

Dr. sc. Kata Ivić*

Zvonimir Jurković, univ. bacc. oec.**

Rozalija Marinković, univ. bacc. oec.**

UDK 025.6 : 164.3
Prethodno priopćenje

IZRADA MODELA NEURONSKIH MREŽA ZA UČESTALOST KORIŠTENJA FAKULTETSKE KNJIŽNICE

SAŽETAK

Neuronske mreže, kao metoda umjetne inteligencije, upotrebljavaju se sve više u današnjem poslovanju u okviru sustava za potporu odlučivanju. Ključan čimbenik za provedbu metode u konačnici je znanje korisnika koji primjenjuje navedenu metodu kako bi mogao kvalitetno prirediti podatke za istraživanje i zatim ih točno tumačiti. Fakultetskim knjižnicama vrlo su vrijedna saznanja i predviđanja dinamike posjeta korisnika, kako bi se rad u knjižnicama mogao nesmetano odvijati bez stvaranja određenih repova čekanja.

Cilj je ovog rada klasificirati studente u jednu od dviju skupina, odnosno prema prosječnim dolascima studenata u fakultetsku knjižnicu tijekom istoga radnog dana. U istraživanje su uključene kontinuirane i kategorijalne varijable, odnosno sedam ulaznih varijabli te jedna izlazna varijabla za testiranje arhitektura. Testirane su dvije arhitekture neuronskih mreža pomoću backpropagation algoritma višeslojne perceptron mreže i izabran je najbolji model. Provedena je analiza osjetljivosti, na temelju koje je otkriveno da najveći utjecaj na izlaznu varijablu ima spol, mogućnost rada u čitaonici te mogućnost korištenja računala. Rezultati modeliranja pokazuju da neuronska mreža dobro uočava međusobne odnose ulaznih i izlaznih varijabli.

Kreirani model ukazuje na mogućnosti i potrebu za primjenom metoda umjetne inteligencije u području knjižničarstva. Izrađeni model pogodan je za daljnja istraživanja u ovom području uz određena poboljšanja, poput uvođenja dodatnih ulaznih varijabli i povećanjem uzorka. Uspješnost modela mjerena je ukupnom stopom klasifikacije.

Ključne riječi: neuronske mreže, fakultetska knjižnica, klasifikacija studenata, višeslojni perceptron, analiza osjetljivosti

* Ekonomski fakultet u Osijeku
** studenti 2. godine diplomskog studija

1. Uvod

Za predviđanje dinamike posjeta knjižnici Ekonomskog fakulteta u Osijeku od velikog je značenja je praćenje statistike korisnika knjižnice i ispunjavanje statističkih obrazaca posjeta korisnika knjižnici. Obradom prikupljenih podataka utvrđuje se:

- tko su korisnici fakultetske knjižnice
- koju vrstu knjižnične usluge trebaju, odn. što tražuju u knjižnici.

Statistički prikupljenim podacima moguće je predvidjeti koja su to vremenska razdoblja tijekom radnog vremena knjižnice najfrekventnija prema broju posjeta korisnika. Pojava nastajanja repova čekanja u knjižnici nepredvidiva je zbog složene dinamike nastavnih procesa i cikličkih potreba korisnika za pojedinom uslugom u knjižnici (literatura za seminarske radove, ispitna literatura – udžbenici, i sl.). Stoga je ključno točno segmentirati studente (vrijeme posjeta knjižnici, godina studija studenta, i sl.) radi dobivanja informacija o tome koji će studenti doći samo jednom u fakultetsku knjižnicu, a koji će studenti imati potrebu za ponovnim dolaskom u knjižnicu istoga radnog dana.

Prediktivni model za ovu svrhu jedne od mogućih metoda jesu neuronske mreže. Model je dobiven usporedbom testiranih neuronskih mreža prema prikupljenim internim statističkim podacima fakultetske knjižnice.

Cilj je ovog rada pronaći uspješniji model te uočavanje mogućnosti poboljšanja testiranih arhitektura za buduće primjene. Jedan od glavnih zadataka koji se želi postići njegovom izradom jest predlaganje testirane metode za buduću uporabu u području knjižničarstva. Također je bitno uočiti koje su varijable najznačajnije za izradu modela.

2. Metodologija istraživanja

Istraživanje o klasifikaciji studenata prema posjetama fakultetskoj knjižnici provedeno je pomoću metode umjetne inteligencije. Jedna od definicija umjetne inteligencije glasi: „Umjetna inteligencija je grana računarske znanosti koja se bavi proučavanjem i oblikovanjem računarskih sustava koji pokazuju neki oblik inteligencije. Takvi sustavi mogu učiti, mogu donositi zaključke o svijetu koji ih okružuje, oni razumiju prirodni jezik te mogu spoznati i tumačiti složene vizualne scene te obavljati druge vrste vještina za koje se

zahtijeva čovjekov tip inteligencije.“ (D.W. Peterson, 1990.)¹ Metoda korištena u obradi podataka za potrebu izrade ovoga rada jest metoda neuronskih mreža.

2.1. Metodologija neuronskih mreža

Znanstvena disciplina kojoj pripadaju neuronske mreže jest umjetna inteligencija. Stoga neuronske mreže pripadaju u inteligentne metode rudarenja podataka.

U literaturi postoje mnoge definicije neuronskih mreža. Jedna od definicija glasi da su neuronske mreže programi ili hardverski sklopovi koji iterativnim postupkom iz prošlih podataka nastoje pronaći vezu između ulaznih i izlaznih varijabli modela kako bi se za nove ulazne varijable dobila vrijednost izlaza.²

U računalstvu i informatici neuronska mreža predstavlja sustav međusobno povezanih elemenata (nazvanih umjetni neuroni) koji služe za raznovrsna izračunavanja, zasnovana na pokušajima oponašanja rada ljudskog mozga. Pretpostavlja se da ljudski mozak ima oko 100 milijardi neurona, od kojih je svaki povezan s 1000 do 10 000 drugih neurona. Iako djelovanje biološkog neurona nije do kraja razjašnjeno, pokazalo se da jednostavni modeli umjetnih neurona i njihovo jednostavno povezivanje može korisno poslužiti pri rješavanju mnogih zadataka u znanosti, tehnici i u drugim primjenama.³ Prema navedenome, samo polazište ideje o neuronskim mrežama potječe iz neuropsihologije. Iz neuropsihologije preuzeta su saznanja o ponašanju živčane stanice koja funkcionira po principu aktiviranja, te su saznanja iskorištena za formiranje matematičkog modela koji se koristi u praksi. U navedenom matematičkom modelu, osnovna matematička jedinica dizajnirana je po uzoru na biološki neuron.⁴ Slikom 1. prikazana je usporedba biološkog i umjetnog neurona.

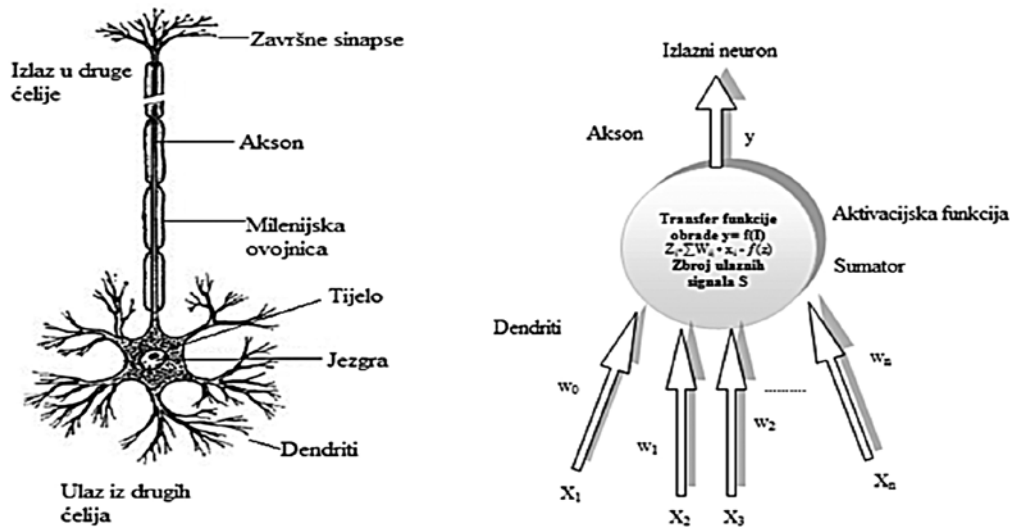
1 Dostupno na: http://www.fer.unizg.hr/_download/repository/UI_1_uvod.pdf (02.02.2011.)

2 Dostupno na: <https://moodle.carnet.hr/mod/resource/view.php?id=23158> (4.2.2012.)

3 Hrvatska enciklopedija, br. 7 Mal – Nj, Leksikografski zavod Miroslav Krleža, Zagreb, 2005., str. 648.

4 Klepac, G., Mršić, L.: Poslovna inteligencija kroz poslovne slučajeve, Liderpress/TimPress, Zagreb, 2006., str. 49.

Slika 1. Usporedba biološkog i umjetnog neurona



Izvor: modificirano prema http://www.byclb.com/TR/Tutorials/neural_networks/ch7_1.htm (19.03.2012.)

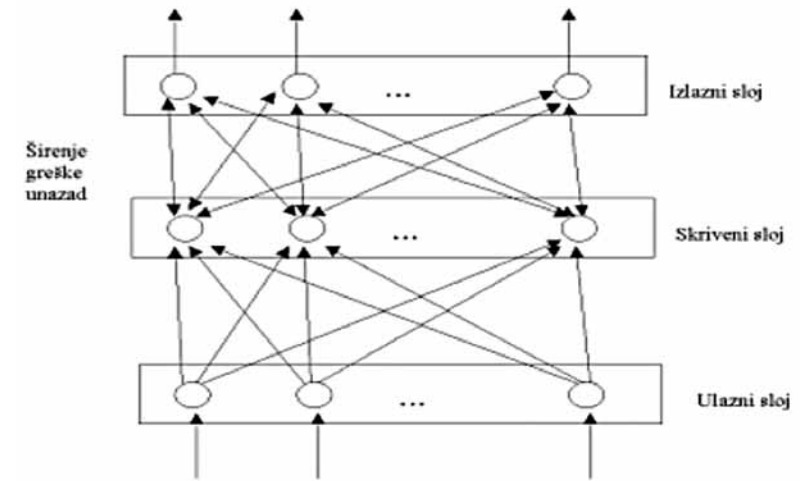
Biološki neuron predstavlja stanicu koja prima od drugih neurona putem dendrita informacije. Zatim ih on obrađuje i šalje impuls putem aksona i sinapsi drugim neuronima u mreži. Kako se jačina sinaptičkih veza mijenja, tako se odvija učenje. Milijuni neurona koji se nalaze u mreži mogu istodobno obrađivati informacije. Za razliku od biološkog neurona umjetni neuron može se definirati kao jedinica za obradu podataka, tj. varijabla. Ona prima ponderirane ulazne vrijednosti od drugih varijabli. Nakon toga, prema određenoj formuli varijabla transformira primljenu vrijednost te šalje izlaz drugim varijablama. Učenje se zatim odvija promjenom vrijednosti „težina“ među varijablama.

S obzirom na broj slojeva postoje brojni algoritmi neuronskih mreža. Stoga se razlikuju dva tipa algoritama neuronskih mreža: dvoslojni i višeslojni. Algoritmi neuronskih mreža prema tipu učenja dijele se na nadgledane i nenadgledane. Kod nadgledanog tipa učenja poznate su

vrijednosti izlaznih varijabli na skupu podataka za učenje mreže (npr. Backpropagation), dok kod nenadgledanog tipa učenja nisu poznate vrijednosti izlaznih varijabli na skupu podataka za učenje mreže.

Najpoznatiji i najčešće upotrebljavani algoritam primijenjen na učenje i testiranje višeslojnih perceptron mreža je algoritam mreža „širenje unatrag“. Algoritam ove mreže (eng. backpropagation) bio je presudan za široku komercijalnu upotrebu ove metodologije, te neuronske mreže učinio široko upotrebljavanom i popularnom metodom u različitim područjima. Njegov prvi kreator bio je Paul Werbos 1974., a usavršili su je 1986. godine Rumelhart, Hinton i Williams. Slikom 2. prikazana je struktura mreže „širenje unatrag“.

Slika 2. Backpropagation mreža



Izvor: <https://moodle.carnet.hr/mod/resource/view.php?id=23158> (01.02.2012.)

Strukturu ove mreže čini ulazni sloj, izlazni sloj i najmanje jedan skriveni sloj. Stvaranje neuronske mreže odvija se kroz tri faze rada (Slika 3.). Najprije se odvija faza učenja ili treniranja mreže. Učenje je proces mijenjanja težina u mreži, a odvija se kao odgovor na podatke izvana koji su predstavljeni ulaznom sloju i u nekim mrežama izlaznom sloju. Podatci koji se predstavljaju izlaznom sloju jesu željene vrijednosti izlaznih varijabli. Prije samog učenja potrebno je definirati model (ulazne i izlazne varijable) te prikupiti podatke iz prošlosti na kojima će se primijeniti mreža. Podatke je potrebno podijeliti na tri uzorka: za treniranje, unakrsnu validaciju i testiranje. Pravila za ovu podjelu nema, osim što se preporučuje da se najveći dio podataka primijeni za treniranje mreže, a manji dio podataka za testiranje i validaciju (npr. 70% za treniranje, 15% za testiranje i 15% za validaciju). Nakon što je definiran model, pripremljeni ulazni podatci i izabran algoritam te pravilo učenja i potrebne funkcije, mrežu treba učiti ili trenirati na pripremljenim podacima iz prošlosti, kako bi ona prepoznala vezu između podataka i bila u mogućnosti na osnovi ulaznih vrijednosti predviđati izlaze. Sama faza učenja je proces podešavanja težina u mreži, koja se odvija u više iteracija ili prolazak kroz mrežu. Jedna iteracija predstavlja učitavanje jednog promatranja iz podataka. Obično se mreža trenira u nekoliko

tisuća iteracija.⁵ U fazi unakrsne validacije mreža nastoji optimirati duljinu treniranja, broj skrivenih neurona i parametre.⁶ Testiranje mreže je treća faza rada neuronske mreže i ona je odlučujuća za ocjenjivanje mreže. Razlika između faze učenja i faze testiranja je u tome što u ovoj fazi mreža više ne uči. Ocjenjivanje mreže obavlja se izračunavanjem greške ili nekog drugog mjerila točnosti (npr. stope točnosti), na način da se izlaz mreže uspoređuje sa stvarnim izlazima. U praksi se koristi mreža s najboljim test rezultatom.

Slika 3. Faze rada neuronske mreže



Izvor: <https://moodle.carnet.hr/mod/resource/view.php?id=23158>, (4.2.2012.)

⁵ Dostupno na: <http://eris.foi.hr/11neuronske/nn-predavanje4.html> (28.01.2012.)
⁶ Dostupno na: <https://moodle.carnet.hr/mod/resource/view.php?id=23158>, (4.2.2012.)

Ne postoji univerzalni model neuralne mreže koji bi bio primjenjiv na sve vrste problema. Izborom skupa podataka za treniranje počinje etapa inicijalizacije mreže koja predstavlja prede-tapu treniranju (učanju) mreže. Ako su rezultati zadovoljavajući, mreža se primjenjuje u praksi, a ako nisu, slijedi povrat na neki od prethodnih koraka i ponavljanje procesa. Katkad je dovoljno promijeniti broj skrivenih slojeva neurona ili pak model same mreže. Modeli neuronskih mreža primjenjuju se u numeričkim analizama u okviru ekonomskih aplikacija koje služe za progno-ziranje određenih vrijednosti nekih pojava, za klasifikaciju pojava stanovite skupine, za planira-nje itd.⁷ S obzirom na to da su neuronske mreže zamišljene da djeluju slično ljudskom mozgu, one se upotrebljavaju u analizi rizika i prognoziranju, kao npr. vrijednosti dionica.⁸

Neuronske se mreže prilikom rudarenja poda-taka često kombiniraju s ostalim metodama ruda-renja podataka. Razlog tome djelomično leži u činjenici da sama metoda neuronskih mreža može vrlo teško jamčiti dobru interpretaciju rezultata, pa se zato moraju koristiti dodatne metode što služe upravo kvalitetnoj interpretaciji rezultata.⁹ Neuronske mreže mogu se koristiti kao alat za segmentaciju. U današnje vrijeme nastaje, razvija se i proučava velik broj sustava koji pokazuju određene značajke ljudske inteligencije kao što je npr. razumijevanje i obrada prirodnog jezika. Iako je zabilježen velik napredak u razvoju inteligentnih sustava od samih početaka do danas, nijedan stroj još uvijek nije u mogućnosti „kopirati“ sva svojstva ljudske inteligencije. Međutim postavljaju se pitanje hoće li to ikada biti moguće.

3. Pregled prethodnih istraživanja

Razvojem inteligentnih sustava posljednja dva desetljeća ostvareno je mnoštvo unaprjeđenja koja su inspirirana biološkim neuronskim sustavom. Neuronske mreže primjenjivale su se za rješavanje niza zadataka – od prepoznavanja uzoraka, predviđanja, dijagnosticanja stanja, softverskih senzora, modeliranja i identifikiranja, vođenja i optimiranja procesa itd.¹⁰ Međutim u

7 Klepac, G., Panian, Ž.: Poslovna inteligencija, Masmedia, Zagreb, 2003., str. 319.

8 Čerić, V., Varga, M.: Informacijska tehnologija u poslovanju, Element, Zagreb, 2004., str. 217.

9 Klepac, G., Panian, Ž.: Poslovna inteligencija, Masmedia, Zagreb, 2003., str. 327.

10 Bolf, N., Jerbić, I.: Primjena umjetnih neuronskih mreža, Kemija u industriji, Vol. 55, br. 11., 2006., str. 457.

online bazama podataka i do danas objavljenim publikacijama i relevantnim časopisima nisu pronađeni radovi čiji su autori proveli istraživanje o klasifikaciji korisnika prema posjetima knjižnici. Stoga su do danas neuronske mreže bile nepoznata metoda umjetne inteligencije u području knjižničarstva kada je u pitanju klasifikacija posjeta njezinih korisnika. Određeni autori koristili su ovu metodu umjetne inteligencije u knjižničarstvu, ali samo za istraživanje u kontekstu pretraživanja podataka i utvrđivanja veza među njima. Konačno, opisani pristup može biti polazna točka za buduća istraživanja klasifikacije studenata u fakultetskim knjižnicama koji može pokazivati i bolje rezultate koristeći veći podatkovni skup. Stoga se u nastavku navodi nekoliko zanimljivih istraživanja pojedinih autora kojima je zajednička samo metoda neuronskih mreža.

Do sada su klasifikaciju studenata, kao npr. potporu navedenom području odlučivanja u obrazovanju i dr. kontekstima, istraživali su brojni autori. Autori Zekić-Sušac, Frajman-Jakšić i Drvenkar (2009.) u svom radu istraživali su predviđanje uspješnosti studenata pomoću ovih inteligentnih metoda. Naime model je kreiran na osnovi demografskih podataka koji su prepoznati kao ključne varijable koje utječu na uspješnost studiranja. Najbolji rezultat dobiven je pomoću višeslojne perceptron mreže. Istraživanje je pokazalo da na uspješnost studenata najviše utječe polaganje kolegija putem sustava kolokviranja, te zatim prisutstvo na vježbama. Ocjene imaju također veliko značenje za studente, kao i primarne stipendije koja zasigurno, potiče studente da budu što uspješniji. U ovom je istraživanju stablo odlučivanja dalo drugačiji prikaz značajnosti varijabli od neuronskih mreža. Zaključak ovog istraživanja ogleda se u tome da je istraživanje potrebno nadopuniti i rezultatima drugih metoda, te bi se na taj način mogao stvoriti inteligentni sustav za potporu odlučivanju u obrazovanju, koji bi u konačnici dao doprinos većoj uspješnosti studiranja.

Kada je riječ o metodi klasifikacije, osim klasifikacije studenata, autori su na primjer istraživali i druge složene operacije kao što je npr. procjenjivanje kvalitete obradive površine, i sl. te došli do zanimljivih rezultata. Vezano uz istraživanje kvalitete obradive površine, istraživači su izradili model neuronske mreže te koristili kontinuirane varijable i mrežu „širenja unazad“. Za ulazne varijable koristili su utjecajne faktore koji utječu na

hrapavost obrađivane površine. Šimunović, Šajić i Lujčić (2009.) koristili su varijable u istraživanju poput materijala, vrste alata, dubine rezanja te brzina rezanja. Za učenje mreže primijenili su Delta pravilo, a prijenosna funkcija koju su koristili jest sigmoidna. Rezultat neuronske mreže dao je strukturu neuronske mreže, koji se sastoji od pet ulaznih neurona, jednog izlaznog neurona, te pet skrivenih neurona. Istraživači su došli do zaključka da bi se izbacivanjem neke od varijabli iz modela, znatno povećala greška, te smanjila stopa točnosti klasifikacije. Stoga se analizom osjetljivosti došlo do zaključka da su sve ulazne varijable izrazito značajne za izradu modela. Izrađeni model predikcije u konačnici je dao rezultat s greškom koja je manje od 5% što potvrđuje prihvatljivost modela u praksi. Ključan rezultat istraživanja odnosi se na činjenicu da se naučena mreža može samostalno koristiti ili integrirati u ERP sustav. Na taj način ubrzao bi se režim obrade površine te bi se skratilo tehnološko vrijeme pripreme.

S obzirom na to da nije pronađen doprinos neuronskih mreža u ovom području, kreirani model u nastavku može biti polazna osnova za daljnja istraživanja kako bi se različiti modeli u buduć-

nosti mogli uspoređivati, a sve u cilju pomoći knjižničarima da učinkovitije isplaniraju proces vođenja i organizacije posla te utvrde tko su za-pravo njihovi korisnici i kakve su njihove navike.

4. Podatci za istraživanje

Podatci za istraživanje prikupljeni su putem interne statistike fakultetske knjižnice Ekonomskog fakulteta u Osijeku za akademsku godinu 2011./2012. među studentima koji su korisnici fakultetske knjižnice. Prikupljeni su i obrađeni podatci na uzorku od 130 studenata diplomskog studija koji su posjetili fakultetsku knjižnicu u prvom tjednu mjeseca studenog 2011. godine. Veličina uzorka znatno je smanjena zbog eliminiranja svih nepotpuno ispunjenih obrazaca. Kreiran je model dolazaka studenta u fakultetsku knjižnicu, i to na temelju izlazne varijable prosječnog broja dolazaka studenata u jednome radnom danu te 7 ulaznih varijabli. Za izrade ovog modela korištene su ulazne varijable koje su prikazane Tablicom 1.

Ulazne varijable izabrane su na temelju pretpo-

Tablica 1. Ulazne varijable modela

Rb.	NAZIV VARIJABLE	TIP VARIJABLE
1.	Datum	Informativna varijabla
2.	Matični broj studenta	Informativna varijabla
3.	Veličina knjižnog fonda	Kontinuirana varijabla
4.	Broj posuđenih knjiga od strane studenta	Kontinuirana varijabla
5.	Broj posuđenih časopisa od strane studenta	Kontinuirana varijabla
6.	Rad u čitaonici	Kategorijalna varijabla
7.	Prosječno vrijeme rada u čitaonici	Kontinuirana varijabla
8.	Upotreba računala	Kategorijalna varijabla
9.	Godina studija	Kontinuirana varijabla
10.	Spol	Kategorijalna varijabla

stavki da utječu na odluku studenta da u jednom radnom danu dođu i da se ponovno vrate u fakultetsku knjižnicu.

Nadalje u radu kreiran je model koji na temelju podataka o studentima diplomskog studija nastoji klasificirati studente na način da se kategoriziraju u jednu od dvije kategorije ili klase: klasa 0 – student posjetio knjižnicu samo jednom u

jednom radnom danu, te klasa 1 – student posjetio fakultetsku knjižnicu u jednom radnom danu 2 ili više puta. Na ovaj način problem je razvrstan u skupine, a cilj modela je ispravno svrstati koji studenti pripadaju u koju klasu. Testiran je i treniran model na sigmoidnoj i hiperboličko-tangentnoj prijenosnoj funkciji u programu Statistica. Izabrani algoritam za obje arhitekture je Backpropagati-

on (MLP).

Varijable datum i matični broj studenta samo su informativne varijable, te one nisu korištene pri izradivanju ovog modela.

Struktura promatranog uzroka kod neuronskih

mreža raspodijeljena je na način da 70% podataka pripada fazi treniranje, 15% podataka odnosi se na testiranje i 15% podataka na validaciju, što je prikazano Tablicom 2.

Kao što se i preporučuje u kreiranju modela

Tablica 2. Struktura ukupnog uzorka i njegova podjela

Faze rada	Broj slučajeva u uzorku	% slučajeva u uzorku
Treniranje	91	70%
Unakrsna validacija	19.5	15%
Testiranje	19.5	15%
Ukupno	130	100%

neuronskih mreža, uzorak je podijeljen na način da se najveći dio podataka odnosi na treniranje, a manji i jednaki dio na validaciju i testiranje.

4.1. Rezultati točnosti klasifikacije najboljeg modela neuronske mreže i analiza značajnosti ulaznih varijabli za promatrani problem

U provedenom istraživanju trenirane su i testirane dvije arhitekture neuronskih mreža i prijenosnih funkcija. Prijenosna funkcija kori-

štena u prvoj arhitekturi je sigmoidna (Logistic), a algoritam MLP (multi-layer perceptron). U drugoj arhitekturi prijenosna funkcija koja je korištena je hiperboličko-tangentna funkcija (Tanh), a algoritam MLP (multi-layer perceptron). Za svaku arhitekturu dobivena je stopa točnosti klasifikacije na uzorku za treniranje i testiranje. Rezultati dobiveni za navedene arhitekture prikazani su Tablicom 3.

Rezultati najboljeg modela dobiveni su u prvom

Tablica 3. Rezultati metode neuronskih mreža na uzorku za validaciju

Rb.	Broj skrivenih neurona i struktura mreže	Prijenosna funkcija	Stopa klasifikacija za klasu 0* (%)	Stopa klasifikacija za klasu 1* (%)	Ukupna stopa klasifikacije (%)
1.	18 (10-18-2)	sigmoidna (Logistic)	69	66	68,42%
2.	15 (10-15-2)	hiperboličko-tangentna funkcija (Tanh)	53	66	57,89%

* 0 = postotak ispravne klasifikacije za studente koji se nisu vratili u fakultetsku knjižnicu

* 1 = postotak ispravne klasifikacije za studente koji su se vratili u fakultetsku knjižnicu

testiranju neuronske mreže. U ovoj arhitekturi mreža je uspješno razvrstala 73,91% slučajeva na uzorku za treniranje, zatim stope klasifikacije za svaku kategoriju posebno iznose 69% i 66% na uzorku za testiranje, te na uzorku za validaciju, program je točno razvrstao 68,42% podataka. Stoga ukupna stopa klasifikacije iznosi 68,42% na

uzorku za testiranje, dok prosječna stopa točnosti iznosi 67,5%. Navedena tri parametra (uspješnost treniranja, testiranja i validacije) pokazuju koliko je jaka veza između onoga što je mreža izračunala i što je trebalo izračunati. Možemo zaključiti da veza i nije vrlo jaka s rezultatom od 0,68 jer dobiveni rezultat treba biti što bliže vrijednosti

od 100%.

Za promatrani model određeno je da mreža koristi dva minimalno skrivena neurona te maksimalno dvadeset skrivenih neurona. Trenirano je dvadeset različitih arhitektura neuronske mreže, od kojih je zadržana ona koja daje najveću stopu točnosti na uzorku za validaciju, što znači da se konačna struktura mreže sastoji od 10 ulaznih neurona, 18 skrivenih neurona u mreži i 2 kategorije jedne varijable u izlaznom sloju.

U nastavku je opisana matrica konfuzije za dobivenu najbolju mrežu. Matrica konfuzije pokazuje broj slučajeva koji su ispravno razvrstani i

broj slučajeva koji nisu ispravno razvrstani u svakoj kategoriji. U Tablici 4. prikazana je matrica konfuzije za najbolju mrežu na uzorku za validaciju, a koji je iznosio 15% od ukupnog uzorka. U matrici se nalazi po jedan stupac i redak za svaku izlaznu klasu. Podaci u redcima su željene klase, a podaci u stupcima su dobivene klase. Brojevi na dijagonali govore o postotku ispravno klasificiranih slučajeva iz svake klase. Cilj je bio dobiti što veću ukupnu stopu klasifikacije.

U Tablici 4. može se vidjeti da je u uzorku za

Tablica 4. Matrica konfuzije za najbolju mrežu

Željene (stvarne) kategorije studenata		
Predviđene kategorije studenata	Klasa 0 – Studenti koji se nisu vratili_BIN-0	Klasa 1 – Studenti koji su se vratili_BIN-1
1.MLP 10-18-2-0	9	2
1.MLP 10-18-2-1	4	4

validaciju postojalo trinaest slučajeva studenata koji se nisu vratili u fakultetsku knjižnicu isti dan i šest slučajeva kada su se studenti vratili u fakultetsku knjižnicu. U prvom stupcu u kojem se nalazi student klasa 0 vidljivo je da je od 13 slučajeva mreža ispravno svrstala 9, dok je 4 po-

grešno svrstala u klasu 1. U stupcu klasa 1, od 6 slučajeva mreža ispravno je svrstala 2, a pogrešno 4 slučaja. Nakon objašnjenja dobivene matrice konfuzije za najbolji model, slijedi opis matrice klasifikacije za najbolji model.

Stoga točnosti klasifikacije najboljeg modela

Tablica 5. Stope točnosti klasifikacije najboljeg modela

Željene (stvarne) kategorije studenata		
	Klasa 0 – Studenti koji se nisu vratili_BIN-0	Klasa 1 – Studenti koji su se vratili_BIN-1
Ukupno	13,00000	6,00000
Broj točno razvrstanih slučajeva	9,00000	4,00000
Broj netočno razvrstanih slučajeva	4,00000	2,00000
Postotak točno razvrstanih slučajeva	69,00000	66,00000
Postotak netočno razvrstanih slučajeva	30,00000	33,00000

na uzorku za validaciju u postotcima dobivena je u Tablici 5., te je iz te tablice također vidljivo da se 13 studenata stvarno nije vratilo u fakultetsku knjižnicu isti dan i da ih je mreža ispravno svrstala 9, a 4 pogrešno. Iz prikazane tablice vidljivo je i kolike su stope točnosti klasifikacije za svaku klasu posebno (u retku postotak točno razvrstanih slučajeva). Od ukupno 13 slučajeva studenata koji se nisu vratili u fakultetsku knjižnicu isti dan, mreža je 69% slučajeva točno razvrstala, dok je postotak netočno razvrstanih slučajeva iznosio

30%. Kod studenta koji su se vratili u fakultetsku knjižnicu isti dan mreža je točno razvrstala 66% slučajeva, dok je postotak netočno razvrstanih slučajeva iznosio 33%.

Kako bi se dobio detaljniji uvid u značaj ulaznih varijabli modela, izvršena je analiza osjetljivosti pri kojoj se analizira utjecaj pojedine ulazne varijable modela na izlaznu varijablu. U Tablici 6. prikazani su dobiveni rezultati za najbolju mrežu.

Grafikonom 1. prikazana je značajnost poje-

Tablica 6. Analiza osjetljivosti

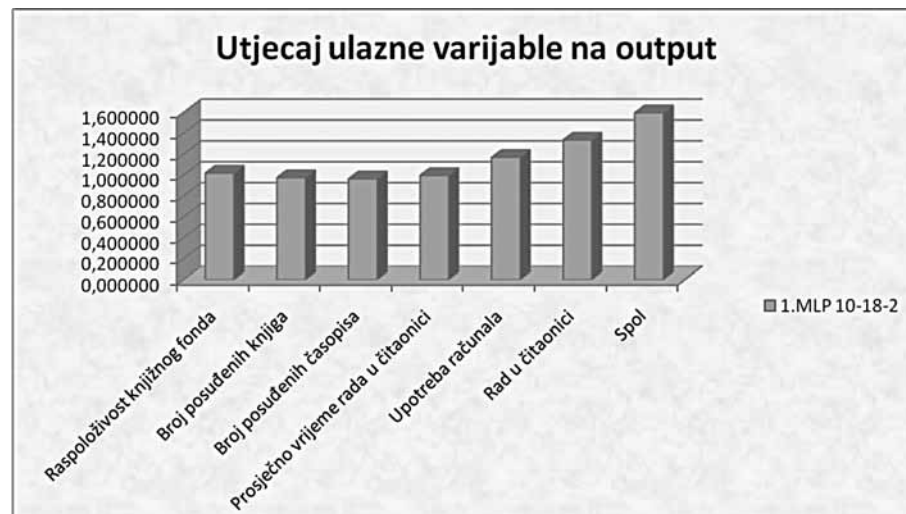
Sensitivity analysis for Prosječan broj dolazaka studenata u jednom radnom danu_BIN (Podatci za učitavanje_Zvonimir.sta) Samples: Validation							
	Raspoloživost knjižnog fonda	Broj posuđenih knjiga	Broj posuđenih časopisa	Prosječno vrijeme rada u čitaonici	Upotreba računala	Rad u čitaonici	Spol
1.MLP 10-18-2	1,016882	0,973115	0,961076	0,993759	1,166843	1,333435	1,593503

dine ulazne varijable modela najbolje neuronske mreže, kroz koeficijente osjetljivosti izlazne

varijable na ulazne.

Analizom osjetljivosti prikazan je odnos

Grafikon 1. Analiza osjetljivosti ulaznih varijabli na izlaznu varijablu Prosječan broj dolazaka studenata u jednom radnom danu_BIN za primjer modela najbolje neuronske mreže



značajnosti varijabli. Prema rezultatima analize osjetljivosti za najbolji model moglo bi se reći kako najveći utjecaj na povratak studenata u fakultetsku knjižnicu ima spol. Nakon toga slijedi mogućnost rada u čitaonici i uporaba računala. Važno je naglasiti da ne postoje velika odstupanja, te nijedna varijabla nije ni dvostruko značajnija od neke druge. Najmanji utjecaj na povratak studenata u fakultetsku knjižnicu prema ovoj analizi ima mogućnost posudbe i pregledavanja sadržaja časopisa. Ovom je analizom potvrđena činjenica da se časopisi (u pisanom, klasičnom obliku) najmanje koriste u fakultetskoj knjižnici, što je već u svojem istraživanju navela skupina autora.¹¹ Današnja praksa studenata je u čitaonici koristiti elektroničku inačicu svih serijskih publikacija.

Provedenu analizu osjetljivosti i značajnosti varijabli pogodno je koristiti za daljnje poboljšanje modela. Poboljšanje se može ostvariti na način da se poveća broj ulaznih varijabli poput kategorijalne varijable seminarska obveza studenta (da/ne) i kontinuirane varijable bi se odnosile na vrijeme putovanja koje je potrebno korisniku da dođe do fakultetske knjižnice. Također bi pogodno bilo povećati uzorak te uključiti u kreiranje modela sve korisnike fakultetske knjižnice. Istraživanjem dobivena analiza osjetljivosti pokazuje da je klasifikacija putem neuronskih mreža vrlo precizna, premda bi u budućim istraživanjima bilo korisno uključiti i druge metode rudarenja poput stabla odlučivanja radi usporedbe dobivenih rezultata.

Zaključak

Rad se bavi kreiranjem modela klasifikacije studenata prema posjetu fakultetskoj knjižnici. Metodološka osnova za izradu ovog rada bila je metoda umjetne inteligencije – metoda neuronskih mreža. Na temelju sedam ulaznih i jedne izlazne varijable izrađene su dvije neuronske mreže s različitim arhitekturama. Njihovom usporedbom, odabran je najbolji model prema rezultatu na uzorku za validaciju.

Na temelju rezultata najbolje neuronske mreže dobivena je analiza osjetljivosti ulaznih varijabli na izlaznu varijablu, matrica konfuzije i matrica klasifikacije.

Veća točnost klasifikacije dobivena je kod

studenata koji su posjetili fakultetsku knjižnicu samo jedanput u jednome radnom danu.

Premda najbolja neuronska mreža nije i dalje preporučljiva za upotrebu u praksi, mreža bi se mogla možda poboljšati mijenjanjem ulaznih varijabli i povećanjem uzorka. Neki od primjera novih ulaznih varijabli koje bi se mogle koristiti za daljnja istraživanja jesu npr. kategorijalna varijabla koja bi se odnosila na dinamiku seminarskih obveza studenta u jednom turnusu i kontinuirana varijabla koja bi pokazivala potrebno vrijeme putovanja korisnika od mjesta stanovanja ili prebivališta do fakultetske knjižnice. Daljnjim poboljšanjima modela, ovaj bi se model mogao primjenjivati kod svih knjižnica u jednom sveučilišnom sustavu. Također se preporučuje u novim, budućim istraživanjima kombinacija metode neuronskih mreža s drugim metodama rudarenja podataka, kako bi se uvidjelo koja metoda daje preciznije i točnije rezultate. Ovakav bi model trebao pomoći knjižničarima fakultetske knjižnice da saznaju:

- koji su korisnici prema svojim značajkama stalni korisnici knjižnice
- koji studenti nemaju potrebu vraćati se više od jedanput dnevno u knjižnicu
- uopće nemaju potrebu za posjetom knjižnici.

Prikupljanjem i pohranjivanjem dobivenih podataka u bazu podataka omogućila bi se određena predviđanja dolaska studenata u knjižnicu.

Ovakav model dobiven u radu nužno je stalno poboljšavati i prilagođavati potrebama potencijalnih korisnika knjižnice.

Neuronske su mreže vrlo prilagodljiva i pregledna metoda umjetne inteligencije, stoga analiza posjeta fakultetskoj knjižnici s točnim podacima o korisnicima knjižnice pomoću inteligentnih metoda svakako može dovesti do kreiranja sustava koji bi omogućili brzu i jednostavnu klasifikaciju studenata u području knjižničarstva.

¹¹ Vidi Ivić, K., Borić, M., Marinković, R.: Fakultetska knjižnica (interaktivno i statistički), Ekonomski vjesnik, Vol. XXIV, broj 2., 2011., str. 405 – 414.

LITERATURA

Knjige

1. Bolj, N., Jerbić, I.: *Primjena umjetnih neuronskih mreža, Kemija u industriji*, Vol. 55, No. 11., 2006., 457 – 468.
2. Čerić, V., Varga, M.: *Informacijska tehnologija u poslovanju, Element, Zagreb, 2004.*
3. *Hrvatska enciklopedija, br. 7 Mal – Nj, Leksikografski zavod Miroslav Krleža, Zagreb, 2005.*
4. Ivić, K., Borić, M., Marinković, R.: *Fakultetska knjižnica (interaktivno i statistički), Ekonomski vjesnik, Vol. XXIV, broj 2., 2011., str. 405 – 414.*
5. Klepac, G., Mršić, L.: *Poslovna inteligencija kroz poslovne slučajeve, Liderpress/TimPress, Zagreb, 2006.*
6. Klepac, G., Panian, Ž.: *Poslovna inteligencija, Masmedia, Zagreb, 2003.*
7. Klepac, G.: *Primjena inteligentnih računalnih metoda u managementu, Sinergija, Zagreb, 2001.*
8. Sinković, V.: *Informacijske mreže, Školska knjiga, Zagreb, 1994*

Internet

9. Dalbelo Bašić, B., *Nastavni materijali za kolegij Umjetna inteligencija, http://www.fer.unizg.hr/download/repository/UII_1_uvod.pdf 01.02.2012.*
10. Zekić-Sušac, M., Frajman-Jakšić, A., Drvenkar, N., *Neuronske mreže i stabla odlučivanja za predviđanje uspješnosti studiranja, http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=73924, 25.01.2012.*
11. ERIS, *Edukacijski repozitorij za Inteligentne Sustave: Neuronske mreže, <http://eris.foi.hr/11neuronske/nn-predavanje4.html> 28.01.2012.*
12. Šimunović, G., Šarić, T., Lujić, R., *Primjena neuronskih mreža u procjenjivanju kvalitete obrađivane površine, http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=60717, 05.02.2012.*
13. Tomašević, M., *Matematičke metode kao čimbenik odlučivanja o uspješnosti menadžmenta, http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=20489*
14. Zekić-Sušac, M., *Nastavni materijali za kolegij Sustavi poslovne inteligencije, Poglavlje 3. Neuronske mreže, <https://moodle.carnet.hr/mod/resource/view.php?id=23158>, 4.2.2012.*
15. Zekić-Sušac, M., Frajman-Jakšić, A., Drvenkar, N., *Neuronske mreže i stabla odlučivanja za predviđanje uspješnosti studiranja, http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=73924, 25.01.2012.*
16. *Chapter 7 Neural Network, Neurophysiological Motivation, http://www.byclb.com/TR/Tutorials/neural_networks/ch7_1.htm 19.03.2012.*

Kata Ivić

Zvonimir Jurković

Rozalija Marinković

DEVELOPMENT OF NEURAL NETWORK MODELS FOR THE FREQUENCY OF USE OF FACULTY LIBRARY

ABSTRACT

Neural network, an artificial intelligence method, is being increasingly used in today's business as part of the decision making support systems. A key factor in the implementation of the method is the knowledge of the person using the mentioned method which will allow that person to gather adequate research data and interpret them accurately. Understanding and predictions of frequency of user visits are valuable to faculty libraries as they allow for their smooth functioning without creating waiting lists. The aim of this paper is to group students in one of two categories, i.e., according to average number of student visits to the library in the course of one day. The research included continuous and categorical variables, i.e., seven input variables and one output variable for testing of architectures. Two neural network architectures were tested using backpropagation algorithms multilayered perceptron network and the best model was selected. A sensitivity analysis was conducted based on which it was discovered that gender, a possibility to work in the library, and a possibility to use a computer have the strongest influence on the output variable. The results of modelling show that the neural network perceives well the relationship between input and output variables. The model created indicates there is possibility and need for application of artificial intelligence methods in the field of library science. The model can be used for further research in this area providing that certain improvements are made, such as introduction of additional input variables and increase of sample size. Model performance was measured by the total classification rate.

Keywords:

Neural networks, faculty library, classification of students, multilayered perceptron, sensitivity analysis