Neizrazito, evolucijsko i neuro računarstvo

6. domaća zadaća

Ivan Skorupan

Fakultet elektrotehnike i računarstva

1. zadatak

NENR

6. domaća zadaća

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

$$x, y \in [-4, 4]$$

> 81 uzorak ; $\{(x,y) : x, y \in [-4,4] ; x, y \in \mathbb{Z}\}$

$$\mu_{A_i}(x) = \frac{1}{1 + e^{b_i(x-a_i)}}$$ => prijenosna funkcija od A_i

$>$ T-norma $\rightarrow \mu_{A_i \cap B_i}(x) = \mu_{A_i}(x) \cdot \mu_{B_i}(x)$

$$\mu_{B_i}(x) = \frac{1}{1 + e^{d_i(x-c_i)}}$$ => prijenosna funkcija od B_i

$A_i(x) \equiv \mu_{A_i}(x)$

$B_i(x) \equiv \mu_{B_i}(x)$

$T(\mu_{A_i}(x), \mu_{B_i}(x)) = \alpha_i$

$$o_k = \frac{\sum_{i=1}^{m} \alpha_i z_i}{\sum_{i=1}^{m} \alpha_i}$$

$z_i = p_i x + q_i y + r_i$

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

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

Za online: $E = E_k$

$a_i \Rightarrow \mu_{A_i} \Rightarrow \alpha_i \Rightarrow o_k$

$b_i \Rightarrow \mu_{A_i} \Rightarrow \alpha_i \Rightarrow o_k$

$c_i \Rightarrow \mu_{B_i} \Rightarrow \alpha_i \Rightarrow o_k$

$d_i \Rightarrow \mu_{B_i} \Rightarrow \alpha_i \Rightarrow o_k$

$p_i \Rightarrow z_i \Rightarrow o_k$

$q_i \Rightarrow z_i \Rightarrow o_k$

$r_i \Rightarrow z_i \Rightarrow o_k$

Izvod za p_i:

$$\frac{\partial E_k}{\partial p_i} = \frac{\partial E_k}{\partial o_k}\,\frac{\partial o_k}{\partial z_i}\,\frac{\partial z_i}{\partial p_i}$$

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

$$\frac{\partial o_k}{\partial z_i} = \frac{\partial}{\partial z_i}\,\frac{\sum_{j=1}^{m} \alpha_j z_j}{\sum_{j=1}^{m} \alpha_j} = \frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j}$$

$$\frac{\partial z_i}{\partial p_i} = x$$

$$\frac{\partial E_k}{\partial p_i} = -x(z_k - o_k)\,\frac{\alpha_i}{\sum_{j=1}^{m} \alpha_j}$$

$$p_i(t+1) = p(t) - \eta\,\frac{\partial E_k}{\partial p_i}$$

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

Izvod za q_i:

$$\frac{\partial E_k}{\partial q_i} = \frac{\partial E_k}{\partial o_k}\,\frac{\partial o_k}{\partial z_i}\,\frac{\partial z_i}{\partial q_i}$$

$$\frac{\partial z_i}{\partial q_i} = y$$

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

Izvod za r_i:

$$\frac{\partial E_k}{\partial r_i} = \frac{\partial E_k}{\partial o_k} \frac{\partial o_k}{\partial z_i} \frac{\partial z_i}{\partial r_i}$$

$$\frac{\partial z_i}{\partial r_i} = 1$$

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

Izvod za a_i:

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

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

$$\frac{\partial o_k}{\partial \alpha_i} = \frac{\sum_{j=1, j\neq i}^{m} \alpha_j (z_i - z_j)}{\left(\sum_{j=1}^{m} \alpha_j\right)^2}$$

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

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

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

$$\boxed{a_i(t+1) = a_i(t) + \eta \cdot (z_k - o_k) \cdot \frac{\sum\limits_{j=1, j\neq i}^{m} \alpha_j (z_i - z_j)}{\left(\sum\limits_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{B_i} \cdot b_i \cdot (1-\mu_{A_i}) \cdot \mu_{A_i}}$$

//

Izvod za b_i:

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

$$\frac{\partial \alpha_i}{\partial \mu_{A_i}} = \mu_{B_i} \quad //$$

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

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

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

Izvod za c_i:

$$\frac{\partial E_k}{\partial c_i} = \frac{\partial E_k}{\partial o_k} \frac{\partial o_k}{\partial \alpha_i} \frac{\partial \alpha_i}{\partial \mu_{B_i}} \frac{\partial \mu_{B_i}}{\partial c_i}$$

$$\frac{\partial \alpha_i}{\partial \mu_{B_i}} = \mu_{A_i} \quad // \quad , \quad \frac{\partial \mu_{B_i}}{\partial c_i} = d_i \cdot \mu_{B_i} \cdot (1-\mu_{B_i}) \quad //$$

$$\boxed{c_i(t+1) = c_i(t) + \eta \cdot (z_k - o_k) \cdot \frac{\sum\limits_{j=1, j\neq i}^{m} \alpha_j (z_i - z_j)}{\left(\sum\limits_{j=1}^{m} \alpha_j\right)^2} \cdot \mu_{A_i} \cdot d_i \cdot (1-\mu_{B_i}) \cdot \mu_{B_i}}$$

Izvod za d_i:

$$\frac{\partial E_k}{\partial d_i} = \frac{\partial E_k}{\partial o_k}\frac{\partial o_k}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial \mu_{B_i}}\frac{\partial \mu_{B_i}}{\partial d_i}$$

$$\frac{\partial \mu_{B_i}}{\partial d_i} = -(x-c_i)\cdot\mu_{B_i}\cdot(1-\mu_{B_i})$$

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

Odlučimo li se koristiti grupnu verziju algoritma, u jednoj ćemo epohi proći kroz sve primjere, za svaki izračunati sve potrebne parcijalne derivacije i njih kroz svaki primjer posumirati kako bismo dobili točan iznos gradijenta funkcije pogreške.

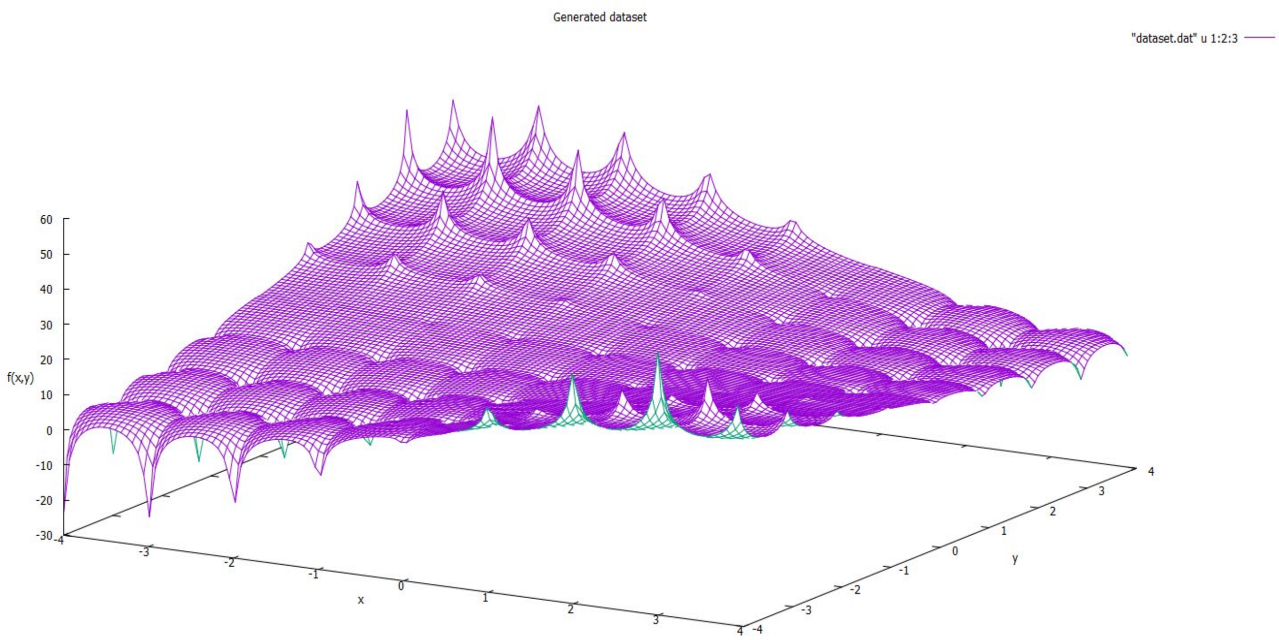
Kod online verzije, u jednoj epohi isto prolazimo kroz sve primjere, ali ovdje ćemo za trenutni primjer izračunati sve potrebne parcijalne derivacije i odmah ažurirati parametre na temelju te aproksimacije gradijenta funkcije pogreške.

2. zadatak

Programsku implementaciju sam stavio na Ferko, kao i ovaj dokument.

3. zadatak

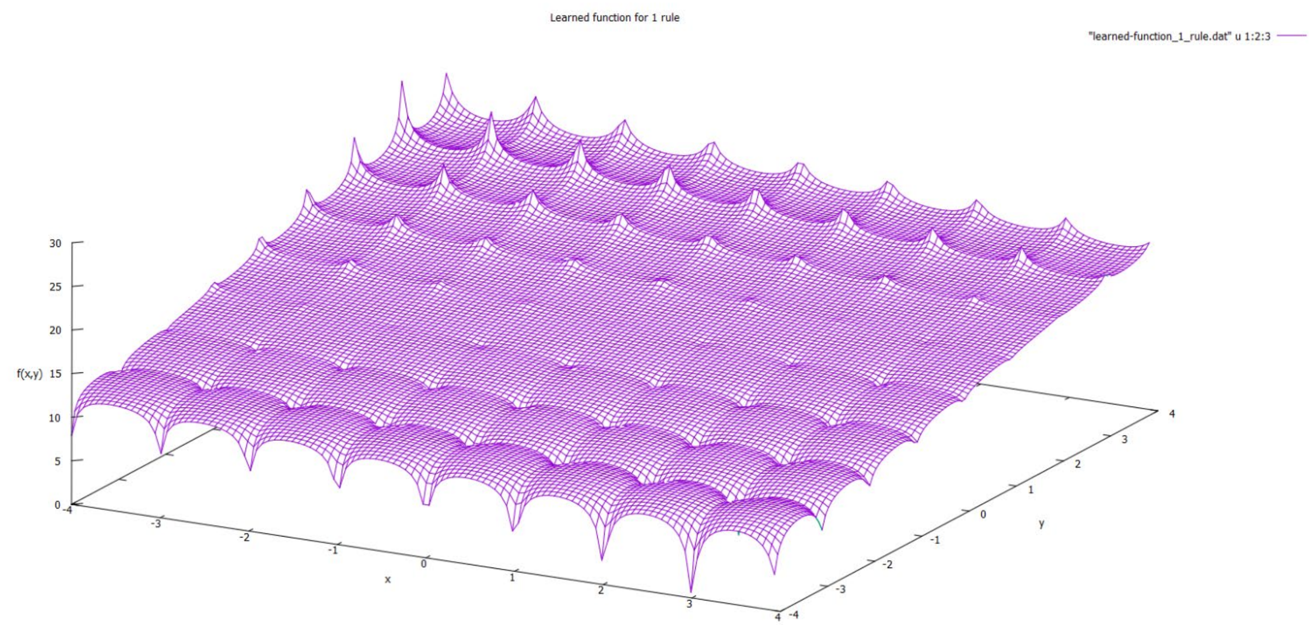
Izgled generiranog skupa podataka je na donjoj slici. Kako je zadana realna funkcija dvije varijable, graf se nalazi u 3-D prostoru. Može se vidjeti da funkcija teži izgledu plohe.



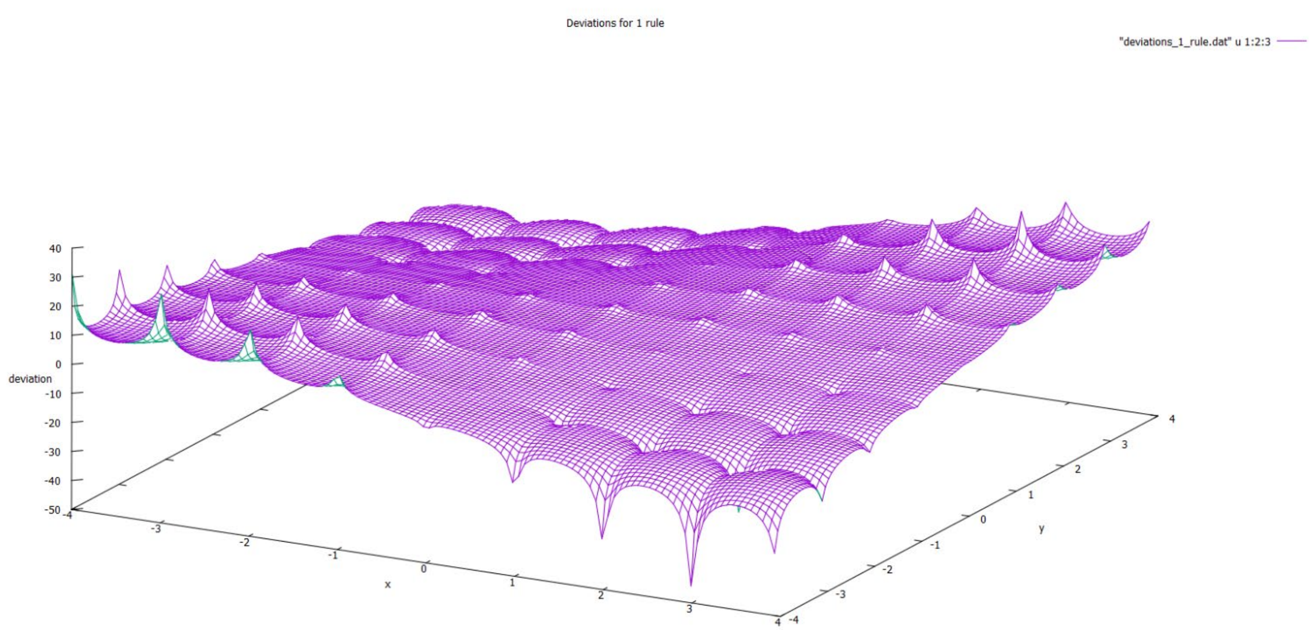
4. zadatak

a) Sustav od samo jednog pravila

Naučena funkcija:

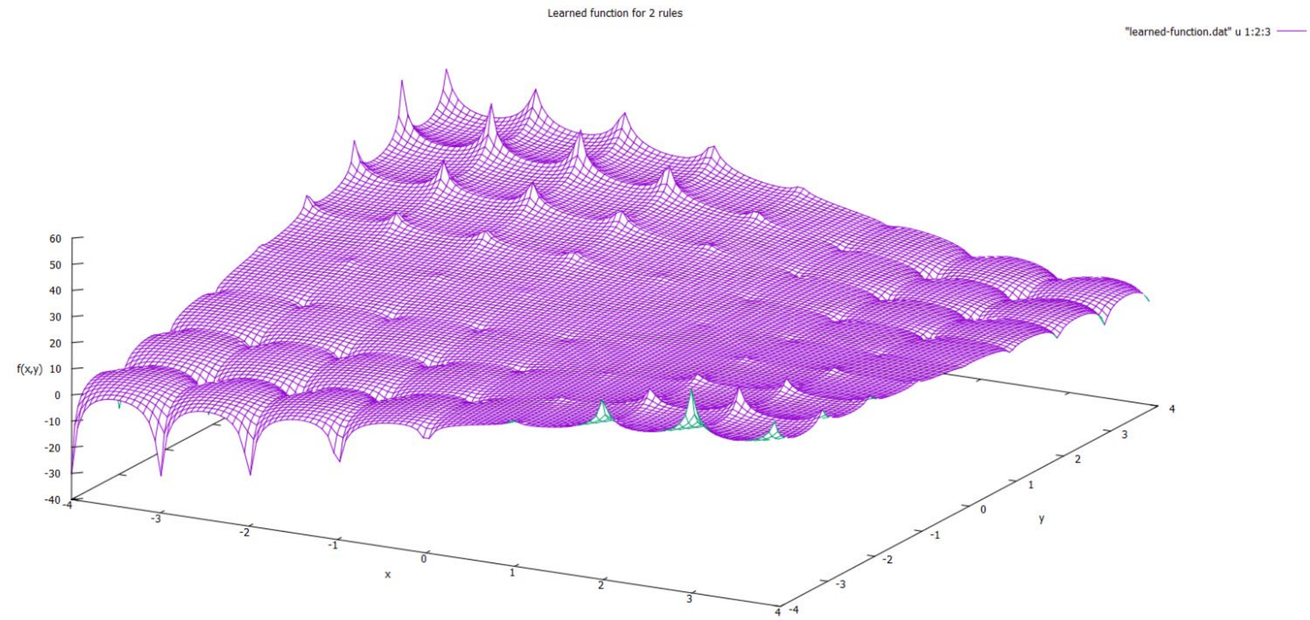


Odstupanja:

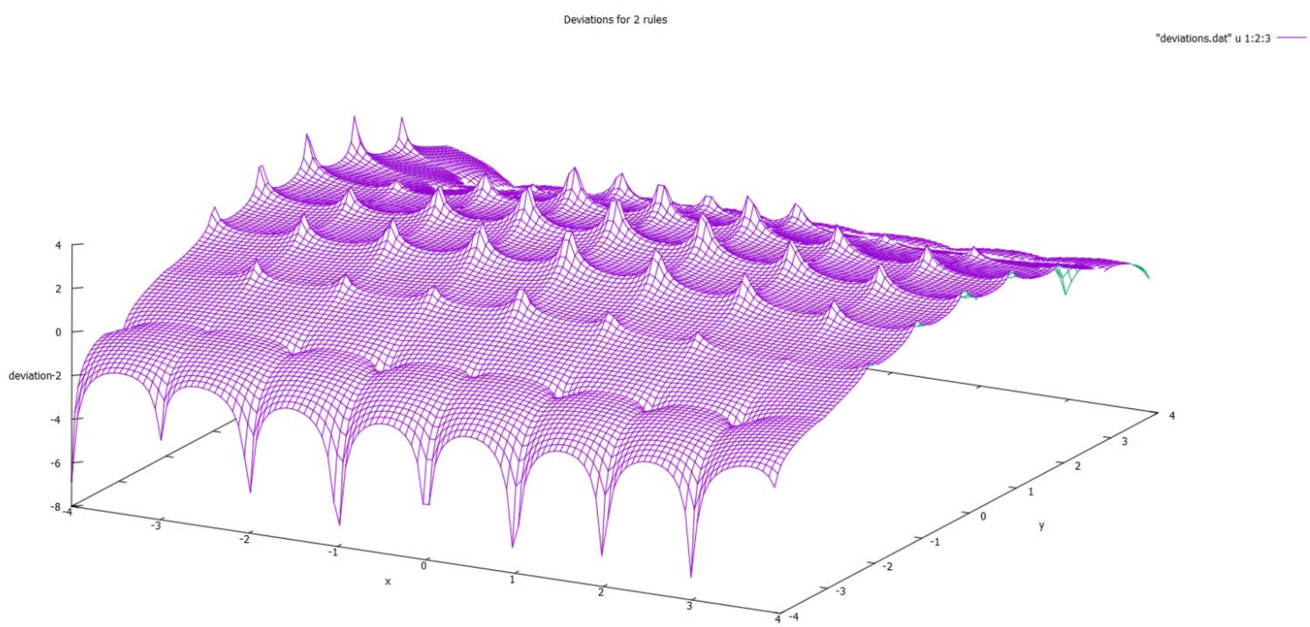


b) Sustav od dva pravila

Naučena funkcija:

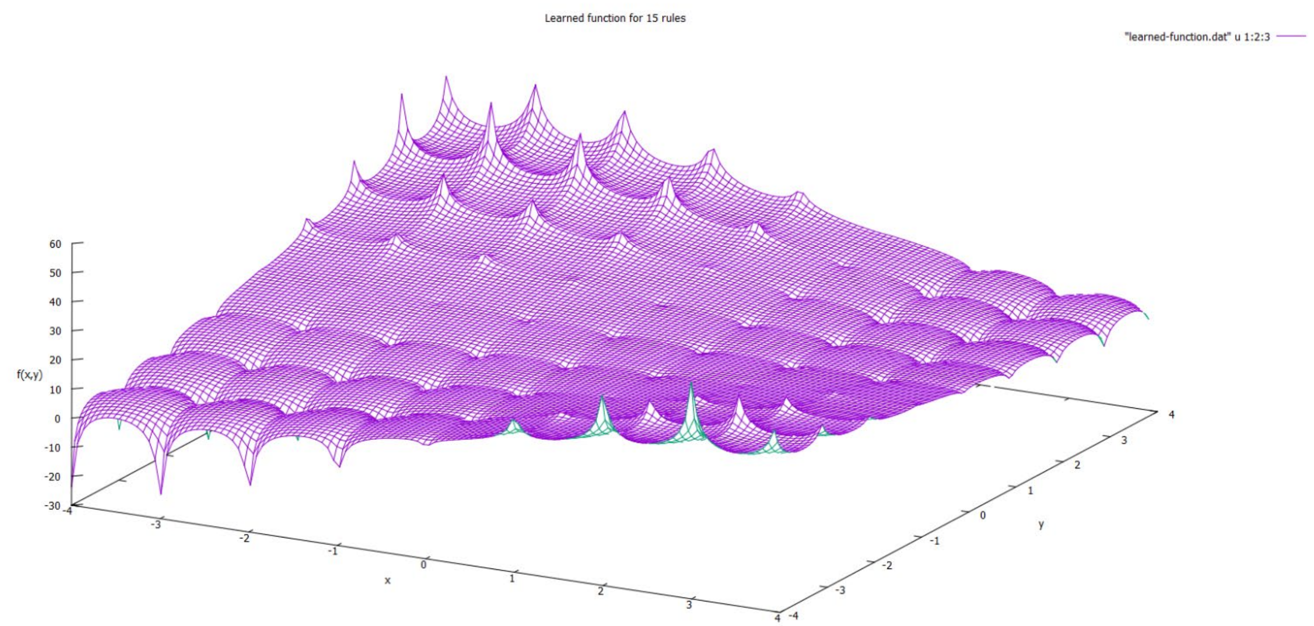


Odstupanja:

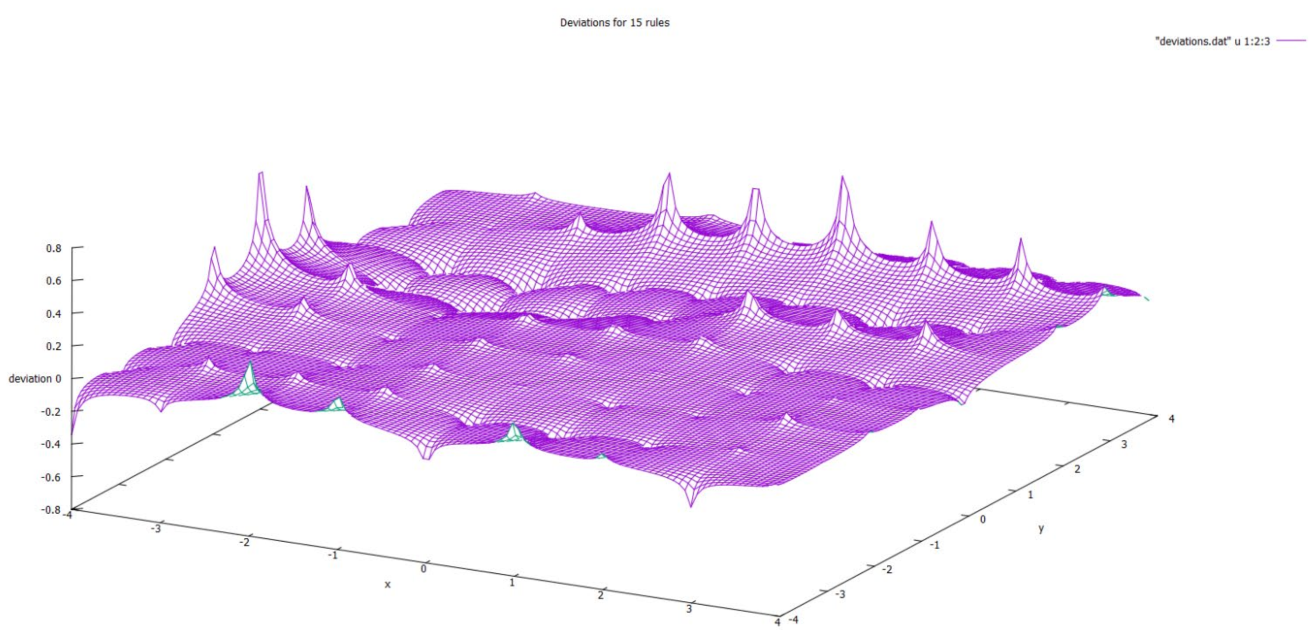


c) Sustav od prikladnog broja pravila (15 pravila kod mene)

Naučena funkcija:



Odstupanja:

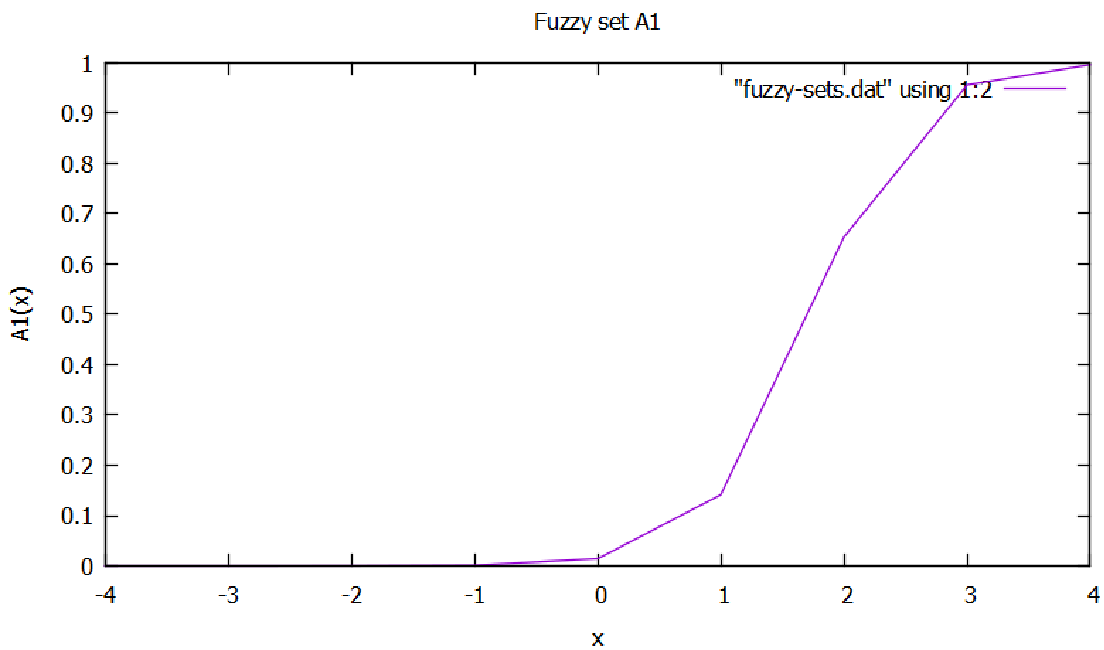


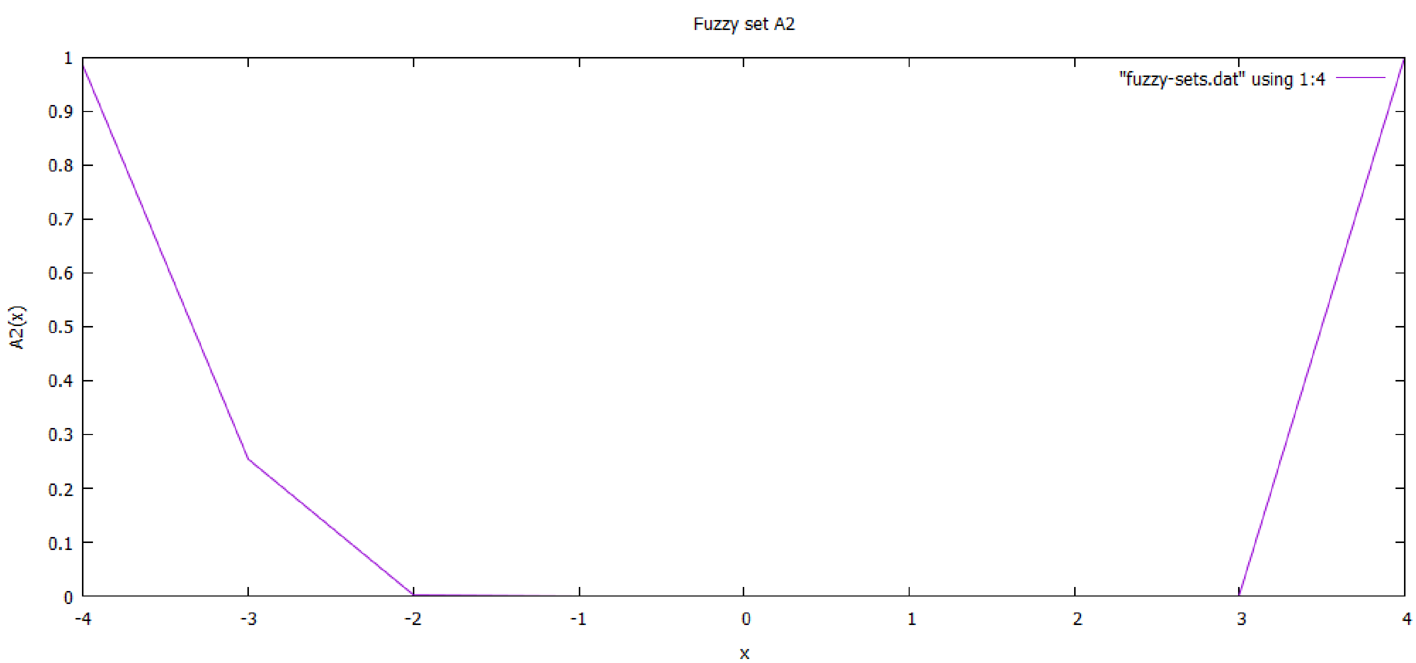
Lijepo se vidi kako za sve veći broj pravila odstupanja teže prema nuli. Za jedno i dva pravila nema značajne razlike u performansama grupne i online verzije algoritma. Za 15 pravila, online algoritam je pokazao bolji uspjeh od grupne verzije. Što se tiče stopi učenja, primijetio sam da je potrebno postaviti veću stopu učenja za parametre P, Q i R, a manje za A i B. Što se povećavao broj pravila, to su stope učenja trebale biti veće kako bi se postigao dobar rezultat. Ponekad mi se dogodilo da su stope učenja za oba skupa parametara bile jednake i da mi je to dalo optimalne rezultate.

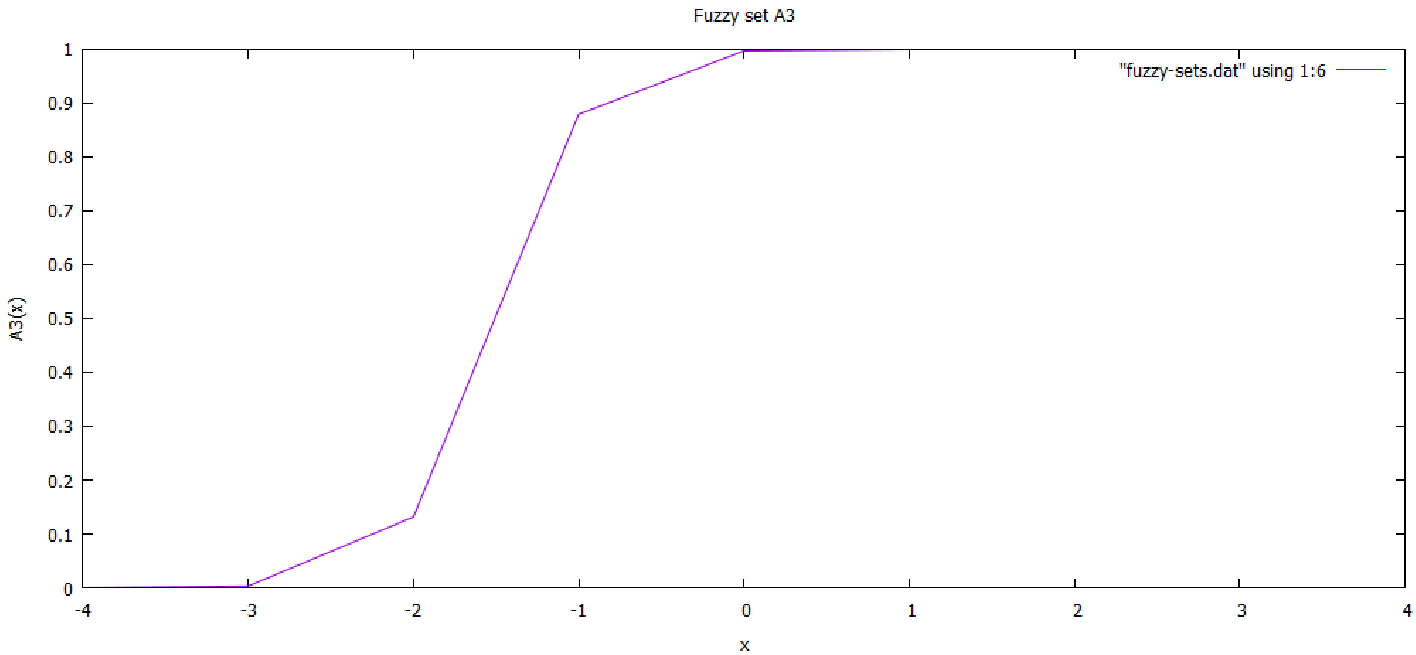
Ovaj zadatak je povezan s točkom 6, koja nije zadatak nego objašnjenje, pa zato nije niti posebno navedena kao zadatak u ovom dokumentu, već je implicitno ugrađena u ovaj zadatak.

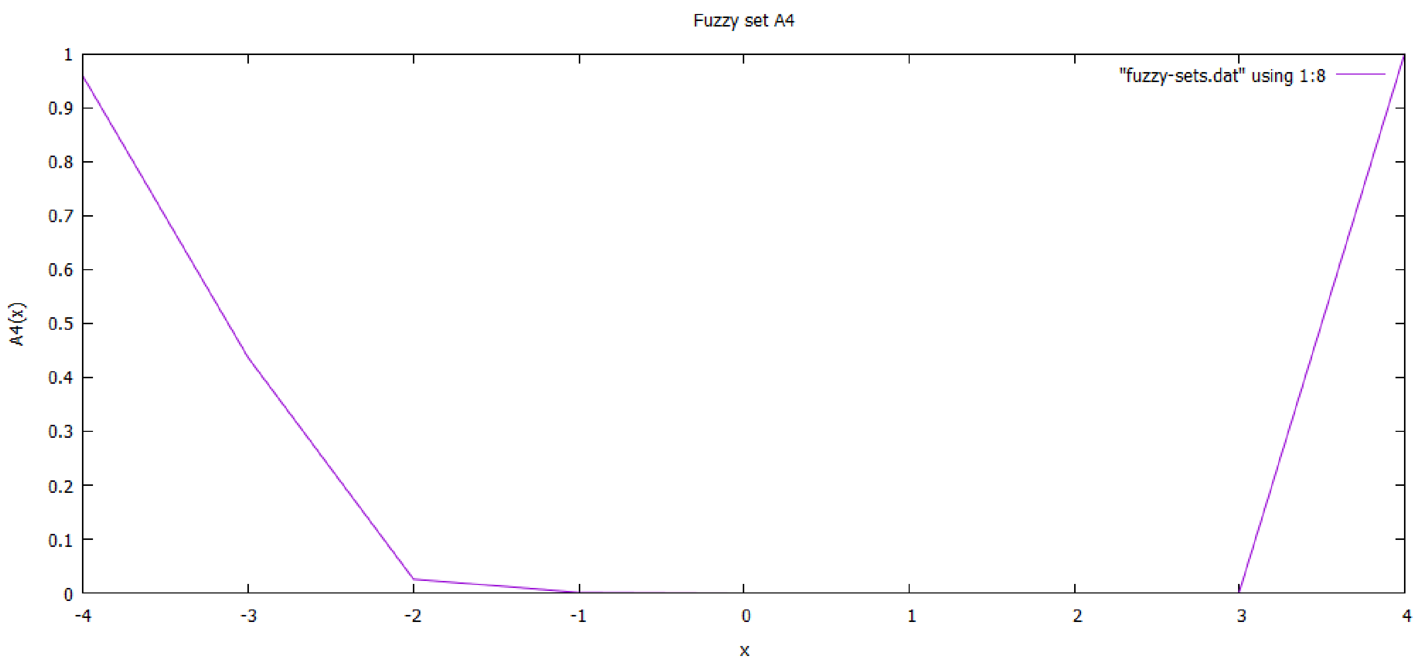
5. zadatak

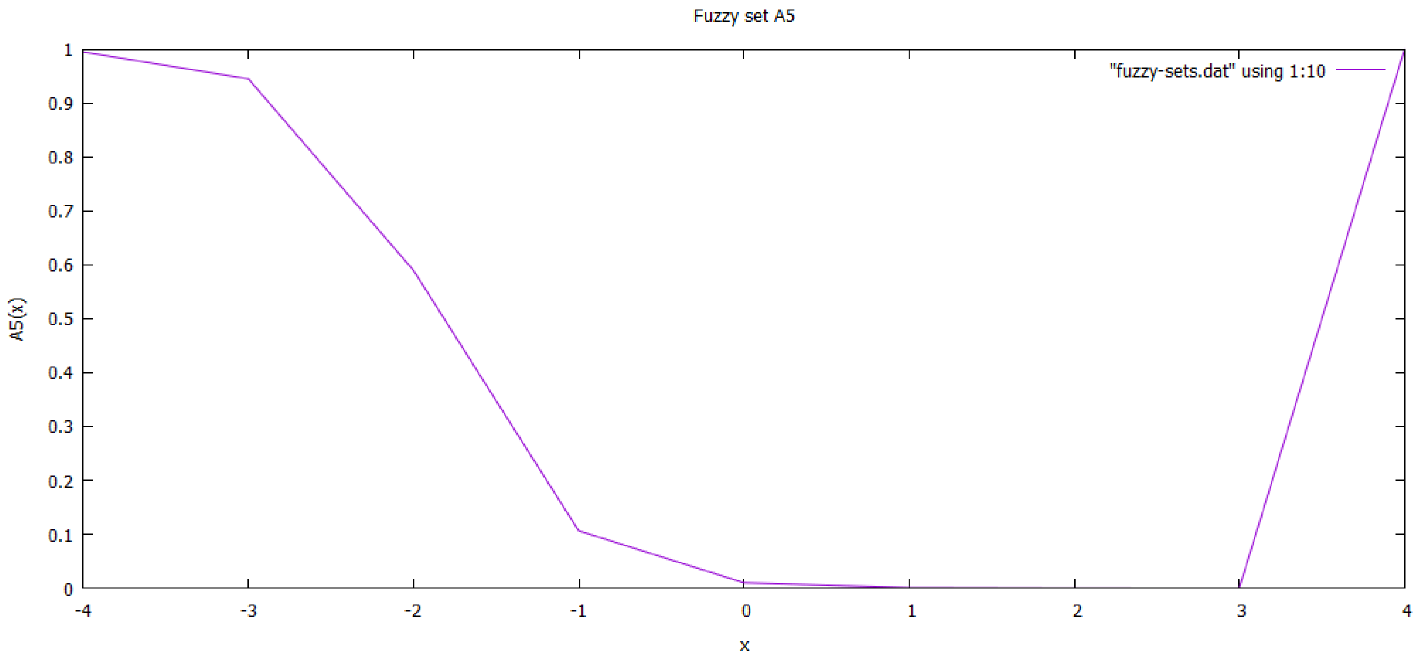
Sustav s kojim sam bio zadovoljan imao je 15 pravila. Međutim, to znači da je on naučio 30 funkcija pripadnosti (A1 do A15 te B1 do B15) i bilo bi previše sve ih ovdje staviti. Zato sam odabrao nekih 10 od tih 30 funkcija pripadnosti koje su prikazane u nastavku.

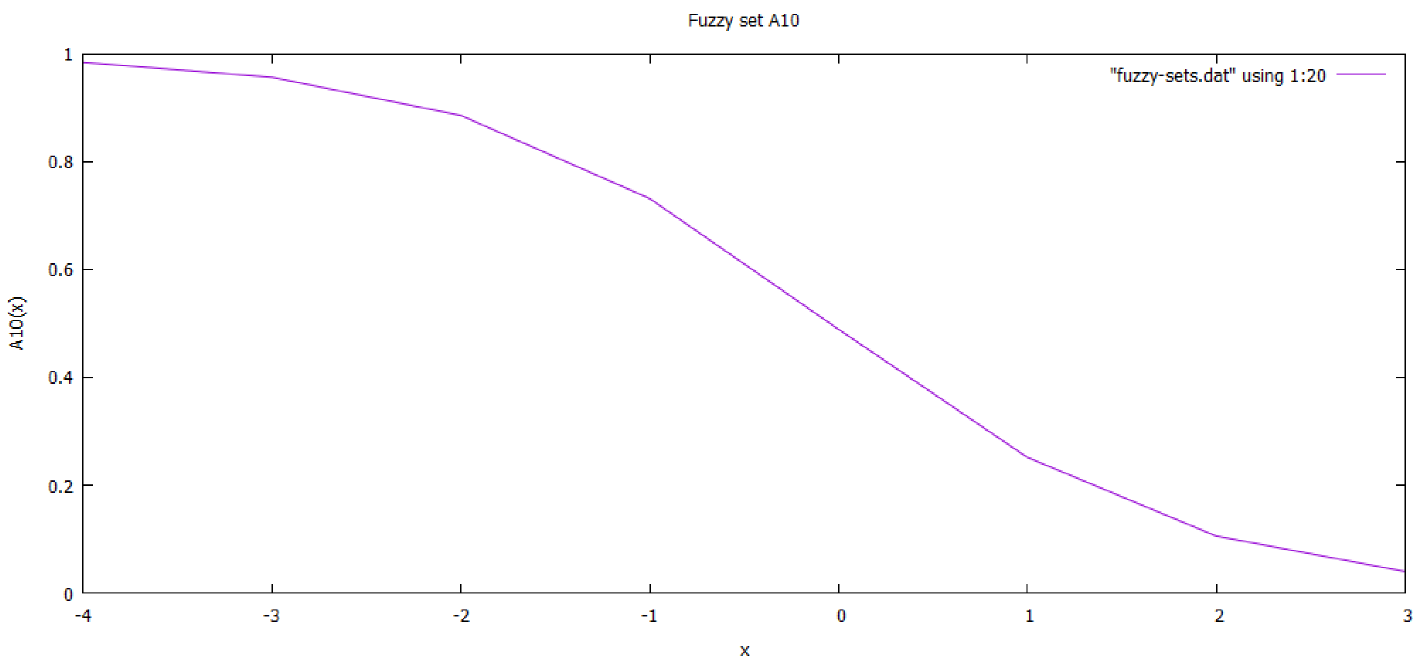


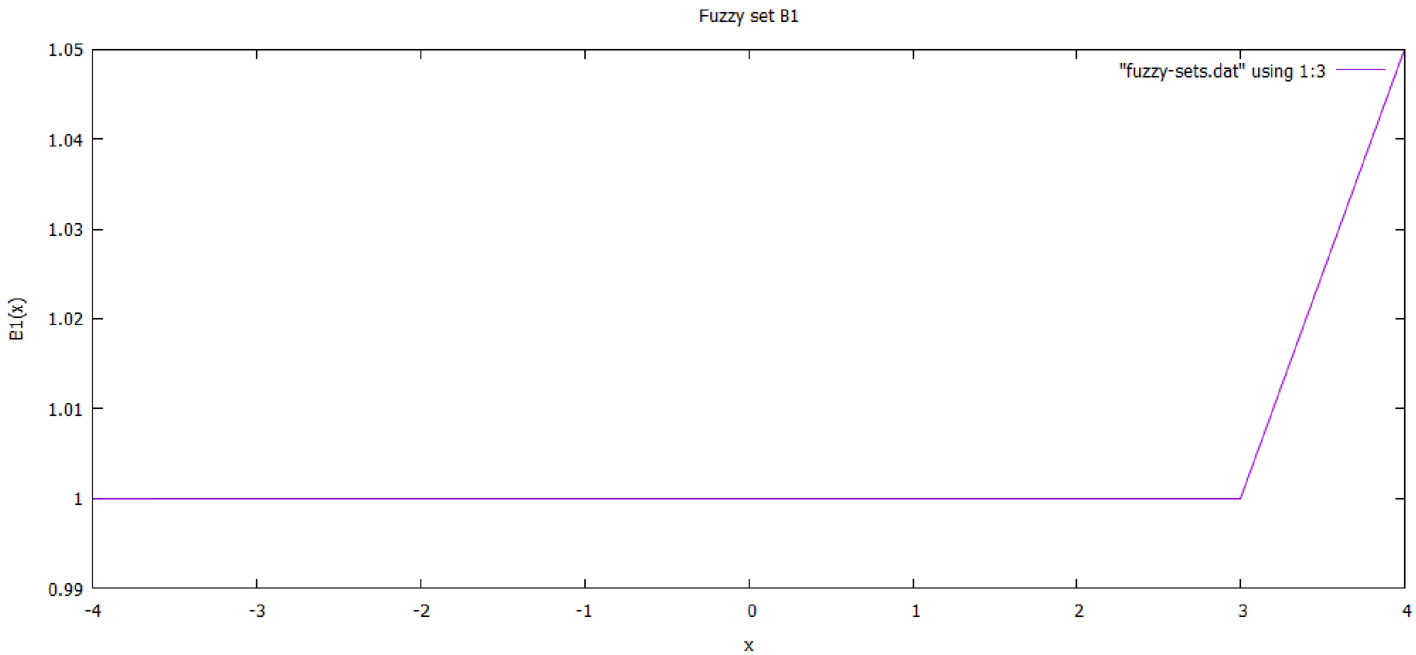



Fuzzy set A3
"fuzzy-sets.dat" using 1:6
A3(x)
x
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4

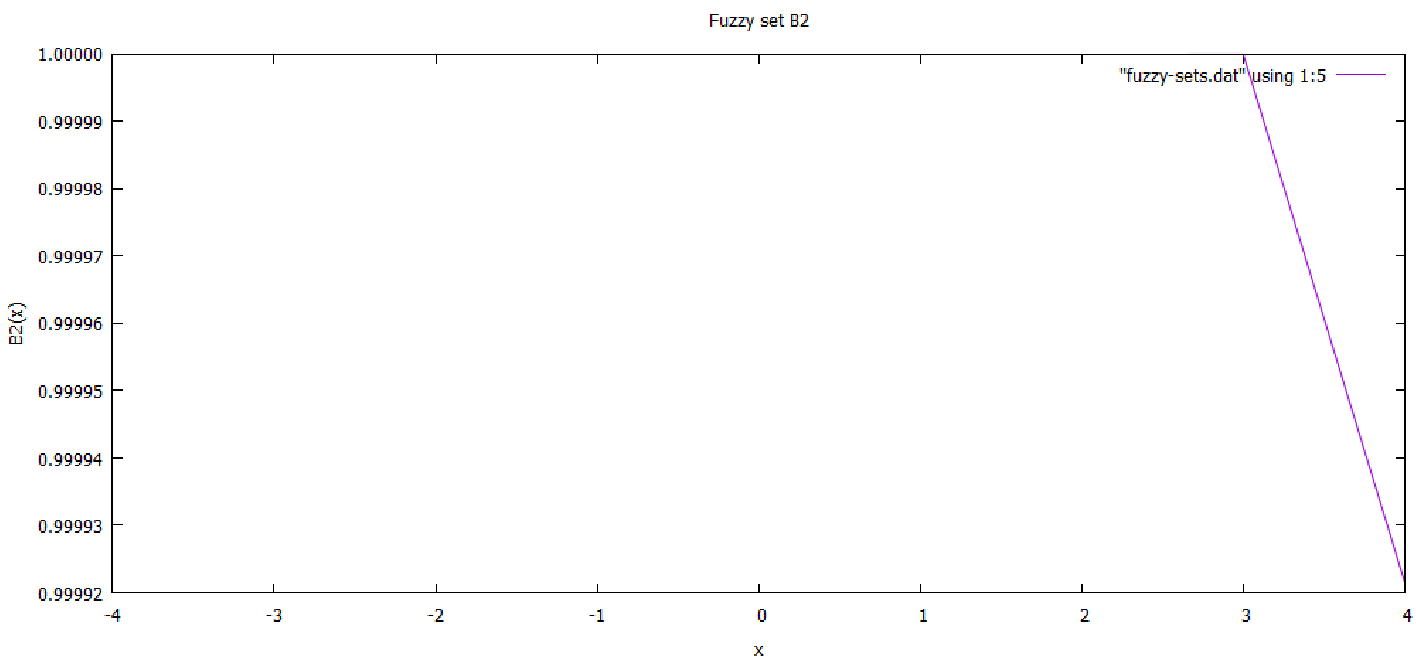

Fuzzy set A4
"fuzzy-sets.dat" using 1:8
A4(x)
x
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4


Fuzzy set A5
"fuzzy-sets.dat" using 1:10
A5(x)
x
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4

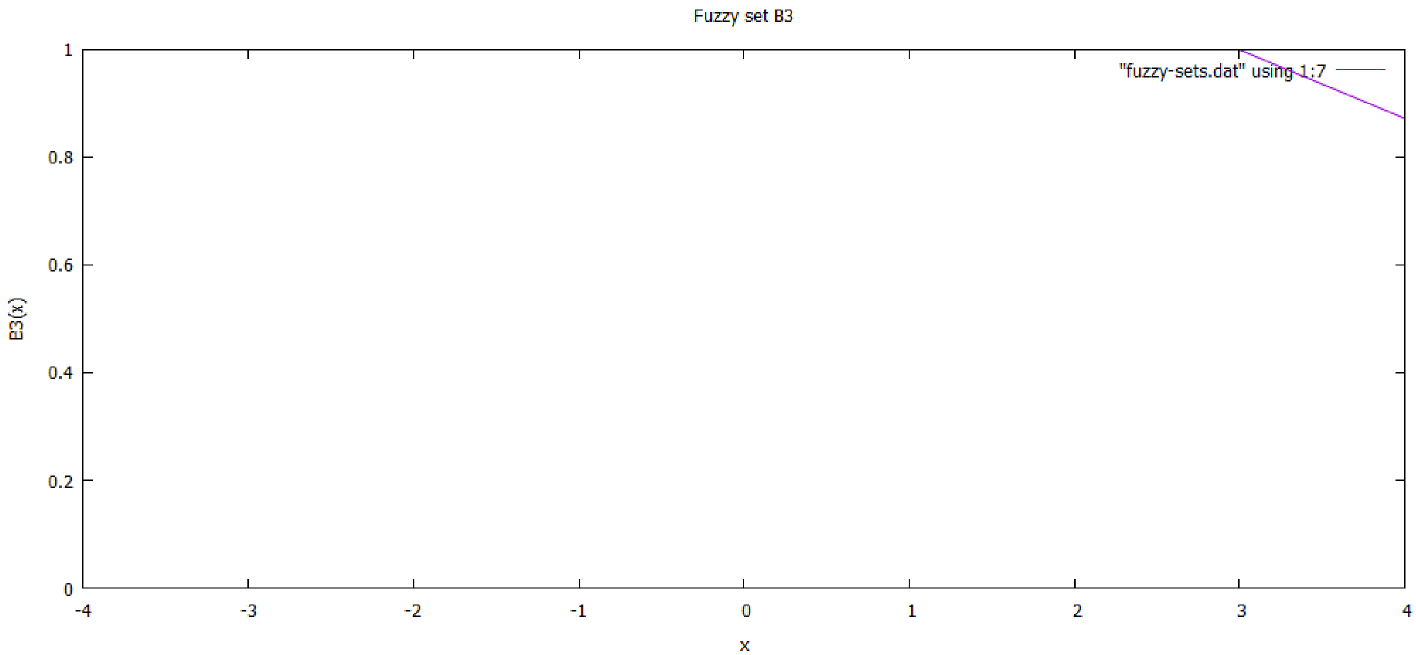

Fuzzy set A10
"fuzzy-sets.dat" using 1:20
A10(x)
x
1
0.8
0.6
0.4
0.2
0
-4
-3
-2
-1
0
1
2
3

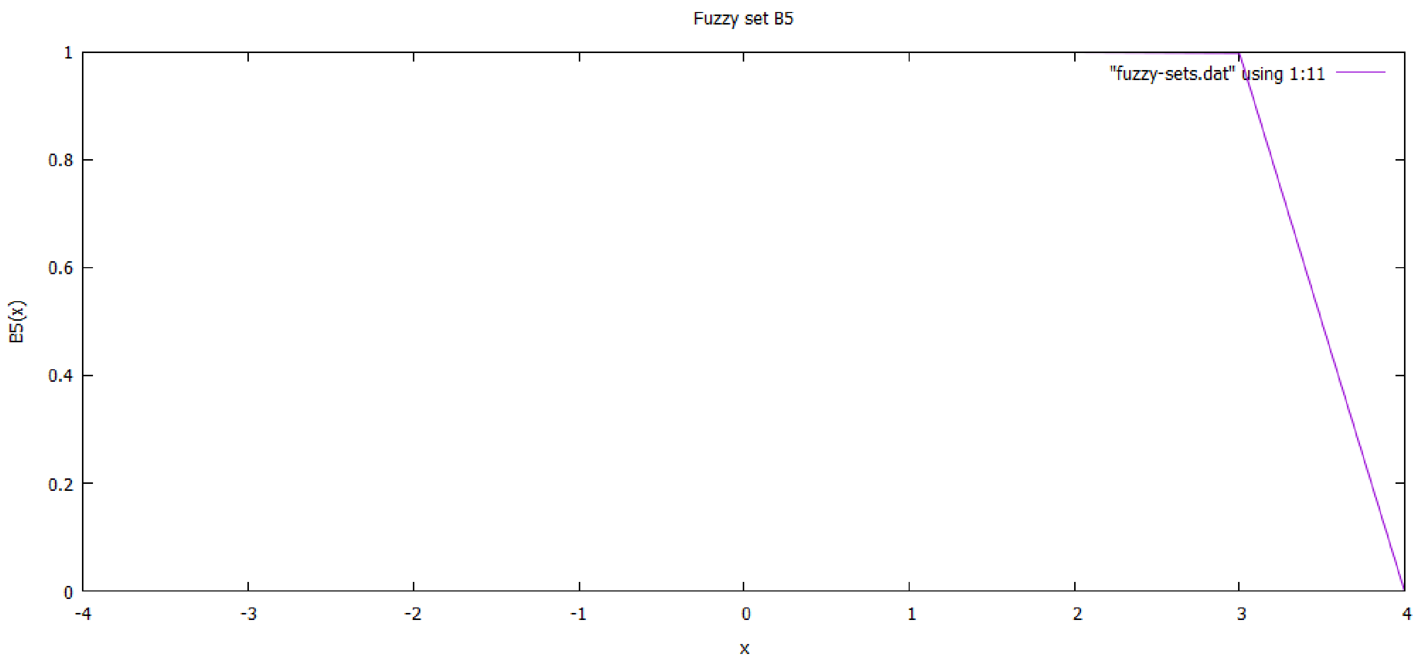


Fuzzy set B1
"fuzzy-sets.dat" using 1:3
B1(x)
1.05
1.04
1.03
1.02
1.01
1
0.99
-4
-3
-2
-1
0
1
2
3
4
x



Fuzzy set B2
"fuzzy-sets.dat" using 1:5
B2(x)
1.00000
0.99999
0.99998
0.99997
0.99996
0.99995
0.99994
0.99993
0.99992
-4
-3
-2
-1
0
1
2
3
4
x


Fuzzy set B3
"fuzzy-sets.dat" using 1:7
B3(x)
x
1
0.8
0.6
0.4
0.2
0
-4
-3
-2
-1
0
1
2
3
4

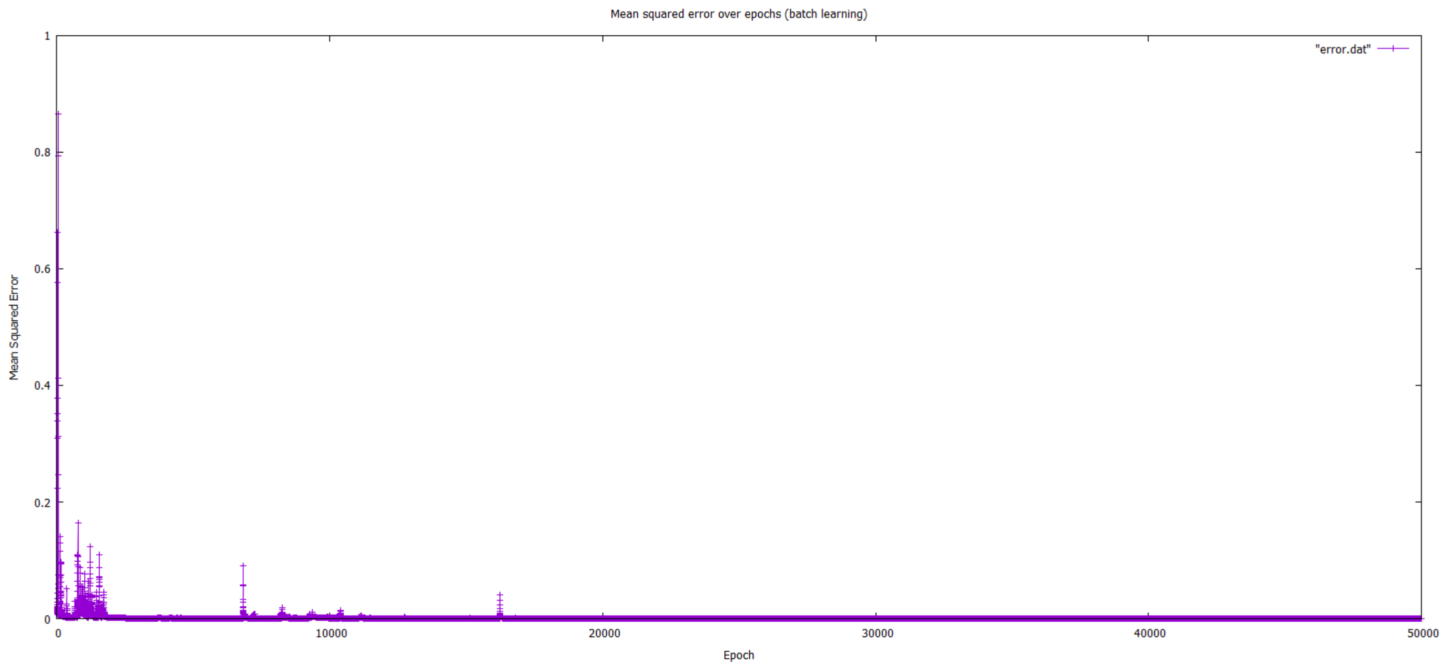

Fuzzy set B5
"fuzzy-sets.dat" using 1:11
B5(x)
x
1
0.8
0.6
0.4
0.2
0
-4
-3
-2
-1
0
1
2
3
4

Pokušat ću dati ime nekim od prikazanih neizrazitih skupova. Na primjer, neizraziti skup A5 mogao bi predstavljati „broj blizu ruba segmenta [-4, 4]“. Neizraziti skup A1 bi mogao predstavljati „broj blizu 4“.

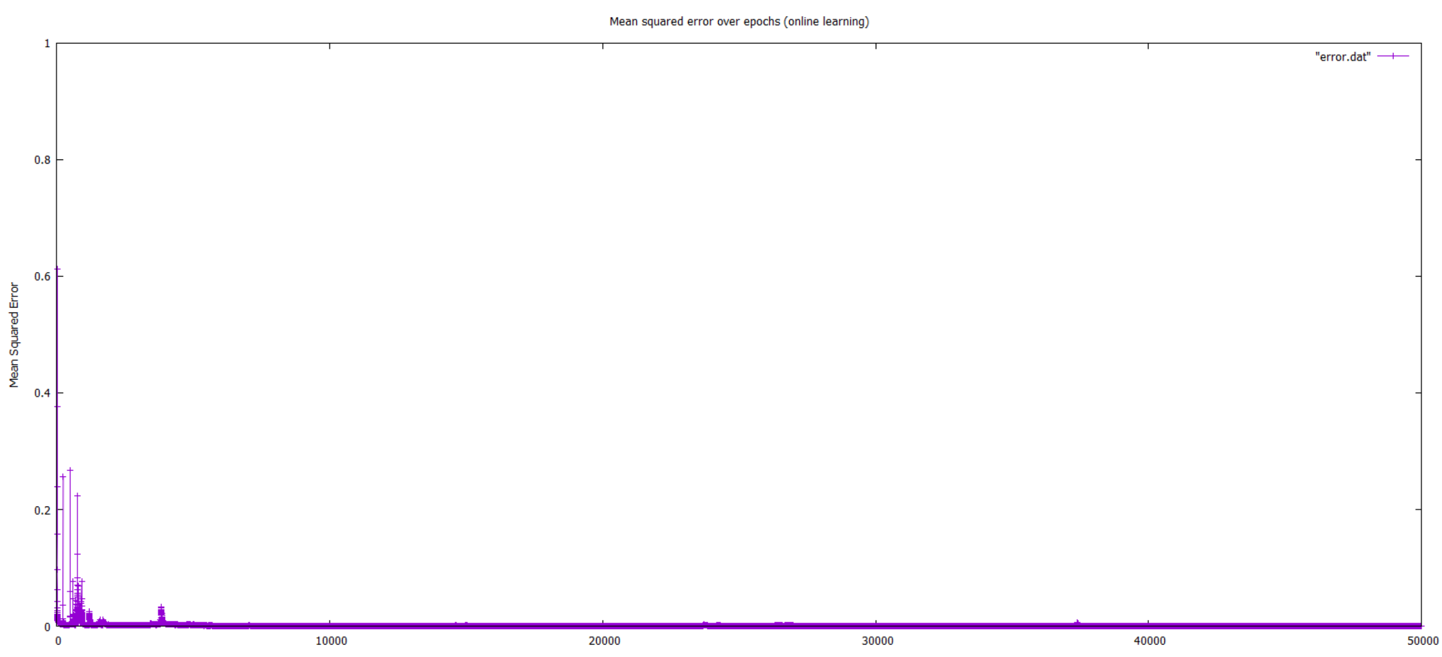
7. zadatak

Sljedeći grafovi prikazuju kretanje srednje kvadratne pogreške kroz epohe obje inačice algoritma. Radilo se sa sustavom od 15 pravila. Vrijednost pogreške je normalizirana na segment [0, 1].

Graf kretanja srednje kvadratne pogreške kroz epohe (grupni algoritam):



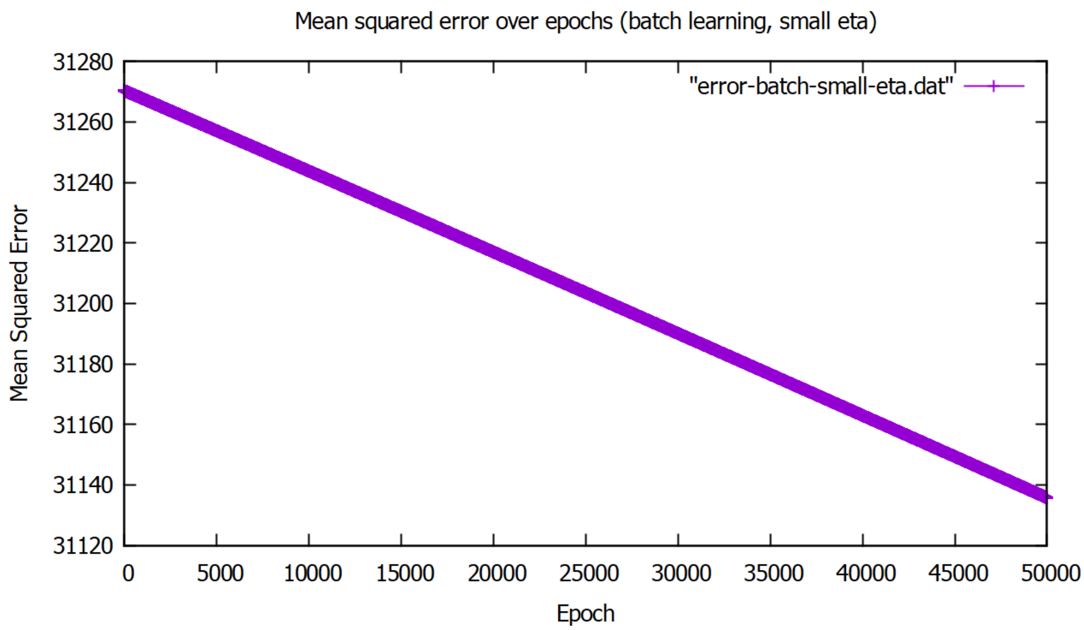
Graf kretanja srednje kvadratne pogreške kroz epohe (online algoritam):

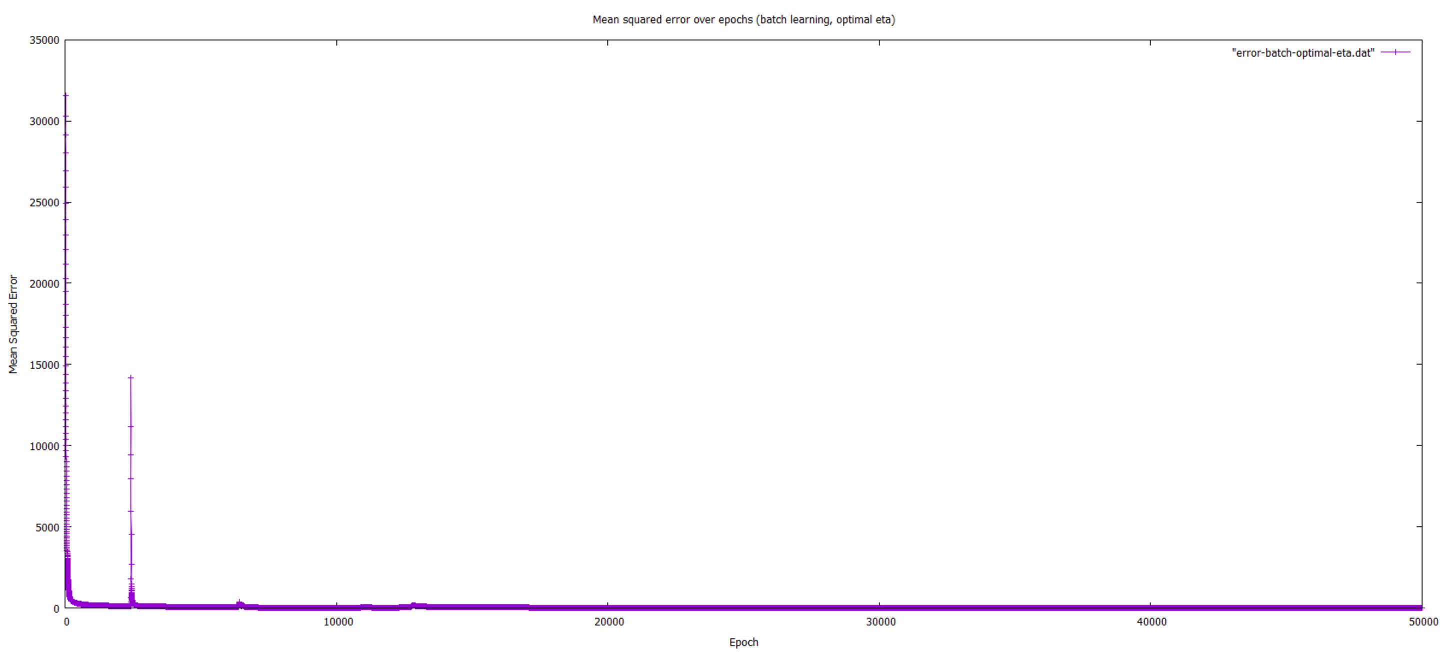


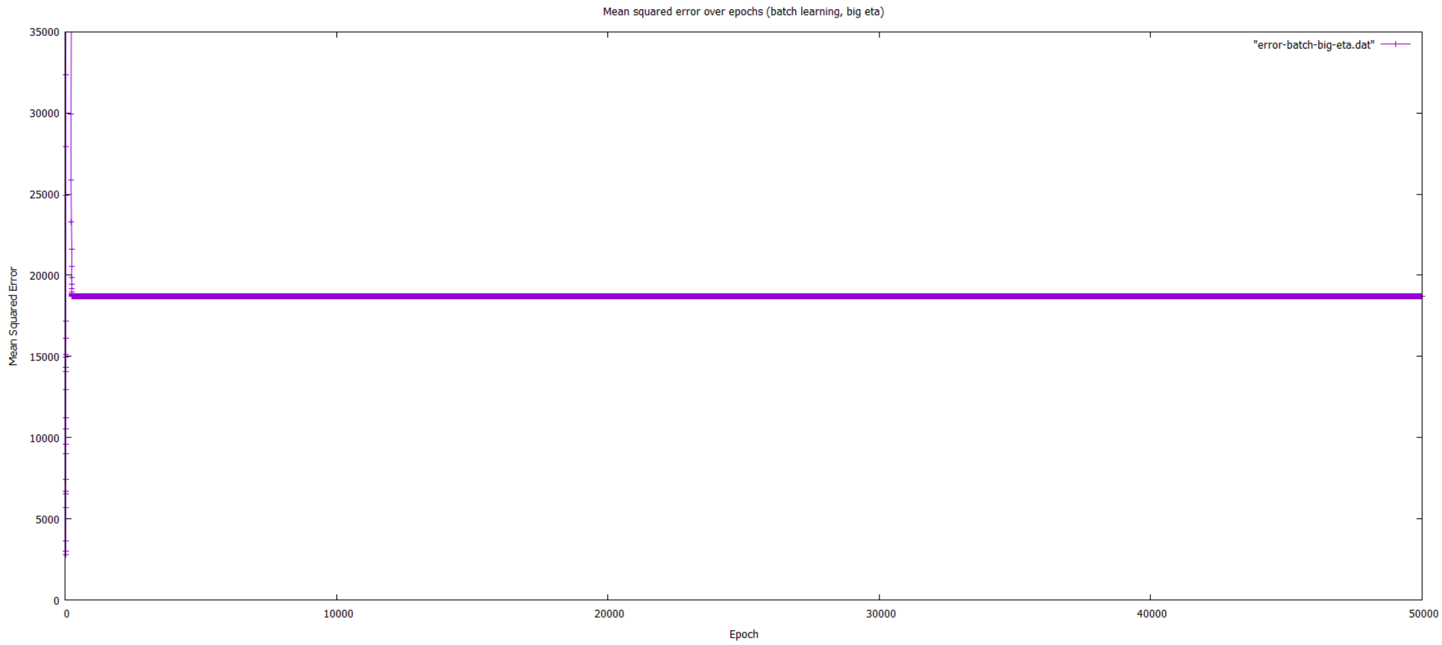
Bismo li htjeli usporediti ova dva grafa, između njih nema nikakve značajne razlike. Obje verzije algoritma su bile podjednako dobre.

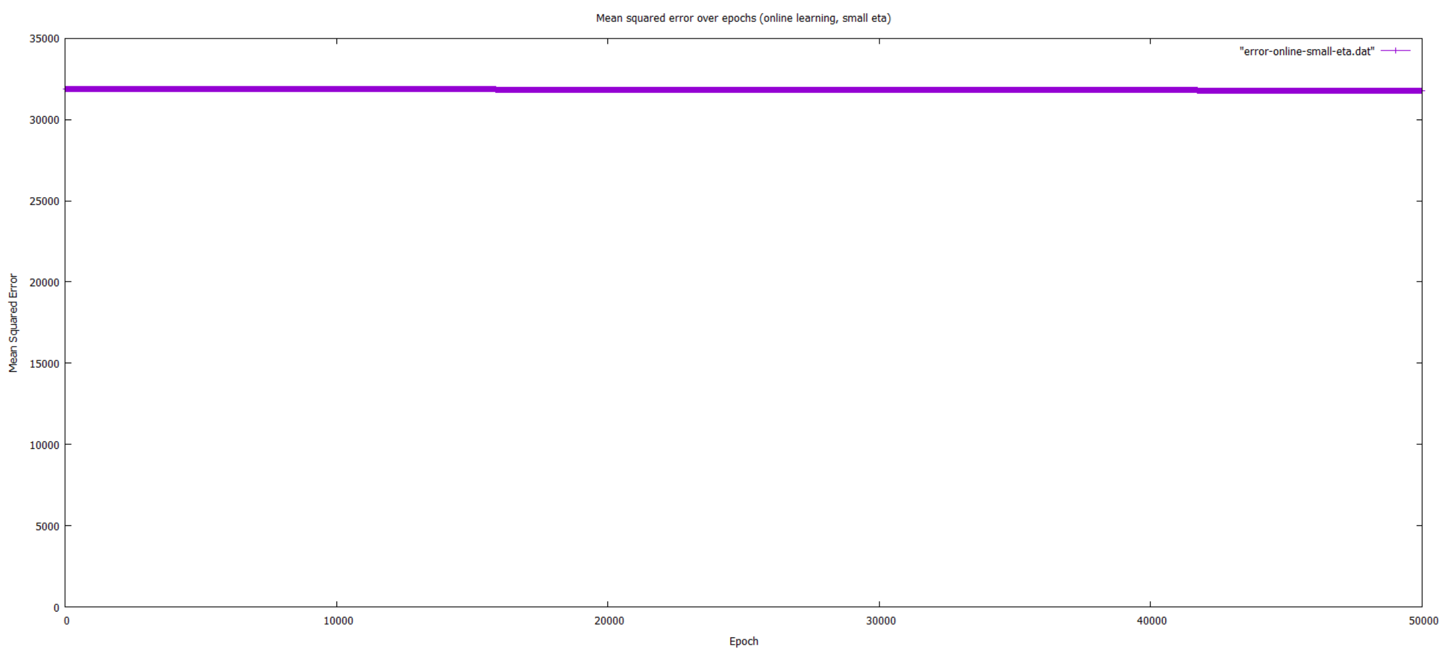
8. zadatak

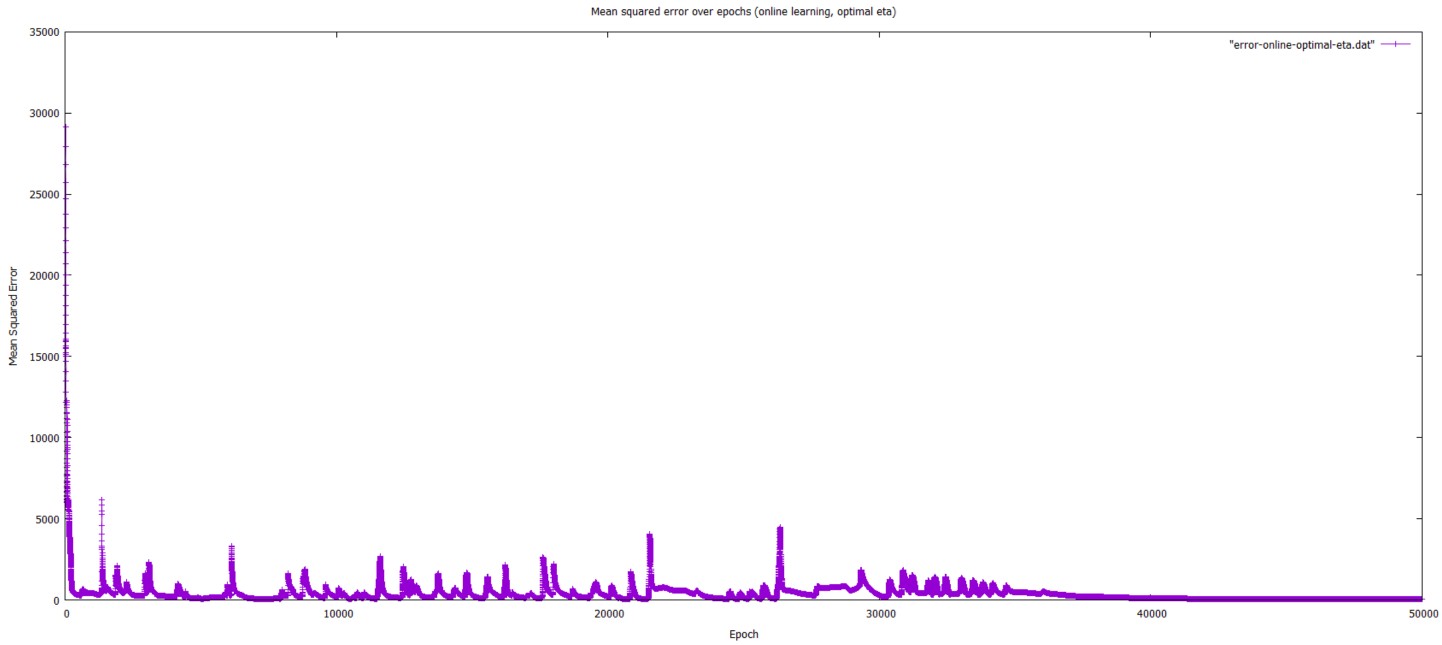
U ovom zadatku prikazujemo grafove kretanja srednje kvadratne pogreške kroz epohe obje verzije algoritma i za tri različite vrijednosti stopa učenja.

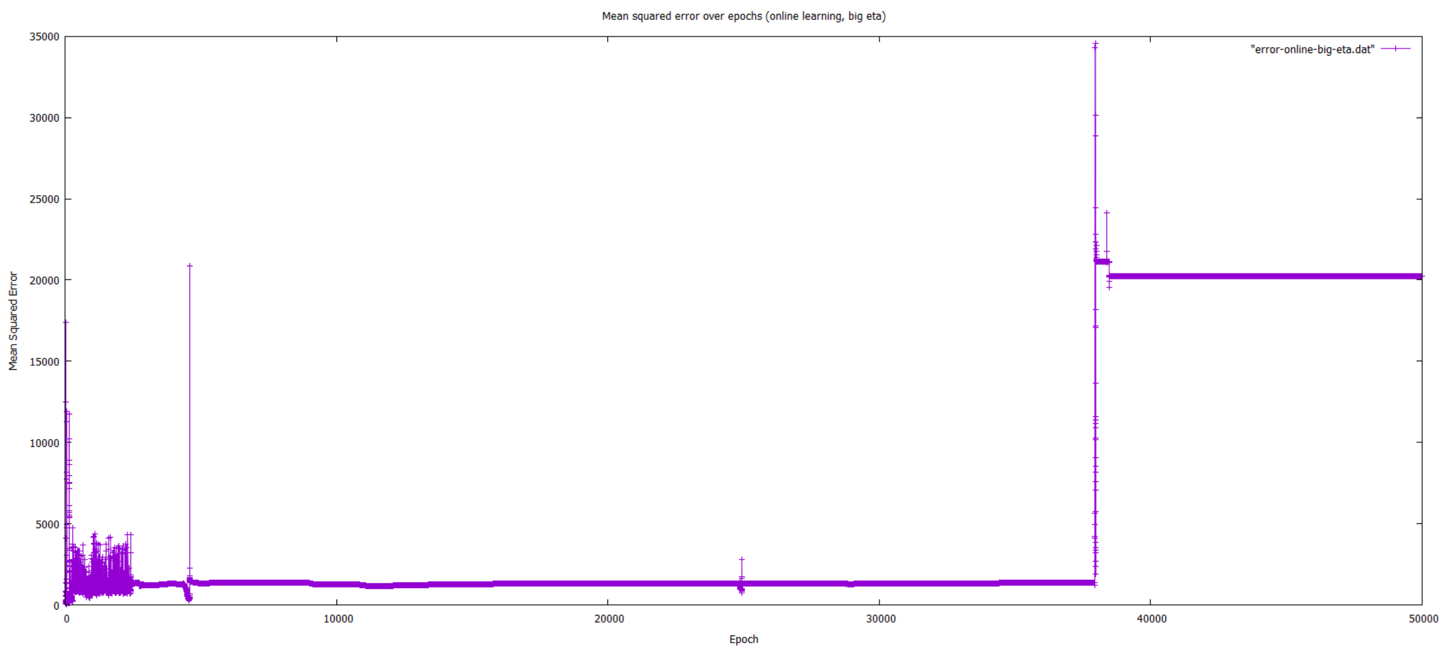





Mean squared error over epochs (batch learning, big eta)
"error-batch-big-eta.dat"
Mean Squared Error
Epoch
0
5000
10000
15000
20000
25000
30000
35000
0
10000
20000
30000
40000
50000


Mean squared error over epochs (online learning, small eta)
"error-online-small-eta.dat"
Mean Squared Error
Epoch
0
5000
10000
15000
20000
25000
30000
35000
0
10000
20000
30000
40000
50000





Može se vidjeti kako za premali iznos stope učenja online algoritam stagnira, a grupni jako sporo konvergira. Za preveliku stopu učenja algoritmi stagniraju, iako se online algoritam bolje ponaša od grupnog u tom slučaju. Da sam stavio još veće stope učenja, oba algoritma bi počela divergirati prema velikim vrijednostima srednje kvadratne pogreške. Za optimalnu vrijednost stop učenja jako se lijepo vidi kako je online verzija algoritma nestabilnija od grupne metode.