Izvještaj za 6. domaću zadaću iz kolegija Neizrazito, evolucijsko i neuroračunarstvo

Autor: Marijo Cvitanović 0036510758

December 17, 2020

1 Uvod

Za rješavanje šeste domaće zadaće bilo je potrebno implementirati neuro-fuzzy sustav koji odgovara ANFIS-u. Točnije, neuro-fuzzy sustav koji koristi zaključivanje tipa 3 (metoda *Takagi-Sugeno-Kang*, *TSK*).

Dodatno, uz implementaciju, bilo je potrebno riješiti nekoliko zadataka, čija rješenja su dostupna u nastavku ovog dokumenta.

2 Izvod postupka učenja

Funkcija, koju želimo da mreža nauči, je oblika:

$$f(x,y) = ((x-1)^2 + (y+2)^2 - 5xy + 3) \cdot cos^2(\frac{x}{5})$$

Na mrežu dovodimo N primjera koji služe kao primjeri za učenje. Primjeri za učenje su oblika:

$$\{((x_1, y_1), z_1), ..., ((x_N, y_N), z_N)\}$$

gdje x_i i y_i predstavljaju ulaze funkcije koju želimo naučiti, a z_i predstavljaju izlaze te funkcije.

Nadalje, izlaz iz sustava neizrazitog upravljanja za k-ti primjer je o_k.

Funkcija pogreške za taj primjer je:

$$E_k = \frac{1}{2}(z_k - o_k)^2 \tag{1}$$

a ako se radi o potpunom gradijentu za sve primjere:

$$E = \sum_{k=1}^{N} E_k$$

Za ažuriranje proizvoljnog parametra ψ koristi se izraz

$$\psi(t+1) = \psi(t) - \eta \cdot \frac{\partial E_k}{\partial \psi} \tag{2}$$

gdje je t trenutna iteracija, a $t+1$ iduća, dok η predstavlja stopu učenja. Za stopu učenja je dobro uzeti različite vrijednosti.

Nadalje, sustav raspolaže sa m pravila oblika:
R_1: Ako x je A_1 I y je B_1 tada z je $p_1x + q_1y + r_1$
...
R_m: Ako x je A_m I y je B_m tada z je $p_mx + q_my + r_m$

Pri čemu su A_i i B_i neizraziti skupovi modelirani na sljedeći način

$$\mu_{A_i}(x) = \frac{1}{1 + e^{b_i \cdot (x - a_i)}} \tag{3}$$

$$\mu_{B_i}(y) = \frac{1}{1 + e^{d_i \cdot (y - c_i)}}$$

Presjek ta dva skupa se modelira pripadajućom t-normom, koja je u ovom slučaju produkt

$$\mu_{A_i \cap B_i}(x) = \mu_{A_i}(x) \cdot \mu_{B_i}(x)$$

Izlaz sustava o_k je definiran

$$o = \frac{\sum_{i=1}^{m} \alpha_i \cdot f_i}{\sum_{i=1}^{m} \alpha_i} \tag{4}$$

α_i predstavlja jakost paljenja i-tog pravila, a $f_i = p_i x + q_i y + r_i$

$$\alpha_i = \mu_{A_i}(x) \cdot \mu_{B_i}(x) \tag{5}$$

Iz svega prethodno navedenog, zaključujem da je potrebno napraviti učenje parametara

$$a_i, b_i, c_i, d_i, p_i, q_i, r_i$$

2.1 Parametar a_i

Prvo, učenje parametra a_i po jednadžbi 2

$$a_i(t+1) = a_i(t) + \frac{\partial E_k}{\partial a_i}$$

Iz jednadžbe 1 vidimo da pogreška E_k ne ovisi o parametru a_i, nego o izlazu neizrazitog sustava o_k. Izlaz neizrazitog sustava ovisi o parametru α_i (jednadžba 4), a α ovisi o μ_{A_i} (jednadžba 5) koji tek ovisi o a_i (jednadžba 3).
Uočavamo lanac $a_i \rightarrow \mu_{A_i} \rightarrow \alpha_i \rightarrow o_k$.
Stoga po lančanom pravilu za deriviranje funkcija možemo pisati

$$\frac{\partial E_k}{\partial a_i} = \frac{\partial E_k}{\partial o_k} \frac{\partial o_k}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial \mu_{A_i}} \frac{\partial \mu_{A_i}}{\partial a_i}$$

Prvo deriviramo $\frac{\partial \mu_{A_i}}{\partial a_i}$ i tako sve parcijalne derivacije od kraja deriviramo pa pomnožimo

$$\frac{\partial \mu_{A_i}}{\partial a_i} = \frac{\partial}{\partial a_i}\left(\frac{1}{1 + e^{b_i \cdot (x - a_i)}}\right)$$

Primijenimo $\frac{1}{a} = a^{-1}$ te pravilo deriviranja $\frac{dx^n}{dx} = nx^{n-1}$ te niz lančanih derivacija

$$\frac{\partial}{\partial a_i}((1+e^{b_i\cdot(x-a_i)})^{-1}) = -1\cdot(1+e^{b_i(x-a_i)})^{-2}\cdot\frac{\partial}{\partial a_i}(1+e^{b_i(x-a_i)}) =$$

Nadalje, koristimo pravilo deriviranja $\frac{de^{cx}}{dx} = ce^{cx}$ te $\frac{dc}{dx} = 0$, odnosno deriviranje konstante je 0.

$$= -1\cdot(1+e^{b_i(x-a_i)})^{-2}\cdot(-b_i)\cdot e^{b_i(x-a_i)}$$

te konačno imamo

$$\frac{\partial\mu_{A_i}}{\partial a_i} = (1+e^{b_i(x-a_i)})^{-2}\cdot b_i\cdot e^{b_i(x-a_i)}$$

što se dalje može srediti

$$\frac{\partial\mu_{A_i}}{\partial a_i} = b_i\frac{e^{b_i(x-a_i)}}{1+e^{b_i(x-a_i)}}\cdot\frac{1}{1+e^{b_i(x-a_i)}} =$$

$$= b_i\cdot\frac{1+e^{b_i(x-a_i)}-1}{1+e^{b_i(x-a_i)}}\cdot\frac{1}{1+e^{b_i(x-a_i)}} =$$

$$= b_i\cdot(\frac{1+e^{b_i(x-a_i)}}{1+e^{b_i(x-a_i)}} - \frac{1}{1+e^{b_i(x-a_i)}})\cdot\frac{1}{1+e^{b_i(x-a_i)}} =$$
$$= b_i\cdot(1-\mu_{A_i})\cdot\mu_{A_i}$$

te konačno imamo da je

$$\frac{\partial\mu_{A_i}}{\partial a_i} = b_i\cdot(1-\mu_{A_i})\cdot\mu_{A_i}$$

Dalje, treba izračunati $\frac{\partial\alpha_i}{\partial\mu_{A_i}}$

$$\frac{\partial\alpha_i}{\partial\mu_{A_i}} = \frac{\partial}{\partial\mu_{A_i}}(\mu_{A_i}\cdot\mu_{B_i}) = \mu_{B_i}$$

Nakon toga treba izračunati $\frac{\partial o_k}{\partial\alpha_i}$

$$\frac{\partial o_k}{\partial\alpha_i} = \frac{\partial}{\partial\alpha_i}\left(\frac{\sum_{i=1}^{m}\alpha_i\cdot f_i}{\sum_{i=1}^{m}\alpha_i}\right)$$

Kako se α_i pojavljuje i u brojniku i u nazivniku, primjenjujemo pravilo za deriviranje razlomka

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}$$

$$\frac{\partial o_k}{\partial \alpha_i} = \frac{\sum_{j=1}^{m} \alpha_j \cdot f_i - \sum_{j=1}^{m} \alpha_j \cdot f_j}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} =$$

$$= \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2}$$

Te konačno, na kraju još je potrebno izračunati derivaciju $\frac{\partial E_k}{\partial o_k}$

$$\frac{\partial E_k}{\partial o_k} = \frac{\partial}{\partial o_k}\left(\frac{1}{2}(z_k - o_k)^2\right) =$$

$$= -(z_k - o_k)$$

Na kraju, konačna derivacija za parametar a_i izgleda

$$\frac{\partial E_k}{\partial a_i} = -(z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot b_i \cdot (1 - \mu_{A_i}) \cdot \mu_{A_i}$$

a izraz za ažuriranje parametra a_i

$$a_i(t+1) = a_i(t) + \eta \cdot (z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot b_i \cdot (1 - \mu_{A_i}) \cdot \mu_{A_i} \quad (6)$$

Prethodni izraz je stohastičko ažuriranje parametra a_i. Gradijentno bi izgledalo

$$a_i(t+1) = a_i(t) + \eta \cdot \sum_{i=1}^{N} \left((z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot b_i \cdot (1 - \mu_{A_i}) \cdot \mu_{A_i} \right) \quad (7)$$

2.2 Parametar b_i

Nakon toga, istu stvar napravimo za parametar b_i. Uočavamo isti lanac $b_i \to \mu_{A_i} \to \alpha_i \to o_k$.
Stoga po lančanom pravilu za deriviranje funkcija možemo pisati

$$\frac{\partial E_k}{\partial b_i} = \frac{\partial E_k}{\partial o_k}\frac{\partial o_k}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial \mu_{A_i}}\frac{\partial \mu_{A_i}}{\partial b_i}$$

Jedino što se razlikuje od parametra a_i je $\frac{\partial \mu_{A_i}}{\partial b_i}$, ostale derivacije su jednake

$$\frac{\partial \mu_{A_i}}{\partial b_i} = \frac{\partial}{\partial b_i}\left(\frac{1}{1+e^{b_i(x-a_i)}}\right) =$$

$$= -1 \cdot (1+e^{b_i(x-a_i)})^{-2} \cdot e^{b_i(x-a_i)} \cdot (x-a_i)$$

Te ekvivalentnim načinom kao za a_i, b_i je

$$\frac{\partial \mu_{A_i}}{\partial b_i} = -(x-a_i) \cdot \mu_{A_i} \cdot (1-\mu_{A_i})$$

Te konačna formula za derivaciju

$$\frac{\partial E_k}{\partial b_i} = (z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot (x-a_i) \cdot (1-\mu_{A_i}) \cdot \mu_{A_i}$$

Odnosno za ažuriranje parametra b_i stohastički

$$b_i(t+1) = b_i(t) - \eta \cdot (z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot (x-a_i) \cdot (1-\mu_{A_i}) \cdot \mu_{A_i}$$

S potpunim gradijentom

$$b_i(t+1) = b_i(t) - \eta \cdot \sum_{i=1}^{N} \left((z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot (x-a_i) \cdot (1-\mu_{A_i}) \cdot \mu_{A_i} \right)$$

2.3 Parametri c_i i d_i

Parcijalne derivacije za c_i su ekvivalentne onima za a_i, kao što su parcijalne derivacije od d_i ekvivalentne onima za b_i

$$c_i(t+1) = c_i(t) + \eta \cdot (z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{A_i} \cdot d_i \cdot (1 - \mu_{B_i}) \cdot \mu_{B_i}$$

$$c_i(t+1) = c_i(t) + \eta \cdot \sum_{i=1}^{N} \left((z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{A_i} \cdot d_i \cdot (1 - \mu_{B_i}) \cdot \mu_{B_i} \right)$$

$$d_i(t+1) = d_i(t) - \eta \cdot (z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{A_i} \cdot (y - c_i) \cdot (1 - \mu_{B_i}) \cdot \mu_{B_i}$$

$$d_i(t+1) = d_i(t) - \eta \cdot \sum_{i=1}^{N} \left((z_k - o_k) \cdot \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (f_i - f_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{A_i} \cdot (y - c_i) \cdot (1 - \mu_{B_i}) \cdot \mu_{B_i} \right)$$

2.4 Parametar p_i

Učenje radimo po formuli 2. Kako bismo odredili $\frac{\partial E_k}{\partial p_i}$ potrebno je uočiti lančanu strukturu. Krenimo redom. E_k ne ovisi direktno o p_i, nego prvo ovisi o o_k po jednadžbi 1

o_k također ne ovisi direktno o p_i nego o f_i po jednadžbi 4

Parametar f_i ovisi o p_i jer je $f_i = p_i x + q_i y + r_r$
Te uočavamo lanac $p_i \rightarrow f_i \rightarrow o_k$

$$\frac{\partial E_k}{\partial p_i} = \frac{\partial E_k}{\partial o_k} \frac{\partial o_k}{\partial f_i} \frac{\partial f_i}{\partial p_i}$$

Krenimo redom

$$\frac{\partial f_i}{\partial p_i} = \frac{\partial}{\partial p_i}(p_i x + q_i y + r_i) = x$$

$$\frac{\partial o_k}{\partial f_i} = \frac{\partial}{\partial f_i} \left(\frac{\sum_{i=1}^{m} \alpha_i \cdot f_i}{\sum_{i=1}^{m} \alpha_i} \right)$$

Pošto se f_i nalazi samo u brojniku i uz i-ti α, jedino će taj α ostati

$$\frac{\partial o_k}{\partial f_i} = \frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j}$$

Te konačno

$$\frac{E_k}{o_k} = -(z_k - o_k)$$

Te izraz za ažuriranje težine p_i jest

$$p_i(t+1) = p_i(t) + \eta(z_k - o_k)\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j} x$$

Odnosno za potpuni gradijent

$$p_i(t+1) = p_i(t) + \eta \sum_{i=1}^{N} \left((z_k - o_k)\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j} x \right)$$

2.5 Parametri q_i i r_i

Ekvivalentno kao p_i se dobije q_i i r_i

$$q_i(t+1) = q_i(t) + \eta(z_k - o_k)\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j} y$$

$$r_i(t+1) = r_i(t) + \eta(z_k - o_k)\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j}$$

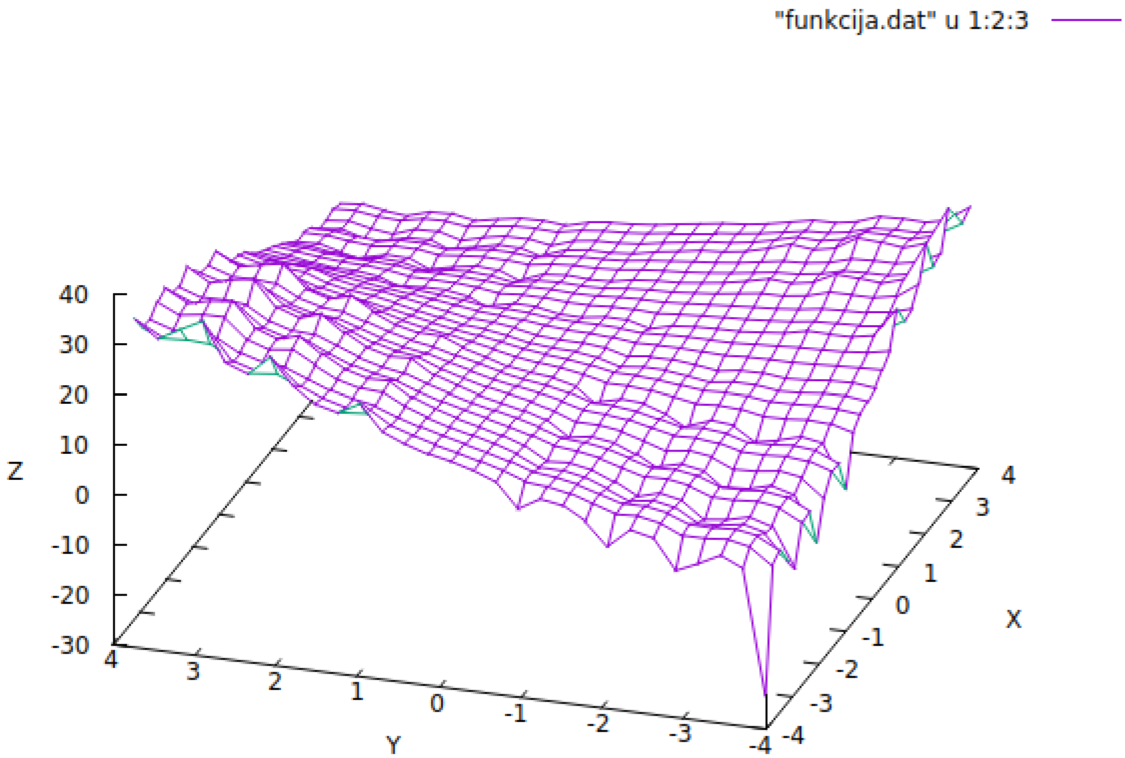
Odnosno za potpuni gradijent

$$q_i(t+1) = q_i(t) + \eta \sum_{i=1}^{N} \left((z_k - o_k)\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j} y \right)$$

$$r_i(t+1) = r_i(t) + \eta \sum_{i=1}^{N} \left((z_k - o_k)\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j} \right)$$

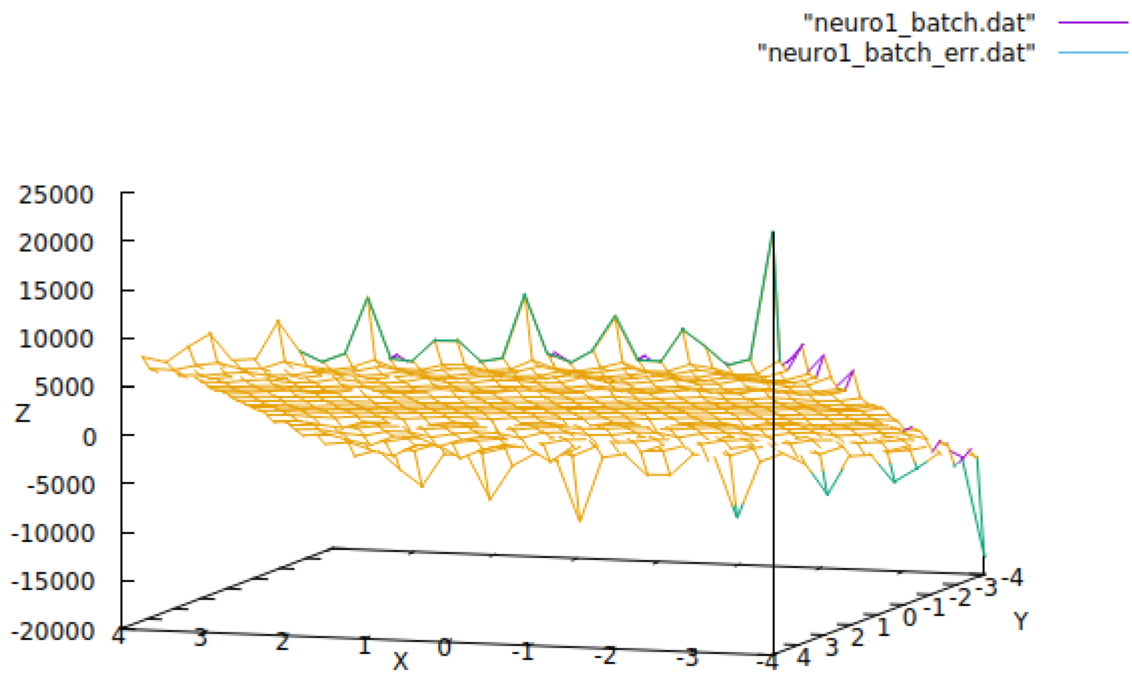
3 Grafovi

3.1 Graf funkcije



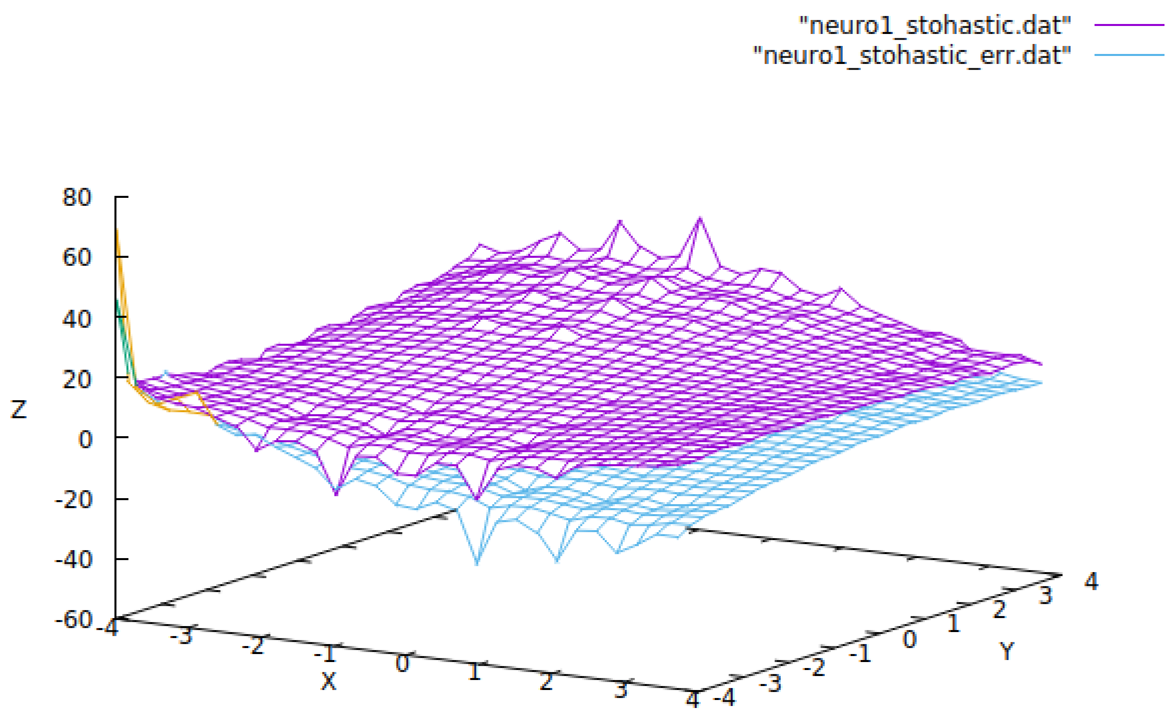
Slika 1: f(x, y)

3.2 Neuro fuzzy izlaz za 1 pravilo - batch



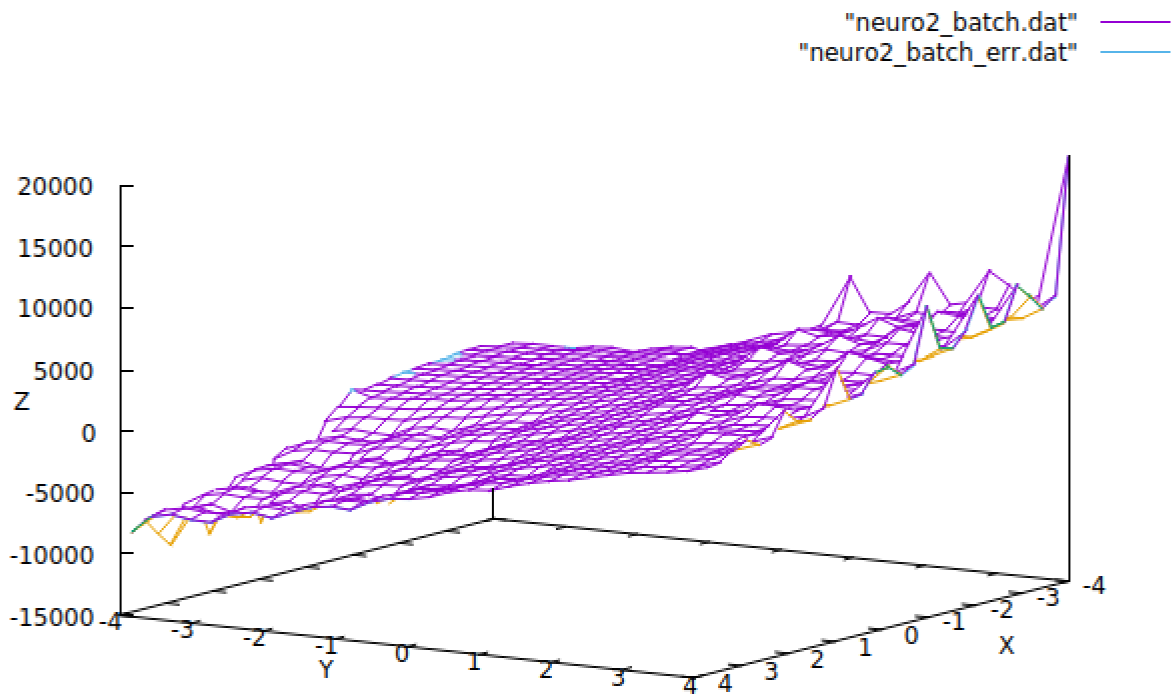
Slika 2: 1 pravilo, izlaz + greška

3.3 Neuro fuzzy izlaz za 1 pravilo - stohastic



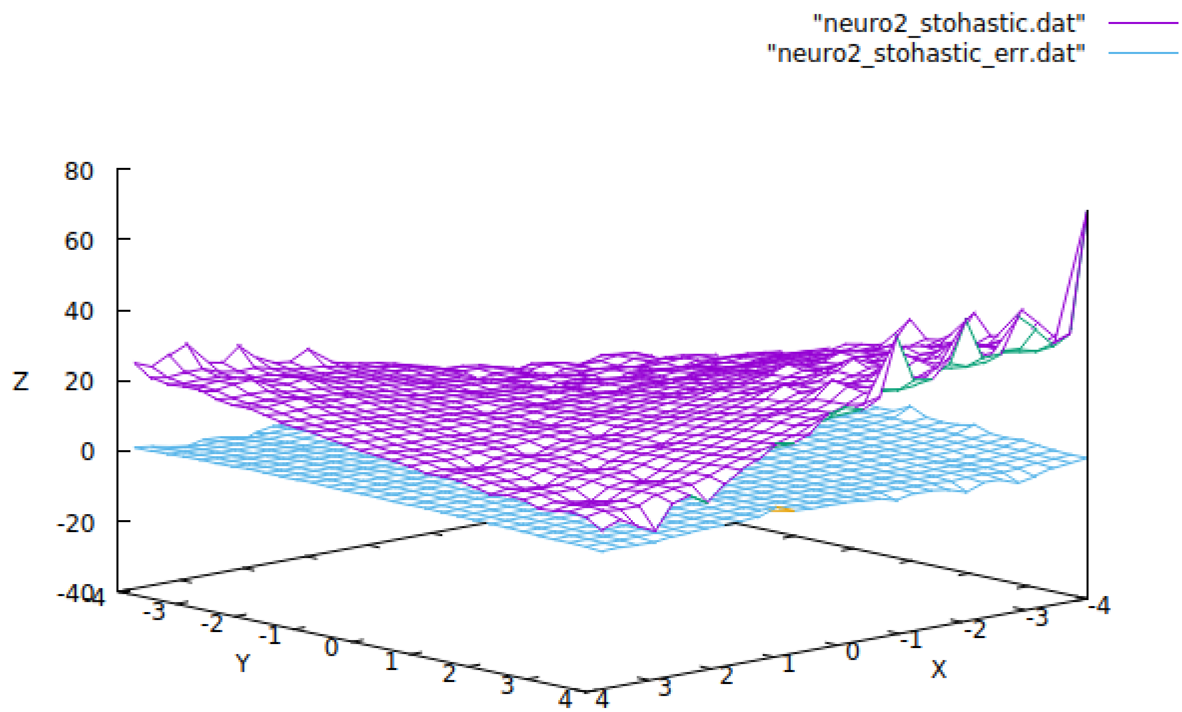
Slika 3: 1 pravilo, izlaz + greška

3.4 Neuro fuzzy izlaz za 2 pravila - batch



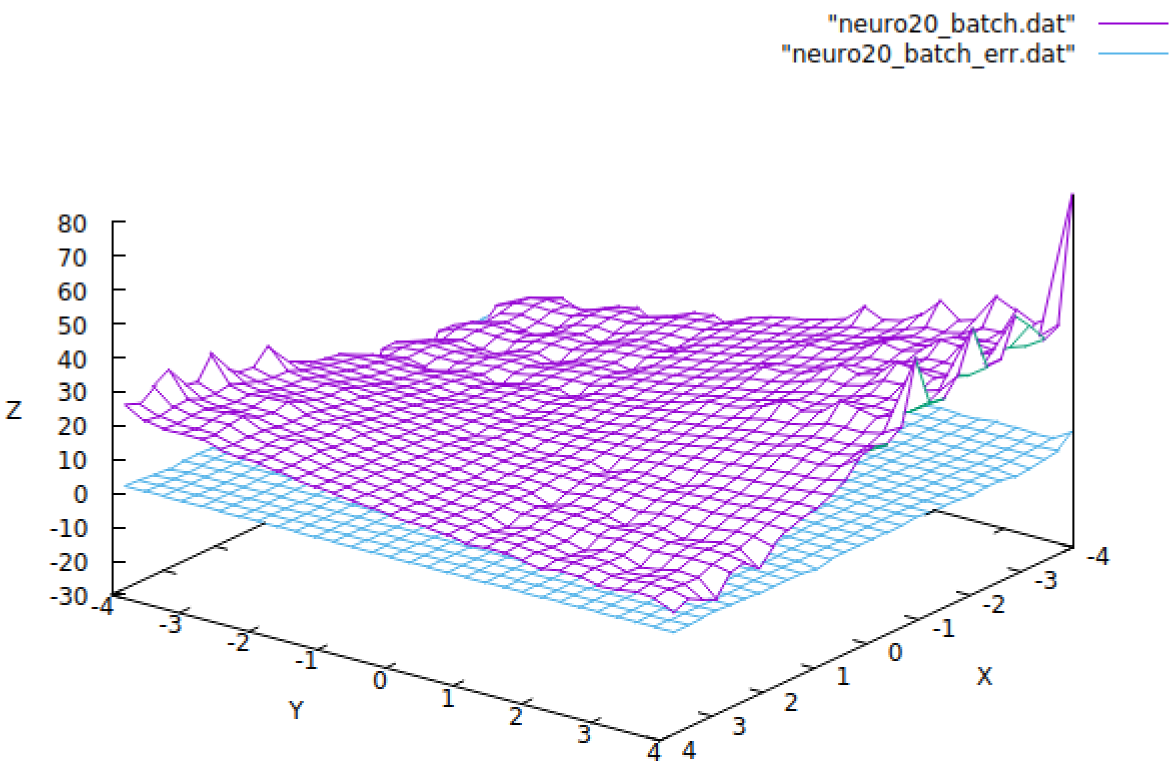
Slika 4: 2 pravila, izlaz + greška

3.5 Neuro fuzzy izlaz za 2 pravila - stohastic



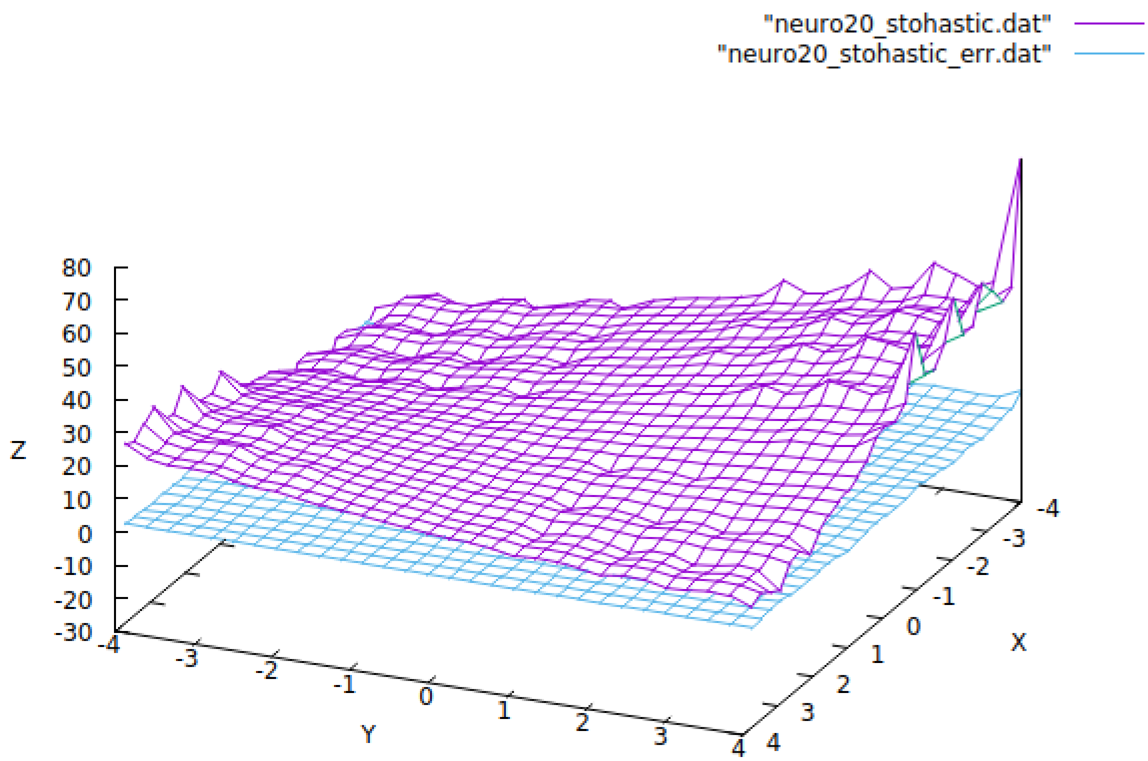
Slika 5: 2 pravila, izlaz + greška

3.6 Neuro fuzzy izlaz za 20 pravila - batch



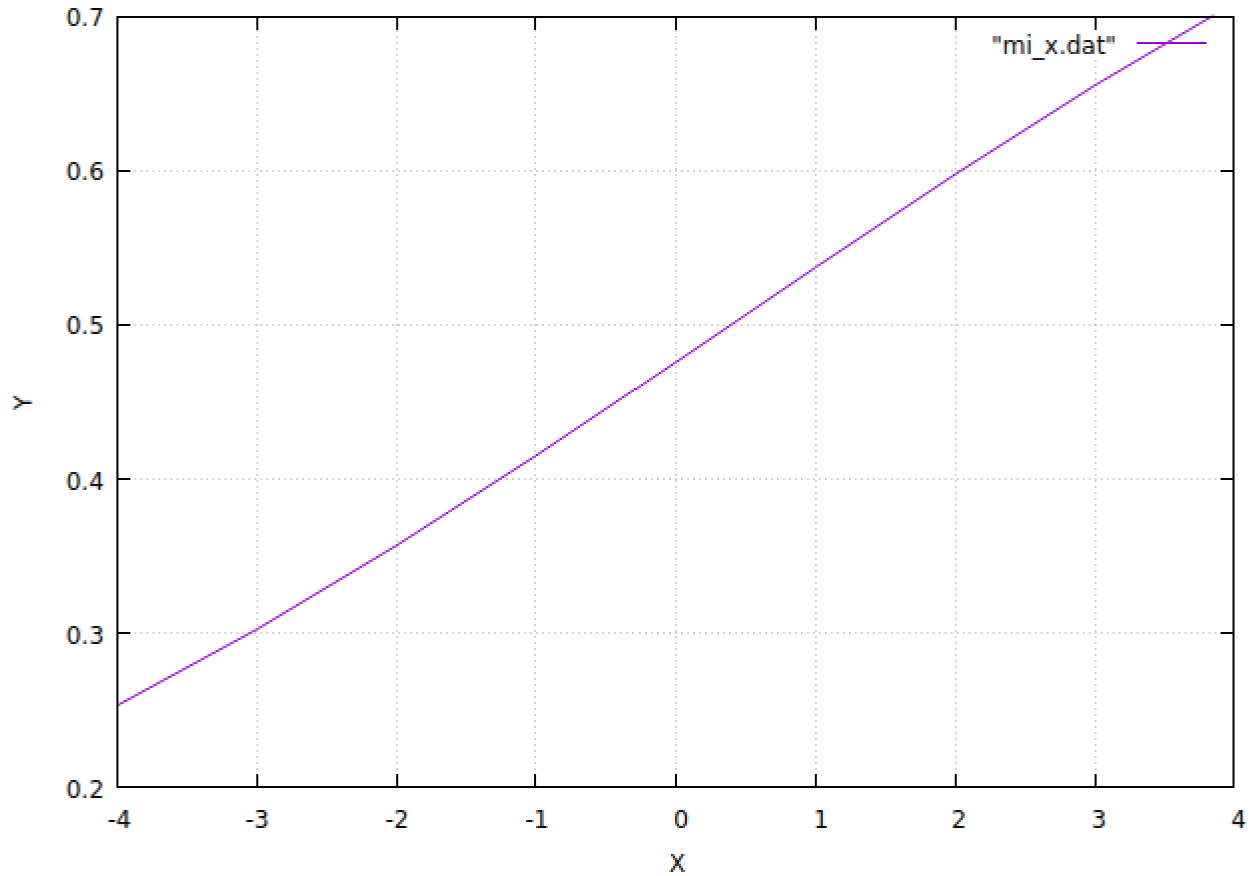
Slika 6: 20 pravila, izlaz + greška

3.7 Neuro fuzzy izlaz za 20 pravila - stohastic



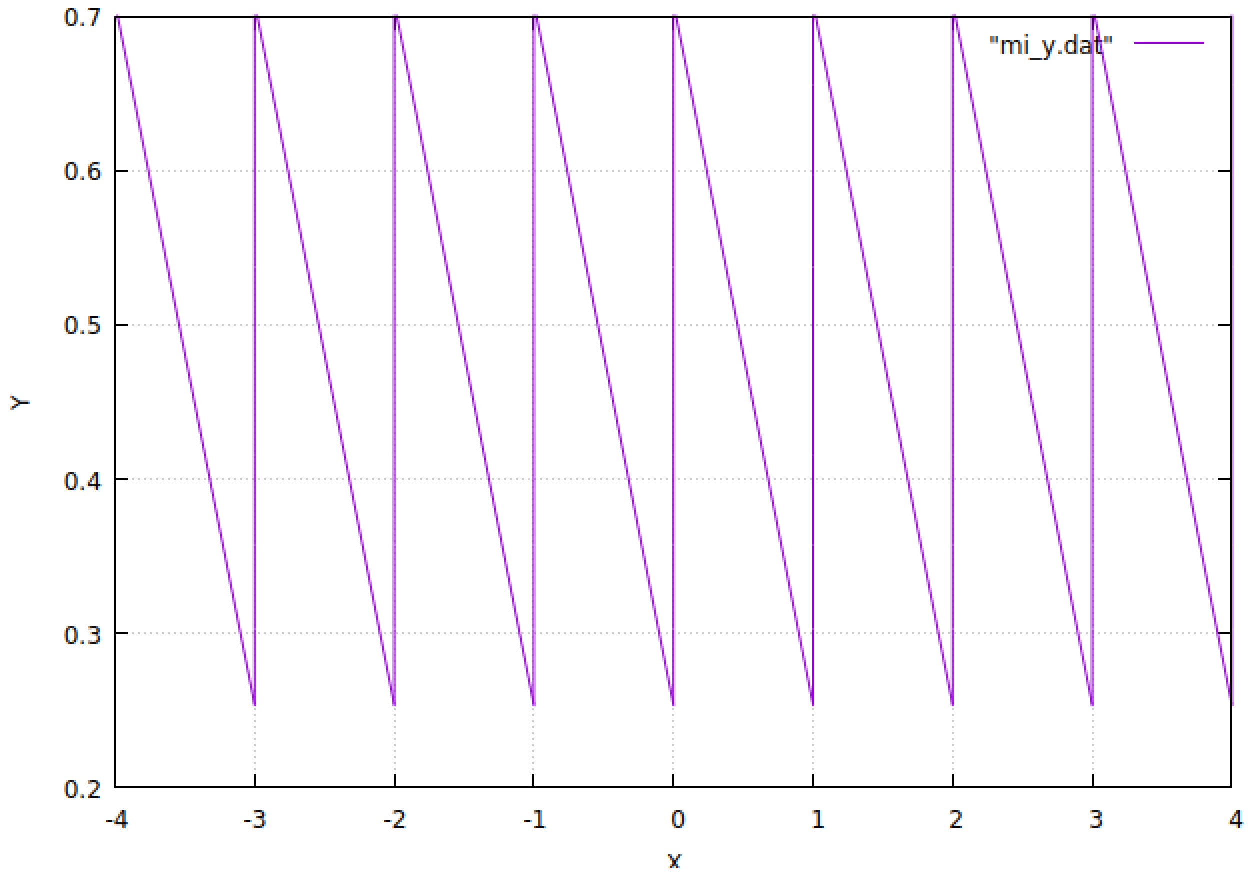
Slika 7: 20 pravila, izlaz + greška

3.8 Pripadnost od x



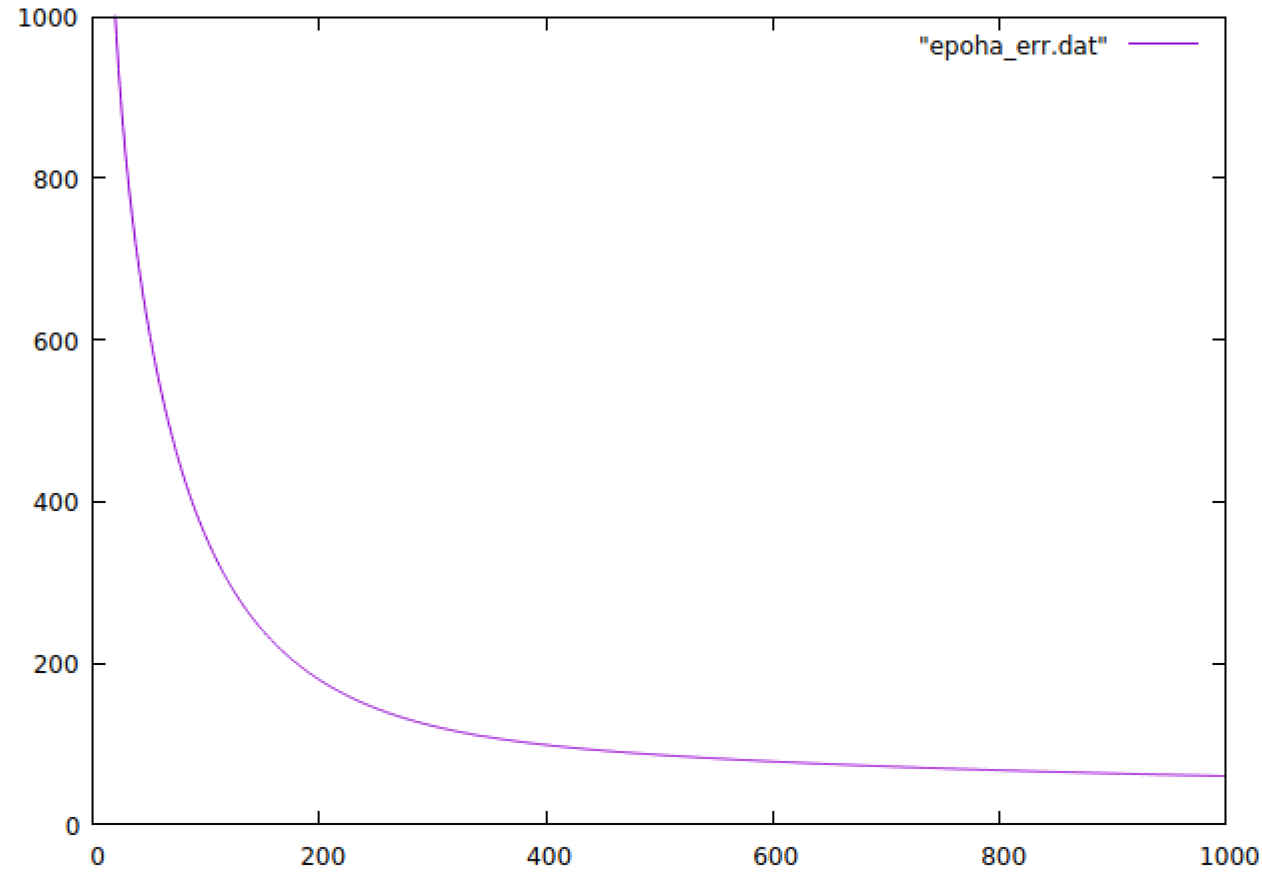
Slika 8: Pripadnost za vrijednost x

3.9 Pripadnost od y



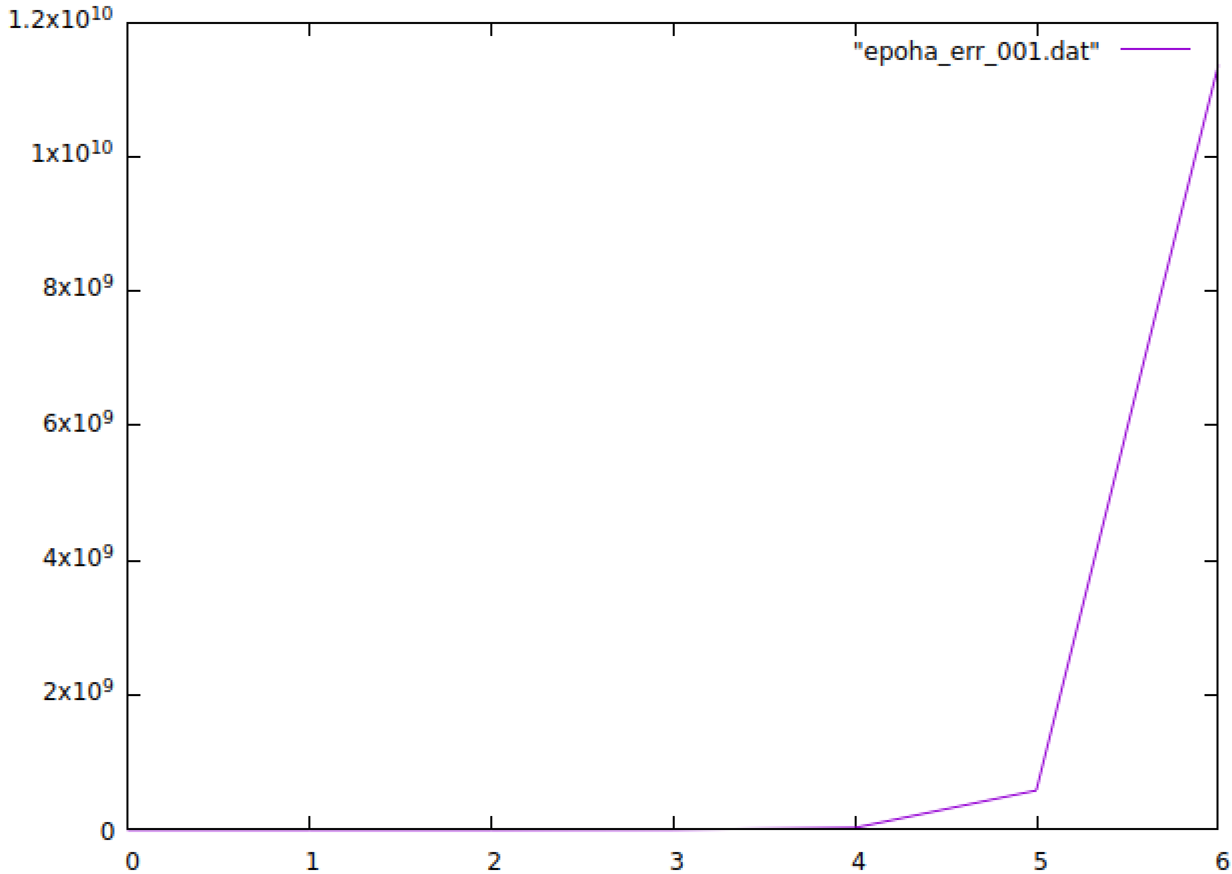
Slika 9: Pripadnost za vrijednost y

3.10 Greška kroz epohe - normalne ete



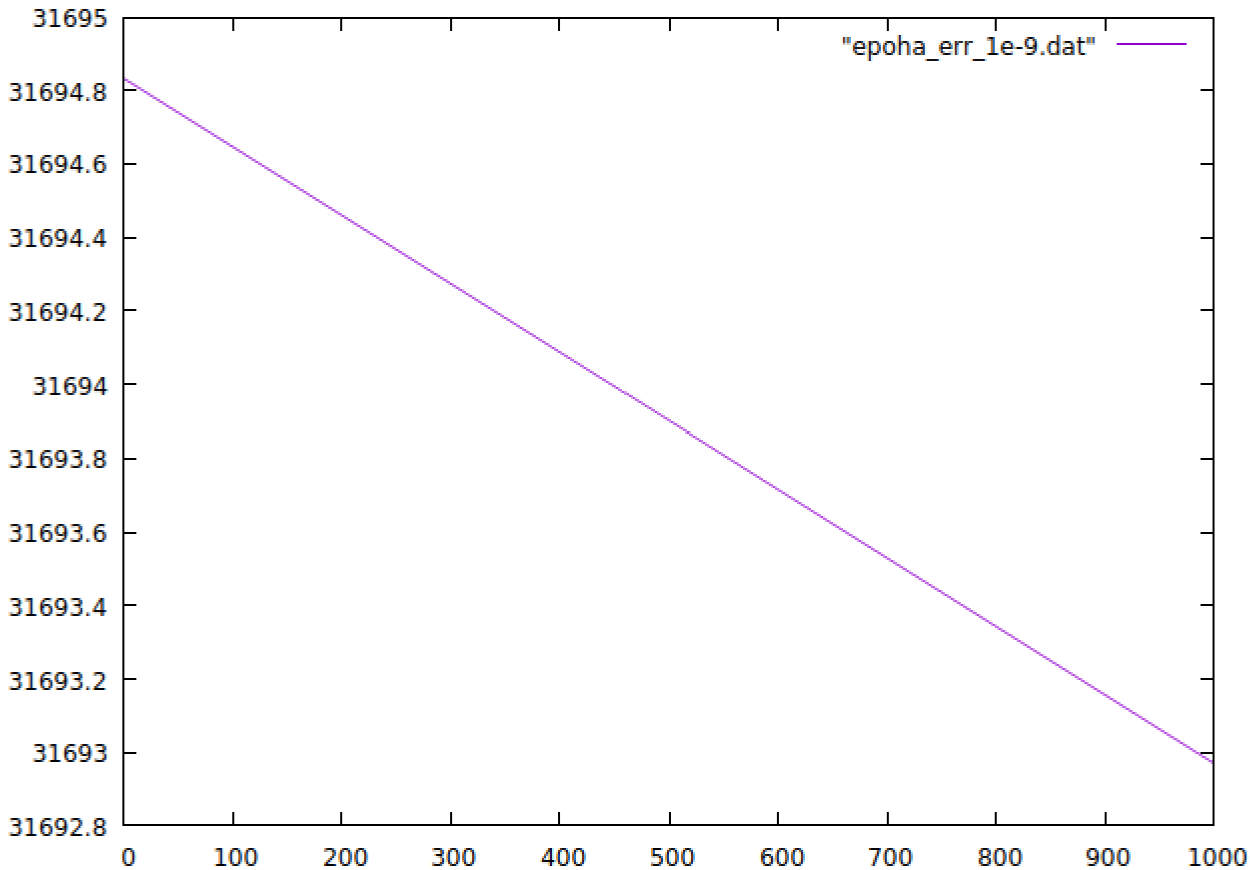
Slika 10: Greške kroz epohe

3.11 Greška kroz epohe - prevelike ete



Slika 11: Greške kroz epohe

3.12 Greška kroz epohe - premale ete



Slika 12: Greške kroz epohe