SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

NEIZRAZITO, EVOLUCIJSKO I NEURO RAČUNARSTVO

6. Domaća zadaća - Sustav ANFIS (engl. *Adaptive Neuro-Fuzzy Inference System*)

0036506587 Darijo Brčina

Zagreb, prosinac 2020.

SADRŽAJ

1. Zadatak 1

Funkcija pogreške za *k*-ti uzorak:

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

gdje y_k predstavlja željeni izlaz za *k*-ti uzorak, dok o_k predstavlja stvarni izlaz neuro-fuzzy sustava.

Ažuriranje parametara u skladu s algoritmom gradijentnog spusta:

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

Promatrat ćemo ANFIS koji koristi zaključivanje tipa 3 (metodu Takagi-Sugeno-Kang). Neuro-fuzzy sustav će imati dva ulaza i jedan izlaz čime će obavljati preslikavanje $\mathbb{R}^2 \to \mathbb{R}$. U općem slučaju, neuro-fuzzy sustav će raspolagati s *m* pravila:

$\mathbb{R}_1$: Ako x je A_1 i y je B_1 tada $z = p_1x + q_1y + r_1$

$\mathbb{R}_2$: Ako x je A_2 i y je B_2 tada $z = p_2x + q_2y + r_2$

...

$\mathbb{R}_m$: Ako x je A_m i y je B_m tada $z = p_mx + q_my + r_m$

Uz zadani skup pravila od *m* pravila, temeljem primjera za učenje sustav treba naučiti ukupno $2 \cdot m$ neizrazitih skupova koji se nalaze u antecedent dijelu pravila i $3 \cdot m$ parametara koji određuju *m* funkcija u konzekvens dijelovima pravila.

Neizraziti skupovi *A* i *B* su definirani sljedećim funkcijama pripadnosti:

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

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

Iz navedenog se može zaključiti da je potrebno naučiti sveukupno $7 \cdot m$ parametara sustava.

Stvarni izlaz neuro-fuzzy sustava je definiran na sljedeći način:

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

gdje je $\pi_i = \alpha_i \cdot \beta_i$. Sada kada znamo kako je sve definirano, možemo izvesti formule za ažuriranje svih parametara tijekom učenja algoritmom gradijenti spust, pa krenimo:

1.
$$\frac{\partial E_k}{\partial a_i} = \frac{\partial E_k}{\partial o_k}\frac{\partial o_k}{\partial \pi_i}\frac{\partial \pi_i}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial a_i}$$
$$\frac{\partial E_k}{\partial o_k} = \frac{\partial}{\partial o_k}(\frac{1}{2}(y_k - o_k)^2) = 2 \cdot \frac{1}{2}(y_k - o_k) \cdot (-1) = -(y_k - o_k)$$
$$\frac{\partial o_k}{\partial \pi_i} = \frac{\partial}{\partial \pi_i}(\frac{\sum_{j=1}^m \pi_j z_j}{\sum_{j=1}^m \pi_j}) = \frac{z_i \sum_{j=1,j\neq i}^m \pi_j - \sum_{j=1,j\neq i}^m \pi_j z_j \cdot 1}{(\sum_{j=1}^m \pi_j)^2} = \frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}$$
$$\frac{\partial \pi_i}{\partial \alpha_i} = \frac{\partial}{\partial \alpha_i}(\alpha_i \cdot \beta_i) = \beta_i$$
$$\frac{\partial \alpha_i}{\partial a_i} = \frac{\partial}{\partial a_i}(\frac{1}{1 + e^{b_i(x-a_i)}}) = (-1) \cdot \frac{1}{(1 + e^{b_i(x-a_i)})^2} \cdot e^{b_i(x-a_i)} \cdot (-b_i) = \alpha_i(1-\alpha_i) \cdot b_i$$
$$\frac{\partial E_k}{\partial a_i} = -(y_k - o_k)\frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}\beta_i\alpha_i(1-\alpha_i)b_i$$
$$\implies a_i(t+1) = a_i(t) + \eta \cdot (y_k - o_k)\frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}\beta_i\alpha_i(1-\alpha_i)b_i$$

2.
$$\frac{\partial E_k}{\partial b_i} = \frac{\partial E_k}{\partial o_k}\frac{\partial o_k}{\partial \pi_i}\frac{\partial \pi_i}{\partial \alpha_i}\frac{\partial \alpha_i}{\partial b_i}$$
$$\frac{\partial E_k}{\partial o_k} = -(y_k - o_k)$$
$$\frac{\partial o_k}{\partial \pi_i} = \frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}$$
$$\frac{\partial \pi_i}{\partial \alpha_i} = \beta_i$$
$$\frac{\partial \alpha_i}{\partial b_i} = \frac{\partial}{\partial b_i}(\frac{1}{1 + e^{b_i(x-a_i)}}) = (-1) \cdot \frac{1}{(1 + e^{b_i(x-a_i)})^2} \cdot e^{b_i(x-a_i)} \cdot (x-a_i) = \alpha_i(1-\alpha_i) \cdot (a_i - x)$$
$$\frac{\partial E_k}{\partial b_i} = -(y_k - o_k)\frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}\beta_i\alpha_i(1-\alpha_i)(a_i - x)$$
$$\implies b_i(t+1) = b_i(t) + \eta \cdot (y_k - o_k)\frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}\beta_i\alpha_i(1-\alpha_i)(a_i - x)$$

3. Po uzoru na ažuriranje parametra *a* uz $\frac{\partial \pi_i}{\partial \beta_i} = \alpha_i$:

$$\implies c_i(t+1) = c_i(t) + \eta \cdot (y_k - o_k)\frac{\sum_{j=1,j\neq i}^m \pi_j(z_i - z_j)}{(\sum_{j=1}^m \pi_j)^2}\alpha_i\beta_i(1-\beta_i)d_i$$

4. Po uzoru na ažuriranje parametra *b* uz $\frac{\partial \pi_i}{\partial \beta_i} = \alpha_i$:

$$\Longrightarrow d_i(t+1) = d_i(t) + \eta \cdot (y_k - o_k) \frac{\sum_{j=1, j \neq i}^{m} \pi_j (z_i - z_j)}{(\sum_{j=1}^{m} \pi_j)^2} \alpha_i \beta_i (1 - \beta_i)(c_i - y)$$

5.

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

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

$$\frac{\partial o_k}{\partial p_i} = \frac{\partial}{\partial p_i} \left(\frac{\sum_{j=1}^{m} \pi_j z_j}{\sum_{j=1}^{m} \pi_j} \right) = \frac{\partial}{\partial p_i} \left(\frac{\sum_{j=1}^{m} \pi_j (p_j x + q_j y + r_j)}{\sum_{j=1}^{m} \pi_j} \right) = \frac{\pi_i}{\sum_{j=1}^{m} \pi_j} x$$

$$\frac{\partial E_k}{\partial p_i} = -(y_k - o_k) \frac{\pi_i}{\sum_{j=1}^{m} \pi_j} x$$

$$\Longrightarrow p_i(t+1) = p_i(t) + \eta \cdot (y_k - o_k) \frac{\pi_i}{\sum_{j=1}^{m} \pi_j} x$$

6.

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

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

$$\frac{\partial o_k}{\partial q_i} = \frac{\partial}{\partial q_i} \left(\frac{\sum_{j=1}^{m} \pi_j z_j}{\sum_{j=1}^{m} \pi_j} \right) = \frac{\partial}{\partial q_i} \left(\frac{\sum_{j=1}^{m} \pi_j (p_j x + q_j y + r_j)}{\sum_{j=1}^{m} \pi_j} \right) = \frac{\pi_i}{\sum_{j=1}^{m} \pi_j} y$$

$$\frac{\partial E_k}{\partial q_i} = -(y_k - o_k) \frac{\pi_i}{\sum_{j=1}^{m} \pi_j} y$$

$$\Longrightarrow q_i(t+1) = q_i(t) + \eta \cdot (y_k - o_k) \frac{\pi_i}{\sum_{j=1}^{m} \pi_j} y$$

7.

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

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

$$\frac{\partial o_k}{\partial r_i} = \frac{\partial}{\partial r_i} \left(\frac{\sum_{j=1}^{m} \pi_j z_j}{\sum_{j=1}^{m} \pi_j} \right) = \frac{\partial}{\partial q_i} \left(\frac{\sum_{j=1}^{m} \pi_j (p_j x + q_j y + r_j)}{\sum_{j=1}^{m} \pi_j} \right) = \frac{\pi_i}{\sum_{j=1}^{m} \pi_j}$$

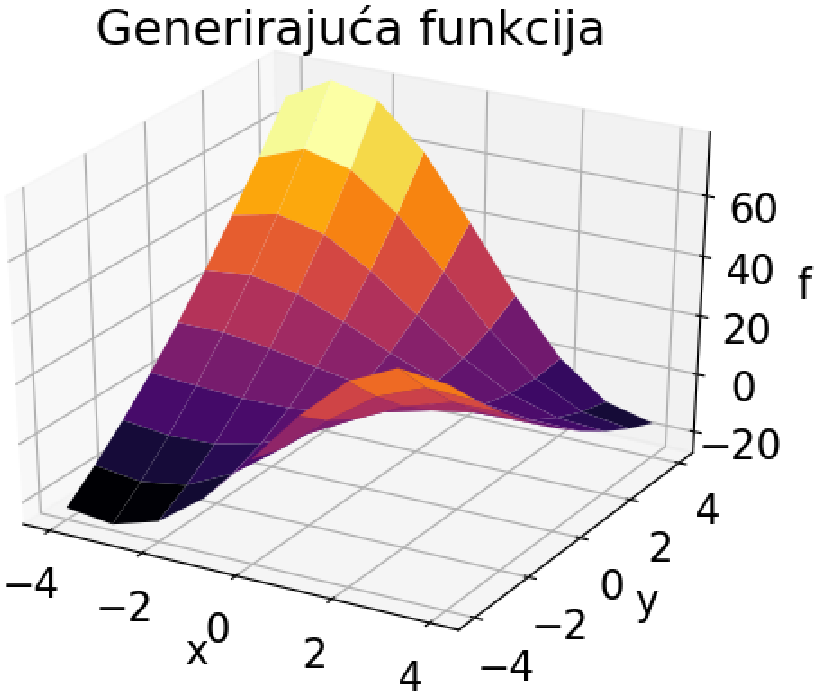
$$\frac{\partial E_k}{\partial r_i} = -(y_k - o_k) \frac{\pi_i}{\sum_{j=1}^{m} \pi_j}$$

$$\Longrightarrow r_i(t+1) = r_i(t) + \eta \cdot (y_k - o_k) \frac{\pi_i}{\sum_{j=1}^{m} \pi_j}$$

Napomena: Ažuriranje parametara je izvedeno za *online* tip učenja, dok bi se za pravi gradijent morala uzimati suma pogrešaka svih uzoraka za učenje tijekom pronalaska parcijalnih derivacija, tj.

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

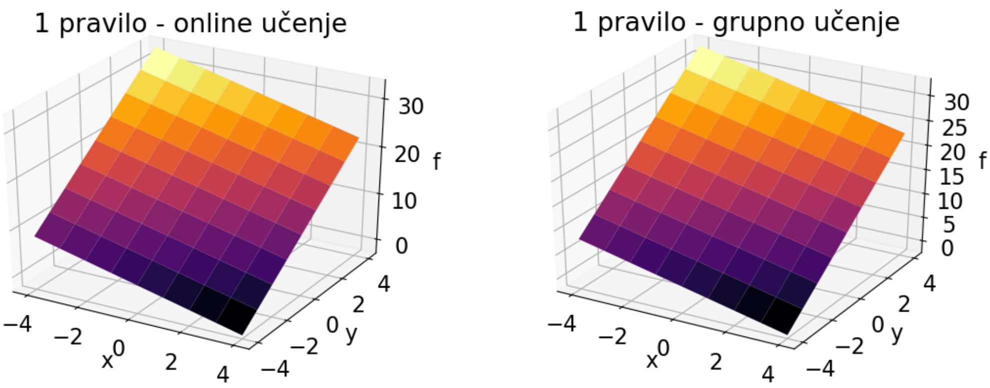
2. Zadatak 3



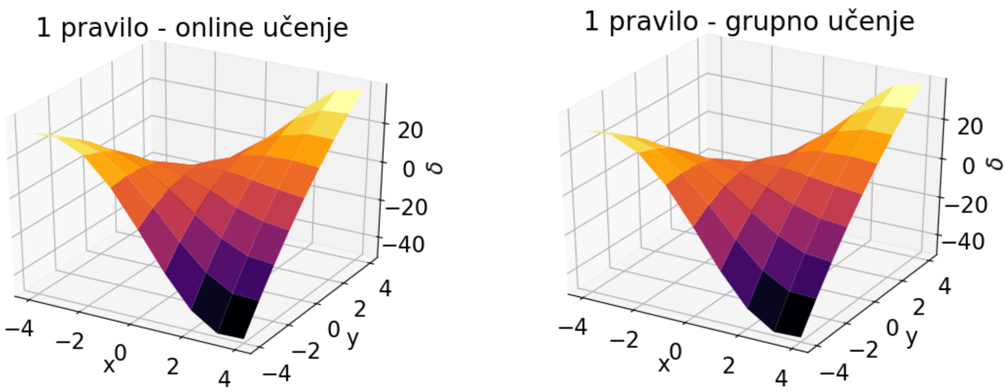
Slika 2.1: Graf generirajuće funkcije

3. Zadatak 4

3.1. Jedno pravilo

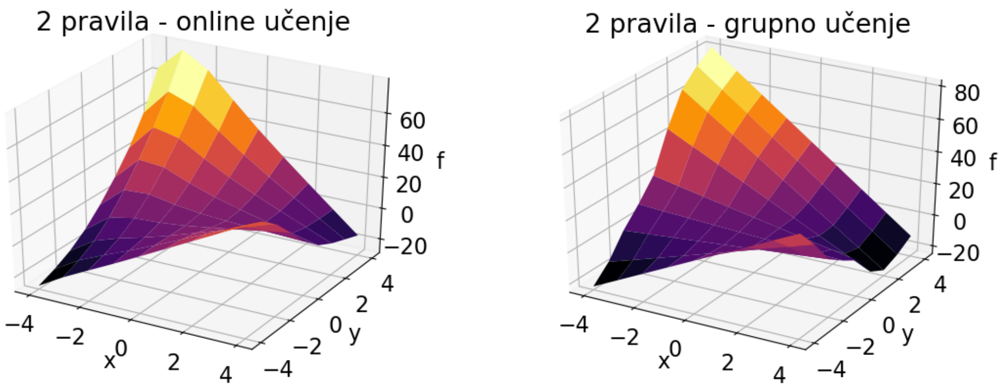


Slika 3.1: Grafovi aproksimirane generirajuće funkcije

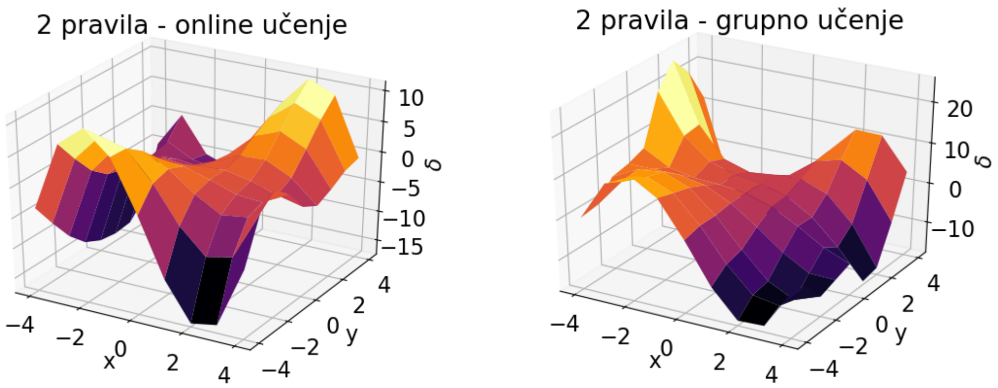


Slika 3.2: Grafovi pogrešaka

3.2. Dva pravila

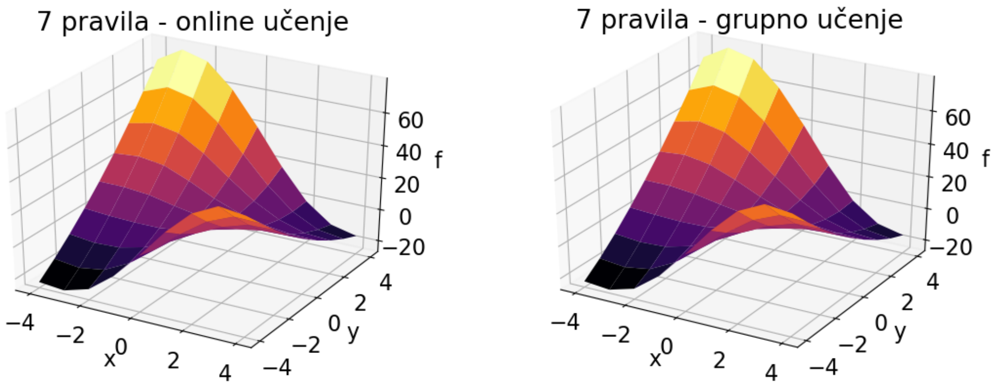


Slika 3.3: Grafovi aproksimirane generirajuće funkcije

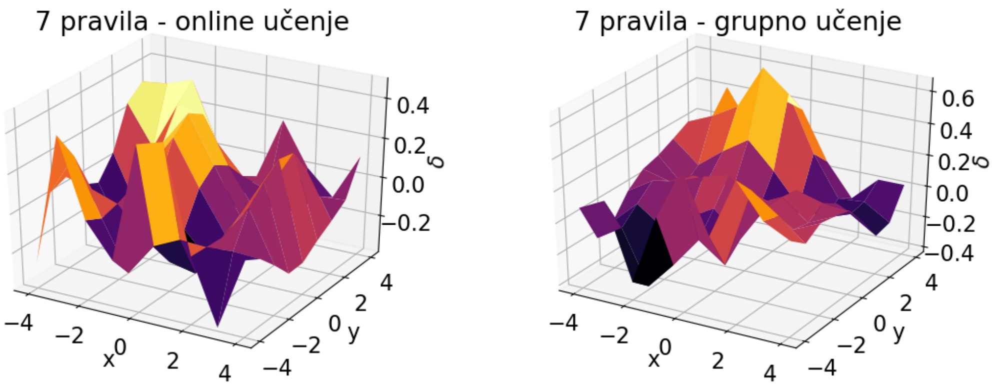


Slika 3.4: Grafovi pogrešaka

3.3. Prikladan broj pravila - sedam

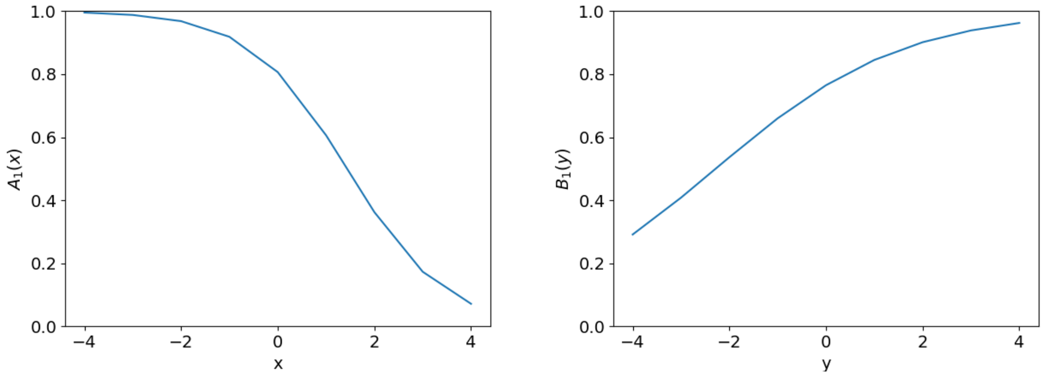


Slika 3.5: Grafovi aproksimirane generirajuće funkcije

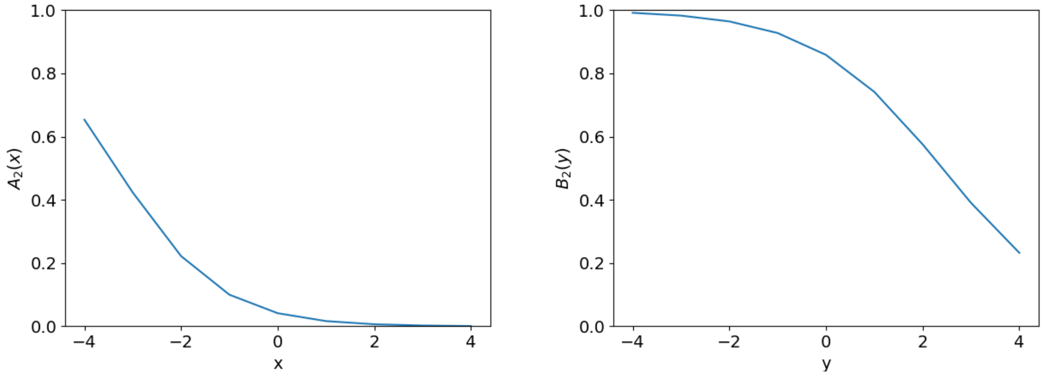


Slika 3.6: Grafovi pogrešaka

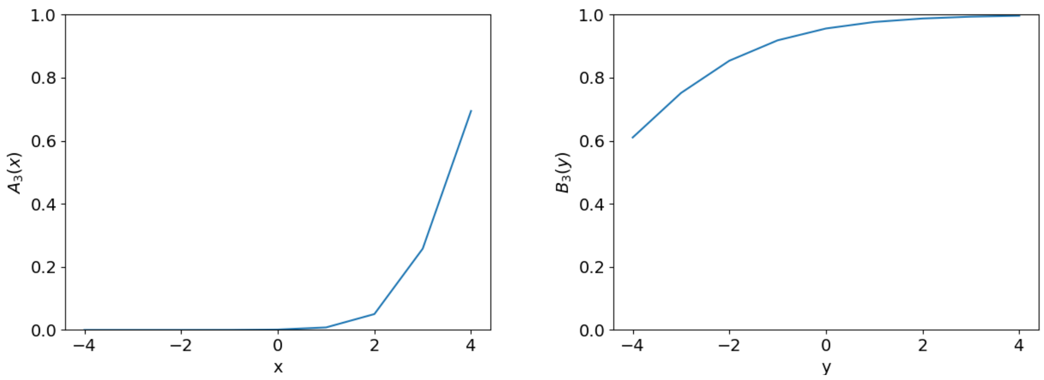
4. Zadatak 5



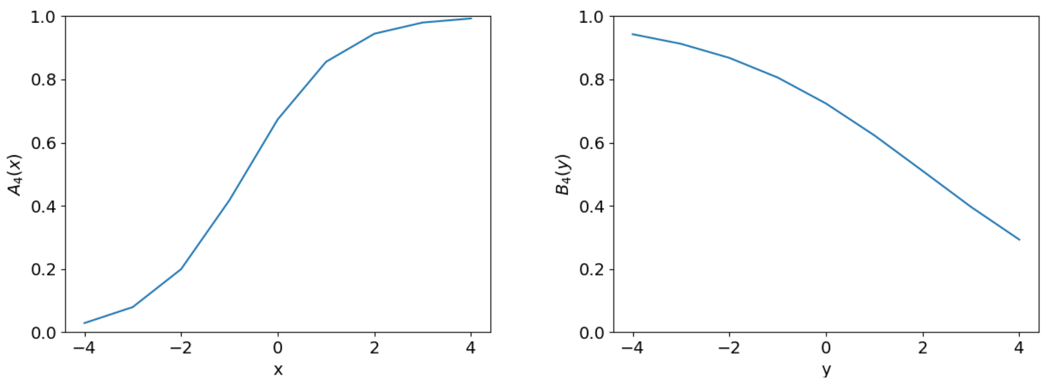
Slika 4.1: Funkcije pripadnosti za prvo pravilo



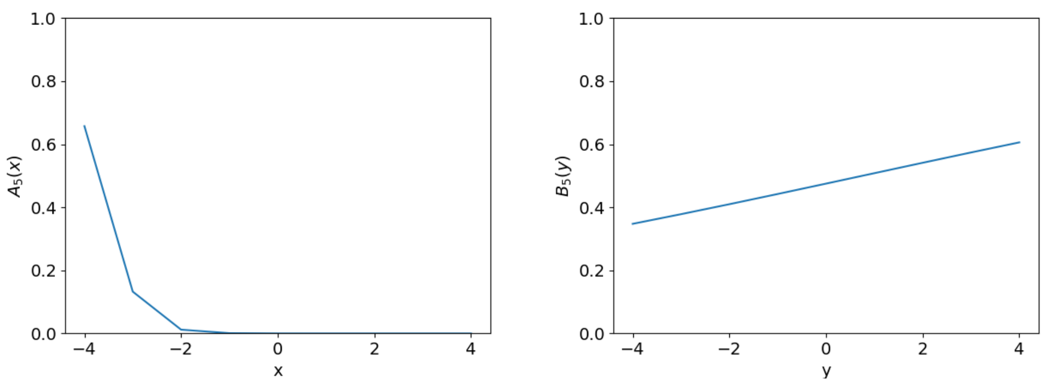
Slika 4.2: Funkcije pripadnosti za drugo pravilo



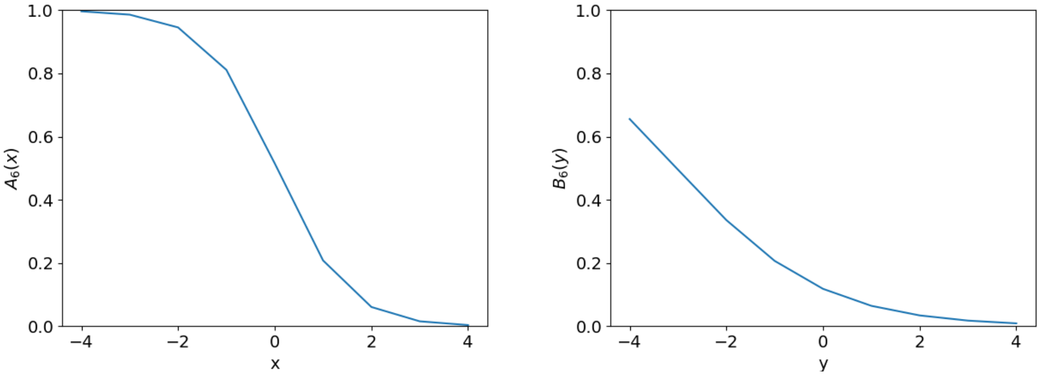
Slika 4.3: Funkcije pripadnosti za treće pravilo



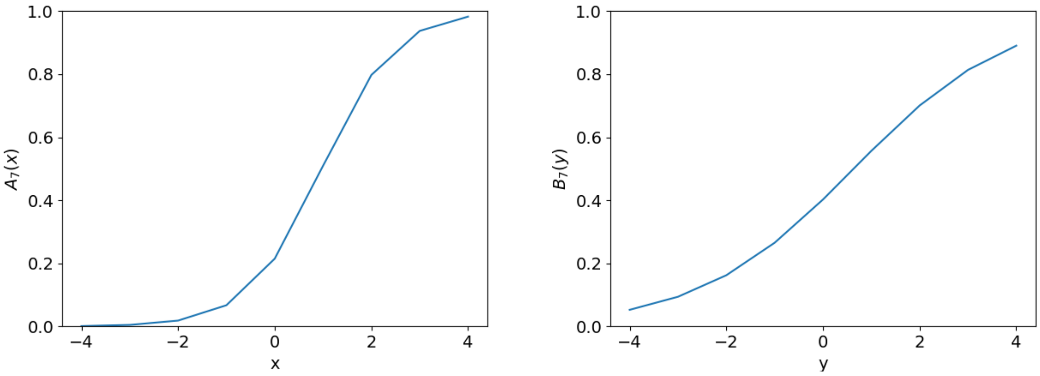
Slika 4.4: Funkcije pripadnosti za četvrto pravilo



Slika 4.5: Funkcije pripadnosti za peto pravilo

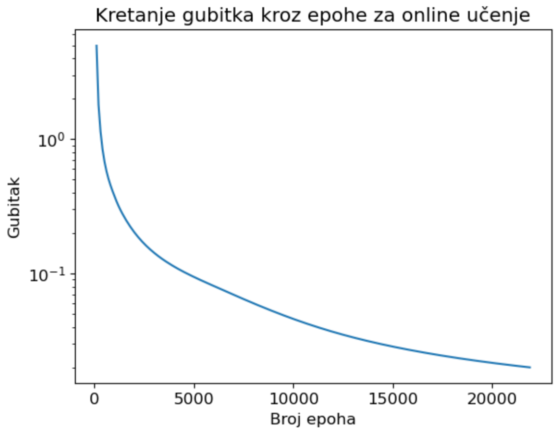


Slika 4.6: Funkcije pripadnosti za šesto pravilo



Slika 4.7: Funkcije pripadnosti za sedmo pravilo

5. Zadatak 7

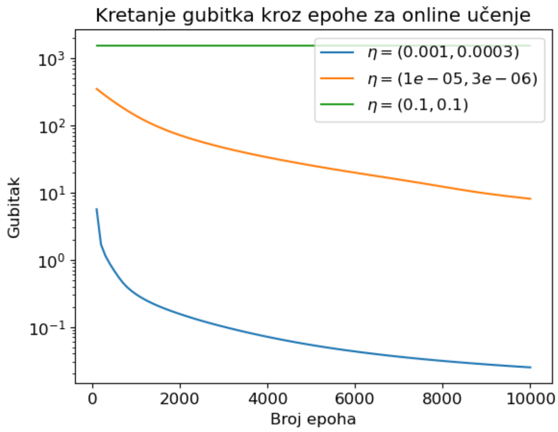


Slika 5.1

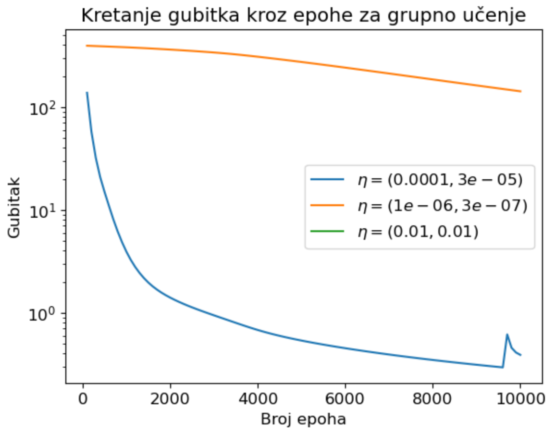
Slika 5.2

Možemo vidjeti da *online* učenje puno brže konvergira nego li grupno, no valja uzeti u obzir i stope učenja. U ovom eksperimentu za *online* učenje su korištene stope učenja 10^{-3} i $3 \cdot 10^{-4}$, dok za grupno 10^{-4} i $3 \cdot 10^{-5}$. Dakle, razlika u redu veličina je 10.

6. Zadatak 8



Slika 6.1



Slika 6.2

Što je stopa učenja manja, to algoritam sporije konvergira, ali konvergira, dok što je stopa učenja veća, algoritam divergira. Također, optimalni iznosi stopa učenje za *online* i grupno učenje razlikuju se za 10 puta.

7. Literatura

M. Čupić, B. Dalbelo Bašić, i M. Golub. *Neizrazito, evolucijsko i neuroračunarstvo*, kolovoz 2013. URL `http://java.zemris.fer.hr/nastava/nenr/knjiga-0.1.2013-08-12.pdf`. [Pristupljeno 15-Prosinac-2020.].