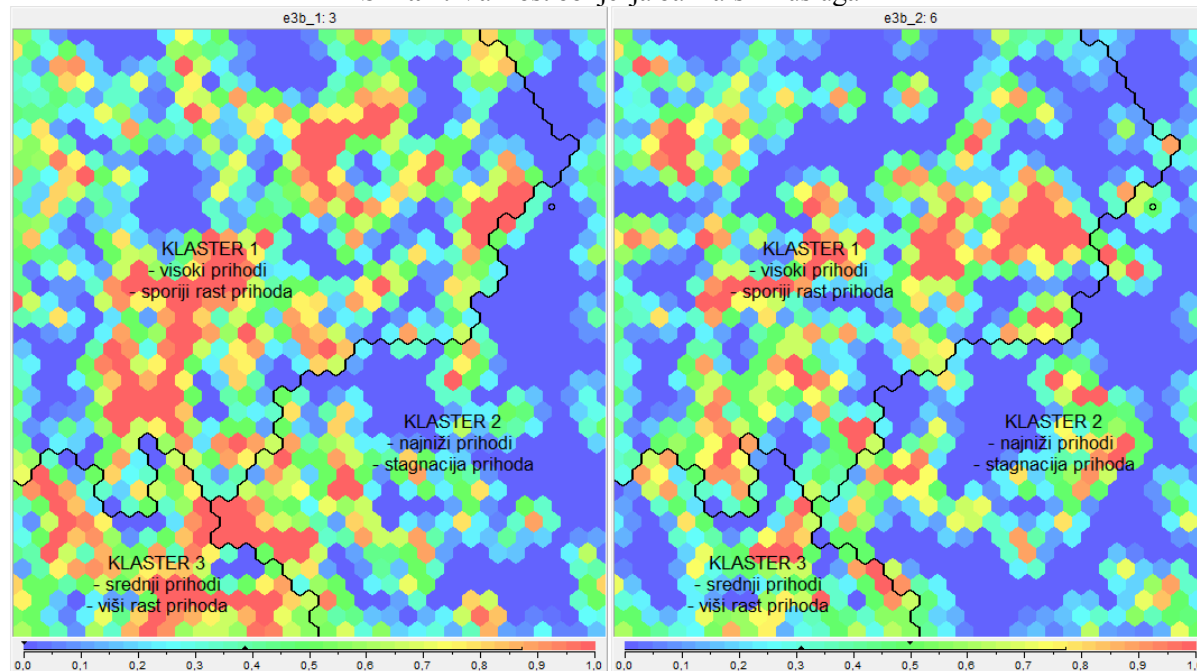
Slika 5: Podrijetlo kapitala



Izvor: izrada autora

Da ne bi bile prikazane samo demografske varijable, na idućoj slici se vidi koje su dvije najvažnije usluge poduzećima. To su točnost obavljanja transakcija (e3b_1:3) i kvaliteta rješavanja zahtjeva i problema (e3b_2:6) (za klustere 1 i 3). Poduzeća u klasteru 2 više cijene široki spektar usluga (e3b_1:1), a druga važna usluga im je opet točnost transakcija.

Slika 7: Važnost obilježja bankarskih usluga



Izvor: izrada autora

Još se može prikazati mapa po pet varijabli koje bi bile značajne za klaster analizu, ali se uz pretpostavku da je koncept samo-organizirajućih mapa moguće shvatiti iz svega dosad prezentiranog, to nije potrebno. Ostale varijable su Dm19 – oblik vlasništva nad poduzećem (državno, mješovito i privatno), k1_1 – kreditne kartice koje posjeduje poduzeće, g1_1 – kriteriji odabira kredita.

Ovime je završen prikaz rezultata klasteriranja i slijedi interpretacija klasteriranja. U tom dijelu rada će se nešto više reći o faktorima (atributima) koji su utjecali na kreiranje klastera, slično kako je to učinjeno pri prikazu rezultata klasteriranja.

5.2. Interpretacija rezultata klasteriranja

Svaki od tri klastera je pomno ispitan i analizirani su podaci unutar svakog pojedinog klastera kako bi se zaključilo kakve su njegove karakteristike. Prilikom interpretacije segmenata korišteni su statistički podaci o podacima unutar odabranog klastera. Dvije vrijednosti koje su ukazivale na razlike među atributima klastera su prosječna vrijednost (engl. *Mean*) i standardna devijacija (engl. *Standard Deviation*). Treba napomenuti kako je izostavljena informacija o veličini tvrtke jer je standardna devijacija u sva tri klastera bila previsoka (redom 1264, 990 i 920).

Klaster 1 je na lijevoj strani svake mape i ima slijedeće karakteristike (svaki od navedenih udjela se odnosi na određeni klaster, a ne na cijelu populaciju):

1. najveći udio uvoza/izvoza u ukupnom godišnjem prometu poduzeća (uvoz 29,6%, a izvoz 16,1%)
2. najveći ukupni godišnji prihod u prošloj godini (od 75 do 300 milijuna, 14,4% unosa),
3. najviše zastupljenih državnih poduzeća (10,9%),
4. najviše kapitala stranog podrijetla (10,2%),
5. najviše planiraju zapošljavanje novih kadrova (48,9%),
6. najviše koriste kreditne kartice (53,4% - American Express),
7. pri odabiru kredita najvažnija kamatna stopa, brzina dobivanja kredita i pojednostavljena procedura dizanja kredita,
8. od banke se zahtijeva visoki stupanj sigurnosti, povoljne kamate na kredite,
9. od bankarskih usluga se zahtijeva visoka točnost transakcija, brzo rješavanje zahtjeva i brzo rješavanje problema.

Dakle, za ovaj klaster je karakterističan stil koji imaju velika poduzeća koja imaju razrađenu strategiju i žele osigurati daljnje nesmetano funkcioniranje poslovanja, poglavito s inozemstvom. Također je ovom segmentu potrebno uskočiti u pomoć s posebno prilagođenim uslugama baš za njih.

Klaster 2 se nalazi na desnoj strani svake od samo-organizirajućih mapa. Karakteristike klastera 2 su:

1. najmanje trgovine s inozemstvom (uvoz 11%, izvoz 6,9%),
2. najmanji prosječni prihod u prošloj godini (do 3,5 mil. kn je zaradilo 54,9% ispitanih),
3. prosječno najviše poduzeća u privatnom vlasništvu (87,9%),
4. najviše zastupljen domaći kapital i minimalno stranog kapitala (94,1%),
5. stagnacija zapošljavanja (48,7%) s neodređenom strategijom zapošljavanja (u najvećoj mjeri neznaju da li će nekoga zaposliti ili otpustiti iduće godine, 8,2%),
6. u najmanjoj mjeri koristi kreditne kartice (37,9% ispitanih ne koristi nijednu karticu),
7. pri odabiru kredita najvažnije stavke su kamatna stopa, brzina dobivanja kredita i jednostavna procedura dobivanja kredita (kao i kod klastera 1),
8. od banke se zahtijeva isto što i u klasteru 1 (visok stupanj sigurnosti, povoljne kamate na kredite i garanciju sigurnosti uloga),
9. od bankarskih usluga se zahtijeva široki spektar usluga i proizvoda, točnost transakcija te brzo i efikasno rješavanje problema.

Iz podataka o ovom klasteru se može vidjeti da su njegovi pripadnici u najvećem broju poduzeća u privatnom vlasništvu, te najmanje poduzeća ima strano podrijetlo kapitala. Ovo su vjerojatno najvećim dijelom obrti, što se može vidjeti i iz najmanjeg prosječnog prihoda u prethodnoj godini. To pokazuje i činjenica da nemaju razrađenu strategiju daljnjeg zapošljavanja ili ne planiraju nikoga zaposliti, što je vjerojatno uzrokovano nepostojanjem službe koja se brine o tom segmentu poslovanja. Značajno je za napomenuti da se ovaj segment razlikuje od ostala dva (većih poduzeća) u činjenici da žele široki spektar usluga od banke, dok velika poduzeća žele visoku točnost transakcija.

Zadnji klaster je najmanji i smješten je u donjem lijevom kutu. Nakon detaljne analize, zaključeno je da klaster 3 ima slijedeće karakteristike:

1. visoki udio uvoza u ukupnom godišnjem prometu (22,1%), ali mali udio izvoza u godišnjem prometu (9,9%),
2. solidne godišnje prihode (u tri prihodovna razreda najviše: 1 – do 3,5 mil. kn, 5 – od 14 do 33 mil. kn, 6 – od 33 do 56 mil. kn),
3. drugi po redu po stupnju privatnog vlasništva, uz tek 9% državnih poduzeća
4. najviše poduzeća s domaćim podrijetlom kapitala (97,5%),
5. stagnacija zapošljavanja u laganoj prednosti pred zapošljavanjem novih zaposlenika (44,9% stagnacija, 39,8% novo zapošljavanje),
6. srednja razina korištenja kreditnih kartica (40,7% - American Express),
7. pri odabiru kredita najvažniji kriteriji su brzina dobivanja kredita, jednostavna procedura ili instrumenti osiguranja i zahtijevanje realnih jamstava od klijenata koji žele dizati kredit,
8. banka bi trebala biti sigurna i stabilna, ima zaduženu osobu za rješavanje problema i odgovore na određene zahtjeve, garantira stabilnost uloga,
9. bankarske usluge koje ovaj klaster zahtijeva su točnost transakcija, brzo rješavanje zahtjeva i brzo rješavanje problema (baš kao i u slučaju klastera 1).

Za ovaj segment je najznačajnije da su tu poduzeća s pretežno domaćim podrijetlom kapitala (97,5%), da pokušavaju ostati na istom broju zaposlenika u malo većoj mjeri nego što će zaposliti nove radnike (15% razlike) te imaju umjerenu potrošnju na kreditnim karticama (40% poduzeća). Kod dobivanja kredita im za razliku od klastera 1 i 2 nije najvažnija kamatna stopa, već je to u prvom redu brzina dobivanja kredita i jednostavna procedura. Osim toga, zahtijevaju da transakcije budu točne, kao i klaster 1.

6. ZAKLJUČAK

Cilj rada je bio pokazati kako se može upotrijebiti koncept samo-organizirajućih mapa u klaster analizi. U prvom redu korištenjem njihova svojstva redukcije dimenzionalnosti, vizualizacije podataka i zbližavanja sličnih podataka na dvo-dimenzionalnoj (ili 3D) mapi, moguće je razumjeti složene međuodnose ulaznih podataka te donositi argumentirane zaključke.

Klasične metode klaster analize pružaju zavidnu razinu točnosti (Mingoti i Lima, 2006), ali njihov problem je manjak mogućnosti automatskog određivanja broja klastera. Određivanje broja klastera ovisi o prethodnom poznavanju analiziranog skupa podataka, a ako analitičar ima obvezu analize složenih baza podataka to mu neće biti lako. Stoga je potreba za automatskim određivanjem pravilnosti (uzoraka) u podacima očigledna. SOM metoda upravo zbog svojih karakteristika navedenih u poglavlju 3, može biti smatrana kao uspješna nadopuna klasičnim metodama klasteriranja.

Pošto su klaster analiza i segmentacija tržišta logički povezane, u radu je spomenuta segmentacija tržišta, prikazano je stanje u bankarskom sektoru poslovnih klijenata i dani su kriteriji segmentacije tržišta. Navedeni kriteriji segmentacije su poslužili za odabir podataka značajnih za analizu u softveru SOMine. Tako je efikasnost SOM-Ward algoritma prikazana na stvarnim podacima i dala je konkretne rezultate segmentacije.

Danas, kad baze podataka bivaju multiplicirane u smislu veličine i složenosti, upotreba alata za analizu složenih podataka je neophodna. Tako su samo-organizirajuće mape u sprezi sa klasičnim tehnikama, kao što je Ward tehnika klasteriranja u slučaju Viscovery SOMine-a, ali i nekim novijim (npr. U*C algoritam (Ultsch, 2007)), idealan alat za klaster analizu.

Segmentacija pomoću SOM-a daje mogućnost raznih kombinacija aktivnosti od strane marketinških stručnjaka kako bi se donijele konkretne odluke o poduzimanju novih koraka u marketingu banaka. Treba napomenuti da je segmentacija u određenu ruku slična tradicionalnoj segmentaciji banaka, no sada se točno može vidjeti mogućnosti za fine prilagodbe u marketinškim aktivnostima.

Isto tako, treba napomenuti da je anketa poduzeća bila provedena 2003. godine i stoga bi bilo vrlo dobro analizirati tržište pomoću novih podataka. Samim time je otvoreno područje za daljnja istraživanja i analizu tržišta pomoću samo-organizirajućih mapa.

LITERATURA

1. Bevanda, A. (2008). Segmentiranje financijskog tržišta u Federaciji Bosne i Hercegovine na temelju čimbenika imidža. *Tržište*, 179-193.
2. Dragičević, M. (2006). Neuronske mreže i analitički hijerarhijski proces u segmentaciji turističkog tržišta. *JEL klasifikacija*
3. Kiang, M. i Fisher, D. (2008). Selecting the right MBA schools-An application of self-organizing map networks. *Expert Systems with Applications*, 946-955.
4. Kiang, M. Y. (2001). Extending the Kohonen self-organizing map networks for clustering analysis. *Computational Statistics & Data Analysis* 38, 161-180.
5. Kohonen, T. (1995). *Self-organizing maps (3rd edition)*. Berlin: Springer-Verlag
6. Kumar, P. R. i Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. *European Journal of Operational Research* 180, 1-28.
7. Lešković, D. (2007). Segmentacija tržišta korištenjem otkrivanja znanja iz baza podataka. *Magistarski rad*
8. Pejić-Bach, M. (2005). Rudarenje podataka u bankarstvu. 181-193.
9. Ultsch, A. (2007). Emergence in self-organizing feature maps. *Proceedings Workshop on Self-Organizing Maps*