

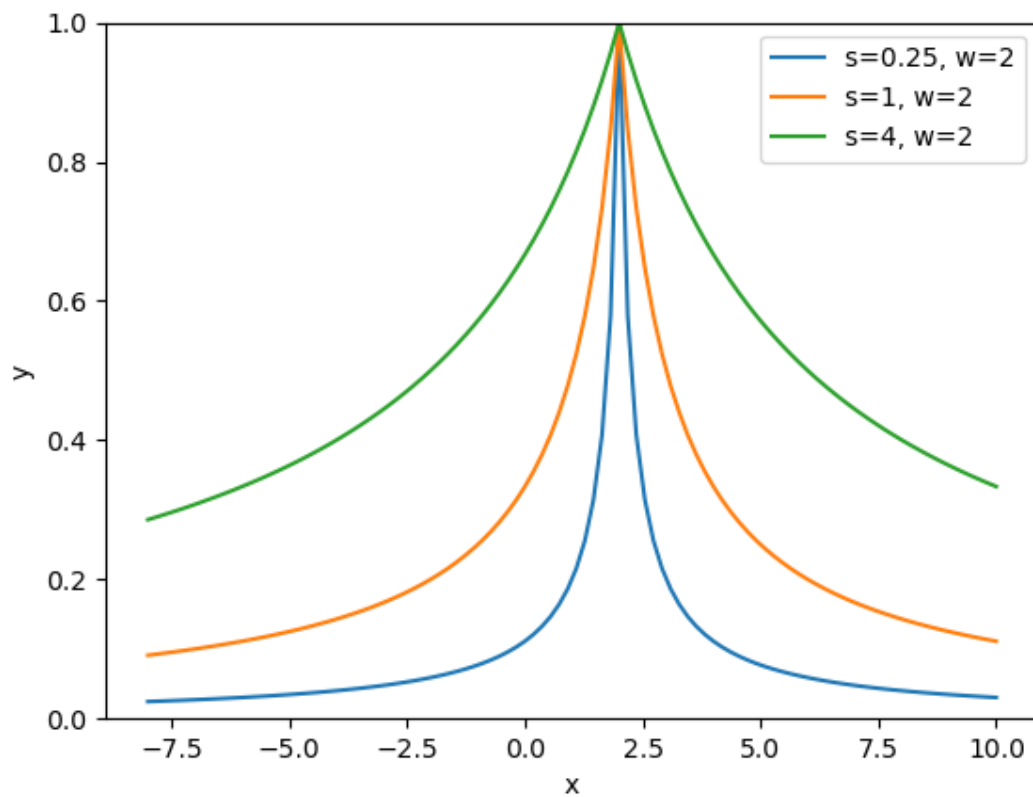
Borna Goljački

Neizrazito, evolucijsko i neuroračunarstvo

## PROJEKTNII ZADATAK 6

GA ANN

Zadatak 1.

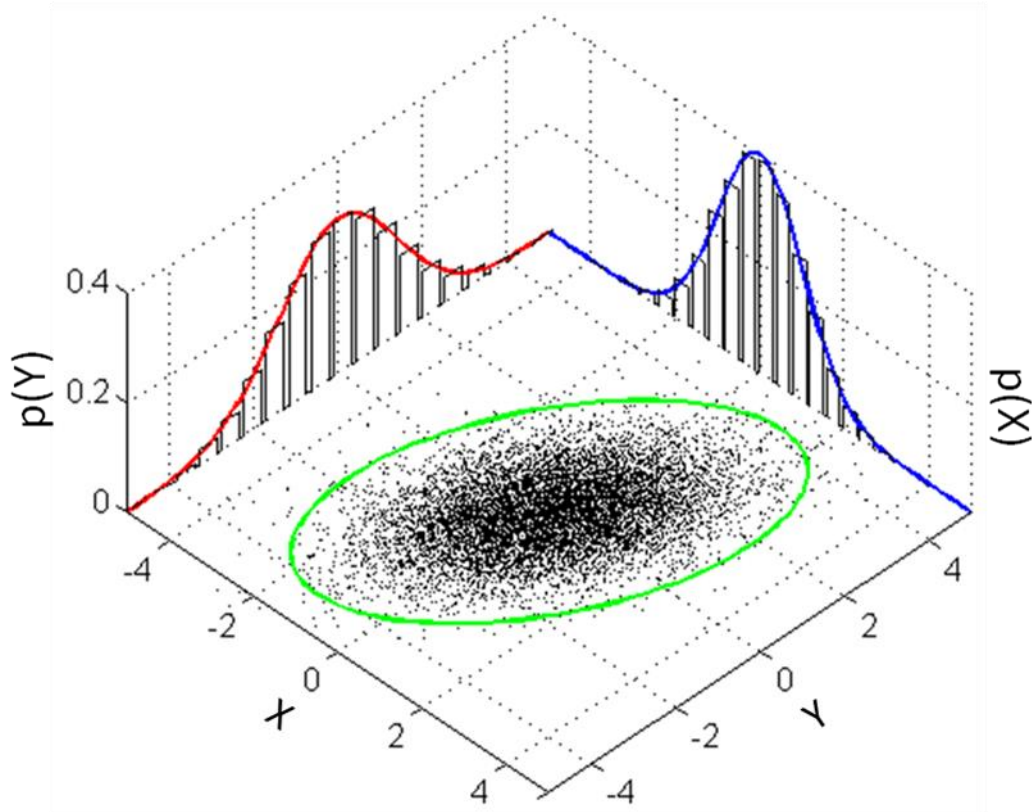


Povećavajući varijablu  $s$  uočavamo da distribucija vrijednosti postaje sve „otvorenija“.

To za posljedicu ima da smo tolerantniji, tj. točke koje su udaljenije smatramo sličnijima nego što bismo da je faktor  $s$  to manji.

U prijevodu,  **$s$  direktno modelira složenost našeg modela**, što je  $s$  veći, to ćemo bolje generalizirati, što je  $s$  manji, to nam se model bolje prilagođava podacima i postaje specijaliziraniji.

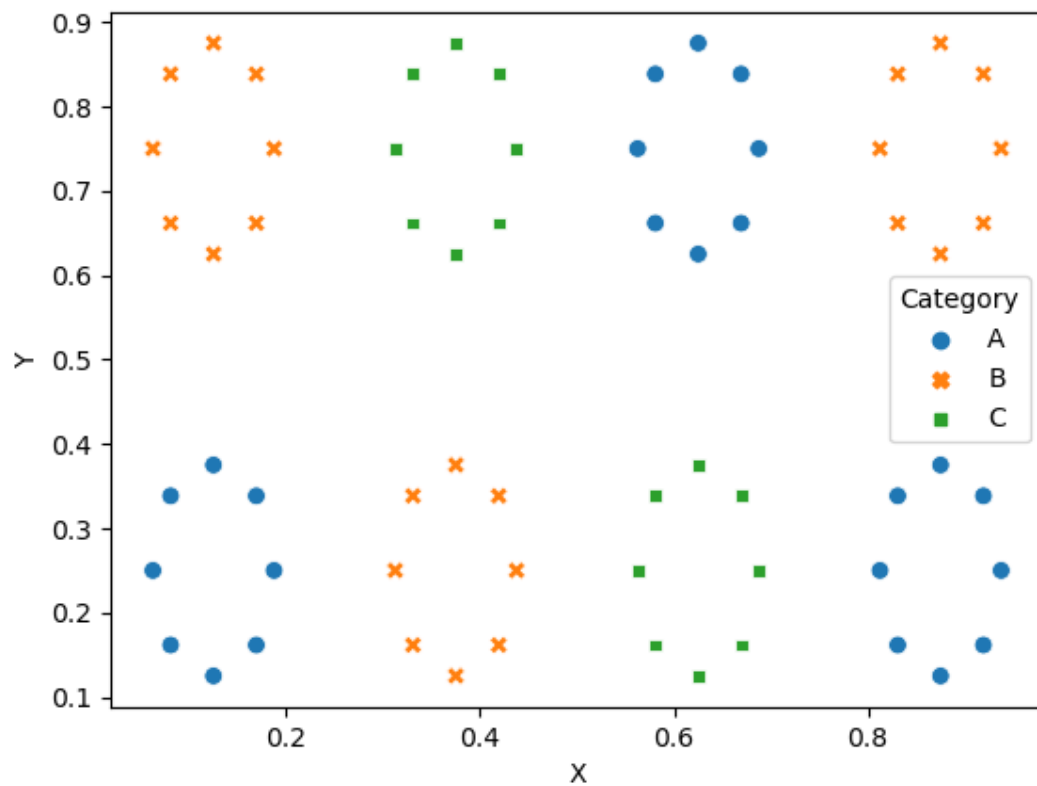
Izlaz neurona koji ima dva ulaza moguće je prikazati 3D grafom. Taj graf ima izgled sličan sljedećem.



Jedina razlika između tog grafa i ovog je ta što ova slika ne prikazuje funkcije sličnosti neurona tipa 1, nego slike normalne distribucije.

Ali, analogno toj slici, možemo zamisliti da je sa svake strane osi zapravo funkcija sličnosti neurona tipa 1, te da bi, mijenjajući parametre  $s_1$  i  $s_2$ , utjecali na „debljinu“ (toleranciju) ili „uskost“ (netoleranciju) samog brda koje točke tvore. Ukratko, povećavajući  $s_1$  i  $s_2$  opet smatramo veći broj točaka sličnijim, nego što bi to bio slučaj da su ti faktori manji, ovo je samo prošireno na 2 dimenzije, gdje imamo slobodu birati hoćemo li više tolerirati ulaz 1 ili ulaz 2.

## Zadatak 2.



Podaci na ovom grafu su grupirani u oblik elipse, tj. njihova  $y$  koordinata je (naoko) 2 puta veća od pripadajuće  $x$  koordinate.

Također je važno za primjetiti da ovaj skup podataka **nije linearno odvojiv** budući da ne postoji hiperravnina koja može ispravno klasificirati svaku od grupa (ne možemo povući pravac koji bi odvojio zelenu, plavu i narančastu grupu, bez da krivo klasificira ijednu od nje).

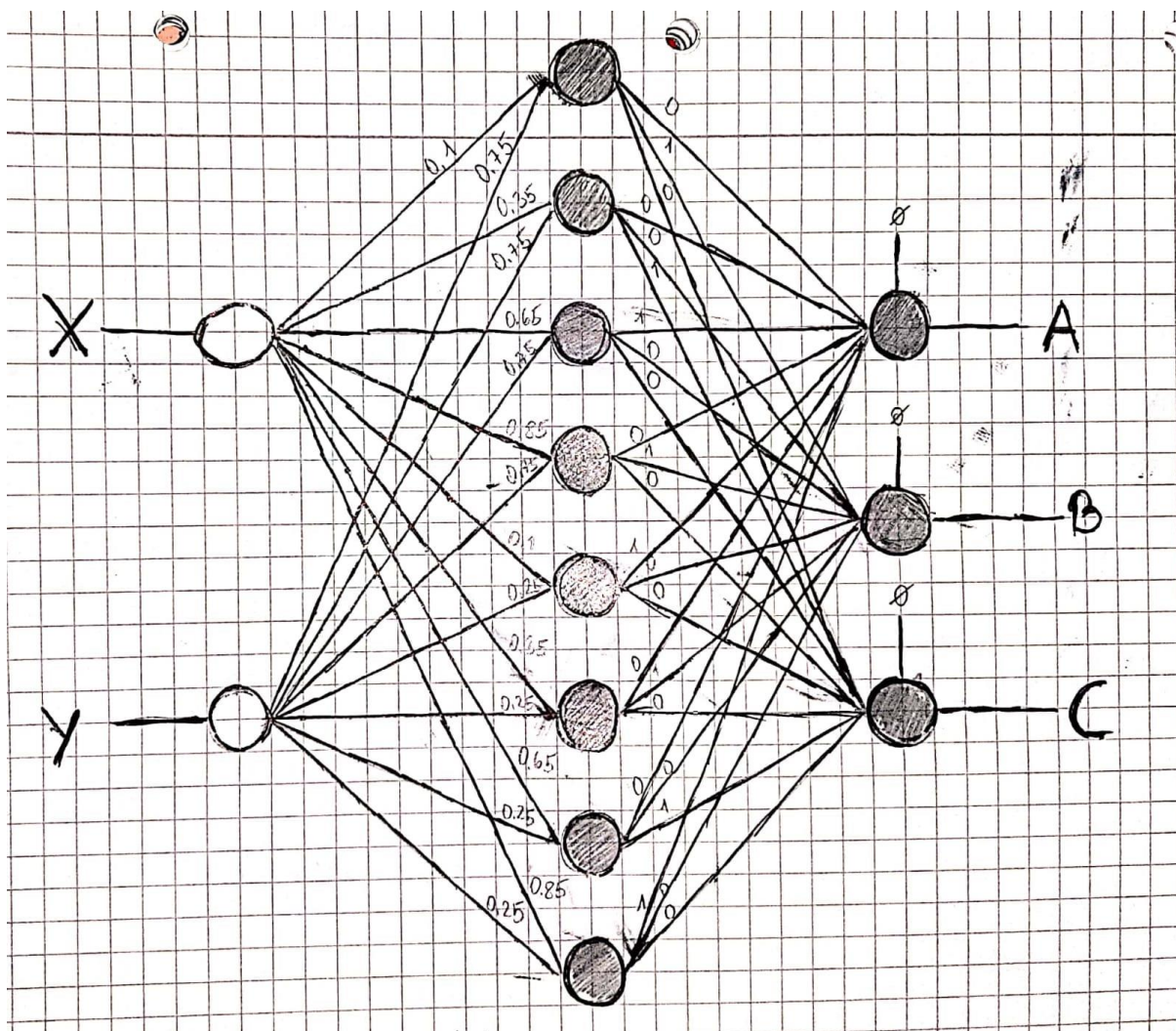
### Zadatak 3.

Kada bih morao ručno odrediti vrijednosti svih parametara (težine) zadane neuronske mreže, postavio bih ih na sljedeći način:

Težine sloja neurona tipa 1 odgovarale bi koordinatama točaka koje se nalaze u središtu svake od elipsa iz skupa podataka.

Neuroni izlaznog sloja bi idealno imali najjasniji mogući odabir; ili je njihov izlaz 0, ili je 1. Tu dužnost imaju neuroni prethodnih slojeva, kao što je i opisano u nastavku.

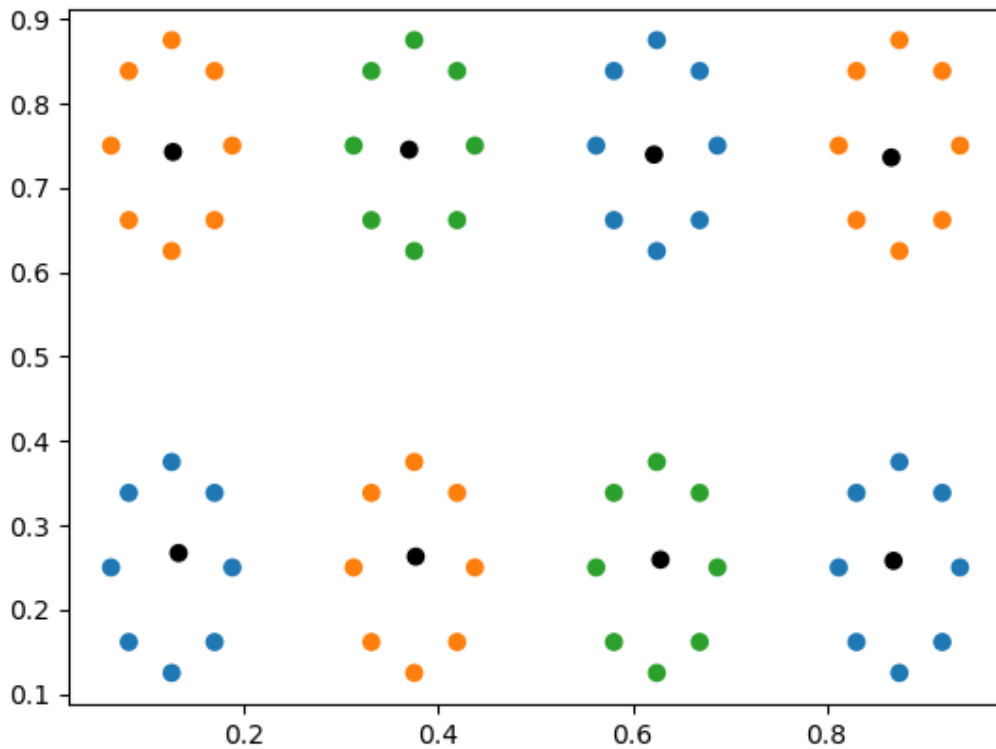
Dužnost neurona ostalih slojeva bila bi tim više aktivirati neurone sloja tipa 1 tako da čim više odgovaraju izlazu, drugim riječima, ako je neuron tipa 1 izgenerirao točku koja se nalazi na središtu same elipse, tada bi svaki sljedeći sloj neurona, ovisno o kategoriji same elipse, imao što veću (ili manju) težinu, u svrhu najveće aktivacije neurona do ispravne klasifikacije (je li nešto uistinu „A“, „B“ ili „C“).



Neuroni tipa jedan bi također u sebi sadržavali parametar  $s$ , koji bi za  $y$  bio dvostruko veći nego za  $x$ .

Zadatak 4.

Neuronska mreža arhitektura 2 x 8 x 3.



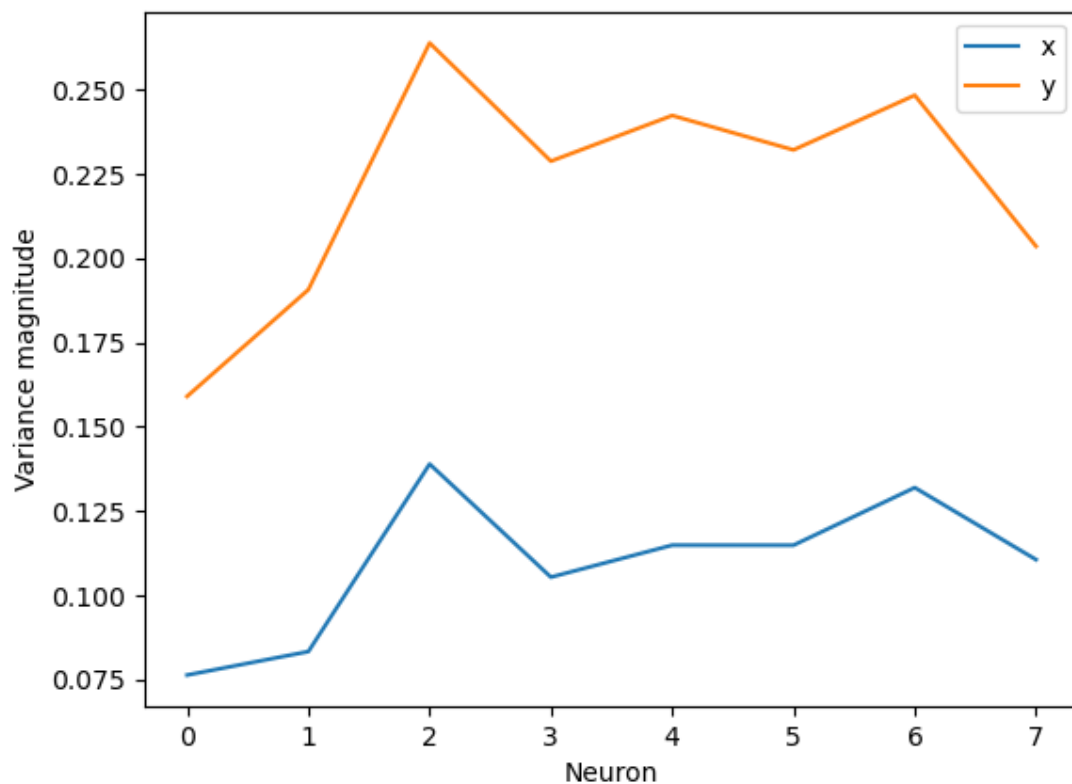
Crne točke predstavljaju točke koje je GA naučio.

Jasno je vidljivo da te točke predstavljaju (okvirni) centar elipsa koju tvori skup podataka.

To je u skladu s očekivanjem budući da je algoritam minimizirao grešku, a greška se dogodila kada predviđene točke ne bi bile slične (dovoljno udaljene) točkama iz skupa podataka, stoga, algoritam je generirao središta samih elipsi.

Na ovoj slici je vidljiv omjer apsolutnih vrijednosti „varijanci“, odnosno parametara  $s$ , na svakom od 8 neurona sloja sličnosti.

Možemo uočiti da je omjer  $y : x$  skoro pa savršeno  $2 : 1$ , što je naravno u skladu s očekivanjem jer naš skup podataka grupiran je u elipse, čija je jedna varijabla dvostruko veća od druge (da su  $1 : 1$  onda bi to značilo da očekujemo kružnice, npr.).

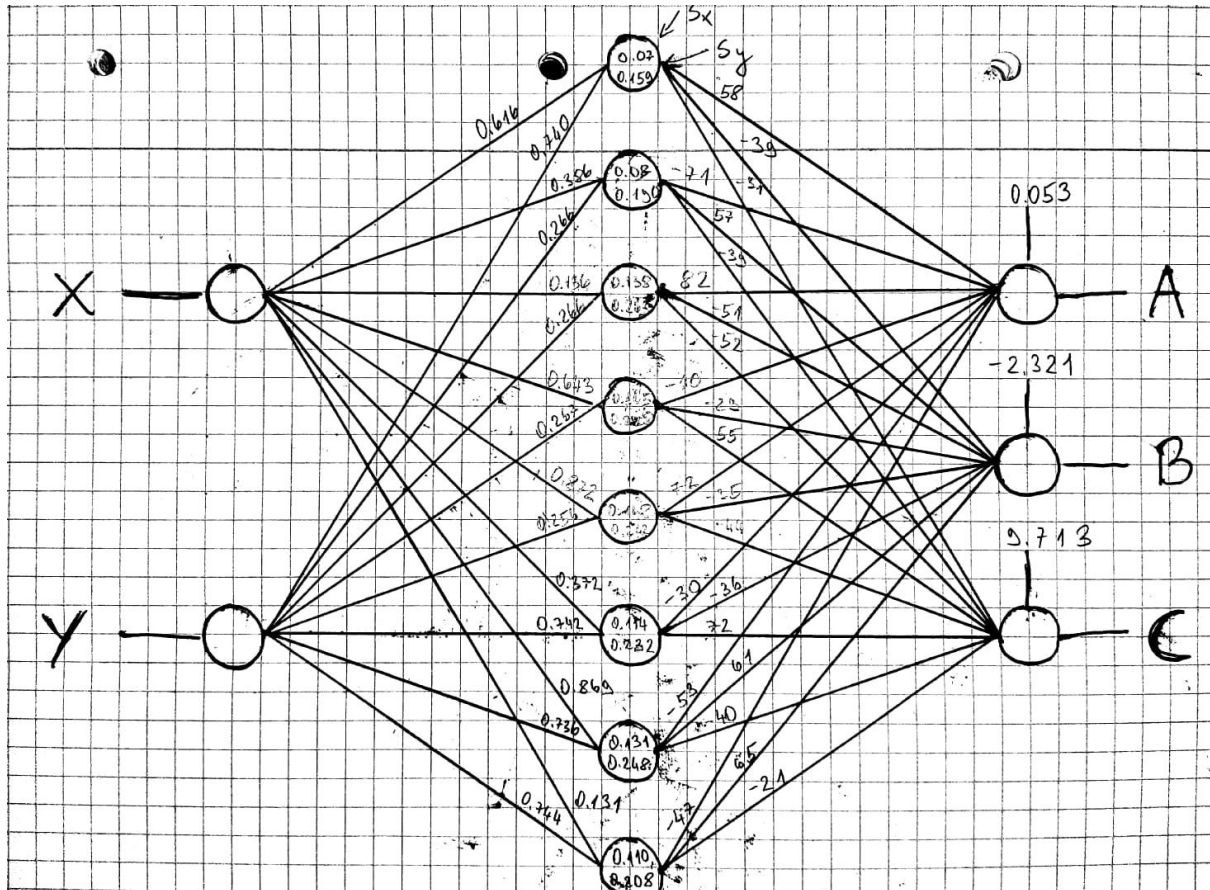


Izgled svih neurona naučene mreže 2 x 8 x 3.

Na slici je vidljivo kako je moje predviđanje iz zadatka 3 bilo ispravno, doduše parametri su optimirani drugačijim redoslijedom.

Svaki od neurona iz sloja sličnosti ima za dužnost tim više aktivirati izlazni sloj što je klasifikacija ispravnija, dakle, prvi neuron iz sloja evidentno aktivira „A“ budući da je njegova težina (58) najveća, dok su mu težine prema neuronima „B“ i „C“ (-39, -31 respektivno) znatno manje.

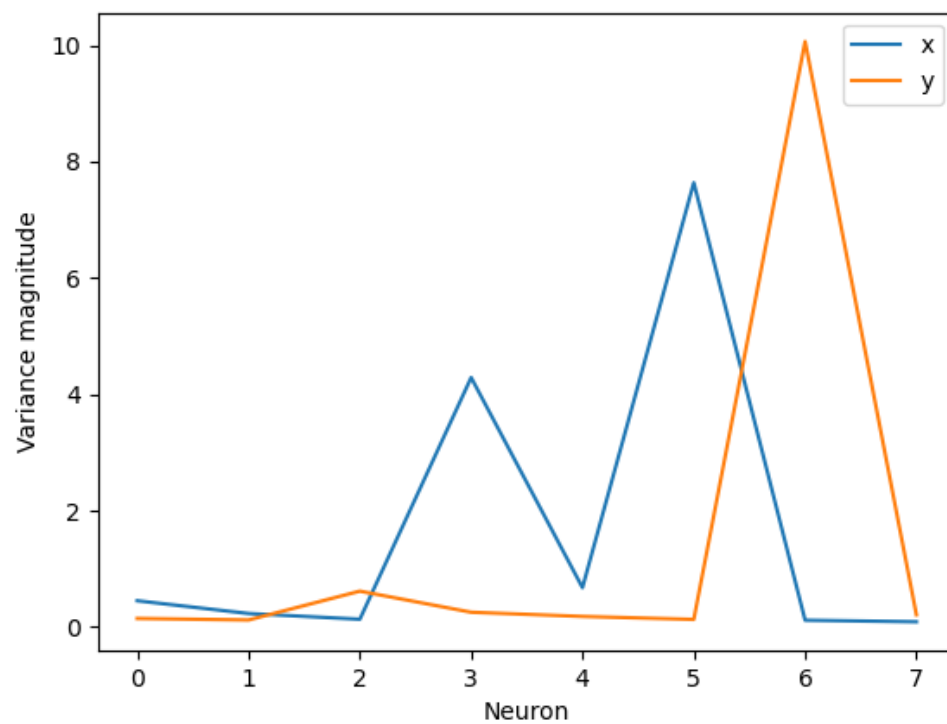
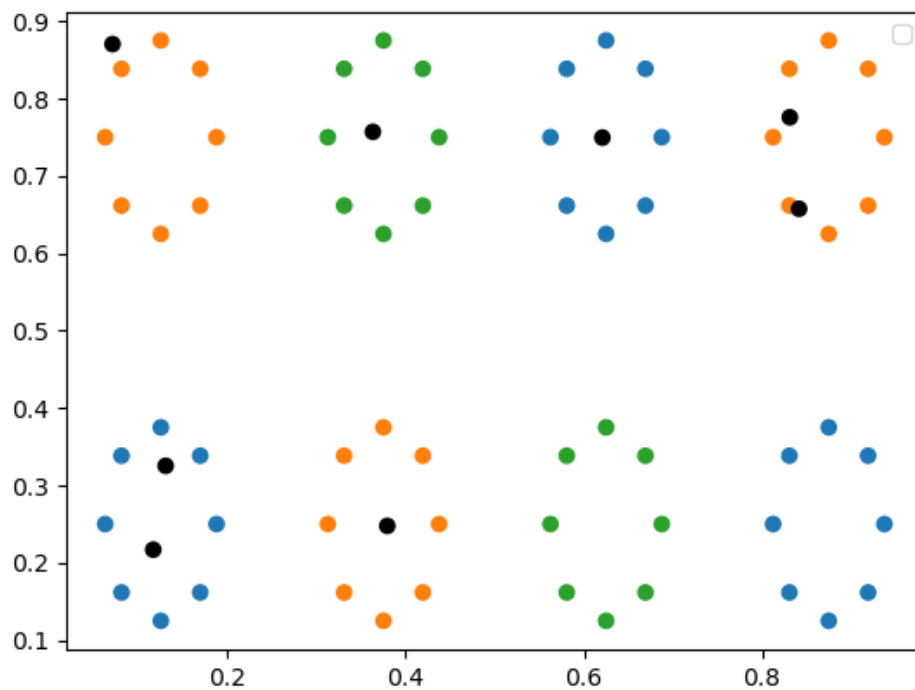
To odgovara mojoj predikciji da bi neuroni na izlazima trebali davati vrijednosti poput (1, 0, 0), iako je ovdje to drastično izraženo.





Zadatak 5.

Mreža arhitekture 2 x 8 x 4 x 3.



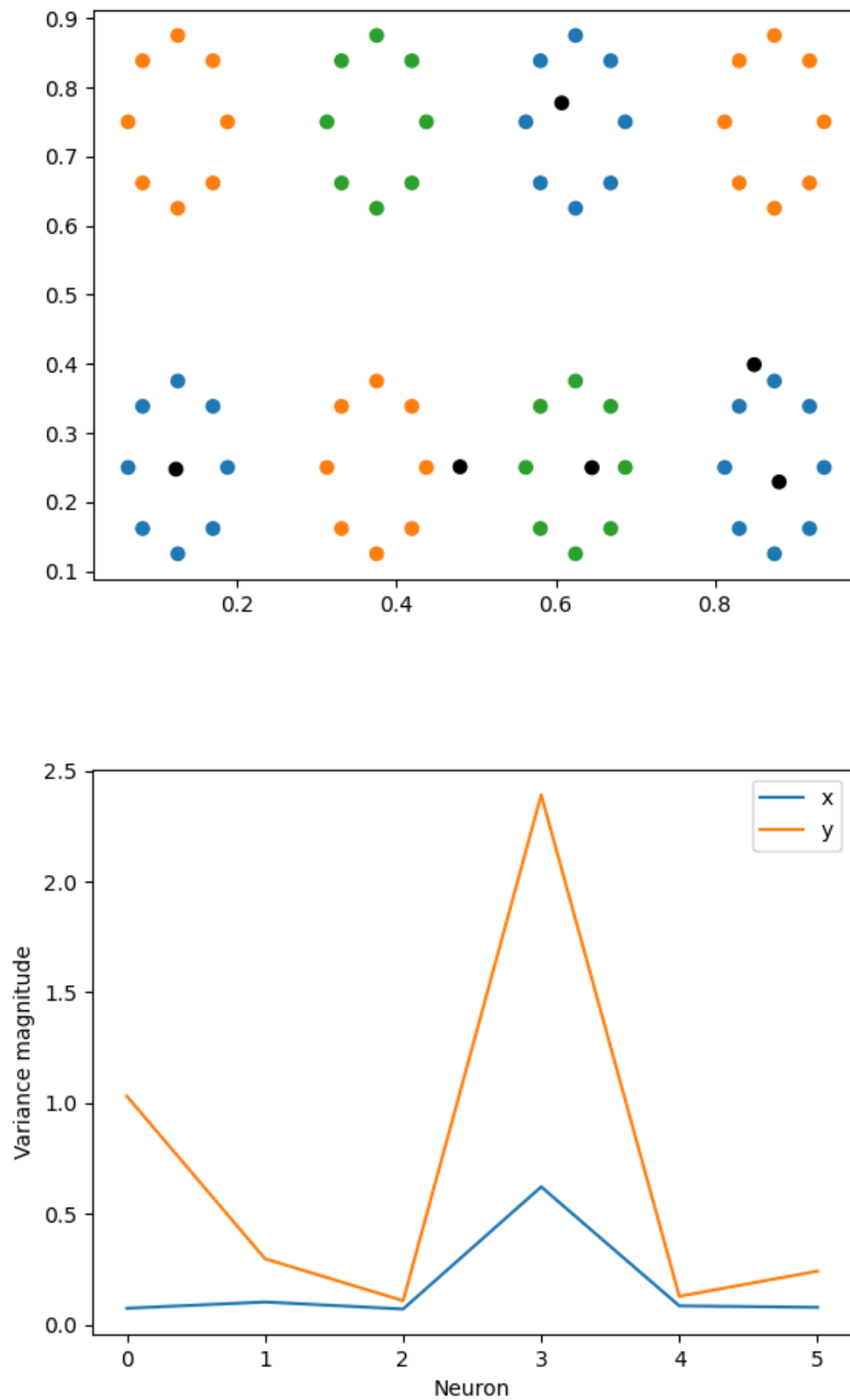
Postupak je trajao kraće, zato što je zbog dodatna 4 neurona imao više prostora „prenaučiti“ se nad podacima, odnosno, brže je smanjio grešku na manju od  $10^{-7}$ .

To je vidljivo na samom grafu gdje je nacrtano 8 točaka koje je mreža izgenerirala (konzistentno kao i prije, obojano crnom) i cijeli skup podataka. **Međutim izgenerirane točke svejedno ne klasificiraju ispravno. Zašto?** Iako je genetski algoritam pronašao mrežu koja uistinu ima veoma malu pogrešku, on nije proizveo točke kojima se opisuju centri elipsa, već naoko besmislene točke na nekim mjestima. Kako je to moguće? Ukratko, zbog parametra  $s$ . **Genetski algoritam je uspio kompenzirati loše koordinate točaka različitim skalama sličnosti te je time postigao „tehnički“ ispravnu minimalnu grešku.** To je primjer mreže koja bi loše generalizirala.

Također je vidljivo da su parametri  $s$  skoro pa kompletno besmisleni, više ne odgovaraju omjeru 2 : 1 u korist  $y$  nad  $x$  već samo nasumično rastu i padaju, kako bi iskompenzirali lošu lokaciju izgeneriranih točaka.

Zadatak 6.

Mreža arhitektura 2 x 6 x 4 x 3.



Kao što je vidljivo iz priloženih grafova, mreža arhitekture  $2 \times 6 \times 4 \times 3$  ispravno ne klasificira podatke.

Razlog tomu je što **mreža nije dovoljno ekspresivna**.

Imajući samo **6** neurona u sloju sličnosti, mreža nije ni teoretski sposobna ispravno klasificirati podatke, jer ne može pronaći centar **8** elipsa.

Ono što smo **izgubili** smanjenjem broja neurona u sloju sličnosti na  $< 8$  je **sposobnost ispravne klasifikacije** i generalizacije same mreže, a tome ne može ni skala sličnosti doskočiti.