

Teo Ljubičić, mag. oec

Mlađi podatkovni analitičar
A1 Hrvatska d.o.o.
Email: tljubicic96@gmail.com
Orcid: <https://orcid.org/0000-0002-1543-7729>

Dr. sc. Marko Hell

Izvanredni profesor
Sveučilište u Splitu
Ekonomski fakultet
Email: marko.hell@efst.hr
Orcid: <https://orcid.org/0000-0003-1685-3561>

PREDIKCIJA USPJEŠNOSTI STUDENATA PRIMJENOM UMJETNIH NEURONSKIH MREŽA¹

UDK / UDC: 004.032.26:378

JEL klasifikacija / JEL classification: C45, I20, I23, O33

DOI: 10.17818/EMIP/2023/2.3

Prethodno priopćenje / Preliminary communication

Primljeno / Received: 30. rujna 2022. / September 30, 2022

Prihvaćeno / Accepted: 6. veljače 2023. / February 6, 2023

Sažetak

Višegodišnja elektronička obrada podataka omogućila je pohranu velike količine podataka koja se danas može iskoristiti za poboljšanje procesa u obrazovanju putem algoritama strojnog učenja. Koristeći se podacima iz sustava Moodle za učenje na daljinu, stvoren je model umjetne neuronske mreže s ciljem predviđanja konačnog ishoda studenta na kraju studija na temelju njegovih zaključnih ocjena prve godine studija. Kroz ukupno tri modela umjetne neuronske mreže pokazana je snaga ovog algoritma, gdje su svi modeli ostvarili vrlo nisku pogrešku, a najbolji rezultati postignuti su modelom umjetne neuronske mreže s dva skrivena sloja od devet neurona, čija je apsolutna greška iznosila 0,1920, a kvadratna pogreška 0,0562. Istraživanje pokazuje da su umjetne neuronske mreže vrlo učinkovite u predikciji konačnog ishoda studenata na temelju ocjena s prve godine studija te da takvi modeli imaju potencijal postati pomoćni alat i sredstvo donošenja odluka u obrazovnim institucijama.

Ključne riječi: strojno učenje, duboko učenje, umjetne neuronske mreže, prediktivno modeliranje, obrazovni sustav.

¹ Ovaj rad proizašao je iz diplomskog rada "Prediktivni model konačnog ishoda studenata primjenom umjetnih neuronskih mreža" koji je pod mentorstvom izv. prof. dr. sc. Marka Hella izradio Teo Ljubičić na Ekonomskom fakultetu Sveučilišta u Splitu

1. UVOD

Rad s podacima u obrazovanju spada u posebnu grana znanosti pod nazivom rudarenje edukacijskih podataka. Ovo se opisuje kao skup tehnika temeljenih na podatkovnoj znanosti, kao što su umjetna inteligencija, rudarenje podataka, upravljanje bazama podataka, strojno učenje i vizualizacija podataka, a cilj mu je izdvajanje korisnih informacija iz sirovih podataka za potrebe odlučivanja obrazovnih institucija (Okewu et al., 2021). Strojno učenje se može definirati kao proces rješavanja problema prikupljanjem skupa podataka i izgradnjom statističkog modela temeljenog na tom skupu (Burkov, 2019). Strojno učenje je razvijeno iz područja umjetne inteligencije (Kelleher, 2019), no razlikuje se po tome što mu nije cilj stvaranje imitacije inteligencije, već samo iskorištavanje računalne snage za obavljanje poslova koji nadilaze ljudske sposobnosti (Shalev-Schwartz i Ben-David, 2014). Među modelima strojnog učenja, posebno se ističu umjetne neuronske mreže zbog svoje velike računalne moći i prilagodljivosti hiperparametara u odnosu na dani problem. Iako umjetne neuronske mreže nisu novost, doživljavaju procvat zbog brzog razvoja računalne tehnologije, jačih grafičkih i procesorskih jedinica te razvoja tehnologije oblaka za pohranu podataka. Zbog toga, umjetne neuronske mreže su u stanju obraditi velike količine podataka i otkriti uzorke koji omogućuju precizne predikcije.

Rast popularnosti online pružatelja tečaja poput Coursera i Udemyja, kao i prelaskom na online nastavu kod nekih visokih škola poput Algebre koja nudi mnoge tečajeve i obrazovanje potpuno online, zajedno s brзом digitalizacijom obrazovanja uslijed pandemije COVID-19 koja je mnoge ustanove natjerala da usvajaju ove nove oblike nastave, najčešće putem Moodlea, omogućilo je dostupnost velikog broja neobrađenih podataka o studentima u obrazovanju. Analiziranjem ovih podataka, obrazovne ustanove bi mogle izvući nove korisne informacije prema kojima bi mogle prilagoditi svoje politike i bolje definirati uspješne parametre, smanjujući broj studenata koji odustaju i prilagođavajući program potrebama i mogućnostima studenata (Okewu, Adewole, Misra, Maskeliunas & Damasevicius, 2021). Uspjeh u prvim godinama studija može biti jedan od najjačih odrednica uspjeha na kraju studija. Vjerojatno je da se interveniranjem kod studenta s prosječno lošijim ocjenama na vrijeme može pomoći, ili se mogu otkriti predmeti s lošijom prolaznošću među studentima i eventualno zahtijevati promjena u njihovom programu i kurikulumu. Također, takvi podaci su lako dostupni obrazovnim ustanovama putem Moodle sustava kojim se gotovo svi koriste.

Cilj ovog rada je razviti optimalnu umjetnu neuronsku mrežu za predviđanje konačnog uspjeha studenata na temelju ostvarenih rezultata u prvoj godini studija. U sljedećem poglavlju objašnjava se proces pripreme podataka za razvoj modela, te osnovni pojmovi umjetnih neuronskih mreža. U trećem poglavlju prikazana je korištena metodologija, dok se četvrtom predstavlja izrada i razvoj konkretnog modela, kao i izbor optimalnog na temelju analize greške.

2. PREGLED LITERATURE

Strojno učenje je znanost programiranja računala koja koristi podatke da bi učila i izvukla korisne informacije (Geron, 2019). U kontekstu velikih podataka naziva se rudarenje podataka, posebno u obrazovanju - rudarenje edukacijskih podataka. Modeli strojnog učenja su prediktivni ili deskriptivni (Alpaydin, 2014). Prediktivni modeli koriste se povijesnim podatcima za predviđanje budućnosti, a deskriptivni modeli za izvlačenje informacija iz trenutnih podataka (Alpaydin, 2014). Prediktivni modeli koriste se nadziranim učenjem, a deskriptivni nenadziranim učenjem (Geron, 2019; Kacapor, 2020).

U posljednjim je godinama proveden velik broj istraživanja na području rudarenja obrazovnih podataka, analitike učenja i algoritama poput umjetnih neuronskih mreža u obrazovanju (Okewu et al., 2021). Predikcija uspjeha studenata popularno je područje u rudarenju obrazovnih podataka (Quinn, Gray, 2019) te se pokazala korisnom za identificiranje studenata s većim rizikom od napuštanja obrazovanja. Ukoliko se studentima pruži pomoć ranije, smanjit će se broj studenata koji zakažu u nastavi i napuštaju obrazovni program (Altaf, Soomro i Rawi, 2019). Quinn i Gray (2019) koristili su se podacima iz Moodlea kako bi predvidjeli uspjeh studenata, analizirajući njihove aktivnosti u nastavi. Isprobali su nekoliko algoritama strojnog učenja, kao što su *random forest*, algoritam najbližih susjeda i *gradient boosting*. Istraživanje je pokazalo da je *random forest* algoritam najbolji model za predikciju uspjeha studenata, s preciznošću od 60,5 % u stvaranju predikcija (Quinn, Gray, 2019). Autori također napominju da su modeli bili precizniji s većim skupom povijesnih podataka.

Shahiri, Husain i Rashid (2015) proveli su komparativnu analizu kako bi procijenili izvedbu modela stabla odlučivanja, naivnog Bayesa, K-najbližih susjeda, vektora podrške i modela umjetne neuronske mreže u predviđanju uspjeha učenika na temelju demografskih informacija i prethodnog tečaja ocjene. Rezultati su pokazali da je model umjetne neuronske mreže postigao najveću preciznost od 98 %, čime je pokazao svoju superiornost u odnosu prema drugim modelima. Umjetne neuronske mreže široko su prihvaćene (Shahiri et al., 2015) (Ersoz Kaya, 2019) zbog svoje snažne računalne snage i sposobnosti prilagodbe hiperparametara kako bi odgovarali specifičnim problemima, što im omogućuje da identificiraju obrasce u podacima bolje od drugih modela. Istraživanje pokazuje da su predikcije na temelju ocjena položenih kolegija ili ocjena ostvarenih na drugim razinama studija najčešći temelj stvaranja predikcija jer su ocjene jedan od glavnih predkazivača uspješnosti na studiju (Altaf i sur., 2019). Osim toga, umjetne neuronske mreže također se mogu koristiti u binarnoj klasifikaciji za određivanje rizičnog statusa učenika (Altaf et al., 2019), pružajući veću pouzdanost i preciznost modela (Quinn i Gray, 2019). Altaf i sur. (2019.) razvili su višeslojni model umjetne neuronske mreže koji se temelji na podacima o studentima na Moodleu kako bi predvidio uspjeh učenika u kolegijima i potrebu za dodatnom podrškom. Rezultati su pokazali da su umjetne neuronske mreže nadmašile ostale modele strojnog učenja u smislu preciznosti. Pavlin-Bernardić i sur. (2016) pokazali su potencijal umjetnih neuronskih mreža u identificiranju i klasifikaciji darovite djece

u hrvatskim osnovnim školama, koristeći se demografskim podacima i rezultatima testiranja kako bi izradili model s 95 % točnosti. S napretkom u strojnom i dubinskom učenju, posebno umjetnim neuronskim mrežama, obrazovne ustanove mogu kategorizirati studente i predvidjeti njihove akademske rezultate u ranoj fazi studija (Susnea, 2010) (Zacharis, 2016) (Jadić i dr., 2010), smanjujući time stope odustajanja, te povećavajući motivaciju (Brocardo, 2017.). Umjetne neuronske mreže također su se pokazale učinkovitima u oblikovanju nastavnog plana i programa obrazovnih ustanova. Naprimjer, Ersoz Kaya (2019) stvorio je model koji je analizirao sklonost predmetima na temelju uspjeha učenika i zaključio da alat može učinkovito mapirati predmete prema razinama znanja učenika.

3. METODOLOGIJA IZRADE MODELA

Ovdje prikazana studija koristila se programskim jezikom Pythonom, koji je trenutno najrašireniji programski jezik u području strojnog učenja zbog svojih iznimnih mogućnosti manipulacije podacima i automatizacije (McKinney, 2018.). Cijeli proces proveden je u razvojnom okruženju, poznatom kao *Jupyter Notebook*, koje nudi jednostavan i jednostavan način za izvršavanje koda i široko ga preferiraju znanstvenici i analitičari podataka. Modeliranje umjetnih neuronskih mreža provedeno je korištenjem biblioteke *TensorFlow* koja nudi specijalizirani skup naredbi posebno prilagođen za zadatke ove prirode.

Skup podataka na kojem će se raditi, uvezen je s Moodle sustava za učenje na primjeru fakulteta, a sadrži sve rezultate ispita koje je student prijavio, uključujući i one koje nije položio. Skup uključuje samo one studente koji su na razini prediplomskog studija, a promatrano razdoblje je od 2006. godine. S obzirom da je cilj istraživanja predvidjeti konačni prosjek studenata na temelju njihovih ocjena s prve godine studija, podaci vezani uz studente koji nisu završili studij nisu relevantni, pa nisu uvršteni u originalni skup. Također, nakon dodatnog čišćenja podataka, gdje su uklonjene vrijednosti s greškama ili prazne vrijednosti, originalni skup podataka sveden je na 119.561 redak koji uključuje ukupno 1924 studenta.

Stupci skupa podataka opisani su na sljedeći način:

- ID - identifikacijski broj studenta
- smjer - studijsko usmjerenje studenta (smjerovi su opisani kraticama a oni su sljedeći: POD, RIR, FM, MAR, IM, MAN)
- status - je li student redovan ili izvanredan student
- sifra_pred - šifra predmeta koji polaže student (npr. EUX000)
- datum_ispita - datum ispita koji polaže student
- ocjena_ispita - ocjena koju je student dobio na ispitu.

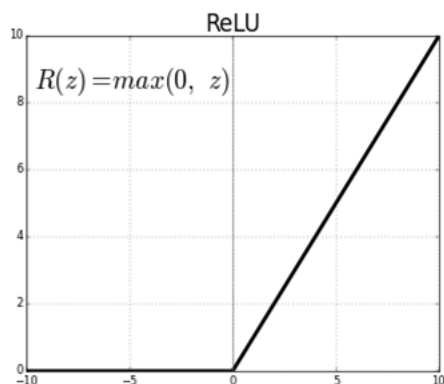
Zbog ograničenosti vrste podataka o svakom studentu koja se pohranjuje na Moodleu, skup podataka se sastoji samo od navedenih općih podataka. Međutim, originalni skup podataka neće ulaziti u sami model umjetnih neuronskih mreža, već je u sljedećim koracima preoblikovan da pristaje spomenutom

problemu i modelu. Zbog toga je napravljeno uklanjanje dvostrukih i nevažnih stavki, ispravljanje strukturalnih grešaka u riječima, uklanjanje ekstremnih vrijednosti (engl. *outlier*) i praznih stavki. Slijedi inženjering značajki na podacima, tj. izbor ulaznih varijabli koje najbolje opisuju razmatrani problem u svrhu uočavanja obrazaca i anomalija, te opisuju uz pomoć sažete statistike i grafičkih prikaza.

Baš kao što je ljudski mozak sačinjen od brojnih individualnih neurona, tako je i umjetna neuronska mreža sastavljena od brojnih umjetnih neurona, tzv. perceptrona. On se temelji na logičkoj jedinici za računanje praga tolerancije na podražaj i vrlo je sličan principu rada bioloških neurona prema kojoj crpi inspiraciju (Geron, 2019). U svojoj osnovi, perceptron klasificira vrijednosti u binarne izraze, odnosno u nule i jedinice, a određuje se matematičkom funkcijom koja definira granicu između te dvije klase.

Svaki umjetni neuron sastoji se od ulaznih vrijednosti koje mogu biti nasumično zadane u početku ili mogu biti određene izlaznim vrijednostima drugih neurona (Haykin, 2009). Ulazna vrijednost (slovo x) množi se s pripadajućim težinskim faktorom (slovo w) koji predstavlja neki broj i pripadajućom pristranosti (eng. Bias, slovo b). Ovisno o konačnoj vrijednosti u odnosu na granicu odluke (eng. *decision boundary*), varijabla se klasificira u jednu od razreda. Ako je izlazna vrijednost perceptrona različita od očekivane izlazne vrijednosti, odnosno ako je podatak krivo klasificiran, perceptron podešava svoje težinske faktore dok podaci nisu dobro klasificirani (Haykin, 2009). U višeslojnoj neuronskoj mreži vrijednost se šalje do sljedećeg neurona kao njegova ulazna vrijednost ukoliko je prag tolerancije zadovoljen (Kelleher, 2019). Na taj način moguće je povezati adekvatan broj umjetnih neurona, a lančanom aktivacijom perceptrona neuronska mreža zatim odabire optimalni ishod na temelju danih podataka. Odnosno, najveća konačna vrijednost izlaznih neurona određuje kojoj će klasi pripadati ulazna varijabla ukoliko je riječ o klasifikacijskom problemu. Ključan korak prilikom procesuiranja informacija neurona je obrada podataka aktivacijskom funkcijom. Svrha aktivacijske funkcije je odrediti hoće li vrijednosti unutar neurona biti prihvaćene u odnosu na granicu odluke (Sykora, Sinko, Vrskova, Kamencay & Hudec, 2019).

U ovom radu će biti korištena aktivacijska funkcija ReLU (*Rectified Linear Unit*) s obzirom da će izlaz neuronske mreže uvijek biti pozitivan broj (konačni prosjek studenta). Ukratko, ReLU je nelinearna aktivacijska funkcija, jednostavna za izračunavanje koja uvodi nelinearnosti u neuronsku mrežu i omogućuje mreži da nauči složenije prikaze podataka, izračunavanjem samo pozitivnog dijela ulaza (Maas, Hannun i Ng, 2013). Također, eliminacija negativnih vrijednosti bolje simulira biološki proces neuronske stanice (Sykora et al., 2019). Danas je ReLU funkcija gotovo potpuno zamijenila logističku funkciju zbog jednostavnosti, veće brzine i učinkovitosti (Haykin, 2009). Karakteristično za ovu funkciju (Slika 1) je što joj je interval između 0 i $+\infty$ gdje ne ograničava pozitivne kontinuirane vrijednosti, dok sve negativne vrijednosti automatski svrstava pod 0. Drugim riječima, sve negativne vrijednosti će automatski svrstati u "neaktivno", odnosno 0, dok će pozitivne vrijednosti ostati nepromijenjene.



Slika 1. ReLU aktivacijska funkcija

U gotovo svim nadziranima modelima strojnog učenja potrebno je odrediti cilj prema kojem će se odrediti uspješnost ili neuspješnost modela u predikciji ili klasifikaciji podataka. Kad je riječ o regresiji, model treba imati funkciju gubitka, gdje je cilj minimizirati tzv. funkciju troška. U ovom istraživanju kriterij za procjenu uspješnosti je srednja apsolutna pogreška (MAE) i srednja kvadratna pogreška (MSE), gdje je u oba slučaja cilj dobiti što manji iznos jer bi to značilo da je i odstupanje očekivanih vrijednosti od stvarnih mala. Srednja apsolutna pogreška dobra je prilikom interpretacije rezultata, no za razliku od srednje kvadratne pogreške nije toliko osjetljiva na promjene pa će se stoga u istraživanju koristiti obje funkcije.

Nakon što su definirani kriteriji procjene uspješnosti modela pristupit će se njihovoj optimizaciji. Iza ovog optimizacijskog procesa stoje dva algoritma: algoritam gradijentnog spusta i algoritam povratnog širenja pogreške. Cilj algoritma gradijentnog spusta je pronaći lokalni minimum svakoj funkciji i time minimizirati pogrešku modela pri stvaranju predikcija postepenim podešavanjem parametara (Geron, 2019), dok algoritam povratnog širenja pogreške na temelju danih izlaznih vrijednosti računa nagibe funkcija počevši od izlazne vrijednosti natrag prema ulaznoj (Geron, 2019). Na temelju toga, algoritam gradijentnog spusta ponovno računa nove, optimizirane parametre, a cijeli proces se ponavlja određeni broj ponavljanja (tzv. epohe) dok se funkcija troška ne minimizira. U ovom istraživanju korišten je optimizacijski algoritam "Adam". Naime, ovaj algoritam se preporuča za većinu modela umjetnih neuronskih mreža jer je jednostavan za implementaciju, ima brže vrijeme izvođenja, male zahtjeve za memorijom i zahtijeva manje podešavanja nego bilo koji drugi algoritam optimizacije (Goodfellow, Bengio & Courville, 2016).

4. STVARANJE MODELA I ANALIZA REZULTATA

4.1. Inženjering značajki

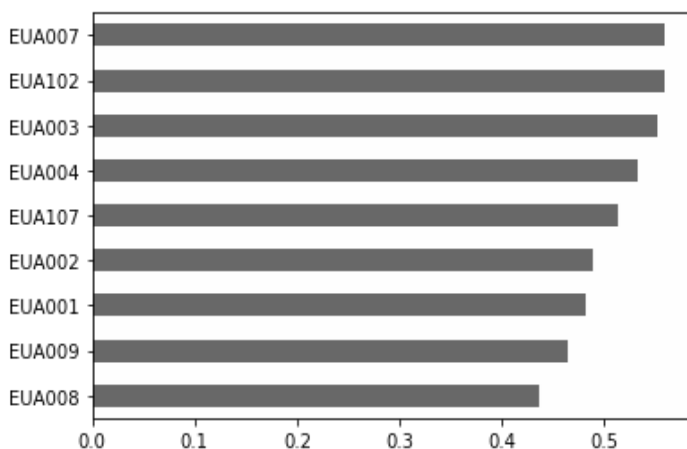
U slučaju ovog rada značajke, odnosno varijable stvorene su agregacijom i kombinacijom više drugih značajki. Pritom je najviše koristilo prijašnje iskustvo autora u stvaranju već sličnih modela, ali i poznavanje problema te samih podataka s kojima se radi.

Uz pomoću biblioteke Pandas napravljena je tablica u kojoj su studenti, odnosno identifikacijski brojevi studenata u redovima, a nazivi predmeta u stupcima. Vrijednosti unutar tablice su agregirani prosjeci ocjena za pojedine predmete. Za ovaj model odabrano je devet glavnih predmeta s prve godine studija, koje su morali položiti svi studenti na fakultetu kako bi mogli napredovati na studiju. Zbog standardizacije predmeta za sve studente su isključeni izborni predmeti kako bi se obuhvatilo najveći broj studenata. Također, pretpostavka je da su to predmeti koji studentima predstavljaju najveći izazov i da ti predmeti mogu biti glavne odrednice uspješnosti studenata na prvim godinama studija. Nova tablica stoga sadrži podatke raspoređene u deset stupaca (devet šifri predmeta i prosječna ocjena studiranja), te 1924 retka (studenta). Prvih devet stupaca predstavlja ulazne varijable modela strojnog učenja. Zadnji stupac koji predstavlja konačni prosjek na kraju treće godine studija je izlazna varijabla modela, tj. ciljana varijabla koju bi model trebao predvidjeti. Konačni cilj je uz pomoć ocjena s prve godine studija predvidjeti konačni prosjek na kraju studija.

4.2. Analiza podataka

Konačni skup podataka sadrži informacije o studentima, njihovom uspjehu na predmetima s prve godine studija i konačnom uspjehu na kraju studija. Predmeti su imenovani prema službenim oznakama predmeta jer su sva imena kolegija, studenata i ustanove zaštićeni od objavljivanja. Stoga se umjesto punih naziva koriste njihove kratice.

Cilj ove analize podataka je uputiti na određene obrasce u ponašanju studenata tijekom ispita koji mogu sugerirati bolji uspjeh tijekom studija. Uz klasičnu deskriptivnu analizu podataka, proučavanje odnosa između varijabli napravljeno je uz pomoć toplinske karte koja jasno prikazuje korelacije među vrijednostima: svjetlija boja ukazuje na jaku korelaciju, a tamnija boja ukazuje na slabu korelaciju. U prosjeku svi predmeti srednje jako do jako pozitivno koreliraju s konačnim prosjekom. Kvadrati u samoj sredini toplinske karte uvijek pokazuju maksimalnu korelaciju jer svaka vrijednost ima savršenu korelaciju sa samom sobom.



Slika 2. Jačina korelacija predmeta s konačnim prosjekom

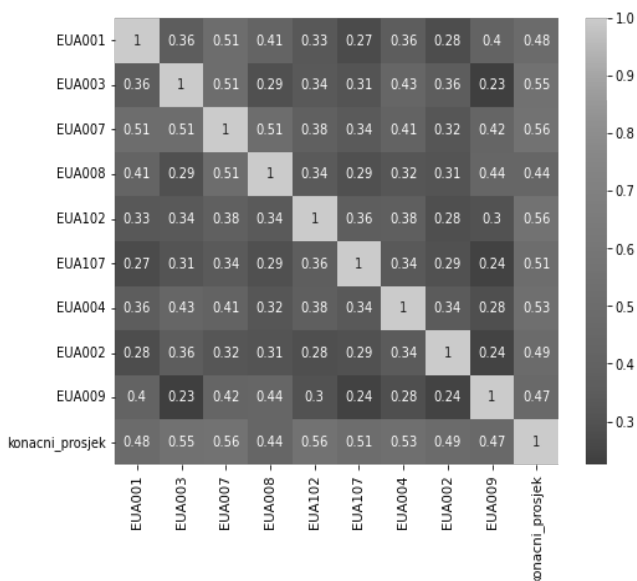
Koeficijent korelacije između konačnog prosjeka i predmeta “EUA102” iznosi 0,56 i ukazuje na jaku pozitivnu korelaciju prema Pearsonovoj jačini korelacije između te dvije vrijednosti², što znači da će povećanje ocjene iz predmeta “EUA102” znatno utjecati na povećanje konačnog prosjeka, što može uputiti na činjenicu da je to jedan od najizazovnijih predmeta za položiti, stoga visoka zaključna ocjena iz tog predmeta može uputiti na bolji ukupni prosjek.

Osim korelacije s konačnim prosjekom, mogu se uočiti korelacije između predmeta, pa tako između predmeta “EUA003” i “EUA007” korelacija iznosi 0,51, što upućuje na jaku pozitivnu korelaciju prema Pearsonovoj skali korelacije čime se može zaključiti da pozitivna ocjena iz jednog predmeta značajno utječe na pozitivnu ocjenu iz drugog predmeta. To može uputiti na činjenicu da su ti predmeti vrlo slični po izvedbenom planu ili vrsti, pa npr. oba predmeta mogu biti matematičko orijentirani predmeti.

Ukoliko se pokaže da postoji vrlo visoka korelacija (iznad 0,7³) određenih predmeta međusobno ili s konačnim ishodom, može se javiti problem multikolinearnosti te bi se morao potpuno ukloniti taj predmet iz modela. Multikolinearnost je pojava visokih međukorelacija između dvije ili više nezavisnih varijabli i u ekstremnim slučajevima može značiti da se varijable mogu opisati kao linearna kombinacija drugih varijabli (James, Witten, Hastie & Tibshirani, 2021). Međutim, u ovom slučaju korelacije nisu previsoke i sve vrijednosti se prihvaćaju.

² <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>

³ <https://datascienceplus.com/multicollinearity-in-r/>



Slika 3. Toplinska karta korelacije predmeta

4.3. Razvoj modela umjetne neuronske mreže

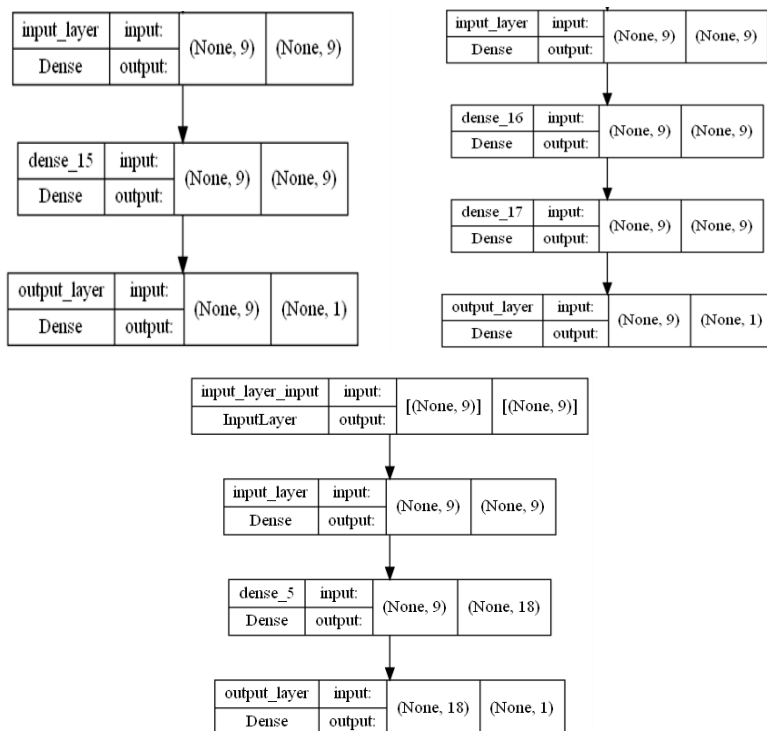
Arhitektura neuronskih mreža rađena je prema Geron (2019), gdje ulazni sloj treba sadržavati onoliko neurona koliko postoji značajki u ulaznim podacima (devet u ovom slučaju), skriveni sloj sadrži proizvoljno mnogo neurona, a izlazni sloj sadrži jedan neuron prema dimenziji (jedan mogući rezultat u ovom slučaju). Sve ulazne varijable promatrane su kroz varijablu X , dok je izlazna varijabla prikazana s Y . Ukupni skup podataka podijeljen je na tri dijela (jedan za treniranje, jedan za testiranje i jedan za validaciju), i to prema omjeru 60/20/20, gdje se 60 % podatka koristi kao skup za treniranje, 20 % kao skup za testiranje i ostalih 20 % kao skup za validaciju..

Cijeli skup podataka podijeljen je u sljedećem omjeru:

- X_{train} (1154 redova i 9 stupaca)
- X_{test} (384 red i 9 stupaca)
- X_{valid} (386 redova i 9 stupaca)
- y_{train} (1154 redova)
- y_{test} (384 red)
- y_{valid} (386 redova).

U svrhu izgradnje modela umjetne neuronske mreže potrebno je dodati slojeve i neurone u neuronsku mrežu, te odrediti funkcija troška i aktivacijsku funkciju čija je svrha opisana u prethodnom koraku. Svaki model će sadržavati ulazni sloj od devet neurona i izlazni sloj od samo jednog neurona. U prvom modelu

koristit će se višeslojna neuronska mreža s jednim skrivenim slojem koji se sastoji od devet neurona. U drugom modelu koristit će se višeslojna mreža s dva skrivena sloja od devet neurona. U trećem modelu će se koristiti višeslojna mreža s dva skrivena sloja od 9 i 18 neurona. Na slici 4. korištenjem prethodno navedenih softvera dan je grafički prikaz izgrađenih neuronskih mreža sa svojim hiperparametrima.

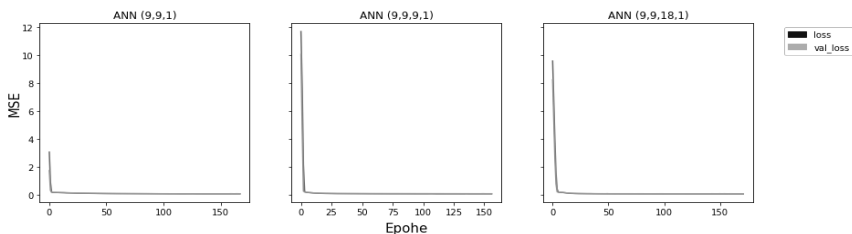


Slika 4. Umjetna neuronska mreža jednim skrivenim slojem, s dva skrivena sloja po 9 neurona i dva skrivena sloja po 9 i 18 neurona

Ulazni slojevi označeni su kao “*Input Layer*”, a izlazni slojevi “*Output Layer*”. Svi slojevi su grafički prikazani pravokutnicima, gdje strelica upućuje na smjer kretanja informacija iz jednog sloja u drugi. Također, prikazane su dvije brojke 9 po svakom pravokutniku jer lijevi broj 9 upućuje na ulaz koji taj sloj dobiva, a desni broj 9 upućuje na izlaz tog sloja. Na sljedećem grafu prikazane su funkcije troška za svaki od prethodno tri razvijena modela umjetnih neuronskih mreža.

Na grafovima slike 5. nalaze se po dvije krivulje, jedna tamnije boje koja predstavlja grešku modela prilikom testiranja i jedna svjetlije boje koja prikazuje grešku modela tijekom validacije. Cilj je da su te dvije krivulje tijekom učenja što

bliže jedna drugoj, što bi značilo da model ne odstupa mnogo od stvarnih vrijednosti. S obzirom da svjetlija krivulja potpuno preklapa tamniju krivulju u ovom slučaju znači da model funkcionira vrlo dobro i da ne postoji problem prenaučnosti.



Slika 5. Krivulje troška tri razvijana modela umjetnih neuronskih mreža

Kako bi se dobili najbolji rezultati, ukupno su napravljena tri modela umjetnih neuronskih mreža s različitim hiperparametrima. Prema tablici 1 vidimo da je model umjetnih neuronskih mreža s jednim skrivenim slojem od devet neurona, ostvario MAE od 0,213912 i MSE od 0,069775. Najbolji je rezultat ostvarila neuronska mreža s dva skrivena slojeva po devet neurona gdje je MAE bio 0,192083 i MSE bio 0,056280, odnosno da je taj model u prosjeku griješio za samo 0,192083. Najkompleksniji model s tri skrivena sloja od po 9, 9, i 18 neurona polučio je lošiji rezultat nego prethodni.

Tablica 1.

Usporedni prikaz greške (MAE i MSE) razvijenih modela

Modeli	MAE	MSE
ANN (9,9,1)	0,213912	0,069775
ANN (9,9,9,1)	0,192083	0,056280
ANN (9,9,18,1)	0,192563	0,058507

To ukazuje da povećavanje kompleksnosti modela kroz povećanje broja slojeva i neurona unutar slojeva ne pridonosi nužno poboljšanju točnosti modela, već suprotno, može dovesti do problema prenaučnosti modela i lošijih performansi. Kao što je vidljivo u ovom primjeru, neuronska mreža s dva skrivena sloja, ali s manjim brojem neurona, ostvarila je bolje rezultate, nego mreža s dva skrivena sloja i većim brojem neurona.

5. ZAKLJUČAK

Svrha ovoga rada bila je ispitati mogućnosti i potencijal umjetnih neuronskih mreža u obrazovanju. Provedeno istraživanje pokazalo je da umjetne neuronske mreže s lakoćom rješavaju problem kao što je predikcija uspješnosti

studiranja na temelju ocjena s prve godine studija. To znači da model uvrštavanjem ocjena prve godine nekog studenta može predvidjeti koliki će biti njegov prosjek na kraju studija. Ovakva mogućnost predikcije, osim studentima, može poslužiti i upravama fakulteta u svrhu upravljanja nastavnim procesom. Unatoč uspješnosti modela u stvaranu predikcije, treba uputiti na ograničenje rada u pogledu ograničenosti samog seta podataka. Može se zaključiti da je do konačnog ishoda vrlo lako doći na temelju ocjena prve godine studija te bi bilo vrlo korisno za buduće istraživanje povećati kompleksnost podataka povećanjem promatranog razdoblja studija (do kraja pete godine studija) ili uključivanjem druge vrste podataka koja dobro nadopunjuje istraživanje.

Daljnja istraživanja usmjerena su prema nadogradnji modela, uključivanjem drugih značajki (stupaca), kao što su: demografski podaci studenata, podaci upisnih rezultata ili neka druga svojstva koje će omogućiti bolju sliku očekivanih učinaka obrazovnog procesa. Primjerice, koristeći se demografskim podacima i obrazovnom pozadinom studenata, na upisu je moguće predvidjeti uspjeh studenata na odabranom smjeru ili prilagoditi program studija ovisno o toj obrazovnoj pozadini.

Činjenica je da se mnoge grane društva danas koriste umjetnim neuronskim mrežama i podacima kako bi poboljšali svoju učinkovitost u poslovanju, stoga ne postoji razlog da i obrazovne institucije ne poduzmu jednak korak.

LITERATURA

Altaf, S., Soomro, W., Rawi, M. I. M. (2019), "Student Performance Prediction using Multi- Layers Artificial Neural Networks: A Case Study on Educational Data Mining", 3rd International Conference on Information System and Data Mining, 59-64. <https://dl.acm.org/doi/10.1145/3325917.3325919>

Burkov, A. (2019), "The Hundred-Page Machine Learning Book". <https://themlbook.com/>

Brocardo, M. L., Traore, I. (2017), "Authorship verification using deep belief network systems", International Journal of Communication Systems 30(12). <https://www.doi.org/10.1002/dac.3259>

Ersoz Kaya, I. (2019), "Artificial Neural Networks as Decision Support Tool in Curriculum Development", International Journal on Artificial Intelligence Tools Vol. 28, No. 4. <https://doi.org/10.1142/S0218213019400049>

Geron, A. (2019), "Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow", O'Reilly

Goodfellow, I., Bengio, Y., Courville, A., (2016), "Deep Learning", The MIT Press

Haykin, S. (2009), "Neural Networks and Learning Machines", Pearson Education Inc.

Jadrić, M., Garača, Ž., Ćukušić, M. (2010), "Student dropout analysis with application of data mining methods", Management, Vol. 15, 2010, 1, pp. 31-46. <https://hrcak.srce.hr/53605>

James, G., Witten, D., Hastie, T., Tibshirani, R. (2021), "An Introduction to Statistical Learning with Applications in R". Springer New York. <https://doi.org/10.1007/978-1-0716-1418-1>

Kačapor, K., Lagumdžija, Z. (2020). "Rudarenje edukacijskih podataka: korištenje klasteriranja za predikciju studentskog uspjeha". 43rd International Convention on Information,

Communication and Electronic Technology pp. 1075 - 1080 https://www.researchgate.net/publication/351871935_Rudarenje_educacijskih_podataka_koristenje_klasteriranja_za_predikciju_studentskog_uspjeha

Kelleher, D. J. (2019), "Deep Learning", The MIT Press. <https://doi.org/10.7551/mitpress/11171.001.0001>

Ljubičić, T. (2022) "Diplomski rad: Prediktivni model konačnog ishoda studenata primjenom umjetnih neuronskih mreža"

McKinney, W. (2018), "Python For Data Analysis: Data Wrangling With Pandas, Numpy And IPython", O'Reilly

Maas, A.L., Hannun, A.Y. and Ng, A.Y. (2013), "Rectifier Nonlinearities Improve Neural Network Acoustic Models". Proceedings of the 30th International Conference on Machine Learning, Atlanta, 16-21 June 2013, 3. https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf

Okewu, E., Adewole, P., Misra, S., Maskeliunas, R., Damasevicius, R. (2021), "Artificial Neural Networks for Educational Data Mining in Higher Education: A Systematic Literature Review", Applied Artificial Intelligence 2021, vol. 35, no. 13, 983-1021. <https://doi.org/10.1080/08839514.2021.1922847>

Pavlin-Bernardić, N., Ravić, S., Matić, I. P. (2016), "The application of artificial neural networks in predicting children's giftedness", *Suvremena psihologija* 19 (2016), 1, 49-59. <https://doi.org/10.21465/2016-SP-191-04>

Quinn, R. J., Geraldine G., "Prediction of student academic performance using Moodle data from a Further Education setting", *Irish Journal of Technology Enhanced Learning* (2019) <https://api.semanticscholar.org/CorpusID:209070178>; <https://doi.org/10.22554/ijtel.v5i1.57>

Rastrollo-Guerrero, J. L., Gómez-Pulido, J. A., Durán-Domínguez, A. (2020), "Analyzing and Predicting Students' Performance by Means of Machine Learning: A Review", *Appl. Sci.* 2020, 10(3), 1042. <https://doi.org/10.3390/app10031042>

Susnea, E. (2010), "Using Artificial Neural Networks In E-Learning Systems", *UPB Scientific Bulletin, Series C: Electrical Engineering*, 91-100. <https://www.researchgate.net/publication/258832960>

Shalev-Schwartz, S., Ben-David, S. (2014), "Understanding Machine Learning: From Theory to Algorithms", Cambridge University Press. <https://doi.org/10.1017/CBO9781107298019>

Shahiri, A., Husain, W., Abdul Rashid, N.. (2015). "A Review on Predicting Student's Performance Using Data Mining Techniques". *Procedia Computer Science*. 72. 414-422. <https://doi.org/10.1016/j.procs.2015.12.157>

Sykora, P., Sinko, M., Vrskova, R., Kamencay, P., Hudec, R. (2019), "Artificial Neural Networks in Educational Process", 17th International Conference on Emerging eLearning Technologies and Applications (ICETA), 751-756. <https://www.doi.org/10.1109/ICETA48886.2019.9040142>

Zacharis, Z., N. (2016), "Predicting Student Academic Performance In Blended Learning Using Artificial Neural Networks", *International Journal of Artificial Intelligence and Applications (IJAIA)*, Vol. 7, No. 5, 17-29. <https://www.doi.org/10.5121/ijaia.2016.7502>

Teo Ljubičić, MEcon

Junior Data Analyst
A1 Hrvatska d.o.o.
Email: tljubicic96@gmail.com
Orcid: <https://orcid.org/0000-0002-1543-7729>

Marko Hell, PhD

Associate Professor
University of Split
The Faculty of Economics, Business and Tourism
Email: marko.hell@efst.hr
Orcid: <https://orcid.org/0000-0003-1685-3561>

STUDENT SUCCESS PREDICTION USING ARTIFICIAL NEURAL NETWORKS

Abstract

Many years of electronic data processing have enabled the storage of a large amount of data that can be used today to improve educational processes through machine learning algorithms. Using data from the Moodle distance learning system, an artificial neural network model was created to predict the final outcome of students at the end of their studies based on their final grades of the first year of study. In three artificial neural network models, the power of this algorithm was demonstrated, where all models achieved a very low error, and the artificial neural network model achieved the best results with two hidden layers of nine neurons, whose absolute error was 0.1920, and the squared error 0.0562. The research shows that artificial neural networks are very effective in predicting the final outcome of students based on the grade from the first year of study and that such models have the potential to become an auxiliary tool and means of decision-making in educational institutions.

Keywords: machine learning, deep learning, artificial neural networks, predictive modelling, education system

JEL classification: C45, I20, I23, O33