# DUBOKO UČENJE I UMJETNE NEURALNE MREŽE

Deeplizard

<https://ch.mathworks.com/discovery/deep-learning.html>

<https://missinglink.ai/guides/neural-network-concepts/complete-guide-artificial-neural-networks/>

<https://arxiv.org/pdf/1511.08458.pdf>

U ovom poglavlju će se objasniti što je to duboko učenje, čime se bavi i koja motivacija leži iza dubokog učenja. Nakon toga će se objasniti što su to standardne umjetne neuralne mreže i duboke umjetne neuralne mreže kako bi se stvorila podloga za upoznavanje s konvolucijskim neuralnim mrežama koje su glavna tema ovog rada. Nadalje, navest će se nekoliko razlika između jedne i druge vrste neuralnih mreža i gdje se one najbolje koriste. Na kraju će se dati kratki, općeniti pregled rada neuralnih mreža kako bi se objasnilo prosljeđivanje podataka kroz mrežu prema naprijed.

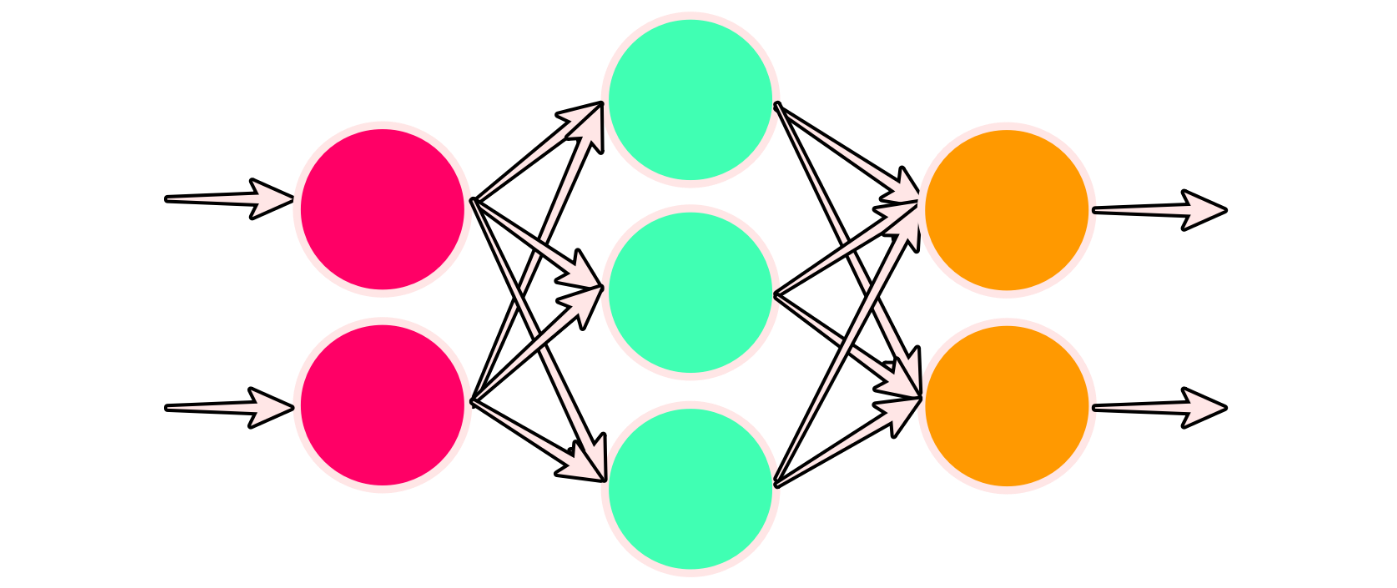
Duboko učenje je polje strojnog učenja koje koristi algoritme ili modele inspirirane strukturom i funkcijama bioloških neuralnih mreža. Razlika između dubokog učenja i strojnog učenja je ta što strojno učenje manualno izvlači relevantne značajke iz slika kako bi ih model naučio ispravno klasificirati. Način rada kod dubokog učenja je da se relevantne značajke iz slika izvlače na automatizirani način.

Modeli, koji se koriste u dubokom učenju, nazivaju se umjetne neuralne mreže (eng. *artificial neural netwroks*, *ANN*) jer dijele nekoliko karakteristika s biološkim neuralnim mrežama. U stranoj literaturi se, uz sami naziv '*artificial neural network*' može naići i na sljedeće nazive:

1. *net*
2. *neural net*
3. *model*.

Pa se tako umjetne neuralne mreže na hrvatskom jeziku mogu nazivati i još:

1. mreža
2. neuralna mreža
3. model.



Slika : shematski prikaz jednostavne umjetne neuralne mreže

Umjetna neuralna mreža je sustav nadziranog učenja inspiriran biološkim neuralnim mrežama kojeg čini veliki broj jednostavnih elemenata koji se nazivaju neuronima, perceptronima ili čvorovima. Neuroni su organizirani u takozvane slojeve (eng. *layers*) i međusobno su povezani vezama. Veze ne povezuju neurone koji se nalaze u istom sloju već međusobno povezuju neurone koji se nalaze u različitim slojevima. Svaka veza između dva neurona transmitira signal od jednog neurona prema drugom. Neuron, koji prima dolazni signal, obrađuje signal kojeg potom šalje neuronima koji se nalaze u sljedećem sloju u mreži.

Općenito, u svakoj neuralnoj mreži, slojevi se mogu podijeliti u tri glavne skupine:

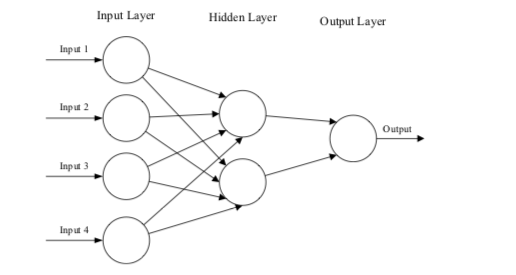
1. ulazni sloj (eng. *input layer*)
2. skriveni sloj (eng. *hidden layer*)
3. izlazni sloj (eng. *output layer*)

Različiti slojevi čine različite transformacije. Tako ulazni sloj na svoje ulaze prima podatke koji se potom propagiraju kroz mrežu. On sam ne izvodi nikakve operacije nad ulaznim podacima. Nadalje, ulazni sloj sadrži po jedan neuron ili čvor za svaku pojedinu komponentu ulaznog podatka koji se prosljeđuje mreži.

Skriveni sloj sadrži proizvoljno odabran broj čvorova. Umjetna neuralna mreža može imati jedan ili više skrivenih slojeva, dok ulaznih i izlaznih slojeva može biti samo po jedan. Nazivaju se skrivenima samo zbog toga što se nalaze između ulaznog i izlaznog sloja.

Izlazni sloj se nalazi na kraju mreže i sadrži po jedan neuron za svaki mogući željeni izlaz što bi ga mreža trebala znati ispravno predvidjeti.

S obzirom da je maloprije spomenuto da ulazni sloj na svoje ulaze prima podatke, ovdje je dobro spomenuti da se ti podaci još mogu nazivati 'ulaznim podacima', 'ulaznim vrijednostima' ili jednostavno 'ulazima'. Nadalje, izlazna vrijednost iz jednog čvora se prosljeđuje na ulaz čvora u sljedećem sloju. To ju onda čini ulaznom vrijednošću u sljedeći čvor. Cilj je svake neuralne mreže ispravno preslikati ulazne podatke, koji se sačinjeni od nekih stvarnih vrijednosti, u ispravne klase predviđanja (eng. *prediction classes*).



Slika : prikaz umjetne neuralne mreže s jednim skrivenim slojem

Osnovna struktura umjetne neuralne mreže može se prikazati pomoću slike. Ulaznom sloju se prosljeđuju podaci u obliku višedimenzionalnog vektora. Ulazni sloj će te podatke proslijediti skrivenom sloju. Skriveni sloj će donositi odluke na temelju ulaznih vrijednosti koje je primio od prethodnog sloja. Neuroni u skrivenom sloju na neki unaprijed određeni način obrađuju podatke te nove izlazne vrijednosti šalju kao ulazne vrijednosti onim čvorovima u sljedećem sloju s kojima su povezani vezama. Izlazni sloj čini konačnu obradu podataka i svoje izlazne vrijednosti pokazuje u obliku vjerojatnosti ili općenite brojčane vrijednosti da određena ulazna vrijednost pripada jednoj od klasa koje su predstavljene neuronima izlaznog sloja.

ZA PRIBACIT SE NA CNN

Ovi biološki inspirirani modeli su sposobni nadmašiti izvođenje starijih modela umjetne inteligencije u uobičajenim zadacima strojnog učenja. Jedan od najimpresivnijih modela dubokog učenja su konvolucijske neuralne mreže (eng. *Convolutional Neural Networks*, *CNN*). Konvolucijske neuralne mreže se obično koriste kako bi riješile teške zadatke prepoznavanja uzoraka na slikama. To čine svojom preciznom, ali jednostavnom arhitekturom.

<https://ch.mathworks.com/discovery/deep-learning.html>

Duboko učenje je područje strojnog učenja koje uči računala činiti ono što ljudi rade po svojoj prirodi: učiti na primjerima. Duboko učenje je ključna tehnologija iza autonomnih vozila. Čini ih sposobnima prepoznati znak stop ili raspoznati pješaka od uličnog svjetla . Duboko učenje je ključna tehnologija i u audio kontroli u uređajima kao što su mobiteli, tableti, televizije i hands-free zvučnici.

Kod dubokog učenja, računalni model uči ispravno klasificirati podatke iz slika, tekstualnih ili zvučnih podataka te ponekad u tome pokazuju bolju izvedbu čak i od ljudi. Modeli su trenirani korištenjem velikog skupa označenih podataka te modeli dubokog učenja, neuralne mreže, sadrže mnogo skrivenih slojeva.

Duboko učenje se prvi put spominje 1980.-ih, ali samo u teoretskom smislu. Dva su razloga zašto je to tako:

1. Duboko učenje zahtijeva velike količine označenih podataka. Na primjer, razvoj autonomnog vozila zahtijeva milijune slika i tisuće sati videa.
2. Duboko učenje zahtijeva znatnu računalnu moć. GPU kartice visokih performansi koje imaju paralelnu arhitekturu koja je efikasna u dubokom učenju. Kada se one udruže s *cluster*-ima računarstva u oblaku (eng. *cloud computing*), znatno se smanjuje vrijeme potrebno za treniranje dubokih neuralnih mreža. Govori se o prvotnom vremenu od nekoliko tjedana do konačnog vremena od nekoliko sati ili manje.

Postoji široki spektar industrija u kojima se primjenjuje duboko učenje. Neke od tih industrija su:

1. industrija autonomnih vozila
2. zračna industrija
3. medicinska industrija
4. elektronika
5. industrijska automatizacija

## NAČIN RADA DUBOKOG UČENJA

Većina metoda dubokog učenja koristi arhitekturu neuralnih mreža. Zbog toga se modeli dubokog učenja nazivaju duboke neuralne mreže (eng. *deep neural networks*).

Pojam 'duboko' se obično odnosi na broj skrivenih slojeva u nerualnoj mreži. Standardne neuralne mreže sadrže 2 ili 3 skrivena sloja dok duvoke neuralne mreže ih mogu imati i 150.

Modeli dubokog učenja su trenirani korištenjem velikog skupa označenih podataka te nauralne mreže uče značajke tih podataka bez potrebe manualnog izvlačenja tih značajki kao što je slučaj kod strojnog učenja.

Jedan od najpopularnijih tipova dubokih neuralnih mreža je poznat pod nazivom 'konvolucijske neuralne mreže' (eng. *convolutional neural networks, CNN, ConvNet*) koje su pokazale izvanrednu izvedbu na 2D podacima kao što su slike. Konvolucijske mreže uče detektirati različite značajke ulaznih podataka korištenjem desetaka ili čak stotina skrivenih slojeva. Svaki skriveni sloj povećava kompleksnost značajki koje model uči. Na primjer, prvi skriveni sloj uči detektirati rubove, dok posljednji skriveni sloj uči detektirati kompleksnije oblike koji su specifični za objekt kojeg model uči raspoznavati.

<https://arxiv.org/pdf/1511.08458.pdf>

Umjetne neuralne mreže su računalni modeli inspirirani strukturom i načinom rada biološke neuralne mreže (kao što je ljudski mozak). Umjetne neuralne mreže su sastoje od velikog broja međusobno povezanih računalnih čvorova (ili neurona). Njihov rad se širi na distributivni način te oni kao kolektiv 'uče' na temelju ulaznih podataka kako optimizirati konačni izlazni rezultat.

Osnovna struktura umjetne neuralne mreže može biti modelirana kako je prikazano na slici. Ulazni podaci se prosljeđuju u mrežu preko ulaznog sloja, obično u obliku nekog više-dimenzionalnog vektora. Ulazni sloj te podatke prosljeđuje skrivenim slojevima. Skriveni slojevi donose neke odluke na temelju podataka koji su im proslijeđeni od prethodnog sloja te odvagati kako stohastička promjena sama po sebi pogoršava ili poboljšava rezultat koji vraća preko izlaznog sloja. Pod ovim se misli da mreža uči.

Duboka mreža je ona koja sadrži više skrivenih slojeva.

Konvolucijske neuralne mreže su primarno bile korištene kako bi se riješili zadaci raspoznavanja uzoraka na slikama. One su analogne tradicionalnim umjetnim neuralnim mrežama na način da se sastoje od neurona koji se optimiziraju kroz proces učenja. I dalje će svaki neuron primiti neki ulaz, izvesti određenu operaciju (na primjer, izračunati neki skalarni produkt koji se potom prosljeđuje nekoj ne-linearnoj funkciji). Na ulazu obično prima neobrađene vektore slike i na izlazu daje neku klasifikacijsku ocjenu. Posljednji sloj će sadržavati funkciju gubitka asociranu s klasama. Sva pravila, koja vrijede kod tradicionalnih umjetnih mreža, vrijede i kod konvolucijskih mreža.

Jedina velika razlika između konvolucijske i tradicionalne umjetne mreže je to što se konvolucijske mreže koriste za raspoznavanje uzoraka u slikama.

Jedno od najvećih ograničenja tradicionalnih umjetnih mreža jest da imaju tendenciju da se muče s računalnom složenošću koju zahtijeva obrada slika. Uobičajeni skupovi podataka strojnog učenja, kao što su MNIST skup podataka u kojem se nalaze skupovi ručno pisanih znamenki veličine tek 28x28, su prikladni za većinu tipova umjetnih mreža. Sa skupom podataka kao što je MNIST, svaki će neuron u prvom skrivenom sloju imati 784 težine (28x28x1 gdje broj 1 označava jedan kanal jer su skupovi MNIST podataka podaci u crno-bijeloj boji) što je izvodljivo za većinu formi umjetnih neuralnih mreža.

<https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-networks-image-classification/>

Algoritmi dubokog učenja za klasifikaciju slika, konvolucijske neuralne mreže, unaprijeđuju mnoge napredne tehnologije i temeljni su subjekt istraživanja u mnogim industrijama od prometa do medicine.

ŠTO JE KLASIFIKACIJA SLIKA?

Klasifikacija slika je proces označavanja slika na temelju predefiniranih kategorija. Taj proces se temelji na nadziranom učenju. Model koji obavlja klasifikaciju slika na svoj ulaz prima slike koje pripadaju specifičnoj kategoriji. Na temelju tog skupa podataka, nakon završenog procesa učenja, model uči klasificirati slike koje pripadaju skupu podataka za testiranje. Model mora moći ispravno predvidjeti kojoj klasi pripadaju te buduće slike te čak i izmjeriti koliko su ta predviđanja bila ispravna.

KORIŠTENJE CNN-OVA ZA KLASIFIKACIJU SLIKA

Pristup korištenja konvolucijski mreža se temelji na pretpostavci da model može ispravno funkcionirati na temelju lokalnog razumijevanja slike. Konvolucijska mreža koristi manje parametara u odnosu na potpuno povezane mreže tako što ponovno iskorištava parametre po nekoliko puta. Dok potpuno povezana neuralna mreža generira težinu za svaki piksel na slici, konvolucijska neuralna mreža generira tek toliko težina kako bi skenirala malo područje na slici u danom trenutku.

Ovaj pristup je koristan za proces treniranja – što mreža ima manje parametara, to će imati bolje izvođenje. Nadalje, kako model zahtijeva manje podataka, brže će učiti.

Još jedna od mogućih primjena konvolucijskih neuralnih mreža su reklame. Na primjer, konvolucijska mreža može jednostavno skenirati korisnikovu stranicu neke društvene mreže, klasificirati slike koje su povezane s modom i tako raspoznati koji je korisnikov modni stil. Na taj način će prodavač moći predlagati relevantnije modne reklame.

Kod dubokih neuralnih mreža, ovaj princip se može iskoristiti za identificiranje lokacija kao što su restorani ili trgovački centri i identificiranje hobija kao što su nogomet ili plesanje.

<https://missinglink.ai/guides/neural-network-concepts/complete-guide-artificial-neural-networks/>

Umjetne neuralne mreže su model nadziranog učenja koje su izgrađene od velikog broja jednostavnih elemenata koji se nazivaju 'čvorovi', 'neuroni' ili 'perceptroni'. Svaki neuron može donijeti neku jednostavnu odluku i proslijediti ju drugim neuronima. Neuroni su organizirani u međusobno povezane slojeve. Neuralna mreža na taj način može oponašati skoro bilo koju funkciju i odgovori na praktički bilo koje pitanje nakon što ju se istrenira dovoljnim brojem uzoraka za treniranje i dovoljnom računalnom moći. 'Plitka' neuralna mreža sadrži samo tri sloja:

1. ulazni sloj koji prima neovisne varijable ili ulaze u model
2. jedan skriveni sloj
3. izlazni sloj koji generira predviđanja.

Duboka neuralna mreža ima sličnu strukturu. Razlika je u tome što ima 2 ili više 'skrivenih' slojeva neurona koji obrađuju ulazne podatke. Iako se i plitke neuralne mreže mogu nositi s kompleksnim problemima, duboke neuralne mreže daju ispravnija predviđanja te su preciznije što im se dodaje više slojeva s neuronima. Najčešće je optimalno imati do 9 ili 10 skrivenih slojeva nakon čega moće predviđanja neuralne mreže počinje opadati. Danas duboke umjetne neuralne mreže imaju najčešće između 3-10 skrivenih neronskih slojeva.

KONVOLUCIJSKE NEURALNE MREŽE

Pokazalo se da su konvolucijske neuralne mreže jako efektivne u rješavanju zadataka koji uključuju podatke koji su zbijeni skupa, primarno u polju računalnog vida. Konvolucijska mreža koristi 3D strukturu s tri skupa neurona koji analiziraju 3 sloja boja u slici – crveni sloj, plavi i zeleni. Analizira jedno po jedno područje slike kako bi identificirala važne uzorke.

Struktura se potpuno povezane neuralne mreže, u kojoj svi neuroni iz jednog sloja komuniciraju sa svim neuronima iz sljedećeg sloja, pokazala neefikasnom tamo gdje je potrebno analizirati slike velikih dimenzija. Konvolucijska mreža koristi tri-dimenzionalne strukture kod kojih neuroni u jednom sloju ne komuniciraju sa svim neuronima u sljedećem sloju. Umjesto toga, svaki skup neurona analizira malo područje ili značajku slike.

Konačan rezultat ovakve strukture je vektor s ocjenama vjerojatnosti. Konvolucijska mreža prvo izvodi konvoluciju (eng. *convolution*), koja 'skenira' sliku. Tako mreća analizira jedan po jedan mali dio slike i kreira 'matricu značajki' (eng. *feature map*) koja sadrži vjerojatnosti da svaka značajka pripada zahtijevanoj slasi. Drugi korak je 'udruživanje' (eng. *pooling*) koji reducira dimenzije svake značajke dok u isto vrijeme drži sačuvanima najvažniije informacije.

Korištenje konvolucijskih mreža:

1. prepoznavanje lica
2. identifikacija i klasifikacija svakodnevnih objekata u slici
3. daje vid robotima i autonomnim vozilima
4. prepoznaje scene i predlaže relevantne opise
5. semantičko parsiranje, klasifikacija rečenica i predviđanje
6. dohvat upita

<https://ch.mathworks.com/discovery/convolutional-neural-network-matlab.html>

Konvolucijska neuralna mreža je jedna od arhitektrua umjetnih neuralnih mreža koja uči direktno iz podataka i tako eliminira potrebu manualnog izdvajanja značajki.

Konvolucijske mreže su osobito korisne za pronalaženje uzoraka u slikama s ciljem prepoznavanja objekata, lica i krajolika. Također znaju biti izrazito djelotvorne i kod klasificiranja ne-vizualnih podataka kao što su audio podaci, podaci s vremenskim serijama ili signalima.

ŠTO ČINI CNN TAKO KORISNOM?

Konvolucijske neuralne mreže su popularne radi tri glavna faktora:

1. Konvolucijske mreže eliminiraju potrebu manualnog izvlačenja značajki – značajke se uče direktno od strane konvolucijske mreže.
2. Konvolucijske mreže daju rezultate prepoznavanja s velikom točnošću.
3. Konvolucijske mreže se mogu ponovno istrenirati za nove zadatke prepoznavanja. Tako omogućuju građenje nove mreže na temelju već postojeće.

NAČINI RADA KONVOLUCIJSKIH NEURALNIH MREŽA

Konvolucijske mreže sadrže desetke ili čak stotine skrivenih slojeva. Svaki sloj uči prepoznavati različite značajke slike. Filteri se primjenjuju na svaku sliku iz skupa podataka za treniranje te je rezultat svake konvoluirane slike prosljeđen kao ulaz sljedećem sloju. Filteri mogu početi kao jednostavni filteri koji raspoznaju najjednostavnije značajke kao što su rubovi i svjetlina. Što se filteri nalaze dublje u mreži, tako raste kompleksnost značajki koje jedinstveno identificiraju neki objekt.

UČENJE ZNAČAJKI, SLOJEVI I KLASIFIKACIJA

Kao i ostale umjetne neuralne mreže, konvolucijska mreža se sastoji od ulaznog sloja, izlaznog sloja i više skrivenih slojeva između ulaznog i izlaznog sloja.

Skriveni slojevi izvode operacije koji izmjenjuju ulazne podatke s namjerom da model nauči značajke koje su specifične za pojedine podatke. Tri najčešće korištena sloja su:

1. Konvolucijski sloj: obrađuje ulaznu sliku skupom konvolucijskih filtera. Svaki od filtera aktivira određene značajke na slici.
2. Aktivacijski sloj: aktivira značajke koje se prosljeđuju sljedećem sloju u mreži.
3. Sloj za udruživanje (eng. *pooling*): pojednostavljuje izlaz tako što izvodi ne-linerano uzorkovanje i tako reducira broj parametara koje mreža treba naučiti.

Ove operacije se ponavljaju kroz desetke ili stotine slojeva. Svaki sloji uči identificirati različite značajke.