PREDIKTIVNI MODEL KONAČNOG ISHODA STUDENATA PRIMJENOM UMJETNIH NEURONSKIH MREŽA

Ljubičić, Teo

SVEUČILIŠTE U SPLIT
EKONOMSKI FAKULTET

DIPLOMSKI RAD

PREDIKTIVNI MODEL KONAČNOG ISHODA STUDENATA PRIMJENOM UMJETNIH NEURONSKIH MREŽA

Mentor:
izv.prof.dr.sc. Marko Hell

Student:
univ.bacc.oec. Teo Ljubičić

Split, rujan, 2022.

SADRŽAJ:

1. UVOD

Današnji čovjek živi u dobu nagle digitalizacije i automatizacije okoline, a tehnologija i umjetna inteligencija postaju dio njegove svakodnevnice i ključan faktor pri donošenju odluka. Područje umjetne inteligencije usko je povezano uz pojam umjetnih neuronskih mreža, algoritam matematičkih operacija koji je u mogućnosti analizirati ogromne količine podataka i uočiti obrasce koji su čovjeku nevidljivi. Pomoću njih moguće je npr. otkriti je li osoba pri većem riziku od kožnih bolesti ili opasnosti od potencijalne prevare prije nego se to dogodi. Unatoč sve široj primjeni ove tehnologije, obrazovanje je grana društva kojom još uvijek pretežito dominiraju ljudi kao faktori donošenja odluka i vođenja procesa. Cilj ovoga rada je istražiti mogućnosti umjetnih neuronskih mreža u kontekstu obrazovanja kao potencijalno sredstvo pružanja podrške pri odlučivanju ili poboljšanju trenutnih procesa.

1.1. Problem istraživanja

Strojno učenje je potpodručje računalnih znanosti koje se bavi izgradnjom algoritama koji se, da bi bili korisni, oslanjaju na zbirku primjera nekog fenomena. Pri fenomenu se može misliti na nešto iz prirode, stvoreno od strane čovjeka ili od drugog algoritma (Burkov, 2019). Najčešće se pri tome misli na nekakav set podataka prema kojima stroj traži uzorke u podacima, odnosno na kojima stroj uči, a zatim na temelju matematičkih formula stvara modele prema kojima može predvidjeti ishode na novom setu podataka koje stroj nikada nije vidio. Strojno učenje se može podijeliti na površno učenje (eng. Shallow learning) i duboko učenje (eng. Deep learning) (Burkov, 2019). U površno učenje se mogu svrstati poznati modeli kao što su linearna regresija, logistička regresija, stabla odlučivanja, dok se u duboko učenje svrstaju oni modeli koji nisu površno učenje[1].

Kada se govori o dubokom učenju, zapravo se govori o umjetnim neuronskim mrežama (eng. Artificial neural networks), modelu koji svoju inspiraciju i naziv uzima od bioloških neuronskih

[1] https://quantdare.com/what-is-the-difference-between-deep-learning-and-machine-learning/

mreža u mozgu (Geron, 2019). Kao što se mozak sastoji od milijuna neurona, međusobno povezanih dendritima i aksonima, tako se model umjetnih neuronskih mreža sastoji od više slojeva umjetnih neurona, međusobno povezani, koji šalju informacije od jednog neurona prema drugome ukoliko postoji određeni prag tolerancije na podražaj (Geron, 2019). Zbog toga, umjetne neuronske mreže su u mogućnosti izvući mnogo informacija iz podataka te otkriti nove odnose i uzorke u podacima koje „površni" modeli nekada ne mogu. Ovaj način rada ih odvaja od drugih modela strojnog učenja i zbog toga se često koriste kod procesa optimizacije, klasifikacije, predviđanja te otkrivanja uzoraka u slikama i video snimkama (Deperlioglu i sur., 2011).

Zbog svoje svestranosti umjetne neuronske mreže koriste se u mnogim sektorima i društvenim grana poput financijskog sektora, sektora medicine, marketinga, obrazovanja, društvenih mreža i drugih[2]. U financijama se često koriste kod predviđanja kretanja cijena dionica ili pak klasifikacije osobe prema tome hoće li uspješno moći otplatiti kredit. U medicini svoju primjenu nalaze kod prepoznavanja uzoraka u slikama i na temelju toga mogu prepoznati maligne tragove na koži, a društvene mreže ga koriste za prilagođavanje sadržaja individualnom korisniku mreže prema onome što taj korisnik preferira vidjeti.[3]

Razvojem informacijskih tehnologija u posljednjih nekoliko godina, a naročito posljedicom utjecaja pandemija COVID-19, velika količina obrazovnog procesa prebacila se na mrežni oblik rada putem raznih sustava za učenje kao što je Moodle. Takav oblik rada zahtijeva pohranu velike količine informacija, a kvalitetnom obradom i korištenjem podataka obrazovne institucije mogu doći do vrijednih informacija koji mogu pomoći u boljem donošenju odluka na temelju tih podataka (Susnea, 2010). Jedan od načina na koji se podaci mogu iskoristiti je kod predviđanja prolaznosti studenta određenog predmeta ili kolegija. Biti u mogućnosti predvidjeti učinak studenta omogućuje obrazovnoj ustanovi da na vrijeme pomognu studentima koji su pri većem riziku od padanja i na taj način mogu smanjiti ukupni broj studenata koji odustanu od obrazovnog programa.

[2] https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/
[3] https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

U posljednjih godina veliki je broj istraživanja provedeno na temu umjetnih neuronskih mreža u obrazovanju. Altaf i sur. (2019) su pomoću višeslojnih umjetnih neuronskih mreža, na temelju podataka studenata s Moodle-a, uspješno stvorili model koji može predvidjeti performanse studenata na predmetima i hoće li određenim studentima trebati dodatna pomoć svladati te predmete. U radu su usporedili rezultate umjetnih neuronskih mreža s drugim modelima strojnog učenja i zaključili da „neuronske mreže, na temelju preciznosti, premašuju ostale klasifikatore."

Pavlin-Bernardić i sur. (2016) su ukazali na potencijal umjetnih neuronskih mreža u identificiranju i klasificiranju nadarene djece u hrvatskim osnovnim školama. Koristeći se demografskim podacima djece i roditelja te rezultatima raznih testova uspješno su stvorili model koji može kategorizirati nadarenu djecu s preciznošću od 95%. Time su ukazali da takav alat može pomoći učiteljima kod donošenja odluka po pitanju nadarene djece pogotovo u školama gdje ne postoji adekvatan broj psihologa i sličnih stručnjaka.

Na temelju navedenih rezultata istraživanja može se doći do zaključka da postoji veliki interes za temu strojnog učenja, dubokog učenja i umjetnih neuronskih mreža u obrazovanju, a naročito se primjeti rast interesa u posljednjih dvadeset godina naglim razvojem računalnih tehnologija.

1.2. Predmet istraživanja

Ovaj rad pobliže će proučiti umjetne neuronske mreže u obrazovnom sektoru, istraživanja koja su se provela na tom području i mogućnosti umjetnih neuronskih mreža u obrazovanju. Danas obrazovne institucije pohranjuju ogromne količine informacija o studentima, njihovim upisnim rezultatima, performansama tijekom studija i o konačnim rezultatima. Koristeći se tehnikama strojnog i dubokog učenja, točnije umjetnim neuronskim mrežama, obrazovne ustanove su u mogućnosti klasificirati studente i na temelju toga predvidjeti njihove konačne ishode već na početku studija (Susnea, 2010). Na temelju toga obrazovne ustanove mogu donijeti odluke koje im mogu pomoći da bolje zadrže studente, spriječe prijevremeno opadanje, donesu odluke koje mogu pomoći boljem usvajanju gradiva i podizanju razine motivacije studenata (Brocardo, 2017).

Prema iznesenim tvrdnjama, istraživački dio rada nastoji ukazati na mogućnosti umjetnih neuronskih mreža u predikciji konačnog prosjeka studenata. Pritom će se prikazati analiza podataka kojom se traže uzorci u podacima, postupak stvaranja input varijabli, postupak modeliranja i analiza konačnih rezultata.

1.3. Cilj istraživanja

Cilj istraživanja je analizirati mogućnosti primjene umjetnih neuronskih mreža u predikciji konačnog ishoda studenata na temelju njihovih ocjena s prve godine studija. Usporednom analizom modela umjetnih neuronskih mreža i modela linearne regresije nastojati će se ukazati na prednosi i moć neuronskih mreža koji ih ističe od drugih modela strojnog učenja. Također, istraživanjem se želi ukazati na potencijal umjetnih neuronskih mreža u obrazovanju.

1.4. Istraživačka pitanja

Prema navedenim problemom i područjem istraživanja postavljaju se istraživačka pitanja na koje rad nastoji odgovoriti.

- Jesu li umjetne neuronske mreže prikladni modeli za predviđanje konačnog ishoda studiranja studenta na kraju treće godine studija na osnovi ocjena s prve godine studija?
- Koje su umjetne neuronske mreže prikladni kao model predikcije konačnog ishoda studija?
- Jesu li umjetne neuronske mreže bolji alat u predikciji konačnog ishoda studija od modela linearne regresije?
- Koje su to karakteristike umjetnih neuronskih mreža koje ih čine boljim prediktivnim alatom od linearne regresije?
- Mogu li modeli umjetnih neuronskih mreža pomoći kod donošenja odluka obrazovnih institucija?

1.5. Metode istraživanja

Nakon predmeta, problema i ciljeva istraživanja važno je navesti metode istraživanja kojima će se obuhvatiti teorijska podloga rada skupa s istraživačkim dijelom rada. Prilikom definiranja metoda istraživanja korišteni su radovi Zelenike (2000) i Tkalac Verčić (2014):

- **Deduktivna metoda** – kod deduktivnog načina se iz općih stavova izvode posebni, pojedinačni zaključci. Ova metoda koristiti će se tijekom cijelog rada a njome će se izvući specifični zaključci o umjetnim neuronskim mrežama i njihovoj ulozi u obrazovanju
- **Metoda analize** – ova metoda biti će najviše korištena kod objašnjenja modela strojnog učenja i njihovih specifičnosti
- **Metoda deskripcije** – ova metoda koristiti će se kako bi se detaljno opisale sve činjenice, pojave ili predmeti navedeni u radu
- **Metoda kompilacije** – tijekom pisanja rada korišteni su različiti izvori, a ponajviše radovi drugih autora. To mogu biti knjige, članci, znanstveni radovi ili sadržaj s internet stranica. Cilj je koristeći se različitim izvorima srodnih tema, stvoriti novi i jedinstveni zaključak
- **Statistička metoda** – ova metoda ponajviše će se koristiti u istraživačkom dijelu rada gdje će biti potrebno analizirati rezultate prediktivnih modela, a pritom će biti korištena statistička metodologija i statistički pojmovi
- **Komparativna metoda** – u istraživačkom dijelu rada uspoređivati će se rezultati modela strojnog učenja prema nekoliko parametara, na temelju toga donijeti će se i zaključak o uspješnosti modela
- **Metoda vizualizacije** – u cijelom radu koristiti će se razni grafovi i tablice zbog boljeg prenošenja informacije čitatelju, a i prema kojima će se jasnije moći donijeti zaključak istraživanja

U istraživačkom dijelu rada će se koristiti programski jezik Python koji je danas najpopularniji programski jezik u sferi strojnog učenja zbog njegove superiorne mogućnosti manipulacije podacima i automatizacije poslova. Cijeli proces će se obaviti u razvojnom okruženju (eng. IDE – Integrated Development Environment) zvanom Jupyter Notebook koji omogućava jednostavno i pregledno izvršavanje linija koda i kojeg preferiraju mnogi podatkovni znanstvenici i analitičari.

Uz Python ujedno dolazi i niz znanstvenih „biblioteka“ (eng. libraries) koji su zapravo unaprijed isprogramirane funkcije koje korisnik treba samo pozvati jednom linijom koda da bi se izvršili. Nek od najpoznatijih su Pandas, namijenjena manipulaciji i analizi podataka[4], NumPy, dizajniranja za stvaranje višedimenzionalnih nizova i matrica za matematičke funkcije na kojima se modeli strojnog učenja oslanjaju[5], te Scikit-learn, biblioteka namijenjen strojnom učenju[6].

Samo modeliranje umjetnih neuronskih mreža obaviti će se bibliotekom „TensorFlow“. TensorFlow je besplatna softverska knjižnica otvorenog koda za strojno učenje i umjetnu inteligenciju. Može se koristiti u nizu zadataka, ali je posebno usmjeren na trening dubokih neuronskih mreža.[7] Radi se o specijaliziranim setom naredbi dizajnirani upravo za zadatke ovakvog tipa.

1.6. Doprinos istraživanja

Istraživanje ovog rada doprinijeti će tako da će prikazati proces modeliranja umjetnih neuronskih mreža u programskom jeziku Python, koristeći se svim potrebnim bibliotekama za strojno učenje. Cijeli model raditi će na podacima studenata i na temelju toga će prikazati mogućnosti predviđanja njihovog konačnog ishoda studija. Na taj način rad može doprinijeti budućim istraživanjima na temu strojnog učenja i umjetnih neuronskih mreža na bazi obrazovnih ustanova u Hrvatskoj gdje se još uvijek nije proveo prevelik broj istraživanja na ovu temu, a naročito na fakultetima gdje se često provode kolegiji koji uključuju statistiku i statističku analizu, ali o umjetnim neuronskim mrežama kao prediktivnim modelima se zapravo uopće ne govori. Također, uspoređujući modele linearne regresije i umjetnih neuronskih mreža rad će nastojati ukazati na prednosti koje ovakvi alati dubokog učenja nose sa sobom.

[4] https://pandas.pydata.org/
[5] https://numpy.org/
[6] https://scikit-learn.org/stable/
[7] https://www.tensorflow.org/

1.7. Struktura diplomskog rada

Prvi dio rada sastoji se od uvoda, definiranja problema i predmeta istraživanja, sastavljanja istraživačkih pitanja prema kojima će se definirati istraživački cilj. Zatim će se opisati što se istraživanjem želi postići, a na kraju i doprinijeti.

Drugi dio rada govoriti će o tome o povijesti stvaranja i razvoja strojnog učenja, što je točno strojno učenje, vrste učenja koji postoje u strojnom učenju, opisati će se problem regresije koji je vrlo čest problem u strojnom učenju i opisati će se model linearne regresije koji će se poslije koristiti u praktičnom radu.

Treći dio ulazi u teorijsku podlogu o umjetnim neuronskim mrežama, njihovoj arhitekturi i usporedbi bioloških i umjetnih neurona, a pritom će se opisati osnovni model višeslojnih mreža bez povratnih veza (eng. Multilayer Feedforward Networks). Poglavlje zatim govori o načinu na koji neuronske mreže uče i procesi koji su pritom uključeni, a na samom kraju govoriti će se o umjetnih neuronskim mrežama u obrazovanju te njihovim koristima i mogućnostima.

U četvrtom dijelu rada će se opisati podaci na kojima će se raditi istraživanje, alati u kojima će se podaci obraditi, programski jezik koji će biti korišten, razvojno okruženje u kojima će se pisati kod i koraci prema stvaranju modela kojima će se predvidjeti konačni prosjek studenata. Zatim će se prikazati cijeli navedeni proces te opisati pojedini koraci. Koristeći se vizualizacijskim alatima stvoriti će se razni grafovi i tablice kako bi se preciznije objasnili ti procesi. Također će se usporediti performanse dvaju modela strojnog učenja prema definiranim statističkim parametrima.

U posljednjem dijelu rada će se interpretirati dobiveni rezultati istraživanja kao i činjenice i spoznaje do kojih se došlo tijekom istraživanja.

2. STROJNO UČENJE

Nagli razvoj informacijskih tehnologija dovela je do preokreta u načinu na koji shvaćamo i gledamo na podatke. Nekoć su podatke prikupljali samo velika poduzeća koja su jedina imala mogućnosti pohranjivati i iskorištavati takve podatke, no danas svaka osoba proizvodi podatke na neki način, bilo to slanjem e-poruka, slikanjem pomoću mobilne kamere ili stvarajući novu objavu na društvenoj mreži. Osim što ljudi stvaraju podatke, oni ih i istovremeno koriste kroz razne prilagođene i personalizirane sadržaje koji im mogu pomoći u stvaranju odluka (Alpaydin, 2014).

U slučaju trgovine robe, koristeći se podacima prethodnih transakcija, poduzeće želi pomoću tih podataka utvrditi koji su proizvodi najpopularniji među kupcima, kada je prodaja najveća te koji se proizvodi često kupuju zajedno. Dok s druge strane kupac želi, svaki put kad kupi nešto, da mu trgovina predloži predmete koje bi se njemu možda svidjele prema njegovim preferencijama i željama u odnosu na prethodnu kupnju. Iako je nemoguće točno predvidjeti što će koja osoba kupiti, prema podacima o prošlim transakcijama možemo uočiti određene uzorke u ponašanju kupaca. Neki proizvodi poput gaziranih pića se možda često kupuju sa slanim čipsom, a njihova popularnost može biti izraženija dok događanja kao što su velika nogometna prvenstva.

Ukoliko postoji određeni uzorak u ponašanju kupaca i ukoliko pretpostavimo da budućnost neće biti mnogo različita od sadašnjosti, može se i pretpostaviti da će se isti uzorak nastaviti ponavljati. Koristeći te uzorke mogu se napraviti i predikcije. Međutim, za to je potreban nekakav algoritam koji će biti u mogućnosti primiti ulazne, sadašnje vrijednosti i pretvoriti ih u izlazne, odnosno buduće vrijednosti. U primjeru trgovine, ulazni podaci mogu biti povijest transakcija individualnih kupaca, a izlazni podaci mogu biti prediktivnog karaktera, poput onoga što će taj kupac sljedeći put kupiti ili deskriptivnog karaktera, gdje se izvlače nove korisne informacije prema karakteristikama kupovina tog kupca (Alpaydin, 2014).

Definirati ulazne i izlazne vrijednosti nije dovoljno da stroj bude u mogućnosti napraviti predikcije, već ono mora iz tih podataka nešto naučiti kako bi mogao pritom sam odrediti najbolji ishod za specifični problem. Na temelju podataka stroj uči i uočava uzorke u podacima, prema kojima radi pretpostavke, a na temelju njih i predikcije (Alpaydin, 2014). Snaga i prednosti strojnog učenja je upravo u otkrivanju tih uzoraka u velikoj gomili neuređenih podataka.

Strojno učenje nalazu upotrebu u brojnim granama znanosti i područjima društva. Primjeri algoritama strojnog učenja mogu se pronaći gotovo svugdje danas, a prosječna osoba dolazi u interakciju svakodnevno s njima. Algoritmi strojnog učenja često se koriste kod[8]:

1. Otkrivanja prijevara – korišteno često u financijskom sektoru gdje algoritam uočava odstupanja u uzorku ponašanja korisnika i prijavljuje to odstupanje kao mogućnost prijevare
2. Virtualnih asistenata – osim što su dostupno na pametnim telefonima, danas postoje razni komercijalni alati koji koriste umjetnu inteligenciju raspoznavanja govora poput popularnog Amazon Echo-a ili Iphone Siri-a
3. Video nadzora – trenutno prepoznavanje lica više subjekata odjednom omogućava bolji sigurnosni nadzor na mjestima poput aerodroma ili gradskih centara
4. Otkrivanja spam mailova – obradom jezika i prepoznavanjem specifičnih riječi u emailu algoritam je u mogućnosti prepoznati neželjeni mail i automatski ga klasificirati kao spam
5. Online službe za korisnike – popularni online pružatelji usluga često imaju stotine tisuća, ako ne i milijuna korisnika, a da bi se svim korisnicima moglo ugoditi potrebni su algoritmi koji služe kao služba za korisnike, poznati kao chatbotovi, koji su u mogućnosti odgovoriti korisnicima na upite i riješiti njihove probleme
6. Preporuka proizvoda – skupljajući informacije o prethodnim transakcijama, pregledanim filmovima, preslužanom glazbom, algoritam je u mogućnosti predložiti novi sadržaj prema ukusu korisnika

Burkow (2019) definira strojno učenje kao proces rješavanja problema prikupljanjem skupa podataka i algoritamskom izgradnjom statističkog modela temeljenog na tom skupu podataka. Ova definicija već uzima u obzir da strojno učenje nije zasebna disciplina već kombinacija više disciplina kao što su matematika, statistika, računalne znanosti i umjetna inteligencija. Doduše, strojno učenje se razlikuje od umjetne inteligencije jer joj cilj nije stvoriti imitaciju inteligencije

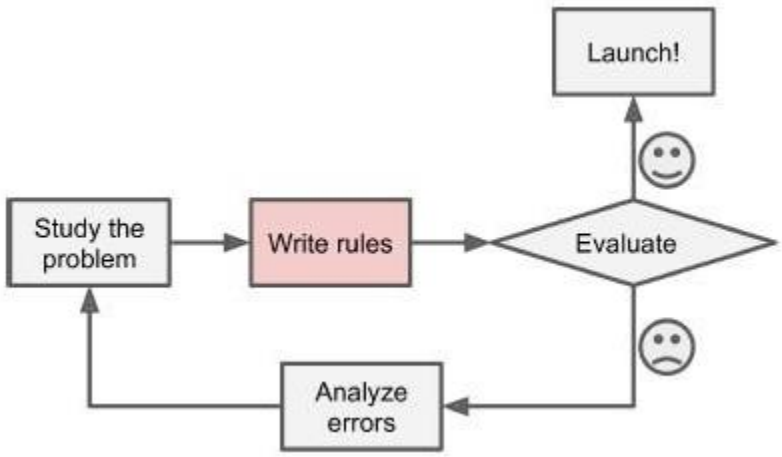
[8] https://medium.com/app-affairs/9-applications-of-machine-learning-from-day-to-day-life-112a47a429d0

već samo iskoristiti snagu računala da bi obavili posao koji seže van ljudskih mogućnosti (Shalev-Schwartz i Ben-David, 2014).

Prema Shalev-Schwarz i Ben-David (2019) dvije situacije zahtijevaju upotrebu strojnog učenja naspram upotrebe klasičnog programa:

1. **Kada je problem previše kompleksan da bi ga se isprogramirao** – odnosi se često na rutinske ljudske aktivnosti poput vožnje ili raspoznavanja lica gdje je potrebno da stroj uči iz „iskustva“ ili na zadatke koji zahtijevaju analizu ogromnih količina podataka
2. **Zadaci koji zahtijevaju prilagodljivost** – dok su klasični programi često fiksni jednom kada ih se napiše, algoritmi strojnog učenja u mogućnosti su prilagoditi svoje izlazne vrijednosti ukoliko dođe do promjene u ulaznim varijablama

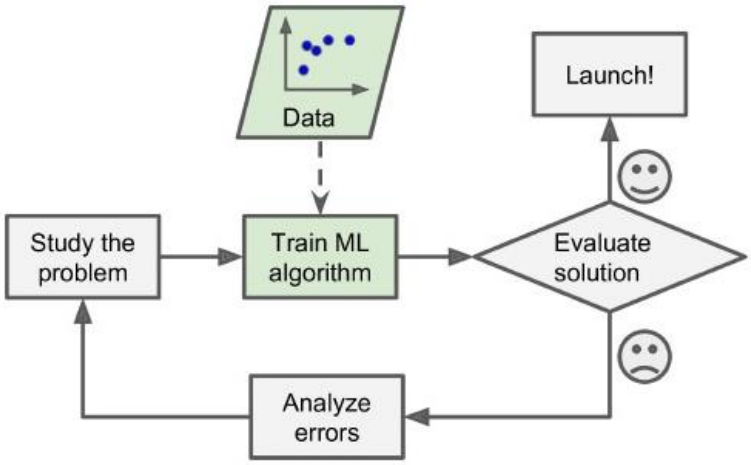
Slike 1 i 2 jasno prikazuju odnos rješavanja problema tradicionalnim programom i programom baziran na strojnom učenju. Prvi slučaj bi zahtijevao iznimno ogromne količine koda i ljudskog rada gdje je ručno potrebno napraviti izmjene svaki put kad se program suoči s novim problemom.



Slika 1: Shema rješavanja problema tradicionalnim programom

Izvor: Geron, A. (2019), Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly, str 5.

U drugom slučaju stroj crpi znanje na temelju seta podataka, gdje na temelju tih podataka stroj uči, rješava problem i zatim mjeri rezultat. Ukoliko nije ispunio očekivanja, stroj ponovno uči na temelju novog znanja i tako u krug dok ne dođe do krajnjeg rješenja.



Slika 2: Shema rješavanja problema upotrebom strojnog učenja

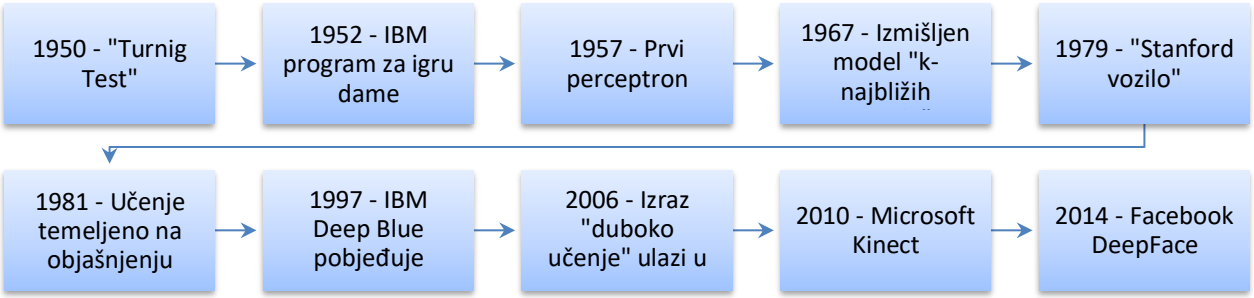
Izvor: Geron, A. (2019), Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly, str 6.

2.1. Povijesni razvoj strojnog učenja

Iako strojno učenje ne spada u potpunosti u područje umjetne inteligencije, a umjetna inteligenciju ne čini samo strojno učenje, za razumijevanje razvoja strojnog učenja potrebno je gledati razvoj umjetne inteligencije od čega je ono nastalo kao vlastita jedinka.

Ljudsko zanimanje za umjetnu inteligenciju sežu sve do daleke povijesti u antičku Grčku kada su drevni filozofi vodili val razmišljanja može li se ljudski um staviti u mehaničko tijelo. U grčkoj mitologiji pojavljuje se lik Talosa, velika brončana čovjekolika figura koju je sagradio grčki bog Hefest, bog kovač i izumitelj, kojem je zatim Zeus podario život[9].

[9] https://news.stanford.edu/2019/02/28/ancient-myths-reveal-early-fantasies-artificial-life/



Slika 3: Povijesni razvoj strojnog učenja

Izvor: Izrada autora prema https://www.forbes.com/sites/bernardmarr/2016/02/19/a-short-history-of-machine-learning-every-manager-should-read/?sh=1db4058815e7

Međutim, pravi, moderni počeci umjetne inteligencije započinju polovicom 20. stoljeća kada Alan Turing objavljuje članak pod nazivom „Computing Machinery and Intelligence“ kada je stvorio tzv. Turing test. U njemu je autor istraživao mogu li strojevi razmišljati, a učinio je to testom gdje stroj i jedna osoba moraju oboje uvjeriti drugu osobu (koja ne zna tko je stroj, a tko stvarna osoba) da su oni zapravo osoba. Druga osoba zatim treba procijeniti tko je od njih osoba, a tko stroj. Ako ne uspije točno pogoditi, stroj pobjeđuje[10].

Dvije godine kasnije Arthur Samuel stvorio je prvi računalni program koji je mogao učiti. Program je bila igra dame, a IBM-ovo računalo je napredovalo u igri što je više igralo, proučavajući koji potezi čine pobjedničke strategije i ugrađujući te poteze u svoj program.[11]

Sljedeći značajni korak u razvoju strojnog učenja je izum prvih neuralnih mreža 1957. godine kada je američki psiholog Frank Rosenblatt dizajnirao perceptron, jednostavan umjetni neuron koji čini dio neuronske mreže. Ovaj izum vrlo je relevantan i danas, a više o perceptronu i umjetnim neuronskim mrežama bavit će se sljedeće poglavlje.

[10] https://en.wikipedia.org/wiki/Computing_Machinery_and_Intelligence
[11] https://www.lightsondata.com/the-history-of-machine-learning/

Unatoč tome što je postojao dugi period gdje je razvoj strojnog učenja usporio zbog manjka interesa investitora i nedostatka kapitala, više od dvadeset godina kasnije, 1979. grupa studenta na Stanfordu izumila je vozilo koje je u mogućnosti kretati se samostalno kroz prostoriju i pritom zaobilaziti sve prepreke, dok je Gerald Dejong 1981. predstavio koncept učenja temeljenog na objašnjenju, gdje računalo analizira podatke o treningu i stvara opće pravilo koje može slijediti odbacujući nevažne podatke. Upravo se moderni modeli strojnog učenja oslanjaju na koncept trening podataka na kojima stroj uči i nalazi uzorke.[12]

U 1990-ima rad na strojnom učenju prešao je s pristupa vođenog znanjem na pristup vođen podacima. Znanstvenici su počeli stvarati računalne programe za analizu velikih količina podataka i izvlačenje zaključaka - ili "učenje" iz rezultata. A već 1997[13]. IBM stvara program koji je uspješno pobijedio tadašnjeg svjetskog prvaka u šahu.

Izraz „duboko učenje“ izmislio je Geoffrey Hinton 2006.[14] godine kako bi objasnio algoritme kojima računala prepoznavaju uzorke u podacima poput slika, videa i zvukova. Time počima revolucija u strojnom učenju i masovno se počinju stvarati proizvodi poput Microsoft Kinect-a koji je u mogućnosti pratiti kretnje i aktivnosti korisnika u stvarnom vremenu koji na taj način komunicira s računalom odnosno igricama na računalu.

Danas sva velika poduzeća poput Amazon-a, Facebook-a ili Microsoft-a koriste strojno učenje i umjetnu inteligenciju za obavljanje zadataka poput analize teksta, filtriranje sadržaja, prepoznavanja govora, slika i videa u svakodnevnoj praksi. Poduzeća poput Google-a čak su objavila besplatne komercijalne alate poput TensorFlow-a koji omogućava korisnicima stvaranje vlastitih modela umjetnih neuronskih mreža. Tehnološkom evolucijom svjedočimo razvoju autonomnih automobila i letjelica, inovacija u medicini koja spašavaju živote mnogih ljudi, ali i razvoju kućanskih uređaja poput robotskih čistača, koji svi koriste algoritme dubokog učenja da bi funkcionirali.

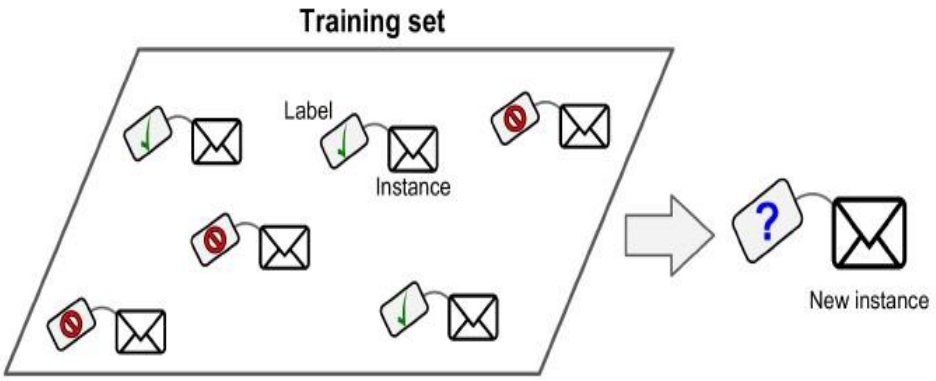
[12] Ibid.
[13] Ibid
[14] https://analyticsindiamag.com/the-history-of-machine-learning-algorithms/

2.2. Vrste učenja

Postoji više načina na koji se mogu kategorizirati sustavi strojnog učenja, no jedan od najvažnijih kategorija opisuje odnos modela strojnog učenja i podataka. Ovisno o vrsti problema, ali i algoritmu koriste se različite vrste učenja. Jedan od načina na koji se sustavi strojnog učenja mogu klasificirati je prema količini i vrsti nadzora koji dobivaju tijekom obuke. Postoje četiri glavne kategorije učenja: nadzirano učenje, nenadzirano učenje, polunadzirano učenje i podržano učenje.

2.2.1. Nadzirano učenje

Budući da učenje uključuje interakciju između učenika (stroja) i okoline (podataka), zadatke učenja mogu se podijeliti prema interakciji navedena dva subjekta, a za objašnjenje načina rada nadziranog učenja uzeti će se primjer detekcije spam mailova. U ovom slučaju set podataka dijeli se na set podataka za trening i set podataka za testiranje. Podaci za trening koji se unose u model uključuju sve varijable i pripadajuća željena rješenja na temelju kojih, kao što naziv govori, model "trenira" odnosno uči i prepoznaje uzorke u podacima. Svi mailovi unutar podataka za trening unaprijed su već definirani kao spam ili kao normalni mailovi. Temeljem tog "treninga" stroj treba osmisliti pravilo za označavanje novih mail poruka, kojeg zatim isprobava koristeći se novim podacima iz seta podataka za testiranje. U takvim slučajevima možemo razmišljati o okolini kao učitelju koji "nadzire" učenika pružajući potrebne informacije (Shalev-Schwartz i Ben-David, 2014).



Slika 4: Shematski prikaz nadziranog učenja

Izvor: Geron, A. (2019), Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly, str 8.

Još jedan tipičan zadatak za nadzirano učenje je predviđanje ciljne numeričke vrijednosti, kao što je cijena automobila, s obzirom na skup značajki (kilometraža, starost, marka itd.) koji se nazivaju prediktori. Ovakva vrste zadataka naziva se regresija, o kojoj će bit više riječi u poglavlju o linearnoj regresiji. Da biste obučili model, potrebno mu je dati mnogo primjera automobila, uključujući njihove prediktore i tražene oznake (npr. cijena).

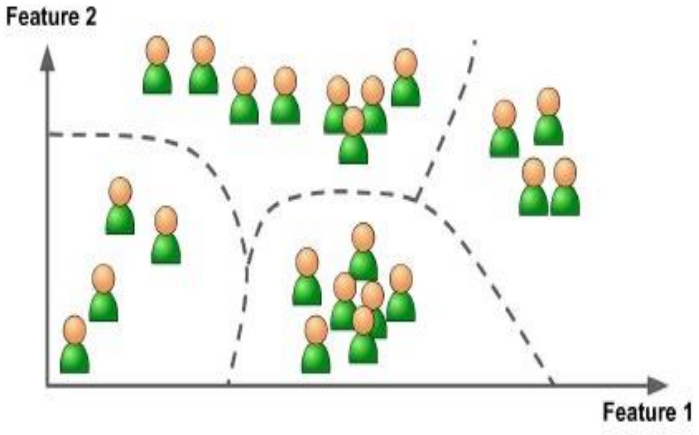
Neki od najvažnijih i najpopularnijih modela nadziranog učenja su:

- Linearna regresija
- Logistička regresija
- K-najbliži susjedi
- Stabla odluke
- Neuronske mreže

2.2.2. Nenadzirano učenje

U nenadziranom učenju, međutim, nema razlike između podataka o treningu i testiranju, odnosno, za cijeli proces koristi se jedan skup podataka. Model obrađuje ulazne podatke s ciljem da dođe do nekog sažetka ili komprimirane verzije tih podataka. Klasteriranje podataka u podskupove sličnih objekata tipičan je primjer takvog zadatak. Ovakav oblik učenja koristan je kada se želi

izvući korisne i do tad nevidljive informacije. Primjerice kod trgovine odjećom, određeni lanac može analizom podataka o kupcima otkriti uzorke u ponašanju. Pa tako mladi ljudi mogu naginjati kupnji ležernog tipa odjeće dok stariji kupci naginju udobnoj odjeći. Ili možda kupci koji kupuju crne hlače skloni su kupnji bijele košulje, a manje skloni kupnji cipela za sport.



Slika 5: Nenadzirano učenje

Izvor: Geron, A. (2019), Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly, str 10.

Još jedan važan zadatak nenadziranog učenja je otkrivanje anomalija - na primjer, otkrivanje neuobičajene transakcije kreditnim karticama za sprječavanje prijevare ili otkrivanje grešaka u proizvodnji (Geron, 2019).

Neki od najvažnijih modela nenadziranog učenja su:

- K-means klasteriranje
- Analiza glavnih komponenti (eng. Principal Component Analysis)
- Stroj s potpornim vektorima (eng. Support Vector Machine)

2.2.3. Polu-nadzirano učenje

U polu-nadziranom učenju, skup podataka sadrži i označene i neoznačene primjere, dakle to je kombinacija nadziranog i nenadziranog učenja.. Obično je količina neoznačenih primjeraka

mnogo veća od broja označenih primjeri. Burkov (2019) ističe kako je cilj polu-nadziranog algoritma učenja isti kao i cilj algoritma nadziranog učenja, a korištenje neoznačenih primjera može pomoći algoritmu učenja da pronađe bolji model. Naime, Burkov (2019) dalje tvrdi da kada se dodaju neoznačeni primjeri u model, dodaju se i nove informacije, unatoč tome što takvi primjeri povećavaju nesigurnost modela.

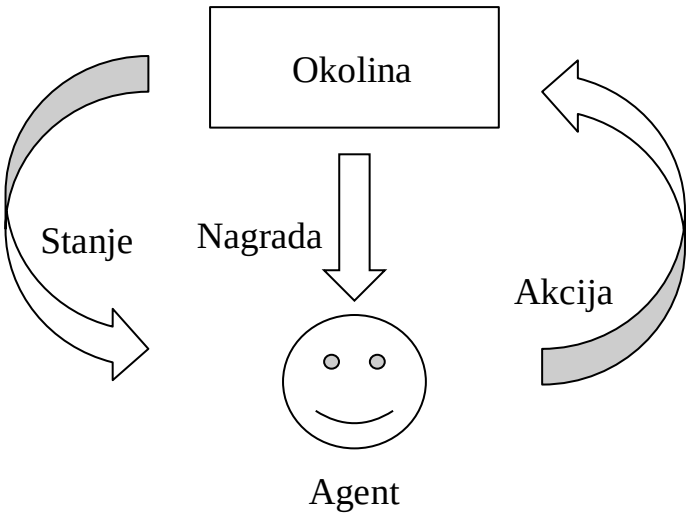
Jedan primjer polu-nadziranog učenja je Googleov servis za slike, tzv. Google Photos. Kada korisnik učita sve svoje obiteljske fotografije na uslugu, servis automatski prepoznaje da se ista osoba A pojavljuje na jednoj skupini fotografija, dok se druga osoba B pojavljuje u drugoj skupini fotografija. Ovo je nenadzirani dio algoritma (klasteriranje). Slijedi da korisnik unese informacije o navedenim osobama na samo nekoliko fotografija i algoritam je u mogućnosti imenovati svakoga na svakoj sljedećoj fotografiji.

2.2.4. Podržano učenje

Podržano učenje je posebna grana strojnog učenja koja uči na podacima generiranom u stvarnom vremenu i pritom uvijek postoji nekoliko jedinstvenih točaka. Prva je stroj koji se naziva *agent* koji se nalazi u *okolini*. Stroj eksperimentira i radi nasumične, ali ograničene, *akcije* te za njih biva nagrađen ili kažnjen ovisno o ishodu. Bolji ishod rezultira većom *nagradom*, što potiče agenta da pamti i ponavlja radnje koje su rezultirale pozitivnih ishodom. Svakom akcijom određeno stanje u okolini se mijenja pri čemu se stroj ponovno treba prilagoditi i učiti. (Alpaydin, 2021). S vremenom stroj više ne proizvodi nasumične radnje, već na temelju naučenog formira svoju strategiju. Stroj mora sam naučiti koja je najbolja strategija, tzv. politika, kako bi s vremenom dobio najveću nagradu. Za razliku od nadziranog učenja gdje se okolina opisuje kao "učitelj", kod podržanog učenja okolina se može smatrati "kritičarom" jer ona nije ta koja uči stroj, već je samo opisuje što se loše radi, a što dobro (Alpaydin, 2021).

Pri donošenju uzastopnih odluka s dugoročnim ciljem, kao u igrama, robotici, upravljanju resursima ili logistici, za pronalaženje rješenja koristi se podržano učenje (Geron, 2019). Jedna od modernih i sve više prisutnih primjena podržanog učenja je kod autonomnih automobila koja su u mogućnosti posve samostalno voziti na cesti bez potrebe intervencije vozača. U ovakvom slučaju

stroj, odnosno algoritam, uvijek crpi nove informacije iz okoline prema kojima korigira svoje pokrete i donosi nove odluke.



Slika 6: Podržano učenje

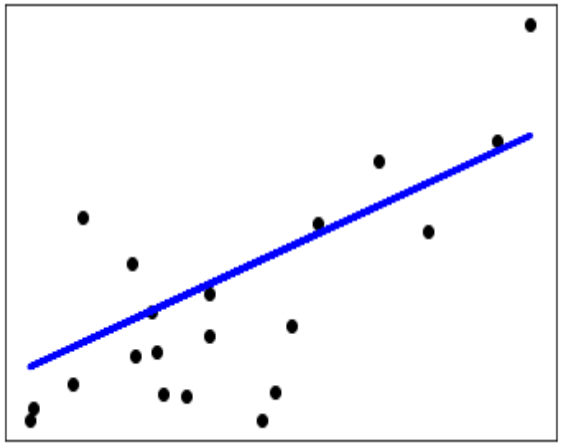
Izvor: Izrada autora prema Alpaydin, E. (2021), Strojno učenje, MATE d.o.o., str. 126.

2.3. Vrste zadataka kod strojnog učenja

Različiti algoritmi strojnog učenja prilagođeni su i namijenjeni različitim vrstama zadataka. Postoji više zadataka koje strojno učenje nastoji riješiti, ali najpoznatiji su regresijski i klasifikacijski problem, kratko opisani u prethodnom poglavlju.

Regresija

Regresija je problem predviđanja oznake stvarne vrijednosti iz neoznačenog uzorka, a cilj regresijskog istraživanja je numerička vrijednost. Dobro poznata upotreba regresije je procjena vrijednosti kuće na temelju atributa kao što su veličina, broj spavaćih soba, lokacija i drugih čimbenika. Najpoznatiji regresijski algoritam je model linearne regresije, koja će detaljnije biti opisana u nastavku.

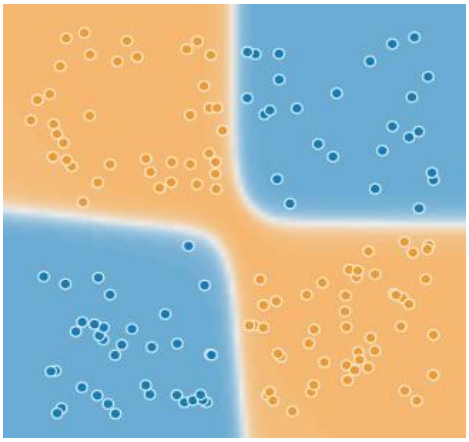


Slika 7: Primjer regresije

Izvor: Izrada autora

Klasifikacija

Problem klasifikacije je kvalitativnog formata te njen cilj nije numerička vrijednosti već određena kategorija, razred ili skupina ciljne varijable. Bračni status osobe, robna marka kupljenog proizvoda, kasni li osoba s plaćanjem duga ili ne ili vrsta raka koji je dijagnosticiran primjeri su kvalitativnih varijabli (James, 2021). Značajke u problemu klasifikacije pripadaju jednoj od ograničenog broja klasa. Binarna klasifikacija se koristi kada postoje samo dvije klasifikacije u skupu. Problem klasifikacije s tri ili više klasa naziva se višeklasna klasifikacija (Bukov, 2019).

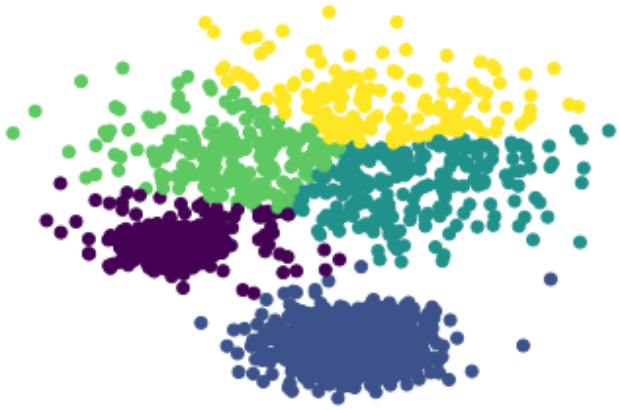


Slika 8: Primjer klasifikacije

Izvor: Izrada autora

Klasteriranje

Klasteriranje je zadatak identificiranja sličnosti između vrijednosti te dodjeljivanje tih vrijednosti njihovim pripadajućim skupinama na temelju sličnosti. Grupe tih vrijednosti nazivaju se klasteri. Klasteriranje se može koristiti kao tehnika istraživačke analize podataka gdje je cilj pronaći grupe koje se prirodno pojavljuju u podacima (Alpaydin, 2021). Klasteriranje je često u upotrebi u poduzećima koja vode evidencije o svojim klijentima prema kojima prilagođavaju sadržaje ili prema kojima segmentiraju svoje klijente kako bi otkrili odbjegle korisnike. U poduzećima je to poznato kao upravljanje odnosa s klijentima (eng. Customer Relationship Management),



Slika 9: Primjer klasteriranja

Izvor: https://www.ml-science.com/k-means-clustering

Predviđanje vremenskih serija

Kada postoji potreba za predviđanjem vrijednosti na temelju podataka vremenske serije, to se naziva problem predviđanja vremenske serije. Vremenska serija je niz numeričkih točaka podataka uzastopnim redoslijedom. Podaci vremenske serije znače da su podaci u nizu određenih vremenskih razdoblja ili intervala. Jedan primjer predviđanja vremenskih serija je nastojanje da se predvidi kretanje dionica na tržištu, algoritam koji je često korišten u svijetu trgovine dionicama. Međutim, predviđanje ovakvih vrijednosti vrlo je težak i često neprecizan zadatak (Marsland, 2009).

Preporuke

Algoritmi strojnog učenja koriste se za stvaranje sustava preporuke (eng. Recommendation Systems) koji funkcioniraju na principu stvaranja generalizacija, odnosno ako osoba A voli film X, a film X sadrži određene čimbenike kao i neki drugi filmovi ili ga je pogledala slična skupina ljudi kao osoba A, tada će prema tim čimbenicima ta osoba sigurno voljeti film Y. Na temelju tih čimbenika stroj treba biti sposoban generalizirati, a cijeli proces se odvija tzv. dekompozicijom matrica (Alpaydin, 2021). Postoje dva pristupa korištena za davanje preporuka: filtriranje temeljeno na sadržaju (eng. content-based filtering) i suradničko filtriranje (eng. collaborative filtering).

2.4. Model linearne regresije

Linearna regresija je metoda za modeliranje odnosa između zavisne varijable i jedne ili više nezavisne varijable, kako bi se otkrila i kvantificirala snaga korelacije među njima. Linearna regresija spada u algoritme nadziranog učenja i jedan je od najstarijih i najpoznatijih statističkih modela koji je još uvijek itekako relevantan. Osim objašnjavanja prirode ovisnosti promatranih pojava na temelju tog analitičkog oblika može se vršiti predviđanje vrijednosti zavisne varijable pri određenim vrijednostima neovisnih varijabli (Pivac, 2010).

Postoje razne vrste modela linearne regresije poput jednostruka, višestruke ili lasso linearne regresije, ali zbog jednostavnosti objasniti će se jednostruka linearna regresija. Grubo rečeno, to je regresijski model za predviđanje kvantitativnog odgovora Y iz jedne varijable prediktora X. Pritom se pretpostavlja da X i Y imaju linearni odnos.

Matematički, ovaj se regresijski model može izraziti kao:

$$Y = \beta 0 + \beta 1 * X$$

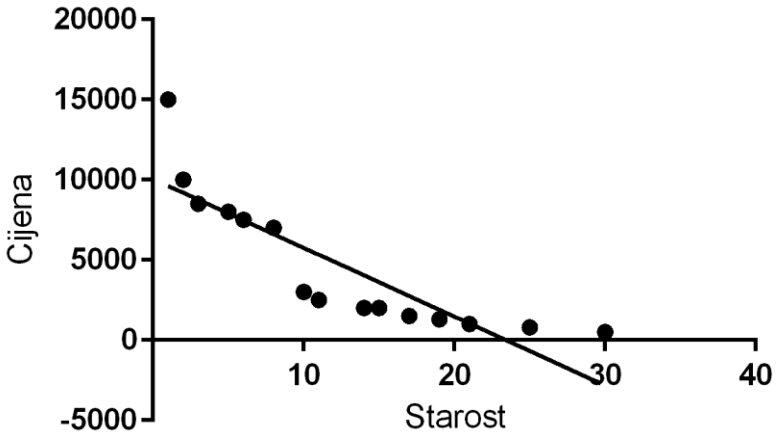
Gdje simbol Y predstavlja zavisnu varijablu, a X predstavlja nezavisnu varijablu. Simbol $\beta 0$ predstavlja konstantu vrijednosti, odnosno vrijednosti kada zavisna varijabla X iznosi 0. Simbol

$\beta 1$ predstavlja regresijski koeficijent i ono pokazuje prosječnu promjenu zavisne varijable kada nezavisna poraste za jednu jedinicu. Ovaj koeficijent interpretira se i kao koeficijent smjera i ono određuje nagib pravca (Pivac, 2010).

Ukoliko se za primjer uzme cijena automobila, cijena može biti pojednostavljeno prikazana kao odnos sa starosti automobila. Uzevši prethodno opisanu formulu, nova formula se može izraziti kao:

$$cijena = \beta 0 + \beta 1 * starost$$

Simboli $\beta 0$ i $\beta 1$ su nepoznanice, a model ih računa kroz analizu podataka o automobila. Ukoliko se proizvoljno odaberu određene vrijednosti starosti i cijena automobila, dobiti će sljedeći model linearne regresije.



Graf 1: Linearna regresija cijene automobila

Izvor: Izrada autora

Starost ima negativnu korelaciju s cijenom automobila, odnosno što je starost veća tj. što je automobil stariji to će cijena biti manja. U ovom slučaju radi se o negativno nagnutoj linearnoj regresiji, te će $\beta 1$ biti negativan.

Pravac treba biti takav tako da prolazi između stvarnih točaka promatranih varijabli i da najbolje tumači vezu između njih, odnosno pravac mora biti takav da odstupanja e_i budu najmanja (Pivac, 2010). Simbol e_i predstavlja i-ti rezidual, odnosno razliku između i-te promatrane stvarne vrijednosti i i-te promatrane očekivanje vrijednosti. Cilj svakog regresijskog modela je minimizirati sumu kvadrata reziduala e_i, odnosno da razlika između očekivanje (predviđene) vrijednosti bude što manja u odnosu na stvarnu vrijednost.

Postoji nekoliko funkcija prema kojoj se mjeri odstupanje očekivanih vrijednosti od stvarnih i prema kojoj se definira uspješnost modela linearne regresije, a to su:

Srednja apsolutna greška (eng. Mean Average Error)

To je prosječna razlika između predviđenih i stvarnih vrijednosti u cijelom ispitanom skupu podataka, a često se koristi kada postoje značajne ekstremne vrijednosti (eng. outliers) u podacima. Uz to vrlo je jednostavan za interpretaciju. Međutim, teže naglašava razlike u grešci između vrijednosti nego druga mjerila[15].

Srednja kvadratna greška (eng. Mean Squared Error)

Predstavlja kvadratnu vrijednost prosječne razlike predviđenih i stvarnih vrijednosti koja je mnogo češća pri analizi regresijskih modela iz razloga što mnogo bolje naglašavaju greške u podacima i kažnjavaju vrlo velike greške.[16] Također, vrijednosti su uvijek pozitivne, pa je stoga bolje što su bliže nuli.

Korijen srednje vrijednosti kvadrata grešaka (eng. Root Mean Squared Error)

Ova metrika dijeli mnoga svojstva sa srednjom kvadratnom greškom jer je to jednostavno njen korijen, ali bolji je u smislu održavanja performansi kada se radi s velikim vrijednostima

[15] https://akhilendra.com/evaluation-metrics-regression-mae-mse-rmse-rmsle/
[16] Ibid.

pogreške[17]. Ujedno je ovo mjerilo mnogo osjetljivije na ekstremne vrijednosti (outlier-e), a kod stvaranja algoritama često je cilj smanjiti utjecaj tih vrijednosti na podatke i na model.

Važno je razumjeti i odabrati pravilni način mjerenja rezultata regresijskog modela jer se prema tome može odrediti radi li model onako kako bi trebao te jesu li predikcije koje model stvara pouzdane. Tijekom praktičnog dijela ovoga rada često će se koristiti navedena mjerila, a prema njima će se stvarati konačni zaključak.

2.5. Duboko učenje

Prema Kelleher (2019) duboko učenje je područje umjetne inteligencije usredotočeno na stvaranje velikih modela neuronskih mreža koji su sposobni donijeti valjanje odluke na temelju danih podataka. Duboko učenje zapravo je nastalo kroz istraživanje umjetne inteligencije u kombinaciji sa strojnim učenjem. Stoga, može se reći da je duboko učenje mnogo uža niša strojnog učenja na što se ono nadovezuje. Odnos između umjetne inteligencije, strojnog učenja i dubokog učenja prikazani je na slici 10.

[17] Ibid.

Slika 10: Odnos umjetne inteligencije, strojnog i dubokog učenja

Izvor: Kelleher, J. D. (2019), Duboko učenje, MATE d.o.o., str 6.

Duboko učenje omogućuje odlučivanje na temelju podataka, pronalaženjem uzoraka u u velikim skupovima podataka koji se zatim precizno preslikavaju u valjane izlazne odluke. Upravo posljednji dio rečenice, koji govori o preslikavanju ulaznih vrijednosti u izlazne Kelleher (2019) najviše naglašava kao najvažnije svojstvo dubokog učenja. Ta preslikavanja mogu biti jednostavne aritmetičke funkcije, slijed pravila poput "ako-onda-inače" ili druge složenije funkcije. Uzorci koje algoritam dubokog učenja izvlači iz podataka funkcije su koje su predstavljenje umjetnih neuronskim mrežama, a upravo su one koncept na koji se duboko učenje usko oslanja.

Sljedeće poglavlje dublje će ući u temu umjetnih neuronskih mreža, a govoriti će se o kratko o njihovom povijesnom razvoju, arhitekturi neuronskih mreža i mogućnostima neuronskih mreža u obrazovanju.

3. UMJETNE NEURONSKE MREŽE

3.1. Kratka povijest umjetnih neuronskih mreža

Kao što je spomenuto, umjetne neuronske mreže čine osnovu dubokog učenja. Kada se kaže duboko učenje ne misli se na sposobnost algoritma da provodi dublje operacije učenja, već se misli na veliki niz slojeva neurona unutar jedne neuronske mreže što ga čini "dubokim". Kelleher (2019) opisuje tijek istraživanja umjetnih neuronskih mreža u tri razdoblja: razdoblje logičkih jedinica tipa prekidača (od 1940-ih do sredine 1960-ih), razdoblje konekcionizma (od početka 1980-ih do sredine 1990-ih) i razdoblje dubokog učenja (od sredine 2000-ih do danas).

Znanost o umjetnim neuronskim mrežama je zapravo prilično stara, a prvi put su spomenuti već 1943. od strane američkih neurofiziologa Warren McCulloha i matematičara Walter Pittsa Naime, u njihovom radu, pomoću matematičke logike, opisan je način na koji ljudi i životinje procesiraju informacije u mozgu putem neurona. Prema njima, svaki ulaz i izlaz u neuronu imao je vrijednost 0 ili 1, a zbroj ulaza predstavljao je prekidačku funkciju za sljedeći neuron. Ako je rezultat zbrajanja bio veći od postavljenog praga, izlaz neurona bi poprimio vrijednost 1, a u protivnom bi poprimio vrijednost 0 (Kelleher, 2019).

Otprilike u isto vrijeme Frank Rosenblatt, američki psiholog, predstavio je svijetu prvi koncept "perceptrona" (Kelleher, 2019), jednostavnog matematičkog izraza umjetnog neurona gdje se prvi put spominju težinski indeksi. Na taj način postavljeni su temelji umjetnih neuronskih mreža koji su i dalje relevantni iako mnogo sofisticiraniji.

Unatoč velikom ushićenju povodom novih otkrića, uslijedilo je "mračno doba" istraživanja o umjetnim neuronskim mrežama zbog rigidnosti osnovnih perceptrona i nedostatka financiranja u to područje (Geron, 2019).

Ranih 1980-ih započima razdoblje konekcionizma kada počima ponovno interest za istraživanje umjetnih neuronskih mreža. Dva su izuma ključna za taj period: Hopfieldove mreže, odnosno prvi modeli višeslojnih perceptrona i algoritam povratnog širenja pogreške (eng. Backpropagation), možda i najvažniji algoritam dubokog učenja (Kelleher, 2019). Na temelju toga otkriveni su

koncepti posve novih tipova umjetnih neuronskih mreža kao što su konvolucijske neuronske mreže (eng. Convolutional Neural Networks) koji su u mogućnosti prepoznavati slike, te povratne neuronske mreže (eng. Recurrent Neural Networks) koje imaju kratkotrajnu memoriju za pamćenje vrijednosti pogodni za obradu prirodnog jezika i strojnog prevođenja.

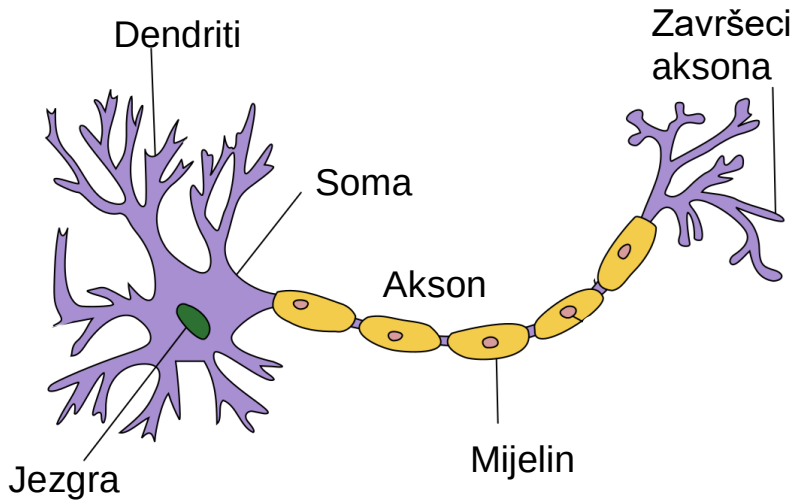
Moderno razdoblje od sredine 2000-ih karakterizira nagli razvoj informatičkih tehnologija i grafičkih jedinica. Sve jača računala rezultirala su širokim i naglim razvojem umjetnih neuronskih mreža, a zbog interneta i sve većeg rasta količine raspoloživih podataka, paralelno je rasla i potreba za algoritmom koji je u mogućnosti obraditi toliku količinu podataka. Razvoj dubokog učenja može se opisati kao zatvoreni krug razvoja hardware-a, rasta količine podataka i sve sofisticiranijih algoritama dubokog učenja (Kelleher, 2019).

3.2. Arhitektura umjetnih neuronskih mreža

Nije moguće objasniti arhitekturu, ulogu i funkciju umjetnih neuronskih mreža bez da se najprije objasni arhitektura neuronskih mreža u mozgu na čiju su inspiraciju i nastali. Najšire gledano, umjetne neuronske mreže su modeli strojnog učenja namijenjeni da oponašaju način na koji prave neuronske mreže funkcioniraju i izvode zadatke (Haykin, 2009).

3.2.1. Biološka neuronska mreža

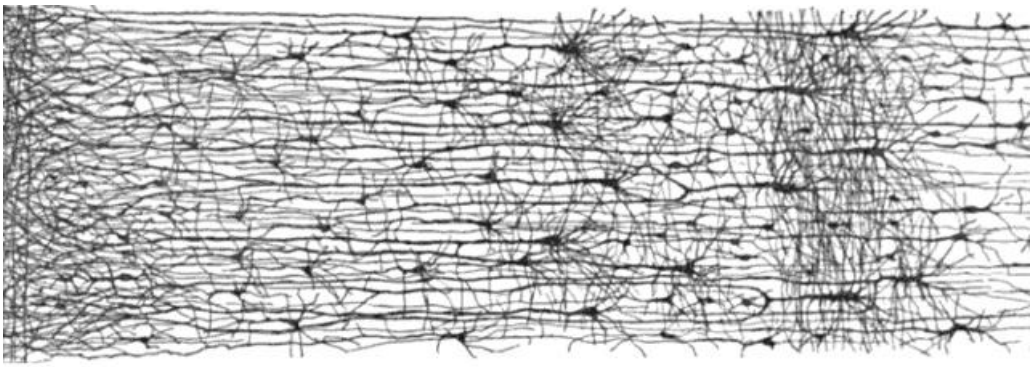
Neuronske mreže sastoje se od brojnih manjih jedinica koji se nazivaju neuroni. Neuron je sastavljen od tijela stanice koje sadrži jezgru i većinu složenih komponenti stanice i mnogih granatih produžetaka koji se nazivaju dendriti, te jedan vrlo dugačak nastavak koji se zove akson. Na samim krajevima aksona postoje sitni produžeci kojima je neuron spojen s dendritima drugog neurona čime se stvara kompleksna mreža unakrsno spojenih neurona.



Slika 11: Arhitektura biološkog neurona

Izvor: https://hr.wikipedia.org/wiki/Neuron

Ove sinapse omogućuju biološkim neuronima primanje kratkih električnih impulsa od drugih neurona, poznatih kao signali. Neuron ispaljuje vlastite signale kada primi dovoljan broj signala od drugih neurona unutar nekoliko milisekundi (Geron, 2019). Smatra se da ljudski korteks sadrži otprilike 10 milijardi neurona i 60 trilijuna sinapsi ili veza. Kao rezultat toga, mozak je izuzetno učinkovita struktura (Haykin, 2009). Slika 12 prikazuje reprezentaciju neuronskih mreža u mozgu.

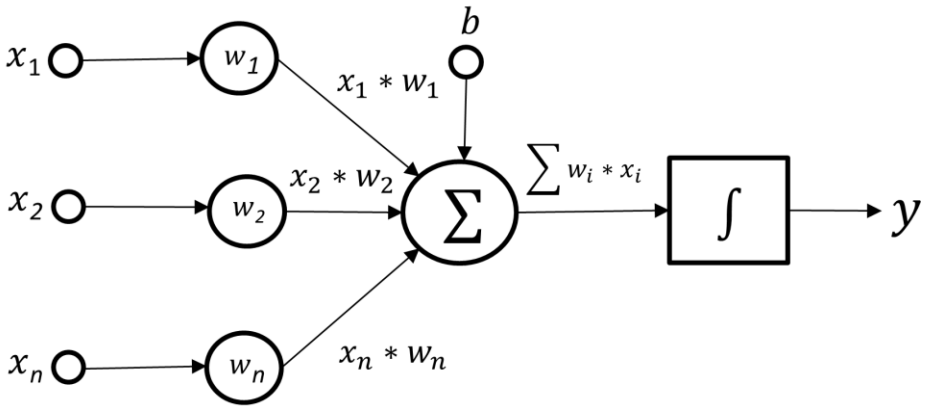


Slika 12: Biološka neuronska mreža

Izvor: https://sensartorg.files.wordpress.com/2017/08/nissli-adult-golgiinfant-cortex.jpg

3.2.2. Umjetna neuronska mreža

Umjetni neuron je jedinica za obradu informacija koja je temeljna za rad umjetnih neuronskih mreža. Kratko je spomenuto u ranijem poglavlju da je jedan od ključnih izuma u razvoju umjetnih neuronskih mreža tzv. perceptron kojeg je izmislio Frank Rosenblatt. Perceptron je jedan od najosnovnijih i najstarijih arhitektura umjetnih neuronskih mreža, a temelj mu leži u logičkoj jedinici za računanje praga tolerancije na podražaj, vrlo slično principu rada bioloških neurona (Geron, 2019). Dok su teorijski koncepti umjetnih neurona bili zasnovani na jednostavnom binarnom obliku, gdje su ulazi i izlazi bili 0 ili 1, perceptron pomoću težinskih faktora zbraja sve ulazne vrijednosti, te prema pragu tolerancije određuje izlaz neurona i njegovu jačinu. Slika 13 prikazuje arhitekturu osnovnog perceptrona.



Slika 13: Arhitektura perceptrona

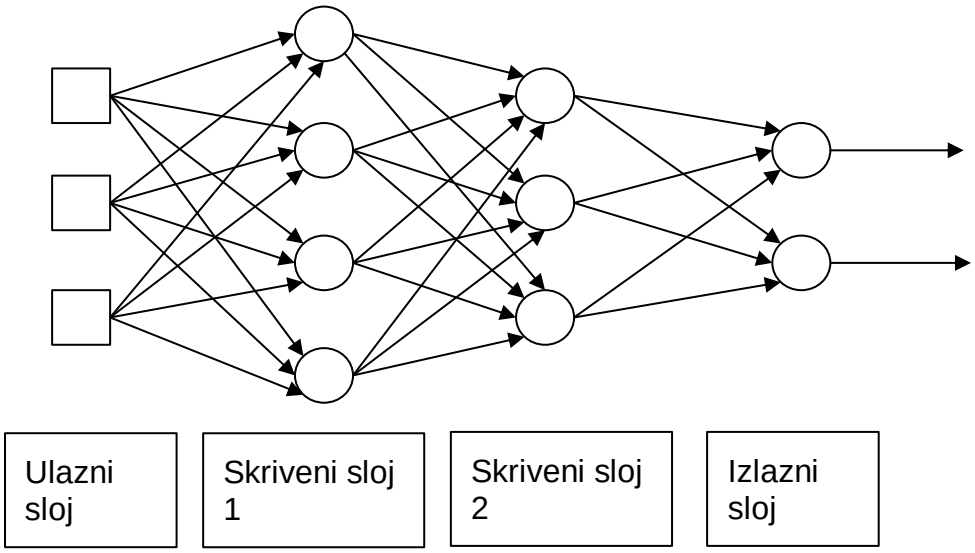
Izvor: Izrada autora prema Haykin, S. (2009), Neural Networks and Learning Machines, Pearson Education Inc., str. 11.

Svaki umjetni neuron sastoji se od njegovih sinapsi, odnosno ulaznih vrijednosti koje poprima početnom, zadanom vrijednosti ili preko drugih neurona. Njegova vrijednost se tada množi s pripadajućim težinama (w) koje se zajedno zbrajaju te neuron dobiva svoju težinsku vrijednost. Težinski faktori omogućuju umjetnoj neuronskoj mreži da ojača ili oslabi vezu između neurona,

čime neuronska mreža može stavljati naglaske na određene vrijednosti koji su od većeg značaja za konačni rezultat.

Svakom neuronu je također pridodana određena pristranost (eng. bias) ovisno o tome je li funkcija pozitivna ili negativna. Pristranost se može koristiti za prilagodbe unutar samog neurona. Na sumu ulaznih vrijednosti neurona se zatim primjenjuje aktivacijska funkcija koja je definirana unaprijed i može poprimiti više različitih funkcija, o kojima će biti više riječi u nastavku. Preko aktivacijske funkcije dolazi se do izlaznog signala koja odlazi u sljedeći sloj neurona ili može biti konačna vrijednost mreže.

Umjetna neuronska mreža sastavljena je od velikog broja neurona, tzv. perceptrona opisanih prethodno. Tipična neuronska mreža dijeli se na više slojeva, jedan ulazni sloj, barem jedan skriveni sloj i jedan izlazni sloj. Kada se kaže skriveni sloj misli se samo na sloj koji nije ni ulazni ni izlazni (Kelleher, 2019). Neuronska mreža može imati proizvoljno veliki broj skrivenih slojeva, a sa svakim novim skrivenim slojem raste i kompleksnost neuronske mreže.



Slika 14: Umjetna neuronska mreža

Izvor: Izrada autora prema Haykin, S. (2009), Neural Networks and Learning Machines, Pearson Education Inc., str. 22.

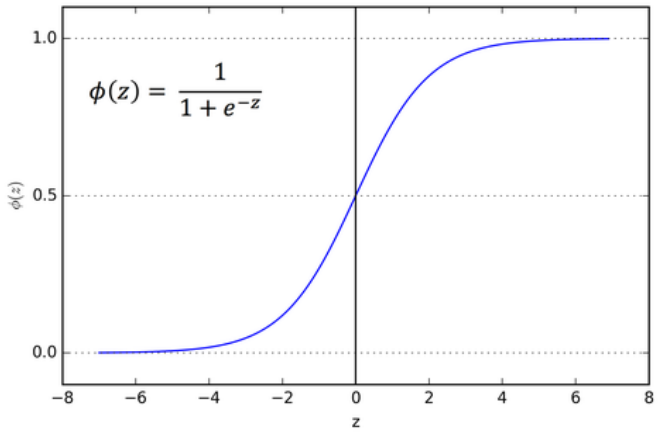
Slika 14 prikazuje umjetnu neuronsku mrežu s 4 sloja, jednim ulaznim, dva skrivena sloja i jednim izlaznim slojem. Kvadrati u ulaznom sloju predstavljaju mjesta unutar neuronske mreže koja služe kao ulazi u mrežu. U tom sloju nema obrade informacija već je u njima pohranjena vrijednost podataka pohranjenih u toj memorijskoj lokaciji. Skriveni slojevi su slojevi neurona koji uzimaju vrijednost prethodnog sloja, odnosno prethodnog neurona i preslikavaju ih u jednu izlaznu vrijednost (Kelleher, 2019) i to na način prethodno opisan u dijelu o perceptronu.

Strelice na slici također pokazuju smjer protoka informacija, koji je u ovom slučaju jednosmjeran. Postoje i vrste umjetnih neuronskih mreža koji procesiraju informacije dvosmjerno, a o tome će biti više u poglavlju o povratnim neuronskim mrežama. Broj neurona u izlaznom sloju varira o vrsti problema s kojima se radi, odnosno o broju potencijalnih rješenja. U ovom slučaju broj rješenja je dva, pa se radi o binarnoj klasifikaciji.

3.3. Vrste aktivacijskih funkcija

Ključan korak prilikom procesiranja informacija neurona je obrada podataka aktivacijskom funkcijom. Svrha aktivacijske funkcije je da odredi hoće li vrijednosti unutar neurona biti prihvaćene u odnosu na prag podražaja, odnosno hoće li neuron biti aktiviran. Aktivacijske funkcije također uvode nelinearnost u neuronske mreže jer ukoliko bi neuroni obrađivali informacije samo težinskim faktorima rezultat bi bio vrlo linearna funkcija koja ne bi imala primjenu u stvarnom svijetu. Aktivacijske funkcije stoga izvode nelinearna preslikavanja ulaza u izlaze prema čemu su u mogućnosti izvoditi kompleksne zadatke, a ovisno o vrsti funkcije najčešće mapiraju vrijednosti između 0 i 1 ili -1 i 1. Neke najpopularnije aktivacijske funkcije su:

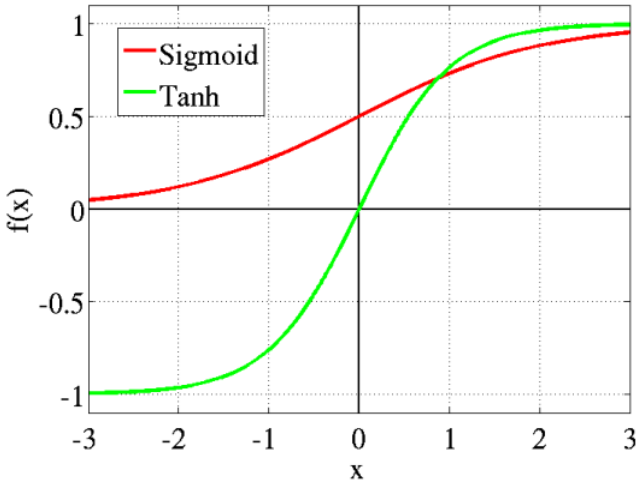
1. **Logistička aktivacijska funkcija** - također zvana Sigmoid funkcija, ova aktivacijska funkcija mapira vrijednosti između 0 i 1, odnosno vrijednost može biti prihvaćena ili odbačena. Ova funkcija se često koristi kod klasifikacijskih zadataka iako je popularnija alternativa tzv. Softmax aktivacijska funkcija kad je u pitanju multiklasna klasifikacija.



Slika 15: Logistička aktivacijska funkcija

Izvor: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

2. **Tanh aktivacijska funkcija** - vrlo slična logističkoj funkciji osim što mapira vrijednosti između -1 i 1 umjesto 0 i 1. Prednost toga je što ova funkcija uzima u obzir negativne vrijednosti i pritom radi manje generalizacije.

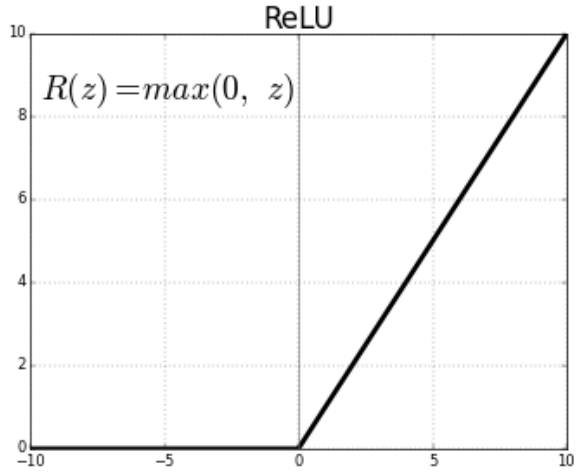


Slika 16: Tanh aktivacijska funkcija

Izvor: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

3. **ReLU (Rectified Linear Unit)** - danas je najpopularniji izbor mnogih podatkovnih znanstvenika i preporuča se kod većine algoritama neuronskih mreža bez povratnih veza

(feedforward neural networks). Karakteristično za ovu funkciju je što joj je interval između 0 i +∞ te pritom uzima u obzir velike i kontinuirane vrijednosti. ReLU funkcija rješava problem rigidnosti logističke ili tanh funkcije koje imaju preveliku osjetljivost u dodjeljivanju vrijednosti prilikom preslikavanja. Problem kod prethodno navedenih funkcija je što sve veće vrijednosti automatski budu dodijeljeni najveću vrijednost 1 ili suprotno, male vrijednosti budu dodijeljeni 0[18]. Danas postoje i varijacije na ReLU funkciju poput **Leaky ReLU** gdje interval vrijednosti umjesto 0, +∞ biva -∞,+∞.



Slika 17: ReLU aktivacijska funkcija

Izvor: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6

3.4. Kako umjetne neuronske mreže uče

Do sada je objašnjenja arhitektura umjetnih neuronskih mreža i uloga njihovih pojedinih dijelova, međutim da bi neuronska mreža bila u mogućnosti izvoditi zadatke visoke kompleksnosti i na kraju učiti na temelju podataka, potrebno je definirati tri važne funkcije odnosno algoritma.

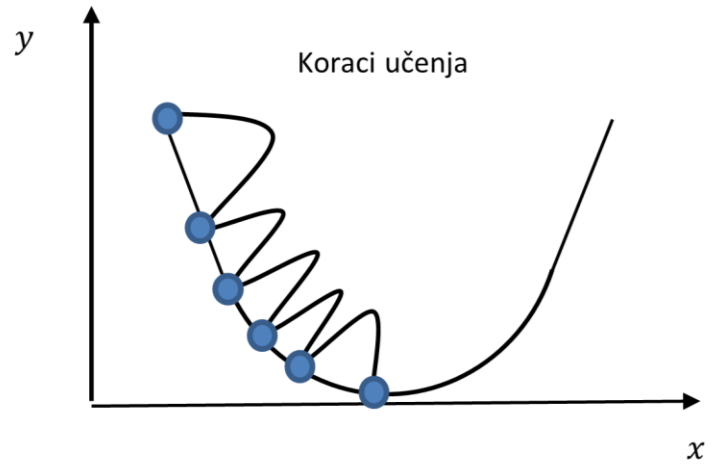
Funkcija troška predstavlja vrlo važan pojam u svijetu strojnog učenja pa tako i kod umjetnih neuronskih mreža. Prilikom stvaranja algoritama strojnog učenja potrebno je odrediti cilj prema

[18] https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/

kojemu se može zaključiti je li algoritam uspješan ili neuspješan u tome što nastoji postići. Taj cilj, ovisno o problemu, može se nastojati maksimizirati ili minimizirati. Preciznost određenih modela, kao kod klasifikacije, nastoji se maksimizirati, dok je kod regresije cilj minimizirati pogrešku između stvarnih i predviđenih vrijednosti.

Gotovo svi algoritmi učenja imaju neku funkciju gubitka, a kod umjetnih neuronskih mreža cilj je minimizirati tzv. funkciju troška kako bismo pronašli najbolji model za određeni problem. Funkcija troška u linearnoj regresiji dana je prosječnim gubitkom, također poznatim kao empirijski rizik. Prosječni gubitak ili empirijski rizik je zbroj svih grešaka (razlike između stvarnih i previđenih vrijednosti) dobivenih primjenom modela na podacima za trening (Burkov, 2019). U slučaju regresijskog problema kao sredstvo mjerila uspješnosti modela koriste se kvadratne funkcije prethodno opisane u poglavlju o linearnoj regresiji, a to su srednja apsolutna pogreška (MAE), srednja kvadratna pogreška (MSE) i/ili korijen srednje kvadratne pogreške (RMSE).

S ciljem optimizacije funkcije troška umjetnih neuronskih mreža koriste se optimizacijske metode kao što je **algoritam gradijentnog spusta**. Opća ideja koja stoji iza gradijentnog spusta je da je to iterativno podešavanje parametara, odnosno postepena optimizacija hiper parametara modela kako bi se minimizirala funkcija troška i stvorio najoptimalniji model. Gradijentno spuštanje ima takav naziv jer ima za cilj pronaći lokalni minimum svake funkcije, a radi to kroz male postepene korake u optimizaciji određeno stopom učenja (Geron, 2019). Što je stopa učenja veća, brža će biti optimizacija modela, ali će preciznost biti manja. S druge strane, što je stopa učenja manja, proces optimizacije će biti znatno sporiji, ali će model biti bolje optimiziran.



Slika 18: Gradijentni spust

Izvor: Izrada autora prema Geron, A. (2019), Hands-on Machine Learning with Scikit-Learn, Keras and TensorFlow, O'Reilly, str 120.

Međutim, nakon što se otkrio lokalni minimum za svaku funkciju kroz cijelu neuronsku mrežu i nakon što se izračunala izlazna vrijednost mreže ne može se još reći da je model optimiziran. Najveća moć umjetnih neuronskih mreža leži u njihovoj sposobnosti samostalne optimizacije, a to radi preko algoritma povratnog širenja pogreške u kombinaciji s gradijentnim spustom. **Algoritam povratnog širenja pogreške** (eng. Backpropagation) jedan je od najvažnijih koncepata u razumijevanju umjetnih neuronskih mreža te on zapravo predstavlja vještinu "učenja" neuronskih mreža (Kelleher, 2019).

Algoritam povratnog širenja pogreške odvija se u dvije faze: širenju prema naprijed i širenju unatrag. U prvoj fazi, širenja prema naprijed, mreža uzima informacije ulaznih podataka, računaju se vrijednosti neurona skrivenih slojeva koji ostaju pohranjeni u memoriji mreže i dobivaju se izlazne vrijednosti prema kojima se računa ukupna pogreška mreže. U fazi širenja unatrag pogreška mreže se prosljeđuje natrag kroz mrežu, pri čemu svaki neuron prima dio krivnje za ukupnu pogrešku mreže, razmjerno njegovoj osjetljivosti pogreške na promjene (Kelleher, 2019). Nakon što se odredila osjetljivost na promjene svakog neurona, algoritam gradijentnog spusta ponovnim širenjem prema unaprijed računa nove pogreške težina i ažurira same težine. Algoritam

povratnog širenja pogreške ponavlja se kroz određeni broj ponavljanja ili "epoha", sve dok se pogreška mreže ne smanji na prihvatljivu razinu.

3.5. Vrste umjetnih neuronskih mreža

Ovisno o vrsti problema i o smjeru protoka informacija danas su u upotrebi česte tri vrste umjetnih neuronskih mreža.

Višeslojna mreža bez povratnih veza (eng. Multilayered Feedforward Network)

Najosnovniji oblik umjetnih neuronskih mreža koji je sastavljen od ulaznog sloja, nekoliko skrivenih slojeva te jednog izlaznog sloja, a protok informacija je bez povrata. Ovaj oblik mreža je opisan u početnom poglavlju o umjetnih neuronskih mrežama pa se neće ulaziti u detalje, ali ovaj tip mreža može se koristit za razne vrste problema strojnog učenja poput regresije ili klasifikacije te jer zbog široke namjene vrlo popularan u svijetu podatkovne znanosti.

Konvolucijske neuronske mreže (eng. Convolutional Neural Network)

Konvolucijske neuronske mreže stvorene su 1980-ih godina prošlog stoljeća sa zadaćom prepoznavanja objekata na slikama. Osnovni cilj ovih mreža bio je raspoznavanje lokaliteta određenih vizualnih značajki koji će zatim biti upotrijebljeni u značajkama višeg reda, odnosno u drugim zadacima (Kelleher, 2019). Prema toma, osnovna zadaća konvolucijskih mreža je funkcija otkrivanja značajki koje mogu prepoznati prisutnost ili odsutnost tih istih značajki na slikama. Kelleher (2019) opisuje da konvolucijske neuronske mreže moraju biti u stanju otkrivati značajke na način *"koji je invarijantan s obzirom na translacije"*, odnosno, da mogu prepoznati uzorke u slikama (lice, nebo, pas) neovisno gdje se ti objekti na slici nalazili ili na kojim se slikama nalazili. Čestu upotrebu nalaze u algoritmima za prepoznavanje lica ili rukopisa.

Povratne neuronske mreže (eng. Recurrent Neural Network)

Povratne neuronske mreže namijenjene su obradi slijednih podataka, a najčešće se koriste u obradi teksta i glasova. Ove mreže najčešće su sačinjene od samo jednog skrivenog sloja neurona, ali sadrže i memorijski međuspremnik koji pohranjuje izlaz tog skrivenog sloja za jedan ulaz i zatim ga vraća nazad u skriveni sloj zajedno sa sljedećim ulazom (Kelleher, 2019). Zbog toga, sve informacije obrađuju se u kontekstu prethodne obrade informacije, odnosno prethodnog ulaza, koji je opet obrađen u kontekstu koji je njemu prethodio. Zbog toga što su tekstovi nizovi riječi i interpunkcijskih znakova, ove mreže idealne su za zadatke obrade teksta. Postoji i još jedna varijacija povratnih neuronskih mreža koji se zovu LSTM (eng. Long Short-Term Memory) koji su u mogućnosti spremati informacije na mnogo duže vremenske periode te su zbog toga pogodne za obradu prirodnog jezika (eng. Natural Language Processing), odnosno kompjutorske obrade jezika u stvarnom vremenu.

3.6. Umjetne neuronske mreže u obrazovanju

Tehnologije umjetnih neuronskih mreža dostupne su i prisutne u gotovo svim djelatnostima i granama ljudskog društva, ali vrlo malo u obrazovanju. Tehnologije poput povratnih neuronskih mreža koriste se prilikom obrade jezika za detekciju plagijarizma u analizi studentskih radova, koriste se u detekciji ružnih riječi i izraza za stvaranje filtera i prilikom stvaranja prijevoda ili transkripcija online predavanja, ali u pogledu unapređivanja kvalitete kolegija, usvajanja gradiva ili poboljšanja kurikuluma, moderne tehnologije strojnog i dubokog učenja nisu pronašle primjenu.

Pandemija COVID-19 uzorovala je nagli razvoj načina na koji komuniciramo preko web sučelja i uvela nove načine na koji izvodimo nastavu. Razvojem sustava za online učenje dovelo je do nakupljanja velike količine informacija u bazama podataka obrazovnih institucija te se otvara pitanje potencijala tih podataka. Obrazovno rudarenje podataka (eng. Educational Data Mining) je pedagoški pristup koji se temelji na podacima te koji koristi tehnike podatkovne znanosti kao što su umjetna inteligencija, rudarenje podataka i baze podataka (Okewu i sur., 2021). Obrazovno rudarenje podatka relativno je novo područje koje se upravo razvilo zbog masovne dostupnosti

podataka i digitalnih alata za njihovu obradu, ali i potrebi za obradom tih podataka u sferi obrazovanja.

Altaf (2017) definira obrazovno rudarenje podataka kao granu istraživanja koja primjenjuje rudarenje podataka, statistiku i strojno učenje na obrazovne podatke. Dakle, osnova je ovog koncepta korištenje alata i tehnika podatkovne znanosti u obrazovnom kontekstu. Cilj obrazovnog rudarenja podataka je izvući korisne informacije iz velike količine podataka koje mogu pomoći u boljem donošenju odluka u vidu obrazovanja te boljem shvaćanju procesa učenja i ponašanja studenata. Koristeći se alatima strojnog učenja, a u kontekstu ovog rada, umjetnim neuronskih mrežama moguće je (Altaf, 2017):

- Predvidjeti performanse studenata kako bi se omogućila pravovremena intervencija i spriječilo prijevremeno opadanje
- Predvidjeti optimalni kurikulum prema studentovim performansama na prijašnjim kolegijima
- Steći uvid u studentov proces učenja ili učiteljeve tehnike predavanja
- Unaprijediti motivaciju studenata i smanjiti troškove institucija kroz optimizaciju kolegija i programa

Zbog svoje kompleksnosti i računalne snage umjetne neuronske mreže su u mogućnosti analizirati ogromne količine podataka i uočiti ključne uzorke u njima na temelju kojih se mogu stvarati predikcije o ponašanju ili ishodu studenata. Zbog toga, umjetne neuronske mreže privlače sve veću pažnju unutar obrazovnog konteksta kao moćno sredstvo analize podataka i stvaranja predikcija (Kehinde i sur., 2021).

Kehinde i sur. (2021) upućuju upravo na potencijal umjetnih neuronskih mreža u svom radu gdje pomoću višeslojnih neuronskih mreža bez povratnih veza, na temelju demografskih podataka studenata, uspješno stvaraju model koji je, s preciznošću od 92%, u mogućnosti predvidjeti konačni ishod studenata.

Zanimljivo istraživanje od Pavlin-Bernardić i sur. (2016) provedeno u hrvatskoj osnovnoj školi za cilj je imalo predvidjeti nadarenost kod djece pomoću višeslojnih neuronskih mreža. Autori su uspješno stvorili klasifikacijski model koji može s visokom preciznošću klasificirati nadarenu

djecu na temelju školskih ocjena, procjene spremnosti za školu i obrazovanju njihovih roditelja. Autori u istom radu pritom navode na mogućnosti umjetnih neuronskih mreža da *"pomognu učiteljima u donošenju odluka, posebno u školama koje imaju manjak psihologa"*.

Ersoz Kaya (2019) je također proveo istraživanje koristeći se umjetnim neuronskim mrežama gdje je predvidio ishod studenata na temelju njihovih ocjena prethodnih kolegija. Na temelju toga zaključio je da se pomoću neuronskih mreža može doći do povezanosti između pojedinih predmeta i studentovog ishoda, na temelju čega se može bolje određivati prioritet i redoslijed kolegija u studijskom programu.

Široka primjena umjetnih neuronskih mreža može se pronaći svugdje, pa tako i u obrazovanju. Međutim, usvajanjem ovih tehnologija kao sredstvo pomaganja pri donošenju odluka, a ne samo kao koristan alat, može donijeti velike promjene u načinu na koji pristupamo učenju i organizaciji školskih programa. Koristeći se raspoloživim podacima u sustavima za online učenje, neuronske mreže su u mogućnosti uočiti uzorke koji ljudi ne mogu, ali na temelju kojih ljudi mogu donositi širi raspon odluka. Istraživanja upućuju da umjetne neuronske mreže mogu biti prikladne za promicanje pametnog obrazovanja, inteligentnog podučavanja i pružanja akademskog savjetovanja (Okewu i sur., 2021).

4. PREDVIĐANJE KONAČNOG ISHODA STUDENATA POMOĆU UMJETNIH NEURONSKIH MREŽA

4.1. Opis problema

Do sada je bila objašnjena teorijska podloga umjetnih neuronskih mreža i modela linearne regresije, njihov princip rada i funkciju svih algoritama koji im pripadaju. U praktičnom dijelu rada nastojat će se ukazati na mogućnosti i moć umjetnih neuronskih mreža u stvaranju predikcije na temelju podataka o studentskim ispitnim rezultatima. Pomoću tih podataka dobiti će se uvid u konačni ishod studenata na kraju treće godina studija, a zatim će se modelom linearne regresije i umjetnih neuronskih mreža usporedno predvidjeti taj ishod na novim, neviđenim podacima.

Usporednom analizom linearne regresije i umjetnih neuronskih mreža dobiti će se bolji uvid u sposobnost i snagu neuronskih mreža u stvaranju predikcija. Stvaranje modela oba algoritma posljednji je dio praktičnog dijela, a prethodi mu dio o opisu podataka, čišćenju podataka i inženjeringu značajki (eng. feature engineering) na podacima.

Konačno, cilj je također ukazati na mogućnosti stvaranja predikcije konačnog uspješnosti studenata na temelju njihove uspješnosti na prvoj godini studija, što može pak omogućiti pravovremenu intervenciju studenata koji su pri većem riziku od ranog opadanja s fakulteta, uočavanju veza između pojedinih kolegija i stoga boljoj alokaciji resursa pojedinim studentima ili poboljšanju programa određenih kolegija.

Obrada podataka raditi će se u Python programskom jeziku koji je danas najpopularniji jezik podatkovnih znanstvenika i analitičara zbog svoje jednostavne sintakse, blage krivulje učenja i mogućnosti lake automatizacije zadataka. Python je ujedno opremljen s nizom programskih biblioteka, koji omogućuju brže i jednostavnije izvođenje koda, specifično dizajnirane za poslovne takvoga tipa. Neki od tih biblioteke namijenjene su manipulaciji podacima i vektorima, neke su pogodne za vizualizaciju, a neke su stvorene specifično za stvaranje modela strojnog učenja i umjetnih neuronskih mreža.

4.2. Alati korišteni u procesu

U nastavku će se opisati najosnovniji alati i programske biblioteke korištene u praktičnom dijelu kako u analizi podataka, stvaranju vizualizacija tako i kod stvaranja konačnih predikcija.

Python i biblioteke temeljene na Pythonu

Python je popularan programski jezik opće namjene koji se može koristiti za širok raspon aplikacija, a njegova filozofija dizajna daje prednost čitljivosti koda korištenjem značajnog uvlačenja. Činjenica da je Python programski jezik otvorenog koda, odnosno da svatko može mijenjati kod i raditi vlastite modifikacije na jeziku, čini ga jednim od programskih jezika s najvećom zajednicom korisnika i najpopularnijim programskim jezikom na svijetu. Bez obzira na njihovu razinu iskustva, programeri iz različitih sredina doprinose jeziku na značajan način. Zbog toga, Python nalazi primjenu u raznim granama IT-a kao što je:

- web razvoj
- razvoj software-a
- znanstveno računalstvo
- strojno učenje

Stoga, praktični dio će se u cijelosti raditi u Python programskom jeziku, a najvažnije biblioteke zasnovane na Python-u namijenjene podatkovnoj znanosti i analizi podataka su: NumPy, Pandas, Matplotlib, Scikit-Learn i TensorFlow.

NumPy

NumPy je biblioteka za programski jezik Python, koja daje podršku za velike, višedimenzionalne nizove i matrice, zajedno s velikom zbirkom matematičkih funkcija za rad s tim nizovima. NumPy predstavlja temelj za sve ostale biblioteke namijenjene strojnom učenju jer se svi podaci prije unošenja u model strojnog učenja pretvaraju u matrice i nizove brojeva.

```
np.random.randn(5,5)
```

```
array([[ 0.70154515,  0.22441999,  1.33563186,  0.82872577, -0.28247509],
       [ 0.64489788,  0.61815094, -0.81693168, -0.30102424, -0.29030574],
       [ 0.8695976 ,  0.413755  ,  2.20047208,  0.17955692, -0.82159344],
       [ 0.59264235,  1.29869894, -1.18870241,  0.11590888, -0.09181687],
       [-0.96924265, -1.62888685, -2.05787102, -0.29705576,  0.68915542]])
```

Slika 19: Primjer NumPy naredbe

Izvor: Izrada autora

Pandas

Pandas je Python biblioteka koja se koristi za rad sa skupovima podataka. Ima funkcije za analizu, čišćenje, istraživanje i manipuliranje podacima pa je stoga učinkovit u tome da neuredne skupove podataka učini čitljivima i relevantnima. Pandas također omogućuje analizu velikih podataka i donošenje zaključaka na temelju statističkih teorija. Pandas je najkorišteniji alat u cijelom praktičnom dijelu jer čišćenje i manipulacija podacima oduzima otprilike 80% vremena na prosječnom zadatku podatkovnog znanstvenika. Ovaj alat omogućuje vrlo jednostavan pregled podataka u obliku tablica odnosno "okvira", a svojim naredbama mogu se vrlo lako iščitati vrijednosti podataka kao što je aritmetička sredina, minimalna i maksimalna vrijednost, standardna devijacija i slično.

```
df[(df['W']>0) | (df['Y']>1)]
```

	W	X	Y	Z
A	2.706850	0.628133	0.907969	0.503826
B	0.651118	-0.319318	-0.848077	0.605965
D	0.188695	-0.758872	-0.933237	0.955057
E	0.190794	1.978757	2.605967	0.683509

Slika 20: Primjer Pandas naredbe

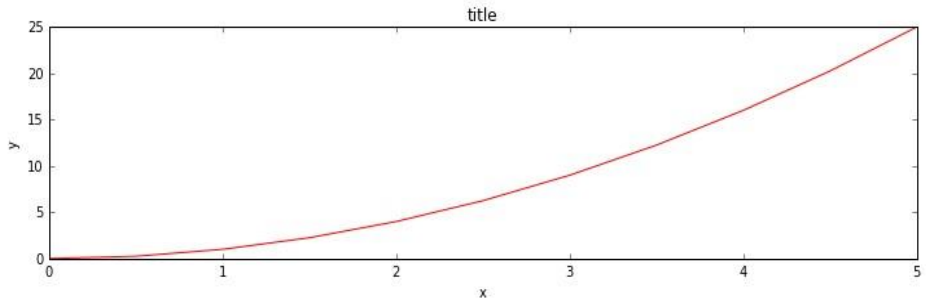
Izvor: Izrada autora

Matplotlib

Matplotlib je sveobuhvatna biblioteka za stvaranje statičkih, animiranih i interaktivnih vizualizacija u Pythonu. Omogućava jednostavno stvaranje grafova s nekoliko linija koda, a koristiti će se upravo pri vizualizaciji podataka tijekom analize.

```
fig, axes = plt.subplots(figsize=(12,3))

axes.plot(x, y, 'r')
axes.set_xlabel('x')
axes.set_ylabel('y')
axes.set_title('title');
```

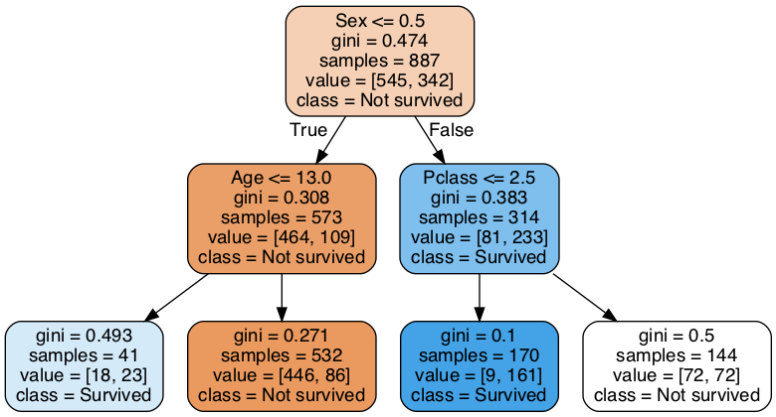


Slika 21: Primjer Matplotlib naredbe

Izvor: Izrada autora

Scikit-Learn

Scikit-Learn fokusiran je na alate za strojno učenje uključujući matematičke, statističke i algoritme opće namjene koji čine osnovu za mnoge tehnologije strojnog učenja. Najvažnija je biblioteka za stvaranje modela strojnog učenja, a koristiti će se prilikom inženjerstva značajki, podjele podataka na setove za trening i testiranje te na kraju stvaranje i optimizaciju modela strojnog učenja.

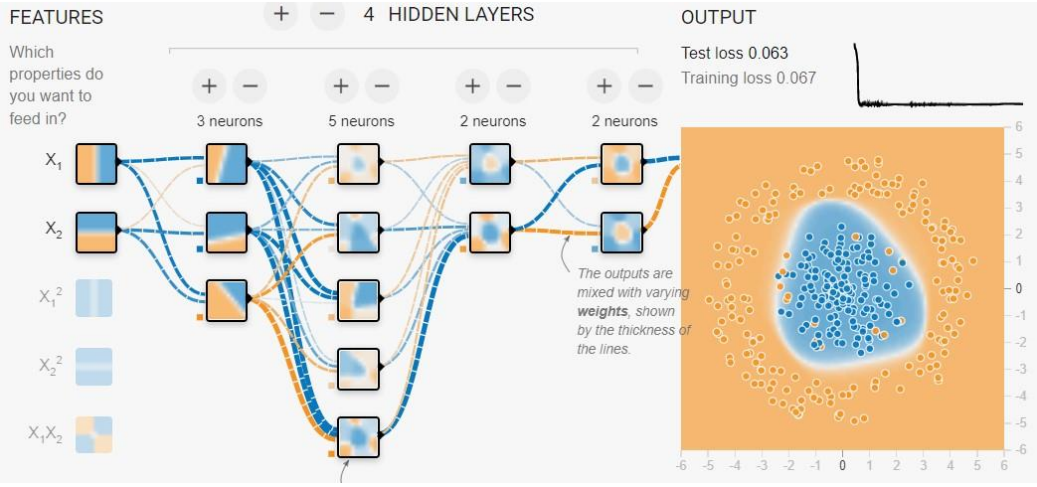


Slika 22: Primjer Scikit-Learn stabla odluke

Izvor: https://towardsdatascience.com/an-introduction-to-decision-trees-with-python-and-scikit-learn-1a5ba6fc204f

TensorFlow

TensorFlow je biblioteka otvorenog koda za numeričko računanje i strojno učenje velikih razmjera, a najviše se bazira na radu s umjetnim neuronskih mrežama. Napravljena je od strane Google-a za vlastitu primjenu no s vremenom je puštena u javnost za slobodnu upotrebu, a sve što je potrebno za rad s umjetnim neuronskim mrežama je internetska veza i grafička kartica ili procesorska jedinica s minimalnim zahtjevima.

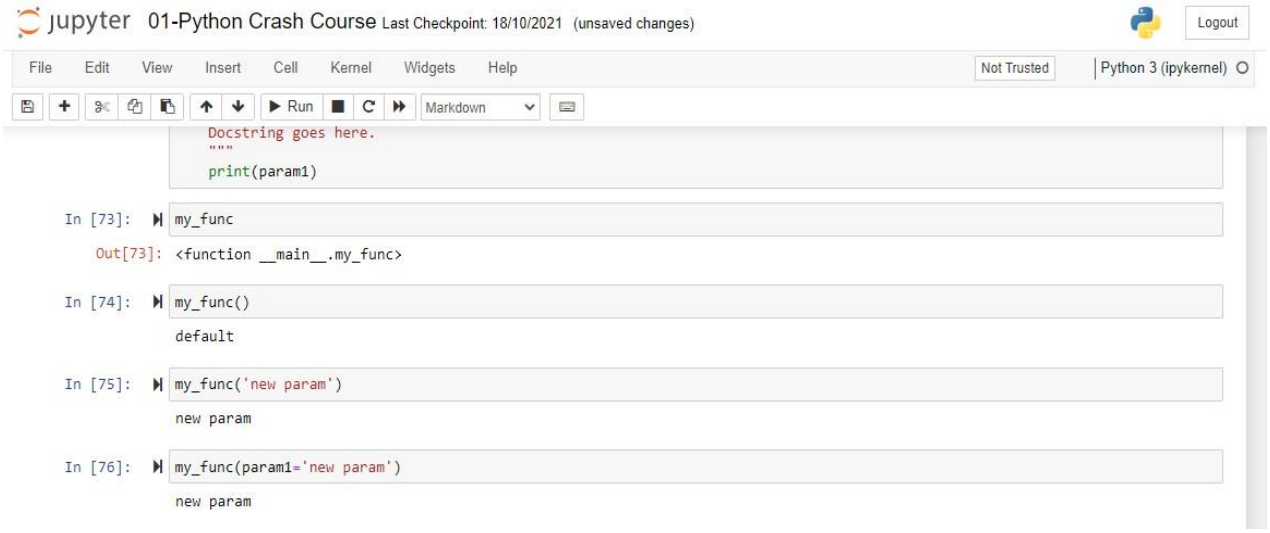


Slika 23: Primjer gradnje umjetne neuronske mreže

Izvor: https://playground.tensorflow.org/

Jupyter Notebook

Jupyter Notebook je integrirano razvojno okruženje, skraćeno IDE (eng. Integrated Development Environment) u kojem se piše kod u Pythonu i pokreću sve radnje. Jupyter Notebook dopušta vrlo pregledan rad s grafovima i tablicama i svaki red koda se može pokretati zasebno kao vlastita jedinica. Zbog svoje jednostavnosti i preglednosti često je omiljen izbor mnogih podatkovnih znanstvenika i analitičara.



Slika 24: Primjer Jupyter Notebook sučelja

Izvor: Izrada autora

4.3. Rad na podacima

Originalni set podataka s kojima će se raditi sadrži informacije studenata nekog proizvoljnog fakulteta, a sve informacije o studentima su nasumično generirane i nisu reprezentativne stvarnim osobama. To se odnosi na informacije o predmetu koji se polaže, broj pokušaja polaganja ispita, datum ispita, ocjene, itd., a studenti su pritom označeni svojim jedinstvenim brojem (ID).

U ovom poglavlju će se opisati proces čišćenja podataka od nepotrebnih simbola i stavki, proces inženjerstva značajki gdje se manipulira podacima kako bi se ih se prilagodilo potrebama problema

i proces analize podataka kojima će se nastojati objasniti veza između varijabli i doći do korisnih spoznaja koji mogu pomoći prilikom gradnje modela.

Originalni set podataka se ukupno sastoji od 7 stupaca i 119561 redaka (7x119561), a nazivi stupaca su sljedeći:

1. Smjer - predstavlja usmjerenje studenta
2. Status - informacija je li student redovan ili izvanredan (1 - redovan, 2 - plaćaju, 3 - izvanredan)
3. ID - identifikacijski broj studenta
4. Sifra_pred - šifra predmeta koji student polaže
5. Datum_ispita - datum ispita kada student polaže
6. Ocjena_ispita - konačna ocjena ispita
7. Ispitni_mjesec - stupac dobiven odvajanjem mjeseca od datuma ispita, na taj način se može samo dobiti uvid u ispitni rok kada se ispit polaže

Radi se o popriličnom velikom skupu podataka, ali zbog toga što je u podacima pohranjen ispitni rezultat svakog pokušaja polaganja svakog studenta na tom fakultetu. Dakle, skup podataka nije sačinjen od toliko velikog broja jedinstvenih studenata, već njegov svaki pokušaj polaganja ispita. Od 119561 redaka u skupu podataka sadržano je ukupno 1924 jedinstvenih studenata te je svaki od tih studenata uspješno diplomirao. S obzirom na opis problema, cilj je predvidjeti konačni prosjek studenata na temelju njihovih ocjena s prve godine studija, te za svrhu ovog rada nisu relevantni studenti koji nisu završili studij. Važno je napomenuti kako ovaj skup podataka ne sadrži još informacije o prosjeku studenata na trećoj godini već će se taj podatak trebati ručno izvući manipulacijom podataka, što će biti opisano u nastavku.

4.3.1. Čišćenje podataka

Prije svega ostaloga podatke je potrebno očistiti od nepotrebnih riječi, simbola ili slova kako bi se ti podaci mogli svesti na jednostavniju i sveobuhvatniju razinu. Proces čišćenja podataka sastoji se od: uklanjanja dvostrukih i nevažnih stavki, ispravljanja strukturalnih pogrešaka u riječima,

uklanjanja neželjenih ekstremnih vrijednosti (outlier-e) i rješavanja praznih stavki. Čišćenje koje je provedeno u ovom setu podataka je:

1. Uklonjeni su simboli "/", "\", "," iz svih varijabli
2. Uklonjeni su razmaci i prazni prostori prije i na kraju teksta iz svih varijabli
3. U stupcu "smjer" prazne vrijednosti su zamijenjene s "Nema usmjerenje"
4. U stupcu "datum_ispita" i "ocjena_ispita" redovi s prazim vrijednostima su potpuno uklonjeni iz podataka
5. U stupcu "ocjena_ispita" redovi s vrijednosti "P" nisu imale značaja pa su potpuno uklonjeni iz podataka

Procesom čišćenja podataka uklonjeno je mnogo izuzetaka u podacima koji su odskakali od većine vrijednosti ili stavki u podacima koji mogu smetati prilikom pravilne analize podataka i gradnje modela. Sljedeći korak je inženjering značajki gdje se od originalnog skupa podataka trebaju transformirati značajke i vrijednosti na način koji će pristajati konačnom modelu.

4.3.2. Inženjering značajki

Inženjerstvo značajki je proces pretvaranja podataka u značajke koje bolje predstavljaju temeljni problem za prediktivne modele, što rezultira poboljšanom preciznošću modela.[19] Dakle, cilj inženjerstva značajki je oblikovati podatke na način koji će bolje odgovarati za problem koji se nastoji riješiti. Citat iz istog izvora to točno opisuje na sljedeći način: "*inženjering značajki je ručno dizajniranje što bi trebali biti ulazni x-ovi*".[20]

Pošto je cilj ovoga rada predvidjeti uspješnost studenata na kraju treće godina studija na temelju ocjena predmeta s prve godine studija potrebno je transformirati originalni skup podataka jer ne sadrži informacije o zaključnim ocjenama predmeta ni zaključnom prosjeku na kraju treće godine. Stoga se pomoću biblioteke Pandas dozivaju naredbe za pivot tablice gdje se u retke pivot tablice

[19] https://machinelearningmastery.com/discover-feature-engineering-how-to-engineer-features-and-how-to-get-good-at-it/
[20] Ibid.

ubacuje stavka “ID”, u stupce se ubacuje stavka “sifra_pred”, a u vrijednosti se ubacuje “ocjena_ispita” s agregatnom funkcijom koja računa aritmetičku sredinu ocjene. Kao rezultat dobiti će se tablica koja prikazuje informacije o svakom studentu i njegovom rezultatu na ispitu za svaki predmet koji je polagao. Prema tome se može lako izračunati njegov konačni prosjek na kraju studija.

Novi skup podataka sastoji se od ukupno 10 stupaca i 1924 retka. Nazivi stupaca predstavljaju 9 glavnih predmeta s prve godine studija nekog proizvoljnog fakulteta, a oni često predstavljaju najveći izazov studenata za položiti i mogu biti dobar indikator uspješnosti na kraju studija. Ovi stupci ujedno predstavljaju **ulazne varijable** modela strojnog učenja. Zadnji stupac predstavlja konačni prosjek studenata na kraju treće godine studija, a dobio se prosjekom ocjena svih predmeta tijekom te tri godine studija. Ovaj stupac će biti **izlazna**, odnosno **ciljna varijabla** koja će se na kraju nastojati predvidjeti. Tablica 1 prikazuje prethodno opisan skup s predmetima i konačnim prosjekom.

sifra_pred ID	EUA001	EUA003	EUA007	EUA008	EUA102	EUA107	EUA004	EUA002	EUA009	konacni_prosjek
989335155	3.0	2.0	2.0	2.0	2.0	2.0	3.0	2.0	3.0	3.028571
994385128	3.0	2.0	3.0	3.0	3.0	3.0	2.0	2.0	5.0	3.294118
997385013	3.0	4.0	4.0	4.0	5.0	2.0	3.0	3.0	5.0	3.852941
985383313	0.0	2.0	0.0	0.0	3.0	4.0	4.0	3.0	0.0	3.454545
987385003	5.0	3.0	3.0	2.0	2.0	2.0	3.0	3.0	2.0	3.085714
...	...	...	...	...	...	...	...	...	...	...

Tablica 1: Skup podataka s ocjenama predmeta prve godine studija i konačnog prosjeka

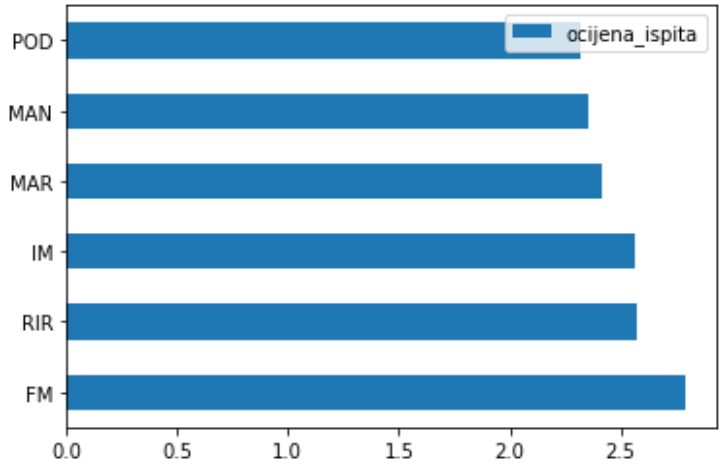
Izvor: Izrada autora

Uspješnim inženjeringom značajki stvoren je skup podataka koji dobro opisuje problem koji se nastoji riješiti, a sljedeći korak je analiza oba skupa podataka.

4.3.3. Analiza podataka

Eksploratorna analiza podataka odnosi se na kritičan proces izvođenja istraživanja podataka kako bi se otkrili obrasci, uočile anomalije i provjerile pretpostavke uz pomoć sažete statistike i grafičkih prikaza.[21] Cilj ove analize podataka je uputiti na određene obrasce u ponašanju studenata tijekom ispita koji mogu sugerirati bolji uspjeh tijekom studija. Također, na temelju tih podataka mogu se stvoriti pretpostavke o konačnom uspjehu studenta već prije nego osoba završi studij.

Uspoređujući smjerove studenata koji pohađaju i njihov konačni prosjek može se primjetiti da su u studenti smjer FM imali veći prosjek od ostalih smjerova koji je iznosio 2.79, dok su najniži prosjek imali studenti smjera POD. Kroz to se može naslutiti da će osoba imati bolju uspješnost na kraju treće godina studija ukoliko je pohađao smjer FM, a lošiju uspješnost ako je pohađao smjer POD. To može biti zbog raznih pretpostavki poput lakšeg programa, manje zahtjevnosti profesora u ocjenjivanju ili bolje obrazovne pozadine studenata tog smjera.

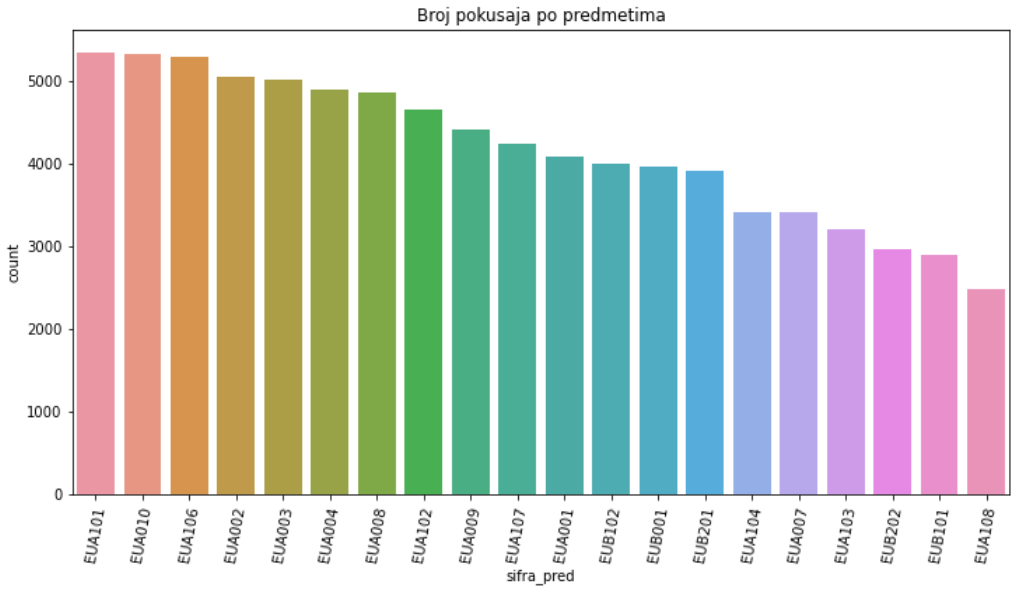


Graf 2: Prosječne ocjene studenata prema smjeru

Izvor: Izrada autora

[21] https://towardsdatascience.com/exploratory-data-analysis-8fc1cb20fd15

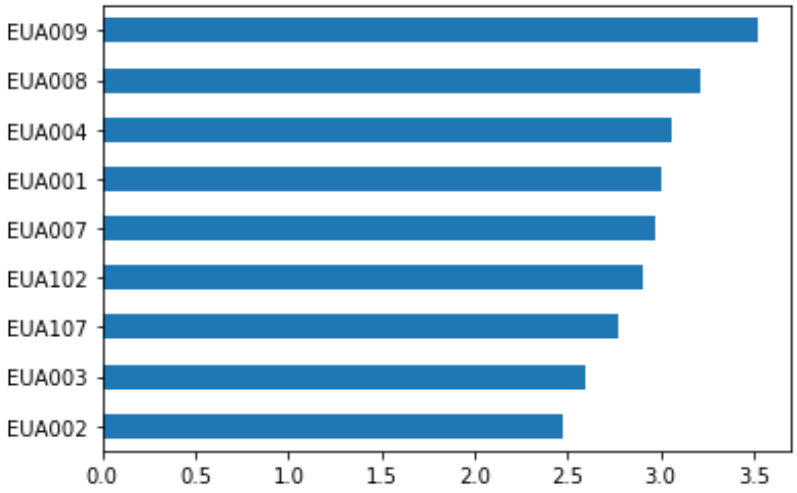
Graf 3 prikazuje analizu izlaznosti na ispite prema predmetima. Prema grafu, najveći broj pokušaja polaganja, odnosno najviše izlazaka na ispitni rok bilo je za predmet “EUA101”, “EUA010” i “EUA106”. Iz toga se može zaključiti da su to predmeti koji su studentima bili najizazovniji za položiti jer im je trebalo najviše ukupnih pokušaja. Fokusirajući se na te predmete tijekom studija može studentu povećati bolju uspješnost na kraju.



Graf 3: Ukupni broj pokušaja polaganja prema predmetima

Izvor: Izrada autora

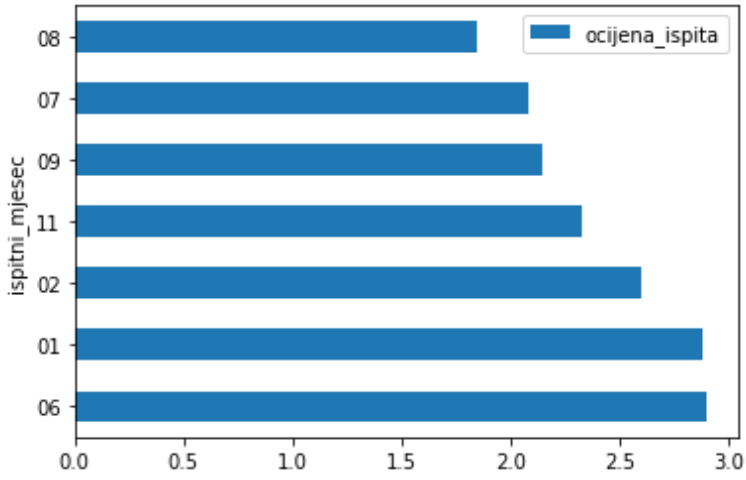
Analizirajući prosječne ocjene predmeta s prve godine studija uočava se uzorak gdje najmanji prosjek imaju isti predmeti koji su imali najveću stopu ukupne izlaznosti što može ponovno uputiti na to da su ti predmeti bili studentima najizazovniji za položiti. Ostvarivanje veće ocjene iz tih predmeta može stoga sugerirati iznadprosječnu konačnu uspješnost studenta na kraju studija.



Graf 4: Prosječne ocjene studenata prema predmetima

Izvor: Izrada autora

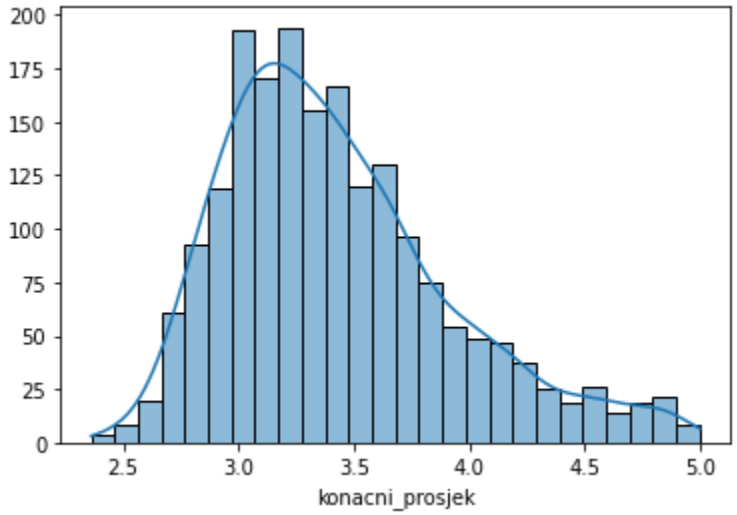
Proučavajući vremensku komponentu u analizi, studenti su imali najbolji prosjek u prvom i šestom mjesecu, odnosno na početku zimskog i ljetnog ispitnog roka, a nešto lošiji prosjek na jesenskom roku. Ovo može uputiti na zaključak da veliki broj studenata bolje prolazi na početku zimskom i ljetnog roka nego studenti koji ostave predmet za kasniji jesenski rok. Zbog toga se može i pretpostaviti da studenti koji ranije riješe predmete imaju bolji konačni uspjeh na studiju.



Graf 5: Prosječne ocjene ispita prema mjesecu polaganja ispita

Izvor: Izrada autora

Iz distribucije vrijednosti konačnog prosjeka studenata se može uočiti da se većina studenata nalazi između prosjeka 2.8 i 3.7, odnosno između ocjene 3 i 4. Može se također zaključiti da je distribucija pretežito podjenako raspodijeljena s blagim nagibom ulijevo, što upućuje na postojanje ekstremnih vrijednosti (eng. outliers) prema većem prosjeku. Tako nešto je važno uočiti jer ekstremne vrijednosti mogu negativno utjecati na performanse modela te je zbog toga dobro proučiti distribucije vrijednosti značajki.



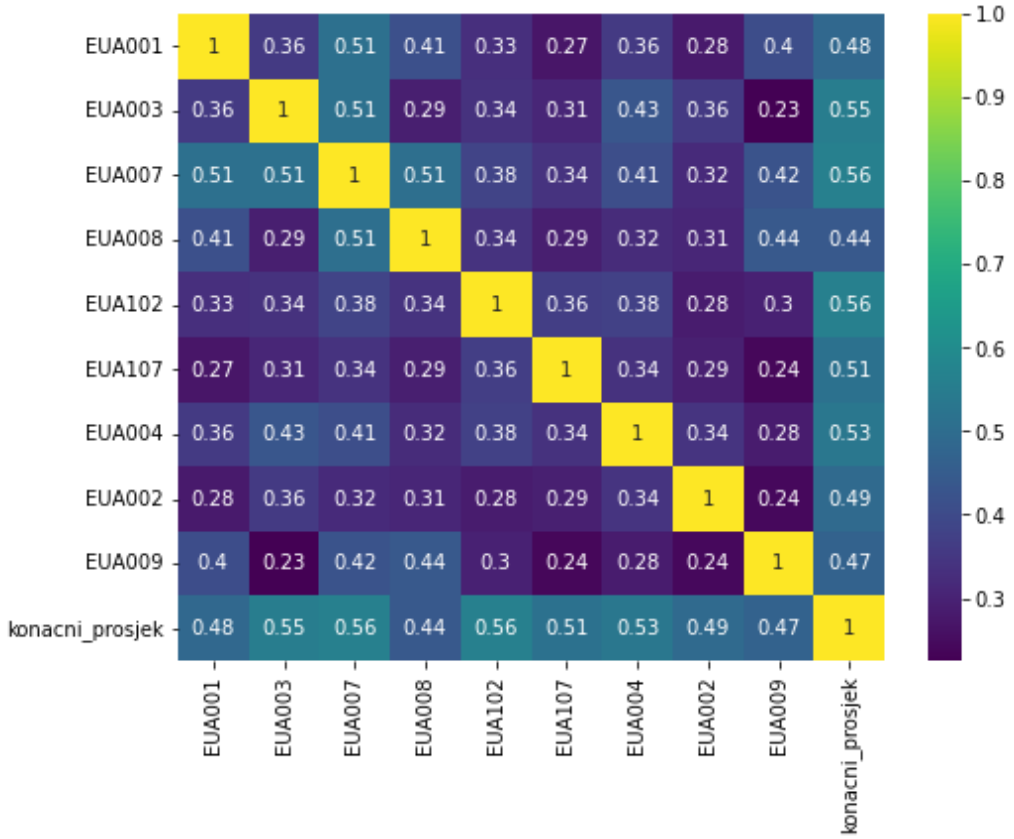
Graf 6: Distribucija vrijednosti konačnog prosjeka

Izvor: Izrada autora

Još jedan važni korak prije stvaranja modela strojnog učenja je proučiti odnose između varijabli. Najbolji način za to napraviti je preko toplinske karte koja jasno prikazuje korelacije među vrijednostima gdje svjetlija boja ukazuje na jaku korelaciju, a tamnija boja 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.

Koeficijent korelacije između konačnog prosjeka i predmeta "EUA102" iznosi 0,56 i ukazuje na jaku pozitivnu korelaciju između te dvije vrijednosti, što znači da će povećanje ocjene iz predmeta

"EUA102" jako utjecati na povećanje konačnog prosjeka. Osim korelacije s konačnim prosjekom mogu se uočiti korelacije između predmeta, pa tako između predmeta "EUA003" i "EUA007" postoji također jaka pozitivna korelacija što upućuje da pozitivna ocjena iz jednog predmeta značajno utječe na pozitivnu ocjenu iz drugog predmeta.



Graf 7: Toplinska karta korelacije predmeta

Izvor: Izrada autora

Ukoliko se pokaže da postoji vrlo visoka korelacija 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. Međutim, u ovom slučaju korelacije nisu previsoke i sve vrijednosti se prihvaćaju.

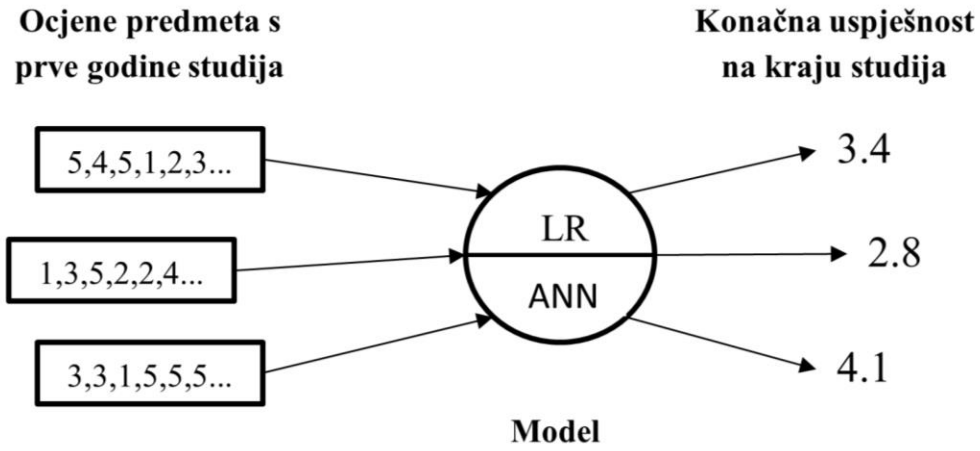
Izvršena je analiza podataka kojom se dobio uvid u podatke s kojima će se raditi u koji će se naknadno uvrstiti u model. Pomoću te analize može se pretpostaviti da:

- Student može imati veću konačnu uspješnost ako studira na smjeru FM
- Student može ostvariti bolji konačni uspjeh ako prioritizira najteže predmete “EUA002”, “EUA003” i “EUA101”
- Student može ostvariti bolji konačni uspjeh ako položi predmete u ranijim ispitnim rokovima nego kasnijim
- Ostvareni uspjeh na predmetima s prve godine studija značajno utječe na konačni prosjek na kraju treće godine

Analizom podataka dobio se uvid u obrasce koji vode do bolje uspješnosti studenata na kraju treće godine. Sljedeći korak je pokretanje modela linearne regresije i umjetnih neuronskih mreža na prethodno stvorenom skupu podataka i analiza rezultata oba modela.

4.4. Testiranje modela

Za stvaranje modela potrebno je imati ulazne i izlazne varijable. U ovom slučaju ulazne varijable su ocjene predmeta s prve godine studija, odnosno svi stupci osim posljednjega koji predstavlja izlaznu varijablu. Ulazne varijable će se grupirati u jednu vrijednost X, dok će izlazna varijabla (ciljna) biti vrijednost y. U svrhu ovog rada provoditi će se nadzirano učenje gdje će postojati podaci za trening na kojem će model učiti i podaci za testiranje na kojem će model izvoditi predikcije.



Slika 25: Shema modela strojnog učenja s ulaznim i izlaznim varijablama

Izvor: Izrada autora

Sljedeći važan korak je takozvani "train test split", odnosno razdvajanje navedenih varijabli na skup podataka za test i skup podataka za trening. Ovo se radi kada ne postoji novi cijeli skup podataka na kojima se može raditi testiranje pa se stoga ti podaci moraju izmisliti, tj. uzima se dio podatka iz ukupnog skupa podataka i dijeli se na više skupova, točnije na četiri skupa podataka:

- X_train (1346 redova i 9 stupaca)
- X_test (578 redova i 9 stupaca)
- y_train (1346 redova)
- y_test (578 redova)

Skupovi podataka y_train i y_test ne sadrže dodatne stupce jer su sačinjeni od samo vrijednosti konačnog prosjeka.

Ukupno će se provesti istraživanje na četiri modela, jedan model linearne regresije i tri modela umjetnih neuronskih mreža s različitim postavkama hiper parametara. Pomoću modela linearne regresije će se postaviti uzorak s kojim će se moći uspoređivati uspješnost svih drugih modela i time će se nastojati ukazati na prednosti umjetnih neuronskih mreža naspram modela linearne regresije. Također, linearna regresije dopušta mnogo bolju interpretaciju podataka nego što to rade umjetne neuronske mreže. Zapravo je jedna od nedostataka neuronskih mreža upravo manjak

mogućnosti interpretacije modela.

Linearna regresija

Stvaranje i pokretanje modela linearne regresije zahtjeva mnogo manje koraka nego modeli umjetnih neuronskih mreža. Sve što je potrebno je pozvati naredbu čime se stvara nova instanca linearne regresije te podatke uvrstiti u model. Nakon što su se skupovi podataka za treniranje uvrstili u instancu za linearnu regresiju, model se pokreće i generiraju se parametri. Kao rezultat modela konstantni član $\beta 0$ iznosi 1.5726, odnosno ako se interpretira, pokazuje da će konačna uspješnost iznositi 1.5726 kada sve nezavisne vrijednosti iznose 0. Interpretacija nema bas smisla jer bi značilo da se može završiti studij, a da se ne položi nijedan predmet s prve godine, međutim u originalnoj kalkulaciji konačnog ishoda su se uzeli svi predmeti s viših godina studija pa je zbog toga i rezultat takav.

Koeficijenti nagiba su sljedeći:

Predmeti	Koeficijenti
EUA001	0.028120
EUA003	0.115426
EUA007	0.052090
EUA008	0.006001
EUA102	0.114507
EUA107	0.100692
EUA004	0.075259
EUA002	0.089570
EUA009	0.064882

Tablica 2: Koeficijenti nagiba modela linearne regresije

Izvor: Izrada autora

Prema tablici se vidi da predmeti “EUA003”, “EUA102” i “EUA107” imaju najveće koeficijente utjecaja na konačni ishod, odnosno povećanje ocjene iz tih predmeta uzrokovat će povećanje konačnog ishoda u prosjeku za **0.11**. Stoga ti predmeti mogu biti i najveći pokazatelj uspjeha studenta na prvim godinama studija.

U ovom radu će se kao funkcija troška koristiti srednja apsolutna pogreška (MAE) i srednja kvadratna pogreška (MSE), a one će ujedno biti glavno sredstvo usporedbe i mjerilo uspješnosti svih modela. Srednja apsolutna pogreška je dobra kada postoje značajni ekstremi (outlieri) i mnogo je jednostavnija za interpretaciju, dok je srednja kvadratna pogreška dobra kada se model suočava s manjim odstupanjima koje je lakše uočiti s ovom funkcijom.

Umjetne neuronske mreže

Pokretanje umjetnih neuronskih mreža u principu funkcionira kao i s linearnom regresijom, međutim karakteristika rada umjetnih neuronskih mreža zahtijeva nekoliko koraka prije pokretanja modela. Kod linearne regresije postojala su dva skupa podataka, jedan za treniranje i jedan za testiranje, dok će se kod neuronskih mreža uvesti još jedan set podataka za validaciju. Skup podataka za validaciju služi samo da kontrolira rad umjetnih neuronskih mreža u predikciji i da pruži nepristrano “mišljenje” o konačnom rezultatu, odnosno nakon što model završi s predikcijama model zatim kontrolira rezultat podacima za validaciju.

Cijeli skup podataka se zatim ponovno dijeli u sljedećem omjeru:

- X_train (865 redova i 9 stupaca)
- X_test (481 red i 9 stupaca)
- X_valid (578 redova i 9 stupaca)
- y_train (865 redova)
- y_test (481 red)
- y_valid (578 redova)

Sljedeći korak se naziva skaliranje (eng. scaling) i on se radi zato što su umjetne neuronske mreže vrlo osjetljive na vrijednosti unutar prostora tj. na udaljenosti između vrijednosti, a u poglavlju o aktivacijskim funkcijama je spomenuto da funkcije uzimaju vrijednosti, najčešće, u intervalu

između 0 i 1 ili -1 i 1.

Mnogi algoritmi strojnog učenja rade bolje kada su ulazne varijable skalirane na jedinstveni raspon. To uključuje algoritme koji koriste ponderirani zbroj ulaza, poput umjetnih neuronskih mreža i algoritme koji koriste mjere udaljenosti, poput k-najbližih susjeda[22].

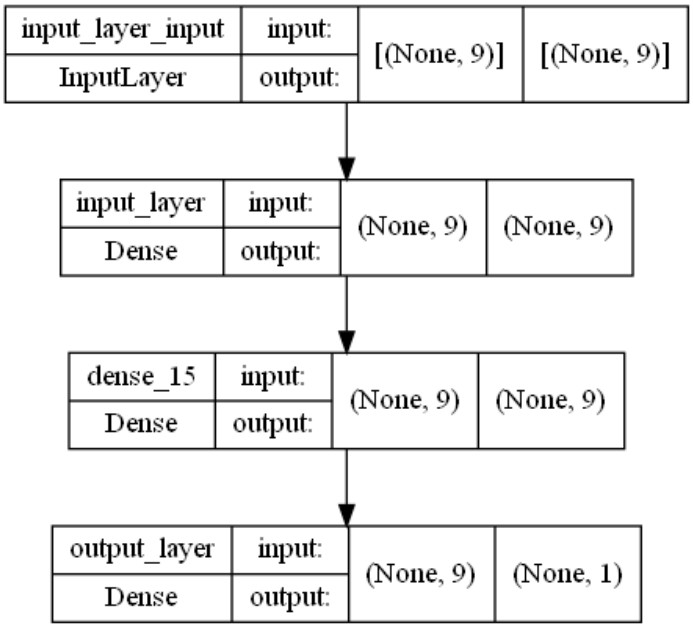
Dvije najpopularnije tehnike za skaliranje numeričkih podataka prije modeliranja su **normalizacija** i **standardizacija**. Normalizacija skalira svaku ulaznu varijablu zasebno na raspon 0-1. Standardizacija skalira svaku ulaznu varijablu zasebno oduzimanjem srednje vrijednosti i dijeljenjem sa standardnom devijacijom kako bi se distribucija pomaknula tako da ima srednju vrijednost od nula i standardnu devijaciju od jedan.

Kod umjetnih neuronskih mreža uvijek se preporuča normalizacija podataka, a biblioteka scikit-learn ima posebnu naredbu za takav zadatak koja se naziva MinMaxScaler. Stoga je potrebno ulazne podatke provući kroz tu naredbu prije stavljanja u model.

Idući korak je izgradnja samog modela umjetne neuronske mreže, a pritom je potrebno dodati slojeve i neurone u neuronsku mrežu. Također se određuje funkcija troška i aktivacija funkcija. Svaki model će sadržavati ulazni sloj od 9 neurona i izlazni sloj od samo jednog neurona.

U prvom modelu koristiti će se višeslojna neuronska mreža sa jednim skrivenim slojem koji se sastoji od 9 neurona. Na slikama su ulazni slojevi nazvani “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 su prikazane 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. Izlaz prethodnog sloja predstavlja također ulaz sljedećeg sloja.

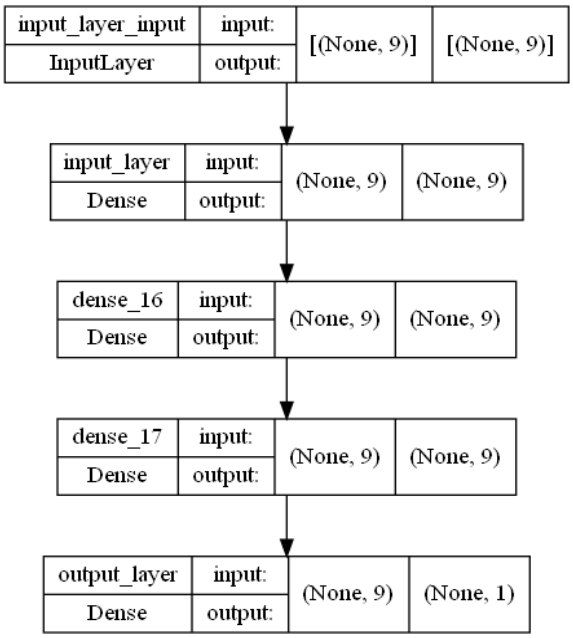
[22] shorturl.at/cdEQU



Slika 26: Umjetna neuronska mreža s jednim skrivenim slojem

Izvor: Izrada autora

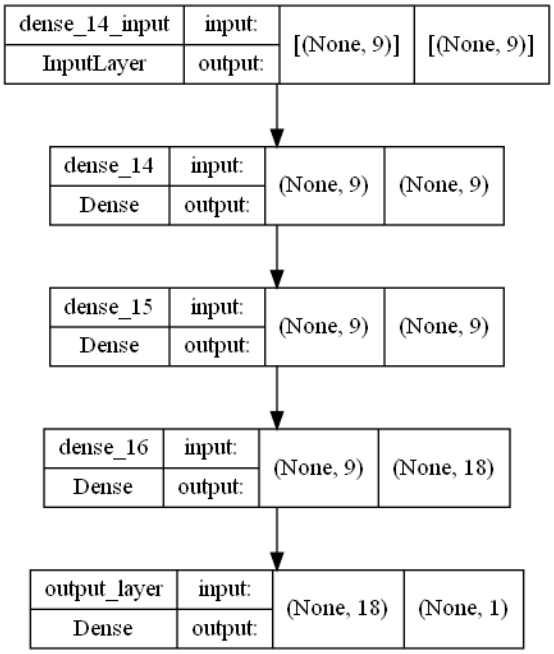
U drugom modelu koristiti će se višeslojna mreža s dva skrivena sloja od 9 neurona.



Slika 27: Umjetna neuronska mreža s dva skrivena sloja po 9 neurona

Izvor: Izrada autora

U trećem modelu će se koristiti višeslojna mreža s dva skrivena sloja od 9 i 18 neurona.



Slika 28: Umjetna neuronska mreža s dva skrivena sloja po 9 i 18 neurona

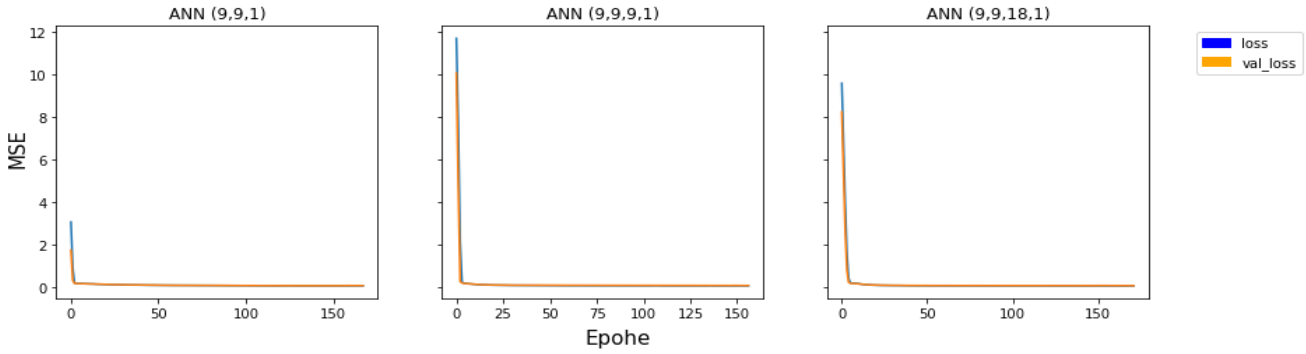
Izvor: Izrada autora

Svakom modelu se postepeno povećava broj slojeva ili neurona kako bi se povećala kompleksnost modela. Arhitektura neuronskih mreža je rađena prema Geron (2019) gdje ulazni sloj treba sadržavati onoliko neurona koliko postoji značajki u ulaznim podacima (9 u ovom slučaju), skriveni sloj sadrže proizvoljno mnogo neurona, a izlazni sloj sadrži 1 neuron prema dimenziji (1 mogući rezultat u ovom slučaju).

Funkcija troška svih modela biti će srednja kvadratna pogreška (MSE) jer će ona najbolje ukazati na fluktuacije u grešci dok model bude učio i srednja apsolutna pogreška (MAE) koja je dobra za interpretaciju rezultata. Aktivacijska funkcija svih modela je funkcija ReLU s obzirom da će izlaz uvijek biti pozitivan broj (konačni prosjek studenta) te je to najpopularnija i najčešće korištena aktivacijska funkcija za razni broj problema. Optimizacijski algoritam koji će biti korišten je tzv. "Adam" optimizator, koji se preporuča kod većine modela umjetnih neuronskih mreža zbog toga što je jednostavan za implementaciju, ima brže vrijeme izvođenja, male zahtjeve za memorijom i

zahtijeva manje podešavanja nego bilo koji drugi algoritam optimizacije[23]. Adam optimizator je nadograđena verzija gradijentnog spusta koji je opisan u prethodnom poglavlju.

Važnost validacijskog skupa podataka doći će do izražaja prilikom pokretanja neuronskih mreža zbog čestog problema kod strojnog učenja koji se naziva **prenaučenost** (eng. overfitting). To je problem koji se odnosi na model koji predobro modelira podatke za trening i pritom generira uzorke u podacima koji možda ne postoje zapravo ili nisu relevantni. Zbog toga model može loše raditi na novim podacima. Problem prenaučenosti (eng. overfitting) kod umjetnih neuronskih mreža može nastati zbog pre kompleksnih i nepotrebno dubokih mreža, a regulira se validacijskim setom.



Graf 8: Krivulje troška umjetnih neuronskih mreža

Izvor: Izrada autora

Graf 8 prikazuje krivulje troška za sva tri model umjetnih neuronskih mreža. Na grafovima se nalaze dvije krivulje, jedna plave boje koja predstavlja grešku modela prilikom testiranja i jedna narančaste 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 narančasta krivulja potpuno preklapa plavu krivulju u ovom slučaju znači da model funkcionira vrlo dobro i da ne postoji problem prenaučenosti.

[23] https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-deep-learning-optimizers/

4.5. Analiza rezultata

Kako bi se dobili najbolji rezultati ukupno su napravljena četiri modela strojnog učenja, jedan model linearne regresije i tri modela umjetnih neuronskih mreža s različitim hiper parametrima. Mjerilo po kojoj će se izvršiti procjena uspješnosti će biti prema srednjoj apsolutnoj pogrešci (MAE) i srednjoj kvadratnoj pogrešci (MSE) gdje je cilj u oba slučaja minimizirati navedeni trošak. Također će se prokomentirati rezultat i uspješnost pojedinog modela u stvaranju predikcije na skupu podataka.

Prema tablici 3 vidimo da je model linearne regresije imao MAE od 0.221440 i MSE od 0.081640, što bi značilo da je u prosjeku ovaj model griješio za 0.221440 konačnog prosjeka. Ovaj rezultat je jako zadovoljavajuć sam od sebe, ali svejedno nije bolje od najlošijeg modela umjetnih neuronskih mreža s jednim skrivenim slojem od 9 neurona, koji je ostvario MAE od 0.213912 i MSE od 0.069775. Najbolji rezultat je ostvarila neuronska mreža s dva skrivena slojeva po 9 neurona gdje je MAE bio 0.192083 i MSE bio 0.056280, odnosno da je taj model u prosjeku griješio za samo 0.192083. Pritom najveću važnost uzima mjerilo MSE koja je ujedno osjetljivije mjerilo i bolje naglašava razliku između modela.

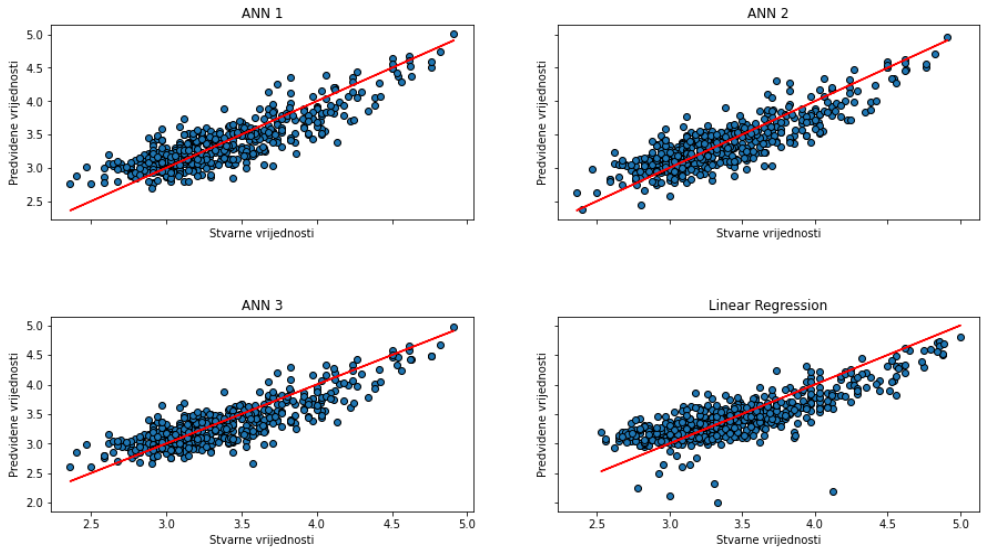
Modeli	MAE	MSE
ANN (9,9,1)	0.213912	0.069775
ANN (9,9,9,1)	**0.192083**	**0.056280**
ANN (9,9,18,1)	0.192563	0.058507
Linearna regresija	0.221440	0.081640

Tablica 3: Srednje apsolutne pogreške i srednje kvadratne pogreške modela

Izvor: Izrada autora

Prema tome, može se zaključiti da je umjetna neuronska mreža s dva skrivena sloja od 9 neurona

najbolje predviđala konačni prosjek studenata na temelju rezultata srednje apsolutne i srednje kvadratne pogreške. Graf 9, na kojoj se mogu vidjeti četiri različita grafa raspršenosti, prikazuje odstupanja pojedinih vrijednosti od regresijske linije te se također može primjetiti mala razlika između pojedinih modela.



Graf 9: Grafovi raspršenosti modela

Izvor: Izrada autora

Prema grafu se može zaključiti da je model linearne regresije vrlo dobro radio na vrijednostima koji su bili u centru distribucije vrijednosti, odnosno model poprilično dobro predviđa vrijednosti između 2.8 i 3.7, ali lošije predviđa veće vrijednosti (prosjek veći od 4.5) nego što su to radile umjetne neuronske mreže, a najbolje je predviđala upravo neuronska mreža s dva skrivena sloja po 9 neurona (ANN 2) što se može vidjeti po malom broju značajnih odstupanja pojedinih vrijednosti (odstupanja ekstrema) od crvene linije.

Unatoč tome što je model linearne regresije napravio vrlo dobar posao predikcije prosjeka studenata i što bi taj model mogao dati zadovoljavajuće rezultate, umjetne neuronske mreže su se svejedno pokazale kao bolji algoritam za stvaranje predikcija konačnog ishoda studenata. Ovo može ukazati na moć umjetnih neuronskih mreža u usporedbi s drugim modelima strojnog učenja.

Treba se također uzeti u obzir mogućnosti dodatnog poboljšanja modela kroz optimizaciju hiper

parametara modela, dubljeg čišćenje podataka, uklanjanje ekstremnih vrijednosti ili povećanja kompleksnosti modela kroz veći skup podataka s više značajki (stupaca). Također je važno ukazati na činjenicu da povećavanje kompleksnosti modela kroz povećanje broja slojeva i neurona unutar slojeva ne pridonosi nužno poboljšanju performansi modela kao što je vidljivo kroz ovaj primjer gdje je neuronska mreža s dva skrivena sloja, ali manjeg broja neurona ostvarila bolje rezultate nego mreža s dva skrivena sloja i većim brojem neurona.

5. ZAKLJUČAK

Umjetne neuronske mreže su još uvijek vruća tema u području podatkovne znanosti i umjetne inteligencije zbog svoje široke primjene, kompleksnosti i prilagodljivosti prema problemima koje rješavaju. Bilo da se radi o traženju malignih dijelova kože na tijelu, autonomne vožnje ili prepoznavanja glasa, umjetne neuronske mreže su tehnologija koja stoji iza toga. Tehnologiju baziranoj na neuronskim mrežama može se pronaći gotovo svugdje danas i u svim granama ljudskog života, ali najmanje u obrazovanju. Unatoč tome što je pandemija virusa COVID-19 ubrzala proces digitalizacije u obrazovanju te iako se tragovi upotrebe umjetnih neuronskih mreža u obrazovanju mogu pronaći u obliku obrade jezika to je još uvijek područje koje je netaknuto tehnološkim rješenjima pri pružanju podrške u odlučivanju.

Analizirajući podatke o studentima, ispitima, profesorima i predmetima umjetne neuronske mreže mogu prepoznati određene uzorke na temelju kojih obrazovne institucije mogu prilagoditi svoju politiku i uvesti bolje promjene u programe. Pomoću modela umjetnih neuronskih mreža moguće je predvidjeti studente koji su pri većem riziku od opadanja i prema tome pravovremeno reagirati ili se mogu uočiti problematični predmeti koji zahtijevaju promjenu u izvođenju nastave. Činjenica je da mnoge grane društva danas koriste moć umjetnih neuronskih mreža i podataka u svoju korist kako bi poboljšali poslovanje pa stoga ne postoji razlog da obrazovne institucije ne poduzmu isti taj korak.

Svrha ovoga rada bio je uputiti na mogućnosti i potencijal umjetnih neuronskih mreža u obrazovanju. U teorijskom dijelu opisana je podloga umjetnih neuronskih mreža i načina na koji one zapravo rade i uče, a kratko se ukazalo na dosadašnju literaturu po pitanju koristi umjetnih neuronskih mreža u obrazovanju. U istraživačkom dijelu napravljen je cijeli postupak čišćenja podataka, inženjeringa značajki i analize skupa podataka studenata jednog proizvoljnog fakulteta, a na samom kraju je na temelju tog skupa napravljena predikcija konačne uspješnosti studenata na kraju treće godine studija na temelju njihovih ocjena s prve godine studija pomoću modela umjetnih neuronskih mreža. Kako bi se naglasila moć umjetnih neuronskih stvoren je i model linearne regresije na istom skupu podataka te se napravila usporedna analiza navedenih modela.

Na temelju provedenog istraživanja nastoji se odgovoriti na istraživačka pitanja postavljena u

samom početku rada.

Jesu li umjetne neuronske mreže prikladni modeli za predviđanje konačnog ishoda studenta na kraju treće godine studija na osnovi ocjena s prve godine studija?

Umjetne neuronske mreže imaju sposobnost rješavanja raznih problema zahvaljujući svojoj snazi obavljanja računalnih operacija i prilagodljivosti problemu. Istraživanje je pokazalo da umjetne neuronske mreže s lakoćom rješavaju problem kao što je predikcija konačnog prosjeka studenata na kraju treće godine na temelju ocjena s prve godine studija. Zapravo, može se i zaključiti da je ovaj problem uvelike prejednostavan za modele neuronskih mreža te da bi bilo poželjno povećati kompleksnost problema uvođenjem većeg skupa podataka s više značajki.

Koje su umjetne neuronske mreže prikladni kao model predikcije konačnog ishoda studija?

Postoji nekoliko vrsta umjetnih neuronskih mreža ovisno o problemu na kojem se radi. Umjetne neuronske mreže mogu biti jednoslojne, višeslojne, konvolucijske ili povratne. Jednoslojne mreže imaju vrlo linearnu karakteristiku te u velikoj većini slučajeva nisu dobri za stvaranje predikcija. Konvolucijske neuronske mreže svoju primjenu nalaze u otkrivanju uzoraka na slikama i videozapisima, dok povratne neuronske mreže služe primarno u analizi teksta i kod prirodne obrade jezika. Zbog toga ni jedna od tih neuronskih mreža nije prikladna za predikciju konačne uspješnosti studija. Istraživanje je stoga pokazalo da su višeslojne neuronske mreže najbolje za takav problem.

Jesu li umjetne neuronske mreže bolji alat u predikciji konačnog ishoda studija od modela linearne regresije?

Istraživanje je pokazalo da su po svim parametrima umjetne neuronske mreže bolji prediktivni alat od modela linearne regresije. Istraživanje je koristilo funkciju srednje apsolutne pogreške (MAE) i srednje kvadratne pogreške (MSE) kao mjerilo uspješnosti gdje je najbolji model umjetne neuronske mreže ostvario MAE od 0.192083 dok je model linearne regresije ostvario MAE od 0.221440. Također, s druge strane isti model neuronske mreže ostvario je MSE od 0.056280 dok je model linearne regresije ostvario MSE od 0.081640. Prema tome se može zaključiti da su

umjetne neuronske mreže bolji prediktivni alat u predikciji konačne uspješnosti studija od linearne regresije.

Koje su to karakteristike umjetnih neuronskih mreža koje ih čine boljim prediktivnim alatom od linearne regresije?

Prednost umjetnih neuronskih mreža nad svim ostalim modelima je njihova prilagodljivost i dubina kompleksnosti izvršenja računalnih operacija. Algoritam gradijentnog spusta skupa s algoritmom povratnog širenja unatrag dopušta umjetnim mrežama da uče i rade korekcije u isto to vrijeme potpuno samostalno.

Mogu li modeli umjetnih neuronskih mreža pomoći kod donošenja odluka obrazovnih institucija?

Na temelju istraživanja, može se zaključiti da su umjetne neuronske mreže u mogućnosti stvoriti predikcije konačne uspješnosti studenata na temelju ocjena s prve godine studija, pa se otvara pitanje daljnjih mogućnosti ovih modela na mnogo složenijim zadacima i problemima. Koristeći se raznim podacima, umjetne neuronske mreže mogu pomoći pri otkrivanju studenata koji su pri većem riziku od opadanja prije nego se to dogodi, otkrivanju nadarenih studenata ili pak uočavanju problematičnih predmeta koji zahtijevaju promjenu u svom programu. Stoga, tehnologija neuronskih mreža može služiti kao sredstvo podrške pri donošenju važnih odluka u obrazovnim institucijama. .

Međutim, unatoč uspješnosti istraživanja u predikciji uspješnosti studenta, treba uputiti u činjenicu da su umjetne neuronske mreže tehnologija koja iziskuje mnogo računalne snage, pogotovo ako su modeli s kojima se radi duboke i kompleksne mreže. Stoga se otvara pitanje potrebe ovako moćnog algoritma za prejednostavne probleme i otvara se mogućnost provođenja dodatnog istraživanja umjetnih neuronskih mreža, ali na temelju većeg i mnogo kompleksnijeg skupa podataka.

LITERATURA

1. Asif, R., Marceron, A., Ali, A. S., Haider, G. N. (2017), Analyzing undergraduate students' performance using educational data mining, Computers & Education 113. Dostupno na: https://www.sciencedirect.com/science/article/abs/pii/S0360131517301124?via%3Dihub [pristupljeno: 20.05.2022.]
2. 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. Dostupno na: https://dl.acm.org/doi/10.1145/3325917.3325919 [pristupljeno: 04.06.2022.]
3. Alpaydin, E. (2021), Strojno učenje: Nova umjetna inteligencija, Mate d.o.o.
4. Burkov, A. (2019), The Hundred-Page Machine Learning Book
5. Brocardo, M. L., Traore, I. (2017), Authorship verification using deep belief network systems, International Journal of Communication Systems 30(12). Dostupno na: https://www.researchgate.net/publication/312503871_Authorship_verification_using_deep_belief_network_systems [pristupljeno: 29.05.2022.]
6. Chaplot, S. D., Rhim, E., Kim, J. (2015), Predicting Student Attrition in MOOCs using Sentiment Analysis and Neural Networks, Seventeenth International Conference on Artificial Intelligence in Education (AIED 2015). Dostupno na: https://www.researchgate.net/publication/277014914_Predicting_Student_Attrition_in_MOOCs_using_Sentiment_Analysis_and_Neural_Networks [pristupljeno: 20.05.2022.]
7. Chen, X., Xie, H., Hwan, G. (2020), A multi-perspective study on Artificial Intelligence in Education: grants, conferences, journals, software tools, institutions, and researchers, Computers and Education: Artificial Intelligence 1. Dostupno na: https://www.sciencedirect.com/science/article/pii/S2666920X20300059 [pristupljeno: 20.05.2022.]
8. Deperlioglu, O., Kose, U. (2011), An educational tool for artificial neural networks, Computers and Electrical Engineering 37 (2011) 392–402. Dostupno na: https://www.sciencedirect.com/science/article/abs/pii/S0045790611000371 [pristupljeno: 20.05.2022.]

9. Ersoz Kaya, I. (2019), Artificial Neural Networks as Decision Support Tool in Curriculum Development, International Journal on Artificial Intelligence Tools Vol. 28, No. 4. Dostupno na: https://www.worldscientific.com/doi/abs/10.1142/S0218213019400049 [pristupljeno: 20.05.2022.]
10. Geron, A. (2019), Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly
11. Grus, J. (2019), Data Science from Scratch, Second Edition, O'Reilly
12. Gallo, C. (2015), Artificial Neural Networks Tutorial, Encyclopedia of Information Science and Technology. Dostupno na: https://www.researchgate.net/publication/261392616_Artificial_Neural_Networks_tutorial [pristupljeno: 21.05.2022.]
13. Haykin, S. (2009), Neural Networks and Learning Machines, Pearson Education Inc.
14. Jadrić, M., Garača, Ž., Ćukušić, M. (2010), Student dropout analysis with application of data mining methods, Management, Vol. 15, 2010, 1, pp. 31-46
15. James, G., Witten, D., Hastie, T., Tibshirani, R. (2021), An Introduction to Statistical Learning with Applications in R
16. Kehinde, J. A., Adeniyi, E. A., Ogundokun, O. R., Gupta, H., Misra, S. (2021), Prediction of Students' performance with Artificial Neural Network using Demographic Traits, Department of Computer Science, Landmark University Omu Aran, Nigeria. Dostupno na: https://arxiv.org/abs/2108.07717 [pristupljeno: 20.05.2022.]
17. Kelleher, J. D. (2021), Duboko učenje, Mate d.o.o.
18. Marion, O., Florence, O., Daramola, O., Roseline, O., Emmanuel, A. (2019), RFID-based human tracking system in tertiary institution, Journal of Engineering and Applied Sciences 14. Dostupno na: http://eprints.lmu.edu.ng/3221/1/RFID-Based%20-2019.pdf [pristupljeno: 29.05.2022.]
19. Marsland, S. (2009), An Algorithmic Perspective, Chapman & Hall/CRC
20. 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.

Dostupno na: https://www.tandfonline.com/doi/full/10.1080/08839514.2021.1922847 [prisupljeno: 20.05.2022.]

21. Omolewa, T. O., Oladele, T. A., Adeyinka, A. A., Oluwaseun, R. O. (2019), Prediction of Student's Academic Performance using k-Means Clustering and Multiple Linear Regressions, Journal of Engineering and Applied Sciences 14. Dostupno na: https://www.researchgate.net/publication/336407561_Prediction_of_Student%27s_Academic_Performance_using_k-Means_Clustering_and_Multiple_Linear_Regressions [pristupljeno: 20.05.2022.]
22. Osmanbegovic, E., Suljic, M. (2012), Data Mining Approach for Predicting Student Performance, Economic Review: Journal of Economics and Business. Dostupno na: http://hdl.handle.net/10419/193806 [pristupljeno: 21.05.2022.]
23. Pavlin-Bernardić, N., Ravić, S., Matić, I. P. (2016), The application of arificial neural networks in predicting children's giftedness, Suvremena psihologija 19 (2016), 1, 49-59. Dostupno na: https://hrcak.srce.hr/176765 [pristupljeno: 05.06.2022.]
24. Pivac, S. (2010), Statističke metode, Ekonomski fakultet u Splitu
25. Shachmurove, Y. (2002), Applying Artificial Neural Networks to Business, Economics and Finance, The City College of the City University of New York and, The University of Pennsylvania. Dostupno na: https://www.researchgate.net/publication/4821288_Applying_Artificial_Neural_Networks_to_Business_Economics_and_Finance [pristupljeno: 20.05.2022.]
26. Susnea, E. (2010), Using Artificial Neural Networks In E-Learning Systems, UPB Scientific Bulletin, Series C: Electrical Engineering. Dostupno na: https://www.researchgate.net/publication/258832960 [pristupljeno: 04.06.2022.]
27. Shahifi, M. A., Husain, W., Rashid A. N. (2015), A Review on Predicting Student's Performance using Data Mining Techniques, Procedia Computer Science 72 (2015) 414 – 422. Dostupno na: https://www.researchgate.net/publication/289991570_A_Review_on_Predicting_Student%27s_Performance_Using_Data_Mining_Techniques [pristupljeno: 20.05.2022.]
28. Shalev-Schwartz, S., Ben-David, S. (2014), Understanding Machine Learning: From Theory to Algorithms, Cambridge University Press

29. Simeunović, V., Preradović, Lj. (2012), Using Data Mining to Predict Success in Studying, Croatian Journal of Education Vol.16; 2/2014 pages: 491-523. Dostupno na: https://www.researchgate.net/publication/297518311_Using_Data_Mining_to_Predict_Success_in_Studying [pristupljeno: 21.05.2022.]
30. Tkalac Verčič, A., Sinčić Ćorić, D., Pološki Vokić, N. (2014), Priručnik za metodologiju istraživanja u društvenim djelatnostima, Ekonomski fakultet u Zagrebu, Zagreb
31. Toth, M., Metzinger, T. C. (2020), Metodologija istraživačkog rada za stručne studije, Veleučilište Velika Gorica. Dostupno na: https://www.vvg.hr/app/uploads/2020/03/METODOLOGIJA-ISTRA%C5%BDIVA%C4%8CKOG-RADA-ZA-STRU%C4%8CNE-STUDIJE.pdf [pristupljeno: 21.05.2020.]
32. Zekić-Sušac, M., Frajman-Jakšić, A. (2009), Neuronske mreže i stabla odlučivanja za predviđanje uspješnosti studiranja, Ekonomski vjesnik: Review of Contemporary Entrepreneurship, Business, and Economic Issues, Vol. XXII No. 2. Dostupno na: https://hrcak.srce.hr/47931 [pristupljeno: 21.05.2022.]
33. Zelenika, R. (2000), Metodologija i tehnologija izrade znanstvenog i stručnog djela, Ekonomski fakultet u Rijeci.

POPIS SLIKA

POPIS TABLICA I GRAFOVA

SAŽETAK

Tehnološke promjene uzrokovane pandemijom COVID-19 dovele su do nagle digitalizacije obrazovanja gdje online nastava i platforme za učenje na daljinu koje postaju dio svakodnevnice. To je uzrokovalo gomilanje velike količine neiskorištenih podataka u bazama podataka obrazovnih institucija i time otvorilo pitanje potencijala rudarenja podataka u korist tih institucija. Pomoću umjetnih neuronskih mreža moguće je otkriti uzorke u podacima i dobiti uvid u ponašanje studenata na temelju kojih obrazovne institucije mogu donijeti bolje odluke. Unatoč tome što se umjetne neuronske mreže već naširoko koriste za rješavanje raznih problema, one su vrlo malo prisutne u obrazovanju. Stoga je cilj ovoga rada istražiti mogućnosti umjetnih neuronskih mreža u obrazovanju provođenjem istraživanja predikcije konačne uspješnosti studenata na kraju treće godine na temelju njihovih ocjena s prve godine studija. Istraživanje je provedeno na ukupno tri modela umjetnih neuronskih mreža, a kako bi se bolje donio zaključak, za usporedbu je korišten i model linearne regresije.

Na temelju mjerila kvadratnog odstupanja vrijednosti, istraživanje je pokazalo da umjetne neuronske mreže mogu vrlo dobro predvidjeti konačnu uspješnost studenta gdje je po rezultatima nadmašilo čak i linearnu regresiju.

Ključne riječi: strojno učenje, duboko učenje, umjetne neuronske mreže, podatkovna znanost

SUMMARY

Technological changes caused by the COVID-19 pandemic have led to a sudden digitization of education, where online classes and distance learning platforms have become part of everyday life. This caused the accumulation of a large amount of unused data in the databases of educational institutions and thus opened the question of the potential of mining the data for the benefit of these institutions. With the use of artificial neural networks, it is possible to discover patterns in data that provide insight into student behavior upon which educational institutions are able to make better decisions. Despite the fact that artificial neural networks are already widely used to solve various problems, they are very little present in education. Therefore, the goal of this work is to investigate the possibilities of artificial neural networks in education by conducting research on the prediction of students' final performance at the end of the third year based on their grades from the first year of study. The research was conducted on a total of three models of artificial neural networks, and in order to reach a better conclusion, a linear regression model was also used for comparison.

Based on the square deviation of the values, the research showed that artificial neural networks can very well predict the final performance of the student, where it even surpassed linear regression according to the results.

Key words: machine learning, deep learning, artificial neural networks, data science