

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA
NEIZRAZITO, EVOLUCIJSKO I NEURO RAČUNARSTVO
Ak. godina 2020./2021.

IZVJEŠTAJ

Sedma domaća zadaća

Jelena Bratulić

Voditelj: *doc.dr.sc. Marko Čupić*

Zagreb, siječanj 2021.

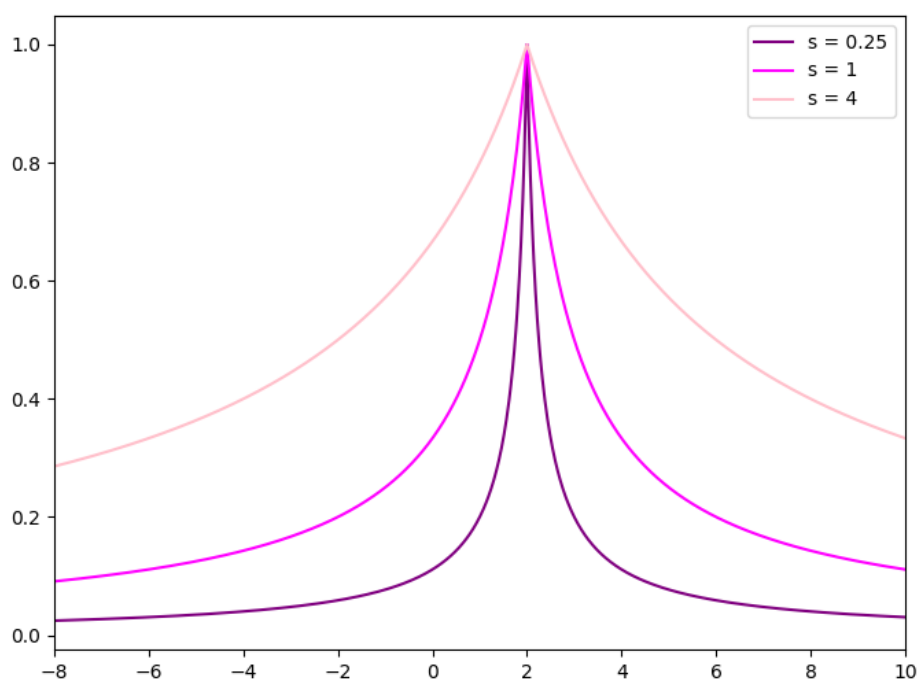
SADRŽAJ

1. Utjecaj parametra s	1
2. Vizualizacija skupa podataka	3
3. Ručno određivanje parametara modela	4
4. Određivanje optimalnih parametara genetskim algoritmom	7
4.1. Arhitektura mreže $2 \times 8 \times 3$	7
4.2. Arhitektura mreže $2 \times 8 \times 4 \times 3$	8
4.3. Arhitektura mreže $2 \times 6 \times 4 \times 3$	10

1. Utjecaj parametra s

U ovoj domaćoj zadaći koristili smo dvije vrste neurona:

- tip 1 - izlaz je mjera udaljenosti, npr. za k ulaza: $y = \frac{1}{1 + \sum_{i=1}^k \frac{|x_i - w_i|}{|s_i|}}$
- tip 2 - izlaz se propušta kroz prijenosnu funkciju sigmoidu, odnosno $y = \frac{1}{1 + e^{-x}}$



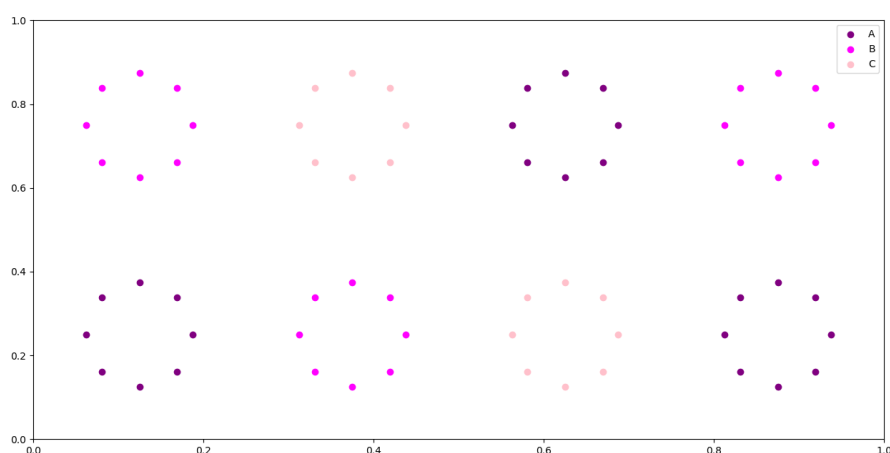
Slika 1.1: Ponašanje izlaza neurona prve vrste za fiksne parametre $w = 2, x \in [-8, 10]$ u ovisnosti o parametru s

Povećavanjem parametra s povećavamo varijancu što uzrokuje širenje zvonkolike krivulje neurona tipa 1. Budući da je neuron prve vrste zapravo mjera udaljenosti, to znači da će šire zvono blaže računati udaljenosti, odnosno neki primjeri koji su više udaljeni će se zapravo prikazati bliži/sličniji uz veću vrijednost parametra s .

Konkretno, za neuron s dva ulaza, predočena zvonkolika krivulja je zapravo mjera sličnosti na temelju udaljenosti. Sada možemo primijetiti kako će šire zvono udaljenije primjere smatrati sličnija nego bi to bilo kod užih krivulja.

Spomenuto ponašanje može se direktno vidjeti i s grafa ukoliko usporedimo krivulje s vrijednostima $s = 0.25$ i $s = 4$. Vidimo da kod manjeg iznosa parametra s pomak po apscisi uzrokuje veću promjenu u vrijednosti same funkcije udaljenosti, nego li je to kod krivulje s većom vrijednosti parametra s .

2. Vizualizacija skupa podataka

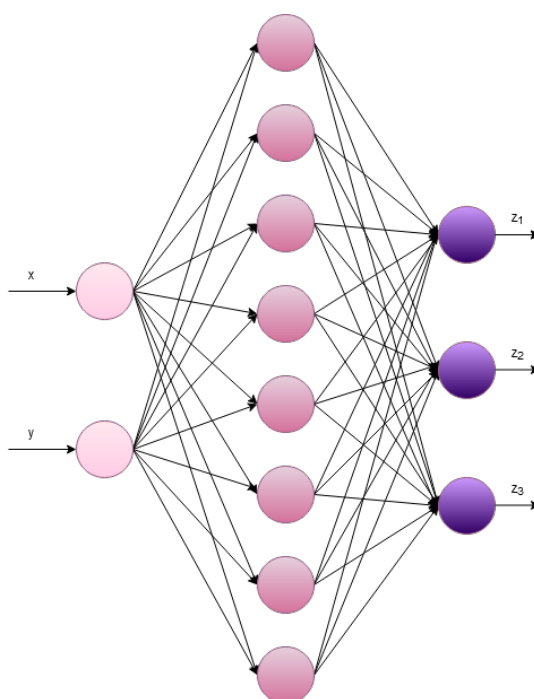


Slika 2.1: Vizualizacija skupa podataka za učenje po pripadnim labelama klasa.

Promatrajući vizualizirane primjere iz danog skupa podataka, vidimo kako podaci nisu linearno odvojivi, odnosno decizijska granica, koja određuje prostore klasa, ne može biti pravac.

Nadalje, možemo primijetiti kako primjeri slijede uzorak grupiranja po kružnicama. Vidimo kako su primjeri s oznakom A grupirani u 3 kružnice sa središtima u aproksimativno $(0.1, 0.25)$, $(0.65, 0.75)$ te $(0.9, 0.25)$. Analogno se mogu procijeniti i središta kružnica za ostale dvije klase B i C.

3. Ručno određivanje parametara modela

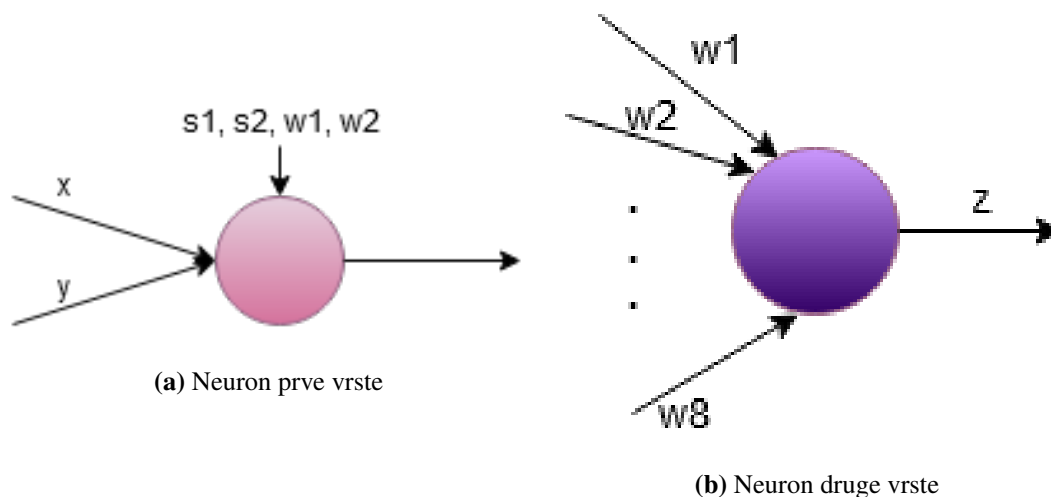


Slika 3.1: Arhitektura mreže $2 \times 8 \times 3$ pri čemu ulazni sloj predstavlja koordinate primjera x i y , skriveni sloj sastoji se od neurona prve vrste, a izlazni sloj od neurona druge vrste.

Na 3.1 je prikazana arhitektura mreže $2 \times 8 \times 3$ pri čemu se ulazni sloj sastoji od neurona koji samo preslikavaju ulazne koordinatne x i y . Skriveni sloj sastoji se od 8 neurona prve vrste, a izlazni sloj čine 3 neurona druge vrste.

Neuroni prve i druge vrste su detaljnije prikazani na 3.2. Primijetimo kako svaki neuron prve vrste ima ukupno 4 parametra - s_1, s_2, w_1, w_2 , a neuron druge vrste u ovom slučaju 8 parametara.

Parametre neurona prve vrste možemo izračunati i bez uporabe cijelog procesa učenja ukoliko pozorno promotrimo vizualizirani skup podataka. Na 2.1 vidimo kako



Slika 3.2: Detaljniji prikaz građe neurona. Parametri prve vrste neurona jesu s_1, s_2, w_1, w_2 , a neurona druge vrste težine w_1, \dots, w_8 .

imamo zapravo 8 kružnih grupacija od kojih 3 odgovaraju klasi A, 3 klasi B i 2 klasi C.

Dakle, svaki neuron iz skrivenog sloja će opisivati jednu od tih kružnih grupacija pri čemu će w_1 i w_2 označavati koordinate središta kružne grupacije, a parametri s_1 i s_2 širenje u smjeru x i y osi. Ukoliko su kružne grupacije pravilne kružnice, parametri s_1 i s_2 će biti jednakog iznosa, no ukoliko grupacija poprima eliptični oblik, jedan od parametara će biti veći od drugog, ovisno o tome je li elipsa izdužena u x ili u y smjer.

Ručno izračunati, odnosno pretpostavljeni, parametri za skriveni sloj dani su u tablici u nastavku.

Klasa	Pozicija	w_1	w_2	s_1	s_2
A	dolje, lijevo	0.12	0.25	0.010	0.018
A	gore	0.62	0.75	0.010	0.018
A	dolje, desno	0.88	0.25	0.010	0.018
B	gore, lijevo	0.12	0.75	0.010	0.018
B	dolje	0.38	0.25	0.010	0.018
B	gore, desno	0.88	0.75	0.010	0.018
C	gore	0.38	0.75	0.010	0.018
C	dolje	0.62	0.25	0.010	0.018

Tablica 3.1: Ručno izračunati parametri skrivenog sloja. Budući da su grupacije blago eliptične, parametar s_2 će biti većeg iznosa od parametra s_1 jer su elipse izdužene po y osi.

Vrijednosti parametara izaznog sloja odredili bismo na sljedeći način. Za prvi izlazni neuron koji određuje oznaku pripadnosti klasi A, s većim bismo težinama uzimali one neurone koji u skrivenom sloju određuju mjeru sličnosti s grupacijama koje pripadaju klasi A. Zapravo možemo reći i da bismo uzimali samo te izlaze iz skrivenog neurona jer za određivanje klase A, nas ne zanima mjera sličnosti s klasom B ili C. Pretpostavljene vrijednosti predložene su u nastavku.

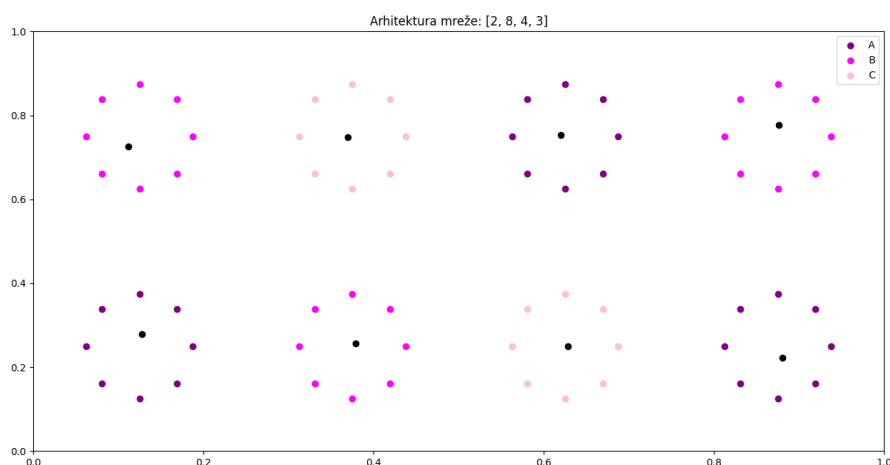
Klasa	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
A	14.05	14.55	14.24	-2.03	-1.02	-1.07	-4.01	-3.01
B	-3.06	-4.01	-4.03	13.52	14.03	13.85	-2.01	-2.02
C	-3.01	-2.02	-4.04	-3.02	-2.03	-1.06	15.06	15.07

Tablica 3.2: Pretpostavljeni parametri izlaznog sloja.

4. Određivanje optimalnih parametara genetskim algoritmom

4.1. Arhitektura mreže $2 \times 8 \times 3$

Optimalni parametri za arhitekturu mreže $2 \times 8 \times 3$ ostvarili su potpuno točnu klasifikaciju uzoraka, a i očekivane vrijednosti parametara w_1, w_2, s_1 i s_2 se podudaraju s optimalnima. U nastavku su prikazani ulazni primjeri s indikacijom razreda te naučeni parametri w .



Slika 4.1: Ulazni primjeri i naučeni parametri w

U nastavku su prikazane i vrijednosti parametara neurona prve vrste kako bismo ih mogli detaljnije usporediti s pretpostavljenim vrijednostima. Optimalni parametri neurona prve vrste se za vrijednosti parametra w uvelike podudaraju s pretpostavljenim vrijednostima. Kod parametara s vrijedi trend da je komponenta za y-os većeg iznosa od komponente za x-os budući da su elipse izdužene po y-osi.

Neuron	Pozicija	w_1	w_2	s_1	s_2
1	dolje, sredina	0.6281	0.2492	-0.114	-0.334
2	gore, sredina	0.6195	0.7531	-0.082	0.2004
3	gore, sredina	0.3694	0.7483	-0.106	-0.287
4	dolje, desno	0.8803	0.2226	-0.097	-0.371
5	gore, desno	0.8758	0.7765	-0.328	-0.181
6	gore, lijevo	0.1121	0.7252	-0.888	0.2466
7	dolje, sredina	0.3786	0.2563	-0.097	0.2069
8	dolje, lijevo	0.1278	0.2793	0.3602	0.1974

Tablica 4.1: Ručno izračunati parametri skrivenog sloja. Budući da su grupacije blago eliptične, parametar s_2 će biti većeg iznosa od parametra s_1 jer su elipse izdužene po y osi.

Arhitektura mreže već je prikazana na 3.1, a naučene težine za neurone druge vrste, koji čine izlazni sloj, dane su u tabličnom prikazu u nastavku. Ukoliko povežemo podatke iz 4.1 i sljedeće podatke, vidimo kako vrijedi trend da neuron ima pozitivnu vrijednost težine za onu labelu koja pripada tom neuronu.

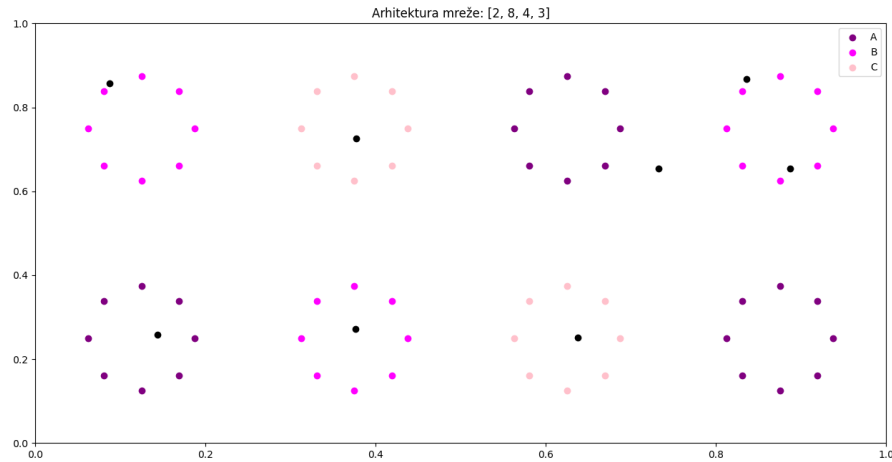
Klasa	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
A	-22.69	35.298	-9.745	12.174	-11.54	-6.420	-20.35	14.734
B	-6.269	-21.87	-24.52	-6.675	14.842	13.674	33.674	-10.78
C	26.417	-12.15	27.358	-5.169	-13.68	-5.031	-11.19	-11.94

Tablica 4.2: Naučeni parametri izlaznog sloja.

Odnosno, iz 4.1 vidimo kako klasi A odgovaraju neuroni 2, 4, 8. U 4.2 vidimo kako ti neuroni imaju pozitivnu vrijednost za ulazeće neurone iz prethodnog sloja koji odgovaraju klasi A, a za ostale su negativni. Analogno vrijedi i za klasu B te C. Ovo potvrđuje pretpostavku od ranije da će izlaznom neuronu za klasu A *biti bitni* samo neuroni iz skrivenog sloja koji direktno mjere udaljenost od elipsi koje opsiuju primjere iz skupa A.

4.2. Arhitektura mreže $2 \times 8 \times 4 \times 3$

Optimalni parametri za ovu arhitekturu ponovno ostvaruju u potpunosti točnu klasifikaciju ulaznih primjera. Zanimljivo je i kako je vrijeme treniranja u ovom slučaju bilo uvelike kraće nego za arhitekturu $2 \times 8 \times 3$.



Slika 4.2: Ulazni primjeri i naučeni parametri w .

Razlika u vremenu učenja modela proizlazi iz činjenice različitog broja parametara čime se mijenja i prostor pretraživanja optimalnih parametara. Odnosno jedinka se sada sastojala od dvadesetak parametara više nego u prošlom slučaju što je dovelo do lakše optimizacije jer nisu trebali svi parametri biti toliko uštimani, a da bi klasifikacija jednako dobro radila.

Iako je u ovom slučaju klasifikacija bila točna, vrijednosti za parametre w se nisu u potpunosti podudarali s onima iz prethodnog slučaja, kao ni s pretpostavljenim vrijednostima. U nastavku su prikazane naučeni parametri neurona prve vrste.

Neuron	Pozicija	w_1	w_2	s_1	s_2
1	gore, desno	0.8871	0.6541	2.1331	3.2248
2	gore, desno	0.7321	0.6534	-6.155	-0.244
3	dolje, sredina	0.3761	0.2723	0.7189	-0.222
4	gore, desno	0.8362	0.8675	6.0792	-0.375
5	gore, lijevo	0.0871	0.8569	-8.808	3.4029
6	gore, sredina	0.3770	0.7255	0.0602	-6.132
7	dolje, sredina	0.6376	0.2506	0.0825	-0.787
8	dolje, lijevo	0.1434	0.2574	0.2772	-5.929

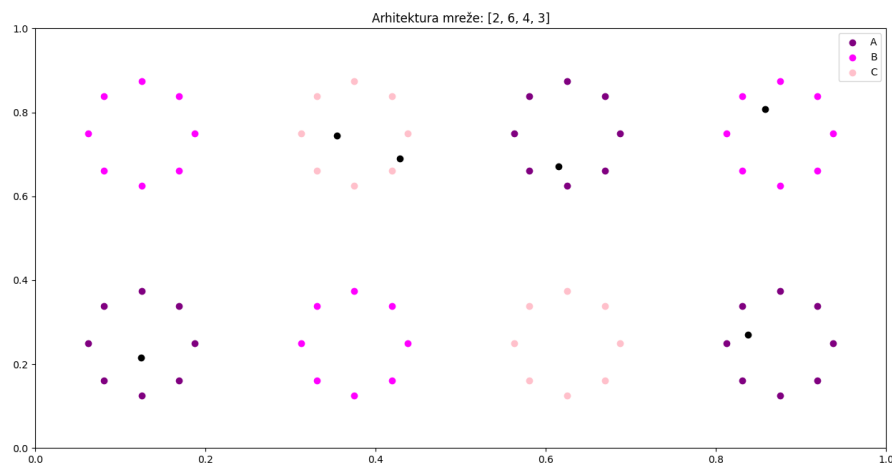
Tablica 4.3: Ručno izračunati parametri skrivenog sloja. Budući da su grupacije blago eliptične, parametar s_2 će biti većeg iznosa od parametra s_1 jer su elipse izdužene po y osi.

Naučene vrijednosti neurona prve vrste se razlikuju od očekivane budući da se iz-

lazi iz tih neurona ne koriste direktno za klasifikaciju nego ulaze u još jedan potpuno-povezani sloj. Time se stvara različiti prostor pretraživanja, ali i drukčiji optimizacijski problem u vidu vrijednosti parametara pa podudaranje s pretpostavljenim vrijednostima nije nužno. Isti razlog istovremeno utječe i na vrijeme učenja koje se očitovalo i u kraćem izvođenju, ali i u manjem broju iteracija koje su potrebne.

4.3. Arhitektura mreže $2 \times 6 \times 4 \times 3$

Spomenuta arhitektura je dala u potpunosti ispravnu klasifikaciju ulaznih primjera. U nastavku su prikazani ulazni primjeri i naučeni parametri w za neurone prve vrste.



Slika 4.3: Ulazni primjeri i naučeni parametri w .

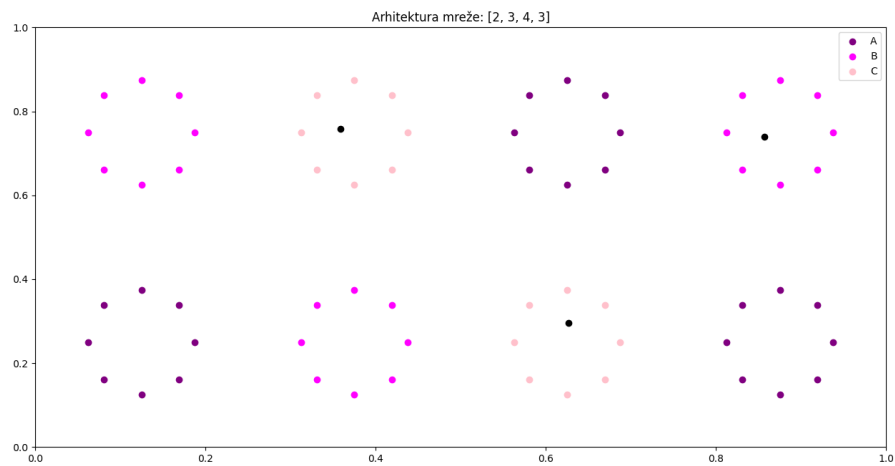
Primijetimo da budući da imamo samo 6 neurona prve vrste, nije moguće ostvariti vrijednosti kao što smo ih dobili za polaznu arhitekturu. Dakle, korištenjem manje od 8 neurona gubimo preklapanje parametara sa stvarnim vrijednostima centroida i odsječaka po x i y osi, no taj gubitak podataka i dalje ne utječe na točnost klasifikacije u vidu jednorećanih oznaka.

Odnosno, možemo reći da smo prije znali kako se model ponaša i što predstavlja svaki sloj mreže, a sada nam to nije poznato ili intuitivno jer nemamo točno 8 neurona koji odgovaraju 8 grupa već ih je manje.

Nadalje, pokušala sam još više smanjiti broj neurona prve vrste. Tako sam i za arhitekturu $2 \times 4 \times 4 \times 3$ dobila potpuno točnu klasifikaciju primjera.

Dodatno sam još isprobala i arhitekturu $2 \times 3 \times 4 \times 3$ koja je također dala potpuno točnu klasifikaciju svih primjera. Naučeni centroidi za spomenutu arhitekturu prika-

zani su na sljedećoj slici. Ovime je pokazano kako nije nužno imati 8 neurona prve vrste kako bi klasifikacija bila potpuno točna.



Slika 4.4: Ulazni primjeri i naučeni parametri w .