

Usporedba unaprijednih neuronskih mreža s jednim i dva skrivena sloja u modeliranju linearnog dinamičkog sustava

Milas, Krešimir

Undergraduate thesis / Završni rad

2015

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:235:176444>

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

Download date / Datum preuzimanja: **2024-11-02**

Repository / Repozitorij:

[Repository of Faculty of Mechanical Engineering and Naval Architecture University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

Završni rad

Krešimir Milas

Zagreb, 2015.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

Završni rad

Usporedba unaprijednih neuronskih mreža s jednim i dva skrivena sloja u modeliranju linearnog dinamičkog sustava

Mentor:
Prof. dr. sc. Dubravko Majetić

Student:
Krešimir Milas

Zagreb, 2015.



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE



Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za završne ispite studija strojarstva za smjerove:
proizvodno inženjerstvo, računalno inženjerstvo, industrijsko inženjerstvo i menadžment, inženjerstvo
materijala i mehatronika i robotika

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum	Prilog
Klasa:	
Ur.broj:	

ZAVRŠNI ZADATAK

Student: **Krešimir Milas**

Mat. br.: 0035186297

Naslov rada na hrvatskom jeziku: **Usporedba unaprijednih neuronskih mreža s jednim i dva skrivena sloja u modeliranju linearnog dinamičkog sustava**

Naslov rada na engleskom jeziku: **Comparison of Feedforward Neural Networks With One and Two Hidden Layers in Modeling of Nonlinear Dynamic System**

Opis zadatka:

U ovom se radu na primjeru učenja linearnog dinamičkog sustava prvog reda treba pokazati kvaliteta učenja i generalizacija statičke neuronske mreže s jednim i dva skrivena sloja neurona. Za kvalitetnu usporedbu treba odabrati više različitih topologija dvoslojne mreže koje će imati isti broj sakrivenih neurona kao i neuronska mreža s jednim skrivenim slojem neurona.

U radu treba načiniti slijedeće:


1. Izvesti matematički model učenja po principu povratnog prostiranja pogreške za statičku neuronsku mrežu s jednim i dva skrivena sloja neurona.
2. Radi dodatnog ubrzanja procesa učenja ugraditi metodu zamaha (momentuma) prvog te drugog reda.
3. Za primjer učenja ponašanja linearnog dinamičkog člana prvog reda odrediti datoteku s podacima za učenje i više različitih datoteka s podacima za testiranje mreže.
4. Usporediti učenje i testiranje naučenih neuronskih mreža s jednim i dva skrivena sloja, pri čemu treba koristiti različite topologije statičke neuronske mreže.
5. Usporediti odzive neuronskih mreža na različite pobudne funkcije.
6. Programsku podršku načiniti u nekom od dostupnih matematičkih programskih paketa.
7. Izvesti zaključke rada.

Zadatak zadan:
25. studenog 2014.

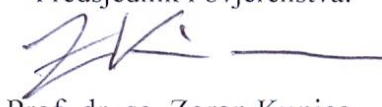
Rok predaje rada:
1. rok: 26. veljače 2015.
2. rok: 17. rujna 2015.

Predviđeni datumi obrane:
1. rok: 2., 3., i 4. ožujka 2015.
2. rok: 21., 22., i 23. rujna 2015.

Zadatak zadao:


Prof. dr. sc. Dubravko Majetić

Predsjednik Povjerenstva:


Prof. dr. sc. Zoran Kunica

Izjavljujem da sam ovaj rad izradio samostalno koristeći znanja stečena tijekom studija i navedenu literaturu.

Srdačno se zahvaljujem svome mentoru, prof. dr. sc. Dubravku Majetiću, na prihvaćanju mentorstva, pružanju savjeta te na strpljenju.

Zahvaljujem se svojoj obitelji na podršci tijekom studiranja.

Također zahvaljujem se svima koji su mi bili podrška tijekom preddiplomskog dijela studija.

Krešimir Milas

Sadržaj

I.	Popis slika.....	7
II.	Popis tablica	8
III.	Popis oznaka.....	9
IV.	Sažetak.....	10
V.	Summary	11
1.	Uvod.....	12
1.1	Teorija sustava.....	12
1.2	Pojam sustava	12
1.3	Identifikacija sustava.....	13
1.4	Umjetne neuronske mreže	13
2.	Statička unaprijedna neuronska mreža s povratnim prostiranjem pogreške.....	17
2.1	Uvod	17
2.2	Model statičke neuronske mreže	19
2.3	Učenje povratnim rasprostiranjem greške	20
2.4	Unaprijedna faza.....	21
2.5	Povratna faza	22
2.6	Promjena težina izlaznog sloja	23
2.7	Promjena težina skrivenog sloja.....	25
2.8	Promjena težina drugog skrivenog sloja.....	27
2.9	Momentum drugog reda	29
3.	Ocjena uspješnosti algoritma.....	30
4.	Modeliranje P1 člana.....	31
4.1	Set za učenje.....	31
4.2	Set za testiranje.....	33
4.3	Usporedba različitih topologija neuronskih mreža.....	36
4.4	Modeliranje P1 člana jednoslojnom mrežom	36
4.4.1	Učenje do zadanog NRMS-a	39
4.4.2	Učenje do zadanog koraka.....	41
4.4.3	Odziv mreže 2-21-1	41
4.5	Modeliranje P1 člana dvoslojnom mrežom.....	45
4.5.1	Učenje do zadanog NRMS-a	46
4.5.2	Učenje do zadanog koraka.....	47
4.6	Odzivi mreža na druge pobudne funkcije.....	48

4.6.1	Nagibne pobudna funkcije.....	48
4.6.2	Sinusna pobudna funkcija.....	51
4.6.3	Negativna step pobudna funkcija	53
4.6.4	Učenje mreža setom za učenje koji sadrži negativni spektar	55
5.	Zaključak.....	57
6.	Literatura	60

I. Popis slika

Slika 2.1 Model neurona	17
Slika 2.2 Tipične aktivacijske funkcije (izvor [2])	18
Slika 2.3 Model statičke unaprijedne neuronske mreže	19
Slika 2.4 Princip učenja neuronske mreže	20
Slika 2.5 Struktura dvoslojne mreže	27
Slika 4.1 Set za učenje	32
Slika 4.2 Normirani set za učenje	33
Slika 4.3 Set za testiranje	34
Slika 4.4 Normirani set za testiranje	34
Slika 4.5 Model povratne veze	35
Slika 4.6 NRMS mreže 2-3-1 za utvrđivanje broja koraka učenja	37
Slika 4.7 Utjecaj momentuma 1. reda na konvergenciju NRMS-a (mreža 2-3-1)	40
Slika 4.8 Utjecaj momentuma 2. reda na konvergenciju NRMS-a (mreža 2-3-1)	40
Slika 4.9 Odziv mreže na $u(n)=0.5, n=1,2...40, u(n)=0, n=41...80$	42
Slika 4.10 Odziv mreže na $u(n)=1, n=1,2...40, u(n)=0, n=41...80$	42
Slika 4.11 Odziv mreže na $u(n)=1.5, n=1,2...40, u(n)=0, n=41...80$	43
Slika 4.12 Odziv mreže na $u(n)=2, n=1,2...40, u(n)=0, n=41...80$	43
Slika 4.13 Odziv mreže na $u(n)=2.5, n=1,2...40, u(n)=0, n=41...80$	44
Slika 4.14 Odziv mreže na $u(n)=3, n=1,2...40, u(n)=0, n=41...80$	44
Slika 4.15 Pojava lokalnog minimuma u učenju mreže	45
Slika 4.16 Pobudne nagibne funkcije	48
Slika 4.17 Odziv mreža na pobudnu funkciju 1	49
Slika 4.18 Odziv mreža na negativnu nagibnu funkciju 2	49
Slika 4.19 Odziv mreža na pozitivnu nagibnu funkciju 3	50
Slika 4.20 Odziv mreža na negativnu nagibnu funkciju 4	50
Slika 4.21 Sinusne pobudne funkcije	52
Slika 4.22 Odziv mreža na sinusnu pobudu 1	52
Slika 4.23 Odziv mreža na sinusnu pobudu 2	53
Slika 4.24 Step pobudna funkcija	54
Slika 4.25 Odziv mreža na step pobudu	54
Slika 4.26 Set za učenje s negativnim vrijednostima	55
Slika 4.27 Odziv mreža na pozitivnu nagibnu pobudnu funkciju	56
Slika 4.28 Odziv mreža na negativnu nagibnu pobudnu funkciju	56
Slika 5.1 Usporedba odziva mreže 2-15-1 i 2-5-7-1	57

II. Popis tablica

Tablica 4.1 Topologije jednoslojnih i dvoslojnih mreža	36
Tablica 4.2 Parametri učenja jednslojne neuronske mreže.....	38
Tablica 4.3 Rezultati učenja do zadanog NRMS-a jednoslojne mreže	39
Tablica 4.4 Rezultati učenja do određenog koraka jednoslojne mreže	41
Tablica 4.5 Parametri učenja dvoslojne mreže.....	46
Tablica 4.6 Rezultati učenja do zadanog NRMS-a dvoslojne mreže	46
Tablica 4.7 Rezultati učenja do zadanog koraka dvoslojne mreže	47
Tablica 4.8 Odzivi mreža na nagibne funkcije	51
Tablica 4.9 Rezultati učenja seta s negativnim vrijednostima.....	56
Tablica 5.1 Usporedba mreže 2-15-1 i 2-5-7-1	58
Tablica 5.2 Usporedba mreže 2-21-1 i 2-8-6-1	58

III. Popis oznaka

Oznaka	Jedinica	Opis
BIAS	-	neuroni bez ulaza izlazne vrijednosti jednake jedan
d_n	-	željeni izlazi mreže
E	-	funkcija cilja
H	-	oznaka skrivenog sloja mreže
I	-	broj ulaznih neurona uvećan za jedan
J	-	broj neurona u skrivenom sloju uvećan za jedan
K	-	broj neurona izlaznog sloja
K_p	-	pojačanje sustava
L	-	broj neurona u drugom skrivenom sloju uvećan za jedan
net	-	vrijednost funkcije sume
NRMS	-	normalizirani korjen srednje kvadratne pogreške
O_k	-	izlazi iz mreže
T	s	vremenska konstanta sustava
T_0	s	vrijeme uzrokovanja
u	-	ulaz sustava
v	-	težine skrivenih slojeva
w	-	težine izlaznog sloja
x	-	vrijednost izlaza identificiranog sustava
y	-	vrijednost izlaza neurona
α	-	koeficijent zamaha prvog reda
β	-	koeficijent zamaha drugog reda
δ	-	parametar algoritma povratnog prostiranja pogreške
η	-	koeficijent brzine učenja mreže
σ	-	standardno odstupanje

IV. Sažetak

U ovom se radu na primjeru učenja modeliranja linearnog dinamičkog sustava prvog reda (identifikacija sustava) pokazala se kvaliteta učenja i generalizacije statičke neuronske mreže s jednim i dva skrivena sloja neurona. Za usporedbu se odabralo više različitih topologija jednoslojne i dvoslojne mreže. Uspoređivale su se mreže s jednakim brojem parametara učenja. Metode po kojima su uspoređivane su: brzina učenja do određenog NRMS-a te učenje do određenog koraka. Učenje se izvodilo po uzorku s zamahom prvog i drugog reda, a algoritam učenja je povratno raspostiranje greške. Nakon učenja gledalo se vrijeme potrebno da se dosegne cilj, postignut NRMS učenja te kvaliteta odziva. Nakon usporedba uzele su se mreže koje su se pokazale najbolje u generalizaciji te ih se pobudilo različitim pobudnim funkcijama (sinusnom i nagibnom) koje nisu bile u setu za učenje. Potom su ponovno uspoređene na temelju mogućnosti generalizacije odziva koje nisu učene.

Ključne riječi:

- neuronske mreže
- umjetna inteligencija
- jednoslojne mreže
- dvoslojne mreže
- učenje po uzorku
- zamah prvog reda
- zamah drugog reda
- generalizacija
- P1 član
- modeliranje sustava
- povratno raspostiranje greške
- pobudne funkcije

V. Summary

This paper dealt with neural network learning of first lag order element. In it the quality of learning and quality of generalization of networks with one and two hidden layers is shown. Multiple topologies of networks with one and two layers were chosen for comparison. Networks with equal number of learning parameters were compared. There were two methods of evaluation of neural networks: learning up to a predetermined error threshold and learning up to predetermined number of epochs. Learning method used was pattern learning with momentum of first and second order. Upon completion of learning process time elapsed, achieved NRMS and quality of response were considered. After comparison, networks that proved best in generalization were introduced to different input functions (sine and ramp) which weren't in original learning set and compared again.

Key words:

- neural networks
- artificial intelligence
- one hidden layer networks
- two hidden layer networks
- pattern learning
- first order momentum
- second order momentum
- generalisation
- first order lag element
- system modelling
- error backpropagation
- input signal

1. Uvod

1.1 Teorija sustava

Opća teorija sustava (ili samo teorija sustava) je interdisciplinarna znanost koja se bavi formulacijom i derivacijom principa koji vrijede općenito za sve sustave. Naime nagli razvoj znanstvenih disciplina, kako prirodnih tako i društvenih, doveo je do potrebe da se nađe podijeljenost među istima [3]. Opća teorija sustava ima za cilj da se postave zajedničke osnove i zajednička metodologija pristupa izučavanja svih znanstvenih disciplina. Opću teoriju sustava je prvi predložio biolog Ludwig von Bertalanffy 30-ih godina, no u znanstvenim krugovima je pojam teorije sustava zaživio tijekom 1950-ih. Danas teorija sustava ima važnu ulogu u suvremenoj znanosti. Ona omogućuje rješavanje mnoštva problema raznih znanstvenih disciplina (npr. tehnika, ekonomija, sociologija, biologija itd.) [11].

1.2 Pojam sustava

Ne postoji jednoznačna i precizna definicija pojma sustav. Mnogi su autori dali svoje definicije tog pojma. Većina definicija sadrži nekoliko istih kategorija. U svakom sustavu postoje elementi, postoje određene veze među elementima, postoji svrsishodna funkcionalna skladna cjelina, postoje realne ili imaginarne granice sustava koje ga izdvajaju od njegova okoliša te postoji određeni odnos sustava s okolišem. Prema tome može se dati sljedeća definicija: sustav je skup objekata (elementi sustava ili podsustavi) koji stoje u određenoj međusobnoj interakciji s okolinom sustava, tako da grade funkcionalnu, svrsishodnu i skladnu cjelinu. Podsustavi se po potrebi mogu razložiti na manje cjeline, dok elementi se ne mogu raščlaniti ili to nije od interesa .

Sustavi mogu biti različitih priroda; tehničkih, bioloških, ekonomskih itd. Primjer tehničkog sustava je termoenergetski blok sa svojim elementima: ložište, isparivač, pregrijači pare i turbina. Čovjek je biološki sustav čiji su elementi organi. Narodna privreda je primjer ekonomskog sustava. Studenti na fakultetu su primjer socijalnog sustava. Postoje i kombinirani sustavi koji su sastavljeni od različitih sustava, tvornica je primjer sustava koji se sastoji od gore navedenih sustava [3].

Glavno obilježje sustava je proces kojim se ulazne veličine (djelovanje okoline na sustav) transformiraju u izlazne (djelovanje sustava na okolinu). Zakonitosti ponašanja nekog sustava se najčešće opisuju matematičkim modelom što predstavlja osnovu za projektiranje njegovog sustava upravljanja ili koristi za simulaciju i predviđanje njegovog budućeg ponašanja [5].

1.3 Identifikacija sustava

Matematički model sustava se određuje teoretskom analizom ili eksperimentalnom analizom ili njihovom kombinacijom. Kod tehničkih sustava se teoretskom analizom može dobiti fizikalni matematički modeli sustava. Postupak dobivanja takvog modela naziva se modeliranje sustava. Eksperimentalna analiza sustava podrazumijeva dobivanje niza vrijednosti ulaza i pripadnih izlaza za niz mjerenja. Taj postupak se naziva identifikacijom sustava. Model sustava dobiven teoretskom analizom daje dobar uvid u unutarnja zbivanja u procesu, no obično je previše složen i ne opisuje točno sustav (npr. smetnje, gubitci itd.). Takav model je neprikladan za projektiranje sustava upravljanja ili za simulacije. Nasuprot takvom modelu je model dobiven eksperimentalno, odnosno identifikacijom sustava, koji ne daje dobar uvid u fizikalnost zbivanja u procesu, ali je jednostavniji i bolje opisuje ulazno-izlazno ponašanje procesa. Te prednosti ga čine pogodnijim za projektiranje sustava upravljanja i simulacije [5].

Postupak identifikacije se odvija u nekoliko osnovnih koraka [1]:

1. Modeliranje – postavljanje strukture različitih modela procesa
2. Procjena parametara (estimacija) – izvođenje eksperimenta (snimanje ulaznih i pripadajućih izlaznih veličina) i procjena parametara procesa
3. Provjera (validacija) – testiranje i odabir najboljeg modela

Postupak je iterativan i treba ga ponavljati dok se ne dobije model koji s propisanom točnošću opisuje promatrani proces što se lako može provesti primjenom neuronskih mreža zbog njihovih osobina kao što su učenje iz primjera, generalizacija, nelinearna aktivacijska funkcija itd [5].

1.4 Umjetne neuronske mreže

Umjetne neuronske mreže su, prema Robertu Hecht-Nielsenu, paralelne distribuirane informacijske strukture koje se sastoje od elemenata procesiranja (neurona) koji su povezani u slojeve jednosmjernim vezama. Pritom neuroni imaju lokalnu memoriju i predstavljaju osnovnu jedinku procesiranja [10]. Tehnologija neuronskih mreža nastala je u pokušaju oponašanja rada životinjskog, odnosno u konačnici ljudskog, mozga s ciljem razumijevanja istog. To podrazumijeva da neuronske mreže trebaju biti sposobne procesirati (prihvaćati, obrađivati, generirati, pohraniti i prenositi) informacije analogno aktivnostima mozga. Osnovna gradivna jedinica mozga je biološki neuron sastavljen od tijela, aksona te mnoštva dendrita. Osnovna funkcija takvog neurona je provođenje električnog impulsa ako mu se akson dovede u stanje dovoljne uzbude, realizirano njegovim trenutačnim stanjem, te utjecajem signala od ostalih neurona preko njegovih dendrita ili njegovog tijela u kratkom

vremenskom intervalu (period latentne sumacije). Sam signal može biti smirujući ili uzbudni. Neuron ispaljuje signal ako je njegova uzbuda veća od smirujućeg utjecaja za kritični iznos, koji predstavlja prag osjetljivosti neurona. Neuroni međusobno spojeni čine neuronsku mrežu, pritom jedan neuron može djelovati na stotine ili tisuće drugih. Neuronske mreže bioloških sustava su sačinjene od velikog broja neurona i vrlo su složene. Tako se za čovjeka procjenjuje da mozak ima preko 100 milijardi neurona. Upravo takva mreža sastavljena od jednostavnijih komponenti, koje su međusobno povezane vezama različite kompleksnosti, ima sposobnost viših funkcija poput razmišljanja, učenja, emocija, spoznaje i percepcije. Dakle više funkcije su ostvarene vezama različite kompleksnosti među elementima, a ne elementima samim. Umjetne neuronske mreže oponašaju gore navedeni biološki sustav, no postojeće umjetne neuronske mreže grade se na osnovi izuzetno pojednostavljenog modela koji ne uzimaju u obzir većinu poznatih funkcija mozga (jer ne postoji odgovarajući matematički model koji bi ih opisao).

Analogno biološkim mrežama osnovna gradivna jedinica umjetnih neuronskih mreža je umjetni neuron. Umjetni neuron se projektira s idejom da oponaša osnovne funkcije (karakteristike prvog reda) biološkog neurona. Tijelo biološkog neurona zamjenjuje se sumatorom, ulogu dendrita preuzimaju ulazi u sumator, izlaz sumatora je akson umjetnog neurona, a uloga praga osjetljivosti bioloških neurona preslikava se na tzv. aktivacijske funkcije. Funkcije sinaptičke veze biološkog neurona s njegovom okolinom preslikavaju se na težinske faktore, preko kojih se i ostvaruje veza umjetnog neurona s njegovom okolinom. Težinski faktor je pozitivan ili negativan broj (u slučaju nule ta veza ne postoji), a može biti i funkcija (varijabilan težinski faktor). Funkcija težinskih faktora je ista kao i sinapsa u biološkom neuronu; povezuje neuron s ostalima. Intenzitet veze ovisi o modulu (iznosu), a karakter ovisi o predznaku težinskog faktora. Izlaz sumatora povezuje se na ulaz aktivacijske funkcije, koja na svom izlazu producira izlaz umjetnog neurona. Aktivacijske funkcije mogu biti linearne i nelinearne. Kod linearnih funkcija, izlaz sumatora se množi s nekim faktorom (pojačanjem) i tako se dobiva izlaz neurona. Nelinearne aktivacijske funkcije mogu poprimiti različite oblike, ali se najčešće koriste: funkcije praga osjetljivosti, sigmoidalne, hiperbolične i harmoničke funkcije. Nelinearne aktivacijske funkcije prevode izlaz sumatora na izlaz neurona preko nelinearnog pojačanja. Opisan umjetni neuron radi na sličnom principu kao i biološki neuron: dobiva signal od prethodnih neurona, množi signale težinskim faktorima, sumira ih te šalje na ulaz aktivacijske funkcije koja daje izlaz neurona. Ovakva struktura zanemaruje mnoge karakteristike biološkog neurona (npr. kašnjenje signala). Ipak i uz ograničenja koja nastaju takvim zanemarivanjima neuronske mreže daju dobre rezultate. Naime, odgovarajućim povezivanjem neurona u mrežu, mogu se ostvariti funkcije koje sam neuron ne može ostvariti. Postoji velik broj vrsta neuronskih mreža koje imaju različita svojstva. No unatoč tome sve neuronske mreže imaju neka zajednička svojstva, koja su svojstvena biološkim neuronskim mrežama, a koja se ne mogu naći kod konvencionalnih računalnih tehnika.

Nekoliko takvih svojstava su:

- Paralelno raspodijeljena obradba informacija (engl. parallel distributed processing).

Za razliku od konvencionalnih računskih tehnika, neuronske mreže prihvaćaju više ulaza paralelno i dobivene informacije obrađuju na raspodijeljen način (engl. distributed processing). Drugim riječima, informacija spremljena u neuronsku mrežu raspodijeljena je na više računskih jedinica, što je potpuno suprotno konvencionalnome spremanju informacija u memoriju gdje je svaka posebna informacija (podatak) spremljena u svoj vlastiti memorijski prostor. Svojstvo raspodijeljenoga spremanja informacija daje neuronskim mrežama više prednosti, od kojih je najvažnija inherentna redundantnost, to jest otpornost na kvar.

- Učenje i adaptacija (engl. learning and adaptive abilities)

Svojstvo učenja i adaptacije čini neuronske mreže sposobnima obrađivati neprecizne i loše ušćuvane podatke u nestrukturiranom i neodređenom okruženju. Na odgovarajući način naučena neuronska mreža ima svojstvo poopćavanja kada se na njezinome ulazu pojave podaci koji nisu bili u uzorku na osnovi kojeg je mreža naučena (svojstvo generalizacije).

- Univerzalni aproksimator

Svojstvo neuronskih mreža da aproksimiraju proizvoljnu kontinuiranu nelinearnu funkciju do željene točnosti njihovo je najvažnije svojstvo sa stajališta modeliranja, identifikacije i upravljanja nelinearnim procesima.

- Viševarijabilni sustavi.

Neuronske su mreže po svojoj strukturi viševarijabilni sustavi što ih čini lako primjenjivim za modeliranje, identifikaciju i upravljanje viševarijabilnim procesima

- Sklopovska implementacija

Više je proizvođača razvilo specijalizirane sklopove za implementaciju neuronskih mreža koji omogućuju paralelnu raspodijeljenu obradbu u stvarnome vremenu.

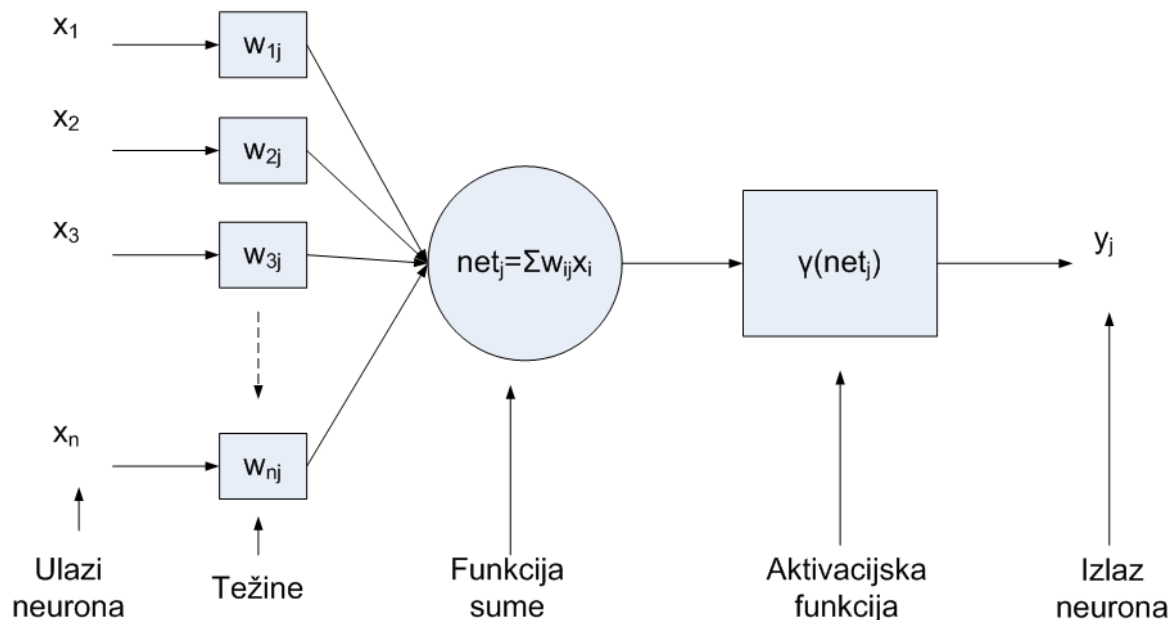
Većina neuronskih mreža prikazanih zahtijeva učenje, to jest primjenu algoritama koji podešavaju iznose sinaptičkih težinskih koeficijenata. Ciljevi učenja mreže ovise o njezinoj primjeni, pa tako i izbor odgovarajućeg algoritma učenja. Iako je razvijen veliki broj raznih algoritama učenja neuronskih mreža, u ovom radu će se koristiti učenje povratnim prostiranjem pogreške. Algoritmi učenja temeljeni na pogrešci često se nazivaju i algoritmi s učiteljem (engl. supervised algorithms) jer zahtijevaju vanjski referentni signal, odnosno učitelja, koji nadgleda ponašanje mreže korigirajući istu dok ne dobije željeno ponašanje mreže. Naime, kod ovog načina učenja najprije se usvoji određena struktura mreže (broj ulaza, broj slojeva, broj neurona, broj izlaza, te broj težina mreže). Usvoje se, obično preko generatora slučajnih brojeva, početne težine (parametri) neuronske mreže. Zatim se na ulaz mreže dovodi skup ulaznih varijabli. Mreža producira određen skup izlaznih varijabli. Skup

izlaznih varijabli uspoređuje se sa željenim skupom izlaznih varijabli. Razlika željenih i stvarnih izlaza neuronske mreže gradi pogrešku mreže, koja se koristi za računanje novih težina (parametara) mreže preko određenog usvojenog algoritma. Cijeli postupak se ponavlja iteracijski dok pogreška mreže ne bude manja od unaprijed zadanog iznosa. Pritom se prema potrebi mijenja struktura mreže (broj neurona, broj slojeva, broj težina). Nakon procesa učenja (treninga, adaptacije) slijedi proces testiranja neuronske mreže. To se radi s novim skupom ulaza mreže koji nije bio sadržan u ulaznom skupu za vrijeme procesa učenja. Mreža sada producira nove izlaze koji se uspoređuju sa željenim izlazima. Pritom se ne mijenjaju parametri (težine) mreže. Iznos pogreške mreže u procesu testiranja služi za ocjenu robusnosti, odnosno generalizacijskih svojstava mreže, tj. sposobnosti mreže da daje zadovoljavajuće izlaze (rezultate) i za skup ulaza kojim nije bila učena [5].

2. Statička unaprijedna neuronska mreža s povratnim prostiranjem pogreške

2.1 Uvod

Statičke neuronske mreže su najčešće korišteni tip neuronskih mreža. Osnovna gradivi element ovih mreža jest statički neuron prikazan na slici. Na slici se vidi da statički neuron ima dvije temeljne funkcije: funkcija sume Σ i aktivacijska funkcija γ . Također osnovna značajka svakog neurona je i posjedovanje više ulaza i samo jedan izlaz. Posebno je dodan jedan ulaz oznake Bias koji je konstantne vrijednosti 1. Njegova funkcija je osigurati učenje mreže. Već i jedan neuron u mreži može aproksimirati jednostavnije funkcije, a aproksimacijska moć mreže se općenito povećava brojem neurona. U statičkim neuronskim mrežama neuroni su vezani na unaprijedni način (često se ove mreže nazivaju i unaprijedne). To znači da svaki neuron može biti povezan s ulazima i/ili s drugim neuronima, ali samo tako da se pri povezivanju ne formiraju povratne veze. Prema tome, statičke mreže ne sadrže dinamičke članove, što ih inherentno čini strukturno stabilnima. Stabilnost je njihova najvažnija prednost u odnosu na dinamičke mreže.



Slika 2.1 Model neurona

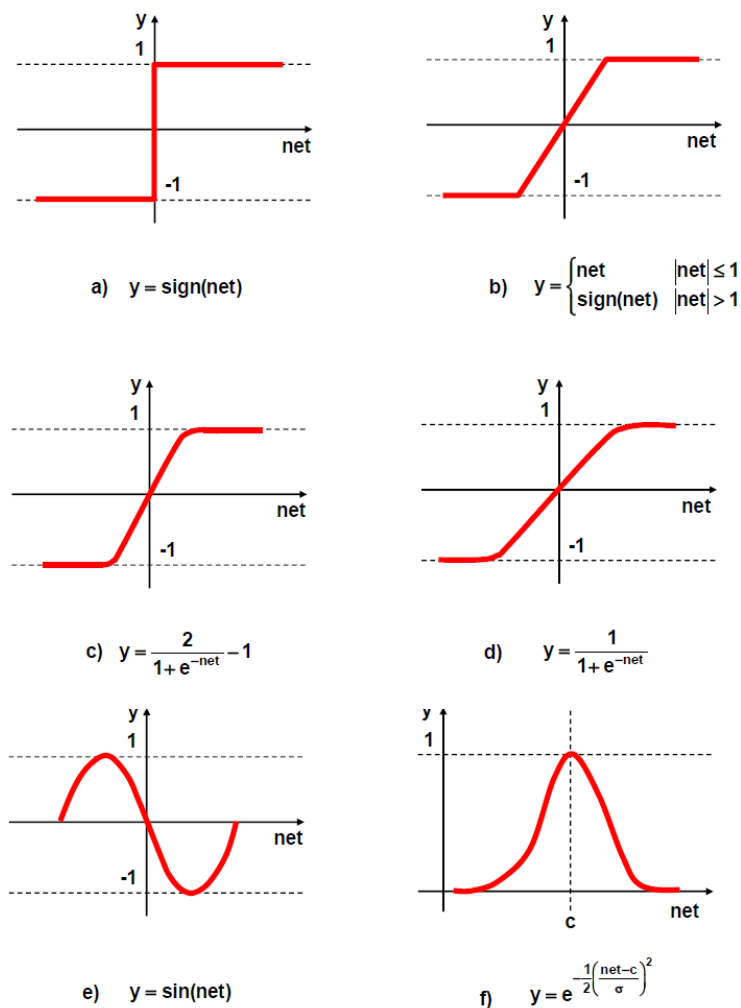
Funkcija sume statičkog neurona daje zbroj umnožaka ulaza neurona i pripadajućih težinskih faktora, izraz funkcije dan je izrazom:

$$net = \sum_{j=1}^J w_j u_j \quad (1)$$

Aktivacijska funkcija γ vrši preslikavanje vrijednosti net u izlaznu vrijednost neurona y prema izrazu:

$$y = \gamma(net) \quad (2)$$

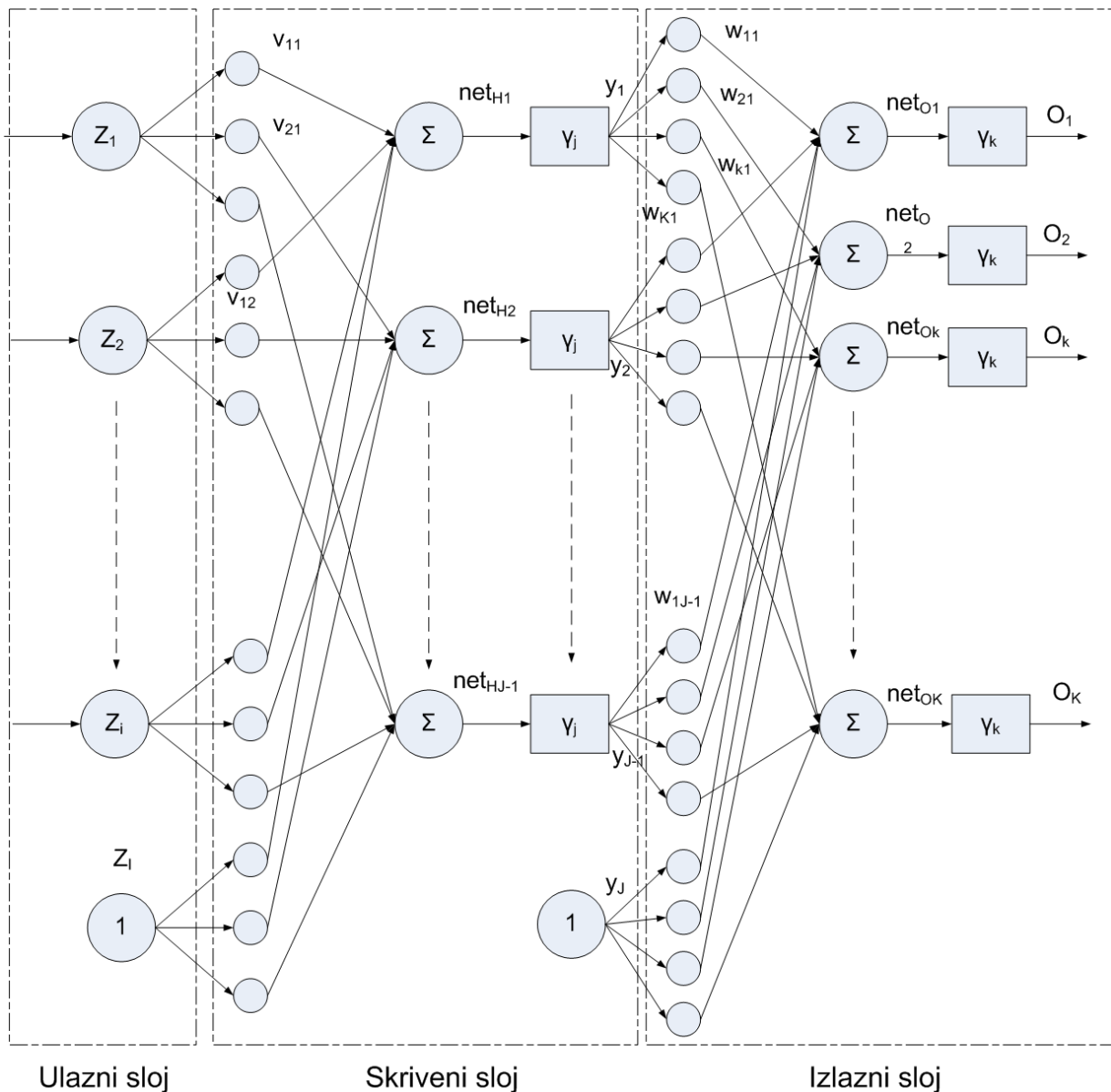
Za aktivacijsku funkciju se uglavnom odabire neka monotona rastuća funkcija sa zasićenjem. Tipični primjeri funkcija dani su na slici 2. Odabir aktivacijske funkcije ovisi o skupu za učenje i o skupu za testiranje naučene neuronske mreže. Pritom treba voditi računa da se vrijednosti učenja i testiranja normaliziraju s obzirom da gotovo sve funkcije imaju zasićenje od te dvije granice [1].



Slika 2.2 Tipične aktivacijske funkcije (izvor [2])

2.2 Model statičke neuronske mreže

Struktura (topologija ili konfiguracija) neuronske mreže se dobiva organiziranjem neurona u slojeve, a potom povezivanja tih slojeva vezama koje su opterećene težinama. Razlikuju se tri tipa slojeva: ulazni, skriveni i izlazni. Ulazni i izlazni slojevi su u direktnoj interakciji s okolinom, dok skriveni nije. Najpoznatiji i najčešće korišteni tip neuronske mreže je statička neuronska višeslojna mreža prikazana na slici 7.3.



Slika 2.3 Model statičke unaprijedne neuronske mreže

Ulazi neurona ulaznog sloja (Z_i) služe za distribuciju signala prema prvom skrivenom sloju neuronske mreže. Ulazni sloj je vezan sa skrivenim vezama koje su opterećene težinama v_{ji} .

Svaki neuron nekog sloja vezan sa svim neuronima prethodnog sloja. Izuzetak su jedino neuroni označeni s *bias*, neuroni bez ulaza izlazne vrijednosti jednake jedinici.

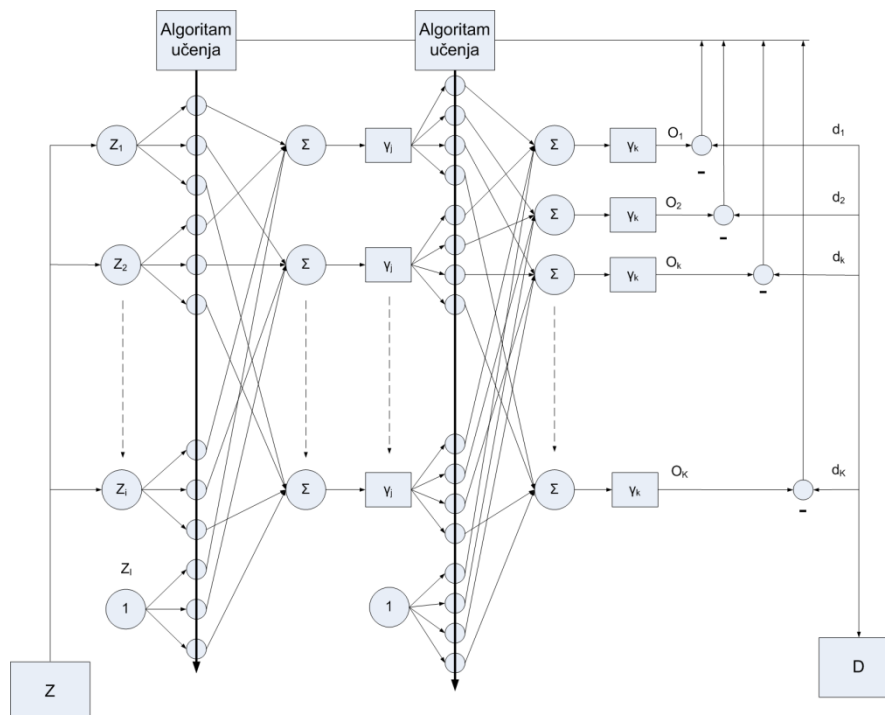
Broj skrivenih slojeva je proizvoljan, no najčešće se radi o jedan do dva sloja. Istraživanja su pokazala da je jedan sloj skrivenih neurona dovoljan za aproksimaciju bilo koje kontinuirane funkcije, jedini uvjet je da taj skriveni sloj ima dovoljan broj neurona.

Broj ulaznih i izlaznih neurona je uvijek definiran problemom koji neuronska mreža mora riješiti. Broj ulaznih neurona mora biti dostatan da prihvati sve ulaze, dok broj izlaznih neurona mora biti dostatan da proizvede sve izlaze.

2.3 Učenje povratnim rasprostranjem greške

Učenje neuronske mreže ima za cilj podesiti težinske koeficijente tako da se izlazi mreže približe poznatim izlazima za odgovarajuće ulaze. Pritom aproksimacija mreže neće i ne treba biti potpuno točna. Kvaliteta aproksimacije ovisi o više parametara kao što su problem učenja, topologija mreže te odabrani algoritam učenja.

Postupak učenja u svakom koraku se sastoji od dvije faze: unaprijedne i povratne. Prva faza je unaprijedna. U unaprijednoj fazi uzimaju se vrijednosti iz zapisnika učenja \mathbf{Z} , te se pomoću njih izračunava odziv mreže \mathbf{O} . Težinski faktori koji opterećuju veze između neurona (\mathbf{W} , \mathbf{V}_1 , \mathbf{V}_2 , ...) se najčešće odabiru slučajno i na takav način da su istog reda veličine kao i ulazi i izlazi.



Slika 2.4 Princip učenja neuronske mreže

2.4 Unaprijedna faza

Račun unaprijedne faze za višeslojnu mrežu počinje, kako je rečeno, s vrijednostima koje se uzimaju iz zapisnika učenja. Vrijednosti seta učenja se dovode na ulaz gdje im se pridodaje jedan ulaz vrijednosti jedan (bias). Potom se svaki ulaz množi s pripadnim težinskim faktorom V_{ji} , te se zbraja s ostalim opterećenim ulazima. Postupak se ponavlja onoliko puta koliko ima neurona u prvom skrivenom sloju, dakle dobije se broj suma jednak broju neurona u skrivenom sloju. Suma **net** se može prikazati izrazom:

$$net_{Hj} = \sum_{i=1}^I v_{ji}Z_i \quad (3)$$

$$j = 1, 2, \dots, J - 1, \quad i = 1, 2, \dots, I$$

Gdje je I jednak broju ulaznih neurona uvećan za jedan, a J je jednak broju neurona u skrivenom sloju uvećan za jedan bias. Izlazi iz skrivenog sloja se dobivaju tako da se sume iskoriste kao argumenti u aktivacijskoj funkciji.

Za nelinearnu sigmoidalnu funkciju vrijedi:

$$y_j = \frac{2}{1 + e^{-net_{Hj}}} - 1 \quad (4)$$

$$j = 1, 2, \dots, J - 1$$

$$y_j = 1, Bias$$

Izlazi prvog skrivenog sloja spojeni su na ulaze slijedećeg sloja. Ako je slijedeći sloj skriveni gore navedeni postupak se ponavlja na analogan način, jedina razlika su ulazi koji se sada ne uzimaju iz seta učenja već se koriste izlazi prethodnog sloja. Ukoliko je slijedeći sloj izlazni ponovno se računaju sume iz ulaza koji su pomnoženi težinskim faktorima izlaznog sloja **W**:

$$net_{Ok} = \sum_{j=1}^J w_{kj}y_j \quad (5)$$

$$k = 1, 2, \dots, K$$

Gdje je K broj neurona izlaznog sloja, odnosno broj izlaza mreže.

Ukoliko je potrebno da mreže izlaza budu veće od 1 može se koristiti linearna aktivacijska funkcija:

$$\begin{aligned} O_k &= K_p \text{net}_{Ok} \\ k &= 1, 2, \dots, K \end{aligned} \quad (6)$$

Gdje je K_p nagib linearne funkcije, za potrebe ovog rada K_p će biti 1.

2.5 Povratna faza

Nakon što je završila unaprijedna faza počinje povratna faza. Na temelju odziva mreže i željenih vrijednosti d izračunava se greška na čijoj se osnovi potom vrši korekcija vrijednosti težinskih faktora. Postupak se vrši za svaki ulazno-izlazni par sve dok se ne postigne greška manja ili jednaka dozvoljenoj pogrešci koja se određuje ovisno o zadatku.

Najčešće korištena funkcija cilja je suma kvadrata pogreške dana izrazom:

$$E = \frac{1}{2} \sum_{n=1}^N (d_n - O_n)^2 \quad (7)$$

Gdje je N broj elemenata u skupu za učenje, tj. Broj podataka učenja iz zapisnika učenja. Rezultat podešavanja koeficijenata je minimiziranje greške E po nekog algoritmu nelinearnog optimiranja.

Forma promjena parametara učenja (težinskih koeficijenata) ϑ dana je izrazom:

$$\vartheta(n+1) = \vartheta(n) + \Delta\vartheta(n) \quad (8)$$

Gdje je n trenutni korak učenja, $\Delta\vartheta(n)$ veličina promjene parametara učenja, $\vartheta=w$ za izlazni sloj, odnosno $\vartheta=v$ za skrivene slojeve, dok je $\vartheta(n+1)$ je nova vrijednost parametara učenja. Pogreška $E(\vartheta)$ se može u okolišu točke aproksimirati ϑ s prva dva člana Taylorova reda:

$$\begin{aligned} E(\vartheta + \Delta\vartheta) &\approx E(\vartheta) + \Delta E(\vartheta) \\ \Delta E(\vartheta) &= \Delta\vartheta^T \nabla E(\vartheta) \\ \nabla E(\vartheta) &= \frac{\partial E(\vartheta)}{\partial \vartheta} \end{aligned} \quad (9)$$

Posljednji izraz se naziva gradijentom pogreške. Da bi se ostvarilo smanjivanje pogreške najvećim mogućim iznosom treba odrediti $\Delta\vartheta$ za koji promjena greške učenja $\Delta E(n)$ poprima najveći negativni iznos, što se ostvaruje uz uvjet:

$$\Delta\vartheta = -\eta\nabla E(\vartheta) \quad (10)$$

gdje je η mjera te promjene, koja se još naziva koeficijentom brzine učenja. Zadaje je učitelj, a obično se poprima vrijednosti između 10^{-3} i 10. Gornji izraz osigurava promjenu težinskih koeficijenata u smjeru najvećeg mogućeg pada ukupne pogreške. Njegovim uvrštavanjem u izraz dobiva se:

$$\vartheta(n+1) = \vartheta(n) - \eta\nabla E(\vartheta(n)) \quad (11)$$

Algoritam minimiziranja pogreške dan gornjim izrazom poznat je pod nazivom algoritam najstrmijeg pada (engl. Steepest Descent), dok se u kontekstu neuronskih mreža naziva algoritam povratnog prostiranja pogreške (engl. Error Back-Propagation Algorithm). Ovakav algoritam je najpoznatiji i najčešće korišteni način promjene parametara učenja (težinskih koeficijenata kod statičkih mreža).

Najveći nedostatak algoritma povratnog prostiranja greške je veliki broj potrebnih iteracija. Da bi se proces učenja ubrzao tj. da bi se smanjio broj potrebnih koraka da greška mreže bude manja ili jednaka od dozvoljene modificira se osnovni algoritam. Jedna od modifikacija je korištenje tzv. momentuma (zamaha) prvog i drugog reda:

$$\Delta\vartheta(n) = -\eta\nabla E(\vartheta(n)) + \alpha\Delta\vartheta(n-1) + \beta\Delta\vartheta(n-2) \quad (12)$$

Gdje n označava trenutnu promenu parametra učenja, dok $n-1$ promjenu parametra učenja u prethodnom koraku, a $n-2$ u koraku prije prethodnog. Vrijednost α je vezana za momentum prvog reda i određuje je učitelj. Vrijednosti se kreću između 0.1 i 0.9. Vrijednost β je vezana za momentum drugog reda i određuje se iz vrijednosti α . Konačni izraz učenja s momentumom jest:

$$\vartheta(n+1) = \vartheta(n) - \eta\nabla E(\vartheta(n)) + \alpha\Delta\vartheta(n-1) + \beta\Delta\vartheta(n-2) \quad (13)$$

2.6 Promjena težina izlaznog sloja

Promjena parametara učenja algoritmom povratnog prostiranja pogreške se u neuronskoj mreži odvija od izlaznog prema ulaznom sloju. Promjena parametara se odvija prema jednadžbi (13) koja se prilagođava težniskim faktorima koji se mijenjaju. Za težinske faktore između zadnjeg skrivenog i izlaznog sloja vrijedi:

$$w_{kj}(n+1) = w_{kj}(n) - \eta\nabla E(\vartheta(n)) + \alpha\Delta w_{kj}(n-1) + \beta\Delta w_{kj}(n-2) \quad (14)$$

Gradijent pogreške za w_{kj} se izračunava prema:

$$\nabla E(n) = \frac{\partial E(n)}{\partial w_{kj}} = \frac{\partial E(n)}{\partial O_k} \frac{\partial O_k}{\partial net_{Ok}} \frac{\partial net_{Ok}}{\partial w_{kj}} \quad (15)$$

J.M. Zurada [8] uvodi karakterističnu vrijednost povratnog prostiranja greške koja se računa prvo za izlazni sloj i glasi:

$$\delta = -\frac{\partial E(n)}{\partial net_{Ok}} \quad (16)$$

Vrijednosti pojedinih članova izraza glase:

$$\frac{\partial E(n)}{\partial O_k} = -(d_k - O_k) \quad (17)$$

$$\frac{\partial O_k}{\partial net_{Ok}} = \gamma_k' = 1 \quad (18)$$

Vrijednost izraza δ izlaznog sloja neurona ima oblik:

$$\delta_{Ok} = d_k - O_k \quad (19)$$

Posljednji član uzastopnih derivacija dobiva se iz izraza i iznosi:

$$\frac{\partial net_{Ok}}{\partial w_{kj}} = y_{2j} \quad (20)$$

Uvrštavanjem izraza (17), (18) i (19) u izraz (15) dobiva se:

$$\nabla E(n) = \frac{\partial E(n)}{\partial w_{kj}} = -(d_k - O_k)y_j = -\delta_{Ok}y_j \quad (21)$$

Ako se izraz (21) uvrsti u izraz (14) dobiva se konačni algoritam promjene težina izlaznog sloja:

$$w_{kj}(n+1) = w_{kj}(n) + \eta \delta_{Ok}y_j + \alpha \Delta w_{kj}(n-1) + \beta \Delta w_{kj}(n-2) \quad (22)$$

2.7 Promjena težina skrivenog sloja

Nakon promjene svih težina w_{kj} izlaznog sloja pristupa se promjeni težina sakrivenih slojeva. Jednadžba po kojoj se vrši adaptacija sakrivenog sloja je izraz (13) prilagođen pripadnom skrivenom sloju:

$$v_{2ji}(n+1) = v_{2ji}(n) - \eta \nabla E(\vartheta(n)) + \alpha \Delta v_{2ji}(n-1) + \beta \Delta v_{2ji}(n-2) \quad (23)$$

Gradijent greške se izračunava na sličan način kao kod izlaznog sloja:

$$\frac{\partial E(n)}{\partial v_{2ji}} = \frac{\partial E(n)}{\partial y_{2j}} \frac{\partial y_{2j}}{\partial net_{2Hj}} \frac{\partial net_{2Hj}}{\partial v_{2ji}} \quad (24)$$

S obzirom da na promjenu svake težine sakrivenog sloja utječu svi neuroni izlaznog sloja (prema slici) prvi razlomak desne strane poprima sljedeći oblik:

$$\begin{aligned} \frac{\partial E(n)}{\partial O_2} = & \frac{\partial E(n)}{\partial O_1} \frac{\partial O_1}{\partial net_{O1}} \frac{\partial net_{O1}}{\partial y_{2j}} + \\ & \frac{\partial E(n)}{\partial O_2} \frac{\partial O_2}{\partial net_{O2}} \frac{\partial net_{O2}}{\partial y_{2j}} + \\ & \frac{\partial E(n)}{\partial O_3} \frac{\partial O_3}{\partial net_{O3}} \frac{\partial net_{O3}}{\partial y_{2j}} + \\ & \vdots \\ & \frac{\partial E(n)}{\partial O_K} \frac{\partial O_K}{\partial net_{OK}} \frac{\partial net_{OK}}{\partial y_{2j}} \end{aligned} \quad (25)$$

Pritom vrijedi:

$$\frac{\partial E(n)}{\partial O_k} = -(d_k - O_k), \quad k = 1, 2, \dots, K \quad (26)$$

$$\frac{\partial O_k}{\partial net_{Ok}} = \gamma_k' = 1, \quad k = 1, 2, \dots, K \quad (27)$$

$$\frac{\partial net_{Ok}}{\partial y_{2j}} = w_{kj}, \quad \begin{aligned} k &= 1, 2, \dots, K, \\ j &= 1, 2, \dots, J-1 \end{aligned} \quad (28)$$

Uvrštavanjem izraza (29), (30) i (31) u izraz (28) dobiva se:

$$\frac{\partial E(n)}{\partial y_{2j}} = - \sum_{k=1}^K (d_k - O_k) w_{kj} \quad (29)$$

Konačni oblik rješenja prvog razlomka na desnoj strani izraza (24) dobiva se uvrštavanjem izraza (19) u izraz (32):

$$\frac{\partial E(n)}{\partial y_{2j}} = - \sum_{k=1}^K \delta_{Ok} w_{kj} \quad (30)$$

Drugi i treći razlomak dobivaju se iz izraza (3) i (4)

$$\frac{\partial y_{2j}}{\partial net_{2Hj}} = \gamma_j' = \frac{1}{2}(1 - y_{2j}^2) \quad (31)$$

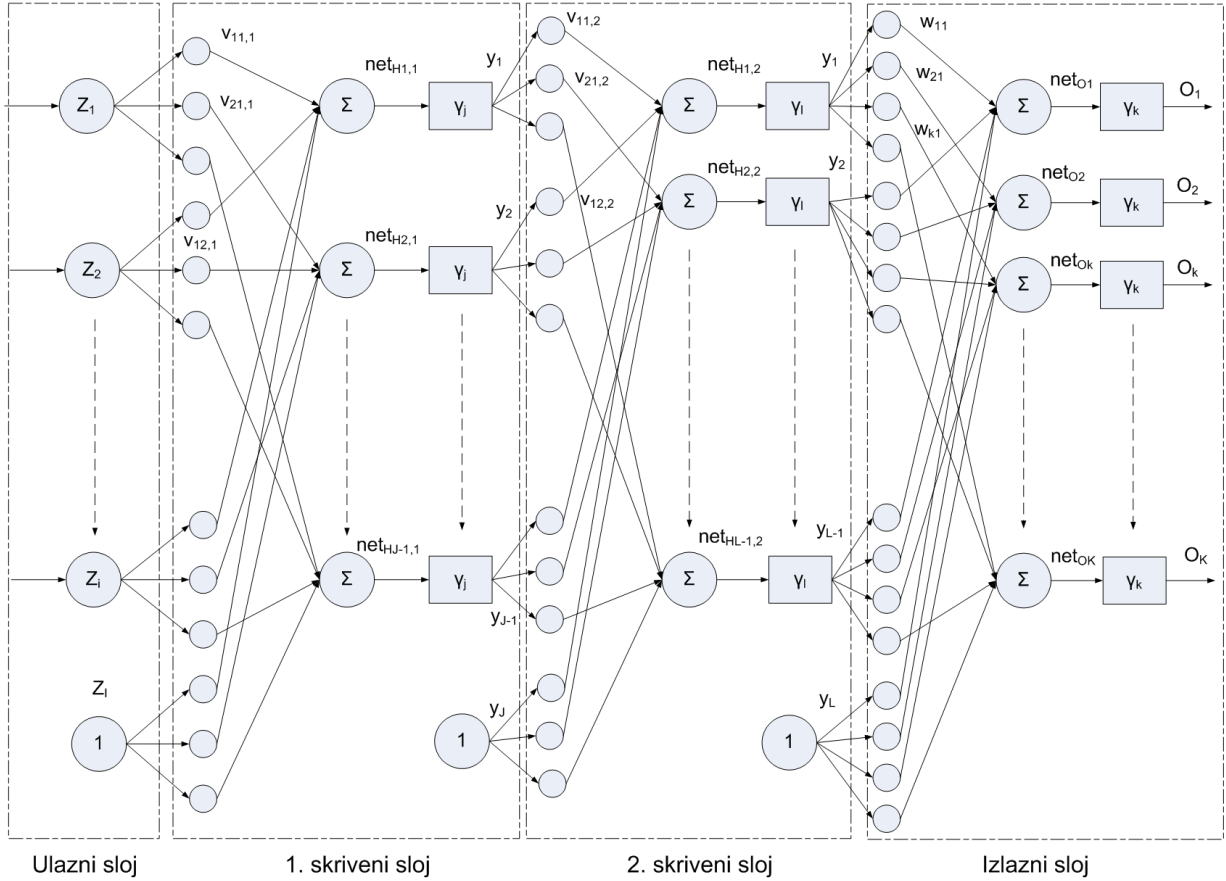
$$\frac{\partial net_{2Hj}}{\partial v_{2ji}} = y_{1j} \quad (32)$$

Kombiniranjem izraza (32), (31) i (30) dobiva se konačni oblik algoritma promjene težinskih koeficijenata posljednjeg sakrivenog sloja:

$$v_{2ji}(n+1) = v_{2ji}(n) + \frac{1}{2}\eta(1 - y_{2j}^2)y_{1j}\left(\sum_{k=1}^K \delta_{Ok} w_{kj}\right) + \alpha\Delta v_{2ji}(n-1) + \beta\Delta v_{2ji}(n-2) \quad (33)$$

2.8 Promjena težina drugog skrivenog sloja

Neuronska mreža može imati još jedan skriven sloj. Takva struktura je prikazana na slici ispod.



Slika 2.5 Struktura dvoslojne mreže

Promjena težina izlaznog skrivenog sloja je jednaka kao i kod jednoslojne mreže. Postupak promjene težinskih koeficijenata drugog skrivenog sloja je jednak postupku promjene koeficijenata skrivenog sloja jednoslojne mreže. Promjena težina prvog sloja (dodanog sloja) je analogna gore opisanom postupku :

$$v_{ji,1}(n+1) = v_{ji,1}(n) - \eta \nabla E(\vartheta(n)) + \alpha \Delta v_{ji,1}(n-1) + \beta \Delta v_{ji,1}(n-2) \quad (34)$$

Gradijent pogreške se izračunava prema izrazu:

$$\frac{\partial E(n)}{\partial v_{ji,1}} = \frac{\partial E(n)}{\partial y_{l,2}} \frac{\partial y_{l,2}}{\partial net_{HL,2}} \frac{\partial net_{HL,2}}{\partial v_{ij1}} \quad (35)$$

Pritom je prvi razlomak jednak:

$$\begin{aligned}
\frac{\partial E(n)}{\partial y_{j,1}} = & \frac{\partial E(n)}{\partial O_1} \frac{\partial O_1}{\partial net_{O1}} \frac{\partial net_{O1}}{\partial y_{1,2}} \frac{\partial y_{1,1}}{\partial net_{H1,2}} \frac{\partial net_{H1,2}}{\partial y_{1,1}} + \\
& \frac{\partial E(n)}{\partial O_2} \frac{\partial O_2}{\partial net_{O2}} \frac{\partial net_{O2}}{\partial y_{2,2}} \frac{\partial y_{2,2}}{\partial net_{H2,2}} \frac{\partial net_{H2,2}}{\partial y_{2,1}} + \\
& \frac{\partial E(n)}{\partial O_3} \frac{\partial O_3}{\partial net_{O3}} \frac{\partial net_{O3}}{\partial y_{3,3}} \frac{\partial y_{2,3}}{\partial net_{H3,2}} \frac{\partial net_{2H3}}{\partial y_{3,1}} + \\
& \vdots \\
& \frac{\partial E(n)}{\partial O_K} \frac{\partial O_K}{\partial net_{OK}} \frac{\partial net_{OK}}{\partial y_{2j}} \frac{\partial y_{2j}}{\partial net_{2Hj}} \frac{\partial net_{2Hj}}{\partial y_{l,1}}
\end{aligned} \tag{36}$$

Članovi gornjeg izraza jednaki su izrazima (29), (30), (31) i izrazu (39):

$$\frac{\partial net_{2Hj}}{\partial y_{1j}} = v_{2ji}, \quad j = 1, 2, \dots, J-1 \tag{37}$$

Uvrštavanjem (38), (39) i (40) u (34) dobiva se:

$$\frac{\partial E(n)}{\partial y_{1j}} = - \sum_{k=1}^K \frac{1}{2} \delta_{Ok} w_{kj} v_{2ji} (1 - y_{2j}^2) \tag{38}$$

Da bi se kompletirao izraz (35) treba jos riješiti dva razlomka na desnoj strani:

$$\frac{\partial y_{1j}}{\partial net_{1Hj}} = \frac{1}{2} (1 - y_{1j}^2) \tag{39}$$

$$\frac{\partial net_{1Hj}}{\partial v_{1ij}} = Z_i \tag{40}$$

Sada se mogu izrazi (40), (38) i (34) kombinirati u konačni izraz za prvi skriveni sloj:

$$\begin{aligned}
v_{1ji}(n+1) = & v_{1ji}(n) + \eta \frac{1}{2} (1 - y_{1j}^2) Z_i \left(\sum_{k=1}^K \frac{1}{2} \delta_{Ok} w_{kj} v_{2ji} (1 - y_{2j}^2) \right) \\
& + \alpha \Delta v_{1ji}(n-1) + \beta \Delta v_{1ji}(n-2)
\end{aligned} \tag{41}$$

2.9 Momentum drugog reda

Kako bi se mogao koristiti momentum drugog reda u izrazima treba odrediti koeficijent β u izrazima (33), (22) i (41).

Koeficijent β je dan izrazom:

$$\beta = \frac{\alpha - 1}{3} \quad (42)$$

3. Ocjena uspješnosti algoritma

Da bi se mogla procijeniti točnost algoritma učenja u rješavanju zadanog problema potrebno je definirati mjeru točnosti. U ovom radu će se za mjeru točnosti koristiti normalizirani korijen srednje kvadratne pogreške, NRMS (eng. Normalised Root Mean Square) dan izrazom:

$$NRMS = \frac{\sqrt{\frac{\sum_{n=1}^N (d_n - O_n)^2}{N}}}{\sigma_{d_n}} \quad (43)$$

Gdje je σ_{d_n} :

$$\sigma_{d_n} = \sqrt{\frac{\sum_{n=1}^N (d_n - \bar{d})^2}{N}} \quad (44)$$

$$\bar{d} = \frac{\sum_{n=1}^N d_n}{N} \quad (45)$$

4. Modeliranje P1 člana

Prvi dinamički član na kojem će biti obavljena analiza učenja mreže s jednim i dva sloja je P1 član dan diferencijalnom jednačinom:

$$T\dot{x} + x = K_p u \quad (46)$$

gdje je T vremenska konstanta, a K_p konstanta pojačanja. Da bi se neuronska mreža mogla naučiti modelirati P1 član izraz (46) mora se diskretizirati.

Diskretizirana funkcija izraza (46) izgleda ovako:

$$x(n) = \frac{T_0}{T + T_0} \left[\frac{T}{T_0} x(n-1) + K_p u(n) \right] \quad (47)$$

U izrazu (47) $x(n)$ i $u(n)$ su trenutne vrijednosti izlazne i ulazne varijable, dok $x(n)$ označava vrijednost izlazne varijable u prethodnom koraku, T_0 je vrijednost perioda uzrokovanja.

Dakle funkcija koju mreža mora naučiti ima dva ulaza i jedan izlaz, prema tome broj ulaznih neurona će biti 2, a izlaznih 1.

4.1 Set za učenje

Set za učenje se sastoji od 160 točaka i predstavlja odziv P1 člana na skokovitu pobudu. Skokovita pobuda koja je korištena da bi se dobio odziv P1 člana dana je matematičkom funkcijom:

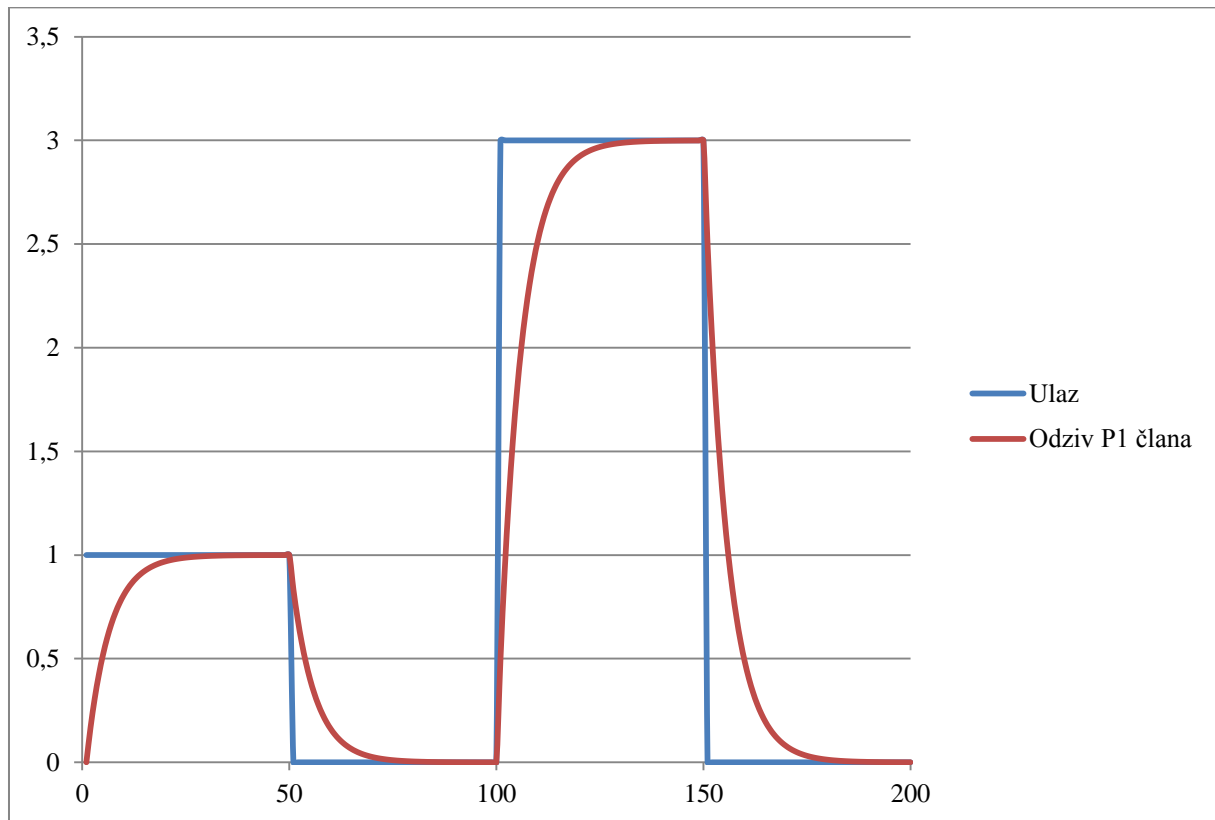
$$\begin{aligned} u(n) &= 1, & n &= 1, 2, \dots, 40 \\ u(n) &= 0, & n &= 41, 42, \dots, 80 \\ u(n) &= 3, & n &= 81, 82, \dots, 120 \\ u(n) &= 0, & n &= 121, 122, \dots, 160 \end{aligned} \quad (48)$$

Da bi se dobio odziv P1 člana, gornji skup točaka je unesen u funkciju (47) uz vrijednosti parametara:

$$\begin{aligned} T &= 1s, \\ T_0 &= 0.2s, \\ K_p &= 1 \end{aligned}$$

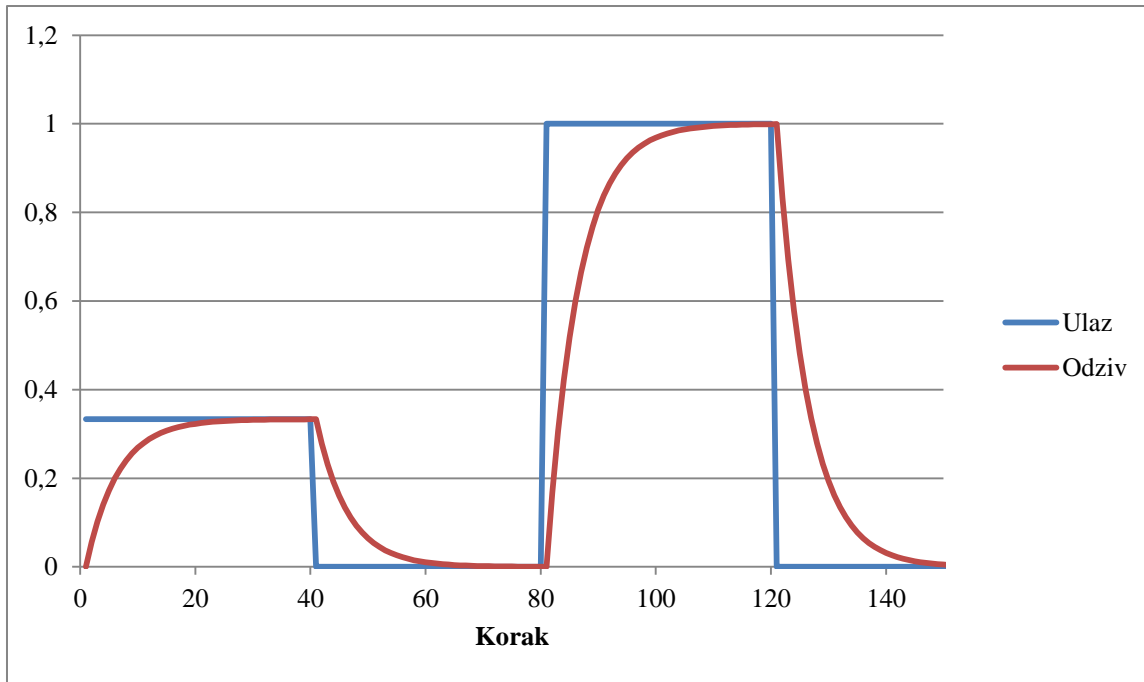
te je dobiven set za učenje prikazan na slici 5. Mreži se kao ulazi pokazuju dvije vrijednosti: $x(n-1)$ i $u(n)$, dok joj se kao željena vrijednost predstavlja $x(n)$. S obzirom da je za generiranje $x(n)$ vrijednosti potrebna vrijednost prethodnog koraka $x(n-1)$ dodana je još jedna točka (161.)

kako bi setovi za učenje i testiranje bili jednaki (ulazne vrijednosti su od točke 1 do 200, a izlazne od 2 do 201, oba seta sada imaju 200 točaka).



Slika 4.1 Set za učenje

Treba uzeti u obzir neuronska mreža ne može proizvesti izlaz veći od jedan. Potrebno je ulaze u mrežu normirati. U ovom radu je to napravljeno tako da su vrijednosti seta za učenje podijeljeni faktorom 3. Naime time se dobiva set za učenje u kojem su ulazne varijable dobro razdvojene. Ulaz koji je bio 3 sada je 1, a ulaz koji je bio 1 sada je 0.333. Mreža brže uči setove za učenje u kojima su jasno odijeljene različite klase. Na slici 6 je prikazan normiran set za učenje .



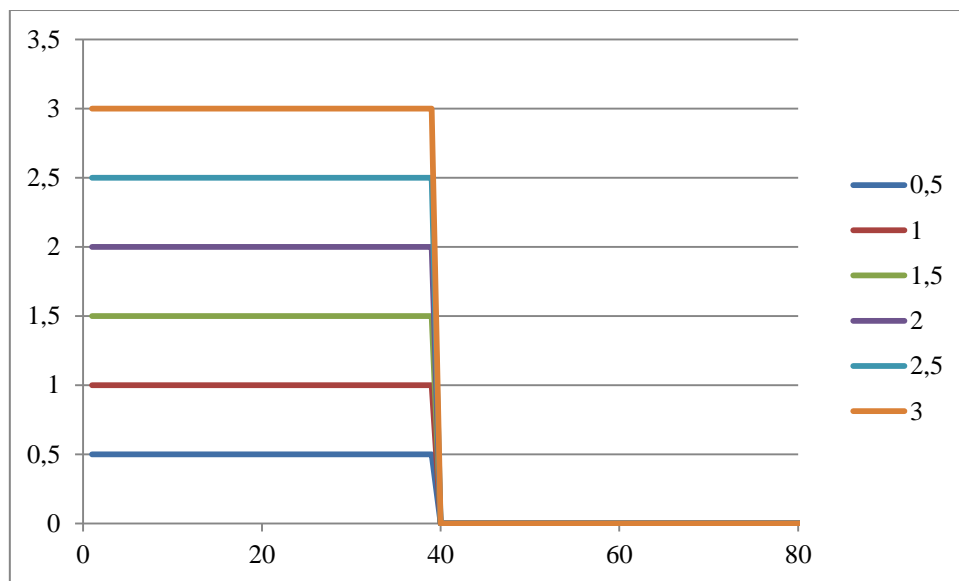
Slika 4.2 Normirani set za učenje

4.2 Set za testiranje

Set za testiranje se sastoji od 6 step signala. Trajanje signala je 40 točaka, a vrijednosti ulaza se mijenjaju od 0.5 do 3 s korakom od 0.5. Ulazi seta za testiranje dani su funkcijom:

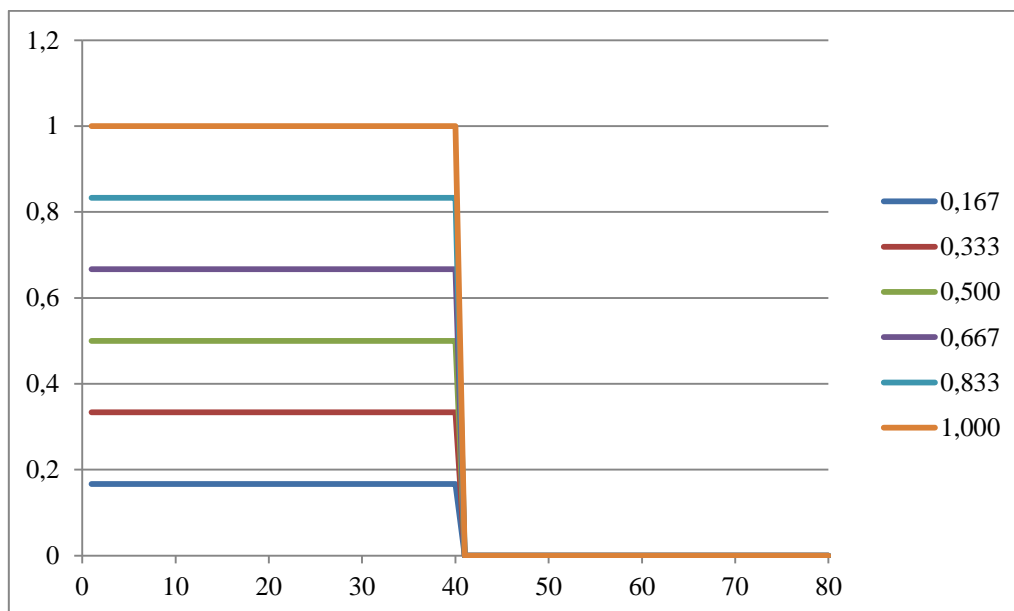
$$\begin{aligned}
 (1) \quad & u(n) = 0.5, & n = 1, 2, \dots, 40 \\
 & u(n) = 0, & n = 41, 42, \dots, 80 \\
 (2) \quad & u(n) = 1, & n = 1, 2, \dots, 40 \\
 & u(n) = 0, & n = 41, 42, \dots, 80 \\
 (3) \quad & u(n) = 1.5, & n = 1, 2, \dots, 40 \\
 & u(n) = 0, & n = 41, 42, \dots, 80 \\
 (4) \quad & u(n) = 2, & n = 1, 2, \dots, 40 \\
 & u(n) = 0, & n = 41, 42, \dots, 80 \\
 (5) \quad & u(n) = 2.5, & n = 1, 2, \dots, 40 \\
 & u(n) = 0, & n = 41, 42, \dots, 80 \\
 (6) \quad & u(n) = 3, & n = 1, 2, \dots, 40 \\
 & u(n) = 0, & n = 41, 42, \dots, 80
 \end{aligned} \tag{49}$$

Set testiranja dobiven uvrštavanjem ulaza (49) u funkciju (47) je dan na slici ispod.



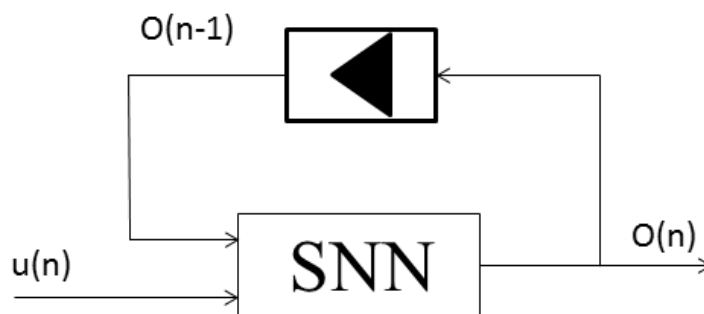
Slika 4.3 Set za testiranje

Set za testiranje se također, na isti način kao i set za učenje, normira. Normirani set za učenje je na slici 8.



Slika 4.4 Normirani set za testiranje

Kod učenja mreži se prikazuju dva ulaza i željeni izlaz. Pri testiranju mreži se prikazuje jedan ulaz, a drugi mreža je sama generirala u prethodnom koraku. Ovo se ostvaruje jediničnim kašnjenjem izlaza spojenog na ulaz (odnosno u programu se to ostvaruje spremanjem dobivene vrijednosti i njenim korištenjem u sljedećem koraku). Shema modela neuronske mreže s povratnom vezom dana je na slici ispod.



Slika 4.5 Model povratne veze

4.3 Usporedba različitih topologija neuronskih mreža

Zadatak rada je usporediti učenje i testiranje naučenih neuronskih mreža s jednim i dva skrivena sloja različitih topologija. Da bi se mogle uspoređivati jednoslojne i dvoslojne mreže one moraju imati isti broj parametara učenja. Uzmu li se sve jednoslojne mreže od najjednostavnije, s jednim skrivenim neuronom, pa do mreže sa sto skrivenih neurona, njih 11 će imati broj parametara učenja koji se može dobiti nekom dvoslojnom mrežom sa do 10 skrivenih neurona u skrivenim slojevima. Najveća odabrana jednoslojna mreža je 2-21-1. Postoje i mreže s više skrivenih neurona, no nastojalo se da broj parametara učenja bude otprilike dva puta manji od broja točaka u setu za učenje. Najmanja jednoslojna mreža je 2-3-1 i ona je ujedno i najmanja odabrana mreža. Ostale dvije jednoslojne mreže su odabrane tako da mreže budu jednako udaljene po broju skrivenih neurona. Dvoslojne mreže su potom birane tako da odgovaraju jednoslojnim. Na tablici ispod su dane odabrane jednoslojne mreže, njihove analogne dvoslojne mreže te broj parametara učenja.

Tablica 4.1 Topologije jednoslojnih i dvoslojnih mreža

Topologija jednoslojne mreže	Analogne dvoslojne mreže	Broj parametara učenja
2-3-1	2-1-3-1	13
2-9-1	2-3-7-1 2-8-2-1	45
2-15-1	2-8-4-1 2-5-7-1	65
2-21-1	2-7-7-1 2-8-6-1	85

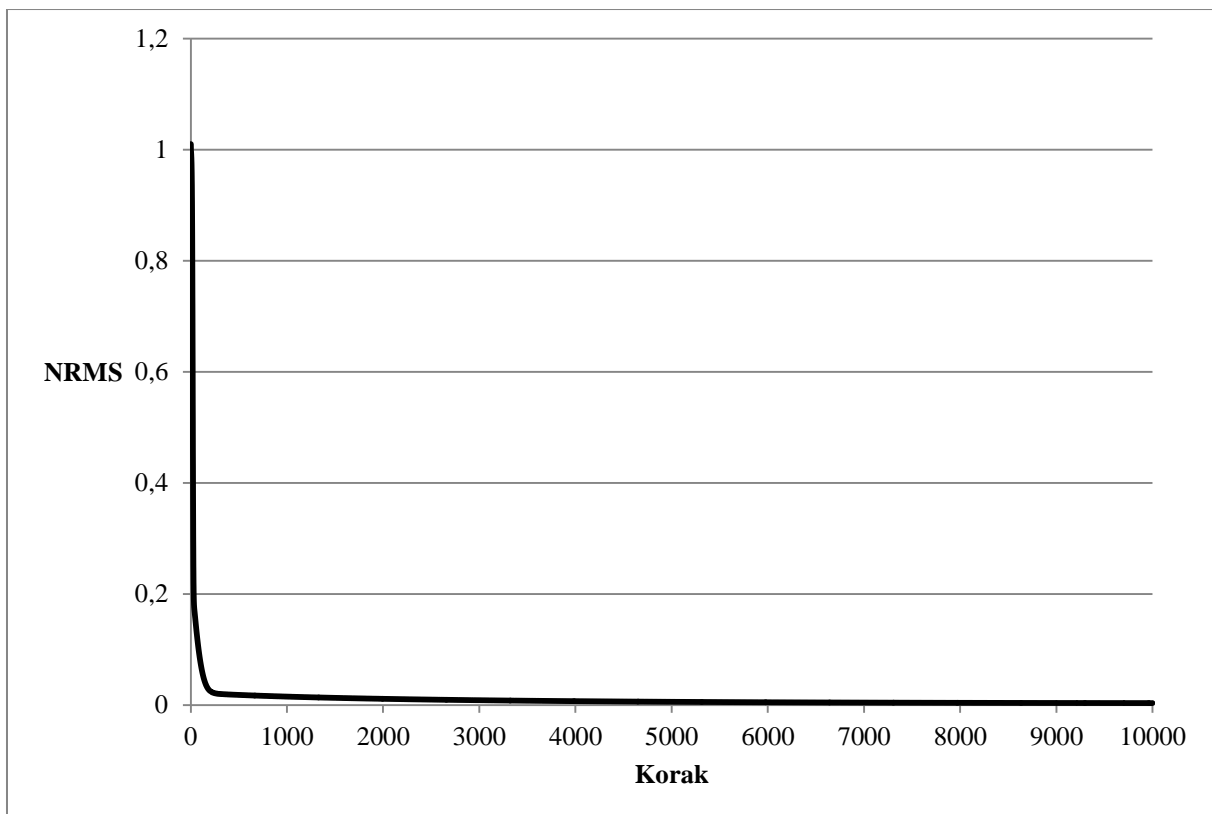
Treba napomenuti da u nijednoj od ovih konfiguracija nisu uračunati Bias neuroni, dakle npr. jednoslojna mreža 2-3-1 će nakon pridodavanja bias neurona imati konfiguraciju 3-4-1 itd.

4.4 Modeliranje P1 člana jednoslojnom mrežom

Prvo će biti obrađene mreže s jednim skrivenim slojem. Broj neurona u skrivenom sloju će se biti prama gornjoj tablici: 3, 9, 15, 21. Testiranje će se vršiti na dva načina. Prvi način će biti da se mreži zada traženi NRMS, a drugi da se mreži da određen broj koraka za učenje. Nakon provedenih učenja promatrati će se njihov učinak u nekoliko kategorija. Za učenje do određenog NRMS-a to će biti broj koraka, vrijeme učenja te će se analizirati generalizacijska svojstva. Kod učenja do određenog broja koraka to će biti postignuti NRMS, vrijeme učenja

te će se također analizirati generalizacijska svojstva. Kod oba pristupa mijenjat će se parametri učenja (koeficijent brzine učenja, momentum prvog i drugog reda) te će se analizirati njihov utjecaj na učenje mreže.

Broj koraka do kojeg mreže uče određen je eksperimentalno. Naime prvoj mreži koja se analizira, mreži 2-3-1 (dakle s tri skrivena neuronom), dan je proizvoljan broj koraka (10000) te je promatrana promjena NRMS-a koja je prikazana na slici 10. Treba napomenuti da mreža uči sigmoidalnom funkcijom po uzorku, koeficijent brzine učenja je $\eta=0.01$, a momentumi su isključeni.



Slika 4.6 NRMS mreže 2-3-1 za utvrđivanje broja koraka učenja

Na slici se vidi da mreža uči brzo do određenog koraka, a potom se učenje jako usporava. NRMS u 1. koraku je 1.0099, u 3000. koraku je 0.0086, a u 10000. koraku je 0.0035. Od 1. do 3000. koraka greška se smanjila 117 puta, a od 3000. koraka do 10000. greška se smanjila 2.5 puta. Vidljivo je da nakon 3000. koraka mreža troši puno vremena da smanji grešku za malen iznos. Za očekivati je da će mreža sa većim koeficijentima brzine učenja i momementumima prije doći u zasićenje. Stoga će se u učenjima gdje je mreži dan određen broj koraka koristiti broj od 5000 koraka (uzet veći broj radi sigurnosti). Isto tako će mrežama koje uče do određenog NRMS-a biti zadan NRMS od 0.008. Koeficijent brzine učenja će biti 0.01. Koeficijent brzine je dobiven eksperimentalno, slično kao i broj koraka, testirano je nekoliko učenja s višim koeficijentima, pošto mreže nisu konvergirale sa uključenim momentumom

prvog i drugog reda ($\alpha=0.9$) koeficijent učenja je smanjivan. Momentum će se mijenjati od 0.4 do 0.8 (korak od 0.2 između vrijednosti) za prvu mrežu (2-3-1), za ostale će se koristiti vrijednost 0.8 te po potrebi smanjivati, te će učenje biti uključeno momentum 2. reda. Momentum je odabran također eksperimentalno, mrežama se prvo dala vrijednost od $\alpha=0.9$ i primjećeno je da mreže sa momentumom 2. reda, iako konvergiraju, osciliraju u konvergenciji. Potom je vrijednost momentuma smanjena na $\alpha=0.8$ te su oscilacije uklonjene. Od aktivacijskih funkcija koristiti će se sigmoidalna aktivacijska funkcija te će sve mreže učiti po uzorku. Početne težine su određene nasumičnim odabirom, u intervalu $[-0.1, 0.1]$, i ne mijenjaju se za promjene parametara učenja. Raspon početnih težina je odabran tako da bude 10 puta manji od ulaznih vrijednosti kako aktivacijska funkcija ne bi ušla u zasićenje.

Tablica 4.2 Parametri učenja jednslojne neuronske mreže

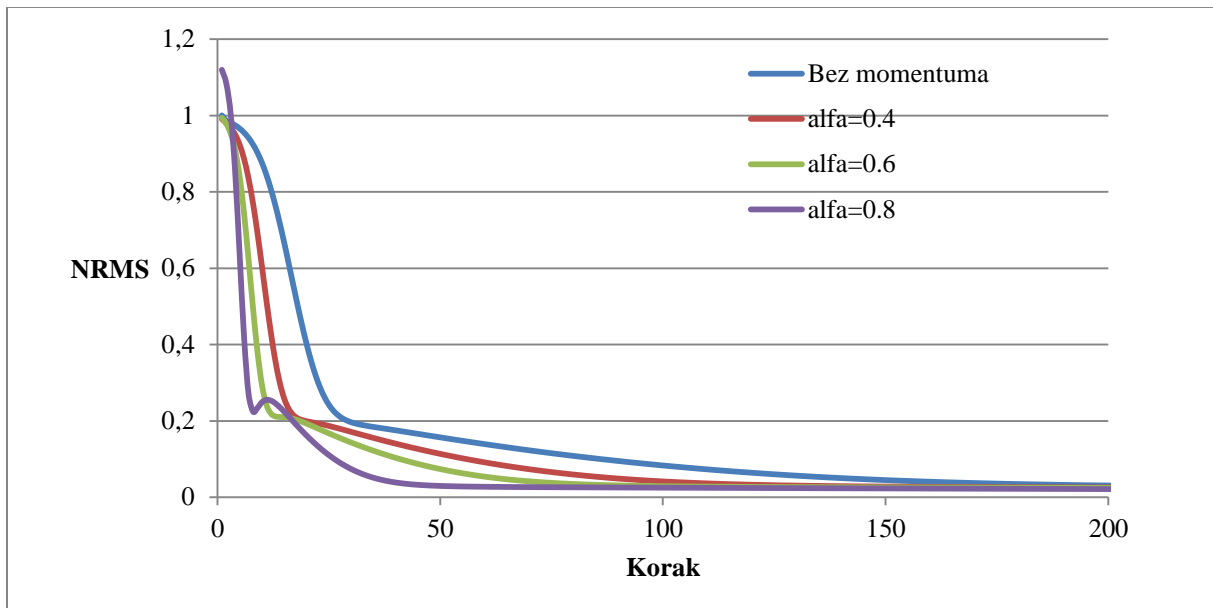
Zadani NRMS	Zadani broj koraka	Početne težine	Koeficijent brzine učenja (η)	Momentum prvog reda (α)
0,008	5000	$[-0,1, 0,1]$	0,01	0,4 0,6 0,8

4.4.1 Učenje do zadanog NRMS-a

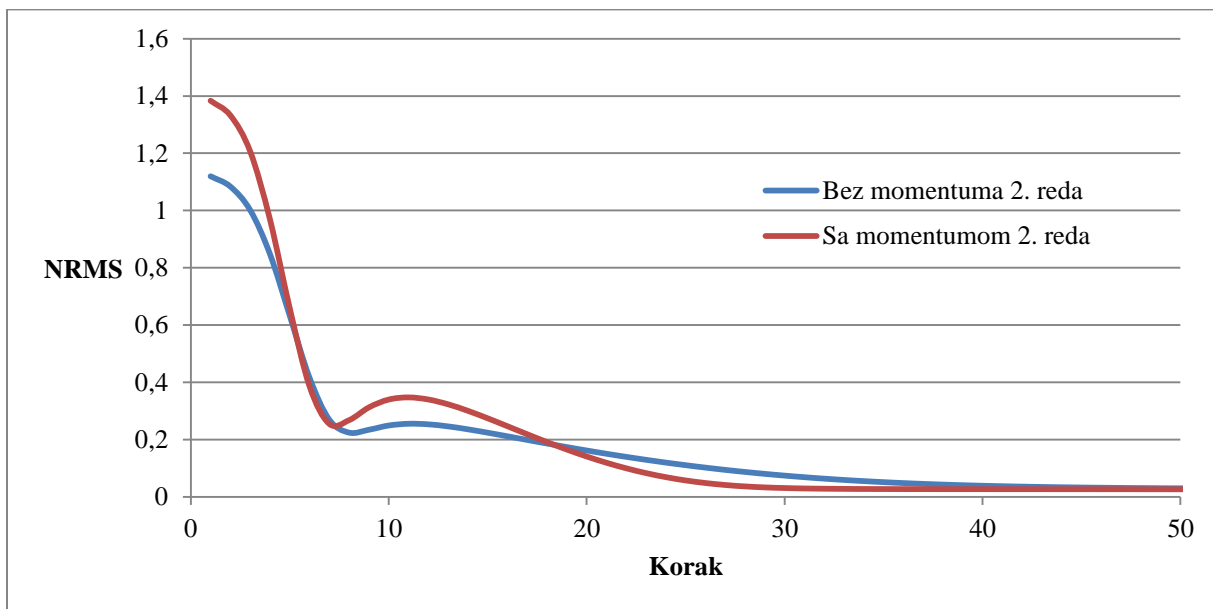
Mreže u prvom ispitivanju moraju učiti do zadanog NRMS-a, koji je, kao što je spomenuto, jednak $NRMS=0.008$. Pritom se mjeri broj koraka i vrijeme koje je mreži potrebno da dosegne taj NRMS. Rezultati učenja su prikazani u tablici ispod. U tablici su također prikazani i utjecaji momentuma prvog i drugog reda na brzinu učenja mreže. Utjecaj momentuma na brzinu učenja je dan i grafovima ispod tablice (slika 9.7 i 9.8).

Tablica 4.3 Rezultati učenja do zadanog NRMS-a jednoslojne mreže

Mreža	Koef. brzine učenja	Momentum 1. reda	Momentum 2. reda	Broj koraka	Vrijeme [s]
2-3-1	0,01	0	Ne	3964	16,7
2-3-1	0,01	0,4	Ne	2552	11,1
2-3-1	0,01	0,6	Ne	1843	8,3
2-3-1	0,01	0,8	Ne	1100	5,3
2-3-1	0,01	0,8	Da	814	4,2
2-9-1	0,01	0,8	Da	753	3,7
2-15-1	0,01	0,8	Da	789	4
2-21-1	0,01	0,8	Da	253	1,5



Slika 4.7 Utjecaj momentuma 1. reda na konvergenciju NRMS-a (mreža 2-3-1)



Slika 4.8 Utjecaj momentuma 2. reda na konvergenciju NRMS-a (mreža 2-3-1)

4.4.2 Učenje do zadanog koraka

Nakon učenja do zadanog NRMS-a, drugo ispitivanje se vrši na način da se mrežama da određeni broj koraka. Nakon što mreža izvede učenje promatraju se postignut NRMS te vrijeme koje je mreži bilo potrebno. Potom se mreža pobudi setom za testiranje. U tablici ispod su dani izmjereni podaci te NRMS odziva mreže na pojedini test (razlika odziva mreže i pravog odziva P1 člana).

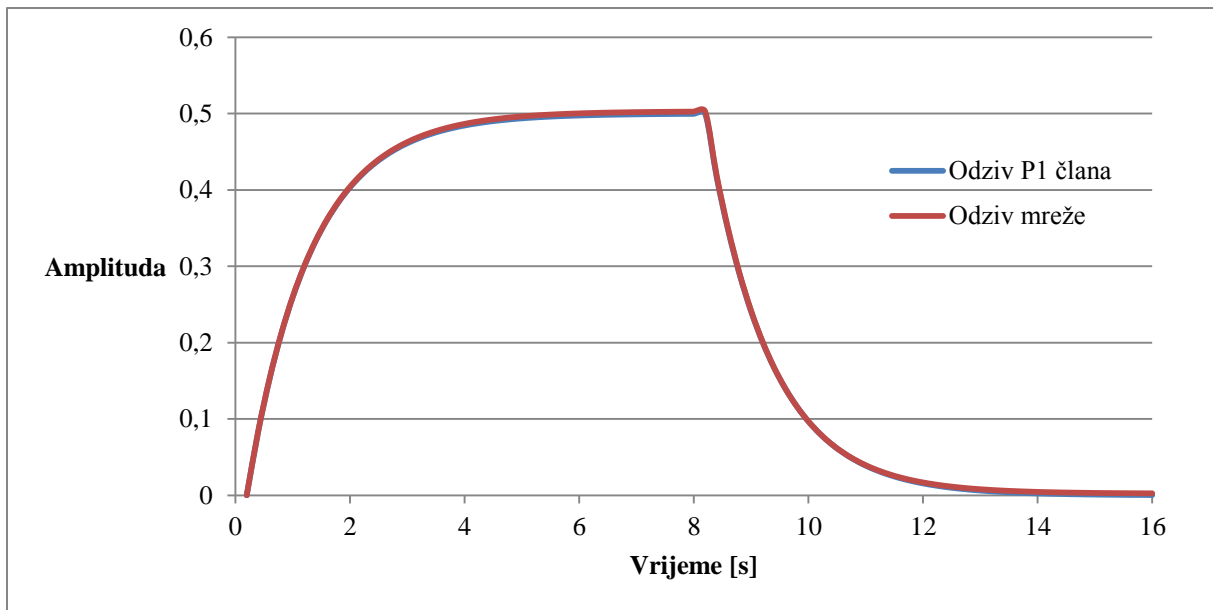
Tablica 4.4 Rezultati učenja do određenog koraka jednoslojne mreže

Mreža	NRMS	Vrijeme potrebno	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
2-3-1	0,0069	20,7	0,1549	0,0644	0,0380	0,0738	0,0683	0,0323
2-9-1	0,0033	22,4	0,0273	0,0442	0,0512	0,0463	0,0312	0,0133
2-15-1	0,0027	21,6	0,0089	0,0296	0,0417	0,0414	0,0296	0,0121
2-21-1	0,0030	21,3	0,0292	0,0425	0,0471	0,0416	0,0275	0,0119

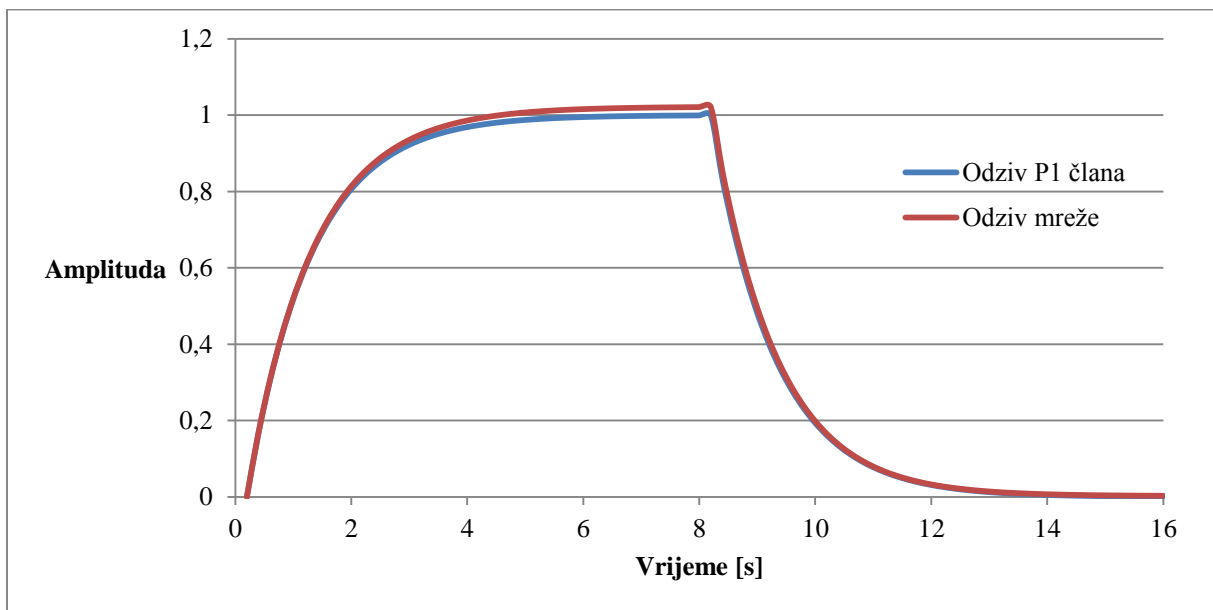
Iz gornjih tablica vidljivo je da momentumi prvog i drugog reda ubrzavaju učenje mreže. Također povećanje broja neurona u skrivenom sloju ubrzava učenje mreže. Što se tiče generalizacije, veći broj neurona bolje uči za jednak broj koraka.

4.4.3 Odziv mreže 2-21-1

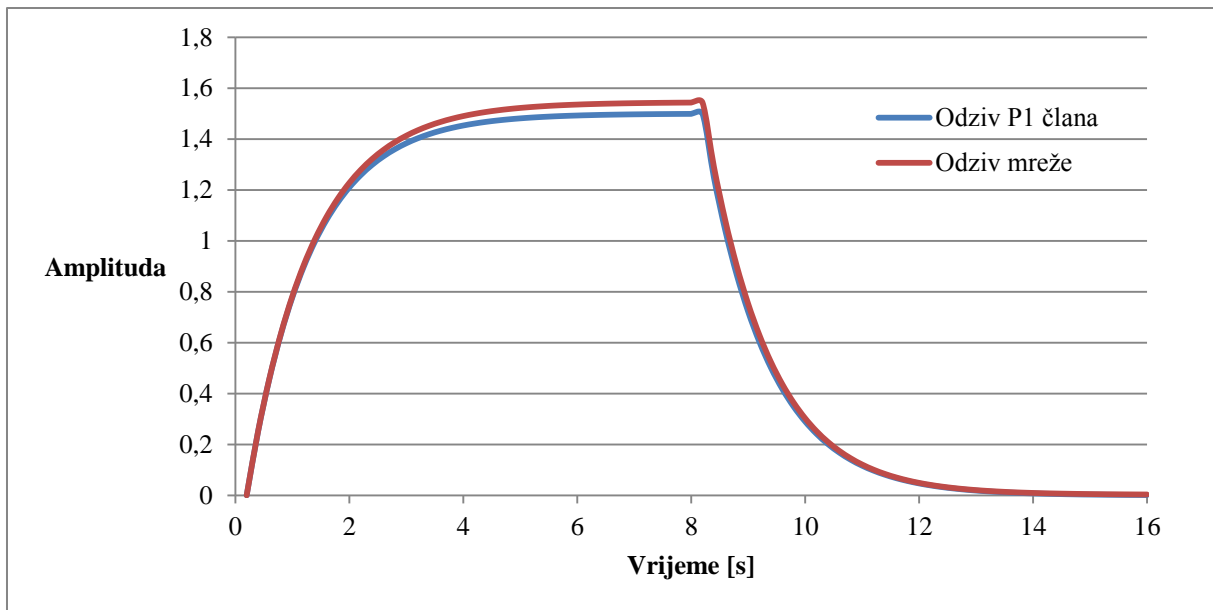
S obzirom da je mreža 2-15-1 imala najbolja generalizacijska svojstva dolje na grafovima je prikazan njen odziv na set za testiranje. Na grafovima je prikazan i odziv P1 člana da se vidi odstupanje od točne vrijednosti.



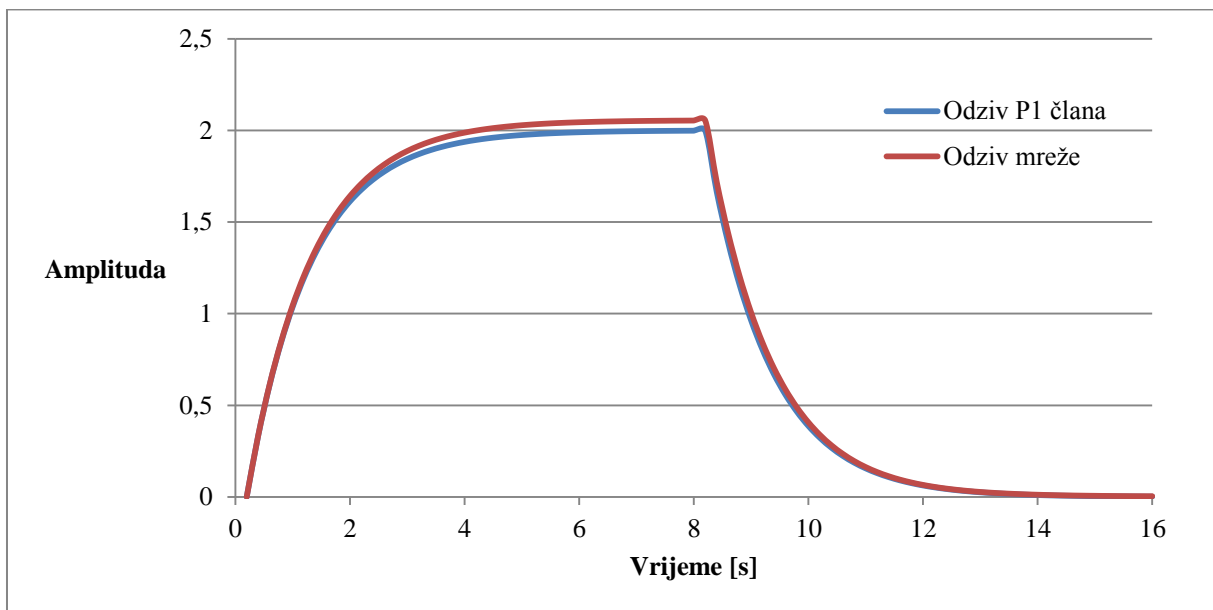
Slika 4.9 Odziv mreže na $u(n)=0.5, n=1,2...40, u(n)=0, n=41...80$



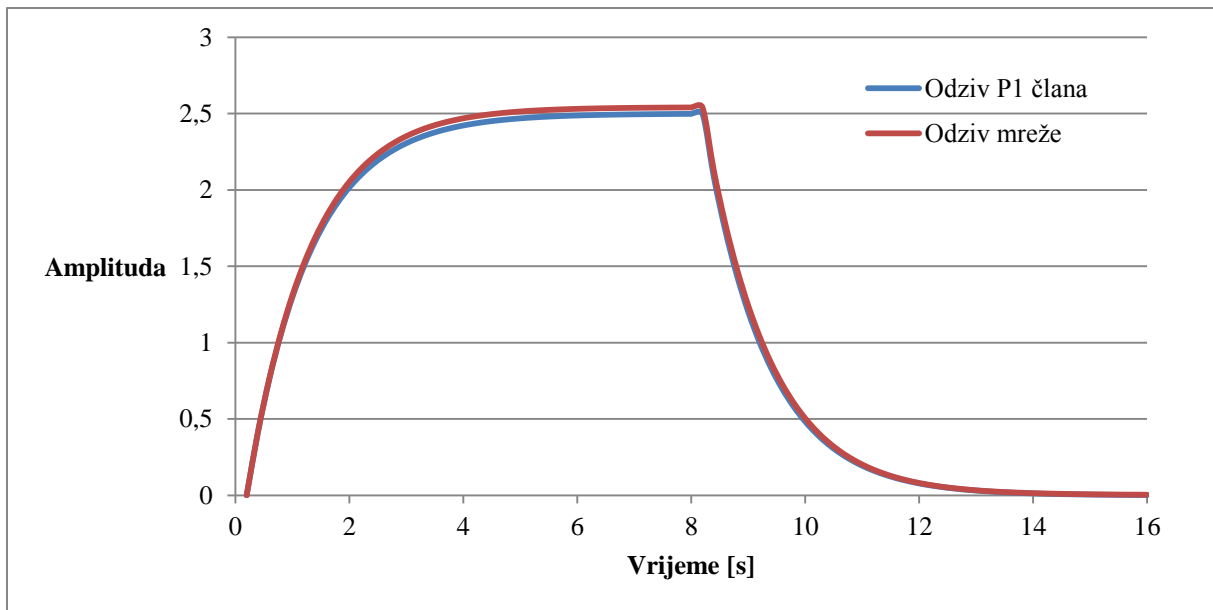
Slika 4.10 Odziv mreže na $u(n)=1, n=1,2...40, u(n)=0, n=41...80$



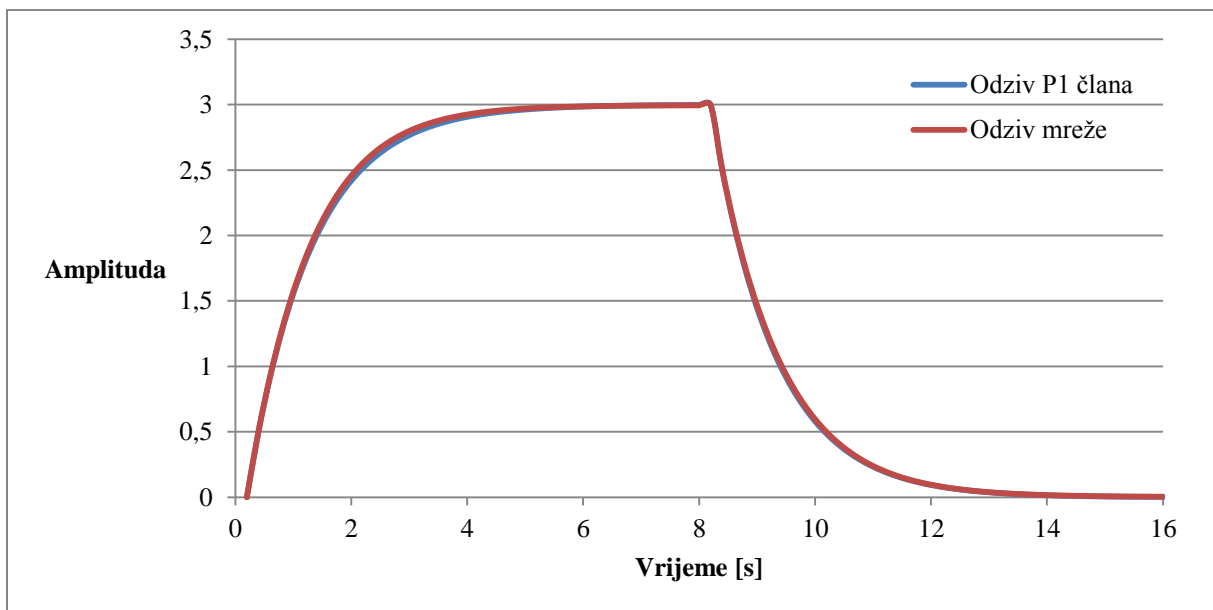
Slika 4.11 Odziv mreže na $u(n)=1.5, n=1,2...40, u(n)=0, n=41...80$



Slika 4.12 Odziv mreže na $u(n)=2, n=1,2...40, u(n)=0, n=41...80$



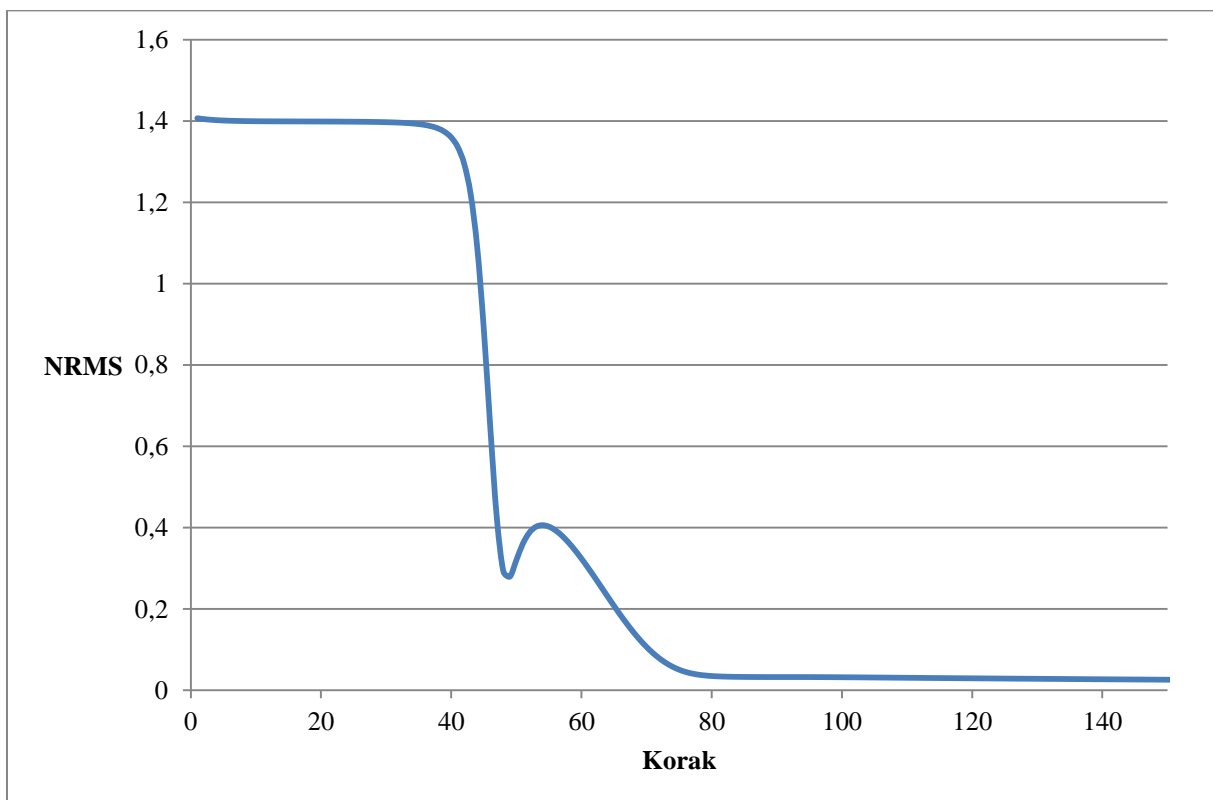
Slika 4.13 Odziv mreže na $u(n)=2.5, n=1,2...40, u(n)=0, n=41...80$



Slika 4.14 Odziv mreže na $u(n)=3, n=1,2...40, u(n)=0, n=41...80$

4.5 Modeliranje P1 člana dvoslojnom mrežom

Dvoslojne mreže će biti testirane na ista dva načina kao i jednoslojne mreže. NRMS do kojeg će mreže učiti je isti kao i kod jednoslojnih mreža ($NRMS=0,008$), te je broj koraka isto tako jednak (5000). Također su isti koeficijent brzine učenja ($\eta=0.01$), te raspon početnih težina i momentum prvog reda. Momentum drugog reda će biti isključivan ako narušava brzinu učenja. Naime mreže kod učenja mogu doći do lokalnog minimuma prikazanog na slici ispod. Prevelik momentum prvog reda i momentum drugog reda može pojačati utjecaj lokalnog minimuma, te produljiti vrijeme učenja. Ovo se može ublažiti smanjivanjem momentuma prvog reda kada je uključen momentum drugog reda. Međutim takva praksa je pokazala slične rezultate kao i kad bi momentum prvog reda ostao isti (jer se njegovim smanjivanjem usporava učenje), a samo se isključio momentum drugog reda. Pritom se uglavnom pokazalo da veći momentum prvog reda, iako može uvesti mrežu u lokalni minimum, ubrzava učenje. Stoga se neće smanjivati momentum prvog reda već će se samo, po potrebi, isključiti momentum drugog reda.



Slika 4.15 Pojava lokalnog minimuma u učenju mreže

Tablica 4.5 Parametri učenja dvoslojne mreže

Zadani NRMS	Zadani broj koraka	Početne težine	Koeficijent brzine učenja (η)	Momentum prvog reda (α)
0,008	5000	[-0,1 0,1]	0,01	0,8

4.5.1 Učenje do zadanog NRMS-a

Kao i kod jednoslojnih mreža prvo ispitivanje će biti učenje mreže do zadanog NRMS-a, a promatrat će se broj koraka i vrijeme učenja.

Tablica 4.6 Rezultati učenja do zadanog NRMS-a dvoslojne mreže

Mreža	Koef. brzine učenja	Momentum 1. reda	Momentum 2. reda	Broj koraka	Vrijeme [s]
2-1-3-1	0,01	0	Ne	21750	164,3
2-1-3-1	0,01	0,8	Ne	14748	109,7
2-1-3-1	0,01	0,8	Da	18707	139,6
2-3-7-1	0,01	0	Ne	6644	47,6
2-3-7-1	0,01	0,8	Ne	6496	45,9
2-3-7-1	0,01	0,8	Da	6683	47,9
2-8-2-1	0,01	0	Ne	2428	17,7
2-8-2-1	0,01	0,8	Ne	740	6
2-8-2-1	0,01	0,8	Da	6428	45,4
2-5-7-1	0,01	0,8	Ne	884	7,2
2-8-4-1	0,01	0,8	Da	1093	8,7
2-7-7-1	0,01	0,8	Da	1027	8,2
2-8-6-1	0,01	0,8	Da	715	6

U gornjoj tablici je pokazano vrijeme učenja mreža. Kod prve tri mreže pokazan je detaljnije proces učenja, sa uključivanjem momentuma prvog i drugog reda. Za zadnje četiri mreže je pokazan samo najbolji odziv i uvjeti u kojima je to postignuto.

Iz gornje tablice se može vidjeti da su dvoslojne mreže sporije od jednoslojnih mreža s jednakim brojem parametara učenja. To se moglo očekivati s obzirom da učenje dvoslojnih mreža zahtjeva učenje dva sloja što troši više procesorskog vremena. Također pokazalo se da su dvoslojne mreže osjetljivije na lokalni minimum što je kod nekih onemogućilo korištenje momentuma drugog reda kao metodu ubrzanja učenja. Jednoslojne mreže su brže prelazile područje lokalnog minimuma uz manji trošak vremena.

4.5.2 Učenje do zadanog koraka

Mreže koje se uče do zadanog koraka imaju iste parametre učenja kao i mreže koje su u prethodnom testu najbrže učile izuzev momentuma drugog reda koji je ovdje isključen za sve mreže.

Tablica 4.7 Rezultati učenja do zadanog koraka dvoslojne mreže

Mreža	NRMS	Vrijeme potrebno	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
2-1-3-1	0,012	40,6	0,3256	0,1604	0,0595	0,1142	0,0974	0,04
2-3-7-1	0,0072	37,4	0,2037	0,0918	0,0381	0,0738	0,0665	0,0279
2-8-2-1	0,0067	37,7	0,1913	0,0843	0,0369	0,0709	0,0637	0,0267
2-5-7-1	0,0054	37,8	0,1511	0,0636	0,0307	0,0585	0,0530	0,0222
2-8-4-1	0,0057	37,8	0,1601	0,0689	0,0314	0,0603	0,0548	0,0231
2-7-7-1	0,0057	37,5	0,1591	0,0684	0,0312	0,0599	0,0545	0,0230
2-8-6-1	0,0047	38	0,1288	0,0509	0,0294	0,0540	0,0481	0,0199

4.6 Odzivi mreža na druge pobudne funkcije

Mreže se mogu testirati i na druge pobudne funkcije kako bi se pokazalo da njihovo učenje nije bilo oponašanje odziva već učenje izraza (47). Mreže će se pobuditi trima pobudnim funkcijama, nagibnom, sinusnom i negativnom step funkcijom. Od jednoslojnih mreža će se testirati mreža 2-15-1, a od dvoslojnih njena analogna mreža 2-5-7-1.

4.6.1 Nagibne pobudna funkcije

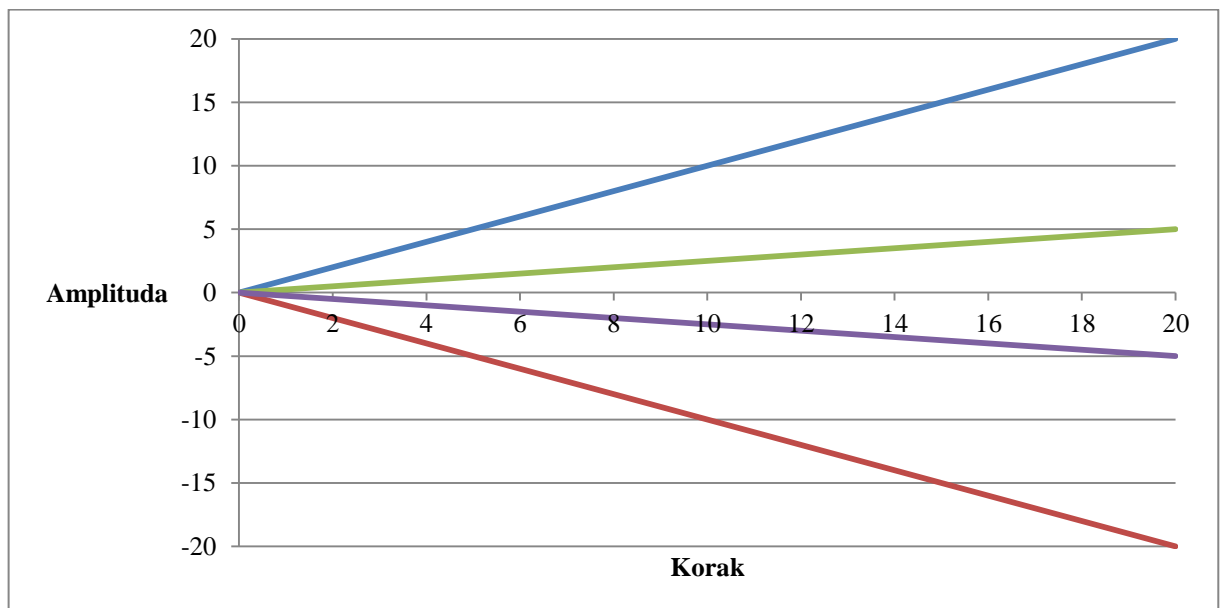
Nagibne funkcije su dane matematičkim izrazom:

$$\begin{aligned} u_1(n) &= \frac{n}{4}, & n = 0, 2, \dots, 20 \\ u_2(n) &= -\frac{n}{4}, & n = 0, 2, \dots, 20 \end{aligned} \quad (50)$$

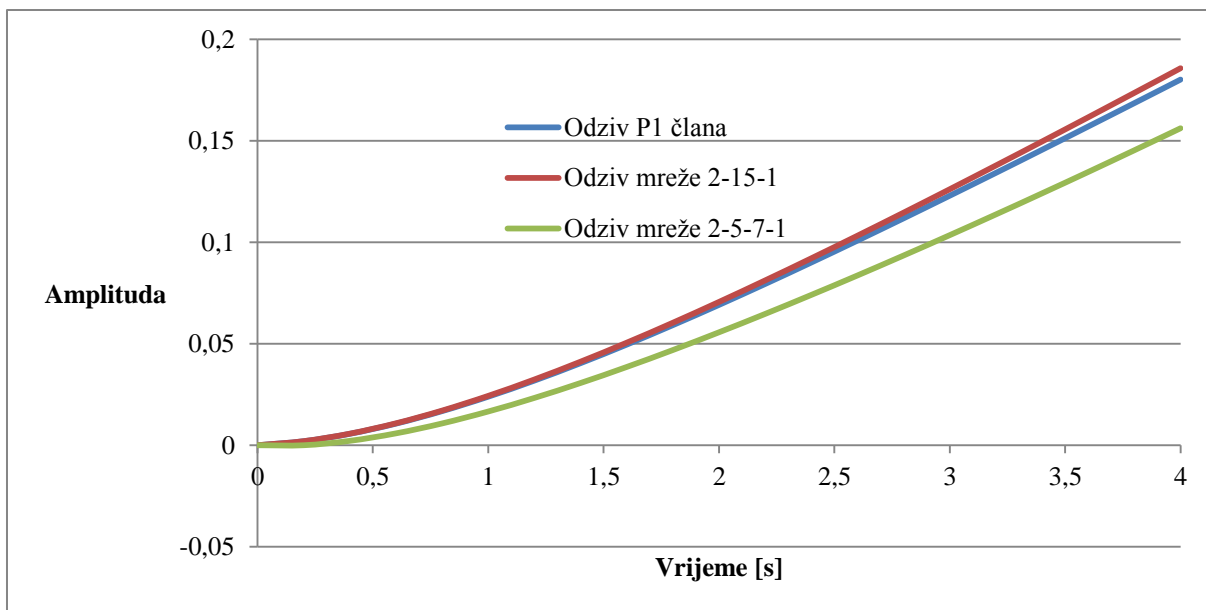
Također uz pozitivnu mreži će se dati i negativna nagibna funkcija:

$$\begin{aligned} u_3(n) &= n, & n = 0, 2, \dots, 20 \\ u_4(n) &= -n, & n = 0, 2, \dots, 20 \end{aligned} \quad (51)$$

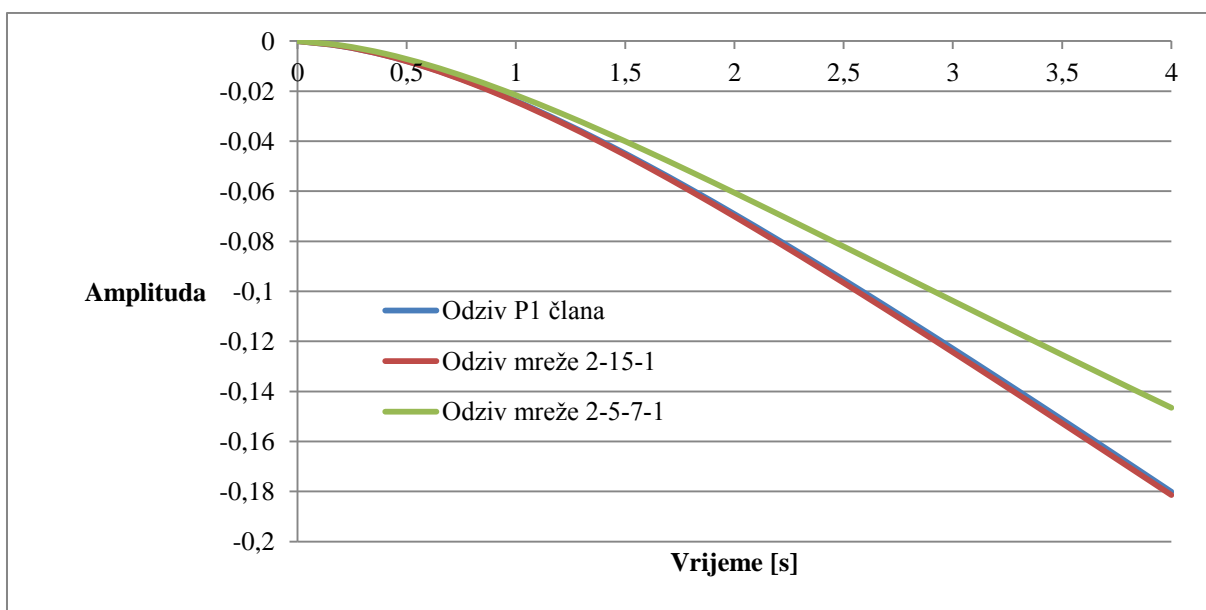
Pobudne funkcije treba normirati i to tako da njihova vrijednost ne prelazi vrijednost 1 jer je tako normiran i set za učenje. S obzirom da nagibne pobudne funkcije imaju 21 točku dijeliti će se faktorom 21.



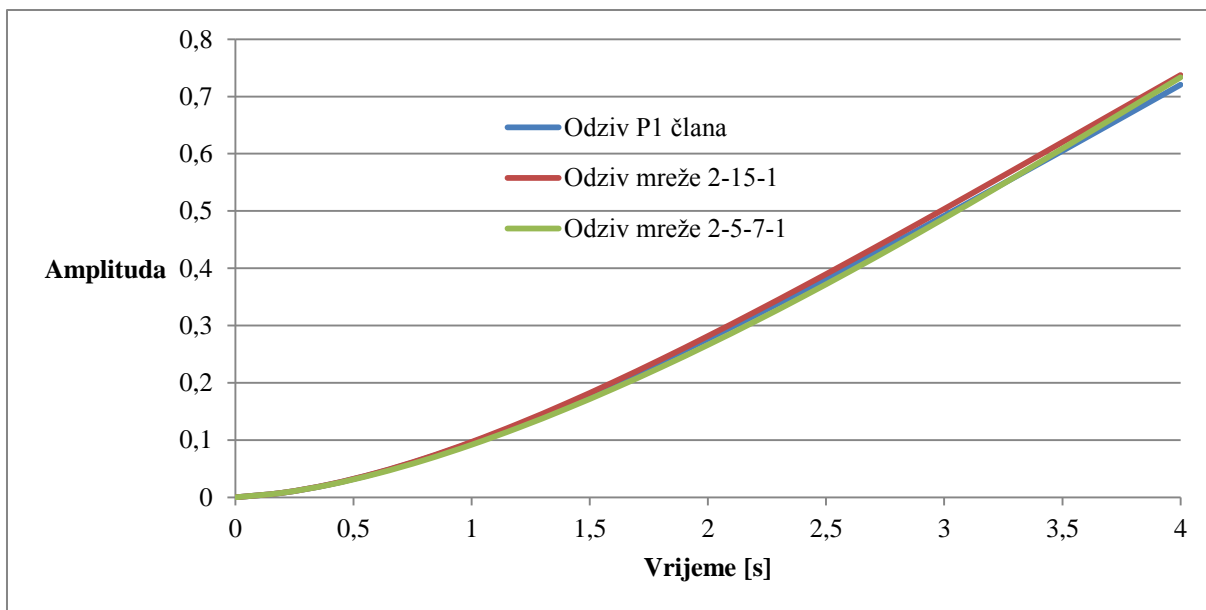
Slika 4.16 Pobudne nagibne funkcije



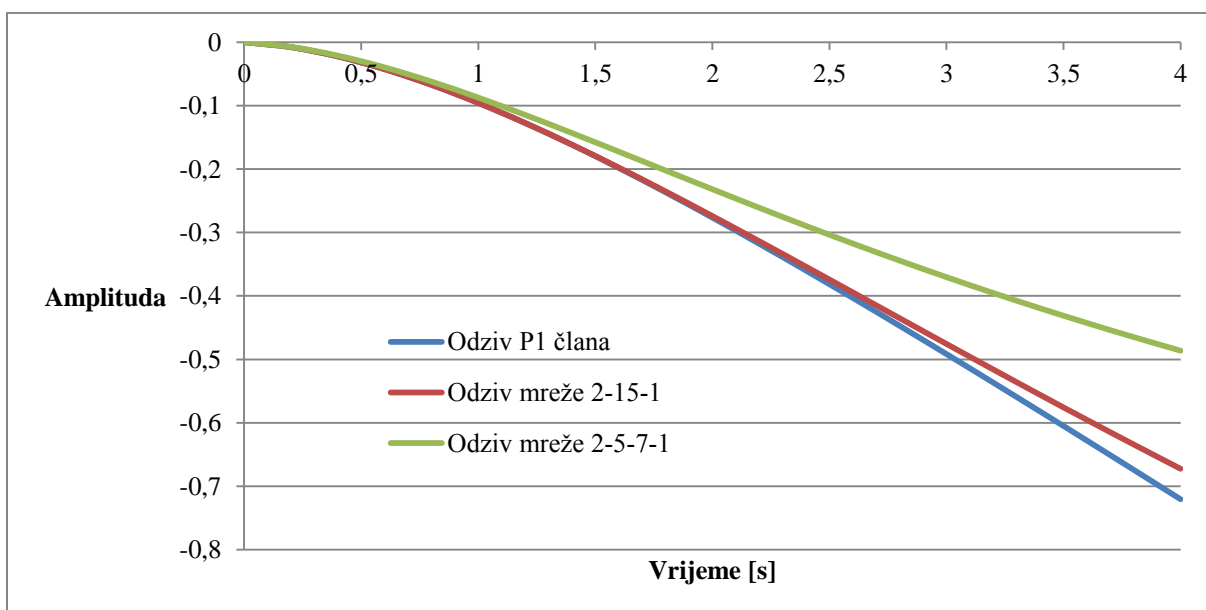
Slika 4.17 Odziv mreža na pobudnu funkciju 1



Slika 4.18 Odziv mreža na negativnu nagibnu funkciju 2



Slika 4.19 Odziv mreža na pozitivnu nagibnu funkciju 3



Slika 4.20 Odziv mreža na negativnu nagibnu funkciju 4

Tablica 4.8 Odzivi mreža na nagibne funkcije

Mreža	Pobudna funkcija	NRMS
2-15-1	Nagibna pozitivna (1)	0,0477
2-5-7-1	Nagibna pozitivna (1)	0,0799
2-15-1	Nagibna negativna (2)	0,0182
2-5-7-1	Nagibna negativna (2)	0,2856
2-15-1	Nagibna pozitivna (3)	0,0407
2-5-7-1	Nagibna pozitivna (3)	0,0319
2-15-1	Nagibna negativna (4)	0,0809
2-5-7-1	Nagibna negativna (4)	0,4708

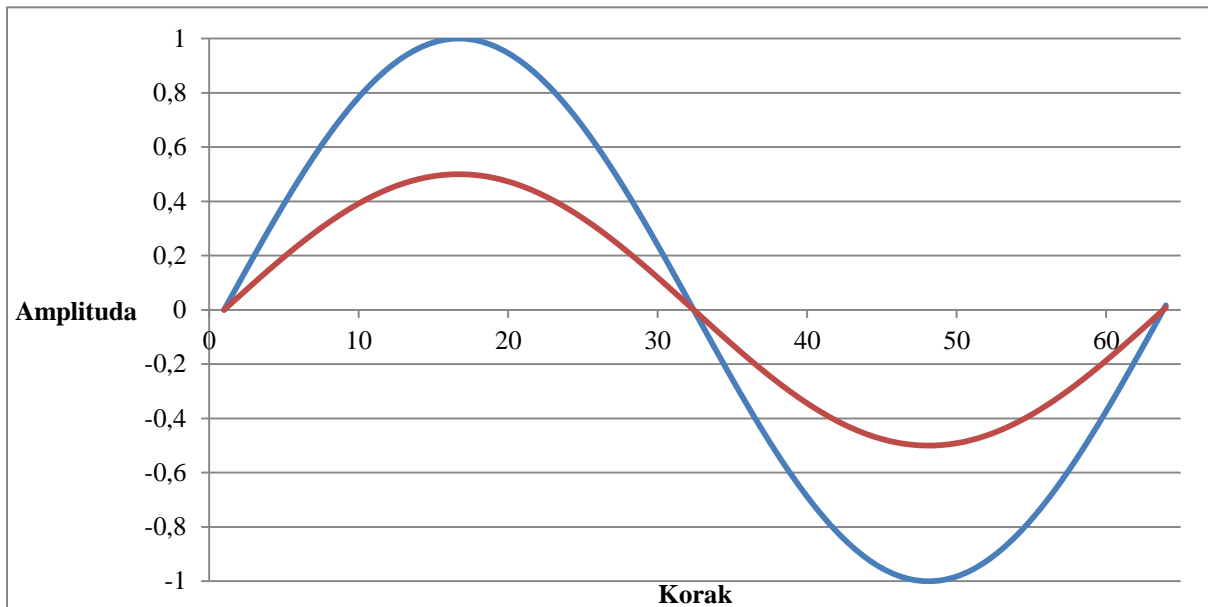
Mreže su bolje generalizirale u pozitivnom spektru u kojem je bio set za učenje. Pritom su i dvoslojne i jednoslojne mreže simale sličan NRMS. U negativnom spektru obje mreže su gore generalizirale nego u pozitivnom spektru, s većom amplitudom je bila i gora generalizacija , ali dvoslojne mreže su za red veličine lošije generalizirale nego jednoslojne mreže.

4.6.2 Sinusna pobudna funkcija

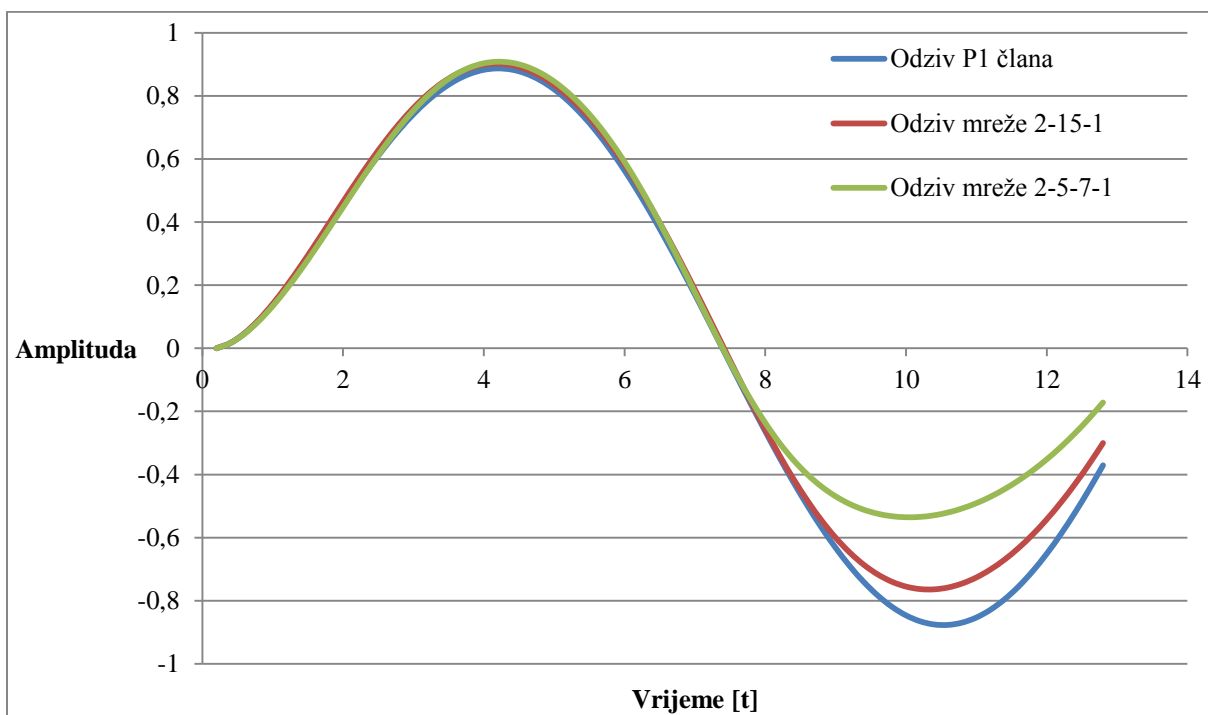
Sinusne funkcije su dane matematičkim izrazima:

$$\begin{aligned}
 u_1(n) &= \sin\left(\frac{n}{100}\right), & n = 0,2, \dots 63 \\
 u_2(n) &= 0.5\sin\left(\frac{n}{100}\right), & n = 0,2, \dots 63
 \end{aligned}
 \tag{52}$$

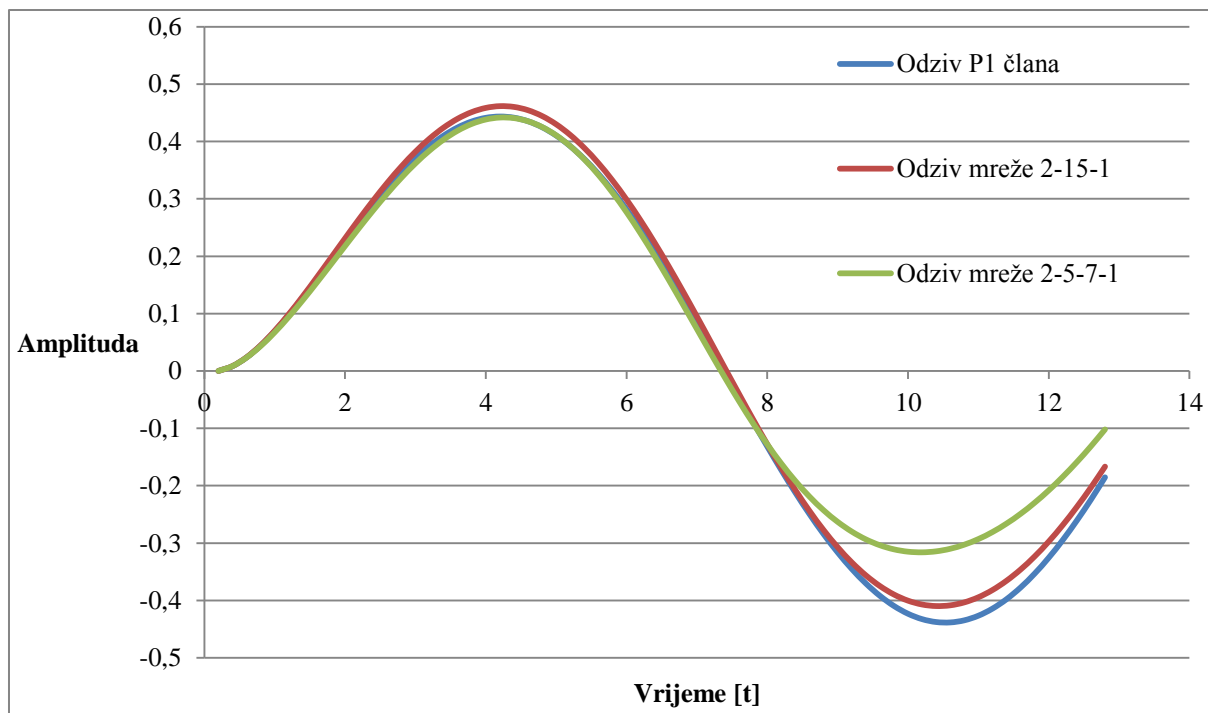
Sinusna funkcija se neće normirati s obzirom da njene vrijednosti ne prelaze vrijednost 1 tj. -1. Izgled pobudnih sinusnih funkcija je dan na slici ispod.



Slika 4.21 Sinusne pobudne funkcije



Slika 4.22 Odziv mreža na sinusnu pobudu 1



Slika 4.23 Odziv mreža na sinusnu pobudu 2

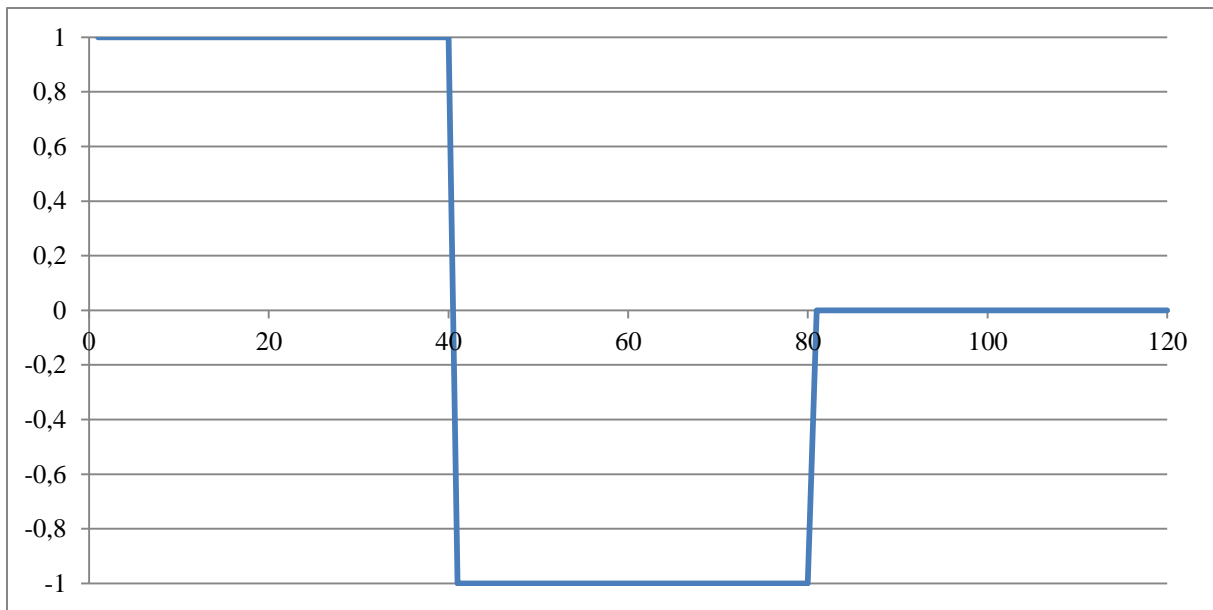
Ponovno se može vidjeti isto, mreže daju dobar odziv u pozitivnom dijelu funkcije, a u negativnom daju lošiji odziv. Opet se dvoslojna mreža pokazala inferiornijom od jednoslojne u negativnom dijelu funkcije. Treba primjetiti i da sa većom amplitudom pobudne funkcija jednoslojna mreža više griješi.

4.6.3 Negativna step pobudna funkcija

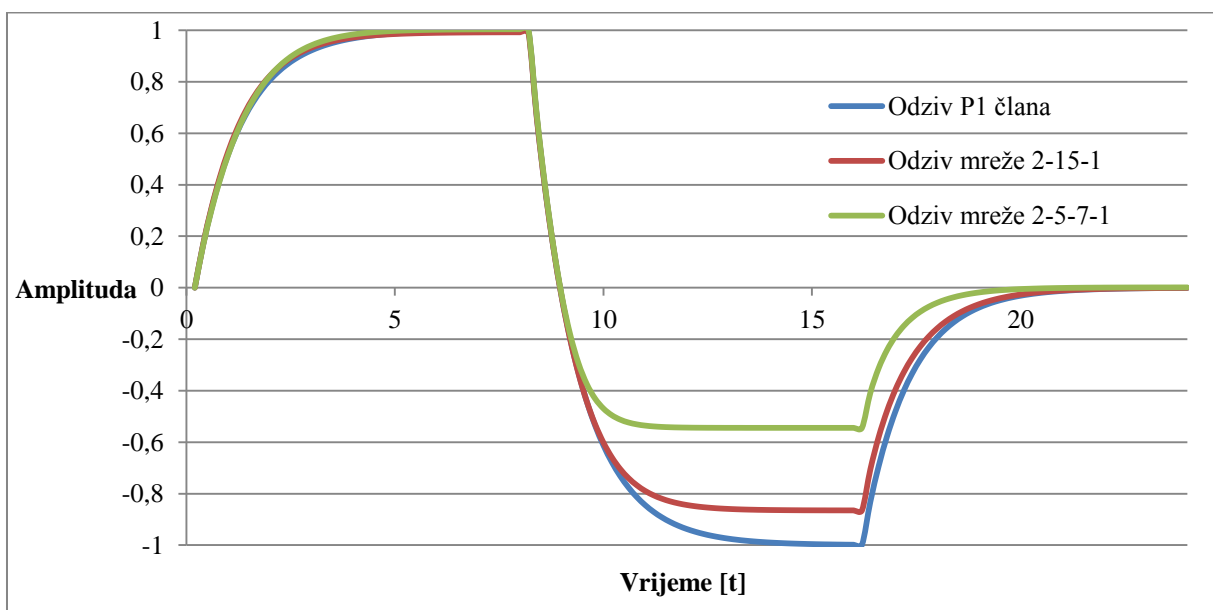
S obzirom da su mreže pokazale lošiju generalizaciju u negativnim dijelovima pobudnih funkcija testirati će se još i odziv na step funkciju, no u negativnom spektru.

$$\begin{aligned}
 u(n) &= 3, & n &= 1, 2, \dots, 40 \\
 u(n) &= -3, & n &= 41, 42, \dots, 80 \\
 u(n) &= 0, & n &= 81, 82, \dots, 120
 \end{aligned}
 \tag{53}$$

Funkcija je jednako normirana kao i kod seta za učenja, dijeljenjem faktorom 3. Pobudna funkcija je prikazana na slici.



Slika 4.24 Step pobudna funkcija



Slika 4.25 Odziv mreža na step pobudu

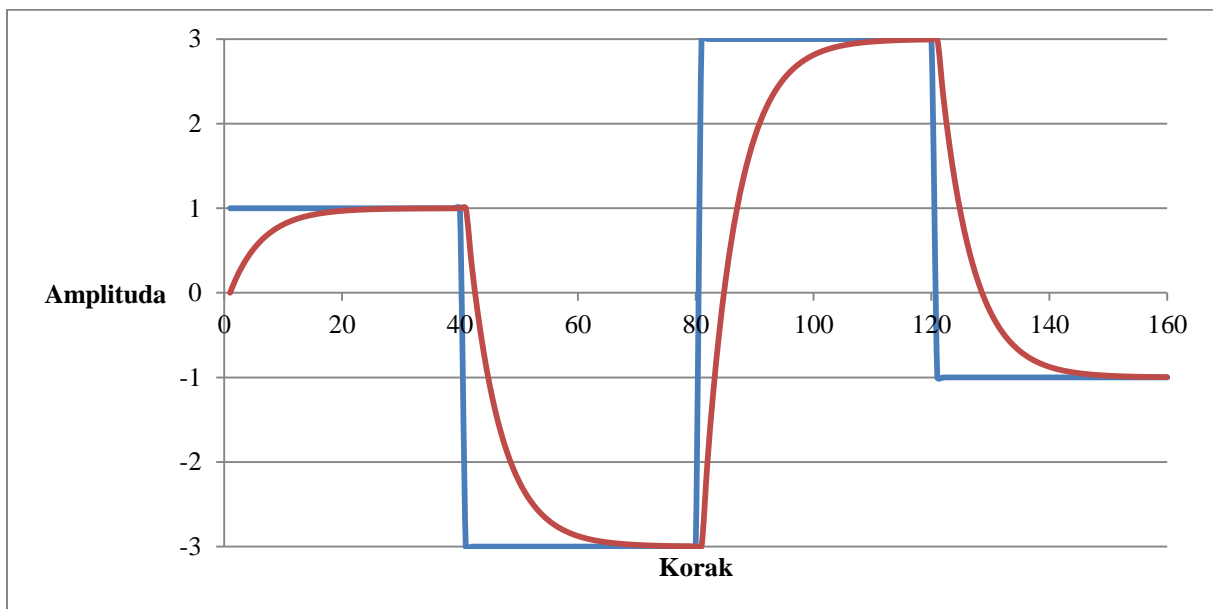
Rezultati su jednaki kao i u prethodnim ispitivanjima. U području u kojem nije bio set za učenje mreže lošije generaliziraju, pritom su dvoslojne lošije od jednoslojnih. Prema tome može se zaključiti da mreže slabije generaliziraju u području u kojem nije bio set za učenje.

4.6.4 Učenje mreža setom za učenje koji sadrži negativni spektar

S obzirom da su mreže opetovano griješile u generalizaciji kada im se dala funkcija u području u kojem nisu učile, mrežama 2-15-1 i 2-5-7-1 će se dati sljedeći set za učenje:

$$\begin{aligned} u(n) &= 1, & n &= 1, 2, \dots, 40 \\ u(n) &= -3, & n &= 41, 42, \dots, 80 \\ u(n) &= 3, & n &= 81, 82, \dots, 120 \\ u(n) &= -1, & n &= 121, 122, \dots, 160 \end{aligned} \quad (54)$$

Set za učenje je normiran kao i set za učenje dan izrazom (48). Nakon što mreže nauče set dati će im se ponovno nagibna funkcija 1 i 2 te će se usporediti odziv

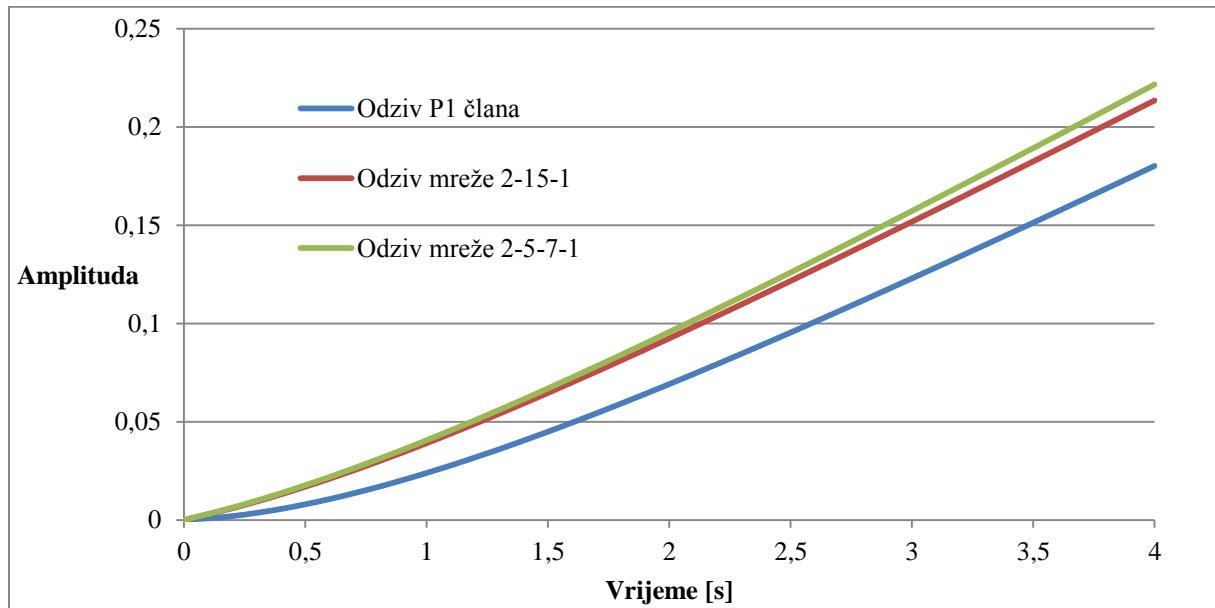


Slika 4.26 Set za učenje s negativnim vrijednostima

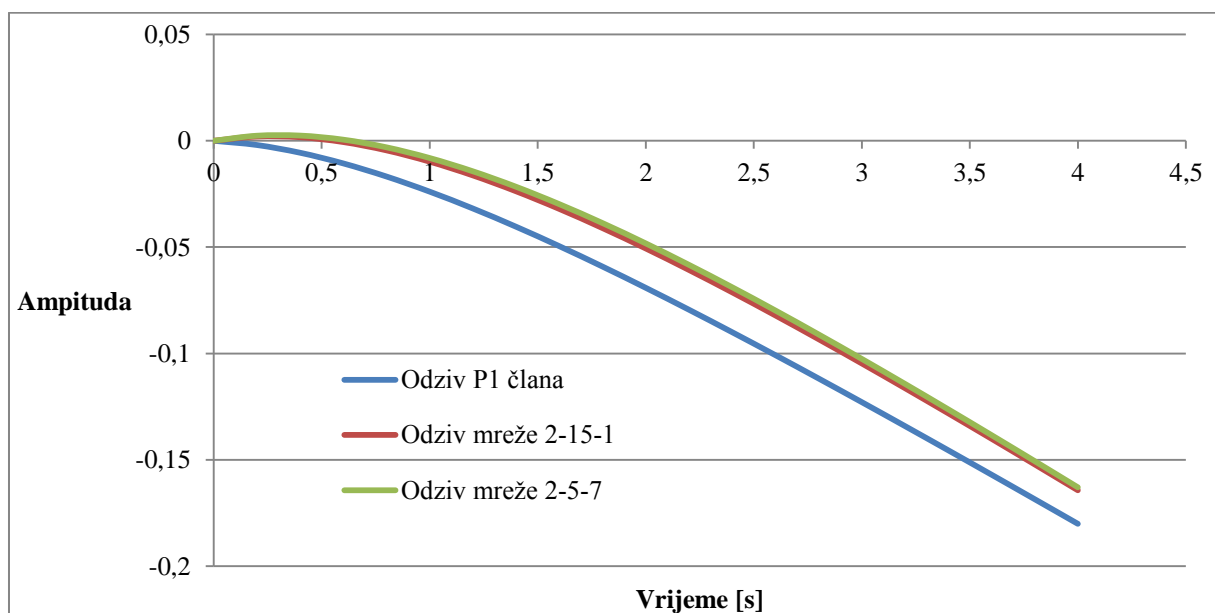
U tablicama ispod su dani rezultati testiranja mreže nagibnim pobudnim funkcijama 1 i 2. NRMS testiranja je lošiji od rezultata generalizacije mreža koje u setu za učenje imale samo pozitivne vrijednosti. Na slici 4.27 i 4.28 dan je odziv mreža.

Tablica 4.9 Rezultati učenja seta s negativnim vrijednostima

Mreža	Pobudna funkcija	NRMS
2-15-1	Nagibna pozitivna (1)	0,4214
2-5-7-1	Nagibna pozitivna (1)	0,4983
2-15-1	Nagibna negativna (2)	0,2856
2-5-7-1	Nagibna negativna (2)	0,3192



Slika 4.27 Odziv mreža na pozitivnu nagibnu pobudnu funkciju

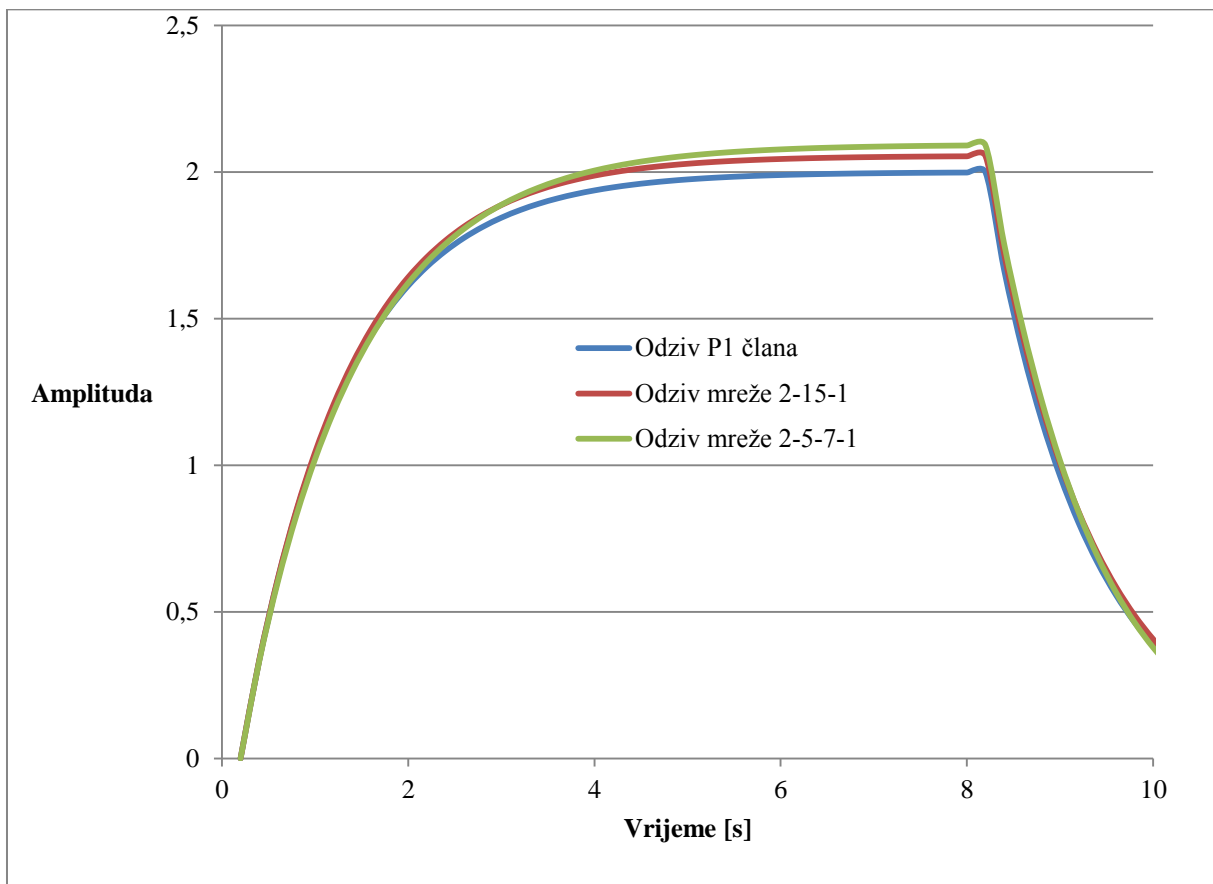


Slika 4.28 Odziv mreža na negativnu nagibnu pobudnu funkciju

5. Zaključak

Jednoslojne mreže su se pokazale brže i bolje u generalizaciji P1 člana nego dvoslojne mreže. Jednoslojne mreže su u svim topologijama bile bolje od svojih analognih dvoslojnih mreža. Dvoslojne mreže se s povećanjem broja parametara učenja približavaju jednoslojnim mrežama u NRMS-u učenja i NRMS-u testiranja no ne mogu ih nadmašiti. Također ne može se odrediti koja topologija dvoslojnih mreža, ona s više neurona u prvom skrivenom sloju ili obrnuto, je bolja u generalizaciji P1 člana.

Na slici ispod je prikazan odziv mreže 2-15-1, koja je bila najbolja od jednoslojnih mreža, odziv mreže 2-5-7-1, najbolje od dvoslojnih mreža s jednakim brojem parametara kao i 2-15-1 te odziv P1 člana.



Slika 5.1 Usporedba odziva mreže 2-15-1 i 2-5-7-1

U tablici ispod su dana usporedba NRMS-a na set za testiranje i vremena učenja.

Tablica 5.1 Usporedba mreže 2-15-1 i 2-5-7-1

Mreža	NRMS	Vrijeme potrebno	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
2-15-1	0,0027	21,6	0,0089	0,0296	0,0417	0,0414	0,0296	0,0121
2-5-7-1	0,0054	37,8	0,1511	0,0636	0,0307	0,0585	0,0530	0,0222

Vidljivo je da je jednoslojna mreža bila bolja u NRMS-u učenja, vremenu učenja i gotovo svim testovima. Ukoliko se uzme najbolja dvoslojna mreža i njena analogna jednoslojna mreža i usporedba se prikaže u tablici:

Tablica 5.2 Usporedba mreže 2-21-1 i 2-8-6-1

Mreža	NRMS	Vrijeme potrebno	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
2-21-1	0,0030	21,3	0,0292	0,0425	0,0471	0,0416	0,0275	0,0119
2-8-6-1	0,0047	38	0,1288	0,0509	0,0294	0,0540	0,0481	0,0199

Zaključak je opet isti, dvoslojne mreže su bile sporije, imale su lošiji NRMS učenja i lošiji NRMS testiranja.

Kada su se mrežama dale pobudne funkcije koje nisu učile, dvoslojne mreže su bile slične u generalizaciji kao i jednoslojne u pozitivnom spektru (u kojem se i set za učenje nalazio), u negativnom spektru su bile daleko lošije u generalizaciji od jednoslojnih mreža.

Prilikom učenja seta koji je imao negativne vrijednosti obje mreže su se pogoršale u generalizaciji, no ipak su jednoslojne mreže imale niži NRMS testiranja nego dvoslojne.

Dakle, glavni zaključak rada bi bio da povećanjem broja skrivenih slojeva nije dobiveno ništa u pogledu brzine učenja, broja koraka učenja ni kvalitete generalizacije mreže.

Treba na kraju napomenuti da ovi zaključci vrijede samo za problem modeliranja P1 člana. Zaključak da su dvoslojne mreže slabije u generalizaciji nego jednoslojne ne može se ekstrapolirati na sve probleme koje neuronske mreže mogu učiti već, kao što je napomenuto, vrijedi samo za ispitani problem.

6. Literatura

- [1] Branko Novaković, Dubravko Majetić, Mladen Široki, *Umjetne neuronske mreže*, Fakultet strojarstva i brodogradnje, Zagreb, 2011.
- [2] Dubravko Majetić, Podloge za predavanja iz kolegija Neuronske mreže, 2014.
- [2] Ludwig von Bertalanffy, *General System Theory*, Braziller, New York, 1969.
- [3] Branko Novaković, *Regulacijski sistemi*, Sveučilišna naklada, Zagreb, 1990.
- [4] Yann LeCun, Leon Bottou, Genevieve B. Orr, Klaus-Robert Müller, „*Efficient BackProp*“, Springer, 1998.
- [5] Ivan Petrović, Mato Baotić, Nedjeljko Perić, *Inteligentni sustavi upravljanja: Neuronske mreže, evolucijski i genetički algoritmi*, skripta, Fakultet elektrotehnike i računarstva, 2011.
- [7] Simon Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice-Hall, Inc., 1999.
- [8] J.M. Zurada, *Artificial Neural Systems*, W.P. Company, USA, 1992.
- [9] James A. Anderson, *An Introduction to Neural Networks*, Massachusetts Institute of Technology, 1995.
- [10] Maureen Caudill, *Neural Network Primer: Part I*, AI Expert, Feb. 1989
- [11] www.enciklopedija.hr