

Kemometrijska klasifikacija psihoaktivnih uzoraka

Tominac, Petra

Master's thesis / Diplomski rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Chemical Engineering and Technology / Sveučilište u Zagrebu, Fakultet kemijskog inženjerstva i tehnologije**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:149:038820>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-29**



Repository / Repozitorij:

[Repository of Faculty of Chemical Engineering and Technology University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE
SVEUČILIŠNI DIPLOMSKI STUDIJ

Petra Tominac

DIPLOMSKI RAD

Zagreb, rujan 2019.

SVEUČILIŠTE U ZAGREBU
FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE
SVEUČILIŠNI DIPLOMSKI STUDIJ

Petra Tominac

KEMOMETRIJSKA KLASIFIKACIJA PSIHOAKTIVNIH UZORAKA

DIPLOMSKI RAD

Voditelj rada: prof. dr. sc. Tomislav Bolanča

Zagreb, rujan 2019.

Ovaj rad izrađen je u Centru za forenzična ispitivanja „Ivan Vučetić“.

Zahvaljujem se mentoru prof. dr. sc. Tomislavu Bolanči na pruženoj prilici da ovaj rad napišem u suradnji sa Centrom za forenzična ispitivanja „Ivan Vučetić“.

Veliko hvala dr. sc. Šimi Ukiću, izv. prof. na strpljenju, nesebičnoj pomoći i razumijevanju pri izradi ovog diplomskog rada.

Posebno hvala Ani, Ivanu, Mileni, Mariji, Kristini i Marku koji su uvijek bili uz mene, tjerali me da dajem sve od sebe i uz osmijeh učinili sve brige lakšima, a život ljepšim.

Najveća hvala svojim roditeljima i braći na razumijevanju i trudu te financijskoj i moralnoj podršci koju su mi pružili kako bih došla do ove diplome.

Sažetak

U ovom radu provedena je klasifikacija uzoraka opojnih droga upotrebom umjetnih neuronskih mreža. Za potrebe istog, najprije je provedena optimizacija mreža: određen je tip mreže, broj neurona u skrivenom sloju, te aktivacijska i funkcija za treniranje. Najboljim se pokazao višeslojni linearni perceptron s 12 neurona u skrivenom sloju, tangens hiperbolnom aktivacijom, treniran funkcijom konjugiranog gradijenta uz visku točnost od 84,93% ispravnog predviđanja na nezavisnom (vanjskom) setu varijabli.

Umjetne neuronske mreže pokazale su se kao brz i pouzdan alat kojim može ostvariti kvalitetna klasifikacija uzoraka opojnih droga te kao takve sigurno imaju svijetlu perspektivu u rasvjetljavanju kriminalnih djelatnosti vezanih uz transfer opojnih droga.

Ključne riječi: umjetne neuronske mreže, klasifikacija podataka, opojne droge

Summary

The classification of samples of narcotic drugs using artificial neural networks is shown in this paper. For the same purpose, network optimization was first implemented: a network type was determined, a number of neurons within the hidden layer, also an activation and training functions were determined. The best solution is multilayer linear perceptron with 12 neurons within the hidden layer, with tangent hyperbolic activation function, trained with conjugated gradient function in the high accuracy of 84.93% of the correct prediction on the independent (external) set of variables.

Artificial neural networks have proven to be a fast and reliable tool that can achieve a high-quality classification of narcotic drugs samples and certainly have a bright perspective on the illicit trafficking of narcotic drugs.

Keywords: artificial neural networks, data classification, narcotics

Sadržaj

1.	Uvod	1
2.	Teorijski dio	2
2.1.	Psihoaktivne tvari	2
2.1.1.	Kokain	2
2.1.2.	Heroin	4
2.1.3.	Karakterizacija psihoaktivnih tvari i profiliranje nečistoća	5
2.2.	Klasifikacija podataka	7
2.3.	Umjetne neuronske mreže	10
2.3.1.	Učenje mreža	12
2.3.2.	Višeslojni linearni perceptron	14
2.3.3.	Mreže s radijalnom funkcijom	17
2.3.4.	Procjena valjanosti modela	17
3.	Eksperimentalni dio	18
4.	Rezultati i rasprava	19
4.1.	RBF	19
4.2.	MLP	19
4.2.1.	Aktivacija eksponencijalnom funkcijom	19
4.2.2.	Aktivacija pravcem $y=x$	21
4.2.3.	Aktivacija logističkom funkcijom	22
4.2.4.	Aktivacija sinusnom funkcijom	23
4.2.5.	Tangens hiperbolna aktivacijska funkcija	25
4.3.	Validacija	26
4.3.1.	RBF	26
4.3.2.	MLP	27
5.	Zaključak	35
6.	Popis pokrata	36
7.	Literatura	37
8.	Životopis	39

1. Uvod

U cilju suprotstavljanja rastućem problemu s psihoaktivnim tvarima (drogama), javlja se sve veća potreba za identifikacijom putova njihove trgovine te prikupljanja podataka o broju izvora psihoaktivnih tvari, lokaciji izvora (jesu li unutar zemlje ili na međunarodnoj razini), distribucijskim točkama i distribucijskoj mreži. Znanost pomoću karakterizacije i profiliranja nečistoća zaplijenjenih psihoaktivnih tvari daje potporu operacijama provedbe zakona. Studije karakterizacija psihoaktivnih tvari pokazale su da je moguće međusobno povezati uzorke, klasificirati materijal iz različitih zapljena u skupine povezanih uzoraka i samim time identificirati podrijetlo uzoraka. Takve se informacije mogu koristiti u dokazne (sudske) svrhe ili se mogu koristiti kao izvor obavještajnih podataka za identifikaciju uzoraka koji mogu imati zajedničko podrijetlo ili povijest. Karakterizacija psihoaktivnih tvari i profiliranje nečistoća također mogu pomoći u utvrđivanju rezultata iz novih nezakonitih laboratorija te u praćenju uobičajenih metoda koje se koriste za proizvodnju droga, što pak može pružiti informacije koje su korisne za održavanje drugih alata za prikupljanje obavještajnih podataka. Naposljetku, karakterizacija droga i profiliranje nečistoća također mogu pružiti popratne dokaze u slučajevima kada se nezakonito proizvedena droga mora razlikovati od one koja je preusmjerena iz dozvoljenih izvora. Sustavna analiza sastava zaplijenjenih lijekova doprinosi boljem razumijevanju trendova zlouporabe droga.

Na svojoj trideset devetoj sjednici, održanoj 1996. godine, UN-ova Komisija za opojne droge prepoznala je potrebu za kohezivnom međunarodnom strategijom na području karakterizacije droga i profiliranja nečistoća. Time je Komisija zatražila od Ureda Ujedinjenih naroda za droge i kriminal (UNODC) da razvije standardne protokole i metode za analizu profila (potpisa) ključnih opojnih droga i psihotropnih tvari. Ciljevi UNODC-s u ovom području su razviti smjernice za karakterizaciju i profiliranje nečistoća ključnih opojnih droga i psihoaktivnih tvari te učiniti smjernice dostupnima nacionalnim forenzičnim laboratorijima s ciljem da im se pruži znanstveni alat za potporu rada policije, kako za dokazne tako i strateške obavještajne svrhe.¹

2. Teorijski dio

2.1. Psihoaktivne tvari

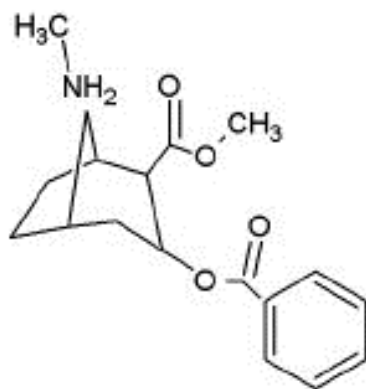
Prema Zakonu o suzbijanju zlorabe droga²: „droga je svaka tvar prirodnoga ili umjetnoga podrijetla, uključivši psihoaktivne tvari, uvrštene u popis droga i psihoaktivnih tvari.“ Odnosno, opojne droge su sintetske ili prirodne psihoaktivne tvari koje izazivaju ovisnost. Često se takve tvari nazivaju i narkoticima. Psihoaktivne tvari su kemijske tvari intenzivnog fiziološkog učinka čija upotreba dovodi do promijene moždane funkcije što rezultira privremenom promjenom percepcije, ponašanja ili raspoloženja svijesti. Psihoaktivne tvari često se koriste u rekreativnoj uporabi, kao poticaj u spiritualne svrhe, ali i kao lijek, ponajviše pri liječenju psihičkih i neuroloških bolesti.³ Ovisno o učincima psihoaktivnih tvari na središnji živčani sustav čovjeka, njegove psihičke procese (osjet, mišljenje, govor, percepcija, raspoloženje) i njegovo ponašanje opojne droge se dijele na depresore, stimulanse i halucinogene.

Depresori su sredstva koja imaju umirujuće ili uspavljujuće djelovanje na središnji živčani sustav, odnosno usporavaju moždane funkcije. Dije se u tri kategorije: sredstva za smirivanje (sedativi, benzodiazepini, hipnotici, neopijatski analgetici, pojedini antiepileptici), alkohol i opijati (opijatske droge kao heroin, morfin, kodein, metadon). Najviše se zlorabljavaju alkohol, sredstva za spavanje i sintetski sedativi.⁴

Stimulansi (psihostimulansi) su spojevi koji aktiviraju, poboljšavaju ili povećavaju funkcioniranje središnjeg živčanog sustava.⁴ Psihostimulansi podižu raspoloženje, održavaju stanje budnosti, proizvode euforiju, uzrokuju povećanje krvnog tlaka i proširenja zjenica, ubrzavaju rad srca, te mogu uzrokovati mučninu i povraćanje. U stimulativna sredstva spadaju kokain, amfetamini, metamfetamini i najpoznatiji stimulans kofein koji stvara blage učinke.⁵

2.1.1. Kokain

Kokain je tropanski alkaloid koji se dobiva iz suhog lišća koke (*Erythroxylon coca*). Najstariji je poznati lokalni anestetik te je služio kao model za sintezu brojnih drugih manje toksičnih sintetičkih lokalnih anestetika.⁶ Kao stimulans, kokain izaziva osjećaj kratkotrajne euforije i energije kao i opasne fizičke efekte poput visokog krvnog tlaka i ubrzanih otkucaja srca.



Slika 1. Struktura kokaina

Kokain se konzumira na tri načina. U obliku praha, kokain se inhalira kroz nos (ušmrkava) ili se otapa u vodi te ubrizgava u krvotok. *Crack* se konzumira pušenjem. *Crack* je čista slobodna baza kokaina u obliku kristalića koji se dobiva posebnom obradom kokaina. Kristal se zagrijava te nastaje para koja se kroz pluća apsorbira u krvotok. *Crack* je dobio ime zbog zvuka pucanja (*crackling sound*) koji se dobiva zagrijavanjem kristala.

Ovisno o načinu konzumiranja kokaina, intenzitet i trajanje će se razlikovati. Ukoliko se kokain ubrizgava ili puši, intenzitet će biti jači, no osjećaj ugone (*high*) trajat će do tri puta kraće nego prilikom ušmrkavanja (5-10 minuta prilikom ubrizgavanja/pušenja, 15-30 minuta prilikom ušmrkavanja). Osobe koje konzumiraju kokain uzet će drogu više puta u kratkom periodu i povećat će dozu kako bi zadržali osjećaj ugone što vodi do ovisnosti. Ovisnost je kronična bolest koja uzrokuje promjene u mozgu ovisnika, a karakterizira ju nekontrolirana potreba za drogom bez obzira na posljedice.⁷

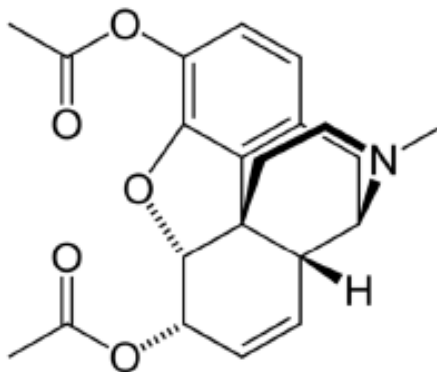
Čistoća zaplijenjene droga tipa kokain u RH u 2014.

Smjese kokaina koje su zaplijenjene u 2014. godini najviše su sadržavale primjese analgoantipiretika fenacetina, lokalnog anestetika i antiaritmika lidokaina, psihostimulansa kofeina, aminokiselinu kreatin, lokalne anestetike benzokain i prokain, šećer laktozu i šećerni alkohol manitol. Određena je i prisutnost antihelminitika levamisola i hidroksizina.

Kvantitativnim vještačenjem je bilo obuhvaćeno 129 predmeta s ukupno 340 uzoraka, pri čemu je minimalni udio kokainske baze iznosio 2,7%, maksimalni 88,7% i prosječni 39,1%. Od tih 129 predmeta, njih 67 s ukupno 90 uzoraka se odnosilo na količine do 1 grama, tzv. „ulične doze“. U tim slučajevima minimalni udio kokainske baze je iznosio 2,7%, maksimalni 78,4% i prosječni 38,4% (Centar za forenzična ispitivanja, istraživanja i vještačenja „Ivan Vučetić“, 2014).⁷

2.1.2. Heroin

Heroin je alkaloid koji se dobiva acetiliranjem morfina. Morfin je najvažniji alkaloid opijuma koji se koristi kao analgetik jer umiruje i umanjuje osjećaj bola, no izaziva ovisnost.⁶ Kao depresor, heroin djeluje na središnji živčani sustav, usporava rad srca, smanjuje krvni tlak i aktivnost crijeva, uzrokuje pospanost i usporava disanje. Pojavljuje se u obliku smeđeg i bijelog praha, a rjeđe u obliku ljepljive crne supstance poznatije pod nazivom „crni katran heroin“.⁸



Slika 2. Struktura heroína

Kao i kokain, heroin se može konzumirati na tri načina: ubrizgavanjem u krvotok, inhaliranjem kroz nos (ušmrkavanjem) ili pušenjem. Neovisno o načinu konzumiranja, heroin veoma brzo dolazi do mozga i uzrokuje zdravstvene rizike kao i veliki rizik da osoba koja ga konzumira postane ovisnik. Konzumiranjem heroína prvo se javlja osjećaj euforije nakon koje dolazi do stanja drijemanja. Heroin se dolaskom do mozga pretvara u morfin koji se veže na opioidne receptore koji se nalaze u mozgu i u tijelu, ponajviše na dijelovima koji su uključeni u percepciju boli i nagrada. Receptori opioida koji se nalaze u moždanom deblu kontroliraju životne procese poput krvnog tlaka, uzbuđenja i disanja. Konzumiranjem i predoziranjem heroinom može doći do prestanka disanja što uzrokuje smrt. Osobe koje konzumiraju heroin imaju povećan rizik zaraze HIV-om ili hepatitisom C, opasnost od puknuća vena, infekcije srčanih zalizaka i ovojnica, bolesti jetre i bubrega te upale pluća.⁸

Čistoća zaplijenjene droga tipa heroin u RH u 2014.

Smjese heroína koje su zaplijenjene u 2014. godini najviše su sadržavale primjese analgoantipiretika paracetamola i psihostimulansa kofeina, a rjeđe šećere saharozu i laktozu, šećerni alkohol manitol. U tragovima nalazili su se anksiolitik diazepam i fungicid griseofulvina.

Kvantitativnim vještačenjem obuhvaćeno je 97 predmeta s ukupno 350 uzoraka, pri čemu je minimalni udio heroinske baze iznosio 0,4%, maksimalni 61% i prosječni 14,5%. Od 97 predmeta, njih 35 s ukupno 48 uzoraka se odnosilo na količine do 1 grama. U tim slučajevima minimalni udio heroinske baze je iznosio 0,4%, maksimalni 18,5% i prosječni 9,9% (Centar za forenzična ispitivanja, istraživanja i vještačenja „Ivan Vučetić“, 2014).⁸

2.1.3. Karakterizacija psihoaktivnih tvari i profiliranje nečistoća

U organskoj kemiji izraz *karakterizacija* često se koristi za opisivanje postupka određivanja točne molekulske strukture organskog spoja. Međutim, u forenzici, pojam *karakterizacija psihoaktivnih tvari* koristi se za određivanje glavnih značajki uzoraka droga, njihovih fizikalnih i kemijskih karakteristika, a također uključuje prisutnost i prirodu primjesa (engl. *cutting agents*) i kvantifikaciju značajnih komponenti uzoraka. Profiliranje nečistoća obično uključuje nekoliko analiza pomoću kojih se dobije detaljna slika (profil) uzorka psihoaktivnih tvari, obično u obliku kromatografskih podataka. Dobiveni kemijski profili, također poznati kao „potpisi“ ili „profili nečistoće“, koriste se u identifikaciji i kvantifikaciji glavnih komponenti u uzorku. Međutim, za razliku od karakterizacije psihoaktivnih tvari, profiliranje nečistoća uključivat će jednu ili više dodatnih analiza koje su namijenjene za ciljanje manjih komponenti u uzorku (komponenti u tragovima). Krajnji cilj analitičara prilikom profiliranja nečistoća je dobivanje profila glavnih i sporednih komponenti u formatima koji mu omogućuju da podatke koristi za usporedbu u svrhu lociranja drugih uzoraka koji imaju slične profile. Kako bi se korištenjem profila nečistoća uzorci mogli usporediti i odrediti veze, treba razmotriti sva moguća sredstva za prikupljanje informacija i podataka, uključujući korištenje fizikalnih svojstava, pakiranja, punila i razrjeđivača.

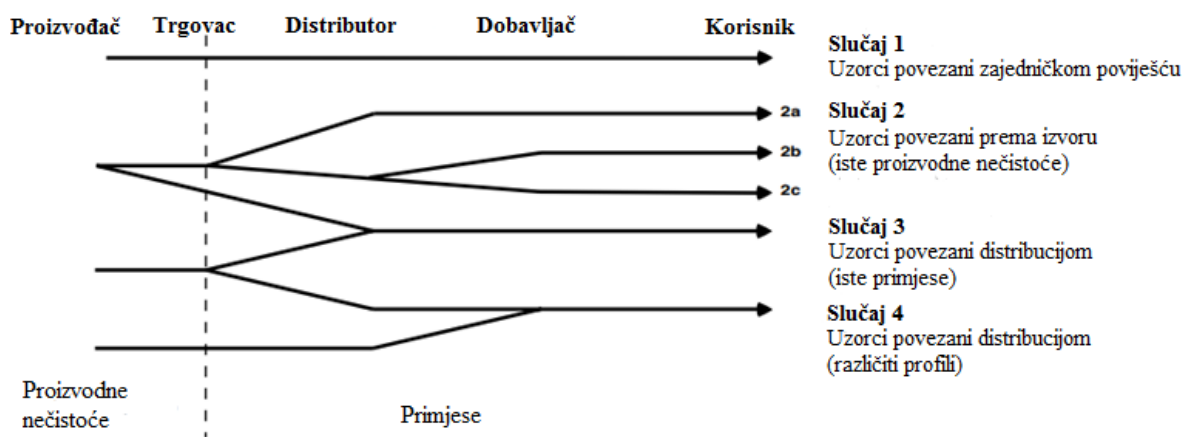
Primjena profiliranja nečistoća uglavnom leži u jednoj od dvije kategorije koje se preklapaju, taktička (dokazna/pravosudna) i strateška (obavještajna) kategorija. Primjer taktičke primjene profiliranja nečistoća je uspostavljanje veza distribucije i/ili trgovine između više zaplijenjenih uzoraka koji su dobiveni na različitim mjestima ili u posjedu različitih pojedinaca. Strateška uporaba može uključivati identifikaciju metoda sinteze koje se koriste u ilegalnim laboratorijima, identifikaciju kemikalija, reagensa i/ili otapala koje koriste ti laboratoriji, znanstveno definiranje i određivanje zemljopisnog podrijetla uzorka.

Kako bi uspješno primijenili rezultate profiliranja nečistoća, sve uključene strane moraju razumjeti i ciljeve i ograničenja tehnike. Prvo i najvažnije, mora se razumjeti da profiliranje onečišćenja droge nije samostalna tehnika, nego znanstveni pristup koji

nadopunjuje informacije o provedbi zakona. Bliska suradnja između laboratorija i osoblja za provedbu zakona nužan je uvjet za uspješnu provedbu ovog posla. Ne samo da forenzički kemičar mora dobro poznavati kemiju svih ciljanih analita, već mora u potpunosti razumjeti svrhu istraživanja kako bi odabrao najprikladniji analitički pristup i pravilno interpretirao rezultate.

Prvi i apsolutno nužan preduvjet za uspješnu provedbu bilo kojeg programa za određivanje podrijetla droga je pribavljanje relevantnih skupova podataka dobivenih iz uzoraka droga poznatog izvora. Takva kompilacija općenito se naziva *autentičnom bazom podataka*. Prikupljanje uzoraka za autentične baze podataka izravno iz polja uzgoja ili izravno iz nezakonitih laboratorija težak je, opasan i skup zadatak koji općenito predstavlja glavnu prepreku stvaranju uspješnog programa određivanja podrijetla za bilo koju nedopuštenu drogu.

Za droge koje su dobivene iz biljaka najčešće je moguće odrediti geografsko podrijetlo ako analitičke tehnike koje se koriste ciljaju isključivo glavne alkaloide. Međutim, postoje iznimke, pogotovo kod uzoraka heroina koji potječe iz jugoistočne Azije. Relativni omjeri alkaloida prisutnih u većini uzoraka heroina jugoistočne Azije znatno se razlikuju od onih u uzorcima heroina koji potječu iz drugih izvora, te se s razumnom sigurnošću, mogu razlikovati jednostavno usporedbom glavnih analiza alkaloida. S druge strane, mnoge uzorke visoko rafiniranog heroina iz jugoistočne Azije nemoguće je čak i uvjetno razlikovati od heroina iz Južne Amerike koristeći samo podatke o alkaloidima.¹



Slika 3. Prikaz lanca opskrbe drogama

Studije karakterizacije psihoaktivnih tvari mogu pružiti informacije korisne za tijela za provedbu zakona o drogama. Mogu se uspostaviti kemijske veze između uzoraka, a materijal

iz različitih zapljena može se svrstati u skupine srodnih uzoraka. Za tijela za provedbu zakona najkorisnije je to što se mogu uspostaviti specifične veze, primjerice, između dobavljača i korisnika, mogu se izgraditi obrasci (mreže) distribucije droga, te se može identificirati izvor, uključujući geografsko podrijetlo uzoraka droge. Točna svrha bilo koje usporedne studije određuje analitički pristup. Forenzični kemičar mora paziti na tumačenje rezultata, uzimajući u obzir osobitosti različitih vrsta droga kao i prisutnost ili odsutnost različitih vrsta nečistoća i primjesa. U lancu opskrbe drogama, koji se sastoji od proizvodnje (izvora), trgovine, distribucije i opskrbe (slika 3), mogu postojati sljedeće veze između dva ili više uzoraka⁹:

- a) uzorci mogu biti povezani zajedničkom poviješću (tj. proizvedeni od istog proizvođača i distribuirani u istom lancu opskrbe, slučaj 1)
- b) mogu biti povezani prema izvoru, ali ne i prema njihovoj distribuciji (slučaj 2)
- c) mogu biti povezani distribucijom, ali potječu iz različitih izvora (slučajevi 3 i 4)

2.2. Klasifikacija podataka

Moderna analitička instrumentacija stvara veliku količinu podataka. Izumom računala, većina multidimenzionalnih podataka pohranjuje se u računalnim instrumentima ili vanjskim bazama podataka. U današnje vrijeme, opsežna primjena kemometrije omogućuje učinkovitu evaluaciju i interpretaciju tih podataka. Glavni ciljevi primjene multivarijatnih metoda u analitičkoj kemiji usmjereni su na grupiranje i klasifikaciju objekata (uzoraka, spojeva ili materijala) kao i na modeliranje odnosa između različitih analitičkih podataka. Neki tipični primjeri uključuju sljedeće:

- Svrstavanje uzoraka u grupe odnosno klastere prema sličnim elementima ili uzorcima materijala s obzirom na usporedivi kemijski sastav i tehnološka svojstva.
- Razvrstavanje uzoraka pomoću analitičkih podataka (spektar, kromatogram ili elementarni uzorak) na temelju poznate klase pripadnosti tih objekata.

Analitički podaci mogu biti uređeni kao podatkovna matrica X od n redaka koji predstavljaju objekte i p stupaca koji predstavljaju svojstva. Objekti mogu biti uzorci, molekule, materijali. Tipične značajke ili varijable tih objekata su elementarni obrasci, spektri, strukturne značajke ili fizikalna svojstva. Matrica podataka $n \times p$ može se napisati na sljedeći način:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \quad (1)$$

Za neke probleme ti se podaci mogu podijeliti u kolone u ovisne i nezavisne varijable, na primjer, za kalibraciju koncentracija na spektrima. Klasa obuhvaća skup objekata koji imaju slične značajke. Uzorak objekta je njegova zbirka karakterističnih značajki.

Formuliranom procjenom podataka nisu nužno korišteni svi objekti i značajke. Neki od dostupnih podataka ne mogu se koristiti u zadanom obliku, stoga je prethodna obrada podataka preduvjet za učinkovitu multivarijatnu analizu podataka.¹⁰

Predobrada podataka

U prvom koraku predobrade podataka, podaci se moraju pregledati i vidjeti jesu li potpuni. Nepotpuni podaci ne ometaju matematičku analizu, no ne smije ih se zamijeniti nulama. Umjesto toga slobodna mjesta treba popuniti srednjom vrijednošću stupca/retka ili u najgorem slučaju, generiranjem slučajnog broja unutar intervala u kojem se nalaze vrijednosti promatranog stupca/retka. Podaci se mogu i ukloniti iz skupa podataka ukoliko su međusobno izrazito korelirani, konstantni ili vidljivo odstupajući. Kako bi se uklonio konstantan pomak podataka, podaci se mogu pomaknuti da budu oko ishodišta koordinatnog sustava. Uobičajeni postupak je tzv. *centriranje središta*, gdje je svaka varijabla, x_{ik} , centrirana oduzimanjem srednje vrijednosti stupca, \bar{x}_k :

$$x_{ik}^* = x_{ik} - \bar{x}_k \quad (2)$$

gdje je i indeks reda, k indeks stupca, a \bar{x}_k vrijednost stupca izračunata prema jednadžbi (3).

$$\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik} \quad (3)$$

Značajke predstavljaju sasvim različita svojstva uzorka ili objekta, tako da se metrika u velikoj mjeri može razlikovati od stupca do stupca. To može značiti različite apsolutne vrijednosti varijabli kao i različite raspone varijabli (varijance). Oba tipa izobličenja utjecat će na većinu statističkih multivarijatnih metoda. Uklanjanje tih razlika može se provesti skaliranjem podataka u slične raspone i varijance. Važne su dvije metode skaliranja koje skaliraju podatke prema opsegu ili prema standardnoj devijaciji (automatsko skaliranje).¹⁰

Podatkovno rudarenje (engl. *data mining*) je termin koji općenito označava „kopanje“ duboko u podatke koji se nalaze u različitim oblicima s ciljem dobivanja informacija i stjecanja znanja o obrascu ponašanja. U procesu rudarenja podataka prvo se sortiraju veliki skupovi podataka, zatim se identificiraju obrasci i uspostavljaju odnosi za analizu podataka i rješavanje problema.¹¹ Podatkovno rudarenje uključuje korištenje sofisticiranih alata kojima se analiziraju podaci kako bi se otkrili nepoznati, ispravni uzorci i odnosi u velikom skupu

podataka. Ovi alati mogu uključivati statističke modele, matematički algoritam i metode strojnog učenja. Prema tome, podatkovno rudarenje se sastoji od prikupljanja i upravljanja podacima, uključuje analizu i predviđanje. Sposobna je obraditi širi spektar podataka od regresije i sve je popularnija te zbog toga je najznačajnija aplikacija za strojno učenje (engl. *machine learning*). Tijekom analize ljudi su često skloni pogreškama i pokušavaju uspostaviti odnose između više značajki što im otežava pronalaženje rješenja za određene probleme. Strojno učenje se često može uspješno primijeniti na te probleme, poboljšavajući učinkovitost sustava i dizajna strojeva.¹²

Postoji nekoliko vrsta klasifikacijskih metoda: stabla odluke (engl. *Decision Trees*), klasifikator k najbližeg susjeda (engl. *k-Nearest Neighbor Method*; KNN)¹², hijerarhijska klaster analiza (engl. *Hierarchical Cluster Analysis*; HCA), analiza glavnih komponenti (engl. *Principal Component Analysis*; PCA), faktorska analiza (engl. *Factor Analysis*; FA), neovisno modeliranje analogije grupa (engl. *Soft Independent Modeling of Class Analogies*; SIMCA), umjetne neuronske mreže (engl. *Artificial Neural Networks*; ANN).¹⁰

Stabla odluke se sastoji od čvorova, grana i listova. Svaki čvor predstavlja jednu značajku rješavanog problema. Grane koje idu iz čvorova predstavljaju svaku moguću vrijednost koju taj čvor može poprimiti i u ovisnosti o vrijednosti tog atributa za određeni ispitni slučaj, putujemo kroz stablo određenom granom. Na kraju se nalaze listovi koji ujedno predstavljaju i razrede na koje možemo klasificirati određene ulazne podatke.¹³

Klasifikator k najbližeg susjeda je metoda klasifikacije temeljena na sličnosti, odnosno, različitosti. Oznaka k odnosi se na broj uspoređivanih najbližih susjeda.¹⁴ Za klasifikaciju nepoznatog objekta, njezina udaljenost, izračunava se na sve objekte. Odabire se minimalna udaljenost i objekt se dodjeljuje odgovarajućoj klasi. Broj susjednih objekata k obično je neparan broj.¹⁰

Hijerarhijska klaster analiza je jedna od mogućih metoda za grupiranje objekata. Objekti se kombiniraju prema njihovim udaljenostima ili sličnostima. Razlikuju se postupci aglomeracije i razdvajanja. Formiranje klasnih podjela temelji se na podjeli cijelog skupa objekata na pojedinačne klastere. Započinje se s pojedinačnim objektima i spaja ih u veće skupine objekata. Za odlučivanje o broju klastera mogu se iskoristiti različiti kriteriji. Vrlo često je poznat broj klastera. U nekim slučajevima, broj klastera se može izvesti iz unaprijed određene mjere udaljenosti ili razlike između klastera.¹⁰

U faktorskoj analizi cilj je izraziti obilježje malim brojem zajedničkih čimbenika te je za svaki faktor dodijeljeno svojstvo koje se ne može izravno promatrati. Primjerice, prilikom

analize zraka u gradu, promet može biti zajednički čimbenik. Faktorska analiza prvotno je razvijena za objašnjavanje psiholoških teorija, poput, opisivanja čimbenika kao što su inteligencija ili pamćenje. Metodom rotacije apstraktni faktori određeni u prvom koraku transformiraju se u interpretativne čimbenike.¹⁰

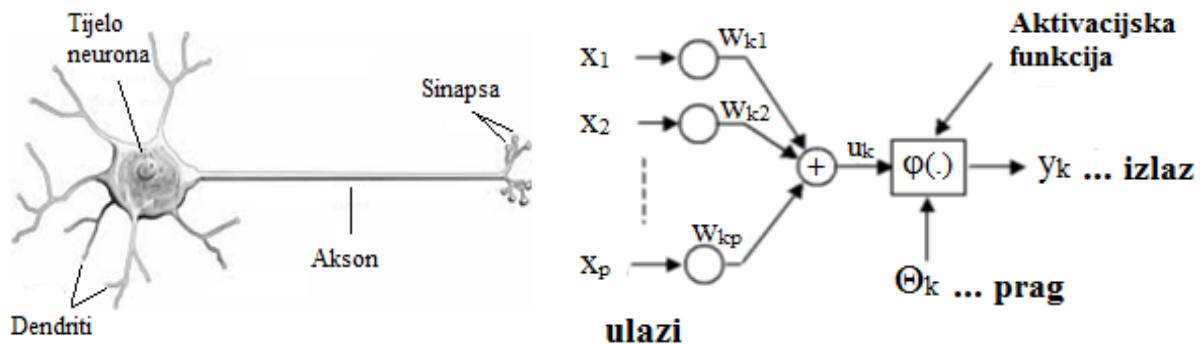
2.3. Umjetne neuronske mreže

McCulloch i Pitts, inspirirani načinom funkcioniranja ljudskog mozga, prvi su opisali model umjetnog neurona čime počinje moderna era neuronskih mreža.¹⁵ Prema Alexanderu i Mortonu¹⁶ umjetna neuronska mreža je masivno paralelni distribuirani procesor koji je dobar za pamćenje iskustvenog znanja. Neuronske mreže mozgu su slične u dva aspekta: znanje stječu kroz proces učenja i koriste međusobne veze između neurona za spremanje znanja. Mozak obrađuje podatke drugačije od konvencionalnih digitalnih računala. Neuronu u mozgu su 5-6 redova veličina sporiji od digitalne logike (ms i ns), ali zbog ogromnog broja neurona (oko 10 milijardi neurona i oko 60 000 milijardi međuspojeva) mozak nadoknađuje brzinu. Mozak je veoma kompleksno, nelinearno, paralelno računalo. Zbog toga je istraživanje neuronskih mreža usmjereno na razvoj struktura mreža koje bi funkcionirale analogno ljudskom mozgu u svrhu rješavanja praktičnih problema. Ukoliko računalo posjeduje sposobnost donošenja zaključaka na temelju određenih činjenica tada se ono smatra inteligentnim. Čovjek radnje poput uočavanja i klasificiranja objekata iz svakodnevnog života obavlja intuitivno što računalima predstavlja prepreku. Ipak, pomoću neuronskih mreža moguće je doći do zadovoljavajućeg rješenja jer neuronske mreže imaju visok stupanj tolerancije grešaka, čak i prilikom analize nejasnih i manjkavih podataka. Upravo zbog te sposobnosti, neuronske mreže koriste se za analizu podataka u ekonomiji, tehničkim, društvenim, prirodnim znanostima i mnogim drugim područjima predviđanja i klasifikacije.¹⁷

Mreža neurona uči i razvija se pomoću treninga na velikom broju primjera i za dobivanje rješenja potrebno je definirano znanje, a ne utvrđivanje na pravilima. Neuronske mreže primaju određene ulazne informacije koje transformiraju u izlazne rezultate i zbog toga se mogu opisati modelom podražaj-reakcija gdje je područje između navedenih ispunjeno formalnim modelima.¹⁷

Efikasnost ljudskog mozga i neuronskih mreža možemo prikazati na dva primjera. Čovjek će lice koje vidi prepoznati u 100-200 ms, dok današnjim računalima trebaju sati i dani da obave taj zadatak, ako ga uopće uspiju obaviti. Također, šišmiš će detektirati udaljenost, veličinu, brzinu, azimut, elevaciju objekta i veličinu raznih dijelova objekta

pomoću eholokacije, a sve operacije odvijaju se u mozgu veličine desetak milimetara. Čak i najbolji sonar i radar nisu toliko točni i precizni kao što je mozak šišmiša.¹⁸



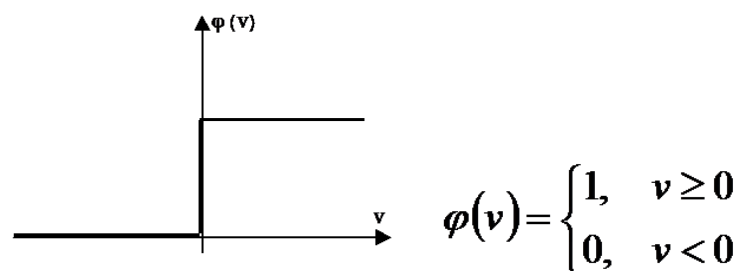
Slika 4. Prikaz biološkog neurona (lijevo) i modela umjetnog neurona (desno)

Na slici 4 dana je usporedba biološkog i umjetnog neurona. Kod biološkog neurona dendriti primaju izlazne informacije s drugih neurona (njihovih završnih nožica). Akson prenosi impulse od dendrita do završnih nožica koje informaciju prenose dalje na dentride susjednih neurona.

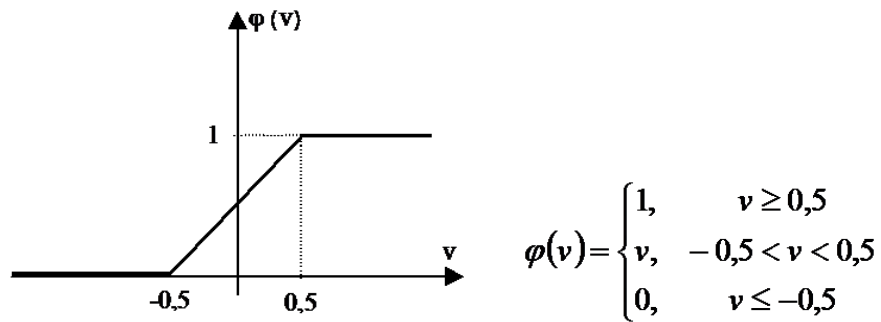
Kod umjetnog neurona umjesto dendrita imamo skup sinapsi, tj. ulaza, od kojih svaki ima svoju važnost tj. težinu, w . Primjerice, signal x_p na p -ulazu u neuron ima težinu w_p . Sumator linearno zbraja ulazne signale pomnožene odgovarajućim težinama. Dobivena suma aktivira se prikladnom aktivacijskom funkcijom koja limitira izlaz neurona, y , na iznos unutar intervala $[0,1]$.

Aktivacijske funkcije dijele se na¹⁸:

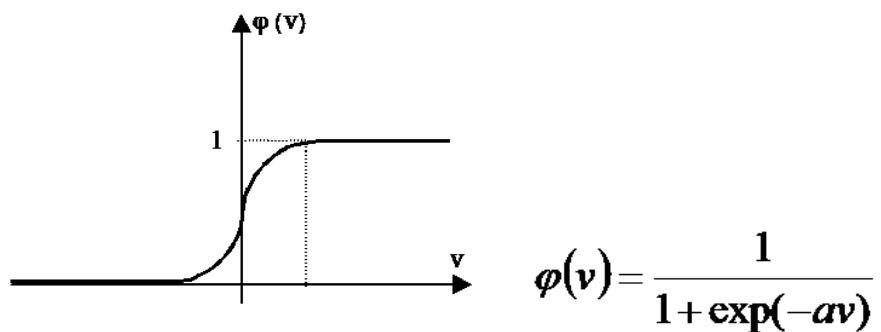
- (a) aktivacijsku funkciju oblika praga koja se koristi za rješavanje klasifikacijske problematike za manji broj klasa (slika 5)
- (b) aktivacijsku funkciju linearnu po odsječcima
- (c) sigmoidnu aktivacijsku funkciju



Slika 5. Aktivacijsku funkciju oblika praga



Slika 6. Aktivacijska funkcija linearna po odsječcima

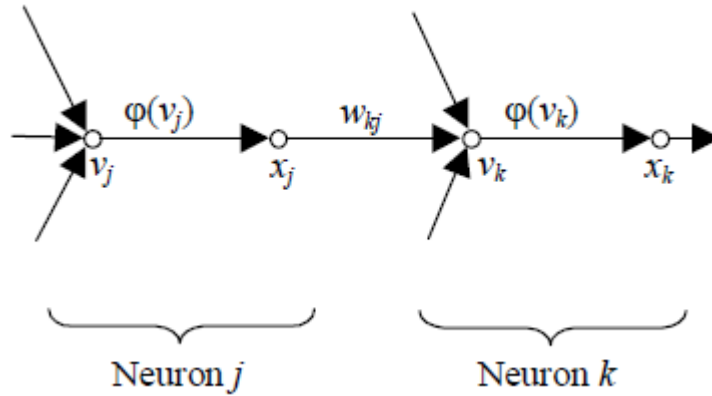


Slika 7. Sigmoidna aktivacijska funkcija

2.3.1. Učenje mreža

Ono što najviše karakterizira neuronske mreže je sposobnost učenja od okoline. Učenje se odvija kroz iterativni proces podešavanja sinaptičkih težina i pragova. Nakon svakog iterativnog procesa, mreža bi trebala imati više znanja o okolini.¹⁹ Dakle, u procesu učenja okolina šalje informaciju elementu za učenje koji će tom informacijom dopuniti bazu znanja. Dodano znanje se ispituje na okolini te se povratna informacija šalje elementu za učenje i na taj način stroj uči pogađanjem i za provjeru znanja prima ocjenu od elementa za učenje.¹⁸

U kontekstu neuronskih mreža učenje je proces adaptiranja slobodnih parametara neuronske mreže kroz kontinuirani proces stimulacije mreže od strane okoline u kojoj se nalazi. Odnosno, okolina stimulira neuronsku mrežu koja se mijenja kao rezultat stimulacije te zbog nastalih promjena neuronska mreža drugačije reagira na okolinu. Način promjene parametara određuje vrstu učenja.¹⁹



Slika 8. Prikaz povezana dva neurona

Na slici 8 prikazana su dva povezana neurona, j i k , gdje su v_j i v_k njihove interne aktivnosti, x_j i x_k izlazi neurona, a $w_{kj}(n)$ je težina sinapse u trenutku n . Ukoliko u trenutku n promijenimo iznos težine sinapse $w_{kj}(n)$ za $\Delta w_{kj}(n)$, dobivamo novu vrijednost težine:

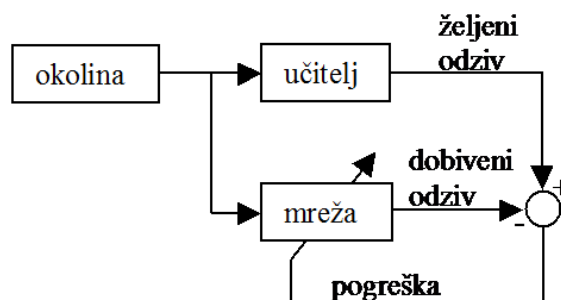
$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (4)$$

Način izračunavanja $\Delta w_{kj}(n)$ određuje algoritam učenja dok odnos neuronske mreže prema okolini određuju paradigme učenja. Algoritam učenja je skup pravila za rješavanje problema učenja. Ne postoji jednostavni algoritam učenja već mnogo različitih algoritama od kojih svaki ima prednosti i mane.¹⁹

Postoje tri osnovne paradigme učenja: učenje pod nadzorom (nadzor vanjskog učitelja), učenje podrškom (engl. *reinforcement learning*, metoda pokušaja i pogreške uz ocjenu suca) i učenje bez nadzora (samoorganizirano - nije potreban učitelj).¹⁹

Učenje pod nadzorom

Glavna karakteristika učenja pod nadzorom (engl. *supervised learning*) je prisutnost vanjskog učitelja koji ima znanje o okolini u obliku parova baza ulaz-izlaz.



Slika 9. Shematski prikaz učenja pod nadzorom

Pogreška je vektorska vrijednost i označava razliku između željenog i dobivenog odziva za neki ulazni vektor. Parametri neuronske mreže mijenjaju se pod utjecajem ulaznih vektora i signala pogreške. Proces se iterativno ponavlja dok mreža ne nauči oponašati učitelja nakon čega učenje završava te učitelj više nije potreban i mreža može raditi bez nadzora.¹⁸ Primjeri algoritama za učenje pod nadzorom su: algoritam najmanjih kvadrata LMS (engl. *least-mean-square*) i algoritam s povratnom propagacijom pogreške BP (engl. *back-propagation*). Nedostatak učenja pod nadzorom je nemogućnost mreže da nauči nove strategije koje nisu pokriven primjerima koji su korišteni za učenje bez vanjskog učitelja. Postoje dva načina izvedbe učenja pod nadzorom: tzv. *on-line* i *off-line* način rada. U *off-line* načinu učenja sustav prvo uči. Nakon završetka učenja, konfiguracija sustava više se ne mijenja i mreža zatim radi na statički način (bez promjena). U *on-line* načinu učenja sustav za vrijeme rada uči i mijenja svoju konfiguraciju u skladu s novim znanjem. Učenje se odvija u stvarnom vremenu odnosno mreža je dinamička.¹⁹

Učenje bez nadzora

Za razliku od učenja pod nadzorom, kod učenja bez nadzora (engl. *unsupervised*) ili samoorganizirajućeg (engl. *selforganized*) učenja, učitelj nije prisutan i ne upravlja procesom učenja. Prisutna je mjera kvalitete znanja koje mreža mora naučiti kao i slobodni parametri koji se optimiziraju obzirom na tu mjeru. Pri učenju bez nadzora koristi se kompetitivno pravilo. Kompetitivni sloj sastoji se od neurona koji se natječu za priliku da odgovore na značajke koji sadrži ulazni vektor. U najjednostavnijem pristupu koristi se *winner-takes-all* strategija.¹⁸

Usporedba učenja sa i bez nadzora

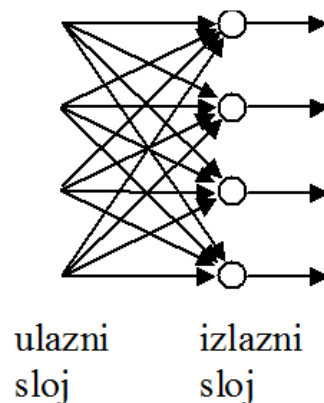
Povratna propagacija pogreške je najčešće korišten algoritam za učenje s nadzorom. Nedostatak ovog algoritma je problem skaliranja gdje s povećanjem broja slojeva mreže i broja neurona vrijeme učenja eksponencijalno raste. Rješenje ovog problema leži u korištenju mreža s više slojeva gdje svaki sloj neovisno uči bez nadzora.¹⁹

2.3.2. Višeslojni linearni perceptron

Topologija mreže određuje način na koji su povezani neuroni. Postoji više vrsta neuronskih mreža: jednoslojne mreže bez povratnih veza ili perceptron (engl. *single-layer feedforward networks*), višeslojne mreže bez povratnih veza ili višeslojni linearni perceptron

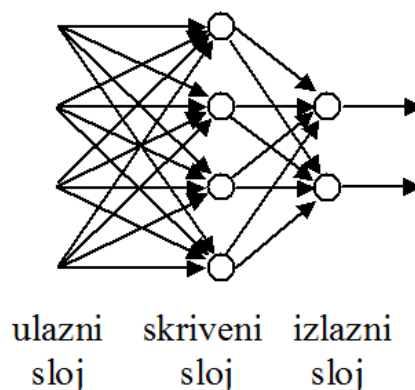
(engl. *multi-layer feedforward networks*, *multilayer perceptron* (MLP)), mreže s povratnim vezama (engl. *recurrent networks*), ljestvičasto strukturirane mreže (enlg. *lattice structures*), samoorganizirajuće mape (engl. *selforganizing map*)...

Perceptron se sastoji od jednog, izlaznog, sloja neurona s obzirom na to da se ulazni sloj neurona ne broji jer u njemu nema procesiranja. Ulazi u mrežu spojeni su na ulaze neurona čiji izlazi predstavljaju i izlaz mreže, dakle nema povratnih veza s izlaza na ulaz.¹⁸

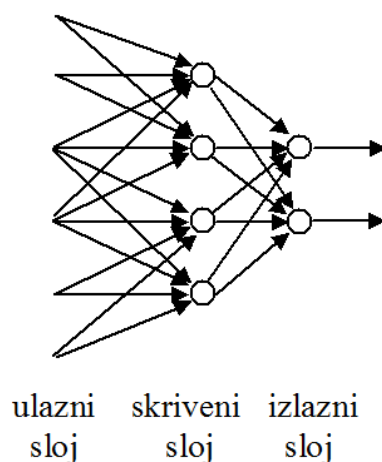


Slika 10. Prikaz jednoslojne neuronske mreže

Višeslojni linearni perceptron (slike 11 i 12) osim ulaznog i izlaznog sloja imaju i jedan ili više skrivenih slojeva neurona. Izlaz neurona iz n -tog sloja predstavlja ulaz u neuron $n+1$ sloja i nema povratnih veza.¹⁸ Iako višeslojni linearni perceptron pokazuje karakteristike koje ga čine poželjnim, ipak posjeduje i neke mane. Visoka nelinearnost može otežati teoretsku analizu ove mreže; korištenje skrivenih neurona otežava vizualizaciju procesa učenja, a proces učenja je teži zbog velikog broja neurona i zbog potrebe da mreža sama odluči što trebaju naučiti skriveni neuroni.²⁰



Slika 11. Prikaz potpuno povezane mreže s jednim skrivenim slojem



Slika 12. Prikaz djelomično povezane mreže s jednim skrivenim slojem

U potpuno povezanoj mreži (slika 11) svaki neuron u nekom sloju povezan je sa svim neuronima u slijedećem sloju. U nepotpuno povezanim mrežama (slika 12) nedostaju neke veze.¹⁸

Višeslojni linearni perceptron koristi se za rješavanje širokog spektra problema gdje se učenje pod nadzorom odvija pomoću algoritma s povratnom propagacijom pogreške (engl. *error back-propagation algorithm*). Glavne karakteristike višeslojnog linearnog perceptrona jesu da model neurona ima nelinearnu izlaznu karakteristiku koja je glatka, mreža je dobro povezana i sadrži jedan ili više skrivenih slojeva neurona (postoji puno sinapsi).²⁰

Tehnike minimalizacije

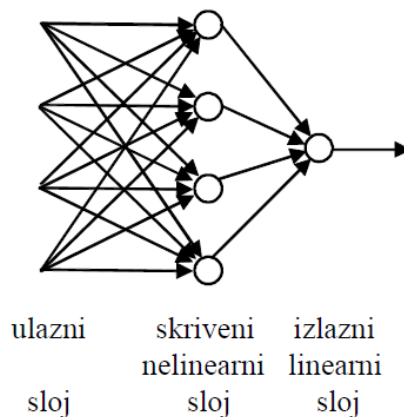
Učenje se sastoji od minimalizacije odstupanja između promatranih i mrežom predviđenih vrijednosti. Da bi se to ostvarilo u neuronskoj mreži, postoji tzv. *kaznena funkcija* takva da ima minimalnu vrijednost samo kada su izlazne vrijednosti jednake ciljnim vrijednostima. Preko nepoznatih parametara minimizira se vrijednost kaznene funkcije. Tehnike minimiziranja mogu se podijeliti u dvije vrste:

- lokalne; traže lokalni minimum,
- globalne; traže globalni minimum.

Globalna minimalizacija je poželjnija, ali je zahtjevnija za računanje. Često je nepraktična zbog složenosti modela i potencijalne veličine skupova podataka.¹⁵

2.3.3. Mreže s radijalnom funkcijom

Kod mreža s radijalnom funkcijom (engl. *radial-basis function networks*; RBF) učenje se interpretira kao problem aproksimacije funkcije s više argumenata u odnosu na višeslojne linearne mreže koje koriste algoritam povratne propagacije pogreške i učenje interpretiraju kao problem optimizacije (minimalizacije srednje kvadratne pogreške). Funkcija ulaz-izlaz je funkcija koju treba aproksimirati i definirana parovima za učenje. Struktura radijalne mreže sadrži ulazni sloj, skriveni sloj koji ima drugačiju ulogu nego kod višeslojnih mreža i izlazni sloj. Transformacija od ulaznog sloja do skrivenog sloja je nelinearna dok je transformacija od skrivenog sloja do izlaznog sloja je linearna.²¹



Slika 14. Prikaz mreže s radijalnom funkcijom

2.3.4. Procjena valjanosti modela

Često nije dovoljno korisna samo spoznaja da u nekom sustavu postoji efekt zbog čega je potrebno moći predvidjeti i procijeniti te efekte, odnosno potrebno je razviti model. Kod teorijskog modela podaci slijede teoriju te se koriste poznati teorijski zakoni i principi. Međutim, empirijski model nije napravljen na temelju teorijskih činjenica. Osnova empirijskog modela je samo fitanje podataka i vrlo često se mora pretpostaviti kako se trend nastavlja i izvan eksperimentalnog područja. Potrebno je minimizirati razliku između izmjerene i predviđene vrijednosti.²² Točnost pokazuje odstupanje rezultata mjerenja i prave, odnosno referentne vrijednosti. Često prava vrijednost nije poznata pa se ne može odrediti točnost mjerenja, no upotrebom statističkih metoda moguće je odrediti interval unutar kojeg se najvjerojatnije nalazi prava vrijednost.²³ Slučajne i sustavne pogreške određuju kvalitetu mjernog procesa.²⁴ Slučajna pogreška određuje preciznost te se ne može odrediti niti ukloniti, ali se može statistički smanjiti. Sustavna pogreška određuje točnost. Mogu se odrediti uzroci sustavne pogreške te se mogu ukloniti.²⁵

3. Eksperimentalni dio

Klasifikacija podataka rađena je u programu *Statistica*, a eksperiment je proveden po načelu *potpunog faktorskog planiranja* (engl. *full factorial design*). Za potrebe izrade modela, eksperimentalni su podaci nasumice podijeljeni u tri seta: set za trening, test i validaciju. Korištene su neuronske mreže s radijalnom funkcijom (RBF) i višestruki linearni perceptron (MLP).

Kod RBF mreža broj neurona mijenjan je od 4 do 12 s korakom 2, mreža je trenirana 200 puta te je predefiniran broj od 200 ciklusa treniranja

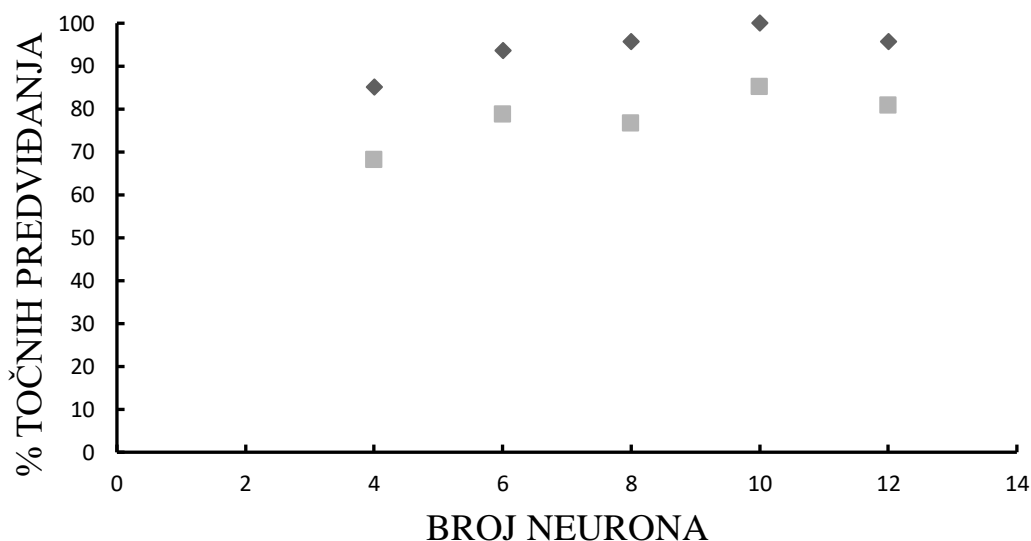
Kod MLP mreža mijenjan je broj neurona, aktivacijska funkcija i trening funkcija. Broj neurona mijenjan je između 4 i 12 s korakom 2. Kao aktivacijske funkcije ispitane su: tangens hiperbolna (*Tanh*), funkcija $x=y$ (*Identity*), logistička (*Logistic*), eksponencijalna (*Exponential*) i sinusna (*Sine*) funkcija, a kao trening funkcije odabrane su: funkcija konjugiranog gradijenta (*Conjugate gradient*), Broyden–Fletcher–Goldfarb–Shanno funkcija (BFGS), i funkcija padajućeg gradijenta (*Gradient descent*). Mreža je trenirana 100 puta, te je predefiniran broj od 500 ciklusa treniranja.

Od ukupno 80 treniranih mreža, odabrana je ona s najboljim značajkama.

4. Rezultati i rasprava

4.1. RBF

Na slici 15 prikazani su podaci predviđeni uporabom mreža s radijalnom funkcijom. Najbolji rezultati na setu podataka za treniranje ostvareni su mrežom s 10 neurona. Ujedno pokazalo se da takva mreža daje i najbolje predviđanje i podataka iz seta za testiranje, čime je potvrđeno da uočena visoka sposobnost predviđanja na setu za testiranje nije bila posljedica pretreniranosti mreže.



Slika 15. Grafički prikaz podataka dobivenih upotrebom mreže s radijalnom funkcijom.

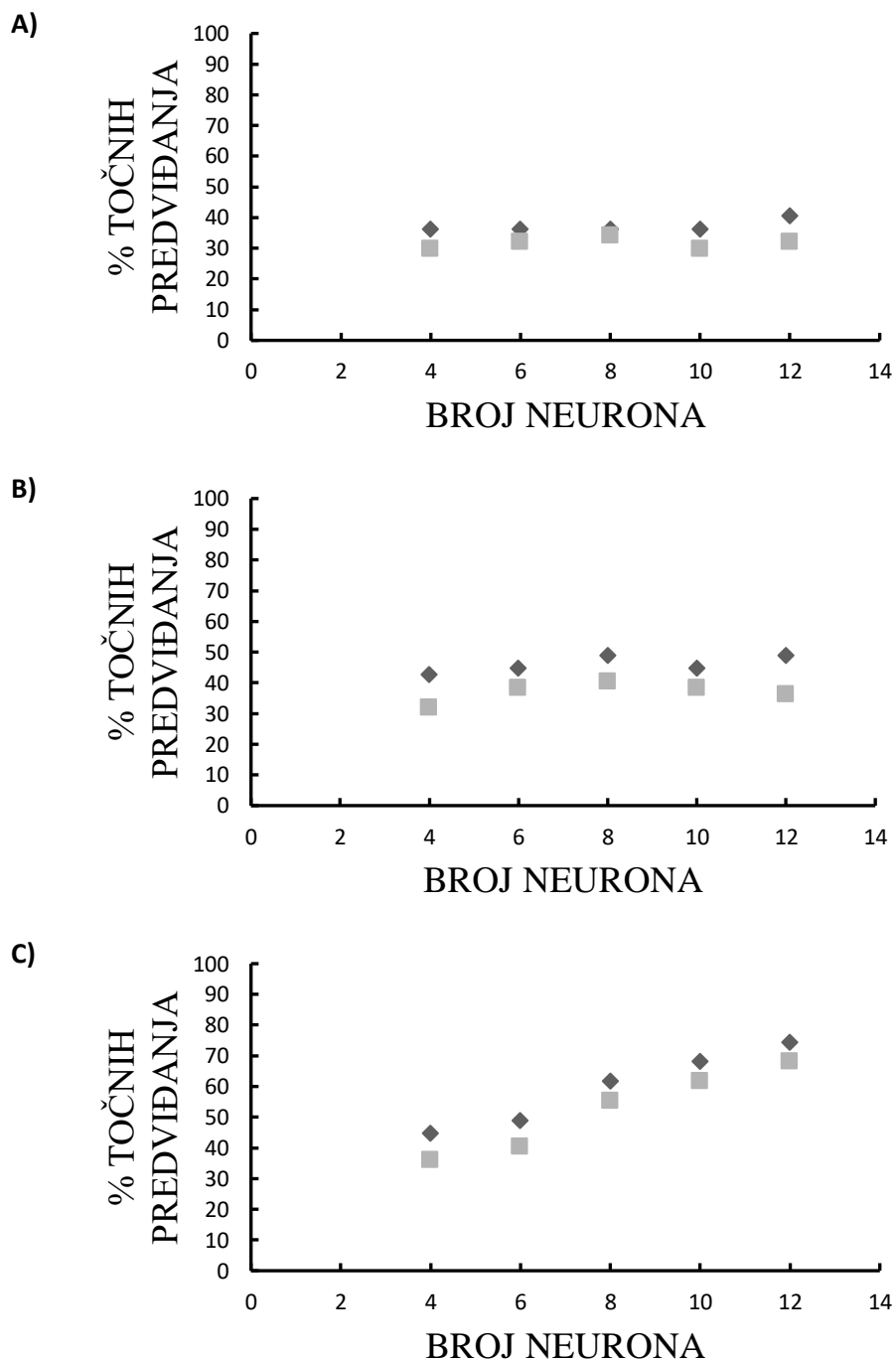
Rombovi predstavljaju podatke za treniranje, a kvadrati podatke za testiranje.

4.2. MLP

4.2.1. Aktivacija eksponencijalnom funkcijom

Na slici 16 prikazani su podaci predviđeni uporabom višeslojne mreže korištenjem eksponencijalne funkcije. Najbolji rezultati na setu podataka za treniranje ostvareni su mrežom s 12 neurona. Uporabom trening funkcije padajućeg gradijenta, pokazalo se da mreža s 12 neurona daje i najbolje predviđanje podataka iz seta za testiranje čime je potvrđeno da uočena visoka sposobnost predviđanja na setu za testiranje nije bila posljedica pretreniranosti mreže. Uporabom trening funkcija BFGS i konjugirani gradijent, najbolje predviđanje

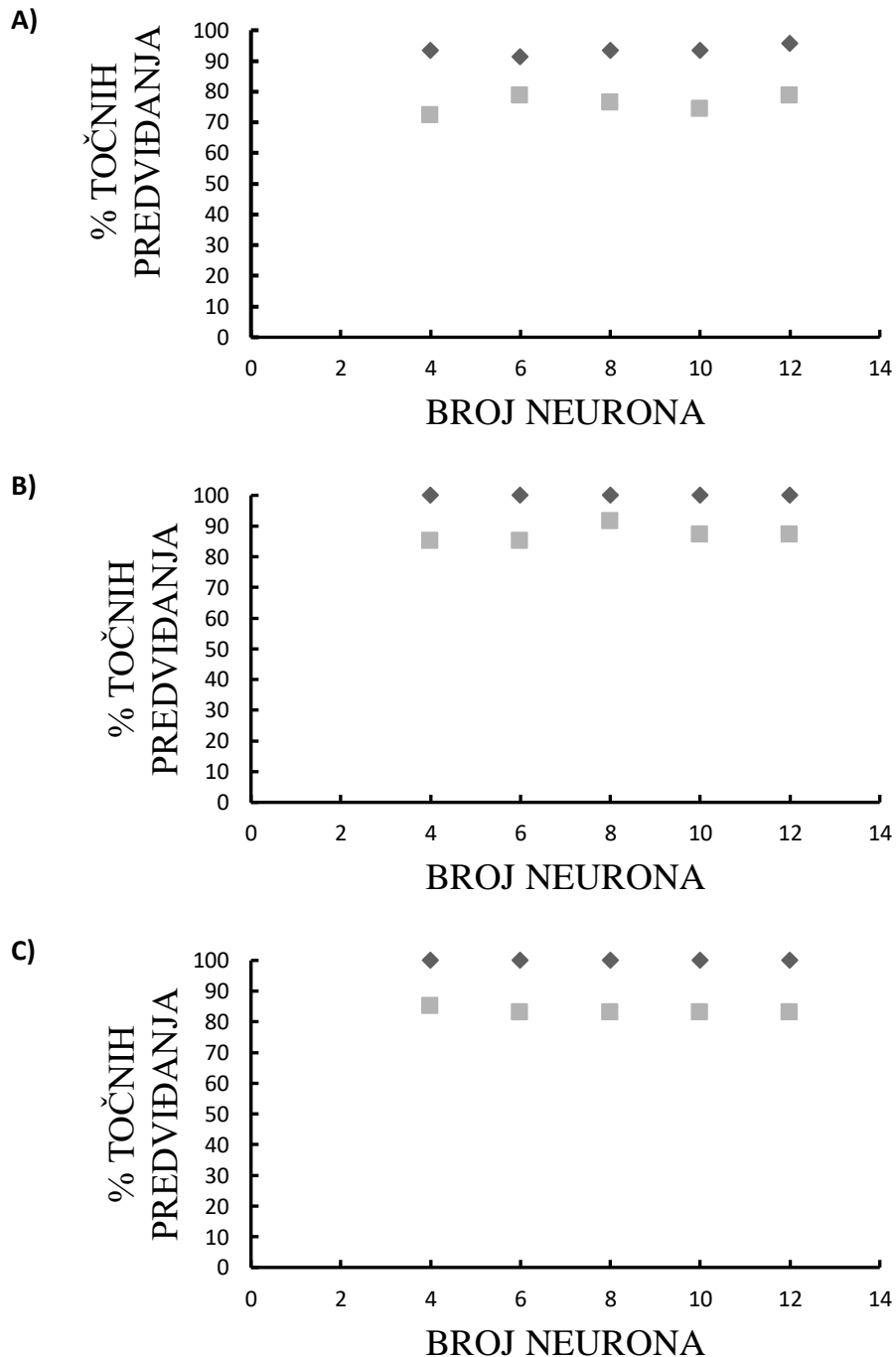
podataka iz seta za testiranje pokazala je mreža s 8 neurona. Na slici 16 B vidljivo je kako do pretreniranosti mreže dolazi nakon upotrebe 10 neurona.



Slika 16. Grafički prikazi rezultata dobivenih korištenjem eksponencijalne aktivacijske funkcije i trening funkcije: A) BFGS, B) konjugiranog gradijenta, C) funkcije padajućeg gradijenta. Rombovi predstavljaju podatke za treniranje, a kvadrati podatke za testiranje.

4.2.2. Aktivacija pravcem $y=x$

Na slici 17 prikazani su podaci MLP mreže uz aktivaciju pravcem $y=x$. Uz BFGS najbolji rezultati na setu za treniranje i testiranje ostvareni su uz 12 neurona. Podjednaki rezultati na setu za treniranje (neovisno o broju neurona) ostvareni su uz konjugirani i padajući gradijent (slike 17 B i C).

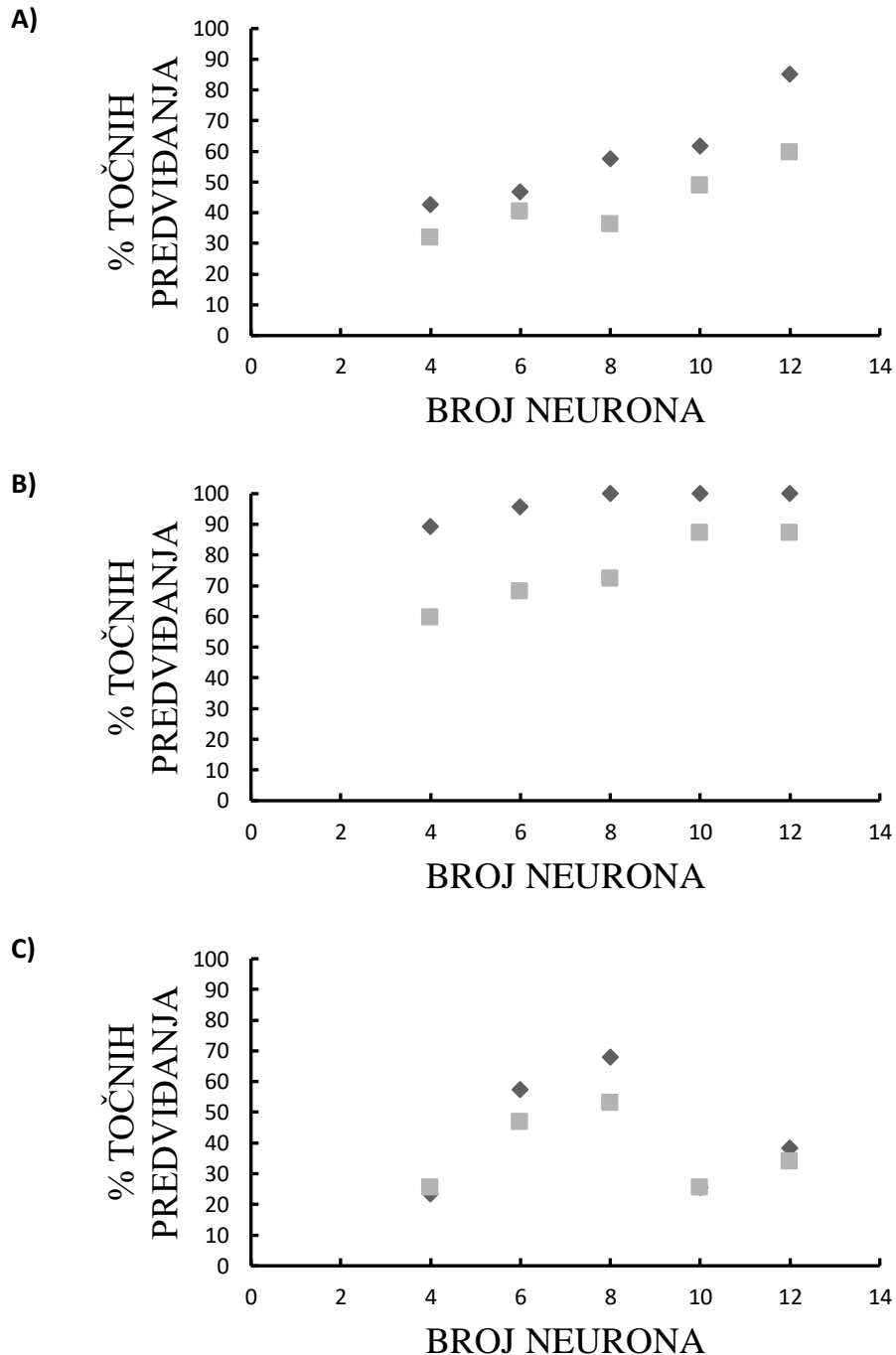


Slika 17. Grafički prikazi rezultata dobivenih korištenjem aktivacijske funkcije $y=x$ i trening funkcije: A) BFGS, B) konjugiranog gradijenta, C) funkcije padajućeg gradijenta.

Rombovi predstavljaju podatke za treniranje, a kvadrati podatke za testiranje.

U istim slučajevima, najbolje predviđanje na setu za testiranje pokazale su mreže s 8 (konjugirani gradijent) i 4 neurona (padajući gradijent). Na slici 17 A vidljivo je pretreniranje mreže za više od 6 neurona, a na slici 17 B za više od 8 neurona.

4.2.3. Aktivacija logističkom funkcijom



Slika 18. Grafički prikazi rezultata dobivenih korištenjem logističke aktivacijske funkcije i trening funkcije: A) BFGS, B) konjugiranog gradijenta, C) funkcije padajućeg gradijenta.

Rombovi predstavljaju podatke za treniranje, a kvadrati podatke za testiranje.

Na slici 18 prikazani su podaci predviđeni uporabom višeslojne mreže korištenjem logističke funkcije.

Uporabom trening funkcije BFGS i konjugirani gradijent, najbolji rezultati na setu podataka za treniranje, a ujedno i najbolje predviđanje podataka iz seta za testiranje, ostvareni su mrežom s 12 neurona.

Uporabom trening funkcije padajućeg gradijenta, pokazalo se da mreža s 8 neurona daje najbolje rezultate na setu podataka za treniranje i najbolje predviđanje podataka iz seta za testiranje.

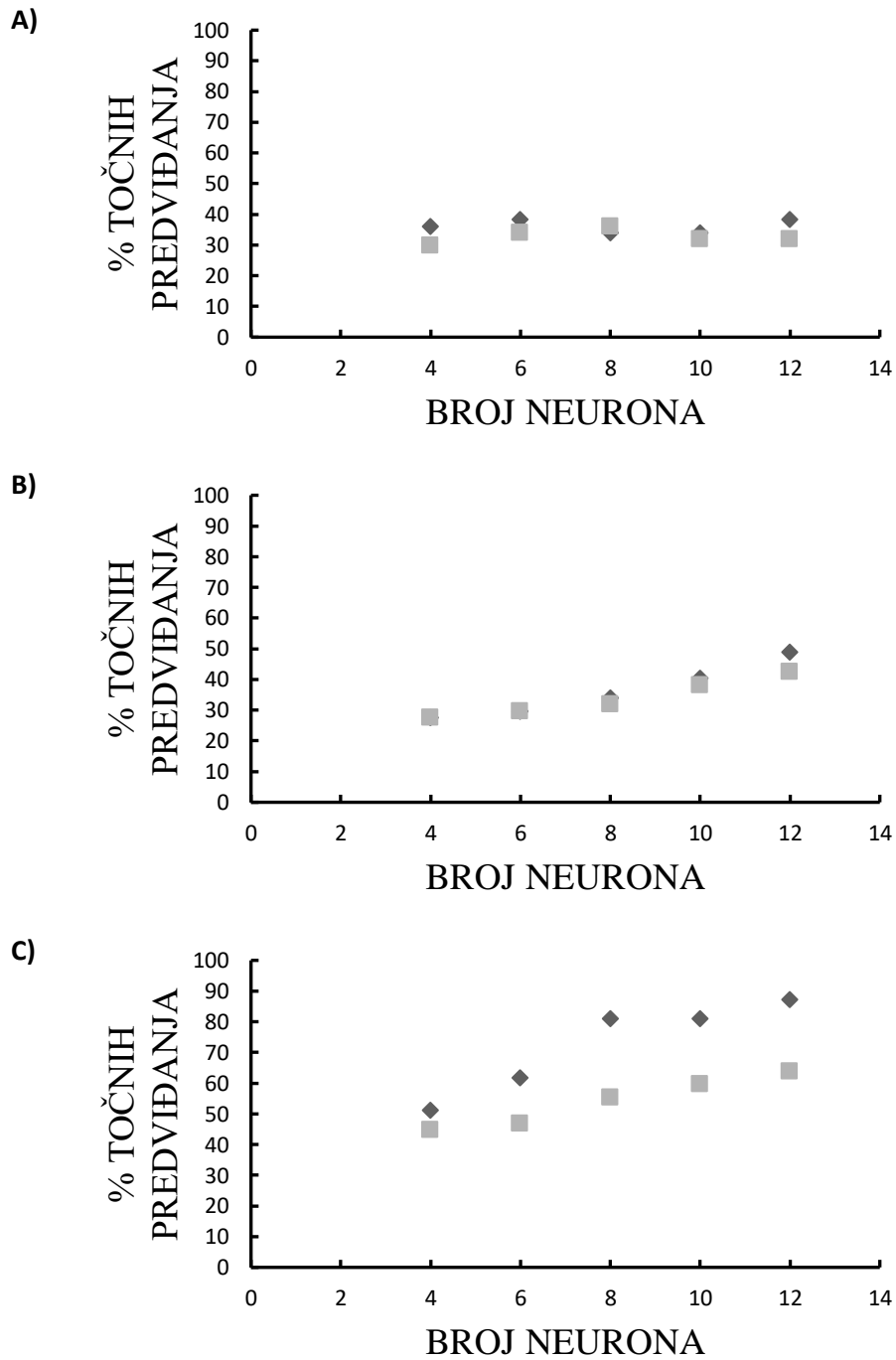
Na slici 18 A, do pretreniranosti mreže dolazi nakon upotrebe 6 neurona. Iz slike 18 C vidljivo je nepravilno ponašanje točnih predviđanja u ovisnosti o broju neurona te se ne može uočiti pretreniranost mreže. Stoga je odlučeno uzeti slučaj s 10 neurona za određivanje validacije jer je taj slučaj pokazao najmanju razliku u setu za treniranje i setu za test.

4.2.4. Aktivacija sinusnom funkcijom

Na slici 19 prikazani su podaci predviđeni uporabom višeslojne mreže korištenjem sinusne aktivacijske funkcije. Najbolji rezultati na setu podataka za treniranje ostvareni su mrežom s 12 neurona za sve tri trening funkcije.

Uporabom trening funkcije BFGS, najbolje predviđanje podataka iz seta za testiranje pokazala je mreža s 6 neurona, dok uporabom trening funkcija konjugiranog i padajućeg gradijenta najbolje predviđanje podataka iz seta za testiranje daje mreža s 12 neurona.

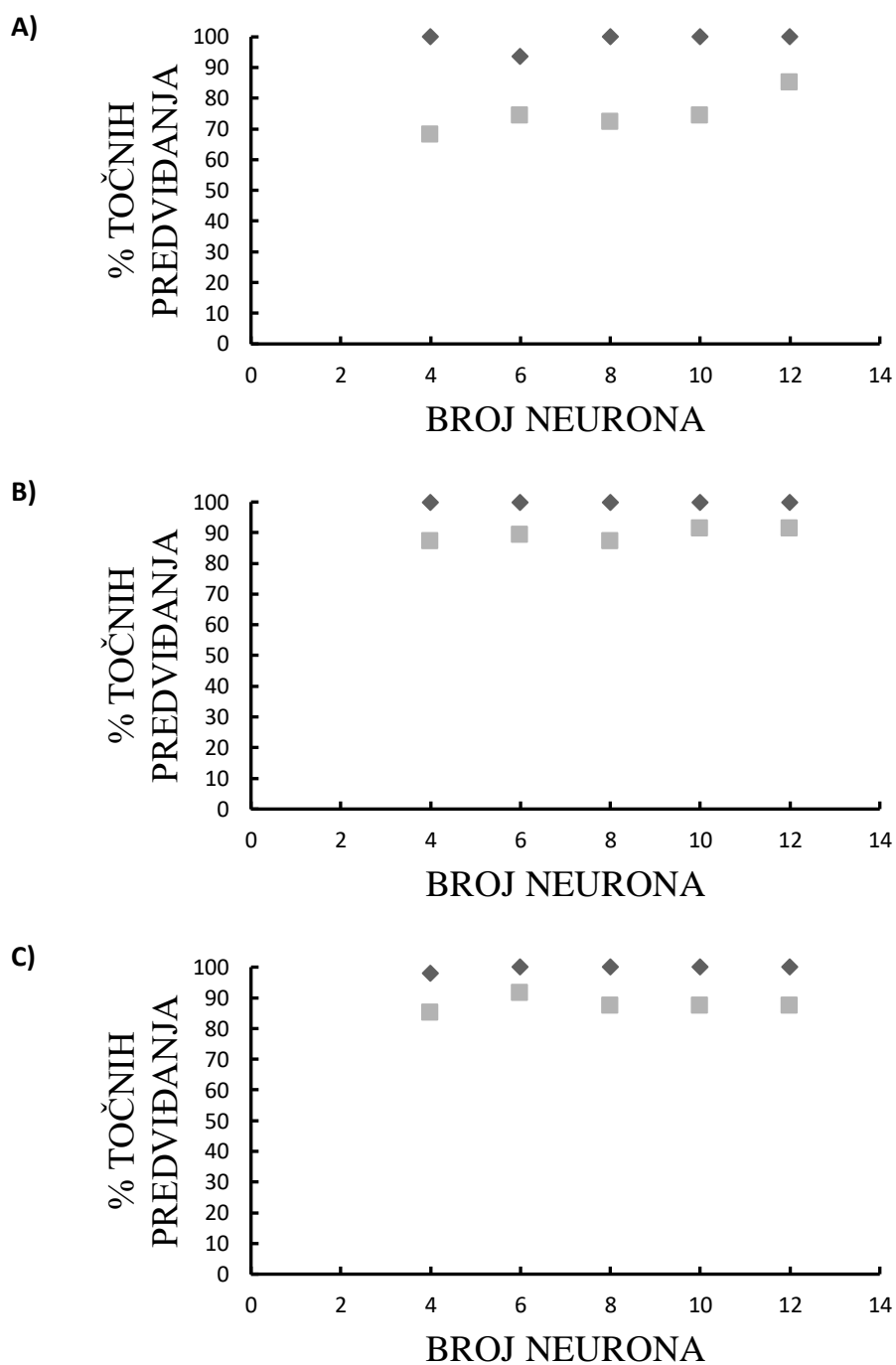
Slika 19 C pokazuje kako do pretreniranosti mreže dolazi odmah nakon uporabe 4 neurona.



Slika 19. Grafički prikazi rezultata dobivenih korištenjem sinusne aktivacijske funkcije i trening funkcije: A) BFGS, B) konjugiranog gradijenta, C) funkcije padajućeg gradijenta.

Rombovi predstavljaju podatke za treniranje, a kvadrati podatke za testiranje.

4.2.5. Tangens hiperbolna aktivacijska funkcija



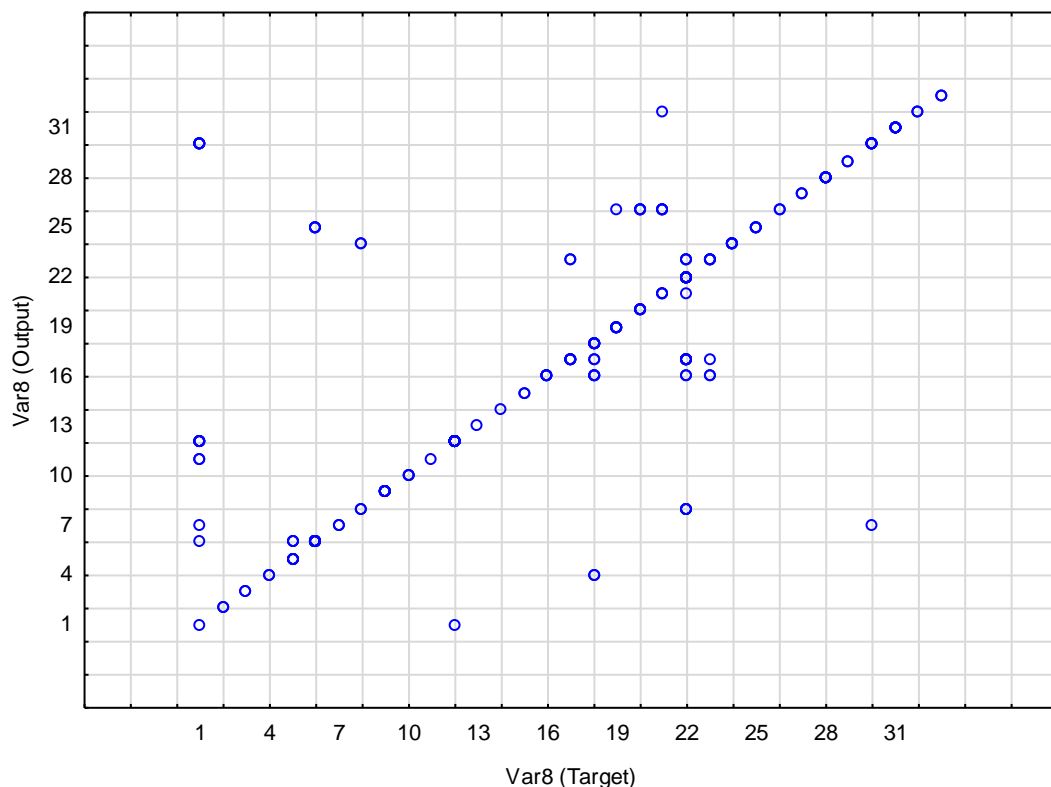
Slika 20. Grafički prikazi rezultata dobivenih korištenjem tangens hiperbolni aktivacijske funkcije i trening funkcije: A) BFGS, B) konjugiranog gradijenta, C) funkcije padajućeg gradijenta. Rombovi predstavljaju podatke za treniranje, a kvadrati podatke za testiranje.

Na slici 20 prikazani su podaci predviđeni uporabom višeslojne mreže korištenjem tangens hiperbolni aktivacijske funkcije. Najbolji mogući rezultati na setu podataka za treniranje ostvareni su mrežom sa svim brojevima neurona za sve tri trening funkcije. Uporabom trening funkcija BFGS i konjugirani gradijent, najbolje predviđanje podataka iz seta za testiranje pokazala je mreža s 12 neurona, dok uporabom trening funkcije padajućeg gradijenta najbolje predviđanje podataka iz seta za testiranje daje mreža s 6 neurona. Na slici 20 A vidljivo je kako do pretreniranja dolazi nakon uporabe 6 neurona.

Višeslojna mreža s 12 neurona, aktivacijskom funkcijom tangens hiperbolni i trening funkcijom konjugirani gradijent dalje najbolje rezultate na setu podataka za treniranje i najbolje predviđanje podataka iz seta za testiranje u odnosu na sve ostale aktivacijske funkcije. Uporabom takve mreže i potvrđeno je da uočena visoka sposobnost predviđanja na setu za testiranje nije posljedica pretreniranosti mreže.

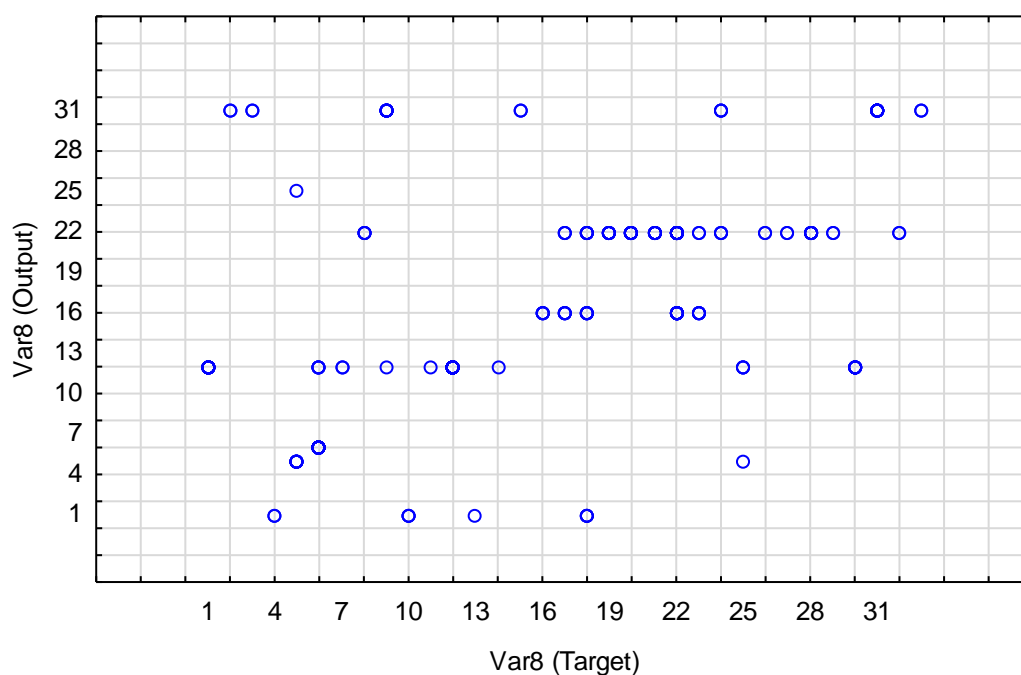
4.3. Validacija

4.3.1. RBF

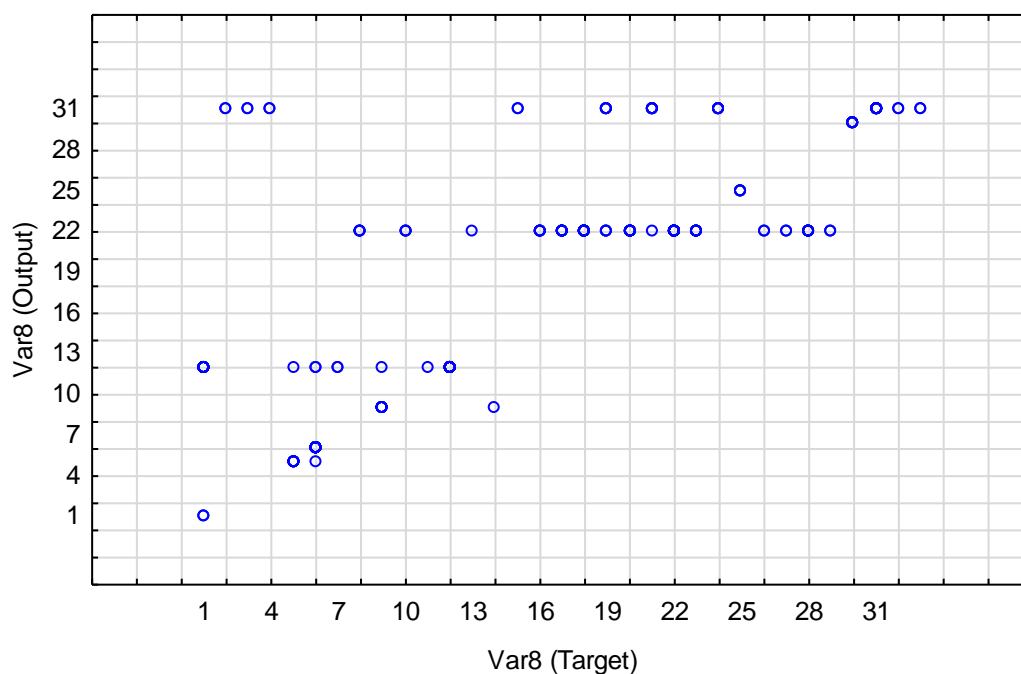


Slika 21. Točnost optimalne RBF mreže (10 neurona): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 74,63%. „Output“ predstavlja predviđene vrijednosti, a „target“ mjerene vrijednosti.

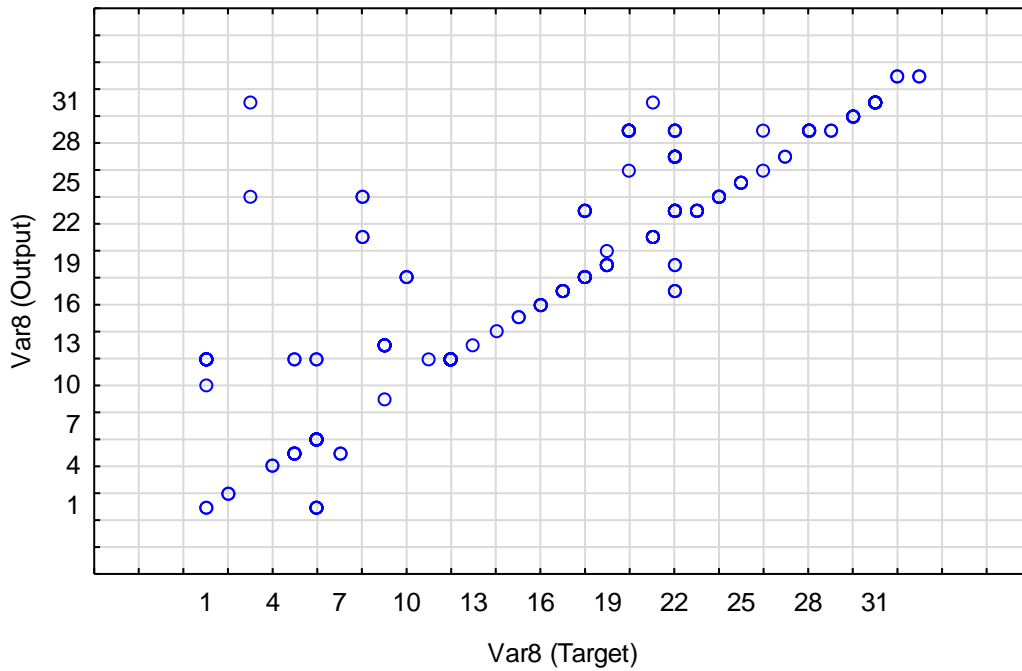
4.3.2. MLP



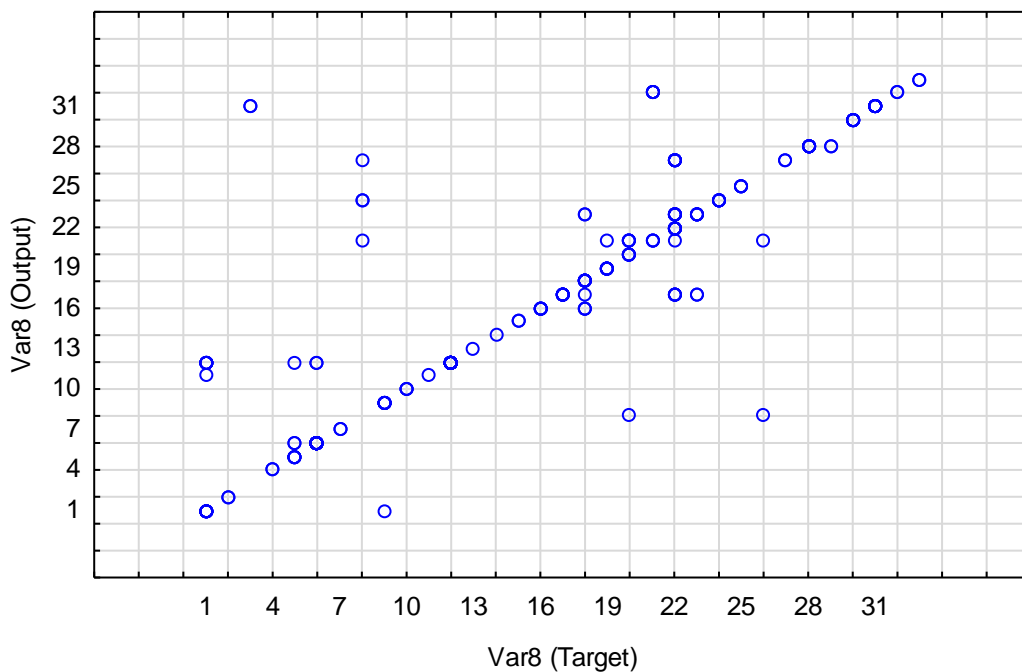
Slika 22. Točnost optimalne MLP mreže (8 neurona, eksponencijalna aktivacijska funkcija, BFGS trening funkcija): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 42,65 %.



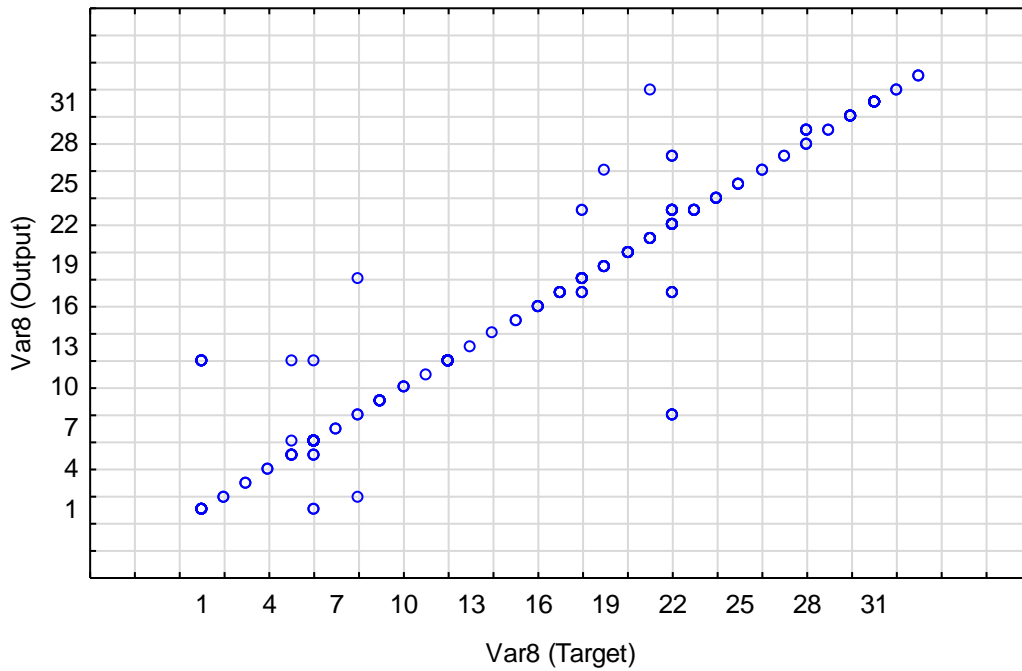
Slika 23. Točnost optimalne MLP mreže (10 neurona, eksponencijalna aktivacijska funkcija, trening funkcija konjugirani gradijent): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 53,68 %.



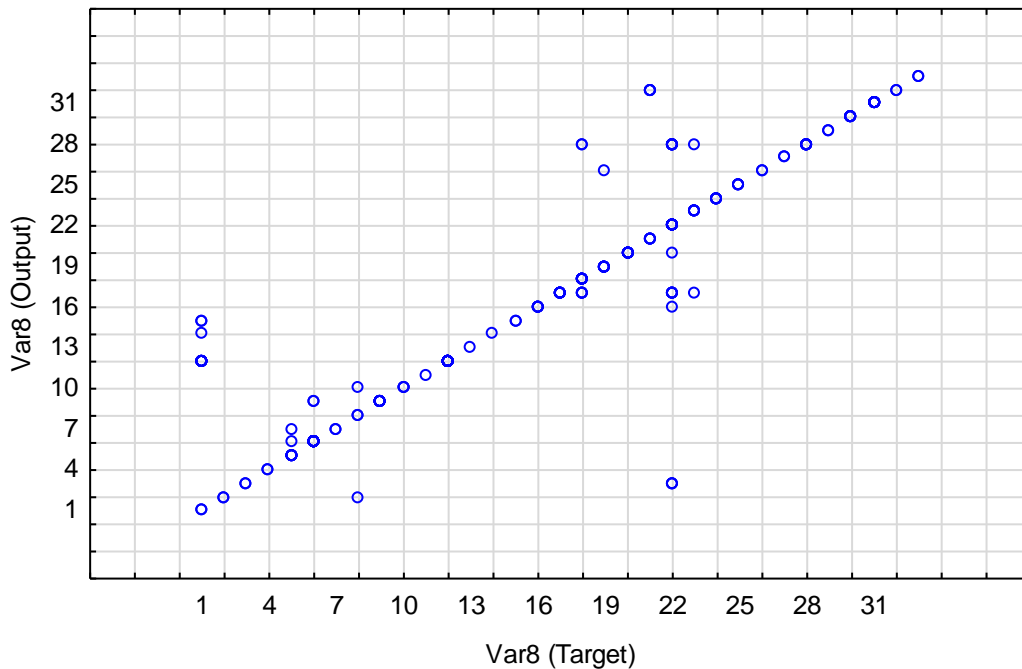
Slika 24. Točnost optimalne MLP mreže (12 neurona, eksponencijalna aktivacijska funkcija, trening funkcija padajućeg gradijenta): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 60,66 %.



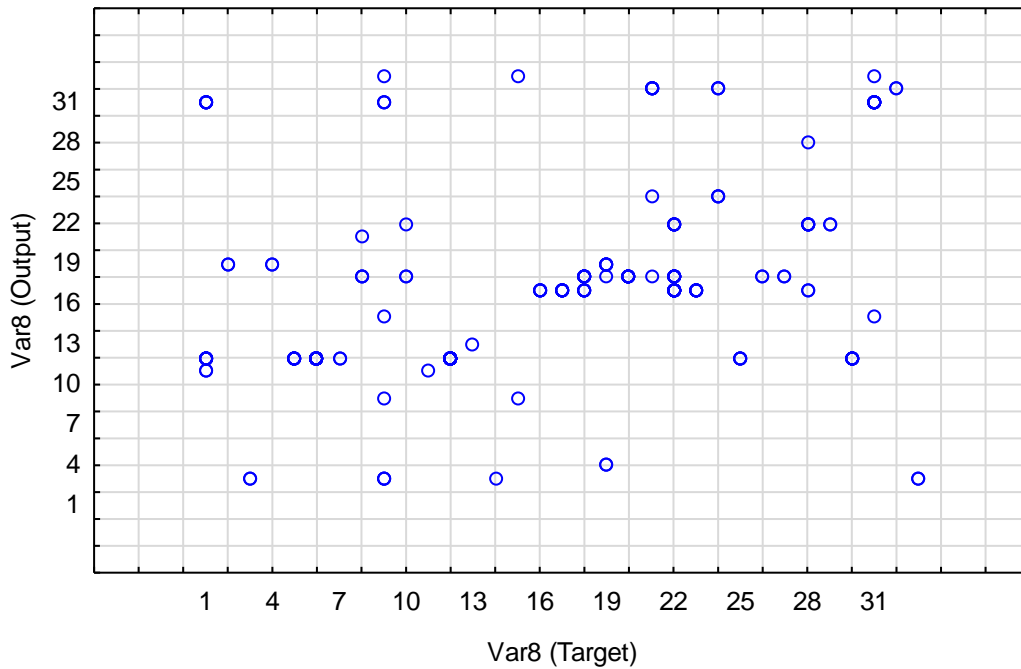
Slika 25. Točnost optimalne MLP mreže (6 neurona, $y=x$ aktivacijska funkcija, trening funkcija BFGS): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 75,74%.



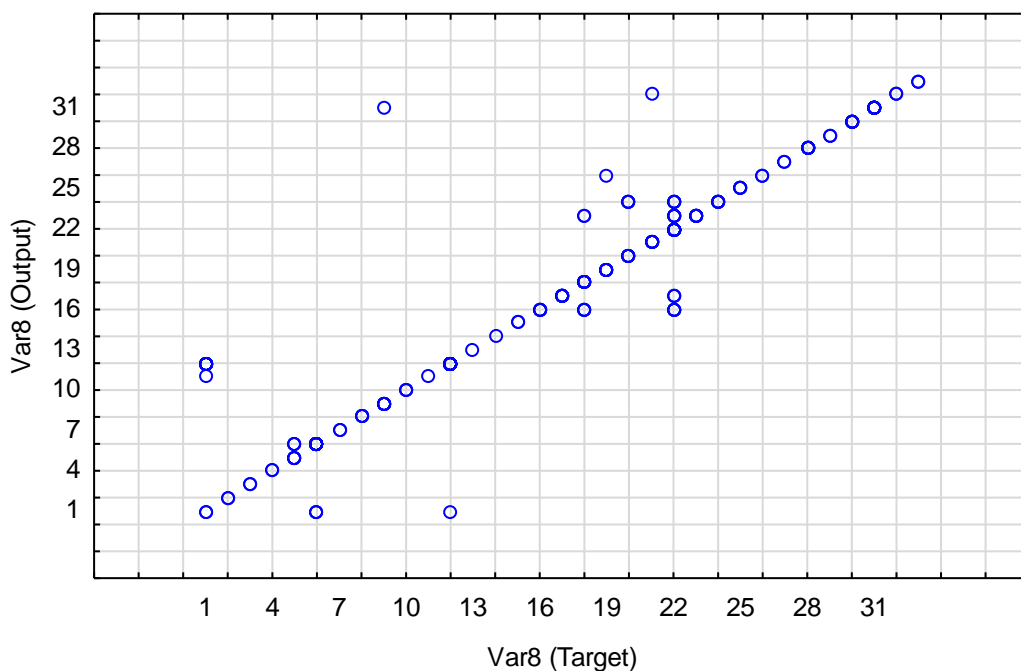
Slika 26. Točnost optimalne MLP mreže (8 neurona, $y=x$ aktivacijska funkcija, trening funkcija konjugirani gradijent): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 81,25 %.



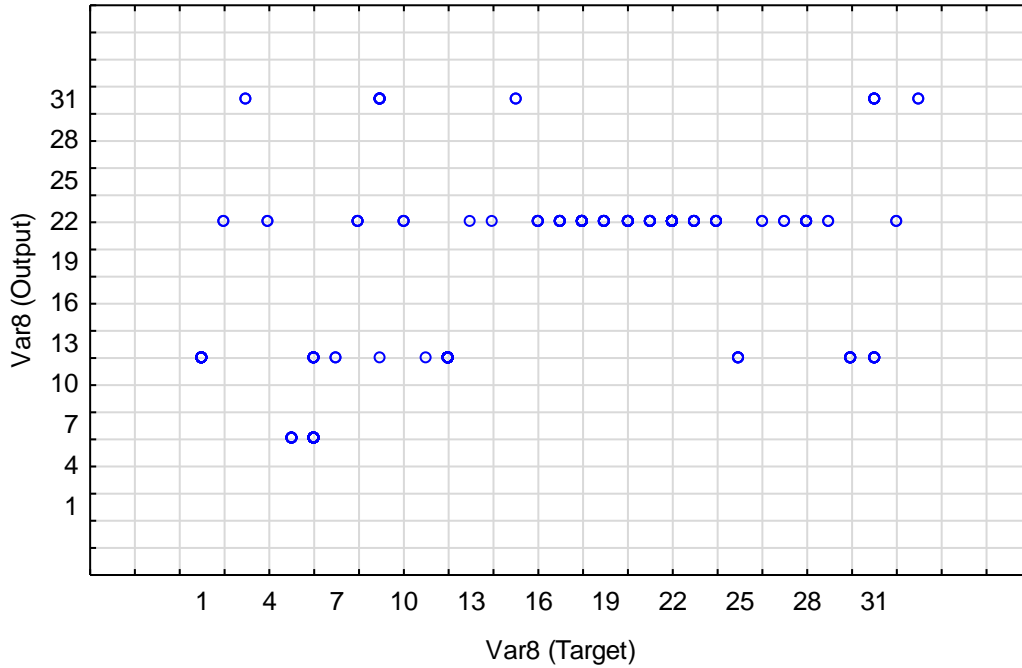
Slika 27. Točnost optimalne MLP mreže (4 neurona, $y=x$ aktivacijska funkcija, trening funkcija padajućeg gradijenta): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 82,35 %.



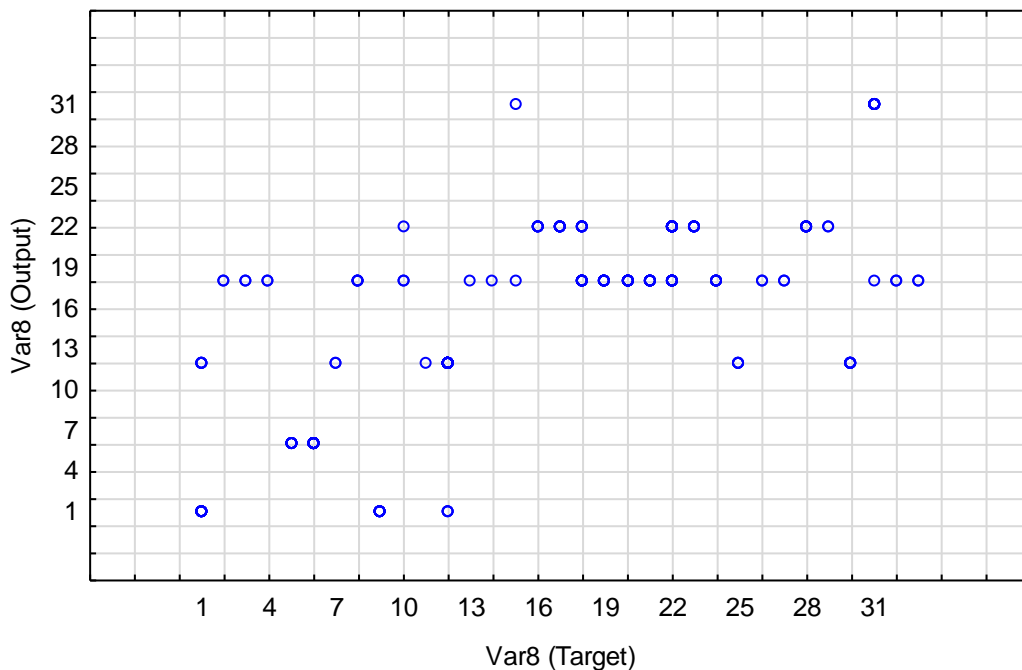
Slika 28. Točnost optimalne MLP mreže (6 neurona, logistička aktivacijska funkcija, trening funkcija BFGS): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 33,09 %.



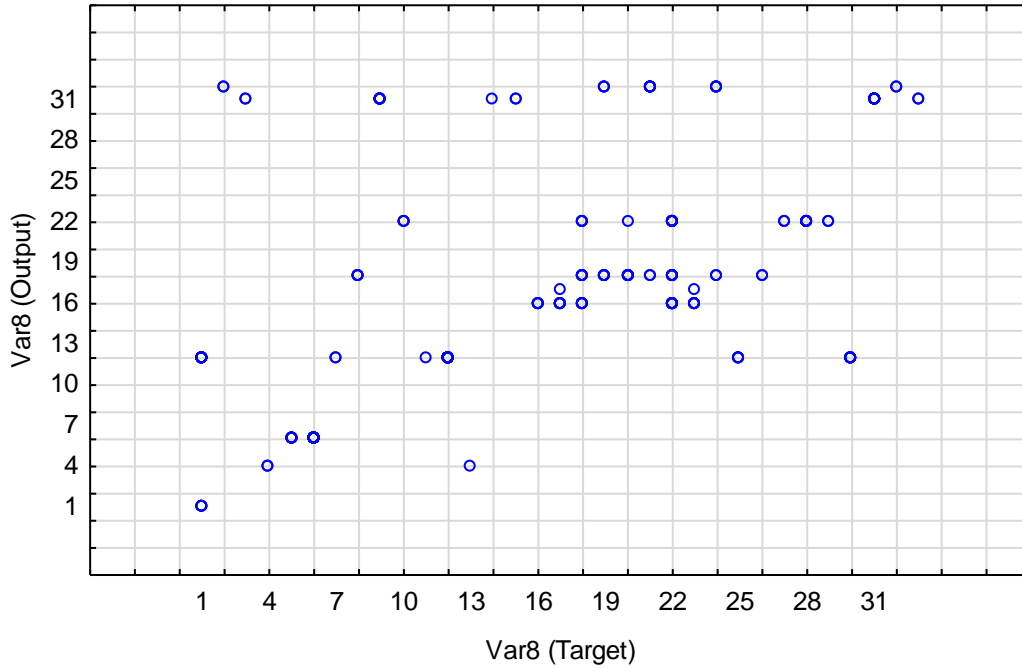
Slika 29. Točnost optimalne MLP mreže (12 neurona, logistička aktivacijska funkcija, trening funkcija konjugirani gradijent): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 80,15 %.



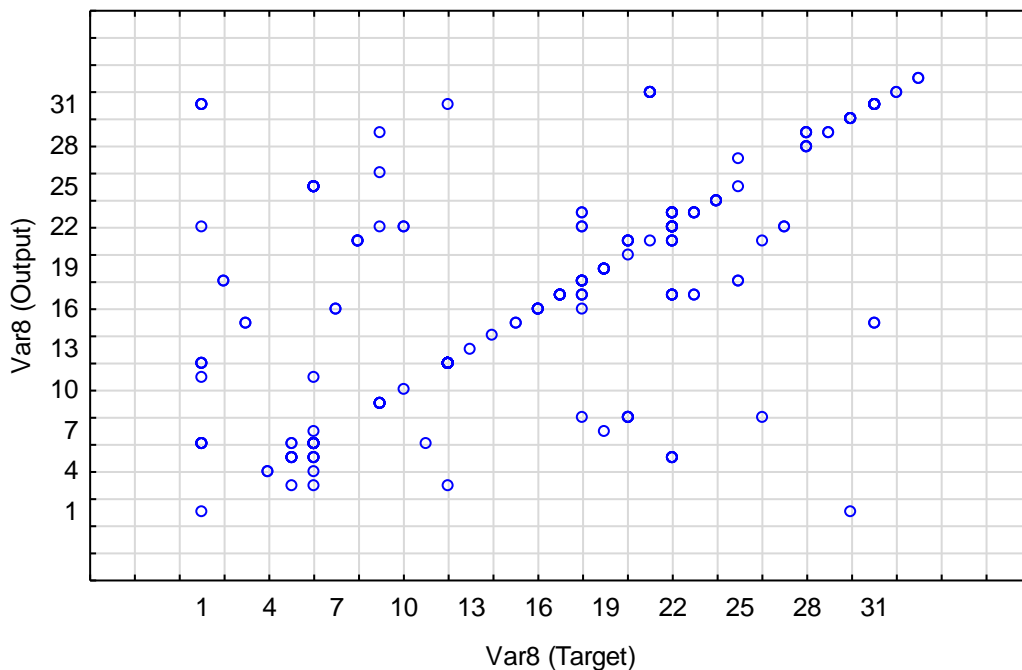
Slika 30. Točnost optimalne MLP mreže (10 neurona, logistička aktivacijska funkcija, trening funkcija padajućeg gradijenta): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 38,24 %.



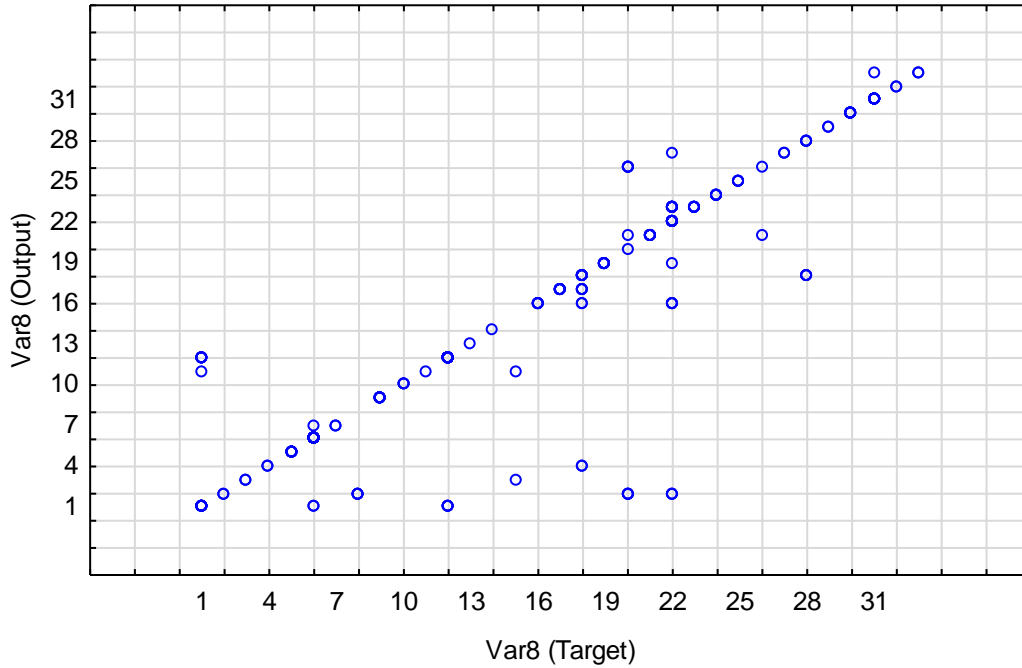
Slika 31. Točnost optimalne MLP mreže (10 neurona, sinusna aktivacijska funkcija, trening funkcija BFGS): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 80,15 %.



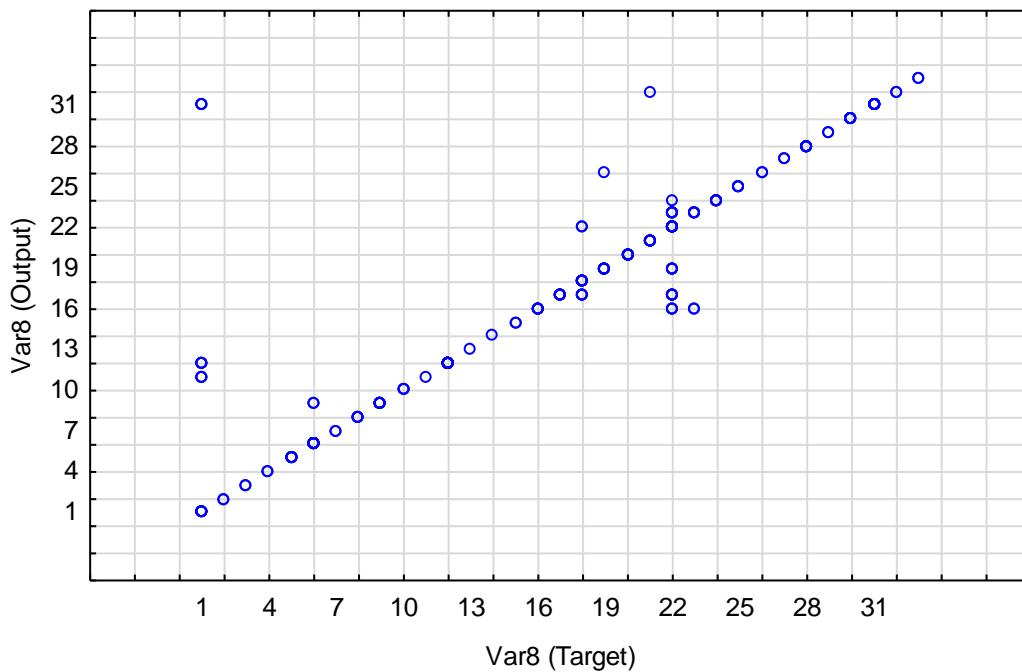
Slika 32. Točnost optimalne MLP mreže (10 neurona, sinusna aktivacijska funkcija, trening funkcija konjugirani gradijent): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 43,01 %.



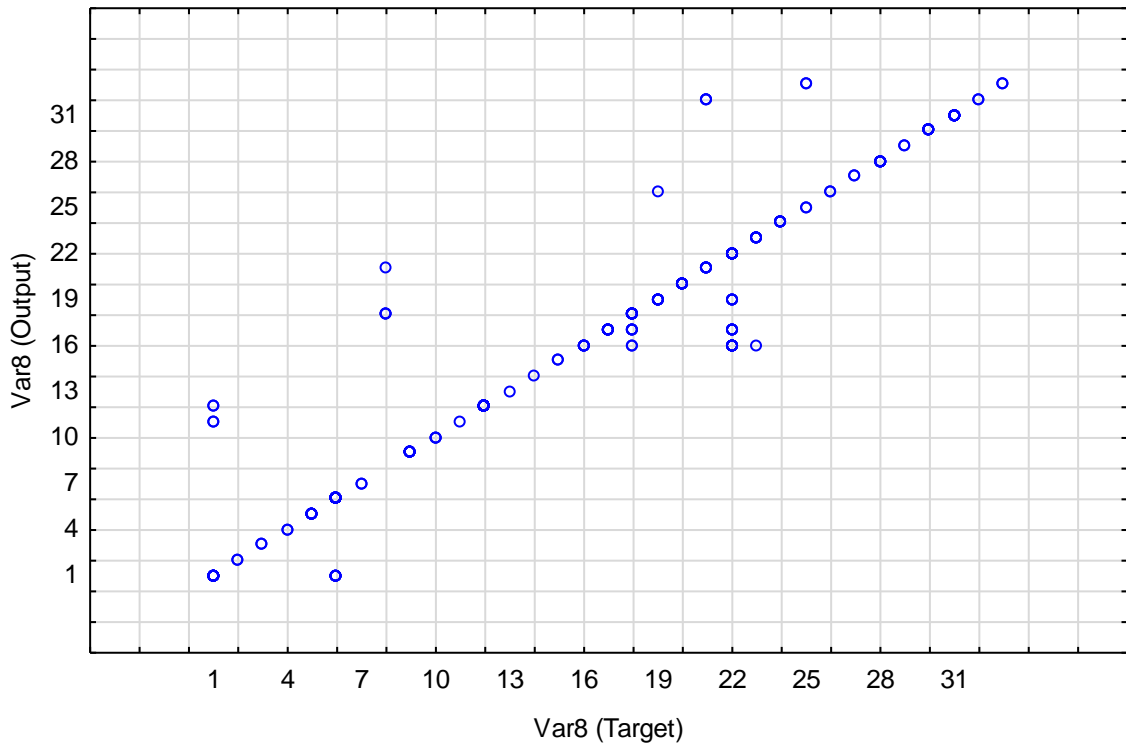
Slika 33. Točnost optimalne MLP mreže (12 neurona, sinusna aktivacijska funkcija, trening funkcija padajućeg gradijenta): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 56,62 %.



Slika 34. Točnost optimalne MLP mreže (12 neurona, tangens hiperbolni aktivacijska funkcija, trening funkcija BFGS): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 77,57 %.



Slika 35. Točnost optimalne MLP mreže (12 neurona, tangens hiperbolni aktivacijska funkcija, trening funkcija konjugirani gradijent): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 84,93 %.



Slika 36. Točnost optimalne MLP mreže (6 neurona, tangens hiperbolni aktivacijska funkcija, trening funkcija padajućeg gradijenta): usporedbe modelom klasificiranih i poznatih grupiranosti uzoraka narkotika. Uspješnost predviđanja iznosi 83,46 %.

Validacijom je pokazano kako su sve upotrijebljene mreže primjenjive na vanjskom setu podataka. Mreža s 12 neurona, aktivacijskom funkcijom tangens hiperbolni i trening funkcijom konjugirani gradijent koje je odabrana kao najbolja mreža ujedno daje i najbolje rezultate validacije, odnosno daje najveću uspješnost predviđanja od 84,93 %.

5. Zaključak

Kako bi borba protiv zlouporabe psihoaktivnih tvari bila brža i uspješnija potrebna je brža i kvalitetnija karakterizacija psihoaktivnih tvari. Poznavanjem profila psihoaktivnih tvari, moguće je pronaći puteve trgovine te u krajnjem slučaju i locirati izvor u kojem se droga ilegalno pripravlja. Upotrebom umjetnih neuronskih mreža postignuta je brza i kvalitetna klasifikacija podataka. Bilo je potrebno odabrati tip mreže, broj neurona, aktivacijsku i trening funkciju koji će dati najbolje rezultate. Upravo je višeslojna mreža s 12 neurona, aktivacijskom funkcijom tangens hiperbolni te trening funkcijom konjugirani gradijent ispunila taj zadatak. Na setu podataka za treniranje, mreža je imala 100% točnih predviđanja, a na setu podataka za testiranje postotak točnih predviđanja iznosio je 91,49%. Uspješnost predviđanja prilikom validacije iznosi 84,93%.

Istraživanja su pokazala da se matematički modeli mogu uspješno primijeniti u klasifikaciji psihoaktivnih tvari te može biti novi moćan alat u suzbijanju kriminala opojnih droga.

6. Popis pokrata

UN	Ujedinjeni narodi
UNODC	Ured Ujedinjenih naroda za droge i kriminal
HIV	Virus humane imunodeficijencije
KNN	Klasifikator k najbližeg susjeda
HCA	Hijerarhijska klaster analiza
PCA	Analiza glavnih komponenti
FA	Faktorska analiza
SIMCA	Neovisno modeliranje analogije grupa
ANN	Umjetne neuronske mreže
MLP	Višeslojni linearni perceptron
RBF	Mreža s radijalnom funkcijom
BFGG	Broyden–Fletcher–Goldfarb–Shanno funkcija

7. Literatura

1. United Nations, Office on Drugs and Crime, Methods for Impurity Profiling of Heroin and Cocaine, Manual for use by National Drug Testing Laboratories, Laboratory and Scientific Section, New York, 2005.
2. https://narodne-novine.nn.hr/clanci/sluzbeni/2001_12_107_1756.html (pristupljeno 23. travnja 2019.)
3. <https://www.scribd.com/doc/82861898/droga> (pristupljeno 23. travnja 2019.)
4. S. Brlas, *Terminološki opisni rječnik ovisnosti*, 5. izdanje, 2. izmijenjeno, Virovitica, 2016.
(http://www.zzjzvpz.hr/hr/sadrzaj/djelatnost/1/publikacije/terminoloski_rjecnik_peto_i_zdanje_drugo_izmijenjeno.pdf)
5. L. Junaković, *Utjecaj zlouporabe psihoaktivnih tvari na osobno i društveno funkcioniranje pojedinca*, Završni rad, Sveučilišni odjel zdravstvenih studija, Sveučilište u Splitu, Split, 2014.
6. S. Raić-Malić, *Alkaloidi*, predavanje iz kolegija „Kemija prirodnih spojeva“, Fakultet kemijskog inženjerstva i tehnologije, Zagreb, 2015.
7. <https://drogeiovisnosti.gov.hr/djelokrug/ovisnosti-i-vezane-teme/droge-i-ovisnost/vrste-droga/kokain/1009> (pristupljeno 23. travnja 2019.)
8. <https://drogeiovisnosti.gov.hr/djelokrug/ovisnosti-i-vezane-teme/droge-i-ovisnost/vrste-droga/heroin/1010> (pristupljeno 23. travnja 2019.)
9. United Nations International Drug Control Programme, *Drug Characterization/Impurity Profiling, Background and Concepts*, Manual for Use by National Law Enforcement Authorities and Drug Testing Laboratories, Scientific Section, New York, 2001.
10. M. Otto, *Chemometrics, Statistics and Computer Application in Analytical Chemistry*, 3. izdanje, Wiley, Verlag, 2017.
11. <https://www.geeksforgeeks.org/basic-concept-classification-data-mining/> (pristupljeno 25. travnja 2019.)
12. T. Nu Phyu, *Survey of Classification Techniques in Data Mining*, Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol I, IMESC 2009, Hong Kong, 2009.
(<https://pdfs.semanticscholar.org/ef65/86c6f6f188f0d9dffef67245a570ed072c9e.pdf>)

13. V. Kolobara, *Rješavanje klasifikacijskih problema evolucijom stabala odluke*, Završni rad, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu, Zagreb, 2016.
14. T. Bolanča, *Klasifikacijske metode*, predavanje iz kolegija „Kemometrija“, Fakultet kemijskog inženjerstva i tehnologije, Sveučilište u Zagrebu
(<https://moodle.srce.hr/2017-2018/course/view.php?id=20853>)
15. R. A. Dunne, *A Statistical Approach to Neural Networks for Pattern Recognition*, Wiley-Interscience, A John Wiley & Sons, Inc., Publication, New Jersey, 2007.
16. S. Lončarić, *Neuronske mreže: Uvod*, predavanje iz kolegija „Neuronske mreže“, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu
([https://www.fer.unizg.hr/_download/repository/01-Uvod\[3\].pdf](https://www.fer.unizg.hr/_download/repository/01-Uvod[3].pdf))
17. S. Dumančić, *Neuronske mreže*, Diplomski rad, Odjel za fiziku, Sveučilište Josipa Jurja Strossmayera u Osijeku, Osijek 2014.
18. T. Bolanča, *Umjetne neuronske mreže*, predavanje iz kolegija „Kemometrija“, Fakultet kemijskog inženjerstva i tehnologije, Sveučilište u Zagrebu
(<https://moodle.srce.hr/2017-2018/course/view.php?id=20853>)
19. S. Lončarić, *Neuronske mreže: Proces učenja*, predavanje iz kolegija „Neuronske mreže“, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu
(https://www.fer.unizg.hr/_download/repository/02-ProcesUcenja.pdf)
20. S. Lončarić, *Neuronske mreže: Višeslojni perceptron*, predavanje iz kolegija „Neuronske mreže“, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu
(https://www.fer.unizg.hr/_download/repository/06-ViseslojniPerceptron.pdf)
21. S. Lončarić, *Neuronske mreže: Radijalne mreže*, predavanje iz kolegija „Neuronske mreže“, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu
([https://www.fer.unizg.hr/_download/repository/07-RadijalneMreze-1s\[1\].pdf](https://www.fer.unizg.hr/_download/repository/07-RadijalneMreze-1s[1].pdf))
22. T. Bolanča, *Osnove modeliranja*, predavanje iz kolegija „Kemometrija“, Fakultet kemijskog inženjerstva i tehnologije, Sveučilište u Zagrebu
(<https://moodle.srce.hr/2017-2018/course/view.php?id=20853>)
23. http://www.phy.pmf.unizg.hr/~pzugec/FP/obrada_rezultata.pdf (pristupljeno 31. svibnja 2019.)
24. E. Mehić, *Određivanje sastavnica mjerne nesigurnosti i validacija korozijskog ispitivanja u slanoj komori*, Završni rad, Fakultet kemijskog inženjerstva i tehnologije, Sveučilište u Zagrebu, Zagreb, 2016.

25. T. Bolanča, *Osnove statistike u kemometriji*, predavanje iz kolegija „Kemometrija“, Fakultet kemijskog inženjerstva i tehnologije, Sveučilište u Zagrebu (<https://moodle.srce.hr/2017-2018/course/view.php?id=20853>)

8. Životopis

Petra Tominac [REDACTED] Pohađala je XV. gimnaziju poznatiju kao MIOC nakon čega je upisala preddiplomski studij Primijenjene kemije na Fakultetu kemijskog inženjerstva i tehnologije Sveučilišta u Zagrebu. Nakon preddiplomskog studija upisala je diplomski studij Primijenjene organske kemije na istom Fakultetu koji i završava. Tijekom studija odradila je praksu na Institutu „Ruđer Bošković“ u laboratoriju za fizikalno-organsku kemiju i Erasmus+ stručnu praksu u sklopu koje je dva i pol mjeseca provela u laboratoriju za organsku kemiju na Sveučilištu Pannonia u Veszpremu (Mađarska). Sudjelovala je u istraživanju koje se provodilo u sklopu znanstveno-istraživačkog projekta „Ciljana izrada prototipa vlaknastog nosača za uzgoj tkivnih stanica kombiniranim elektroispredanjem –COMBOELECTROSPUN“, IP-2016-06-6878, financiran od Hrvatske zaklade za znanost. Za sudjelovanje u tom istraživanju dobila je 2017. Rektorovu nagradu za rad pod nazivom „Izrada i karakterizacija elektroispredenih nosača s uzrokovanom topografijom za uzgoj ljudskih stanica kože (ili oka)“. Završni rad „Fotokatalitička razgradnja metilenskog plavila na titanatnim nanocjevčicama“ izradila je na Zavodu za anorgansku kemijsku tehnologiju i nemetale pod vodstvom prof. Stanislava Kurajice Uz studiranje radila je studentske poslove u DM-drogerie markt d.o.o i u ZARA d.o.o. Aktivno je volontirala u Udruzi Mudrica i sudjelovala u župnim instrukcijama za učenike osnovnih i srednjih škola.