

VEŠTAČKE NEURONSKE MREŽE U PREDVIĐANJU USPEŠNOSTI STUDENATA U PROGRAMIRANJU

Nebojša Stanković¹, Marija Blagojević¹

¹Univerzitet u Kragujevcu, Fakultet tehničkih nauka Čačak, Čačak, Srbija

Kratak sadržaj: U radu je dat pregled mogućnosti veštačkih neuronskih mreža za predviđanje uspešnosti studenata. Takođe, dat je predlog modela neuronskih mreža uz definisanje ulaznih i izlaznih parametara. Rad prikazuje način utvrđivanja kompozitne mere uspešnosti u programiranju.

Ključne reči: programiranje, uspešnost, veštačke neuronske mreže

ARTIFICIAL NEURAL NETWORKS IN PREDICTING STUDENTS' SUCCESS IN PROGRAMMING

Abstract: The paper provides an overview of the capabilities of artificial neural networks to predict student performance. Also, a model of neural networks is proposed, with input and output parameters defined. The paper presents a way to determine a composite measure of programming success.

Key Words: programming, success, artificial neural networks

1. UVOD

U današnjem digitalnom svetu bez programiranja skoro da ništa nije moguće. Posedovanje veština programiranja je neophodno i one predstavljaju ključnu kompetenciju za sve sfere ljudskog delovanja.

Programski jezici postoje preko sedamdeset godina. Osnovni mehanizmi njegove primene nisu se menjali. U današnjem digitalnom svetu biti uspešan inženjer, ekonomista, biznismen,... zahteva korišćenje računarskog modelovanja (istovremeno poznavanje određenih naučnih modela i programskih jezika). Za rešavanje određenog problema neophodna je ideja i transformacija te ideje u formalni model. Stručnjak, koji ima tu sposobnost (naučnik, praktičar), želi da primeni ovaj model kao IT rešenje ili da proveri njegovu validnost pomoću računarske simulacije. Zbog toga, stručnjak kontaktira drugu osobu/tim koja ima sposobnost da komunicira sa računarima pomoću programskih jezika.

U tom smislu, savremeni kurikulumi studijskih programa u oblasti informacione tehnologije obuhvataju proučavanje različitih programskih jezika i koncepata programiranja. Posebna pažnja posvećena je analizi uspešnosti u programiranju putem različitih metoda i tehnika. Uspešni studenti u programiranju lakše dolaze do prakse u IT kompanijama, a zatim i do posla. Upravo iz tog razloga, kao i zbog činjenice da je u Srbiji neophodno što više stručnjaka iz pomenute oblasti svrha istraživanja odnosi se na ispitivanje mogućnosti primene veštačkih neuronskih mreža za predviđanje uspešnosti u programiranju.

Cilj rada se odnosi na definisanje polaznih postavki u istraživanju, kroz definisanje ulaznih parametara u mreži kao i definisanje parametara koji se odnosi na uspešnost studenata. Takođe, u ovom radu će biti predstavljen model neuronskih mreža.

2. PREGLED SRODNIH ISTRAŽIVANJA

Postoji veliki broj radova koji se bave analizom uspešnosti studenata. U odnosu na srodna istraživanja u ovom radu predložen je originalni model koji proističe iz višedecenijskog rada u oblasti programiranja sa studentima.

Bucos i Drăgulescu u [1] koriste tehnike obrazovnog majninga podataka u cilju otkrivanja znanja u okruženju u kome studenti uče. Pomenute tehnike koriste i kako bi predvideli buduće trendove i ponašanja studenata. Autori su u [1] kreirali pet modela za klasifikaciju kako bi identifikovali model sa najboljim performansama. Modeli su kreirani u cilju identifikovanja studenata koji su "rizični" kao i u cilju predviđanja rezultata studenata.

Kovačić u [2] istražuje u domenu sociodemografskih varijabli (godine, pol, obrazovanje, nacionalnost, radni status) kao i varijable koje se odnose na okruženje u kome uče (kurs) u cilju predviđanja uspešnih i neuspešnih studenata. Ovo istraživanje u svojim rezultatima prikazuje nacionalnost i program kursa kao karakteristike koje najviše utiču na uspešnost/neuspešnost studenata.

Istraživanje prikazano u [3] je fokusirano na implementiranje Naïve Bayes klasifikatora kako bi se poboljšala preciznost predviđanja uspešnosti studenata. Uspešnost predviđanja je preko 90%. Osim toga, autori upućuju na veoma lak način interpretacije dobijenih rezultata. Poseban značaj ogleda se u identifikaciji studenata koji su u većem riziku za odustajanje.

U [4] je prikazano istraživanje koje integriše podatke o studentima koji se odnose na demografske podatke, ponašanje u okviru kurseva, kao i njihovu uspešnost. Nakon toga se primenjuju klasične statističke metode ali i metode majninga podataka kako bi se na najbolji način identifikovali prediktori i uspešnost studenata.

3. DEFINISANJE PARAMETARA MREŽE

Učesnici istraživanja su studenti Fakulteta tehničkih nauka u Čačku, četvorogodišnjih studija na studijskom programu Informacione tehnologije. Istraživanje je obavljeno u drugom semestru školske 2018/19. godine. U istraživanju je učestvovalo ukupno 159 studenata na sve 4 godine, po godinama studiranja: prva godini 63 studenta, druga godina 34, treća godina 32 i četvrta godina 30 studenta. Kao rezultat ovog istraživanja kreiran je model veštačke neuronske mreže koji dovoljno predviđa uspešnost studenata u programiranju, utvrđujući faktore i prediktore uspešnosti u usvajanju znanja i veština programiranja.

Parametri neuronske mreže definisani su kroz ulazne i izlazne parametre. U tabeli 1 su definisane kategorije parametara veštačke neuronske mreže, kao i izvori ulaznih parametara iz kojih su oni preuzeti/dobijeni.

Tabela 1: Pregled ulaznih i izlaznih parametara u veštaškoj neuronskoj mreži

Parametri	Kategorije	Izvor podataka
Ulazni parametri	Sociodemografski parametri	Interna evidencija fakulteta
	Parametri koji se odnose na prethodno obrazovanje	Interna evidencija fakulteta, anketni upitnik
	Parametri koji se odnose na trenutno obrazovanje	Interna evidencije fakulteta, evidencije nastavnika/asistenata, anketni upitnik (subjektivne procene studenata)
Izlazni parametar	Uspešnost u programiranju	

Od ulaznih parametara u mreži se koriste:

- Sociodemografski parametri:* pol i veličina mesta u kome je završena srednja škola. Istraživanje je obuhvatalo 51 studenta ženskog pola (32%) i 108 studenta muškog pola (68%). Mesta iz kojih dolaze učesnici istraživanja su klasifikovana na mala (manje od 20.000 stanovnika – 11%), srednja (20.000-100.000 – 43%) i velika (više od 100.000 stanovnika – 46%).
- Parametri koji se odnose na prethodno obrazovanje:* postojanje IT odeljenja u mestu u kome je ispitanik završio srednju školu, poeni iz srednje škole, poeni i predmet koji je polagan na prijemnom ispitu, formalno i neformalno obrazovanje vezano za IT (formalno - obrazovanje u osnovnoj, srednjoj školi, na fakultetu, neformalno - obrazovanje van škole), razlog upisa studijskog programa Informacione tehnologije. U tabelama 2-4 data je statistika o pojedinim parametrima koji se odnose na prethodno obrazovanje.

Tabela 2: Statistika uspeha iz srednje škole i prijemnom ispitu učesnika istraživanja (po godinama studija)

Godina studiranja	Bodovi srednja škola (16-40)	Bodovi prijemni ispit (0-60)	Predmet polaganja	
			Informatika	Matematika
Prva	36,36	41,46	76%	24%
Druga	34,48	46,08	88%	12%
Treća	35,29	45,88	78%	22%
Četvrta	33,10	50,52	90%	10%
Svega	35,13	44,78	82%	18%

Tabela 3: Formalno i neformalno obrazovanje za određene programske jezike (po veličini mesta)

Veličina mesta	Formalno obrazovanje (%)								Neformalno obrazovanje (%)							
	VBA	Delphi	C	C++	C#	Java	Python	Matlab	VBA	Delphi	C	C++	C#	Java	Python	Matlab
Veliko (>100.000)	16,4	5,5	97,3	93,2	53,4	65,8	2,7	45,2	5,5	5,5	24,7	21,9	19,2	41,1	11,0	2,7
Malo (<20.000)	11,8	5,9	82,4	76,5	29,4	52,9	5,9	29,4	17,6	0,0	35,3	23,5	23,5	29,4	5,9	5,9
Srednje (20.000-100.000)	14,5	18,8	95,7	85,5	27,5	52,2	1,4	27,5	5,8	4,3	44,9	36,2	23,2	49,3	11,6	7,2
Svega	15,1	11,3	95,0	88,1	39,6	58,5	2,5	35,8	6,9	4,4	34,6	28,3	21,4	43,4	10,7	5,0

Tabela 4: Razlozi upisa na smer IT (po polu)

Pol	Sticanje opšteg programerskog znanja	Sigurno zaposlenje, dobra zarada	WEB i grafički dizajn	Administriranje računarskih mreža i baza podataka	Održavanje računarskih mreža
Ženski	37,25%	43,14%	15,67%	3,92%	0,00%
Muški	49,07%	38,89%	9,26%	2,78%	0,00%
Svega	45,28%	40,25%	11,32%	3,14%	0,00%

- *Parametri koji se odnose na trenutno obrazovanje:* godina studiranja, prosečna ocena studiranja, procenat pohađanja nastave/vežbi; uspešnosti rešavanja sekvencijalnih zadataka; uspešnosti rešavanja kompleksnih zadataka, preferirani tip programiranja; nivo znanja za određene programske jezike; nivo programerske spremnosti (osposobljenosti) za rad u industriji, stil učenja po Kolbu. U tabelama 5-7 data je statistika o pojedinim parametrima koji se odnose na trenutno obrazovanje

Tabela 5: Statistika uspešnosti (po godinama studiranja)

Godina studiranja	Broj studenata	Pohađano više od 50% nastave/vežbi	Procenat položenih predmeta	Prosek programerskih predmeta	Ukupna prosečna ocena
Prva	63	94%	100%	8,35	7,86
Druga	34	98%	97%	8,14	7,87
Treća	32	88%	83%	8,36	7,96
Četvrta	30	84%	91%	7,71	7,67
Svega	159	92%	94%	8,11	7,84

Tabela 6: Spremnost za programiranjem (po godinama studiranja)

Godina studiranja	Ne znam	Samo uz nečiju pomoć	Samo jednostavne programe	Složen program na jednom PJ	Složen program na više PJ
Prva	1,5%	19,9%	23,9%	45,7%	9,1%
Druga	5,88%	20,59%	32,35%	26,47%	14,71%
Treća	0,00%	25,00%	21,88%	31,25%	21,88%
Četvrta	0,00%	20,00%	33,33%	33,33%	13,33%
Svega	1,89%	20,75%	27,04%	36,48%	13,84%

Tabela 7: Stilovi učenja po Kolbu (po polu)

Stil učenja	Ženski	Ženski %	Muški	Muški %	Svega	Svega %
Akomodator (aktivista)	14	27,5%	15	13,9%	29	18,2%
Asimilator (teoretičar)	11	21,6%	26	24,1%	37	23,3%
Diverger (refleksivan)	9	17,6%	19	17,6%	28	17,6%
Konverger (praktičar)	14	27,5%	30	27,8%	44	27,7%
Nediferenciran (neodređen)	3	5,9%	18	16,7%	21	13,2%

Izlazni parametar je uspešnost u programiranju. Ovaj parametar sadrži više mera i to: procenat položenih ispita, prosečna ocena iz predmeta koji se odnose na programiranje, rezultati testa kojim se meri uspešnost iz programiranja.

Test kojim se meri uspešnost u programiranju predstavlja instrument koji su razvili autori i koji uzima u obzir sve relevantne parametre iz oblasti programiranja na smeru IT. Kreiranje ovakvog instrumenta proističe iz decenijskog rada sa studentima u ovoj oblasti. Osim rezultata koje pruža pomenuti instrument uzeti su u obzir i rezultati koje studenti postižu u redovnom radu u okviru ispita.

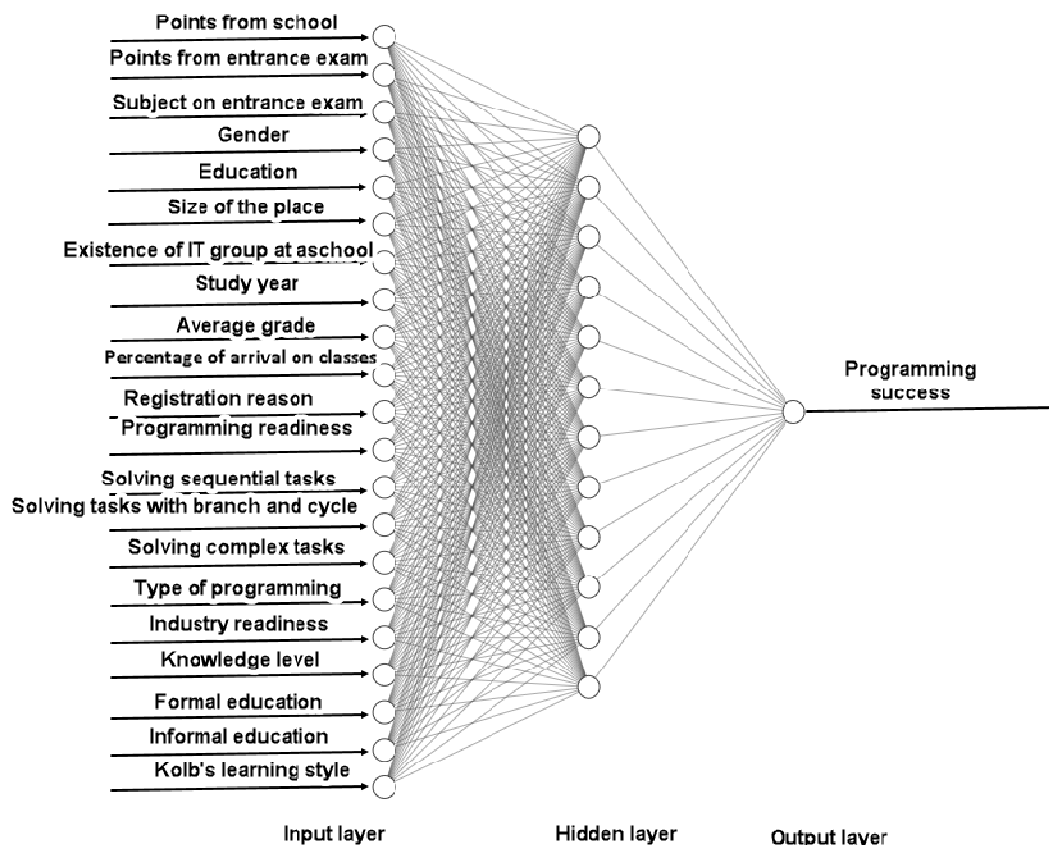
4. MODEL NEURONSKIH MREŽA

Na osnovu definisanih ulaznih i izlaznih parametara definisana je struktura neuronske mreže i prikazana je na slici 1.

Ulazni parametri su unapred definisani dok je izlaz iz mreže uspešnost studenata. Korišćen je jedan skriveni sloj. Predviđeni način treniranja mreže je backpropagation [5]. Tokom faze treniranja podaci za treniranje nalaze se u ulaznom sloju. U ovom algoritmu svaki čvor u skrivenom sloju dobija ulazne podatke iz svakog čvora u ulaznom sloju koji se množe sa odgovarajućim težinama i zatim sabiraju. Izlaz iz skrivenog čvora je nelinearna transformacija rezultujuće sume. Slično, svaki čvor u izlaznom sloju dobija ulaz iz svih čvorova iz skrivenog sloja koji se množe sa odgovarajućim težinama i zatim sabiraju. Izlaz iz izlaznog čvora je nelinearna transformacija rezultujuće sume.

Zatim se dobijene vrednosti na izlazu upoređuju sa ciljanim izlaznim vrednostima. Ciljane izlazne vrednosti su one vrednosti kojima se pokušava podučiti neuronska mreža.

Zatim se računa greška između dobijenih izlaznih vrednosti i ciljanih vrednosti i vraća unazad preko skrivenog sloja. Ovaj postupak se zove povratni postupak kroz skriven sloj. Greška se koristi da bi se korigovala snaga konekcije između čvorova, odnosno na osnovu nje se ažuriraju težine između ulaznog i skrivenog sloja, i skrivenog i izlaznog sloja.



Slika 1. Model neuronskih mreža

5. ZAKLJUČAK

Imajući u vidu predstavljen model može se zaključivati o polaznim pretpostavkama u cilju predviđanja uspešnosti u programiranju. Prema datom modelu i trenutno raspoloživim podacima o studentima zaključuje se da je model kompletan i da je uzeo u obzir sve relevantne podatke koji mogu imati uticaj na uspešnost studenata.

Uspešnost studenata je takođe kompletna mera jer uzima u obzir sve parametre koji mere uspešnost. Jedna od njih je i instrument koji su autori razvili a koji meri uspešnost u programiranju.

Rezultati predviđanja uspešnosti u programiranju su predmet budućeg istraživanja. Nakon implementacije predloženog modela utvrdiće se relevantnost predloženih ulaznih parametara kao i potencijalna optimizacija istih.

6. LITERATURA

- [1] Marian Bucos, Bogdan Drăgulescu, *Predicting student success using data generated in traditional educational generated in traditional educational*, TEM Journal. Volume 7, Issue 3, Pages 617-625, ISSN 2217-8309, DOI: 10.18421/TEM73-19, August 2018.
- [2] Zlatko J. Kovačić, *Predicting student success by mining enrolment data*, Research in Higher Education Journal, Vol. 15, pp 1-20, Mar 2012.
- [3] Fermin Ornelas, Carlos Ordonez, *Predicting student success: A Naïve Bayesian application to community college data*, Technology, Knowledge and Learning, October 2017, Volume 22, Issue 3, pp 299–315.
- [4] Denise Nadasen, Alexandra List, *Predicting four-year student success from two-year student data*, Big Data and Learning Analytics in Higher Education: Current Theory and Practice, Springer International Publishing, 2016, pp 221-236.
- [5] Casper Hansen, *Neural Networks: Feedforward and Backpropagation Explained & Optimization*, retrieved from: <https://mlfromscratch.com/neural-networks-explained/#/>, last access January 16 2020.