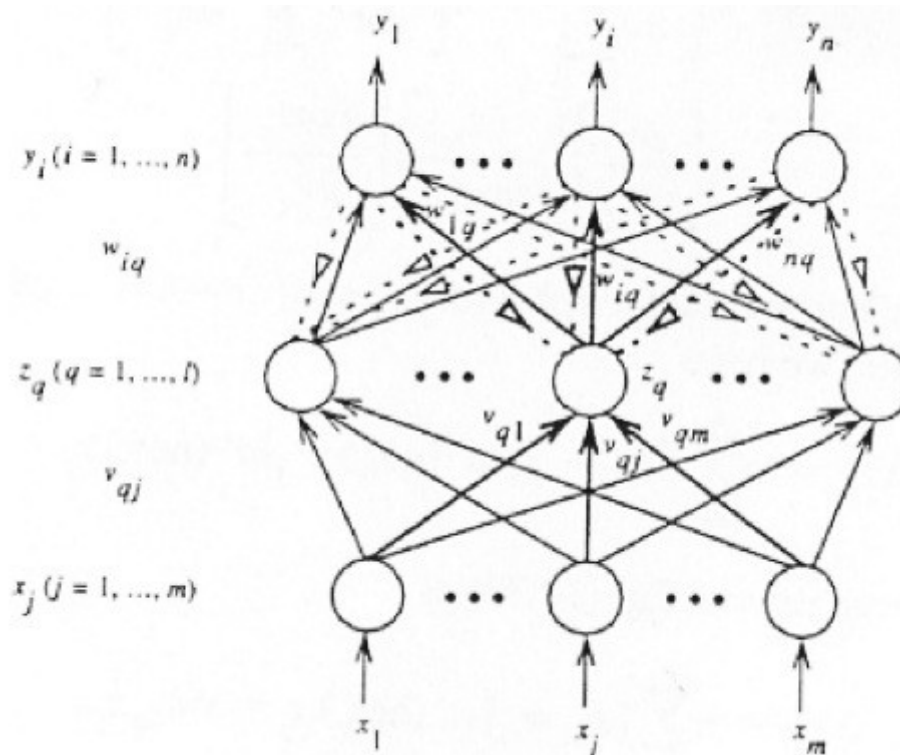


# 1. Nacrtati i objasniti algoritam propagacije greške unazad (back propagation).



Backpropagation je algoritam ucenja primenjen kod viseslojnih feedforward mreza ciji neuroni imaju kontinualnu diferencijabilnu aktivacionu funkciju. Prvo se zadaju ulazni parametri i ocekivani izlazi. Racunanje se obavlja u dva koraka: prvo se racunaju izlazi svakog sloja do zadnjeg, izlaznog sloja, a zatim se generise signal greske koji se salje od izlaznog sloja ka prethodnim slojevima i popravljaju se tezine. Tako se greska salje za sve slojeve unazad po jedan sloj i racuna, sve do ulaznog sloja i popravljaju se tezine.

# 2. Definisati pojam kresanja neuralne mreže, i kada se koristi.

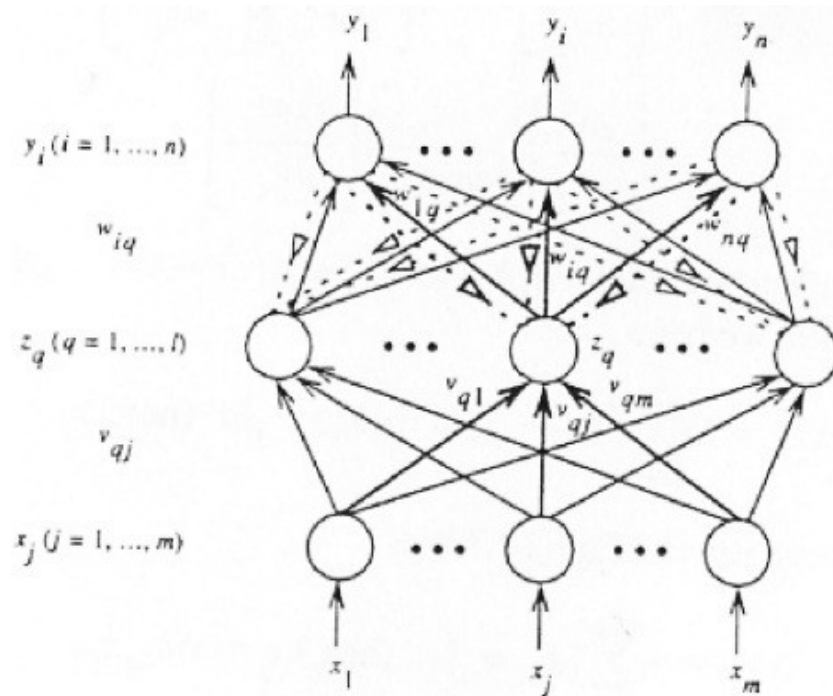
Kresanje je ukidanje određenog broja neurona i konekcija i da pri tome redukovana mreza radi ispravno. Koristi se kada je potrebno da se smanji kompleksnost same mreze (povećanje performansi) i da se poboljša generalizacija. Može se početi obucavanje mreze koja je prevelika za dati problem, to se uvidi i krene se sa uklanjanjem neurona (veza) da bi se svela na dovoljnu velicinu.

# 3. Koja dva tipa implementacije algoritma back propagation postoje, i koje su razlike.

1) Inkrementalni režim rada – apdejtujemo pojačanja posle svakog odbirka (izračunavamo apdejt i apdejtujemo pojačanja za svaki uzorak)

2) Batch mod(p.p.) – primenimo za sve uzorke, pa onda izračunamo  $\Delta$  nad svim uzorcima pa tek onda apdejtujemo pojačanja

4. **Nacrtati i objasniti algoritam propagacije greške unazad (back propagation), batch tipa.**



Postupak je isti kao za inkrementalni mod u pod 1, samo što se ovde prvo primene svi ulazi, izracunaju se izlazi mreze i tek onda se radi racunanje greske. Dakle gore smo imali situaciju da dodje jedan ulaz, izracuna se do izlaza i onda se racuna greska i salje unazad, pa se tako nastavlja dalje. Ovde se greska ne racuna posle svakog ulaza, nego prvo prodju svi ulazi i njihove greske sa saberu i samo jednom se updatuju tezine.

5. **Definisati preobučavanje, kako nastaje i koje su metode borbe protiv preobučavanja i ukratko ih objasniti.**

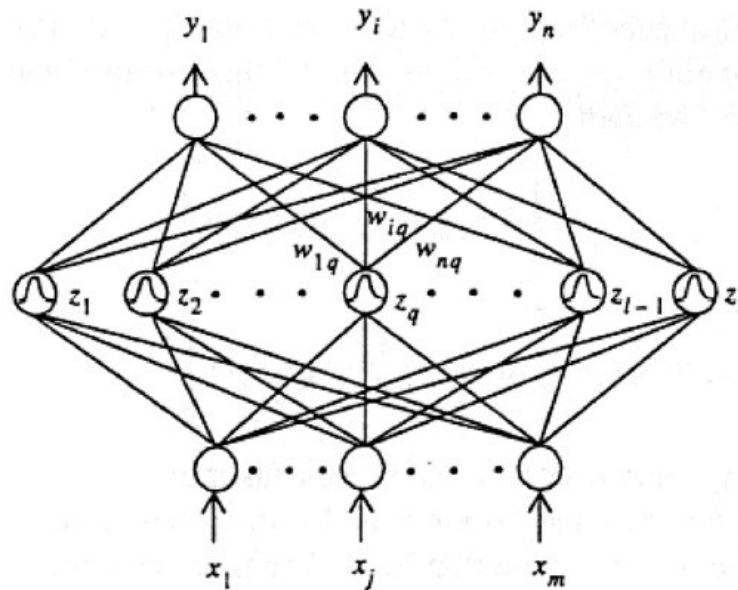
Kresanje mreze oznacava rucno uklanjanje cvorova i grana koji najmanje doprinose gresci. Weight decay oznacava softversko nestajanje odredjene grane – teжина grane teži nuli, sem ako se ne desi pojacanje te grane.

Da se mreza ne specijalizuje i da za slican set podataka radi skup za obucavanje mora da bude dovoljno veliki (veliki broj podataka) i da na celom opsegu imamo dovoljno odbiraka (da bude regularan). Ako na skupu za obucavanje imamo male promena na ulazu one ne uticu puno na promenu na izlazu (ne bi trebalo). Ako utice puno (prelazi u drugu klasu) to znaci da je potrebno preobucavanje date mreze (imamo suvise slozenu mrezu za dati problem).

1) Slabljenje pojacanja - Pojacanja koja uticu na performansu (ne uticu) izbacujemo  
 $w_{ij}(k+1) = (-\eta \frac{dE}{dw_{ij}})(k) + \omega_{ij}(k) \Leftrightarrow E' = E + \gamma \sum w_{ij}^2 \in [0, 1]$   
Svako takvo pojacanje preko 3 dovodimo do 0

2) Potkresivanje veze i čvorova – Ukidamo čvorove i veze(grane) koje najmanje doprinose grešci.  
Ako izbacimo neki čvor ili granu i greška se ne povećava ili se povećava malo onda možemo izbaciti čvor ili granu.

## 6. Struktura neuralnih mreža zasnovanih na funkcijama radijalne osnove (Radial basis function networks).



Predstavlja neuralnu mrežu sa tri sloja čvorova:

1. ulaznim,
2. izlaznim i
3. skrivenim (jezgrima)

Ulazni sloj služi da se prihvate ulazi u mrežu, skriveni ima aktivacionu funkciju sa radijalnom osnovom (npr.  $e^{-n^2}$ ) i izlaz je neuron koji sve to sabira. Tezine su povezane sa ulaznog sa svih neurona na sve neurone skrivenog sloja i svi sa skrivenog su povezani sa svim sa izlaznog sloja.

Svaki čvor skrivenog sloja predstavlja jednu f-ju radijalne osnove sa pridruženom centralnom pozicijom i širinom.

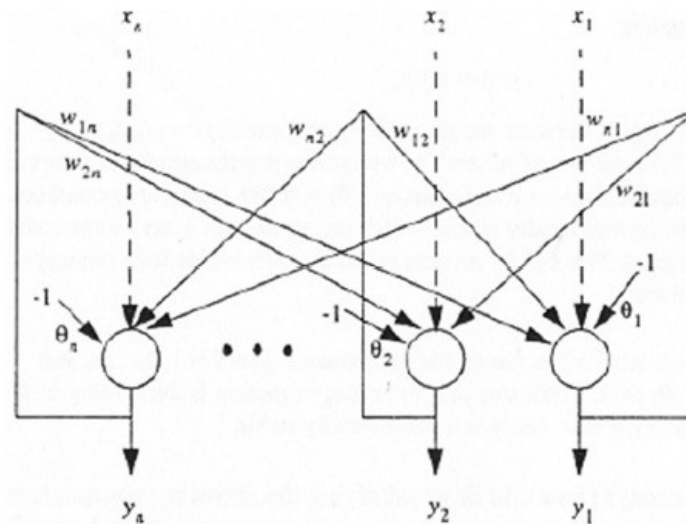
## 7. Definisati i objasniti uticaj konstante obučavanja na proceduru obučavanja mreže.

Konstanta obučavanja označava koliko će jako nove promene biti primenjene na trenutne tezine.

Što je  $\eta$  veće algoritam brže konvergira ako je preveliko divergira eksperimentalno od  $10^{-3}$  do  $10$ . Povećavamo  $\eta$  u toku obučavanja ako greška opada u više iteracija konzistentno.

Prevelike vrednosti dovode do velikih promena tezina (overshooting) što prouzrokuje nemogućnost da se dođe do lokalnog minimum, dok premale vrednosti previše sporo konvergiraju ka minimumu.

## 8. Definirati Hopfield-ove mreže i skicirati njihovu strukturu.



Hopfield-ova mreža je jednoslojna rekurentna mreža (vraća izlaze na ulaze). Radi u diskretnom i kontinualnom vremenu.

Asinhrono pravilo apdejtovanja – apdejtujemo  $y$  i odmah promenimo to novo  $y$  za računanje nekog drugog  $y$ .

Sinhrono pravilo apdejtovanja – novu vrednost  $y$  promenimo tek u sledećoj iteraciji. Izlazi mogu biti unipolarni, bipolarni, digitalni. Ulazi se dovode samo jednom u mrežu (prvi put), a posle se sklone i mreža onda radi sama. Izlazi iz svakog neurona se vraćaju na ulaze svih drugih neurona sem na sebe. Pojačanja su simetrična  $w_{ij}=w_{ji}$  i važi  $w_{ii}=0$ .

Pravilo apdejtovanja

$$y_i = \text{sgn}(\sum_{j=1, j \neq i}^n w_{ij} y_j + x_i - \theta_i)$$

## 9. Kako se određuju pojačanja kod Hopfield-ove asocijativne memorije, kada je potrebno zapamtiti određen broj oblika?

Bipolarni binarni vektori:

$$(1) W = \sum_{k=1}^p x_k (x_k)^T - pI \Leftrightarrow (2) w_{ij} = \sum_{k=1}^p x_{ik} x_{jk} - \delta_{ij} \quad w_{ii} = 0$$

(1) Oduzima se  $p$  uzoraka puta  $I$  da bi na dijagonalama matrice bile nule ( $I$  je jedinična matrica).

(2) Nema oduzimanja jer smo kod ovog naglasili da je  $w_{ii}$  jednako 0.

Unipolarni binarni vektori:

$$w_{ij} = \sum_{k=1}^p (2x_{ik} - 1)(2x_{jk} - 1) \quad i \neq j \quad w_{ii} = 0$$

## 10. Koje su vrste obučavanja neuralnih mreža zasnovanih na funkcijama radijalne osnove.

RBFN se zasniva na hibridnom obučavanju:

- 1) supervizijsko učenje za izlazni sloj
- 2) nesupervizijsko za ulazni sloj

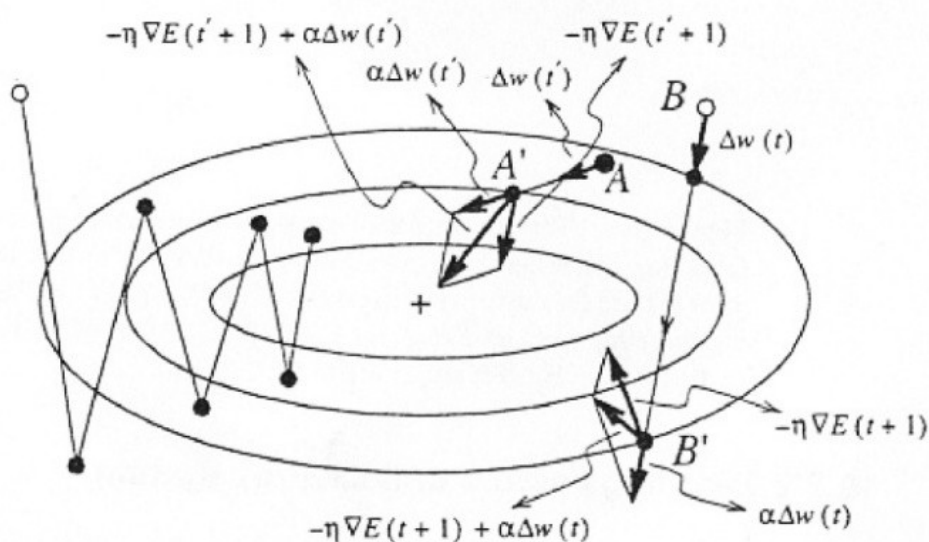
1) Za izlazni sloj se može primeniti delta pravilo obučavanja

$$\Delta w_{iq} = \eta (d_i - y_i) Z_{iq}$$

(imamo tačan odgovor  $d$ )

2) nesupervizijsko se zasniva na tome da se nađu  $m$  i  $q$  .  $m$  i  $q$  se nalazi primenjujući Honen-ovo pravilo obučavanja  $\Delta m$  najbliže  $= \eta(x - m$  najbliže  $)$ , nađemo centar najbliži  $x$ -y i njega menjamo, ostale ne diramo, dok se  $\sigma$   $q$  od hoc  $\sigma$   $q = |m$   $q$  -  $m$  najbliže  $|/\gamma$ . Može se trenirati preko back propagation algoritma i time postati u potpunosti supervizijska.

## 11. Objasniti pojam momentum-a, koje su prednosti njegove primene i skicirati primer.



Poboljšavamo pojačanje tako što kada pronalazimo apdejt pojačanja uzimamo dosta i od prethodnog. Bez momentuma putanja ima veće oscilacije (leva u odnosu na desnu). Poboljšava usmeravanje ka cilju i ako smo na pravom putu i ako smo preterali sa  $\eta$ .

Povećava (poboljšava) konvergenciju. Primenjuje se kod batch moda.

$$\Delta w(t) = -\eta \Delta(\text{obrnuto})E(t) + \alpha \Delta w(t-1) \quad \alpha \in [0, 1]$$

Uzima se obično  $\alpha=0.9$

**12. Definisati nesupervizijsko obučavanje neuralnih mreža zasnovanih na funkcijama radialne osnove, i objasniti kako ono funkcioniše.**

Nesupervizijsko se zasniva na tome da se nađu  $m$  i  $q$  .  $m$  i  $q$  se nalazi primenjujući Hohen-ovo pravilo obučavanja  $\Delta m$  najbliže  $=\eta(x-m$  najbliže  $)$ , nađemo centar najbliži  $x-y$  i njega menjamo, ostale ne diramo, dok se  $\sigma$  od hoc  $\sigma =|m$  i  $q$  -m najbliže  $|/\gamma$ . Može se trenirati preko back propagation algoritma i time postati u potpunosti supervizijska.

**13. Definisati preobučavanje, i objasniti metodu ranog zaustavljanja. Zašto se koristi metoda ranog zaustavljanja?**

Preobucavanje predstavlja pojam koji oznacava previse dobru obuku mreze za obucavajuće podatke, dok će greška biti veoma velika za neke druge vrednosti ulaznih podataka (npr test podaci). Javlja se kada ima previse parametara koji se mogu obuciti, pa mreza pocinje dobro da uci, ali lose da generalizuje.

**Metoda ranog zaustavljanja se koristi kao metod borbe protiv overfittinga (pojava kada se prekompleksna mreza specijalizuje za ulazne podatke sa obucavanja i ne ume vise dobro da generalizuje za podatke sa simulacije). Realizuje se tako sto se uporedo sa obucavanjem mreze gleda greška na nekom nezavisnom skupu podataka. Sve dok greška za njih ne raste, to znaci da mreza kreće da se specijalizuje za trening podatke, a to je potrebno izbeci**

Metoda ranog zaustavljanja se realizuje na sledeci nacin: paralelno sa obucavanjem mreze na podacima za obuku, mreza se testira sa nekim drugim podacima i sve dok se greška smanjuje, vrsi se obucavanje na osnovnom skupu. Kada greška pocne da raste, obucavanje se prekida.

**14. Koje su prednosti, a koji nedostaci Hopfield-ovih mreža?**

Brzo i lako obucavanje i jako dobro prepoznavanje oblika su prednosti.

Nedostaci su ogranicenost onoga sto se prepoznaje, neplanirana stabilna stanja i nesigurno vraćanje (odnosi se na Hopfieldovu memoriju, kada je prenatrpana sa razlicitim oblicima, ako je mala razlika izmedju oblika, može se desiti da se vrati pogresan oblik).

## CRTANJE

15. **Nacrtati neuralnu feedforward neuralnu mrežu koja treba da klasifikuje četiri različita oblika, a postoje merenja boje i težine. Mreža poseduje dva skrivena sloja sa po tri neurona. Predložiti aktivacione funkcije neurona u svim slojevima, i objasniti kodiranje izlaznih podataka.**

Ulazni sloj ima tri neurona sa linearnim aktivacionim funkcijama I na prvi je povezana velicina, drugi boja, treci tezina. Skriveni sloj ima tri neurona, izlazni ima cetiri neurona. Aktivacione funkcije za skriveni I izlazni sloj su npr. sigmoidne fje. Prvi neuron u izlaznom sloju ako je 1, onda je to prvi oblik, ako je drugi neuron 1, drugi oblik itd.

16. **Potrebno je napraviti neuralnu mrežu koja će aproksimirati funkciju koja ima tri maksimuma. Skicirati neuralnu mrežu zasnovanu na funkciji radijalne osnove, čiji je ulaz x koordinata, a izlaz vrednost funkcije. Kako izgledaju sve funkcije aktivacije ovih neurona?**

Postoji ulazni sloj sa jednim neuronom (aktivaciona funkcija je linearna), sledeci sloj je skriveni sloj sa tri neurona sa aktivacionom funkcijom sa radijalnom osnovom. Izlazni sloj je jedan neuron sa linearnom(nisam siguran) aktivacionom funkcijom. Ulazni sloj (taj jedan neuron) ima po jednu granu ka svakom od tri neurona sakrivenog sloja, i svaki neuron skrivenog sloja ima po jednu granu ka izlaznom neuronu.

17. **Definisati Hopfield-ovu mreže sa četiri neurona, napisati matricu pojačanja i skicirati strukturu. Koje probleme može da rešava Hopfield-ova mreža.**

Hopfieldova mreza je mreza gde je izlaz svakog neurona povezan na ulaz svih ostalih neurona. Pocetni ulazi se pustaju jednom, nakon cega se uklanjaju. Tezine su simetricne. Izlazi se update-uju sve dok mreza ne postane stabilna.

Matrica pojacanja se dobija formulom:

suma(od k=1 do n)  $X_i \cdot X_{iT} - p_i$

Moze da resava problem u kojima je potrebno prepoznati ulazni vektor, npr kada je potrebno prepoznati oblike (slika se sastoji od piksela, pa zapamti crne piksele).

18. Nacrtati neuralnu feedforward neuralnu mrežu koja treba da klasifikuje četiri različita oblika, a postoje merenja veličine, boje i težine. Mreža poseduje jedan skriveni sloj sa tri neurona. Predložiti aktivacione funkcije neurona u svim slojevima, i objasniti kodiranje izlaznih podataka.

Tri ulazna neurona za velicinu, boju i tezinu, skriveni sloj ima tri neurona, a izlazni sloj ima dva neurona koju mogu da daju cetiri razlicita oblika binarnim vrednostima.

Y1Y2:

- 00 – prvi oblik
- 01 – drugi oblik
- 10 – treci oblik
- 11 – cetvrti oblik

19. Koje su prednosti povećavanja vrednosti konstante obučavanja, a koji su loše posledice?

20. Kako se određuju pojačanja kod Hopfield-ove asocijativne memorije, kada je potrebno zapamtiti 5 oblika?

21. Nacrtati mrežu koja za ulaz ima  $x$  i  $y$  koordinate, izlaz su joj dve funkcije  $z_1$  i  $z_2$ .