

Univerza
v Ljubljani
Fakulteta
*za gradbeništvo
in geodezijo*

*Janova 2
1000 Ljubljana, Slovenija
telefon (01) 47 68 500
faks (01) 42 50 681
fgg@fgg.uni-lj.si*



Univerzitetni program Gradbeništvo,
Konstrukcijska smer

Kandidatka:
Nataša Šinkovec

Uporaba umetnih nevronske mreže za oceno trdnosti lesenih elementov

Diplomska naloga št.: 3008

Mentor:
prof. dr. Goran Turk

Somentor:
izr. prof. dr. Dejan Zupan

Ljubljana, 28. 3. 2008

IZJAVA O AVTORSTVU

Podpisana **NATAŠA ŠINKOVEC** izjavljam, da sem avtorica diplomske naloge z naslovom:
**»UPORABA UMETNIH NEVRONSKIH MREŽ ZA OCENO TRDNOSTI LESENIH
ELEMENTOV«.**

Izjavljam, da prenašam vse materialne avtorske pravice v zvezi z diplomsko nalogo na UL,
Fakulteto za gradbeništvo in geodezijo.

Ljubljana, 15.03.08

Podpis:

BIBLIOGRAFSKO – DOKUMENTACIJSKA STRAN IN IZVLEČEK

UDK:	004:624.011.1:624.07(043.2)
Avtor:	Nataša Šinkovec
Mentor:	izr. prof. dr. Goran Turk, univ. dipl. inž. grad.
Somentor:	doc. dr. Dejan Zupan, univ. dipl. inž. mat.
Naslov:	Uporaba umetnih nevronske mreže za oceno trdnosti lesenih elementov
Obseg in oprema:	81 str., 33 pregl., 5 gr., 24 sl., 42 en.
Ključne besede:	umetne nevronske mreže, raztros podatkov, karakteristike lesa, trdnost lesa.

Izvleček

Pri načrtovanju konstrukcij ima poleg pravilnega projektiranja in izvedbe konstrukcije, ključen pomen konstrukcijski material. Ta mora imeti zadovoljivo trdnost, ki zagotavlja ustrezno nosilnost konstrukcije. Za določitev lastnosti lesenih elementov se uporabljajo destruktivne metode, kar pomeni, da preizkušane obremenimo do porušitve, in nedestruktivne metode. V tej diplomski nalogi smo za oceno trdnosti lesenih elementov uporabili umetne nevronske mreže. Uporabili smo bazo podatkov 293 lesenih preizkušancev, za katere smo poznali gostoto in sedem elastičnih modulov, ki so bili dobljeni na podlagi nedestruktivnih metod preizkušanja, ter trdnost, dobljeno na osnovi destruktivne metode. Naključno smo izbrali 250 preizkušancev za učenje umetnih nevronske mreže, preostalih 43 pa za testiranje mreže. Obravnavali smo pet primerov, ki so se med seboj ločili v izbiri vhodnih podatkov. Za gradnjo in učenje nevronske mreže uporabimo fortranski program NTR 2003 in program, ki smo ga zasnovali na osnovi knjižnice nevronske mreže v programskem paketu Matlab. Med sabo primerjamo uspešnost različnih mreže na osnovi več statističnih kazalcev. Ukvarjamo se tudi z analizo vpliva raztrosa podatkov in predstavimo postopek združevanja učnih parov.

BIBLIOGRAPHIC – DOCUMENTALISTIC INFORMATION

UDC: 004:624.011.1:624.07(043.2)
Author: Nataša Šinkovec
Supervisor: Assoc. Prof. Goran Turk, Ph. D.
Co-Supervisor: Assist. Prof. Dejan Zupan, Ph. D.
Title: The use of artificial neural networks for timber strength estimation
Notes: 81 p., 33 tab., 5 gr., 24 pic., 42 eq.
Key Word: artificial neural networks, data dissipation, characteristics of timber, strength of structural timber.

Abstract

In design of various engineering structures structural material together with proper planning and construction realisation is of key importance. Structural material should have satisfactory strength that enables a proper load carrying capacity. To determine the characteristics of wooden elements destructive and nondestructive methods are used. In destructive methods the tested materials are loaded until failure occurs. Artificial neural networks have been used in this diploma thesis to estimate the strength of structural timber. 293 wooden specimens were taken into consideration. For all specimens the density and seven elastic modules were measured by indestructive testing methods and the strength was based on the destructive method. 250 specimens were randomly chosen for the training of artificial networks and 43 were left for testing procedure. Five examples, in which different input data is used, were taken into closer consideration, the fifth example also considers different output data. For generating and learning of neural networks two programs were used: (i) a fortran code NRT2003 and (ii) code based on Neural Network Library in Matlab. The efficiency of various neural networks is compared through several statistical quantities. The effect of data dissipation is also considered and a method for data integration is presented.

ZAHVALA

Iskrena hvala mentorju izr. prof. dr. Goranu Turku in somentorju doc. dr. Dejanu Zupanu za pomoč in zavzetost pri izdelavi diplomske naloge. Zahvaljujem se tudi mami, Božotu, Juretu in noni za vso podporo v času študija ter vsem, ki so mi kakorkoli pomagali in naredili trenutke na faksu nepozabne.

KAZALO VSEBINE

1	UVOD	1
2	OBRAVNAVANI PODATKI	3
2.1	Kratek opis metod preizkušanja	6
2.1.1	Vizualno razvrščanje lesa	6
2.1.2	Metode frekvenčnega odziva	7
2.1.3	Longitudinalno širjenje valov	8
2.1.4	Ultrazvok	10
2.1.5	Določanje statičnega modula elastičnosti	12
2.1.6	Destruktivna metoda preizkušanja	15
3	UMETNE NEVRONSKE MREŽE	16
3.1	Kratek opis razvoja umetnih nevronske mreže	16
3.2	Teoretične osnove umetnih nevronske mreže	17
3.2.1	Biološki nevron	17
3.2.2	Umetna nevronska mreža	18
3.2.2.1	Razdelitev umetnih nevronske mreže	25
3.2.3	Večnivojska usmerjena umetna nevronska mreža	28
3.2.4	Pravilo delta	29
3.2.5	Posplošeno pravilo delta	32
4	UPORABA UMETNIH NEVRONSKIH MREŽ ZA OCENO TRDNOSTI LESA	37
4.1	Opis parametrov, ki vplivajo na učenje	37
4.1.1	Največja dovoljena napaka	37
4.1.2	Velikost učnega koraka	37
4.1.3	Maksimalno število iteracij	37
4.1.4	Geometrija mreže	38
4.1.5	Izhodne funkcije	38
4.2	Priprava podatkov na učenje	38

4.3	Statistični kazalci uspešnosti nevronske mreže	39
4.4	Program NTR2003	40
4.4.1	Primer 1	43
4.4.2	Primer 2	46
4.4.3	Primer 3	51
4.4.4	Primer 4	55
4.4.5	Primer 5	58
4.4.6	Rezultati	61
4.5	Knjižnica umetnih nevronske mreže v programu Matlab	63
4.5.1	Primer 1	66
4.5.2	Primer 2	68
4.5.3	Primer 3	70
4.5.4	Primer 4	72
4.5.5	Primer 5	74
4.5.6	Rezultati	76
5	ZAKLJUČEK	77
6	VIRI	79
7	PRILOGE	81

KAZALO PREGLEDNIC

- Preglednica 1: Koeficienti korelacije uporabljenih lastnosti gradbenega lesa.
- Preglednica 2: Vrednosti uporabljenih parametrov pri učenju.
- Preglednica 3: Vrednosti statističnih kazalcev za mreže brez nenaučenih učnih parov.
- Preglednica 4: Vrednosti uporabljenih parametrov pri učenju.
- Preglednica 5: Primer posrednega združevanja učnih parov.
- Preglednica 6: Primer neposrednega združevanja učnih parov.
- Preglednica 7: Vrednosti statističnih kazalcev.
- Preglednica 8: Vrednosti uporabljenih parametrov pri učenju.
- Preglednica 9: Primer posrednega združevanja učnih parov.
- Preglednica 10: Primer neposrednega združevanja učnih parov.
- Preglednica 11: Vrednosti statističnih kazalcev.
- Preglednica 12: Vrednosti uporabljenih parametrov pri učenju.
- Preglednica 13: Primer posrednega združevanja učnih parov.
- Preglednica 14: Primer neposrednega združevanja učnih parov.
- Preglednica 15: Vrednosti statističnih kazalcev.
- Preglednica 16: Vrednosti uporabljenih parametrov pri učenju.
- Preglednica 17: Primer posrednega združevanja učnih parov.
- Preglednica 18: Primer neposrednega združevanja učnih parov.
- Preglednica 19: Vrednosti statističnih kazalcev.
- Preglednica 20: Uporabljeni parametri pri učenju in statistični kazalci.
- Preglednica 21: Uporabljeni parametri pri učenju in statistični kazalci pri večjem številu iteracij.
- Preglednica 22: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.
- Preglednica 23: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.
- Preglednica 24: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.

- Preglednica 25: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.
- Preglednica 26: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.
- Preglednica 27: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.
- Preglednica 28: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.
- Preglednica 29: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.
- Preglednica 30: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.
- Preglednica 31: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.
- Preglednica 32: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.
- Preglednica 33: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.

KAZALO GRAFIKONOV

- Grafikon 1: Prikaz raztrosa za f v odvisnosti od E_{f-g} in ρ .
- Grafikon 2: Prikaz raztrosa za f v odvisnosti od E_{f-g} in E_{Sylv} .
- Grafikon 3: Prikaz raztrosa za f v odvisnosti od ρ in E_{Tno} .
- Grafikon 4: Učenje s štirimi vhodno-izhodnimi učnimi pari.
- Grafikon 5: Učenje z dvestopetdesetimi vhodno-izhodnimi učnimi pari.

KAZALO SLIK

- Slika 1: Dva načina podpiranja pri metodi frekvenčnega odziva.
- Slika 2: Osnovni sistem merjenja longitudinalnega širjenja valov.
- Slika 3: Inštrument Timber Grader MTG.
- Slika 4: Shema delovanja enosmernega ultrazvočnega sistema.
- Slika 5: Ultrazvočna naprava Sylvatest.
- Slika 6: Shema delovanja upogibne naprave v tekočem traku.
- Slika 7: Shema preizkusa za določitev lokalnega modula elastičnosti po SIST EN 408.
- Slika 8: Shema preizkusa za določitev globalnega modula elastičnosti po SIST EN 408.
- Slika 9: Shema biološkega nevrona.
- Slika 10: Shema delovanja umetne nevronske mreže.
- Slika 11: Pragovna oz. stopničasta funkcija.
- Slika 12: Eksponentna sigmoidna funkcija.
- Slika 13: Tangentna sigmoidna funkcija.
- Slika 14: Linearna funkcija.
- Slika 15: Shema enonivojske usmerjene nevronske mreže.
- Slika 16: Shema večnivojske usmerjene nevronske mreže.
- Slika 17: Shema večnivojske nevronske mreže s povezavami nazaj.
- Slika 18: Ploskev napake pri večnivojskih usmerjenih nevronskih mrežah.
- Slika 19: Vhodni podatek nenaučenega učnega para podoben naučenemu.
- Slika 20: Vhodni podatek nenaučenega učnega para podoben drugemu nenaučenemu.
- Slika 21: Testni par podoben učnemu paru.
- Slika 22: Testni par izven meja, ki jih postavljajo učni pari.
- Slika 23: Prikaz prenaučeniosti nevronske mreže.
- Slika 24: Shema umetne nevronske mreže pri kateri je vsak nivo nevronov povezan z vsemi naslednjimi.

1 UVOD

Les je eden najstarejših materialov, ki se uporabljajo v gradbeništvu. Uporablja se tako za konstrukcijske namene (nosilne sisteme), kot za nekonstrukcijsko uporabo (pohištvo, talne obloge, itd.). Njegova uporaba je v 18. stoletju začela upadati, predvsem zaradi nevarnosti požara in zaradi slabih izvedb konstrukcij, ki so vodile v propadanje lesa. Še bolj pa se je zmanjšala poraba lesa v gradbeništvu z uporabo jekla in armiranega betona.

Pri nas je raba tega obnovljivega in ekološko čistega materiala omejena predvsem na strešne konstrukcije stanovanjskih objektov. Vendar se v današnjem času, predvsem na račun uporabe lepljenih lesenih nosilcev in povečane gradnje montažnih hiš, njegova uporaba nekoliko povečuje.

Kot prednosti konstrukcijskega lesa lahko izpostavimo sledeče značilnosti: majhna lastna teža, ugodno razmerje med lastno težo in trdnostjo, dobra toplotna izolativnost, energetska varčnost in estetska vrednost. Ena od najosnovnejših zahtev za materiale, ki se uporabljajo v gradbeništvu, je dovolj visoka trdnost, s katero je omogočena zadostna nosilnost konstrukcije. Napovedati trdnost je pri lesu zaradi njegovih neponovljivih, neenakomernih in nenapovedljivih lastnosti težko. Na to kažejo trdnosti lesenih preizkušancev, ki so različne že za elemente istega debela, kaj šele za elemente različnih debel iste vrste dreves.

Za reševanje težav pri napovedovanju trdnosti mnoge države po svetu, predvsem pa največje proizvajalke in uporabnice lesa za konstrukcijske namene, kot so Združene Države Amerike, Kanada, Avstralija in države EU, izvajajo vrsto različnih raziskav za čimboljše ugotavljanje trdnosti lesa na osnovi lastnosti, dobljenih iz neporušnih raziskav. Eden od alternativnih modelov napovedovanja glavnih lastnosti (trdnost, elastični modul) lesenih preizkušancev so tudi umetne nevronske mreže.

Umetne nevronske mreže so sposobne reševati probleme, ki jih težko opišemo z matematičnimi izrazi. Hkrati pa z analizo določene količine podatkov poiščejo vzorce, relacije in lastnosti objektov v problemskih situacijah, kjer te zveze niso dobro poznane. Kot že

njihovo ime pove, je osnovna ideja delovanja umetnih nevronske mreže simulacija delovanja nevronov v bioloških možganih. Zato se lahko od njih pričakuje, da za reševanje problemov uporabljajo podobne procese kot človeški možgani: asociiranje, posploševanje in podobno. Njihova največja slabost je nezmožnost obrazložitve sprejetih odločitev (uteži) v procesu obdelovanja podatkov. Tako se lahko mreža natančno prilagodi učnim podatkom, njeni rezultati pa ne bodo verodostojni za nove podatke. Zato je izrednega pomena preizkus mreže na kontrolnih podatkih.

Seveda bo vsaka metoda napovedovanja trdnosti tem uspešnejša, čim natančnejši bodo podatki. Pri lesenih preizkušancih pa se srečamo z velikim raztrosom meritev, kar izredno vpliva na natančnost napovedi z nevronske mreže. V diplomski predstavi predstavimo učenje umetnih nevronske mreže s programom NTR2003 in knjižnico nevronske mreže v programskem paketu Matlab. Analiziramo različne geometrije nevronske mreže, uspešnost učenja mreže in uspešnost posploševanja mreže na kontrolnih podatkih. Zaradi velikega raztrosa meritev in posledično slabše uspešnosti učenja, se ukvarjamo tudi z združevanjem učnih podatkov.

2 OBRAVNAVANI PODATKI

Obravnani podatki so rezultati porušnih in neporušnih preiskav na 293 lesenih preizkušancih kvadraste oblike. Nastali so v okviru aplikativnega projekta Metode razvrščanja gradbenega lesa po trdnosti (šifra projekta: L2-6147-0792; trajanje: 01.07.2004-30.06.2007).

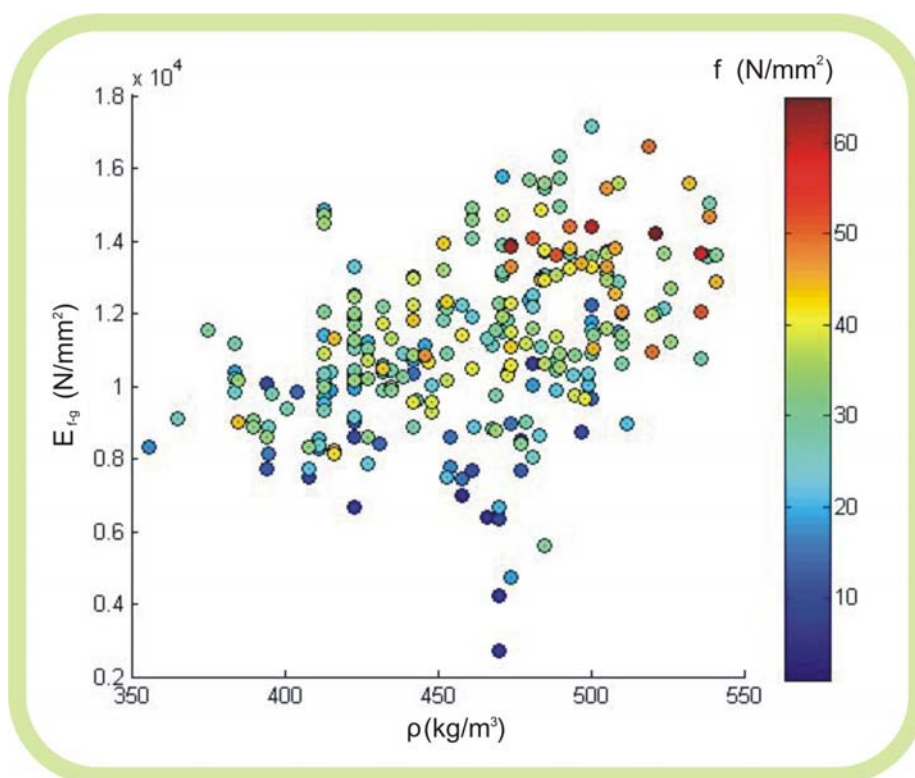
Uporabljeni podatki so:

- gostota pri 12% vlažnosti lesa, ρ [kg/m³];
- lokalni elastični modul (pokončni), E_{e-l} [N/mm²];
- globalni elastični modul (pokončni), E_{e-g} [N/mm²];
- lokalni elastični modul (ležeči), E_{f-l} [N/mm²];
- globalni elastični modul (ležeči), E_{f-g} [N/mm²];
- dinamični elastični modul, E_{Tno} [N/mm²];
- dinamični elastični modul, E_{Zag} [N/mm²];
- dinamični elastični modul, E_{Sylv} [N/mm²];
- upogibna trdnost lesa, f [N/mm²].

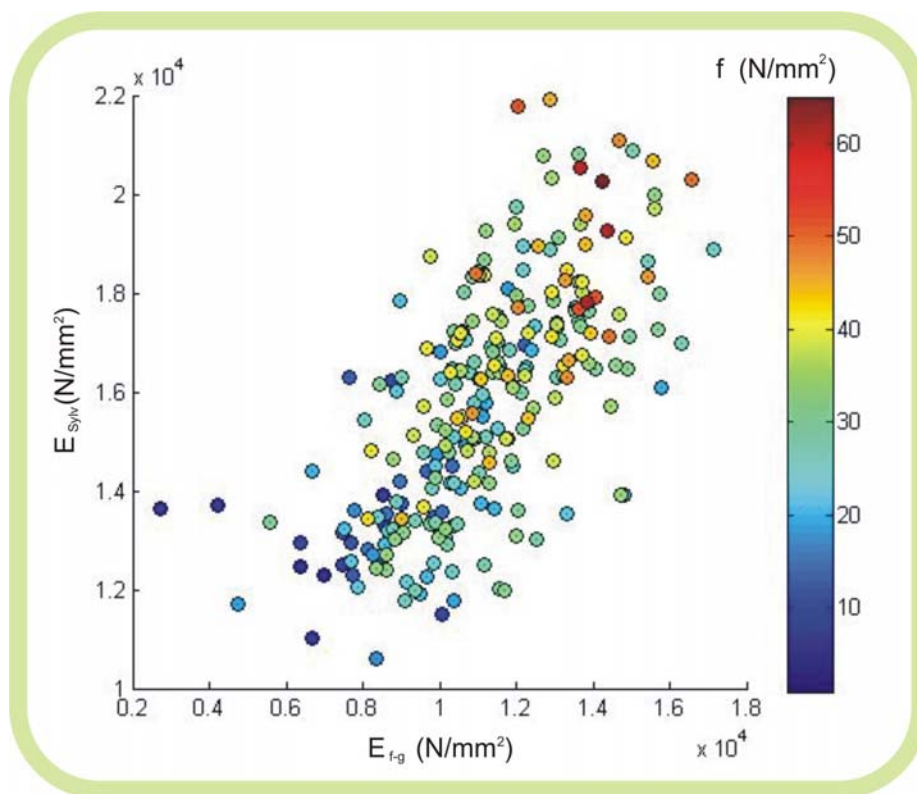
Elastični moduli E_{e-l} , E_{e-g} , E_{f-l} in E_{f-g} so dobljeni s statičnim obremenilnim preizkusom. Indeks e označuje elastične module izmerjene na preizkušancih, ki so bili med preiskavo postavljeni pokončno, indeks f pa se nanaša na med preizkusom ležeče preizkušance. Indeksi l in g označujeta dva načina, s katerima merimo elastični modul: lokalno, na osnovi upogibka na delu preizkušanca med koncentriranima obtežbama s konstantnim momentom in globalno, na osnovi največjega upogibka celotnega preizkušanca. Elastični moduli z indeksi Tno , Zag in $Sylv$ so dinamični moduli elastičnosti. Indeks Tno označuje elastični modul, določen v sodelovanju z inštitutom TNO na Nizozemskem, indeks Zag se nanaša na elastični modul, dobljen na Zavodu za gradbeništvo, indeks $Sylv$ pa pomeni dinamični modul, določen s pomočjo merilne naprave Sylvatest.

Za uporabljene lastnosti gradbenega lesa je značilen velik raztros. Iz grafikonov 1-3 razberemo, da imajo preizkušanci s podobnimi lastnostmi za gostoto in elastične module zelo

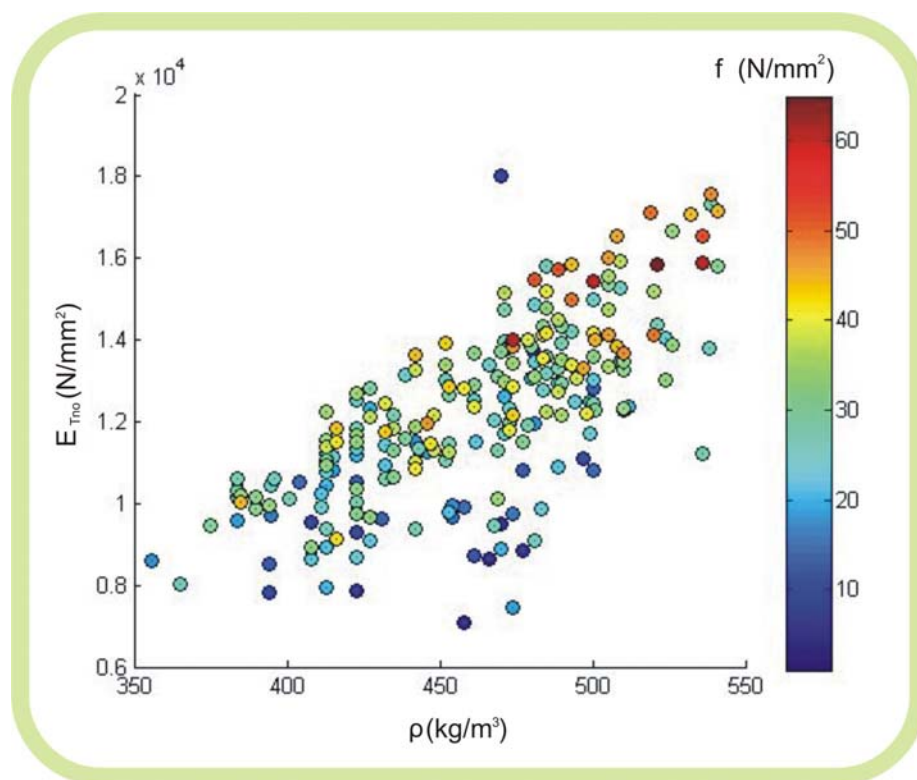
različne trdnosti. Poleg tega iz preglednice 1 vidimo, da imajo posamezne lastnosti preizkušancev slabo korelacijo, še posebej ρ in f .



Grafikon 1: Prikaz raztrosa za f v odvisnosti od E_{f-g} in ρ .



Grafikon 2: Prikaz raztrosa za f v odvisnosti od E_{f-g} in E_{Sylv} .



Grafikon 3: Prikaz raztrosa za f v odvisnosti od ρ in E_{Tno} .

Preglednica 1: Koeficienti korelacije uporabljenih lastnosti gradbenega lesa.

	f [N/mm ²]	ρ [kg/m ³]	E_{e-l} [N/mm ²]	E_{e-g} [N/mm ²]	E_{f-l} [N/mm ²]	E_{f-g} [N/mm ²]	E_{Tno} [N/mm ²]	E_{Zag} [N/mm ²]	E_{Sylv} [N/mm ²]
f [N/mm ²]	1								
ρ [kg/m ³]	0.2951	1							
E_{e-l} [N/mm ²]	0.7086	0.4484	1						
E_{e-g} [N/mm ²]	0.7160	0.4863	0.8728	1					
E_{f-l} [N/mm ²]	0.5202	0.3664	0.5392	0.5682	1				
E_{f-g} [N/mm ²]	0.5004	0.4157	0.6371	0.7242	0.8006	1			
E_{Tno} [N/mm ²]	0.5407	0.7124	0.7110	0.7759	0.6658	0.7471	1		
E_{Zag} [N/mm ²]	0.4741	0.7838	0.6039	0.5680	0.5113	0.4557	0.7968	1	
E_{Sylv} [N/mm ²]	0.5026	0.8270	0.6494	0.7268	0.6072	0.6700	0.9015	0.8847	1

Pri nedestruktivnih metodah gre za ugotavljanje mehanskih ali fizikalnih lastnosti preizkušanca, ne da bi s tem vplivali na končno uporabnost. Delijo se v tri skupine: vizualne, mehanske in kemijske. Podrobnejši pregled in opis metod preizkušanja je v diplomskem delu Mitje Plosa.

2.1 Kratek opis metod preizkušanja

2.1.1 Vizualno razvrščanje lesa

Vizualne metode temeljijo zgolj na človeškem opazovanju in so zato tudi najstarejši način ugotavljanja raznih napak v lesu in razvrščanja lesa. Različni načini vizualnega ocenjevanja po svetu so odraz različnih lokalnih vrst dreves, geografskih značilnosti, tradicionalne gradnje in različnih zahtev.

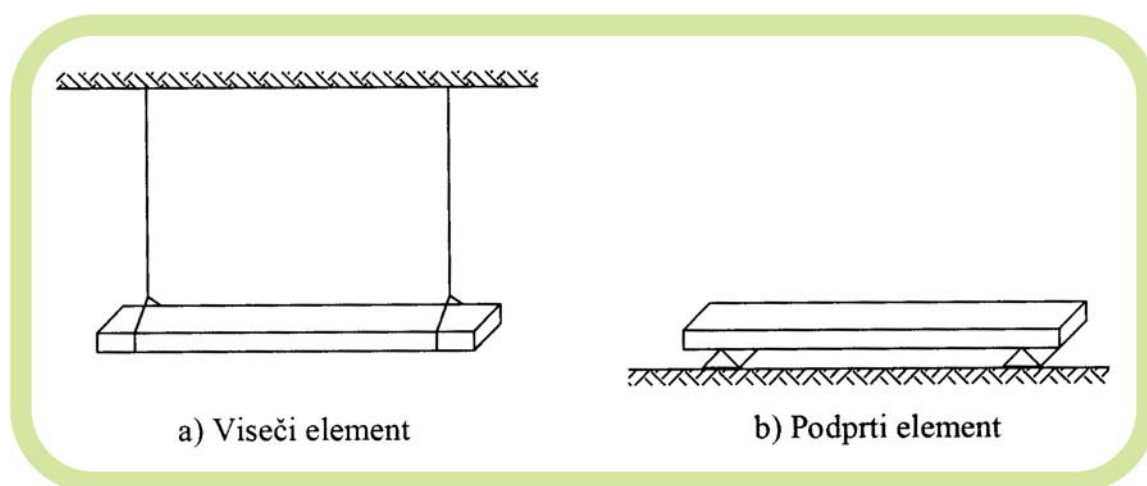
Z vizualnimi metodami se na preprost način ugotavljajo napake (zunanje poškodbe, razkroj, poškodbe vlaken, razpoke, itd.), ki vplivajo na trdnost lesa. Nato se na osnovi standardov les

razvrsti v trdnostne razrede. Vendar pa so tovrstne metode omejene predvsem na zunanje površine in zato ne omogočajo celostnega pregleda preizkušanca.

V zadnjem času se uveljavljajo tudi avtomatizirani načini vizualnega razvrščanja.

2.1.2 Metode frekvenčnega odziva

Osnova metodam frekvenčnega odziva oz. metodam transverzalnih vibracij je teorija, ki pravi, da so dinamične mehanske lastnosti odvisne od disipacije in lastne frekvence. Pri omenjenih preizkusih je možnih več načinov podpiranja. Tu sta navedena dva, prvi je prostoviseč, drugi pa prostoležeč (slika 1). Pri prvem je element obešen na obeh krajnih koncih, pri drugem pa je element na obeh koncih podprt kot prostoležeč nosilec.



Slika 1: Dva načina podpiranja pri metodi frekvenčnega odziva.

Ločimo dva načina vzbujanja vibracij: vsiljeno, kjer je element periodično obremenjen in prosto vibriranje, kjer je element obremenjen hipno in opazujemo amplitudo nihanja, ki s časom upada. Pri vsiljenem vzbujanju vibracij se z vsiljeno periodično silo s funkcijo $P_0 \sin \omega t$ določi resonančna frekvenca. Z njo se lahko izračuna dinamični modul elastičnosti za viseč element E_F po enačbi:

$$E_F = \frac{f_r^2 ml^3}{12.65I}, \quad (1)$$

kjer je:

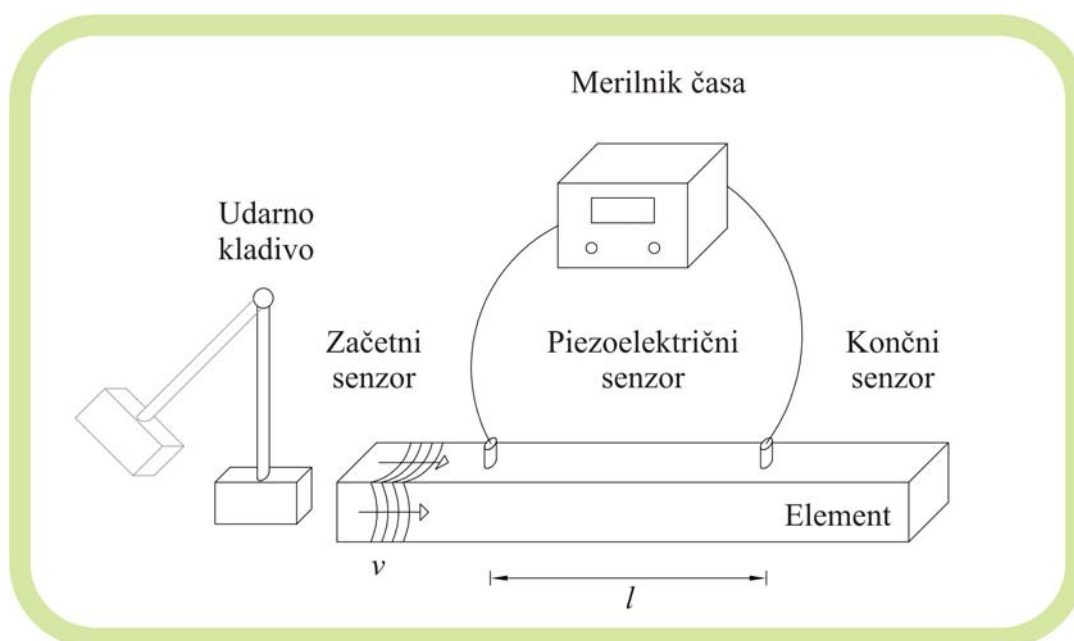
- f_r . . . resonančna frekvenca,
- m . . . masa elementa,
- l . . . dolžina elementa,
- I . . . vztrajnostni moment prereza.

Ravno tako se z resonančno frekvenco določi dinamični modul elastičnosti E_P za podprti element po enačbi:

$$E_P = \frac{f_r^2 ml^3}{2.46I}. \quad (2)$$

2.1.3 Longitudinalno širjenje valov

Metoda longitudinalnega širjenja valov je ena najhitrejših nedestruktivnih metod ocenjevanja trdnosti. Temelji na empiričnih ugotovitvah, da valovi potujejo počasneje pri lesu slabše kakovosti. Torej nižje hitrosti potovanja valov označujejo slabši material. Osnovni sistem merjenja longitudinalnih valov sestavljata dva senzorja, kladivo, ki povzroča valove, in merilnik, ki meri čas potovanja valov (slika 2).



Slika 2: Osnovni sistem merjenja longitudinalnega širjenja valov.

Iz izmerjene hitrosti v potovanja valov in gostote ρ lahko izračunamo dinamični modul elastičnosti E_{DIN} po enačbi:

$$E_{\text{DIN}} = \rho v^2, \quad (3)$$

kjer je:

ρ . . . gostota elementa,

v . . . hitrost potovanja valovanja, ki se izračuna z enačbo:

$$v = \frac{l}{t}, \quad (4)$$

kjer je:

l . . . razdalja med senzorjema,

t . . . čas potovanja vala.

Poleg zgornje metode je poznana še metoda dvojnega preleta. Pri tem načinu je princip merjenja enak kot zgoraj, le da sta vzbujevalnik in senzor na isti strani elementa.

Eden izmed izvajalcev preizkusa z longitudinalnim širjenjem valov je inštitut TNO na Nizozemskem, ki je sodeloval pri razvoju inštrumenta Timber Grader MTG. Omenjeni inštrument je osnovan na zgoraj omenjeni metodi dvojnega preleta, je enostaven za uporabo in priročen, saj je v napravi vgrajeno tako kladivo kot senzor (slika 3).

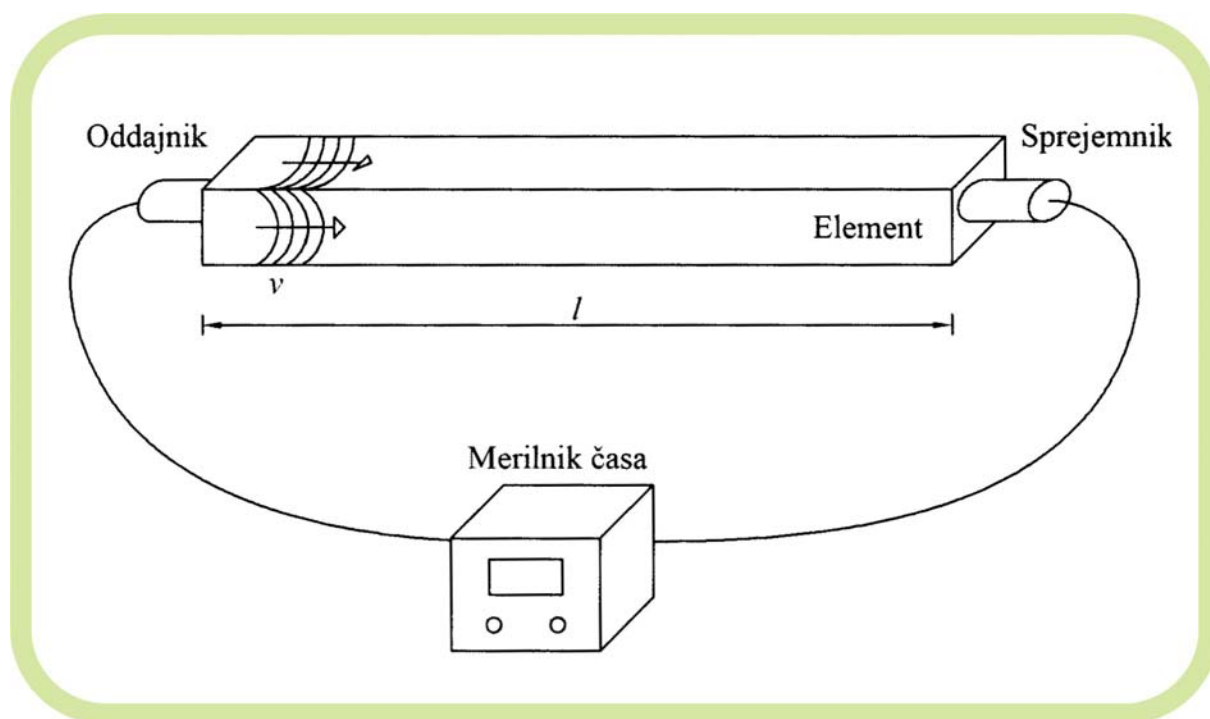


Slika 3: Inštrument Timber Grader MTG.

2.1.4 Ultrazvok

Ultrazvočna metoda je podobna metodam longitudinalnega širjenja valov, le da deluje pri višjih frekvencah. Velikokrat je zaradi natančnosti meritev uporabljena pri drugih materialih, kot so beton, jeklo in keramika, za odkrivanje raznih napak, razpok in drugih anomalij. Pri lesu se ultrazvok največ uporablja pri vezanih ploščah in lepljenem lesu zaradi zmožnosti, da se ultrazvočni signal koncentrira na majhne površine. Osnova te metode je merjenje hitrosti potovanja ultrazvočnega signala skozi material, ta pa je povezana z gostoto in elastičnim modulom, ki ga lahko izračunamo enako kot pri longitudinalnem širjenju valov po enačbah (3) in (4).

Dve najbolj pogosti ultrazvočni metodi sta podobno kot pri longitudinalnem širjenju valov enosmerna in dvosmerna metoda z odbojem oz. metoda dvojnega preleta. Osnovni sistem enosmernega ultrazvoka je sestavljen iz oddajnika in sprejemnika, ki sta nameščena na nasprotnih koncih elementa, in iz merilnika časa (slika 4).



Slika 4: Shema delovanja enosmerne ultrazvočne sistema.

Ko ultrazvočni signal pri potovanju skozi element naleti na anomalijo, se delno odbije nazaj k oddajniku, sprejemnik pa zato prejme oslabljen signal. Iz teh dveh signalov se nato ugotovijo napake v materialu. Za razliko od enosmerne metode ima dvosmerna metoda z odbojem le en senzor, ki deluje istočasno kot sprejemnik in oddajnik impulzov.

Na potovanje ultrazvočnega signala vplivajo poleg gostote tudi druge lastnosti, kot je vlaga in smer vlaken. Signal potuje hitreje po vzdolžnih vlaknih, zato je iz spremembe hitrosti potovanja signala možno ugotoviti tudi smer vlaken v materialu.

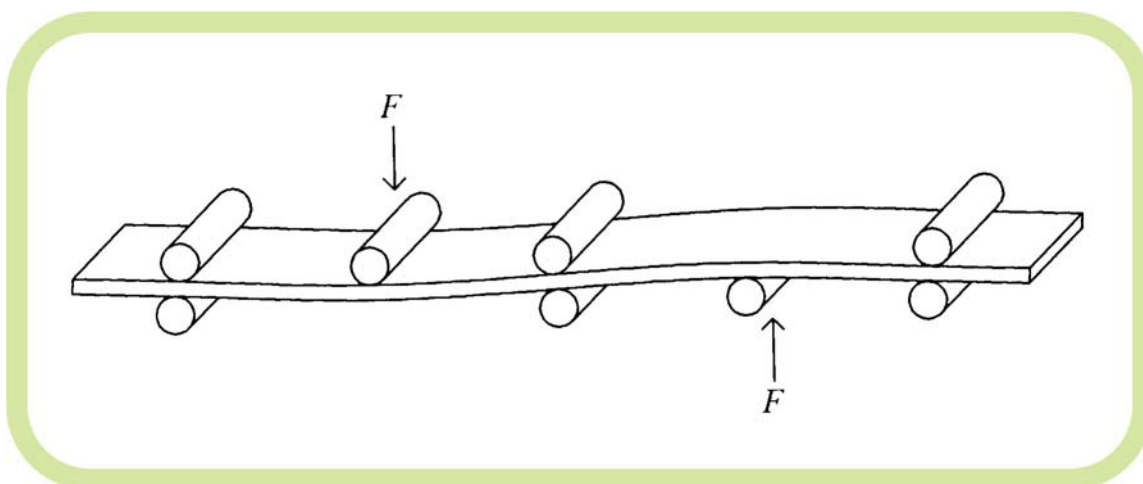
Eden izmed ultrazvočnih inštrumentov je Sylvatest, ki je namenjen samo za les. Naprava vsebuje dva senzorja, ki sta koničaste oblike. Njuna glava se vstavi v predhodno izvrtane luknjice na elementu (slika 5).



Slika 5: Ultrazvočna naprava Sylvatest.

2.1.5 Določanje statičnega modula elastičnosti

Upogibna metoda za ugotavljanje statičnega modula elastičnosti je v svetu precej v uporabi, saj je enostavna, natančna in hitra. Z merjenjem upogibka ob poznavanju dimenzij elementa lahko izrazimo elastični modul. Obstajata dva načina: pri prvem se meri pomik pri določeni sili, pri drugem se meri sila, ki bo povzročila določen pomik. Upogibno metodo je možno tudi enostavno umestiti v sistem tekočega traku, kjer se element spusti skozi sistem valjev, merimo pa pomik med valji (slika 6).



Slika 6: Shema delovanja upogibne naprave v tekočem traku.

Upogibna metoda omogoča določitev lokalnega in globalnega modula elastičnosti. Element je lahko postavljen pokončno ali ležeče. Z uporabo standarda SIST EN 408 je postopek naslednji:

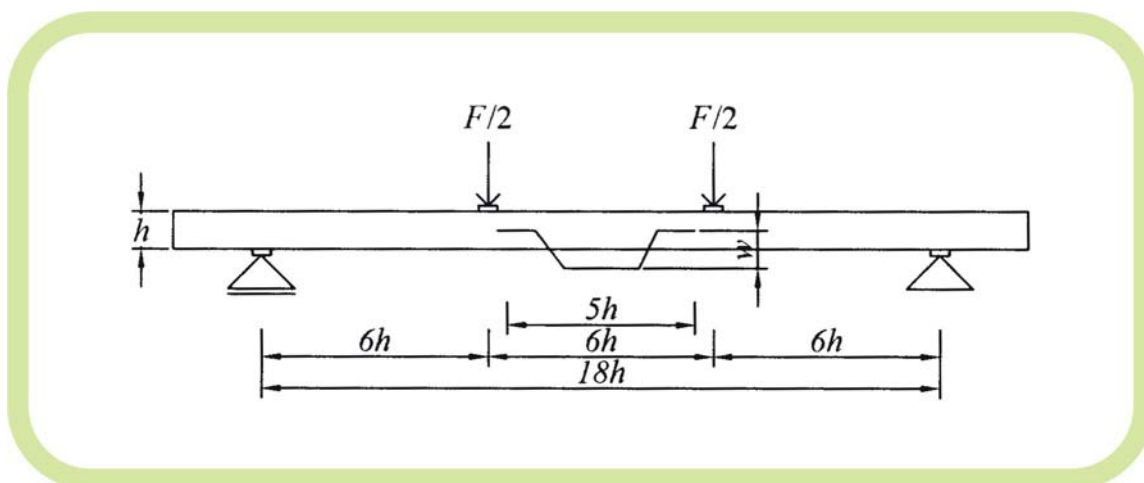
- element je podprt kot prostoležeči nosilec,
- bočna podpora za preprečitev uklona,
- element se na razponu $18h$ obremeni simetrično z dvema točkovnima silama,
- konstantno povečevanje sile,
- maksimalna sila je lahko največ $0.4 F_{\max}$,
- merjenje pomika w na sredini razpona elementa na razdalji $5h$,
- izračun modula elastičnosti.

Lokalni modul elastičnosti za preizkušane pravokotnega prereza se izračuna po naslednji enačbi:

$$E_l = \frac{3al_1^2(F_2 - F_1)}{4bh^3(w_2 - w_1)}, \quad (5)$$

kjer je:

- $F_2 - F_1$. . . sprememba sile,
- $w_2 - w_1$. . . sprememba pomikov,
- b . . . širina elementa,
- h . . . višina elementa,
- a . . . razdalja med silo in podporo,
- l_1 . . . razdalja med silama.



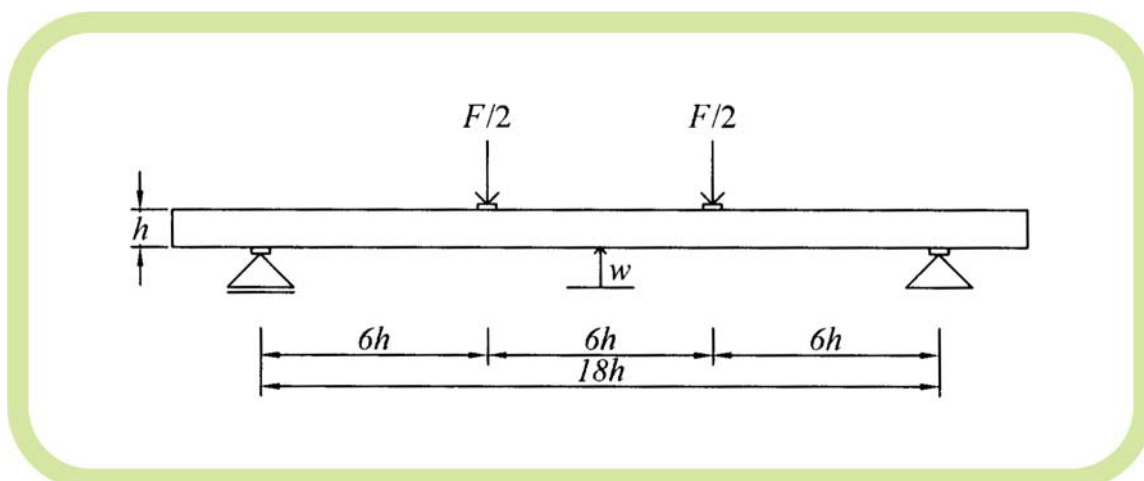
Slika 7: Shema preizkusa za določitev lokalnega modula elastičnosti po SIST EN 408.

Določitev globalnega statičnega modula elastičnosti je podobna, le da so pomiki merjeni na celotnem elementu (slika 8). Enačba za izračun je naslednja:

$$E_g = \frac{l^3(F_2 - F_1)}{bh^3(w_2 - w_1)} \left(\left(\frac{3a}{4l} \right) - \left(\frac{a}{l} \right)^3 \right), \quad (6)$$

kjer je:

- l . . . razdalja med podporama,
- a . . . razdalja med silo in podporo,
- b . . . širina elementa,
- h . . . višina elementa.



Slika 8: Shema preizkusa za določitev globalnega modula elastičnosti po SIST EN 408.

2.1.6 Destruktivna metoda preizkušanja

Za določanje trdnosti lesenih elementov se uporabljajo destruktivne metode, kjer se preizkušane obremeni do porušitve. Postopek preizkusa je enak kot pri določanju statičnega modula elastičnosti, le da se v tem primeru obremenitev povečuje do porušitve. Na osnovi izmerjene sile pri porušitvi in na osnovi dimenzij prereza izračunamo upogibno trdnost lesenega preizkušanca po enačbi:

$$f = \frac{18F_{\max}}{hb}, \quad (7)$$

kjer je:

- f . . . upogibna trdnost lesa,
- F_{\max} . . . obremenitev ob porušitvi,
- h . . . višina elementa,
- b . . . širina elementa.

3 UMETNE NEVRONSKE MREŽE

3.1 Kratek opis razvoja umetnih nevronske mreže

Matematični model živčne celice oz. nevrona sta prva predstavila McCulloch in Pits leta 1943. Ta model, ki sta ga poimenovala pragovni element, je skupaj z vhodnimi in izhodnimi povezavami oz. utežmi osnovni gradnik umetne nevronske mreže.

Leta 1949 je Hebb podal učno pravilo za postopno iskanje vrednosti uteži in pragovnega elementa, ki bodo ustrezale našim zahtevam in tako je model nevrona vzbudil zanimanje številnih raziskovalcev. Deset let pozneje je Rosenblatt modelu nevrona dodal povratno zanko in to poimenoval perceptron. Ta oznaka se nanaša na poenostavljen model biološkega mehanizma, kjer izmenjava informacij poteka preko percepcije oz. zaznave. Perceptron v najosnovnejši obliki sestavljata vhodni in izhodni nivo nevronov, nevroni posameznega nivoja med sabo niso povezani. Tak koncept predstavlja enega izmed prvih adaptivnih sistemov. Tok informacij v perceptronu je usmerjen od vhodnega proti izhodnemu nivoju in nikoli obratno, zato se lahko tako mrežo poimenuje tudi usmerjena umetna nevronska mreža (angl. feed forward neural network).

Okoli leta 1969 sta Minsky in Papert pokazala, da tako preprost perceptron z dvema nivojema nevronov ni sposoben rešiti že preprostih primerov. Eden takih primerov je "ekskluzivni ali" (angl. Exclusive-OR). Funkcija XOR zahteva dva vhodna nevrona in enega izhodnega. Vhodna nevrona morata biti povezana z izhodnim na tak način, da se izhodni nevron aktivira le takrat, ko je aktiven le eden od vhodnih nevronov. Ta kritika je močno vplivala na nadaljnji razvoj umetnih nevronske mreže.

Ponovno je to področje zaživelo šele leta 1986, ko so David Rumelhart, Geoffrey Hinton in Ronald Williams predstavili učni algoritem za določanje vrednosti uteži v večnivojskih usmerjenih umetnih nevronske mrežah. Ta učni algoritem se imenuje posplošeno pravilo delta in temelji na iterativnem spreminjanju vrednosti uteži na podlagi čimmanjše razlike med željenim in z umetno nevronske mrežo izračunanim signalom. Ime za učenje takih

večnivojskih usmerjenih mrež je tudi učenje umetne nevronske mreže z vzratnim širjenjem napake (angl. error back-propagation algorithm).

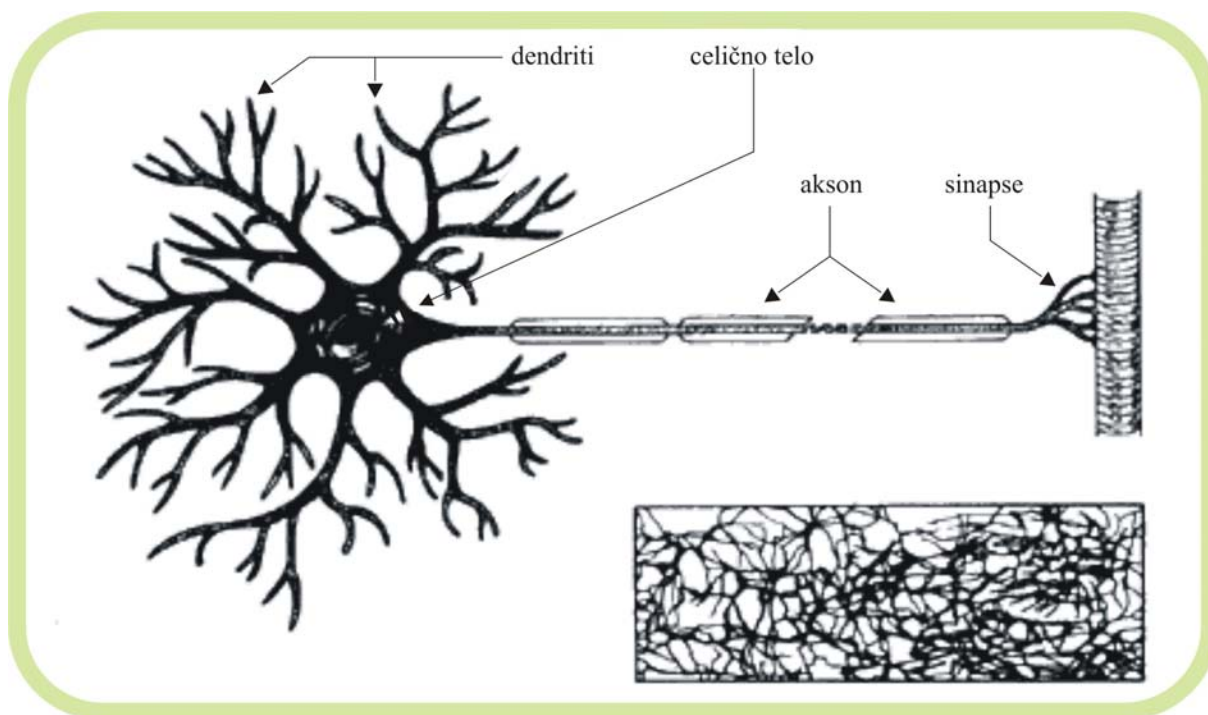
Kasneje se je področje umetnih nevronske mreže še naprej razvijalo v smislu različnih struktur mrež. V zadnjem času se je njihova uporaba razširila na vrsto različnih področij, med drugim tudi na gradbeništvo.

3.2 Teoretične osnove umetnih nevronske mreže

Področje umetnih nevronske mreže je eno od področij umetne inteligence. Že samo ime pove, da osnovna ideja umetnih nevronske mreže izhaja iz delovanja človeških možganov. Nanje umetne mreže spominjajo iz večih vidikov: njihova sestava iz nevronov in povezav med njimi, učenje, s katerim pridobivajo znanje, shranjevanje znanja v povezavah med nevroni oz. utežeh, sposobnost asociiranja, t.j. zmožnost reševanja problemov ob nepopolnih podatkih, zmožnost generiranja, to pomeni da mreža zna izračunati izhodne vrednosti tudi za take podatke, ki pri učenju niso sodelovali, in neobčutljivost na napake vezi in nevronov, torej tudi če je nek nevron ali vez odstranjena, bo mreža še vedno sposobna delovati.

3.2.1 Biološki nevron

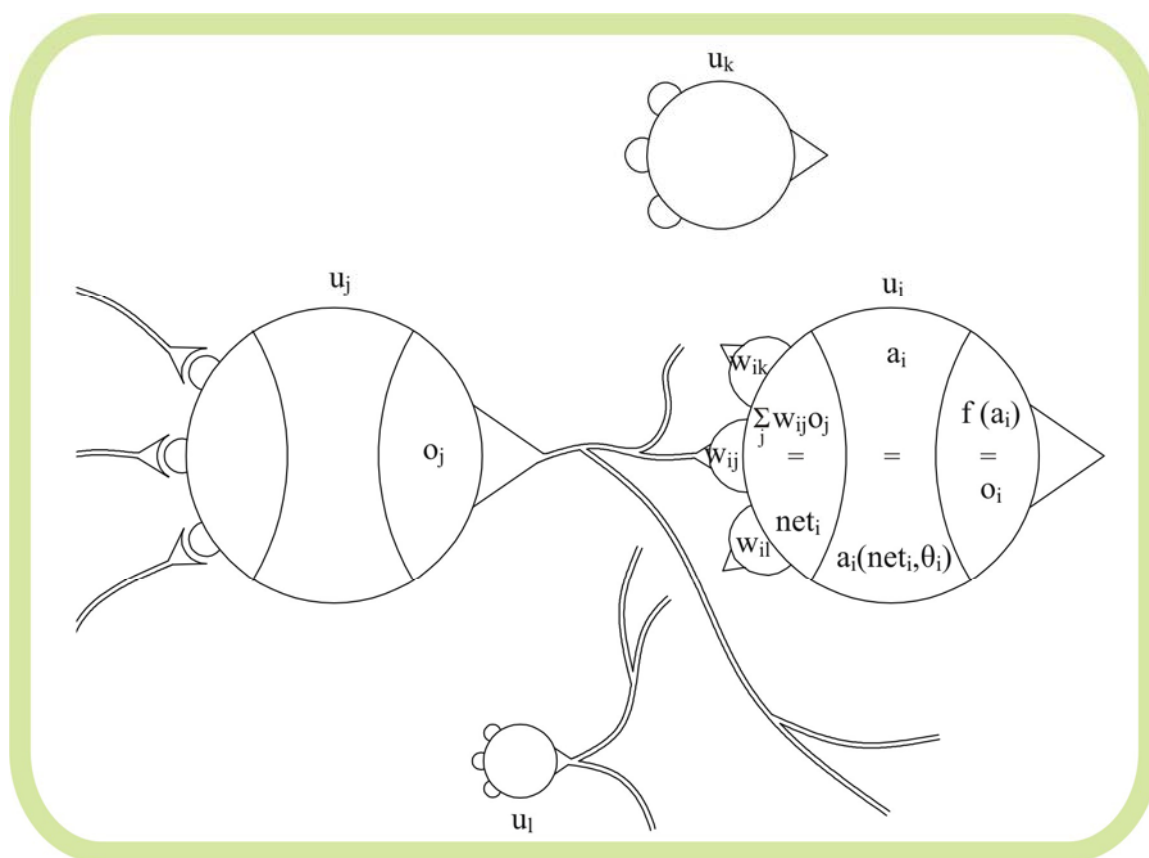
Osnovna celica živčnega sistema je nevron. Ocenjujejo, da ima človek okoli sto milijard nevronov, ki so med seboj povezani s številnimi vezmi in tvorijo biološko mrežo oz. možgane. Živčna celica je sestavljena iz celičnega telesa, iz običajno enega daljšega izrastka-aksona, ki se konča s sinapso in iz večih krajših izrastkov-dendritov. Sinapse imajo določeno stanje oz. prepustnost za signale, ravno tako pa imajo neko stanje nevroni. Nevron od dendrita preko sinapse prejme signal, ki ga obdelata. Če je ta obdelan signal večji od določenega praga tega nevrone, to nevron "vžge" in ga pošlje po aksonu do drugih celic oz. nevronov. Če signal ni večji od praga nevrone, ne pride do "vžiga" in signal ne potuje naprej.



Slika 9: Shema biološkega nevrona.

3.2.2 Umetna nevronska mreža

Umetne mreže so sicer manjše in manj komplicirane od bioloških, a delujejo na podoben način kot biološke mreže. Gradniki umetnih nevronske mreže so umetni nevroni oz. krajše nevroni, ki so med seboj povezani z vezmi imenovanimi uteži. Preko povezav sprejme nevron signal od drugih nevronov, signal ojača ali oslabi in ga pošlje naprej drugim nevronom. Kako močno so nevroni med seboj povezani določajo vrednosti uteži (slika 10).



Slika 10: Shema delovanja umetne nevronske mreže.

kjer je:

- a_i . . . stanje aktivnosti i -tega nevrona v določenem času, t.j. vrednost v i -tem nevronu,
- u_i . . . i -ti nevron,
- u_j . . . j -ti nevron,
- u_k . . . k -ti nevron,
- u_l . . . l -ti nevron,
- w_{ij} . . . utež povezave med i -tim in j -tim nevronom,
- w_{ik} . . . utež povezave med i -tim in k -tim nevronom,
- w_{il} . . . utež povezave med i -tim in l -tim nevronom,
- net_i . . . vrednost vhodnega signala v i -tem nevronu,
- θ_i . . . prag i -tega nevrona,
- f_i . . . funkcija, ki pretvori stanje aktivnosti nevrona v izhodni signal o_i , ali izhodna funkcija,
- o_i . . . izhodni signal i -tega nevrona,

o_j . . . izhodni signal j -tega nevrona.

Nevron. Nevron je osnovna procesna enota za obdelavo informacij. Nevroni, ki nastopajo na vhodu mreže, so vhodni nevroni. Nevroni, ki ležijo na koncu mreže, so izhodni nevroni, vmes pa so skriti nivoji s poljubnim številom skritih nevronov. Naloga i -tega nevrona je, da izhodni signal, ki ga oddaja k -ti nevron, pomnoži z ustrežno utežjo. To stori še za vse ostale nevrone, s katerimi je povezan. Nato vse te produkte med sabo sešteje, kar predstavlja neto vhodni signal nevrona, ki ga primerja s pragom. Zatem preko izhodne funkcije generira izhodni signal.

Povezave med nevroni. Posamezni nevroni v umetni nevronske mreži so med seboj povezani s povezavami, v katerih je shranjeno naučeno znanje. Povezava med nevronoma u_i in u_j je opisana z utežjo w_{ij} . Moč in smer povezave je določena z absolutno vrednostjo uteži. Negativen predznak pomeni, da nevron u_j zavira delovanje nevrona u_i . Pozitivno število w_{ij} pomeni, da nevron u_j spodbuja delovanje nevrona u_i . Če pa je vrednost uteži enaka 0, nevrona u_i in u_j med seboj nista povezana.

Stanje aktivnosti. To je stanje oz. vrednost posameznega nevrona v nevronske mreži v določenem trenutku. Stanje nevrona lahko opišejo diskretne vrednosti 1 ali 0, lahko pa so te vrednosti zvezne npr. na intervalu med 0 in 1. Stanje aktivnosti je odvisno od neto vhodnega signala in praga θ .

Vhodni signal. Vhodni signal net_i nevrona u_i je vsota produktov posameznih uteži w_{ij} , ki povezujejo nevron u_i z ostalimi nevroni, z izhodnimi vrednostmi o_j teh nevronov. Ker so nevroni združeni v mreže, imamo več nevronov, več neto vhodnih signalov in več uteži. Zato lahko govorimo o vektorjih vhodnih signalov in vektorjih uteži.

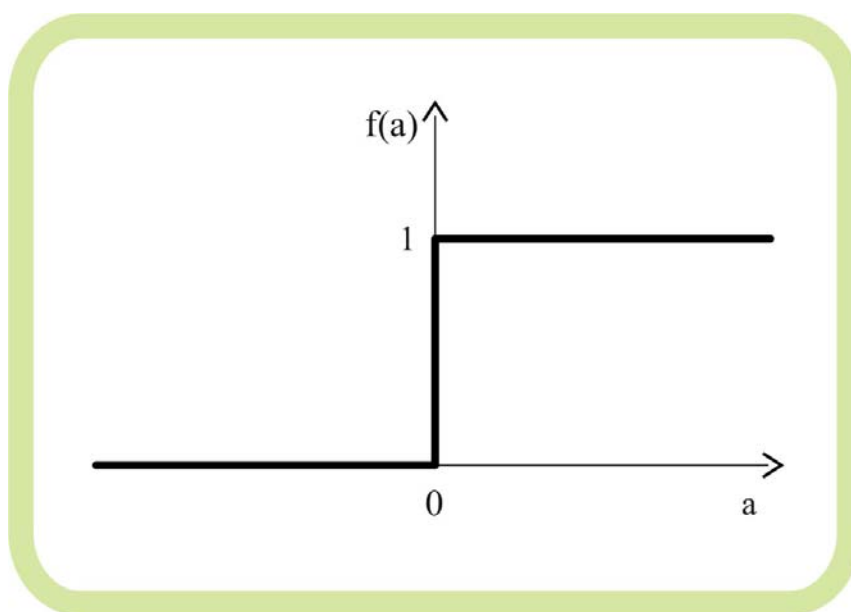
Izhodna funkcija. Njena naloga je, da pretvori stanje aktivnosti nevrona u_i v izhodni signal o_i ; velja $o_i = f(a_i) = f(net_i, \theta_i)$. Z njo skrčimo vrednost izhodnega signala na območje nekega intervala. Vsak nevron v mreži lahko vsebuje svojo izhodno funkcijo. Najpogosteje se uporabljata naslednji izhodni funkciji:

- pragovna oz. stopničasta funkcija:

$$f(a) = \begin{cases} 1 & \dots a \geq 0 \\ 0 & \dots \text{sicer} \end{cases}, \quad (8)$$

kjer je $a = net_i - \theta_i$.

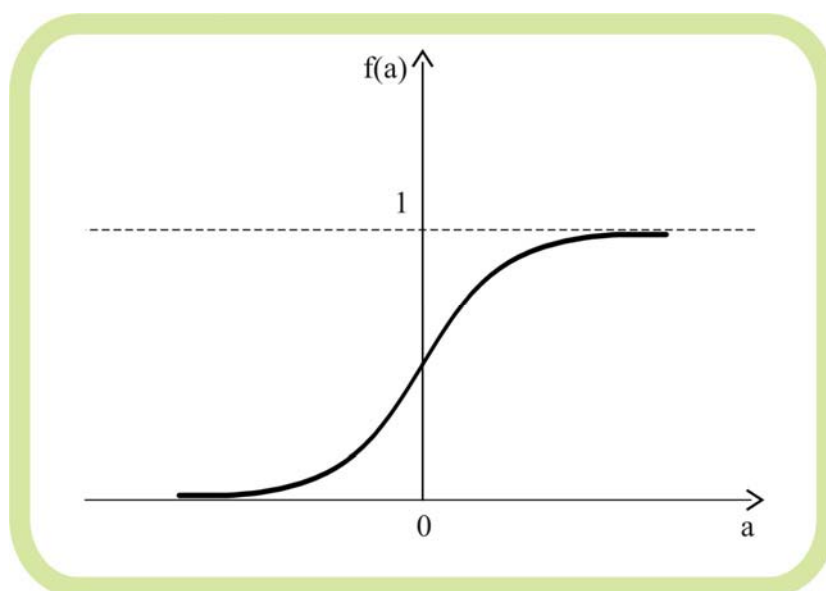
Pragovna funkcija vrne diskretne izhodne vrednosti. In sicer vrednost 1, če je stanje aktivnosti a pri nekem nevronu nenegativno in vrednost 0, če je negativno.



Slika 11: Pragovna oz. stopničasta funkcija.

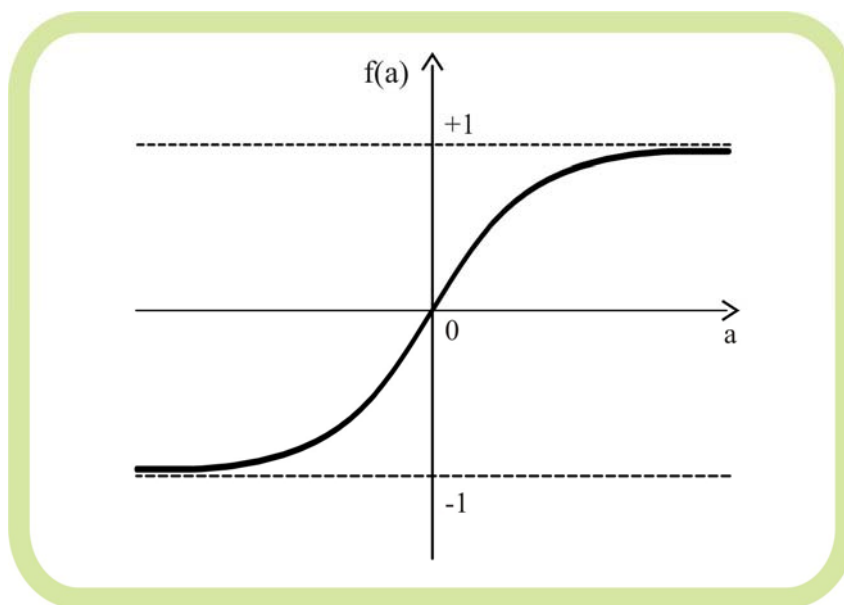
- sigmoidna funkcija (v Matlabu je to funkcija *logsig*):

$$f(a) = \frac{1}{1 + e^{-a}} \quad (9)$$



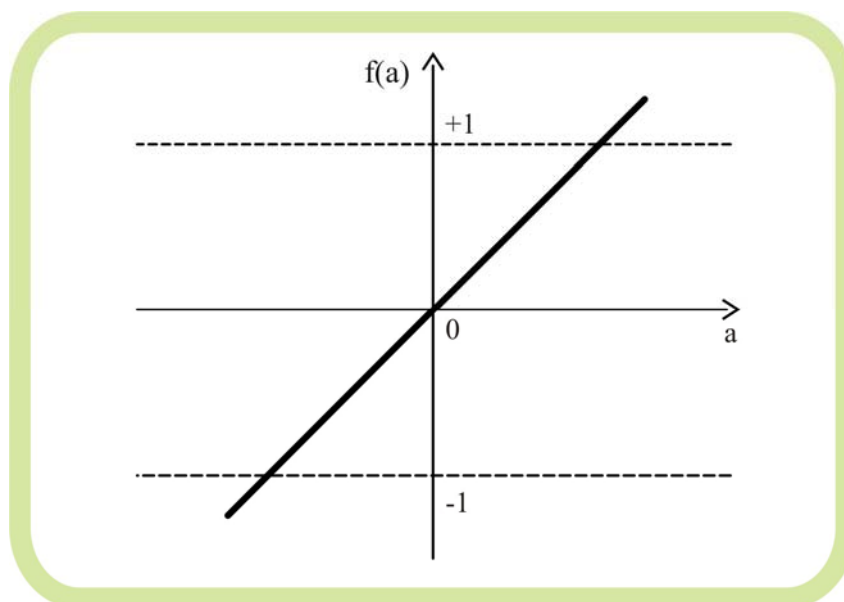
Slika 12: Eksponentna sigmoidna funkcija.

Sigmoidna izhodna funkcija je zvezna, zvezno odvedljiva in strogo naraščajoča funkcija. Pri velikih pozitivnih vrednostih na abscisi se funkcija asimptotično približuje številu 1, pri velikih negativnih vrednosti pa ničli. Torej ima tudi sigmoidna funkcija zalogo vrednosti na intervalu $[0,1]$. Vrsta sigmoidne funkcije je tudi tangentna sigmoidna funkcija (slika 13), ki ima za razliko od eksponentne sigmoidne funkcije zalogo vrednosti na intervalu $[1,1]$. V Matlabu je to funkcija *tansig*.



Slika 13: Tangentna sigmoidna funkcija.

Poleg omenjenih izhodnih funkcij se uporabljajo tudi linearne funkcije (slika 14). V Matlabu je to funkcija *purelin*.



Slika 14: Linearna funkcija.

Učenje. Posebnost pri umetnih nevronske mrežah je, da jih ne programiramo, ampak jih učimo. Proces učenja pomeni spreminjanje povezav med nevroni. Spreminjanje lahko poteka na tri načine:

- ukinitve obstoječih povezav,
- vzpostavitev novih povezav,
- sprememba vrednosti uteži.

Nevronska mreža ima "spomin", kjer hrani vse informacije pridobljene med učenjem. Učenje zahteva učne in testne podatke, kjer so vhodni in pripadajoči izhodni podatki poimenovani vhodno-izhodni pari. Učni so tisti, na katerih se mreža uči, spreminja povezave oz. vrednosti uteži med nevroni. Nevronska mreža na osnovi vhodnih podatkov postavi uteži na povezavah, ki se primerjajo s pragom nevrona. Postavitev uteži v naslednjem nivoju je odvisna od vrednosti uteži na predhodnem nivoju. Tako dobimo na izhodnih nevronih končne vrednosti, ki jih primerjamo z dejansko vrednostjo izhodnih podatkov. Če bo z mrežo izračunana vrednost znotraj določenih meja, bomo z vrednostmi uteži in pragov nevronov zadovoljni, v nasprotnem primeru postopek ponovimo, začne se nova iteracija. Testni vhodno-izhodni pari se uporabijo za ugotavljanje, kako dobro zna nevronska mreža posplošiti naučene zveze.

Pri učenju lahko pride tudi do prenaučenosti nevronske mreže, kar pomeni, da je mreža sicer sposobna zelo natančno reproducirati vhodno-izhodne učne pare, ni pa sposobna dovolj dobro napovedati vrednosti za nove podatke. To pomeni, da za vhodno-izhodne pare, ki pri učenju niso bili vključeni, mreža ne bo sposobna generalizacije, torej ne bo zmožna izračunati izhodne vrednosti, ki bi bile podobne dejanskim.

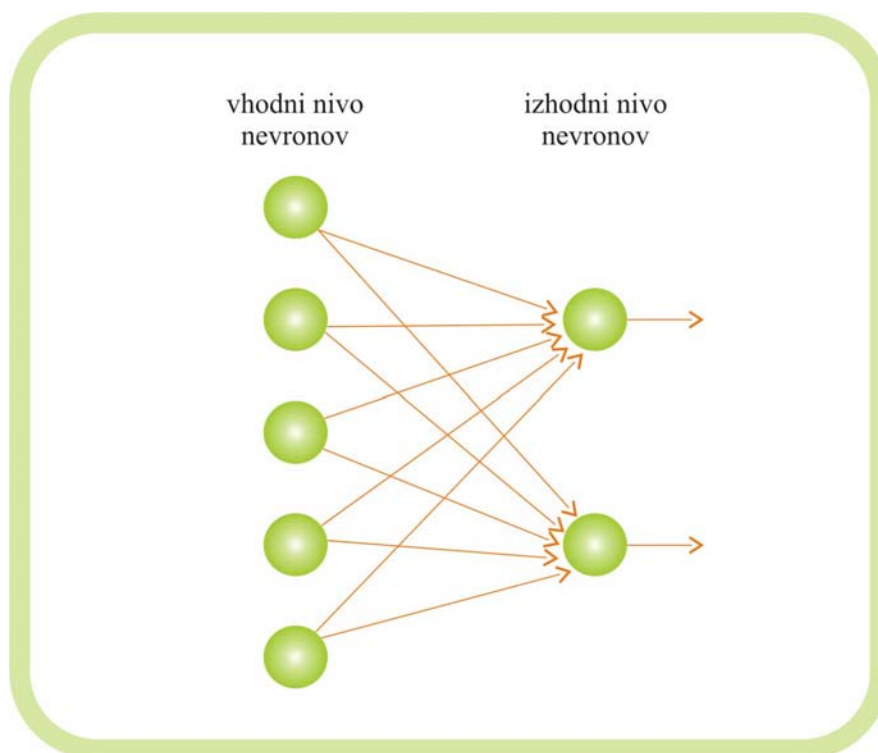
Način učenja je odvisen od tipa mreže in od učnega algoritma, ki ga mreža uporablja za učenje. Največkrat je kot učni algoritem uporabljeno pravilo delta, imenovano tudi Widrow-Hoffovo učno pravilo, ki je izpeljanka Hebbovega pravila. Pravilo delta temelji na razliki, ki jo običajno označimo z oznako δ , med dejanskimi in z nevronske mreže izračunanimi vrednostmi. Ta učni algoritem oz. njegovo nadaljno posplošitev uporabljajo tudi večnivojske usmerjene nevronske mreže, zato je podrobneje prikazan kasneje.

3.2.2.1 Razdelitev umetnih nevronske mreže

Umetne nevronske mreže lahko razdelimo glede na več kriterijev:

Geometrija. Geometrija mreže se nanaša na število nevronov na vhodu v mrežo, na izhodu iz mreže, na število skritih nivojev in na število nevronov na posameznem skritem nivoju:

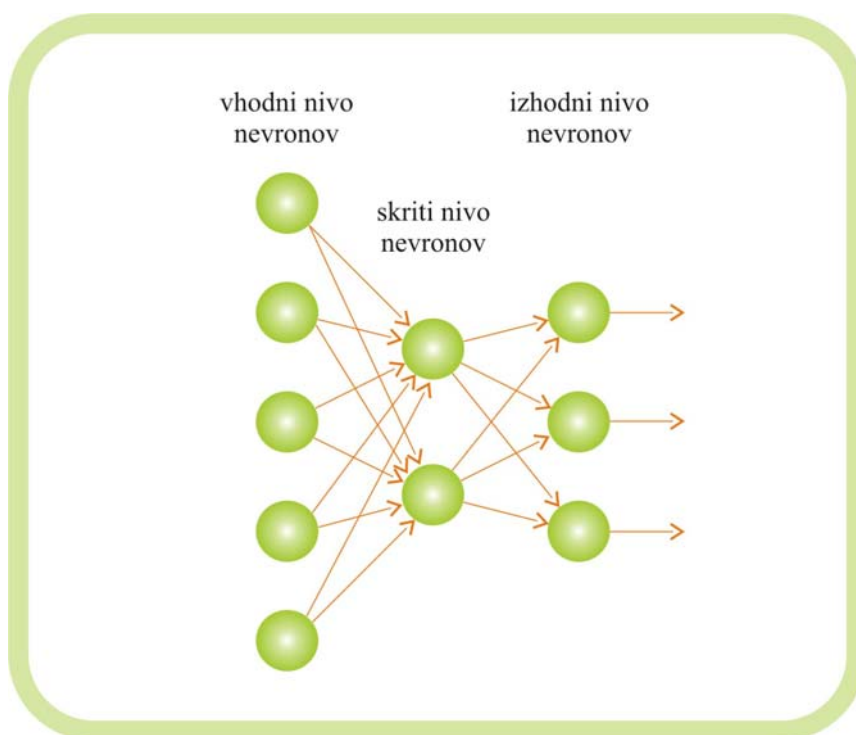
- Enonivojska usmerjena nevronska mreža (angl. single-layer feedforward network): Zanja je značilno, da je sestavljena iz enega vhodnega in enega izhodnega nivoja nevronov. Vhodnemu nivoju so posredovani vhodni podatki, ki jih ta pošlje preko vezi izhodnemu nivoju nevronov, v katerem poteka procesiranje in izračuna izhodne vrednosti. Poimenovanje mreže kot enonivojske se nanaša na izhodni sloj nevronov, ker le v njem poteka procesiranje. Oznaka usmerjena pa se nanaša na potek informacij le v eni smeri in sicer od vhodnih nevronov do izhodnih in ne obratno.



Slika 15: Shema enonivojske usmerjene nevronske mreže.

- Večnivojska usmerjena nevronska mreža (angl. multi-layer feedforward network):

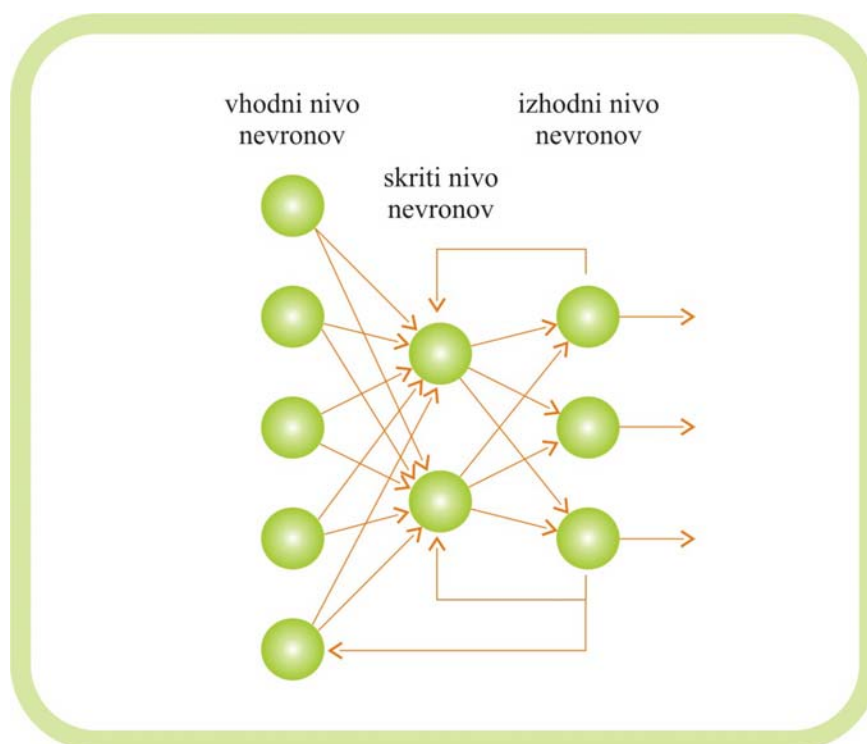
Ta tip mreže ima poleg vhodnega in izhodnega nivoja še enega ali več skritih nivojev, na katerih se nahajajo skriti nevroni. Nevroni na prvem skritem nivoju prejmejo informacije od vhodnih nevronov in jih oddajo naprej izhodnemu nivoju ali naslednjemu skritemu nivoju, če je mreža več nivojska. Torej predstavljajo izhodni signali prvega skritega nivoja vhodne signale drugega skritega nivoja itd.



Slika 16: Shema večnivojske usmerjene nevronske mreže.

- Nevronska mreža s povezavami nazaj (angl. recurrent network):

Zanjo je značilno, da ima za razliko od zgornjih dveh mrež vsaj eno vzvratno povezavo med nivoji nevronov (slika 17).



Slika 17: Shema večnivojske nevronske mreže s povezavami nazaj.

- Kohonenova umetna nevronska mreža:
Kohonenova nevronska mreža je dvonivojska. Sestavljena je iz vhodnega nivoja in Kohonenovega nivoja s povezavami samo naprej. Zanj je značilno samoorganizirano, kompetitivno in nenadzorovano učenje. To pomeni, da se mreža sama organizira tako, da je sposobna dati koristne informacije. Primerna je za zvezne vhodne podatke in reševanje problemov aproksimacije neznane funkcije.
- Hopfieldova umetna nevronska mreža:
Hopfieldova nevronska mreža shrani učne vzorce kot množico stabilnih stanj. Ko damo na vhod testni vzorec, nevronska mreža konvergira k tistemu stabilnemu (naučenemu) stanju, ki leži najbližje vhodu. Vsak nevron lahko zavzame samo dve bipolarni stanji 1 ali -1, pri čemer stanje 1 pomeni prisotnost, stanje -1 pa odsotnost neke lastnosti.
- Hammingova umetna nevronska mreža:
Hammingova umetna nevronska mreža se uporablja pri binarnih vhodnih podatkih in pri problemih razvrščanja v razrede.

Način učenja. Ločimo dva načina učenja:

- **Nadzorovano učenje:**

Pri tovrstnem učenju so mreži podani vhodni in izhodni podatki. Skušajo se učiti tako, da je izračunana vrednost izhodnega podatka čim bolj podobna dejanski vrednosti.

- **Nenadzorovano učenje:**

Podani so le vhodni podatki, izhodnih ni. Ta način učenja mreže se lahko primerja s človekovim načinom učenja hoje, dihanja, itd.

Vrsta vhodnih podatkov. Ločimo diskretne-binarne vhodne podatke, ki zavzemajo vrednost 1 ali 0 in zvezne vhodne podatke. Primer diskretnih vhodnih podatkov je npr. digitalni zvok ali slika. Običajnejši so zvezni vhodni podatki, ki lahko zavzamejo poljubno vrednost na nekem območju (temperatura, vlaga, pomiki, hitrosti, itd.). Vsi obravnavani podatki o lastnostih lesa so zvezni podatki.

3.2.3 Večnivojska usmerjena umetna nevronska mreža

Večnivojske usmerjene nevronske mreže so bile razvite v osemdesetih letih in so danes najpogosteje uporabljene nevronske mreže, saj omogočajo reševanje problemov iz najrazličnejših področij.

Tovrstna nevronska mreža je sestavljena iz vhodnih, izhodnih in iz dveh ali več skritih nivojev, ki vsebujejo poljubno število skritih nevronov. Geometrija take mreže je prikazana v poglavju 3.2.2.1. Število vhodnih nevronov je odvisno od števila vhodnih spremenljivk oz. podatkov. Število izhodnih nevronov je odvisno od števila izhodnih spremenljivk oz. od števila spremenljivk, ki jih hočemo izračunati z mrežo. Zanj je značilno nadzorovano učenje po algoritmu z vzratnim širjenjem napake (angl. error-backpropagation algorithm), ki temelji na posplošenem pravilu delta (poglavje 3.2.5). Omenjeni algoritem je torej postopek, po katerem se mreža uči. Kakšen je ta postopek? Signal potuje od vhodnega nevrone preko vseh skritih nivojev do izhodnega nivoja, ki poda izračunano izhodno spremenljivko. Nato na podlagi razlike med dejansko in z mrežo izračunano vrednostjo spreminja uteži v vzratni smeri, torej od izhodnega nivoja proti vhodnemu.

Poleg značilne geometrije imajo večnivojske nevronske mreže še naslednje tri pomembne lastnosti:

- **Zveznost in odvedljivost.** Za izhodno funkcijo je pomembno, da je zaradi učnega algoritma zvezno odvedljiva in strogo naraščajoča. Običajno se uporablja sigmoidna funkcija, ki je nelinearna.
- **Hierarhija nivojev.** Pomembna je zaradi potovanja signala. Le ta mora vedno potovati v smeri od najnižje ležečih nivojev (vhodni nivo) do najvišje ležečega nivoja (izhodni nivo).
- **Povezanost nivojev.** Vsak nevron v nivoju je povezan z vsakim nevronom naslednjega nivoja. Nevroni istega nivoja med seboj niso povezani.

3.2.4 Pravilo delta

Umetna nevronska mreža se uči na podlagi učnega algoritma, ki določa, kako se spreminjajo uteži v nevronih. Pri učenju posredujemo na vhodnem nivoju vhodne podatke, mreža izračuna rezultate in jih primerja z dejanskimi vrednostmi na izhodnem nivoju. Nato spreminja posamezne uteži na osnovi učnega algoritma. Večnivojska umetna mreža se uči na podlagi posplošenega pravila delta. Najprej si oglejmo učni algoritem, ki ga uporabljajo umetne nevronske mreže s samo vhodnim in izhodnim nivojem nevronov, torej mreže brez skritih nivojev. Algoritem imenujemo pravilo delta in temelji na formuli:

$$\Delta_p w_{ji} = \eta (t_{pj} - o_{pj}) i_{pi} = \eta \delta_{pj} i_{pi} \quad , \quad (10)$$

kjer je:

$\Delta_p w_{ji}$. . . sprememba uteži, ki se zgodi ob upoštevanju p -tega vhodno- izhodnega para,

η . . . velikost učnega koraka,

t_{pj} . . . željena vrednost izhodnega nevrone oz. spremenljivke j za p -ti vhodno-izhodni par,

o_{pj} . . . z umetno nevronske mreže izračunana vrednost izhodne spremenljivke j za p -ti vhodno-izhodni par,

i_{pi} . . . vrednost vhodnega nevrona oz. vhodne spremenljivke i za p -ti vhodno-izhodni par,

δ_{pj} . . . razlika med željeno in z umetno nevronske mreže izračunano vrednostjo izhodne spremenljivke j za p -ti vhodno-izhodni par.

Definirajmo še mero za napako E_p , kot vsoto kvadratov odstopanj med željenimi in z umetno nevronske mreže izračunanimi vrednostmi za p -ti vhodno-izhodni par na izhodnem nivoju umetne nevronske mreže:

$$E_p = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2. \quad (11)$$

Napaka celotnega niza vhodno-izhodnih parov na izhodnem nivoju je potem:

$$E = \sum E_p. \quad (12)$$

Iščemo minimum napake v odvisnosti od uteži, zato E_p odvajamo po w_{ji} . Kadar so izhodne funkcije linearne, velja:

$$-\frac{\partial E_p}{\partial w_{ji}} = k \delta_{pj} i_{pi}, \quad (13)$$

kjer je k naklon linearne izhodne funkcije.

Odvod je torej sorazmeren $\Delta_p w_{ji}$, glej enačbo (10).

Dokažimo zgornjo trditev. Z verižnim pravilom lahko odvod v (13) razbijemo na produkt odvoda napake glede na izhodno vrednost j -tega nevrona in odvoda izhodne vrednosti j -tega nevrona glede na ustrezno utež:

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ji}}, \quad (14)$$

kjer prvi odvod desne strani zgornje enačbe pove, kako se spreminja napaka s spreminjanjem izhodne vrednosti j -tega nevrona, drugi odvod pa pove, za koliko se spremeni izhodna vrednost j -tega nevrona pri spremembi uteži w_{ji} . Iz (11) sledi:

$$\frac{\partial E_p}{\partial o_{pj}} = -(t_{pj} - o_{pj}) = -\delta_{pj}, \quad (15)$$

kar pomeni, da je prispevek j -tega nevrona k napaki E_p sorazmeren razliki δ_{pj} . Kadar so izhodne funkcije linearne:

$$o_{pj} = k \sum_i w_{ji} i_{pi} + n, \quad (16)$$

in velja:

$$\frac{\partial o_{pj}}{\partial w_{ji}} = i_{pi} \cdot k. \quad (17)$$

Če enačbe (15 – 17) upoštevamo v enačbi (14), velja:

$$-\frac{\partial E_p}{\partial w_{ji}} = \delta_{pj} i_{pi} \cdot k.$$

Odvod napake celotnega niza vhodno-izhodnih parov je potem enak:

$$\frac{\partial E}{\partial w_{ji}} = \sum_p \frac{\partial E_p}{\partial w_{ji}}. \quad (18)$$

Dokazali smo, da je sprememba uteži Δw_{ji} sorazmerna odvodu napake, zato lahko rečemo, da pravilo delta temelji na gradientu napake. Izračunane popravke prištejemo utežem in postopek ponavljamo do željene natančnosti.

3.2.5 Posplošeno pravilo delta

Za večnivojske mreže je potrebno osnovno pravilo delta prilagoditi tako, da bo veljalo tudi za uteži med nevroni na skritih nivojih. Zato so Rumelhart in sodelavci podali **posplošeno pravilo delta**.

Pri posplošenem pravilu delta je smiselno uporabiti nelinearne izhodne funkcije. Poleg tega pa to pravilo deluje samo pri odvedljivih in nepadajočih izhodnih funkcijah. Vsem tem zahtevam zadoščajo sigmoidne funkcije.

Najprej zapišimo izraz za celotni neto vhodni signal nevrona u_j :

$$net_{pj} = \sum_i w_{ji} o_{pi} \quad , \quad (19)$$

Izhodni signal pa je:

$$o_{pj} = f_j (net_{pj}) \quad , \quad (20)$$

kjer je f_j odvedljiva, nepadajoča funkcija na j -tem skritem nivoju.

Pri posplošenem pravilu delta vrednosti uteži spreminjamo v sorazmerju z odvodi napake E_p po uteži:

$$\Delta_p w_{ji} \propto - \frac{\partial E_p}{\partial w_{ji}} \quad , \quad (21)$$

kjer je E_p napaka, ki smo jo vpeljali že pri pravilu delta (glej (11)). Enako lahko tudi tukaj odvod v enačbi (21) razdelimo na dva dela. Prvi del predstavlja spremembo napake v odvisnosti od vhodnega signala, drugi del pa ponazarja spremembo vhodnega signala nevrona v odvisnosti od spremembe ustrezne uteži:

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial net_{pj}} \frac{\partial net_{pj}}{\partial w_{ji}}. \quad (22)$$

Drugi člen lahko zapišemo kot:

$$\frac{\partial net_{pj}}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} \sum_k w_{jk} o_{pk} = o_{pi}. \quad (23)$$

Če z δ_{pj} označimo:

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}}, \quad (24)$$

dobimo:

$$-\frac{\partial E_p}{\partial w_{ji}} = \delta_{pj} o_{pi}. \quad (25)$$

Da bi napako E_p manjšali v smeri gradienta, moramo uteži spreminjati podobno kot pri pravilu delta:

$$\Delta_p w_{ji} = \eta \delta_{pj} o_{pi}, \quad (26)$$

vendar je δ_{pj} povsem drugače določen. Ugotovili so, da lahko δ_{pj} izračunamo s pošiljanjem signala napake v mreži v vzvratni smeri. Torej je za določitev signala napake potrebna

vzvratna pot, ki se začne na izhodnem nivoju nevronov in konča na vhodnem nivoju nevronov umetne nevronske mreže. Za izračun δ_{pj} po enačbi (24) uporabimo verižno pravilo za odvajanje. Prvi del predstavlja spremembo napake v odvisnosti od izhodnega signala, drugi del ponazarja spremembo izhodnega signala nevrna v odvisnosti od sprememb v vhodnem signalu:

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}} = -\frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial net_{pj}}. \quad (27)$$

Drugi člen dobimo neposredno iz (19):

$$\frac{\partial o_{pj}}{\partial net_{pj}} = f'_j(net_{pj}) \quad (28)$$

in predstavlja odvod izhodne funkcije nevrna.

Pri prvem členu ločimo dva primera. V prvem primeru predpostavimo, da je nevron u_j izhodni nevron. Torej iz definicije E_p sledi:

$$\frac{\partial E_p}{\partial o_{pj}} = -(t_{pj} - o_{pj}). \quad (29)$$

Za izhodne nevrone tako velja:

$$\delta_{pj} = (t_{pj} - o_{pj}) f'_j(net_{pj}). \quad (30)$$

Račun spremembe uteži je za izhodne nevrone drugačen kot za skrite. Pri izhodnih nevronih je izračun napake preprost, ker vsakemu pripada željena izhodna vrednost. Če u_j ni izhodni nevron, uporabimo verižno pravilo za odvajanje:

$$\sum_k \frac{\partial E_p}{\partial net_{pk}} \frac{\partial net_{pk}}{\partial o_{pj}} = \sum_k \frac{\partial E_p}{\partial net_{pk}} \frac{\partial}{\partial o_{pj}} \sum_i w_{ki} o_{pi} = \sum_k \frac{\partial E_p}{\partial net_{pk}} w_{kj} = - \sum_k \delta_{pk} w_{kj}, \quad (31)$$

Za neurone na ostalih nivojih smo ugotovili:

$$\delta_{pj} = f'_j(net_{pj}) \sum_k \delta_{pk} w_{kj}. \quad (32)$$

Z enačbama (30) in (32) torej izračunamo δ_{pj} , ki so podlaga za nadaljnji izračun sprememb uteži po enačbi (26). Zgornji postopek za večnivojske umetne nevronske mreže se imenuje posplošeno pravilo delta.

Povzemimo glavne enačbe, ki določajo posplošeno pravilo delta:

$$\Delta_p w_{ji} = \eta \delta_{pj} o_{pi}, \quad (33)$$

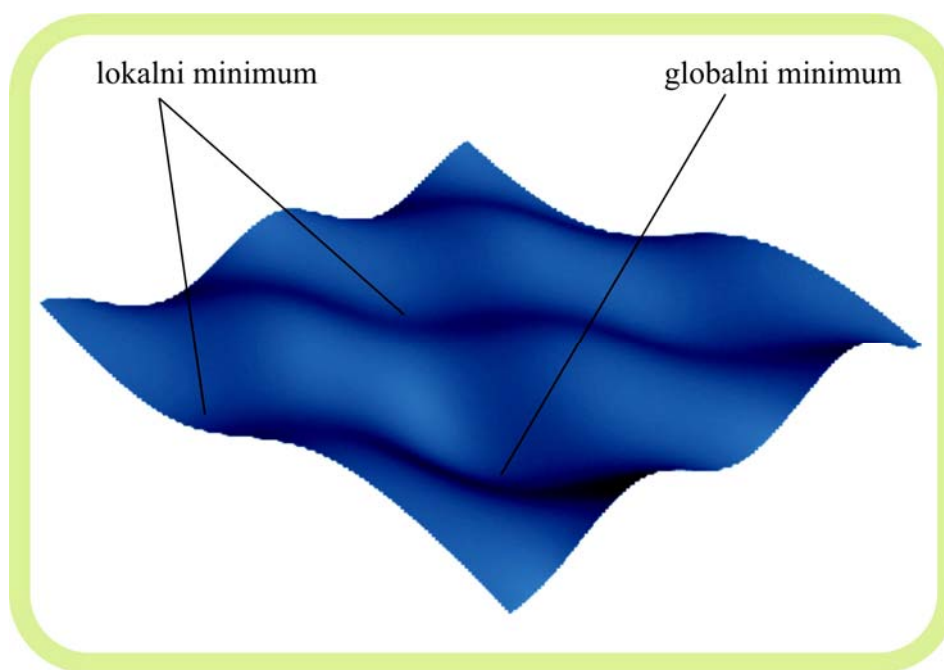
$$\delta_{pj} = \begin{cases} f'_j(net_{pj})(t_{pj} - o_{pj}); & \text{za uteži ob izhodnih nevronih,} \\ f'_j(net_{pj}) \sum_k \delta_{pk} w_{kj}; & \text{za uteži med skritimi neuroni.} \end{cases}$$

Sprememba uteži je torej sorazmerna odvodu napake nevrona, ki jo različno določimo za izhodne in ostale nivoje nevronov.

Uporaba posplošenega pravila delta zajema dve fazi. V prvi fazi vhodnim nevronom umetne nevronske mreže posredujemo vhodni podatek vhodno-izhodnega para. Vhodni neuroni pričnejo širiti signal po mreži in dobimo vrednosti izhodnih signalov o_{pj} . Izhodne vrednosti primerjamo z željenimi in glede na njuno razliko določimo signal napake δ na izhodnih nevronih. V drugi fazi vzvratno pošiljamo signal napake nevronom na skritih nivojih in spreminjamo uteži tudi na teh nivojih. Pri tem je zanimivo, da lahko določeno število uteži fiksiramo in njihovih vrednosti ne spreminjamo. Pri večnivojskih umetnih nevronske mrežah

gradient napake ne vodi vedno k minimumu, saj lahko poleg globalnega minimuma obstaja tudi več lokalnih. Veliko možnosti je, da mreža med učenjem naleti na minimum, ki je lokalni in ne globalni, kar pomeni, da vrednosti uteži niso optimalne.

Metoda padajočega gradienta zahteva, da je učni korak η ustrezne velikosti. Če je korak velik, bo velika tudi sprememba uteži v eni iteraciji in mreža se bo hitreje učila. Vendar to velja le na območju, ki je dovolj oddaljeno od globalnega minimuma. Ko se na ploskvi, ki opisuje napako (slika 18) približujemo minimumu in smo od njega oddaljeni za nekaj dolžin koraka, ga z velikim korakom hitreje dosežemo. Vendar je velika verjetnost, da ga bomo s prevelikim η zgrešili. Zato je dobro, če na ožjem območju globalnega minimuma učni korak zmanjšamo.



Slika 18: Ploskev napake pri večnivojskih usmerjenih nevronske mrežah.

4 UPORABA UMETNIH NEVRONSKIH MREŽ ZA OCENO TRDNOSTI LESA

Za določitev trdnosti lesa iz rezultatov neporušnih preizkusov smo izbrali večnivojske umetne nevronske mreže z učnim algoritmom vzratnega širjenja napake. Uporabili smo fortranski program NTR2003 in knjižnico nevronske mreže v komercialnem programu Matlab. Opozorimo, da uporabljeni algoritmi niso povsem primerljivi, saj se podprogrami v obeh programih razlikujejo v več podrobnostih. Rezultate prikažemo za vsako programsko orodje posebej.

4.1 Opis parametrov, ki vplivajo na učenje

Na uspešnost učenja nevronske mreže vplivajo mnogi parametri. Naštejmo in na kratko opišimo najpomembnejše.

4.1.1 Največja dovoljena napaka

Velikost dovoljene napake pomeni razliko med dejanskimi in z nevronske mreže izračunanimi izhodnimi vrednostmi podatkov oz. rezultatov.

4.1.2 Velikost učnega koraka

Pri izbiri velikosti učnega koraka v splošnem velja, da manjša velikost učnega koraka zahteva večje število iteracij in obratno. Vendar obstaja možnost, da se pri velikem učnem koraku umetna nevronska mreža ne nauči dobro, ker ne najde minimuma napake.

4.1.3 Maksimalno število iteracij

Na osnovi vhodnih podatkov izračuna umetna nevronska mreža vrednost izhodnega podatka v več iteracijah. Vrednost izhodnega podatka je odvisna od vrednosti uteži, določeni pri predhodni iteraciji. Nato mreža primerja izračunane izhodne vrednosti z željenimi in na

podlagi teh razlik po posplošenem pravilu delta spremeni vrednosti uteži. (Reščič, 2004). Z maksimalnim številom iteracij omejimo število ciklov učenja.

4.1.4 Geometrija mreže

Geometrija umetne nevronske mreže se nanaša na število vhodnih in izhodnih nevronov ter na število skritih nivojev in število nevronov na posameznem skitem nivoju. Če je mreža premajhna, ne bo sposobna predstaviti željene funkcije, če je prevelika, lahko pride do prekomernega prileganja (angl. overfitting). Ponavadi optimalne geometrije mreže ne poznamo, zato je za posamezen problem potrebno narediti mnogo parametričnih študij, s katerimi najdemo ustrezne dimenzije mreže. Obstajajo alternativne metode, ki temeljijo na učenju večje mreže kot je potrebno, nato pa se s pomočjo obrezovanja (angl. pruning) odstranijo odvečni deli mreže. Lahko se uporabijo tudi premajhne mreže in se jim nato dodajo povezave, ki omogočijo boljšo generalizacijo. Možna je tudi kombinacija obeh metod.

4.1.5 Izhodne funkcije

Opisali smo jih že v poglavju 3.2.2. Za naše primere uporabimo eksponentno sigmoidno izhodno funkcijo pri programu NTR2003 in različne funkcije pri knjižnici mrež v Matlabu.

4.2 Priprava podatkov na učenje

Učenju umetne nevronske mreže je namenjenih 250 naključno izbranih preizkušancev, testiranju pa preostalih 43 preizkušancev, za katere poznamo 9 lastnosti (ρ , E_{e-l} , E_{e-g} , E_{f-l} , E_{f-g} , E_{Tno} , E_{Zag} , E_{Sylv} in upogibna trdnost lesa f). Dejansko pri nobeni vrsti preiskav ne merimo vseh omenjenih parametrov istočasno. Zato diplomska naloga obravnava 5 različnih primerov zvez med vhodnimi in izhodnimi podatki, ki se med seboj razlikujejo glede na to, kateri podatki so bili uporabljeni kot vhodni podatki in kateri kot izhodni.

- V primeru 1 kot vhodni podatki nastopajo: ρ , E_{e-l} , E_{e-g} , E_{f-l} , E_{f-g} , E_{Tno} , E_{Zag} , E_{Sylv} , kot izhodni podatek pa upogibna trdnost lesa f . Teh 9 lastnosti enega preizkušanca, ki so razdeljene na 8 vhodnih spremenljivk in eno izhodno spremenljivko predstavlja en

vhodno-izhodni par. Če je ta par namenjen učenju se imenuje učni vhodno-izhodni par, če je uporabljen za testiranje, se imenuje testni vhodno-izhodni par.

- V primeru 2 nastopajo kot vhodni podatki: ρ , E_{f-g} , E_{Tno} , kot izhodni podatek pa upogibna trdnost lesa f . Prve tri lastnosti oz. spremenljivke in upogibna trdnost predstavljajo en vhodno-izhodni par učnih podatkov oz. testnih, če so namenjeni testiranju.
- V primeru 3 nastopajo kot vhodni podatki: ρ , E_{f-g} , E_{Zag} , kot izhodni podatek upogibna trdnost lesa f . Prve tri lastnosti predstavljajo vhodni del, upogibna trdnost pa izhodni del vsakega vhodno-izhodnega para podatkov.
- V primeru 4 nastopajo kot vhodni podatki: ρ , E_{f-g} , E_{Sylv} , kot izhodni podatek upogibna trdnost lesa f . Prve tri lastnosti so vhodni, upogibna trdnost pa izhodni podatek.
- V primeru 5 nastopajo kot vhodni podatki: ρ , E_{f-g} , E_{Tno} , kot izhodni podatek pa upogibna trdnost lesa f in E_{e-l} . Prve tri in drugi dve lastnosti predstavljajo en vhodno-izhodni par učnih podatkov.

4.3 Statistični kazalci uspešnosti nevronske mreže

Kako uspešna je bila umetna nevronska mreža pri učenju ugotovimo, ko ji posredujemo nize testnih parov. Mreža na osnovi vrednosti uteži, izračunanih v zadnji iteraciji učenja, izračuna izhodne vrednosti za testne podatke in jih primerja z dejanskimi vrednostmi. Razlike med dejanskimi in izračunanimi rezultati bomo prikazali z naslednjimi statističnimi kazalci:

- povprečna razlika: $\bar{\delta} = \frac{\sum_i \delta_i}{n}$, (34)

- povprečno absolutno odstopanje razlik: $a = \frac{\sum_i |\delta_i - \bar{\delta}|}{n}$, (35)

▪ standardni odklon: $s = \sqrt{\frac{\sum_i (\delta_i - \bar{\delta})^2}{n}}$, (36)

▪ koeficient korelacije: $r_{XY} = \frac{S_{XY}}{S_X S_Y}$, (37)

kjer je S_{XY} kovarianca, S_Y in S_X sta standardni deviaciji vzorca. Izračunamo jih na naslednji način:

$$S_{XY} = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{n}, \quad (38)$$

$$S_X = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}, \quad (39)$$

$$S_Y = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}}. \quad (40)$$

V izrazih (34 – 40) pomenijo:

δ_i . . . razliko med dejanskimi in z umetno nevronske mreže izračunanimi izhodnimi vrednostmi za i -ti vhodno-izhodni testni par,

n . . . število vhodno-izhodnih testnih parov,

X_i . . . dejansko vrednost izhodnega podatka i -tega testnega para,

\bar{X} . . . povprečje izhodnih podatkov X_i ,

Y_i . . . z nevronske mreže izračunano vrednost izhodnega podatka i -tega testnega para,

\bar{Y} . . . povprečje izhodnih podatkov Y_i , izračunanih z nevronske mreže.

4.4 Program NTR2003

Program NTR2003 je namenjen učenju in testiranju večnivojskih usmerjenih umetnih nevronske mreže. Proces učenja poteka na osnovi algoritma z vzratnim širjenjem napake, ki temelji na posplošenem pravilu delta.

Za zagon programa je potrebno pripraviti tri datoteke:

- datoteke s splošnimi podatki s končnico *.inp*,
- datoteke z učnimi podatki s končnico *.dat*,
- datoteke s testnimi podatki s končnico *.tst*.

V datoteki s splošnimi podatki podajamo naslednje parametre:

- dovoljena velikost napake,
- velikost učnega koraka,
- maksimalno število iteracij,
- multiplikacijski faktor,
- geometrija umetne nevronske mreže:
 - število nevronov vhodnega nivoja,
 - število nevronov izhodnega nivoja,
 - število skritih nivojev,
 - število nevronov na posameznem skitem nivoju.

V datoteko z učnimi podatki podamo nize vhodno-izhodnih učnih podatkov. V datoteko s testnimi podatki podamo nize vhodno-izhodnih testnih podatkov. V obeh datotekah je potrebno v prvih stolpcih podati vhodne podatke vhodno-izhodnih parov, tem sledijo stolpci z izhodnimi podatki vhodno-izhodnih parov podatkov.

Rezultate program shrani v dve datoteki:

- datoteka z utežmi s končnico *.wgt*.
- datoteka s končnico *.out*.

V tekstovni datoteki s končnico *.out* se nahajajo podatki, kot so čas procesiranja, število nenaučenih učnih parov vhodno-izhodnih podatkov, velikost napake in statistični kazalci, ki povedo, kako uspešna je umetna nevronska mreža na testnih podatkih.

V programu NTR2003, zaradi lažjega obvladovanja velikosti posameznih spremenljivk in večje stabilnosti računa, vse vrednosti količin preslikamo na interval [0.1, 0.9]. To dosežemo z normiranjem podatkov po preprosti linearni formuli:

$$U_i = 0.1 + \frac{(X_i - X_{\min})}{(X_{\max} - X_{\min})} \cdot 0.8, \quad (41)$$

kjer je:

U_i . . . vrednost normirane spremenljivke,

X_i . . . vrednost spremenljivke,

X_{\min} . . . vrednost najmanjše spremenljivke v nizu vseh spremenljivk,

X_{\max} . . . vrednost največje spremenljivke v nizu vseh spremenljivk.

Postopek učenja mreže v NTR2003 je iterativen in poteka na osnovi spreminjanja uteži med nevroni. To pomeni, da večkrat zapored procesira podane nize vhodno-izhodnih podatkov in si skuša "zapomniti" čimveč le-teh. Na podlagi "spomina" bo nato umetna nevronska mreža zmožna izračunati za neznan vhodni podatek, ki je podoben učnim, ustrezno izhodno vrednost. Na učenje umetne nevronske mreže vpliva vrsta parametrov, kot so: maksimalno število iteracij, velikost učnega koraka, geometrija mreže, multiplikacijski faktor, maksimalna velikost napake in tudi sami podatki, na katerih se mreža uči. Poudarek v diplomskem delu je ravno na zadnjem parametru, torej na podatkih. Število iteracij, velikost dovoljene napake in velikost učnega koraka so uporabljeni z upoštevanjem ugotovitev v diplomskem delu D. Reščiča (2004), saj se pri obeh delih obravnava podobne podatke, t.j. lastnosti lesa. Glede geometrije umetne nevronske mreže ni mogoče točno definirati, katera ustreza konkretnim podatkom. Pri vseh petih primerih smo se odločili za obravnavo mrež z dvema ali tremi skritimi nivoji. Ko je program NTR2003 izračunal vse izbrane tipe mrež, smo za nadaljno obravnavo izbrali tisto, ki je imela najmanjše število nenaučenih učnih parov. Nato smo pri izbrani mreži izločili nenaučene pare. Za vsak nenaučen učni par smo izračunali norme vhodnih podatkov glede na ostale nenaučene učne pare podatkov in glede na naučene učne pare podatkov. S tistim učnim parom (nenaučenim ali naučenim), s katerim je nenaučen par tvoril najmanjšo normo vhodnih podatkov, smo ga združili. Združitev je potekala v smislu izračuna povprečnih vrednosti vhodnih in izhodnih podatkov. Nato so ti novi združeni podatki predstavljali nov učni par, ki je nadomestil tista dva, ki sta tvorila minimalno normo. Na tem principu je temeljilo vseh pet primerov.

4.4.1 Primer 1

Prva skupina obravnavanih podatkov je sestavljena iz 8 vhodnih podatkov in enega izhodnega (poglavje 4.2). Uporabili smo mreže z dvema in mreže s tremi skritimi nivoji. Poleg tega smo preizkusili različna števila maksimalnih iteracij učnega procesa, velikost učnega koraka je bila 0.1 ali 0.5, absolutna velikost dovoljene napake normiranih podatkov pa 0.1 ali 0.05. Vse uporabljene kombinacije so predstavljene v preglednici 2. Za izbrane vrednosti parametrov smo se odločili na osnovi ugotovitev diplomske naloge D. Reščiča (2004). Zanimalo nas je, katera izmed mrež se bo naučila največ vhodno-izhodnih učnih parov z namenom nadaljnje obravnave tovrstne mreže. Iz rezultatov je razvidno, da se je več vrst mrež naučilo vse učne pare, zato tu učnih podatkov nismo dodatno obdelali.

Preglednica 2: Vrednosti uporabljenih parametrov pri učenju.

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	dovoljena napaka normiranih podatkov	št. nenaučenih učnih parov
8-20-15-1	10000	0.5	0.1	52
8-20-15-1	50000	0.5	0.1	36
8-20-15-1	100000	0.5	0.1	33
8-20-15-1	10000	0.1	0.1	51
8-20-15-1	50000	0.1	0.1	0
8-20-15-1	100000	0.1	0.1	0
8-20-15-1	100000	0.5	0.05	0
8-20-15-1	10000	0.1	0.05	0
8-20-20-1	10000	0.5	0.1	11
8-20-20-1	50000	0.5	0.1	0
8-20-20-1	100000	0.5	0.1	0
8-20-20-1	10000	0.1	0.1	54
8-20-20-1	50000	0.1	0.1	0
8-20-20-1	100000	0.1	0.1	0
8-20-20-1	50000	0.1	0.05	0
8-20-20-1	100000	0.1	0.05	0
8-20-20-1	10000	0.5	0.05	2
8-20-20-1	50000	0.5	0.05	0
8-20-15-15-1	10000	0.5	0.1	23

se nadaljuje ...

... nadaljevanje

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	dovoljena napaka normiranih podatkov	št. nenaučenih učnih parov
8-20-15-15-1	50000	0.5	0.1	61
8-20-15-15-1	100000	0.5	0.1	97
8-20-15-15-1	10000	0.1	0.1	6
8-20-15-15-1	50000	0.1	0.1	0
8-20-15-15-1	100000	0.1	0.1	0
8-20-15-15-1	100000	0.5	0.05	0
8-20-15-15-1	10000	0.1	0.05	0
8-30-20-20-1	10000	0.5	0.1	0
8-30-20-20-1	50000	0.5	0.1	0
8-30-20-20-1	100000	0.5	0.1	0
8-30-20-20-1	10000	0.1	0.1	3
8-30-20-20-1	50000	0.1	0.1	0
8-30-20-20-1	100000	0.1	0.1	0
8-30-20-20-1	50000	0.1	0.05	76
8-30-20-20-1	100000	0.1	0.05	85
8-30-20-20-1	10000	0.5	0.05	127
8-30-20-20-1	50000	0.5	0.05	0
8-30-20-20-1	100000	0.5	0.05	0
8-8-7-7-1	10000	0.5	0.1	17
8-8-8-1	50000	0.5	0.1	8
8-30-20-20-1	5000	0.5	0.1	18
8-30-20-20-1	1000	0.5	0.1	66
8-30-20-20-1	100	0.5	0.1	81

Pri tistih mrežah, ki so se naučile vseh 250 nizov učnih parov, smo preverili, kako so se odzvale na testne pare. Parametre učenja skupaj s statističnimi kazalci uspešnosti učenja prikazujemo v preglednici 3. Najmanjša vrednost za povprečno absolutno odstopanje razlik trdnosti je 8.869 (normirano povprečno absolutno odstopanje razlik trdnosti se giblje med 20% in 35%), za standardni odklon pa 11.418 (mreža z geometrijo 8-30-20-20-1). Opozorimo, da sta dokaj veliki vrednosti. Koeficient korelacije za to mrežo je največji izmed vseh, vendar še vedno vrednost 0.4931 ne pomeni velike linearne odvisnosti med željenimi in z mrežo izračunanimi vrednostmi. To pomeni, da se je mreža naučila vseh 250 učnih parov z zahtevano natančnostjo. Vendar ko so ji bili predstavljeni testni pari, so bile vrednosti med

dejanskimi in z nevronske mreže izračunanimi rezultati precej različne. Zavedamo se, da je les nehomogen material in posledično imajo njegove lastnosti velik raztros (poglavje 2), kar se tudi odraža v velikosti napake med dejanskimi in izračunanimi izhodnimi vrednostmi. Po drugi strani pa dobra naučenost učnih in slabi rezultati na testnih parih kažejo na pretirano prilagoditev mreže učnim parom (prenauchenost) in slabšo sposobnost posploševanja.

Preglednica 3: Vrednosti statističnih kazalcev za mreže brez nenaučeni učni parov.

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	dovoljena napaka normiranih podatkov	št. nenauč. učni parov	povprečno absolutno odstopanje razlik	stand. odklon	koef. korelac.
8-20-15-1	50000	0.1	0.1	0	12.061	16.080	0.279
8-20-15-1	100000	0.1	0.1	0	12.061	16.080	0.279
8-20-15-1	100000	0.5	0.05	0	12.409	15.898	0.179
8-20-15-1	10000	0.1	0.05	0	12.409	15.898	0.179
8-20-20-1	50000	0.5	0.1	0	11.656	15.901	0.243
8-20-20-1	100000	0.5	0.1	0	11.656	15.901	0.243
8-20-20-1	50000	0.1	0.1	0	12.036	15.863	0.351
8-20-20-1	100000	0.1	0.1	0	12.036	15.863	0.351
8-20-20-1	50000	0.1	0.05	0	14.727	17.921	0.279
8-20-20-1	100000	0.1	0.05	0	14.727	17.921	0.279
8-20-20-1	50000	0.5	0.05	0	14.304	18.212	0.156
8-20-15-15-1	50000	0.1	0.1	0	13.055	16.424	0.334
8-20-15-15-1	100000	0.1	0.1	0	13.055	16.424	0.334
8-20-15-15-1	100000	0.5	0.05	0	14.037	17.881	0.147
8-20-15-15-1	10000	0.1	0.05	0	14.037	17.881	0.147
8-30-20-20-1	10000	0.5	0.1	0	8.869	11.418	0.493
8-30-20-20-1	50000	0.5	0.1	0	8.869	11.418	0.493
8-30-20-20-1	100000	0.5	0.1	0	8.869	11.418	0.493
8-30-20-20-1	50000	0.1	0.1	0	10.788	13.293	0.445
8-30-20-20-1	100000	0.1	0.1	0	10.788	13.293	0.445
8-30-20-20-1	50000	0.5	0.05	0	12.461	16.733	0.203
8-30-20-20-1	100000	0.5	0.05	0	12.461	16.733	0.203

Vhodno-izhodni pari učni in testni podatkov so v prilogi A.

4.4.2 Primer 2

V primeru 2 so kot vhodni podatki obravnavani: ρ , E_{f-g} , E_{Tno} , izhodni podatek oz. rezultat je trdnost f . Tudi v tem primeru smo preizkusili več tipov mrež, vendar se za razliko od prvega primera nobena nevronska mreža ni naučila vseh učnih parov. V vseh primerih je bila velikost dopustne napake normiranih podatkov enaka 0.1. Za nadaljnjo obravnavo je izbrana tista, ki ima najmanj nenaučenih učnih parov (44). To je mreža z geometrijo 3-30-20-20-1, z maksimalnim številom iteracij 100000 in velikostjo učnega koraka 0.1.

Preglednica 4: Vrednosti uporabljenih parametrov pri učenju.

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučenih učnih parov
3-20-15-1	10000	0.5	85
3-20-15-1	50000	0.5	77
3-20-15-1	100000	0.5	71
3-20-15-1	10000	0.1	98
3-20-15-1	50000	0.1	77
3-20-15-1	100000	0.1	66
3-20-20-1	10000	0.5	84
3-20-20-1	50000	0.5	64
3-20-20-1	100000	0.5	60
3-20-20-1	10000	0.1	98
3-20-20-1	50000	0.1	80
3-20-20-1	100000	0.1	56
3-20-15-15-1	10000	0.5	97
3-20-15-15-1	50000	0.5	93
3-20-15-15-1	100000	0.5	95
3-20-15-15-1	10000	0.1	87
3-20-15-15-1	50000	0.1	65
3-20-15-15-1	100000	0.1	71
3-30-20-20-1	10000	0.5	76
3-30-20-20-1	50000	0.5	79
3-30-20-20-1	100000	0.5	86
3-30-20-20-1	10000	0.1	80
3-30-20-20-1	50000	0.1	74
3-30-20-20-1	100000	0.1	44
3-8-8-1	10000	0.5	85

se nadaljuje ...

... nadaljevanje

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučenih učnih parov
3-8-7-7-1	10000	0.5	76
3-8-7-7-1	50000	0.1	75
3-8-8-1	50000	0.1	88
3-8-7-1	10000	0.5	90
3-8-7-1	50000	0.5	79
3-8-7-1	100000	0.5	79
3-8-7-1	10000	0.1	97
3-8-7-1	50000	0.1	79
3-8-7-1	100000	0.1	84
3-8-8-1	50000	0.5	83
3-8-8-1	100000	0.5	82
3-8-8-1	10000	0.1	95
3-8-8-1	100000	0.1	79
3-8-7-7-1	50000	0.5	96
3-8-7-7-1	100000	0.5	95
3-8-7-7-1	10000	0.1	87
3-8-7-7-1	100000	0.1	84
3-9-8-8-1	10000	0.5	81
3-9-8-8-1	50000	0.5	92
3-9-8-8-1	100000	0.5	91
3-9-8-8-1	10000	0.1	88
3-9-8-8-1	50000	0.1	75
3-9-8-8-1	100000	0.1	72
3-30-20-20-1	5000	0.1	95
3-30-20-20-1	1000	0.1	102
3-30-20-20-1	100	0.1	117

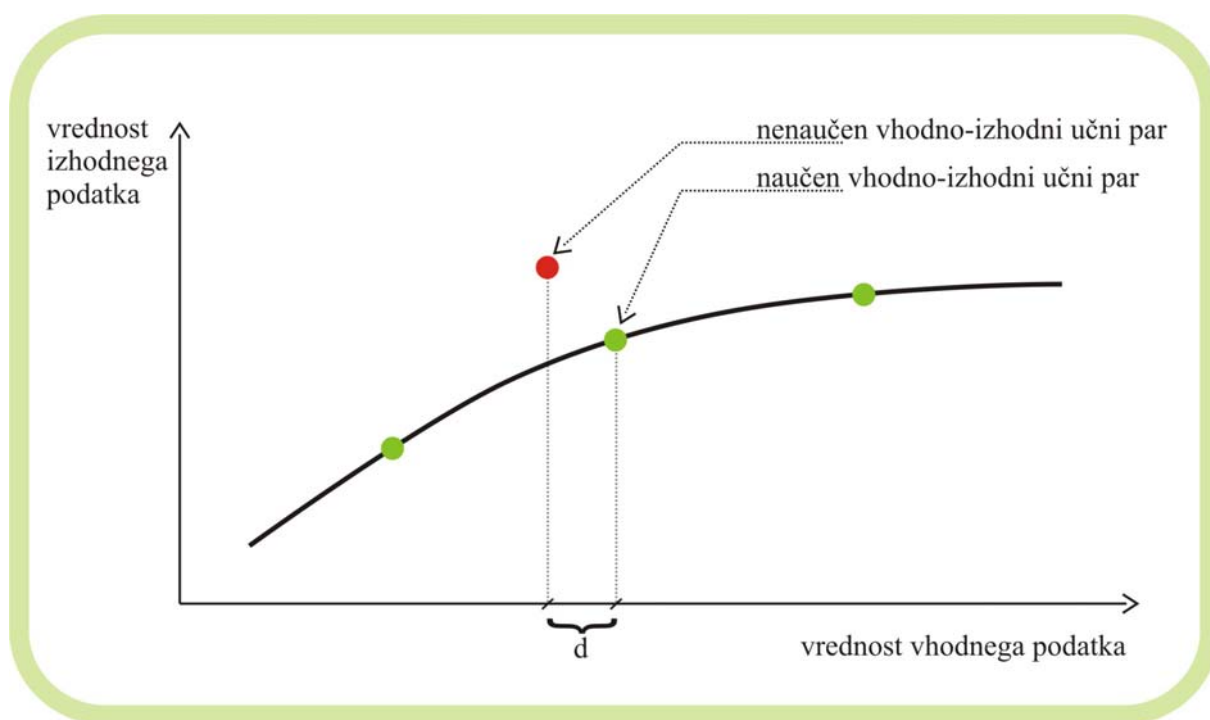
Zanima nas, kako podatki vplivajo na učenje. Če imata 2 nenaučena para precej različne vhodne podatke in podoben izhodni podatek, je normalno, da to mrežo "zmoti" pri učenju. Enako velja, če ima nenaučen par podobne vhodne podatke kot naučen učni par. Kako smo to preverili? Iz niza učnih parov smo izpisali 44 nenaučenih in jih razvrstili glede na največjo razliko med dejansko vrednostjo izhodnega podatka in z nevronske mreže izračunano. Nato smo za vsak nenaučen par izračunali norme glede na naučene učne pare (slika 19) in glede na ostale nenaučene učne pare (slika 20).

Za račun norme parov smo uporabili formulo:

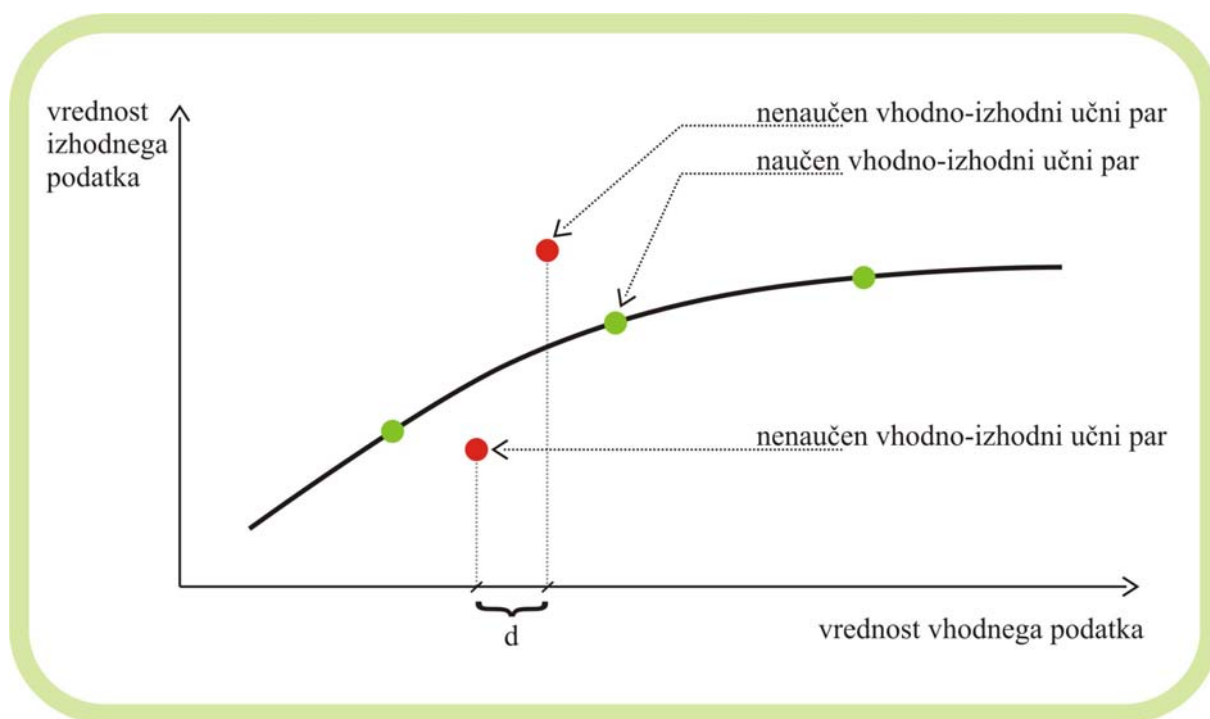
$$N = (W_p - W_r)^2 + (Y_p - Y_r)^2 + (Z_p - Z_r)^2 = d^2, \quad (42)$$

kjer je:

- N . . . norma dveh točk,
- (W_p, Y_p, Z_p) . . . vrednosti točke p ,
- (W_r, Y_r, Z_r) . . . vrednosti točke r ,
- d . . . razdalja med točkama.



Slika 19: Vhodni podatek nenaučenega učnega para podoben naučenemu.



Slika 20: Vhodni podatek nenaučene učnega para podoben drugemu nenaučenu.

Za vsak nenaučeni par izračunamo minimalno normo, ki pokaže, s katerim od naučenih in preostalih nenaučeni učni parov so si vhodni podatki najbližje. Tako dobimo 2 učna para, ki ju nadomestimo z novim. Ta vsebuje povprečne vhodne in izhodni podatek prvotnih dveh učni parov. Če je med seboj posredno (preglednica 5) ali neposredno (preglednica 6) povezanih več parov, jih združimo na enak način.

Preglednica 5: Primer posrednega združevanja učni parov.

indeks para nenaučeni učni podatkov	pripadajoči par učni podatkov	norma vhodni podatkov
94	109 nenaučeni	0.001055
206	214 nenaučeni	0.000023
191	206 nenaučeni	0.000199
28	94 nenaučeni	0.002834
109	214 nenaučeni	0.000276
214	206 nenaučeni	0.000023

Preglednica 6: Primer neposrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
3	230 nenaučen	0.000188
230	3 nenaučen	0.000188

Po združevanju učnih parov dobimo nov vhodno-izhodni niz 211 učnih parov, ki jih podamo programu NTR2003 v datoteko *.dat* in program ponovno zaženemo. Umetna nevronska mreža se nauči vse "nove" učne pare.

Poglejmo kako se na združevanje podatkov odzovejo statistični kazalci:

Preglednica 7: Vrednosti statističnih kazalcev.

št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
pred združevanjem	13.347	17.449	0.03
po združevanju	11.294	15.505	0.3

Iz preglednice 7 je razvidno, da sta se povprečno absolutno odstopanje razlik (normirano povprečno absolutno odstopanje razlik trdnosti je pred združevanjem približno 30%, po združevanju podatkov pa 25%) in standardni odklon zmanjšala, kar pomeni, da se je učenje z združevanjem učnih parov izboljšalo. Ravno tako se je izboljšal koeficient korelacije, a je še vedno zelo majhen. Tudi v primeru združevanja učnih podatkov sta povprečno absolutno odstopanje razlik in standardni odklon testnih podatkov velika. Ugotovimo lahko, da se je po združevanju mreža uspešno naučila učne podatke, sposobnost posploševanja pa ni zelo dobra. Razlog je v veliki razpršenosti podatkov in prenačenosti mreže.

Vhodno-izhodni učni in testni pari podatkov pred združevanjem ter izračun norm in novi učni pari po združevanju učnih podatkov so prikazani v prilogi B.

4.4.3 Primer 3

Vhodni podatki so ρ , E_{f-g} in E_{zag} , izhodni pa je zopet trdnost lesa f . Dopustna absolutna napaka normiranih parov je 0.1. Največ učnih parov se je naučila mreža z geometrijo 3-20-15-1, z maksimalnim številom iteracij 50000, z velikostjo učnega koraka 0.1. Nenaučeni je ostalo 57 učnih parov od skupno 250 (preglednica 8).

Preglednica 8: Vrednosti uporabljenih parametrov pri učenju.

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučeni učnih parov
3-8-7-1	10000	0.5	97
3-8-7-1	50000	0.5	79
3-8-7-1	100000	0.5	70
3-8-7-1	10000	0.1	92
3-8-7-1	50000	0.1	87
3-8-7-1	100000	0.1	77
3-8-8-1	10000	0.5	95
3-8-8-1	50000	0.5	85
3-8-8-1	100000	0.5	80
3-8-8-1	10000	0.1	90
3-8-8-1	50000	0.1	76
3-8-8-1	100000	0.1	69
3-8-7-7-1	10000	0.5	94
3-8-7-7-1	50000	0.5	115
3-8-7-7-1	100000	0.5	97
3-8-7-7-1	10000	0.1	90
3-8-7-7-1	50000	0.1	79
3-8-7-7-1	100000	0.1	77
3-9-8-8-1	10000	0.5	92
3-9-8-8-1	50000	0.5	102
3-9-8-8-1	100000	0.5	97
3-9-8-8-1	10000	0.1	91
3-9-8-8-1	50000	0.1	75
3-9-8-8-1	100000	0.1	72
3-20-15-15-1	50000	0.1	68
3-20-15-1	50000	0.1	57

se nadaljuje ...

... nadaljevanje

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučenih učnih parov
3-30-20-20-1	100000	0.1	73
3-20-15-1	5000	0.1	91
3-20-15-1	1000	0.1	96
3-20-15-1	100	0.1	116

V tem primeru so na enak način kot v primeru 2 izračunane norme nenaučenih učnih parov glede na naučene (slika 19) in glede na ostale nenaučene učne pare (slika 20). Nato so podatki posredno (preglednica 9) ali neposredno (preglednica 10) združeni med sabo glede na minimalne medsebojne norme.

Preglednica 9: Primer posrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
144	157 naučen	0.00305
211	157 naučen	0.0027
182	211 nenaučen	0.0053

Preglednica 10: Primer neposrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
202	4 nenaučen	0.00085
4	202 nenaučen	0.00085

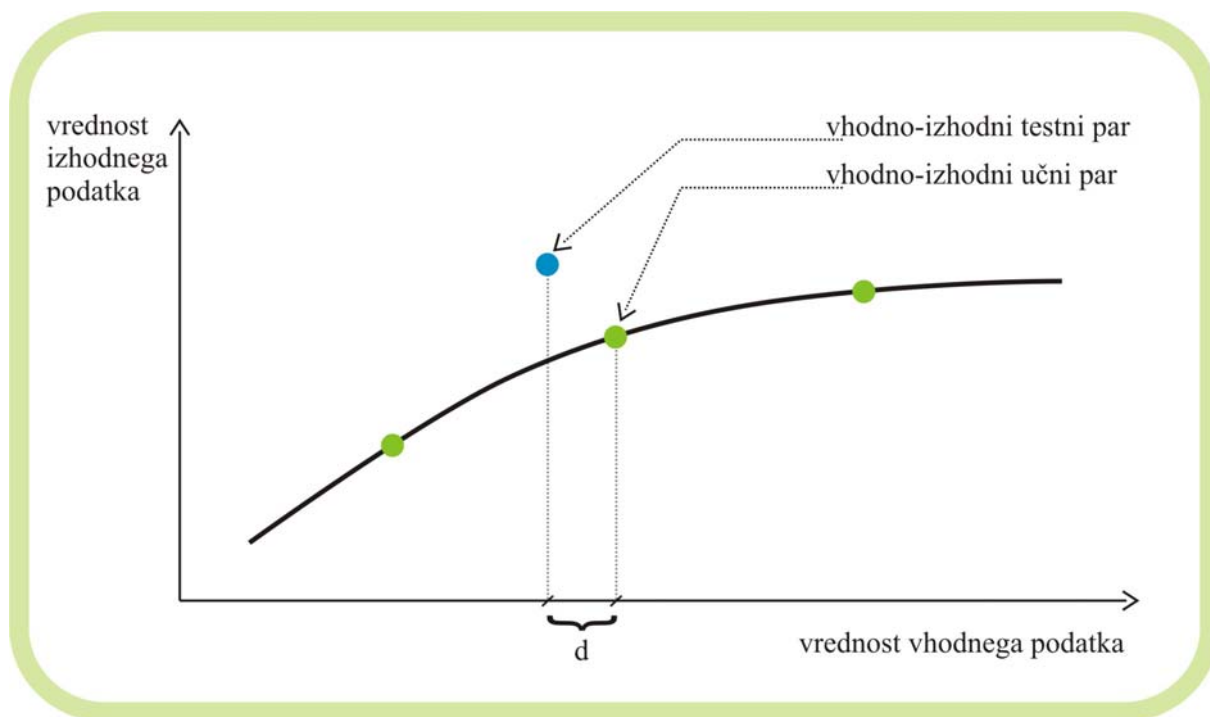
Umetni nevronske mreže podamo nove združene učne pare; namesto prvotnih 250 jih je sedaj 196 združenih. Število nenaučenih učnih parov se je po pričakovanju zmanjšalo; od 196 učnih parov je ostalo nenaučenih le 9. Tako smo ponovno potrdili, da razpršenost podatkov pomembno vpliva na število nenaučenih učnih parov.

Vendar je še pomembnejša od števila naučenih podatkov sposobnost posploševanja, za kar uporabimo testne pare. Iz preglednice 11 je razvidno, da sta velikosti povprečnega absolutnega odstopanja razlik in standardnega odklona po združevanju učnih parov celo večji kot pred združevanjem. Zato skušamo te vrednosti zmanjšati oz. hočemo, da bo mreža učinkovitejša. Namesto združevanja uporabimo preurejanje ali izločanje testnih parov, ki najbolj odstopajo.

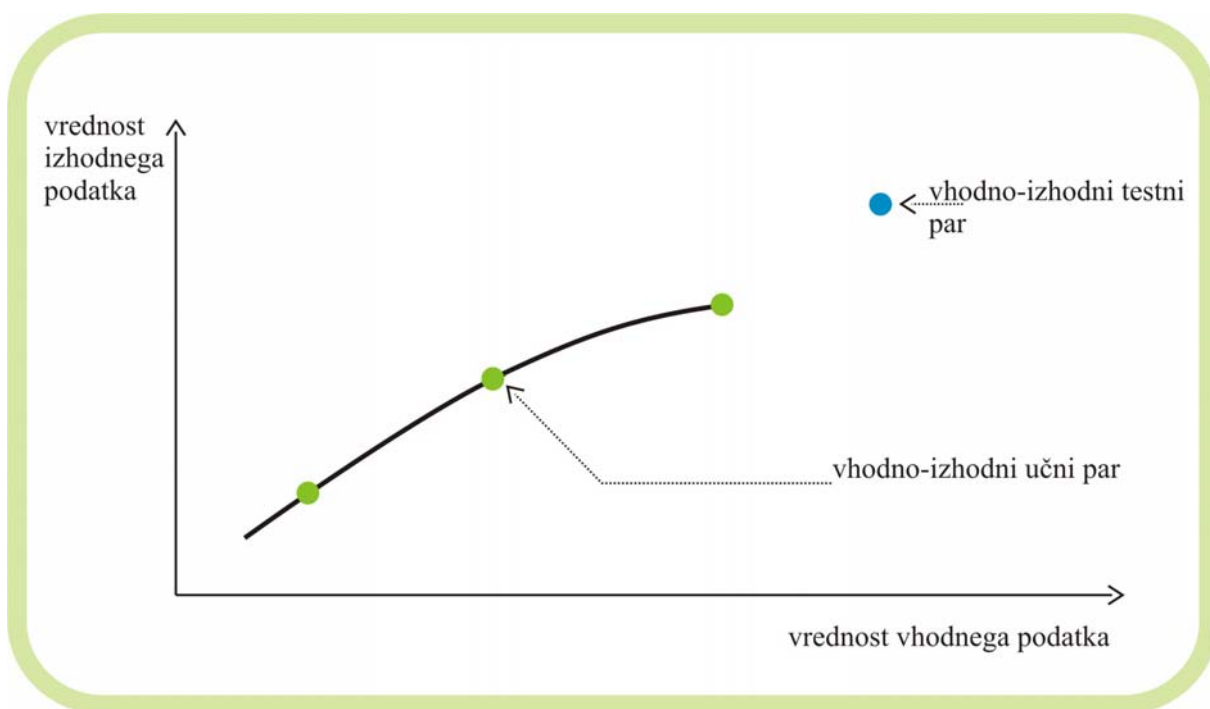
Testne pare podatkov razvrstimo glede na največje razlike med dejanskimi vrednostmi in vrednostmi izračunanimi z nevronske mreže. Nato za vsak testni par izračunamo norme z vsemi 196 učnimi pari in pogledamo, kateri ima minimalno. Če je testni vhodni podatek blizu enemu od učnih vhodnih podatkov (slika 21) pomeni, da imata podobne vhodne vrednosti in precej različno izhodno vrednost. Zato tak testni par izločimo iz računa. Če pa se izkaže, da je minimalna norma testnega podatka še vedno dovolj velika, potem tak testni par vključimo med učne. Izmed učnih parov pa enega vključimo med testne (slika 22).

Preglednica 11: Vrednosti statističnih kazalcev.

	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
pred združevanjem	9.384	12.747	0.348
po združevanju	11.850	14.211	0.267
(a)	12.095	14.378	0.265
(b)	9.914	12.695	0.363
(c)	9.013	11.731	0.350



Slika 21: Testni par podoben učnemu paru.



Slika 22: Testni par izven meja, ki jih postavljajo učni pari.

Pod točko (a) iz preglednice 11 je bil iz niza testnih parov izločen par z minimalno normo. Ker to ni izboljšalo učinkovitosti učenja pomeni, da je minimalna norma testnega para velika in obravnavamo primer, prikazan na sliki 22. Zato pod točko (b) testni par z največjo minimalno normo vključimo med učne pare, od tod pa med testne pare vključimo par, ki da z omenjenim testnim parom minimalno normo. Tudi to ne izboljša učinkovitosti učenja. Pod točko (c) izvedemo isti postopek kot pod (b), le da med testne pare vključimo tistega učnega, ki da z izključenim testnim parom maksimalno normo. Ta poteza se izkaže za uspešnejšo, saj sta povprečno absolutno odstopanje razlik (normirano povprečno absolutno odstopanje razlik trdnosti je pred združevanjem približno 22%, po združevanju pa 20%) in standardni odklon manjša kot pred učenjem z združevanjem učnih parov. Koeficient korelacije je za spoznanje večji. Ti posegi in njihovi rezultati kažejo na problematiko velikega raztrosa podatkov in širok razpon vrednosti, ki jih zavzamejo podatki. Nakazujejo tudi potrebo po večjem številu učnih podatkov.

Vhodno-izhodni učni in testni pari podatkov pred združevanjem ter izračun norm in novi učni pari po združevanju učnih podatkov so prikazani v prilogi C.

4.4.4 Primer 4

V primeru 4 so vhodni podatki ρ , E_{f-g} , E_{Sylv} , izhodna vrednost pa trdnost lesa f . Ponovno je obravnavanih več nevronske mreže in največ od vseh 250 učnih parov se jih je naučila mreža z geometrijo 3-8-7-7-1, z maksimalnim številom iteracij 50000, z velikostjo učnega koraka 0.1. Velikost dovoljene napake je bila v vseh primerih 0.1. Nenaučeni je ostalo 66 učnih parov (preglednica 12).

Preglednica 12: Vrednosti uporabljenih parametrov pri učenju.

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučeni učnih parov
3-8-7-1	10000	0.5	89
3-8-7-1	50000	0.5	78
3-8-7-1	100000	0.5	76

se nadaljuje ...

... nadaljevanje

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučenih učnih parov
3-8-7-1	10000	0.1	100
3-8-7-1	50000	0.1	98
3-8-7-1	100000	0.1	90
3-8-8-1	10000	0.5	93
3-8-8-1	50000	0.5	81
3-8-8-1	100000	0.5	85
3-8-8-1	10000	0.1	96
3-8-8-1	50000	0.1	85
3-8-8-1	100000	0.1	69
3-8-7-7-1	10000	0.5	93
3-8-7-7-1	50000	0.5	96
3-8-7-7-1	100000	0.5	102
3-8-7-7-1	10000	0.1	99
3-8-7-7-1	50000	0.1	66
3-8-7-7-1	100000	0.1	71
3-9-8-8-1	10000	0.5	87
3-9-8-8-1	50000	0.5	97
3-9-8-8-1	100000	0.5	91
3-9-8-8-1	10000	0.1	96
3-9-8-8-1	50000	0.1	67
3-9-8-8-1	100000	0.1	73
3-20-15-15-1	50000	0.1	77
3-20-15-1	50000	0.1	87
3-30-20-20-1	100000	0.1	107
3-20-15-1	50000	0.1	87
3-8-7-7-1	5000	0.1	97
3-8-7-7-1	1000	0.1	101
3-8-7-7-1	100	0.1	116

Tudi v tem primeru, kot v prejšnjih dveh, izračunamo norme podatkov. Podatke posredno (preglednica 13) in neposredno združujemo (preglednica 14) glede na minimalne norme, ki jih tvorijo nenaučeni z naučenimi (slika 19) in preostalimi nenaučenimi učnimi pari (slika 20). Po združevanju dobimo 189 vhodno-izhodnih učnih parov, ki jih predstavimo nevronske mreži. Po ponovnem učenju ostane nenaučenih 27 novih učnih parov.

Preglednica 13: Primer posrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
95	185 nenaučen	0.00034
17	95 nenaučen	0.00075
185	95 nenaučen	0.00034

Preglednica 14: Primer neposrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
128	29 nenaučen	0.00055
29	128 nenaučen	0.00055

Po združevanju sta povprečno absolutno odstopanje razlik in standardni odklon (preglednica 15) zopet večja kot pred združevanjem. Zato učinkovitost povečamo s preurejanjem testnih podatkov, kot smo opisali v primeru 3. Torej preverimo, zakaj so dejanske in z mrežo izračunane vrednosti preveč različne. Vzrok je lahko v zelo podobnih vhodnih vrednostih testnega in učnih podatkov in preveliki razliki v izhodni vrednosti (slika 21). Lahko pa se testni podatek nahaja izven območja učnih podatkov (slika 22).

Preglednica 15: Vrednosti statističnih kazalcev.

	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
pred združevanjem	9.665	12.019	0.256
po združevanju	10.186	13.55	0.312
(a)	9.910	13.32	0.337
(b)	9.414	12.204	0.20
(c)	10.458	13.22	0.25
(d)	9.553	12.30	0.136

Pod točko (a) iz preglednice 15 je bil iz niza testnih parov izločen par z minimalno normo. Ker to ni izboljšalo učinkovitosti učenja pomeni, da je minimalna norma vhodnih podatkov

testnega para velika in obravnavamo primer prikazan na sliki 22. Zato pod točko (b) testni par z največjo minimalno normo vključimo med učne pare. Od tod med testne vključimo tisti učni par, ki da z omenjenim testnim maksimalno normo. Ta način se izkaže za najučinkovitejši, vseeno pa povprečno absolutno odstopanje razlik trdnosti in standardni odklon ne padeta pod vrednosti, ki jih imata pri prvotnem učenju pred združevanjem podatkov. Pod točko (c) izvedemo isti postopek kot pod (b), le da med testne pare vključimo tistega učnega, ki da z izključenim testnim parom najmanjšo minimalno normo. Tudi ta poteza se izkaže za neuspešno, saj sta povprečno absolutno odstopanje razlik trdnosti in standardni odklon večja kot pred učenjem z združevanjem učnih parov. Preizkusimo še testni par z največjo minimalno normo vključiti k učnim parom in nobenega izmed učnih parov ne vključimo h testnim (točka (d)). Tudi pri tej varianti statistična kazalca nista manjša kot v načinu (b). Ravno tako pri koeficientu korelacije ne pride do večjih izboljšav (normirano povprečno absolutno odstopanje razlik trdnosti je pred in po združevanju približno 22%). Torej je pri danih vhodno-izhodnih podatkih združevanje učnih parov vplivalo na število nenaučenih parov, vendar učenje samo ni bilo uspešnejše. To pomeni, da mreža ni bila sposobna iz naučenih parov posplošiti "ugotovitev" na testne pare z drugačnimi vrednostmi.

Vhodno-izhodni učni in testni pari podatkov pred združevanjem ter izračun norm in novi učni pari po združevanju učnih podatkov so prikazani v prilogi D.

4.4.5 Primer 5

Zadnji primer obravnava kot vhodne podatke ρ , E_{f-g} , E_{Tno} , izračunati želimo upogibno trdnost lesa f in lokalni elastični modul E_{e-l} . Velikost dopustne absolutne napake normiranih vrednosti je 0.1. Med vsemi uporabljenimi mrežami je pri mreži z geometrijo 3-20-15-2, z maksimalnim številom iteracij 50000 in z velikostjo učnega koraka 0.1, ostalo nenaučenih najmanj učnih parov. Vseeno je kar 94 učnih parov imelo razliko med dejanskimi in z mrežo izračunanimi rezultati večjo od velikosti dovoljene napake. Mrežo 3-20-15-2 smo izbrali za nadaljnje študije.

Preglednica 16: Vrednosti uporabljenih parametrov pri učenju.

geometrija mreže	maksimalno št. iteracij	velikost učnega koraka	št. nenaučenih učnih parov
3-8-7-2	10000	0.5	122
3-8-7-2	50000	0.5	120
3-8-7-2	100000	0.5	110
3-8-7-2	10000	0.1	134
3-8-7-2	50000	0.1	118
3-8-7-2	100000	0.1	112
3-8-8-2	10000	0.5	123
3-8-8-2	50000	0.5	111
3-8-8-2	100000	0.5	103
3-8-8-2	10000	0.1	134
3-8-8-2	50000	0.1	107
3-8-8-2	100000	0.1	97
3-8-7-7-2	10000	0.5	123
3-8-7-7-2	50000	0.5	139
3-8-7-7-2	100000	0.5	140
3-8-7-7-2	10000	0.1	126
3-8-7-7-2	50000	0.1	107
3-8-7-7-2	100000	0.1	102
3-9-8-8-2	10000	0.5	116
3-9-8-8-2	50000	0.5	143
3-9-8-8-2	100000	0.5	148
3-9-8-8-2	10000	0.1	120
3-9-8-8-2	50000	0.1	100
3-9-8-8-2	100000	0.1	113
3-20-15-2	50000	0.1	94
3-20-15-15-2	50000	0.1	97
3-30-20-20-2	100000	0.1	112
3-20-15-2	5000	0.1	133
3-20-15-2	1000	0.1	132
3-20-15-2	100	0.1	137

Zopet izračunamo norme med nenaučenimi in naučenimi učnimi pari (slika 19) ter norme med samimi nenaučenimi učnimi pari (slika 20). Nato pare med seboj združimo posredno (preglednica 17) in neposredno (preglednica 18).

Preglednica 17: Primer posrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
93	147 nenaučen	0.00049
195	147 nenaučen	0.00017
34	158 nenaučen	0.00137
57	147 nenaučen	0.00026
158	93 nenaučen	0.00086
147	195 nenaučen	0.00017

Preglednica 18: Primer neposrednega združevanja učnih parov.

indeks para nenaučenih učnih podatkov	pripadajoči par učnih podatkov	norma vhodnih podatkov
95	185 nenaučen	0.00034
17	95 nenaučen	0.00075
185	95 nenaučen	0.00034

Po združevanju povezanih podatkov nastane 171 učnih parov in po predstavitvi le-teh umetni nevronske mreže (geometrije 3-20-15-2, maksimalno število iteracij je 50.000, 0.1 velikostjo učnega koraka) ostane nenaučenih 39 novih učnih parov.

Spremembe statističnih kazalcev pri učenju z združitvijo so prikazane v preglednici 19 (normirano povprečno absolutno odstopanje razlik trdnosti je pred združevanjem približno 25%, po združevanju okoli 20%; normirano povprečno absolutno odstopanje razlik lokalnega elastičnega modula je pred združevanjem približno 20%, po združevanju okoli 15%). V primerjavi z učenjem pred združevanjem učnih parov sta povprečno absolutno odstopanje razlik in standardni odklon nekoliko manjša, koeficient korelacije se je pri trdnosti malo poslabšal, pri modulu elastičnosti pa izboljšal. To pomeni, da je bilo učenje nekoliko uspešnejše kot prvotno brez združevanja učnih parov. Zato preurejanja testnih podatkov nismo izvedli.

V tem primeru je bilo združevanje podatkov uspešno, vseeno pa posploševanje mreže na testne podatke ni zelo dobro.

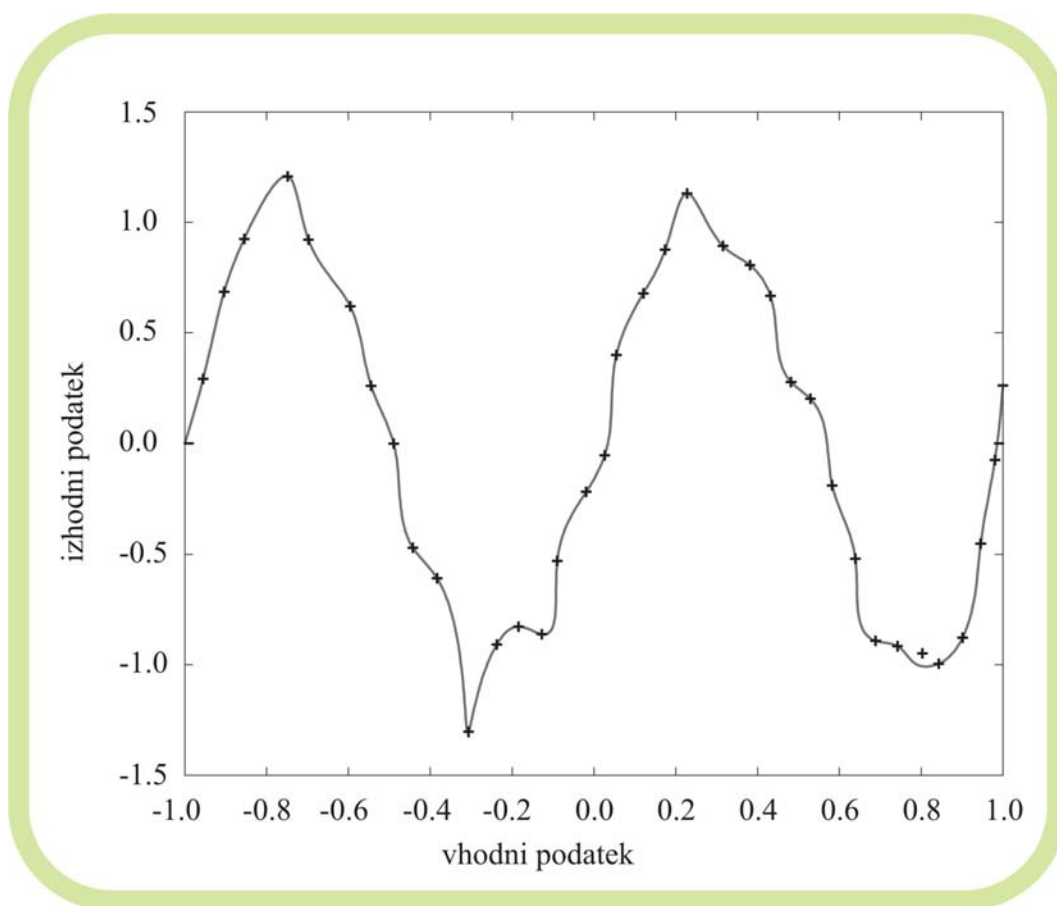
Preglednica 19: Vrednosti statističnih kazalcev.

	<i>f</i>			<i>E_{e-l}</i>		
	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
pred združevanjem	10.766	13.683	0.279	2314.27	2938.3	0.32
po združevanju	10.333	13.656	0.239	2024.66	2680.69	0.41

Vhodno-izhodni učni in testni pari podatkov pred združevanjem ter izračun norm in novi učni pari po združevanju učnih podatkov so prikazani v prilogi E.

4.4.6 Rezultati

Opozorili smo že, da je za podatke značilen velik razstros, saj so to lastnosti lesa, dobljene na preizkušancih, ki so se nahajali na različnih delih drevesa, na drevesih iz različnih rastišč ali celo na drevesih različnih vrst. V odvisnosti od vhodnih podatkov so se umetne nevronske mreže naučile različno število učnih parov. Najbolj izstopa prvi primer, kjer se je več različnih mrež naučilo vse učne nize podatkov že pred združevanjem parov. Če pogledamo statistične kazalce, ki kažejo uspešnost učenja na testnih parih, vidimo, da učenje ni bilo preveč učinkovito. To kaže tudi na problem prenaučeniosti nevronske mreže. Ta problem se pojavlja pri vseh petih primerih, vendar je pri prvem najbolj izrazit. Na sliki 23 je shematično prikazano, kako se lahko mreže točno naučijo vse učne podatke. Točke označujejo dejanske izhodne vrednosti, polna črta ponazarja učenje umetne mreže. Iz poznejšega testiranja se je izkazalo, da bi bilo problem prenaučeniosti smiselno rešiti z upoštevanjem manjšega števila maksimalnih iteracij.



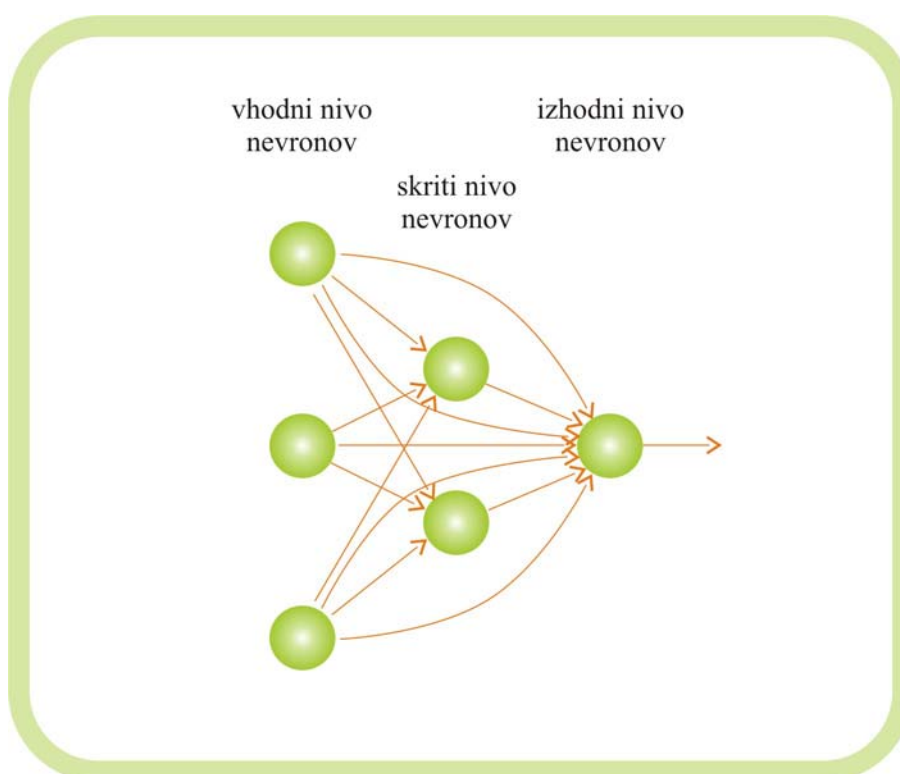
Slika 23: Prikaz prenaučeni nevrnske mreže.

Združevanje podatkov vpliva predvsem na število nenaučenih učnih parov, saj je njihov delež pri vseh petih primerih po združevanju podatkov manjši. Statistični kazalci se ne spremenijo drastično. V splošnem lahko iz prikazanih rezultatov zaključimo, da je z združevanjem učnih parov mogoče naučiti umetno nevronske mrežo nekoliko več učnih parov, vendar rezultati na testnih podatkih niso bistveno boljši, v nekaterih primerih so celo slabši. Vseeno je potrebno pri prikazu statističnih kazalcev, ki kažejo na uspešnost učenja poudariti, da so bili izračunani na testnih podatkih, za katere je tako kot za učne značilen velik raztros.

4.5 Knjižnica umetnih nevronske mreže v programu Matlab

Za učenje in testiranje umetnih nevronske mreže smo preizkusili tudi knjižnico nevronske mreže v programskem paketu Matlab. Zaradi primerjave s programom NTR2003 je izbrani proces učenja mreže v programu Matlab adaptiven in temelji na posplošenem pravilu delta, kombiniranim z momentno metodo pri računu popravkov uteži. To pomeni, da program obdela vsak vhodno-izhodni učni par posebej in se ga skuša naučiti na podlagi razlike med dejansko in izračunano vrednostjo izhodne spremenljivke.

Zaradi slabšega obnašanja večnivojskih usmerjenih umetnih nevronske mreže, smo za račun v programu Matlab uporabili umetne nevronske mreže, kjer je vsak nivo povezan z vsemi naslednjimi (slika 24).



Slika 24: Shema umetne nevronske mreže pri kateri je vsak nivo nevronov povezan z vsemi naslednjimi.

Pripravili smo lasten program, ki za učne in kontrolne podatke pripravi mrežo in ustrezno ovrednoti rezultate. Uporabili smo vgrajeno mrežo *newcf* z adaptivnim učenjem (funkcija *adapt*).

Za zagon programa je potrebno pripraviti tri matlabove datoteke s končnico *.m*:

- datoteka s splošnimi podatki,
- datoteka z učnimi podatki,
- datoteka s testnimi podatki.

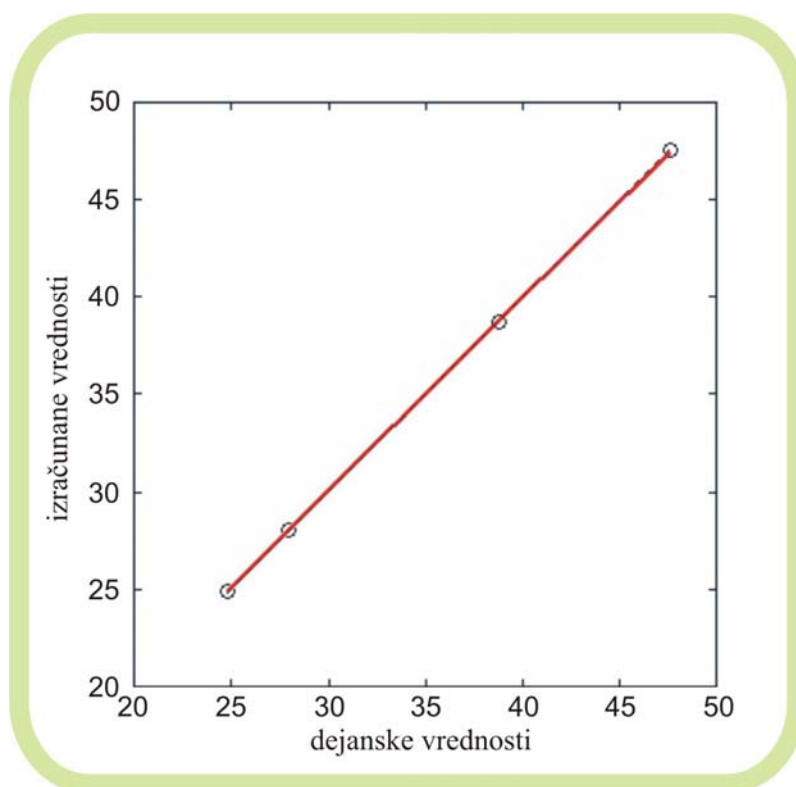
V matlabovi datoteki s splošnimi podatki so zapisani ukazi, ki omogočijo učenje na učnih parih in testiranje na testnih parih ter parametri, ki določajo posamezen tip mreže:

- velikost dovoljene napake,
- maksimalno število iteracij,
- izhodna funkcija na posameznem nivoju,
- geometrija umetne nevronske mreže:
 - število nevronov vhodnega nivoja,
 - število nevronov izhodnega nivoja,
 - število skritih nivojev,
 - število nevronov na posameznem skritem nivoju.

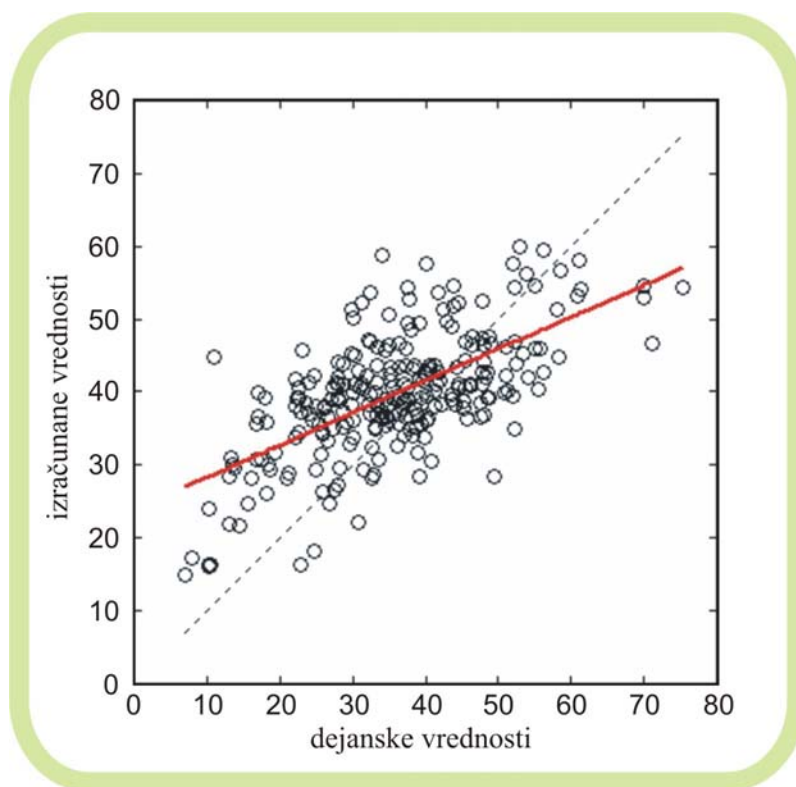
Po procesiranju posamezne mreže program kot rezultat poda:

- št. nenaučenih učnih vhodno-izhodnih parov,
- povprečno absolutno odstopanje razlik,
- standardni odklon,
- koeficient korelacije.

Testiranje umetnih nevronske mreže v Matlabu je pokazalo, da so se različni tipi mrež sposobni zelo dobro naučiti vse učne vhodno-izhodne pare, če je le teh zelo malo (na grafikonu 4 so rezultati učenja štirih vhodno-izhodnih parov). Če pa smo mrežo učili na velikem številu vhodno-izhodnih parov, je bilo učenje manj uspešno (grafikon 5). Vzrok za slabšo uspešnost učenja pri večjem številu vhodno-izhodnih parov je v velikem raztrosu uporabljenih podatkov (grafikoni 1-3).



Grafikon 4: Učenje s štirimi vhodno-izhodnimi učnimi pari.



Grafikon 5: Učenje z dvestopdesetimi vhodno-izhodnimi učnimi pari.

Velikost dovoljene napake je enaka kot pri uporabi programa NTR2003, maksimalno število iteracij je pri vseh mrežah 100. Za tako maksimalno število iteracij smo se odločili na podlagi večkratnega izračuna istega tipa mreže. Ker program na začetku naključno izbere uteži mreže, se rezultati učenja ustalijo šele po zadostnem maksimalnem številu iteracij. Mreže, dobljene po 100 korakih adaptivnega učenja, pa so se le minimalno razlikovale, ne glede na različne začetne vrednosti uteži. Opozorimo, da je število iteracij učenja bistveno manjše kot pri programu NTR2003, kjer smo uporabili priporočila drugih avtorjev.

Za učenje je bilo ravno tako kot pri programu NTR2003 uporabljenih 250 učnih vhodno-izhodnih parov, za testiranje pa preostalih 43 vhodno-izhodnih testnih parov. Zopet je obravnavanih 5 različnih primerov, ki se razlikujejo glede na uporabljene vhodne spremenljivke (preglednica 20 do preglednica 29).

Uporabljali smo različne geometrije mrež in različne izhodne funkcije, ki so prikazane na slikah (12 – 14).

4.5.1 Primer 1

V primeru 1 kot vhodni podatki nastopajo: ρ , E_{e-l} , E_{e-g} , E_{f-l} , E_{f-g} , E_{Tno} , E_{Zag} , E_{Sylv} , kot izhodni podatek pa upogibna trdnost lesa f (poglavje 4.4.1). Z različnimi mrežami smo obravnavali 250 učnih parov in 43 testnih parov.

Preglednica 20: Uporabljeni parametri pri učenju in statistični kazalci.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koefficient korelacije
8-5-1	logsig	126	8.558	11.04	0.55
8-8-8-1	logsig	130	9.581	12.21	0.269
8-4-4-1	logsig	129	9.199	11.88	0.29
8-5-5-5-1	logsig	122	8.627	11.127	0.46
8-20-15-15-1	logsig	123	8.458	10.941	0.56
8-5-1	tansig	61	7.036	9.143	0.665

se nadaljuje ...

... nadaljevanje

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
8-8-8-1	tansig	63	7.067	9.040	0.67
8-4-4-1	tansig	64	6.654	8.739	0.699
8-5-5-5-1	tansig	62	7.143	6.698	0.713
8-20-15-15-1	tansig	61	7.451	9.439	0.66
8-5-1	purelin	73	7.030	8.953	0.686
8-4-5-4-1	purelin	68	6.942	8.913	0.690
8-4-4-1	purelin	67	6.961	8.912	0.689
8-3-5-1	purelin	67	6.98	8.926	0.689
8-2-2-1	purelin	65	6.948	8.892	0.689

Umetne nevronske mreže izračunane v Matlabu so uspešnejše pri učenju kot mreže izračunane s programom NTR2003, kljub temu, da so se naučile manj učnih parov. Ker se uspešnost učenja meri s statističnimi kazalci na testnih parih in ne na osnovi števila nenaučenih učnih parov, ta rezultat pomeni, da so se mreže v programu NTR2003 "prenaučile" glede na učne pare. Zato pri kontroli uspešnosti učenja na testnih parih, niso bile sposobne posplošiti svojega znanja.

Zanimivo je, da so v Matlabu uspešne skoraj vse mreže, slabše se obnašajo le mreže z izhodnimi funkcijami *logsig*. Te se za obravnavani primer izkažejo kot neprimerne.

Preglednica 21: Uporabljeni parametri pri učenju in statistični kazalci pri večjem številu iteracij.

geometrija mreže	funkcije med skritimi nivoji	maks. št. iteracij	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standard. odklon	koeficient korelacije
8-5-5-5-1	tansig	1000	25	8.317	10.782	0.562
8-5-5-5-1	tansig	5000	0	13.186	17.471	0.317

Iz preglednice 21 je razvidno, da se z večanjem števila maksimalnih iteracij umetna nevronska mreža nauči več učnih parov. Pri posploševanju na testnih parih je manj uspešna.

4.5.2 Primer 2

V primeru 2 so kot vhodni podatki obravnavani: ρ , E_{f-g} , E_{Tno} , izhodni podatek oz. rezultat je trdnost f . Tudi v tem primeru smo z različnimi mrežami obravnavali 250 učnih parov in 43 testnih parov. Nato smo na istih mrežah uporabili še učne podatke, ki smo jih združili na osnovi minimalne norme učnih parov (glej poglavje 4.4.2).

Preglednica 22: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-5-1	logsig	130	9.602	12.225	-0.423
3-8-8-1	logsig	130	9.599	12.221	0.303
3-4-4-1	logsig	130	9.570	12.182	0.251
3-5-5-5-1	logsig	129	9.599	12.221	-0.399
3-20-15-15-1	logsig	128	9.381	11.827	0.255
3-5-1	tansig	107	8.514	11.324	0.393
3-8-8-1	tansig	97	8.788	11.368	0.391
3-4-4-1	tansig	107	8.538	11.236	0.406
3-5-5-5-1	tansig	105	8.675	11.292	0.397
3-20-15-15-1	tansig	109	8.801	11.434	0.433
3-5-1	purelin	105	8.539	11.26	0.397
3-4-5-4-1	purelin	104	8.493	11.296	0.398
3-4-4-1	purelin	105	8.533	11.2597	0.399
3-3-5-1	purelin	101	8.533	11.258	0.399
3-2-2-1	purelin	102	8.519	11.272	0.397

Preglednica 23: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-5-1	logsig	101	9.519	12.141	0.206
3-8-8-1	logsig	104	9.601	12.216	0.162

se nadaljuje ...

... nadaljevanje

geometrija mreže	izhodne funkcije na skritinivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-4-4-1	logsig	101	9.378	11.985	0.303
3-5-5-5-1	logsig	104	9.598	12.219	-0.120
3-20-15-15-1	logsig	102	9.568	12.184	0.307
3-5-1	tansig	85	8.753	11.3923	0.375
3-8-8-1	tansig	84	8.776	11.358	0.385
3-4-4-1	tansig	83	8.899	11.412	0.368
3-5-5-5-1	tansig	88	8.877	11.453	0.364
3-20-15-15-1	tansig	73	8.717	11.511	0.391
3-5-1	purelin	80	8.613	11.277	0.390
3-4-5-4-1	purelin	78	8.629	11.254	0.393
3-4-4-1	purelin	75	8.634	11.235	0.395
3-3-5-1	purelin	79	8.733	11.266	0.388
3-2-2-1	purelin	76	8.613	11.276	0.39

Ravno tako kot v primeru 1, so se tudi v tem primeru, umetne nevronske mreže v Matlabu uspešneje naučile učne pare kot mreže v programu NTR2003, saj so vsi trije statistični kazalci boljši. Združevanje učnih parov je sicer zmanjšalo število nenaučenih parov, a na uspešnost učenja ni vplivalo. Ponovno ugotovimo, da so izhodne funkcije *logsig* za naš problem neustrezne, saj dobimo celo negativne koeficiente korelacije. Veliko število nenaučenih parov je odraz raztrosa podatkov in ilustrira dejansko stanje ter kaže na sposobnost posploševanja.

Preglednica 24: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.

geometrija mreže	funkcije med skritimi nivoji	maks. št. iteracij	št. nenaučenih učnih parov	povpreč. absolutno odstop. razlik	standardni odklon	koeficient korelacije
3-2-2-1	purelin	1000	98	8.540	11.268	0.396
3-2-2-1	purelin	5000	96	8.542	11.267	0.396

Iz preglednice 24 je razvidno, da se z večanjem števila maksimalnih iteracij umetna nevronska mreža nauči več učnih parov. Pri posploševanju na testnih parih ni opaziti večjega poslabšanja.

4.5.3 Primer 3

V primeru 3 nastopajo kot vhodni podatki: ρ , E_{f-g} , E_{Zag} , izhodni podatek je zopet upogibna trdnost lesa f . Na enak način kot v primeru 2 so se v tem primeru mreže učile najprej na nezdruženih in nato na združenih učnih parih (združevanje parov je opisano v poglavju 4.4.3).

Preglednica 25: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-5-1	logsig	129	9.601	12.224	-0.269
3-8-8-1	logsig	129	9.598	12.221	-0.302
3-4-4-1	logsig	129	9.409	11.938	0.255
3-5-5-5-1	logsig	129	9.597	12.218	0.023
3-20-15-15-1	logsig	129	9.655	12.161	0.151
3-5-1	tansig	104	8.569	11.383	0.387
3-8-8-1	tansig	105	8.715	11.837	0.344
3-4-4-1	tansig	100	8.772	11.686	0.358
3-5-5-5-1	tansig	105	8.726	11.739	0.362
3-20-15-15-1	tansig	106	9.198	12.192	0.344
3-5-1	purelin	103	8.692	11.691	0.351
3-4-5-4-1	purelin	105	8.828	11.849	0.337
3-4-4-1	purelin	106	8.700	11.741	0.350
3-3-5-1	purelin	104	8.814	11.879	0.338
3-2-2-1	purelin	104	8.6698	11.694	0.353

Preglednica 26: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-5-1	logsig	103	9.602	12.225	-0.259
3-8-8-1	logsig	102	9.598	12.221	-0.234
3-4-4-1	logsig	97	9.317	11.835	0.251
3-5-5-5-1	logsig	103	9.595	12.215	0.074

se nadaljuje ...

... nadaljevanje

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koefficient korelacije
3-20-15-15-1	logsig	99	9.586	12.174	0.178
3-5-1	tansig	67	8.609	11.552	0.371
3-8-8-1	tansig	65	8.651	11.784	0.356
3-4-4-1	tansig	61	8.846	11.759	0.353
3-5-5-5-1	tansig	68	8.768	11.837	0.355
3-20-15-15-1	tansig	69	8.587	11.322	0.402
3-5-1	purelin	66	8.683	11.649	0.351
3-4-5-4-1	purelin	68	8.647	11.691	0.353
3-4-4-1	purelin	69	8.665	11.730	0.350
3-3-5-1	purelin	68	8.654	11.649	0.353
3-2-2-1	purelin	69	8.673	11.643	0.352

Tudi v tem primeru so se nevronske mreže v Matlabu naučile manj učnih parov kot mreže v programu NTR2003, vendar je bilo njihovo učenje glede na testne podatke uspešnejše. Ponovno se funkcija *logsig* izkaže za neprimerno. Združevanje učnih parov na uspešnost učenja ni vplivalo, zmanjšalo pa je število nenaučenih parov.

Preglednica 27: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.

geometrija mreže	funkcije med skritimi nivoji	maks. št. iteracij	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koefficient korelacije
3-4-4-1	tansig	1000	95	8.724	11.630	0.367
3-4-4-1	tansig	5000	91	8.850	11.645	0.393

Iz preglednice 27 je razvidno, da se z večanjem števila maksimalnih iteracij umetna nevronska mreža nauči več učnih parov. Pri posploševanju na testnih parih ni opaziti večjega poslabšanja.

4.5.4 Primer 4

V tem primeru so vhodni podatki ρ , E_{f-g} , E_{Sylv} , izhodna vrednost je zopet trdnost lesa f . Ponovno je obravnavanih več nevronske mreže, ki se učijo najprej na nezdruženih podatkih in nato na združenih (združevanje učnih parov je prikazano v poglavju 4.4.4).

Preglednica 28: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-5-1	logsig	129	9.603	12.225	-0.3799
3-8-8-1	logsig	129	9.599	12.221	-0.382
3-4-4-1	logsig	131	9.347	11.953	0.301
3-5-5-5-1	logsig	129	9.595	12.216	0.131
3-20-15-15-1	logsig	131	9.276	11.893	0.230
3-5-1	tansig	110	8.60	11.315	0.391
3-8-8-1	tansig	109	8.837	11.683	0.356
3-4-4-1	tansig	109	8.505	11.284	0.402
3-5-5-5-1	tansig	105	8.6297	11.423	0.391
3-20-15-15-1	tansig	97	9.0495	11.895	0.405
3-5-1	purelin	108	8.453	11.197	0.411
3-4-5-4-1	purelin	111	8.472	11.201	0.412
3-4-4-1	purelin	109	8.465	11.186	0.412
3-3-5-1	purelin	111	8.448	11.199	0.412
3-2-2-1	purelin	106	8.450	11.2096	0.412

Preglednica 29: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-5-1	logsig	87	9.372	11.961	0.265
3-8-8-1	logsig	88	9.395	12.062	0.225
3-4-4-1	logsig	87	9.588	12.210	0.205
3-5-5-5-1	logsig	85	9.172	11.735	0.291

se nadaljuje ...

... nadaljevanje

geometrija mreže	izhodne funkcije na skritih nivojih	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-20-15-15-1	logsig	85	9.284	11.905	0.228
3-5-1	tansig	65	8.512	11.3296	0.390
3-8-8-1	tansig	67	8.671	11.497	0.3696
3-4-4-1	tansig	59	8.742	11.597	0.364
3-5-5-5-1	tansig	64	8.395	11.194	0.404
3-20-15-15-1	tansig	75	9.0697	11.784	0.347
3-5-1	purelin	70	8.426	11.222	0.411
3-4-5-4-1	purelin	68	8.439	11.185	0.412
3-4-4-1	purelin	70	8.458	11.162	0.411
3-3-5-1	purelin	70	8.428	11.1995	0.4095
3-2-2-1	purelin	68	8.434	11.210	0.411

Ponovno so se nevronske mreže v Matlabu naučile manj učnih parov kot mreže v programu NTR2003, vendar so se izkazale za uspešnejše na testnih podatkih. Za najprimernejšo izhodno funkcijo se izkaže funkcija *purelin*. Združevanje učnih parov na uspešnost učenja ni vplivalo.

Preglednica 30: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.

geometrija mreže	funkcije med skritimi nivoji	maks. št. iteracij	št. nenaučenih učnih parov	povprečno absolutno odstopanje razlik	standardni odklon	koeficient korelacije
3-2-2-1	purelin	1000	107	8.454	11.202	0.4097
3-2-2-1	purelin	5000	106	8.454	11.203	0.4096

Iz preglednice 30 je razvidno, da se z večanjem števila maksimalnih iteracij umetna nevronska mreža nauči več učnih parov. Pri posploševanju na testnih parih ni opaziti večjega poslabšanja.

4.5.5 Primer 5

Zadnji primer obravnava kot vhodne podatke ρ , E_{f-g} , E_{Tno} , izračunati pa želimo upogibno trdnost lesa f in E_{e-l} . Tudi tu smo mreže učili na nezdruženih in na združenih učnih parih (glej poglavje 4.4.5).

Preglednica 31: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov.

geomet. mreže	izhodne funkc. na skrit. nivojih	št. nenauč. učnih parov	f			E_{e-l}		
			povpr. abs. odstop. razlik	standard. odklon	koefic. korelac.	povpr. abs. odstop. razlik	standard. odklon	koefic. korelac.
3-5-2	logsig	238	9.5597	12.188	0.298	1884.2	2545.4	-0.506
3-8-8-2	logsig	237	9.589	12.207	0.286	1818.9	2475.6	0.247
3-4-4-2	logsig	238	9.60	12.221	-0.038	1880.6	2542.2	-0.137
3-5-5-5-2	logsig	238	9.363	11.929	0.272	1879.2	2539.6	0.231
3-20-15-15-2	logsig	240	9.598	12.217	0.075	1826.1	2455.8	0.257
3-5-2	tansig	209	8.840	11.399	0.373	1719.5	2284.4	0.488
3-8-8-2	tansig	217	8.753	11.398	0.381	1777.1	2367.5	0.442
3-4-4-2	tansig	216	8.662	11.238	0.401	1709.5	2272.4	0.484
3-5-5-5-2	tansig	217	8.7697	11.3996	0.377	1691.6	2253.9	0.496
3-20-15-15-2	tansig	217	8.720	11.288	0.412	1717.8	2266.9	0.506
3-5-2	purelin	224	8.555	11.252	0.398	1659.1	2246.8	0.501
3-4-5-4-2	purelin	223	8.556	11.256	0.396	1662.1	2249.5	0.50
3-4-4-2	purelin	222	8.534	11.270	0.397	1671.5	2256.1	0.501
3-3-5-2	purelin	225	8.555	11.257	0.396	1654.2	2243.3	0.501
3-2-2-2	purelin	221	8.551	11.249	0.398	1653.6	2242.8	0.502

Preglednica 32: Uporabljeni parametri pri učenju in statistični kazalci po združitvi učnih parov.

geomet. mreže	izhodne funkc. na skrit. nivojih	št. nenauč. učnih parov	f			E_{e-l}		
			povpr. abs. odstop. razlik	standard. odklon	koefic. korelac.	povpr. abs. odstop. razlik	standard. odklon	koefic. korelac.
3-5-2	logsig	149	9.621	12.245	-0.421	1729.5	2412.6	0.325
3-8-8-2	logsig	148	9.610	12.234	-0.396	1842.8	2547.7	0.2795
3-4-4-2	logsig	150	9.1199	11.681	0.338	1741.2	2423.1	0.348
3-5-5-5-2	logsig	146	9.572	12.191	0.302	1740.1	2406.2	0.337
3-20-15-15-2	logsig	144	9.578	12.185	0.235	1658.4	2336.9	0.395
3-5-2	tansig	150	8.527	11.211	0.403	1658.9	2222.4	0.501
3-8-8-2	tansig	154	8.946	11.469	0.392	1664.5	2240.4	0.501
3-4-4-2	tansig	152	8.773	11.428	0.376	1607.1	2171.4	0.527
3-5-5-5-2	tansig	151	9.172	11.737	0.352	1603.4	2177.9	0.525
3-20-15-15-2	tansig	153	9.312	11.804	0.366	1665.7	2226.4	0.518
3-5-2	purelin	151	8.505	11.276	0.397	1630.7	2226.9	0.504
3-4-5-4-2	purelin	147	8.472	11.296	0.397	1672.3	2261.1	0.504
3-4-4-2	purelin	149	8.506	11.278	0.3998	1638.9	2237.1	0.501
3-3-5-2	purelin	146	8.4596	11.353	0.398	1695.1	2279.7	0.502
3-2-2-2	purelin	149	8.505	11.272	0.397	1626.7	2222.5	0.504

Rezultati zopet kažejo na večjo uspešnost učenja nevronske mreže v Matlabu. V tem primeru je število nenaučenih učnih parov še nekoliko večje kot pri ostalih primerih. Verjetno je temu vzrok, da se morajo mreže naučiti izraziti dva izhodna podatka. Združevanje podatkov ne vpliva na statistične kazalce uspešnosti učenja. Zelo pomembna je izbira izhodne funkcije. *Tansig* in *purelin* sta primerni, *logsig* je povsem neprimerna.

Preglednica 33: Uporabljeni parametri pri učenju in statistični kazalci pred združitvijo učnih parov pri večjem številu iteracij.

geomet. mreže	funkc. med skrit. nivoji	maks. št. iteracij	št. nenauč. učnih parov	f			E_{e-l}		
				povpr. abs. odstop. razlik	standard. odklon	koefic. korelac.	povpr. abs. odstop. razlik	standard. odklon	koefic. korelac.
3-4-4-2	tansig	1000	123	8.756	11.395	0.383	1751.6	2320.6	0.463
3-4-4-2	tansig	5000	97	9.758	12.918	0.303	2032	2556.1	0.398

Iz preglednice 33 je razvidno, da se z večanjem števila maksimalnih iteracij umetna nevronska mreža nauči več učnih parov. Pri posploševanju na testnih parih je opaziti manjše poslabšanje.

4.5.6 Rezultati

V splošnem geometrija mreže pri računu v Matlabu nima izstopajočega vpliva. Pri izhodnih funkcijah na skritih nivojih se je izkazalo, da ima na število nenaučenih učnih parov najmanjši vpliv funkcija *logsig* in ravno tako sta pri vseh treh statističnih kazalcih dali boljše rezultate funkciji *tansig* in *purelin*. Koeficient korelacije je v nekaterih primerih pri funkciji *logsig* celo negativen, kar pomeni, da se mreža ni pravilno naučila učnih parov. V primerjavi s programom NTR2003 so se umetne nevronske mreže v programu Matlab naučile manj učnih parov, vendar so bile pri kontroli njihovega "znanja" na testnih podatkih uspešnejše. To pomeni, da štetje nenaučenih podatkov ni najboljšo merilo za učinkovitost učenja pri tolikšnem raztrosu podatkov. Mreže v Matlabu so sposobne svoje "znanje" bolje posplošiti kot mreže v programu NTR2003, kjer je prihajalo do prenaučnosti. Po združevanju učnih parov se je nekoliko zmanjšal delež nenaučenih učnih parov. Na statistične kazalce, ki pokažejo uspešnost učenja, pa združevanje nima posebnega vpliva. Torej so mreže računane s programom Matlab neodvisne od podatkov, ki so združeni v smislu minimalne medsebojne norme. To seveda ne pomeni, da drugačne metode združevanja ne bi omogočile boljšega učenja, vendar je to področje nadaljnjih raziskav. Glede na velik raztros podatkov smo lahko z dobljenimi rezultati zadovoljni.

5 ZAKLJUČEK

Osnovni namen diplomske naloge je bil ugotoviti uspešnost učenja umetnih nevronske mreže ob zelo razpršenih podatkih, pri tem pa ugotoviti vpliv združevanja učnih vhodno-izhodnih podatkov. Za učenje nevronske mreže smo uporabili program NTR2003 in knjižnico nevronske mreže v programu Matlab.

Iz rezultatov programa NTR2003 je v splošnem razvidno, da se te mreže naučijo veliko učnih parov, vendar so pri posploševanju znanja na testne pare slabe. Združevanje podatkov, v smislu minimalne norme učnih parov, najbolj vpliva na delež nenaučenih učnih parov, saj se ta po združitvi parov zmanjša. Tudi uspešnost učenja se za malenkost izboljša, razen pri primeru 4, kjer niti vključitev določenega testnega para med učne ni izboljšala učenja mreže. Torej ima združevanje podatkov v smislu minimalnih norm le manjši vpliv na uspešnost učenja mreže. Zaradi ugotovitve o prenaučeni mreži, smo program NTR2003 preizkusili še z manjšim številom iteracij (5.000, 1.000 in 100 namesto 10.000). Rezultati so pokazali, da bi se taka mreža obnašala boljše, saj ni prišlo do prenaučeni. Naučila se je manj učnih parov, vendar je bila pri posplošitvi znanja na testne pare boljše.

Iz rezultatov, dobljenih s programom Matlab, je razvidno, da imajo na učenje mreže velik vpliv izhodne funkcije na skritih nivojih. Združevanje podatkov v splošnem zmanjša delež nenaučenih učnih parov, uspešnost učenja je od združevanja učnih parov neodvisna.

Zaključimo lahko, da sta pri posploševanju na testne pare oba programa relativno slabo uspešna. Razlog je v veliki razpršenosti podatkov. Združevanje zmanjša razpršenost in s tem število nenaučenih parov, vendar seveda uspešnost mreže na testnih podatkih ne more biti bistveno boljše. Paziti moramo tudi na problem prenaučeni, kjer sicer naučimo mrežo učnih zvez do zahtevane natančnosti, testni podatki, pa so le slabo opisani z utežmi take mreže.

Opozoriti je potrebno, da je možno v programu Matlab uporabiti tudi drugačne tipe mreže, ki imajo vgrajene določene "varovalke" za preprečevanje prenaučeni. Ena takih "varovalk" je predčasna ustavitev učenja mreže (angl. early stopping), kjer program določen del učnih

podatkov uporabi za učenje, preostale pa za kontrolo. Ko prične napaka kontrolnih učnih podatkov naraščati, se učenje mreže ustavi in s tem se prepreči prenaučenost mreže. Take "varovalke" so vgrajene v mrežah, ki uporabljajo za učenje funkcijo *train*. To je način učenja, kjer se mreža skuša učiti učne podatke kot celoto in ne vsakega posebej, kot to zahteva funkcija *adapt*.

Osnovni namen diplomske naloge je dosežen. Vendar je potrebno poudariti, da na uspešnost učenja vpliva še veliko dejavnikov, ki presegajo obseg tega dela. Uspešnost učenja je merjena na naključno določenih testnih podatkih, ki so ravno tako kot učni, izpostavljeni velikemu raztrosu. Za še boljše rezultate bi bilo smiselno za učenje preizkusiti večje število vhodno-izhodnih parov in preizkusiti druge metode združevanja.

6 VIRI

Ambrožič, T., 2001. Aplikacija umetnih nevronske mreže v napovedovanju ugrezanja zaradi podzemnega rudarjenja. Doktorska disertacija. Ljubljana, Univerza v Ljubljani, FGG: 109 f.

Bronštejn, I. N., Semendjajev, K. A., Musiol, G., Mühlig, H., 1997. Matematični priročnik. Ljubljana, Tehniška založba Slovenije: 967 f.

Dobnikar, A., 1990. Nevronske mreže: Teorija in aplikacije. Radovljica, Didakta: 157 f.

Haykin S., 1994. Neural Networks: A Comprehensive Foundation. New York, Macmillan College publishing Company, Inc.: 216 f.

Kartam, N., Flood, I., Garrett, J. H., 1997. Artificial neural networks for civil engineers: fundamentals and applications. New York, American society of civil engineers: 216 f.

Plos, M., 2006. Nedestruktivne metode za razvrščanje gradbenega lesa po trdnosti. Diplomski naloga. Ljubljana, Univerza v Ljubljani, FGG: 59 f.

Demuth, H., Beale, M., Hagan, M., 2005. Neural network toolbox 5. User's guide. Natick, Mathworks.

Reščič, D., 2004. Strojno razvrščanje lesa z uporabo nevronske mreže. Diplomski naloga. Ljubljana, Univerza v Ljubljani, FGG: 63 f.

Rumelhart, D. E., McClelland, J. L., the PDP Research Group, 1986. Parallel Distributed Processing, Volume 1: Foundations. Cambridge, The MIT Press: 501 f.

Topping, B. H. V., Bahreininejad, A., 1997. Neural computing for structural mechanics. Edinburg, Saxe-Coburg publications: 176 f.

Turk, G., Logar, J., 1996. Uporaba umetne inteligence v gradbeni konstruktivi. Gradbeni vestnik 35, 5-6-7: 147-153.

Turk, G., 2003. Program NTR2003. Ljubljana.

Zupan, D., 2007. Programski jezik Matlab, Kratek tečaj. Ljubljana, Univerza v Ljubljani, FGG.

<ftp://ftp.sas.com/pub/neural/faq.html#questions> (19.02.2008)

7 PRILOGE

- Priloga A: Primer 1: učni in testni podatki.
- Priloga B: Primer 2: učni in testni podatki pred združevanjem učnih parov, izračun norm, učni in testni podatki po združitvi učnih parov.
- Priloga C: Primer 3: učni in testni podatki pred združevanjem učnih parov, izračun norm, učni in testni podatki po združitvi učnih parov.
- Priloga D: Primer 4: učni in testni podatki pred združevanjem učnih parov, izračun norm, učni in testni podatki po združitvi učnih parov.
- Priloga E: Primer 5: učni in testni podatki pred združevanjem učnih parov, izračun norm, učni in testni podatki po združitvi učnih parov.
- Priloga F: Program, ki smo ga zasnovali na osnovi knjižnice nevronske mreže v programskem paketu Matlab.

Vse priloge so na CD-ju.