

# **NEURALNE MREZE**

## **Izvestaj o prvom projektnom zadatku**

Filip Kojic 2018/0285

Ivan Jevtic 2018/0550

## **1.Zadatak**

Vrednosti dobijene na osnovu indeksa su:

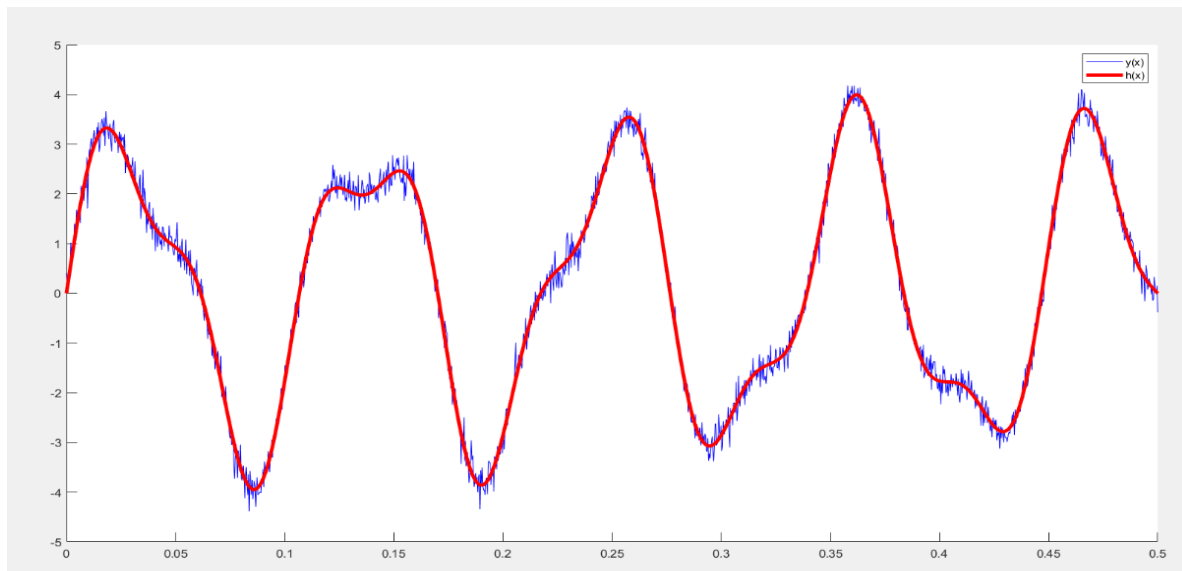
$A = 1$ ,  $B = 3$ ,  $f_1 = 20$ ,  $f_2 = 9$ , pa je:

$$h(x) = \sin(40 * \text{PI} * X) + 3\sin(18 * \text{PI} * x),$$

$$y(x) = h(x) + \text{std} * \text{randn}(1, N)$$

$$\text{std} = 0,2$$

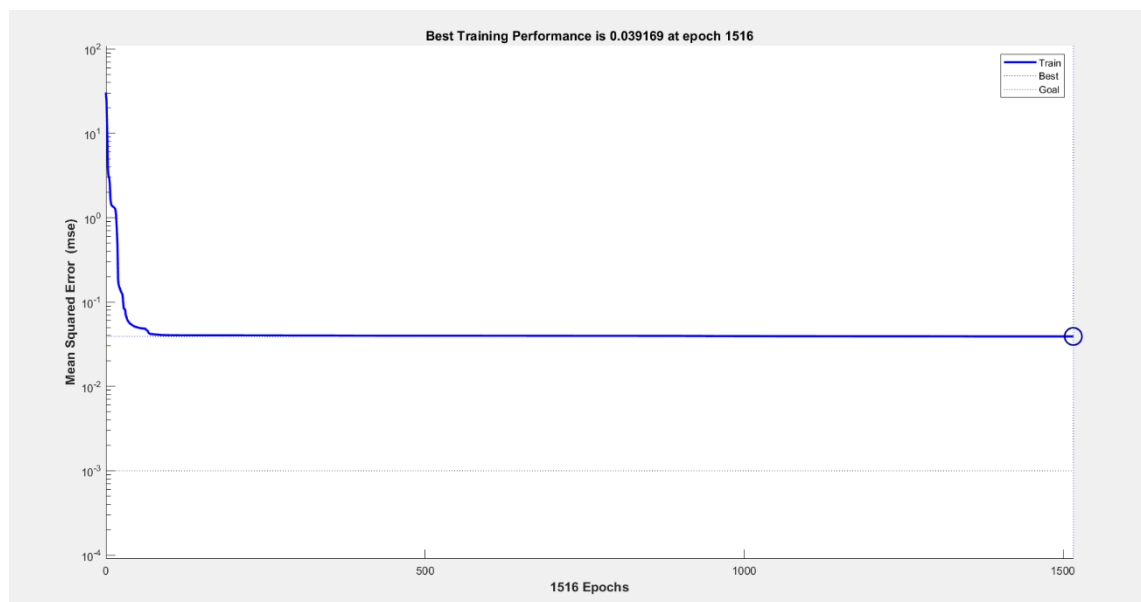
Grafik funkcija  $h(x)$  i  $y(x)$ :



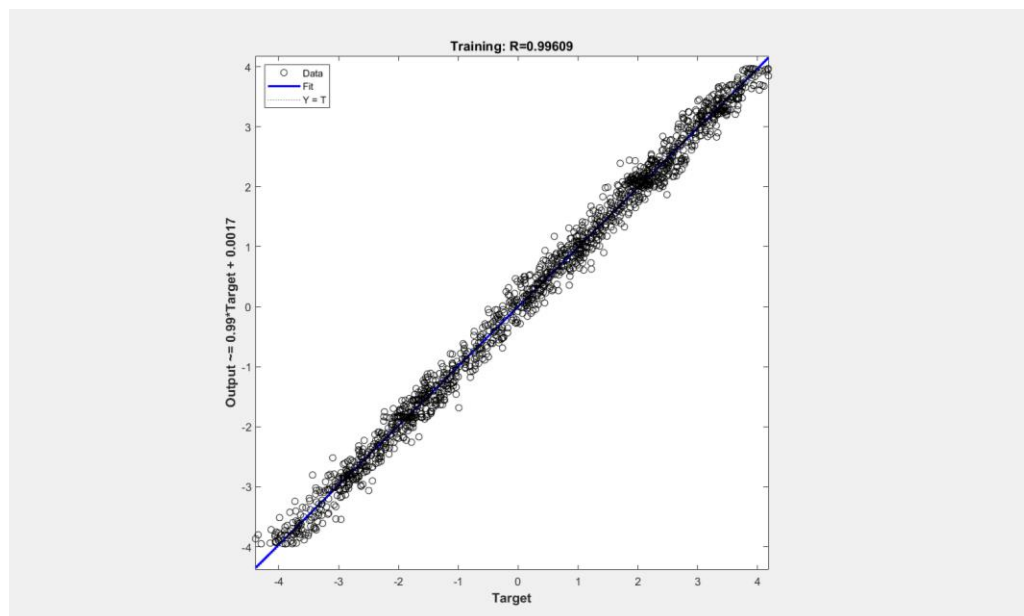
Kreirana je neuralna mreza sa 2 skrivena sloja koji imaju 10 i 6 neurona.

Aktivaciona funkcija u skrivenim slojevima je tansig,a u izlaznom sloju purelin.

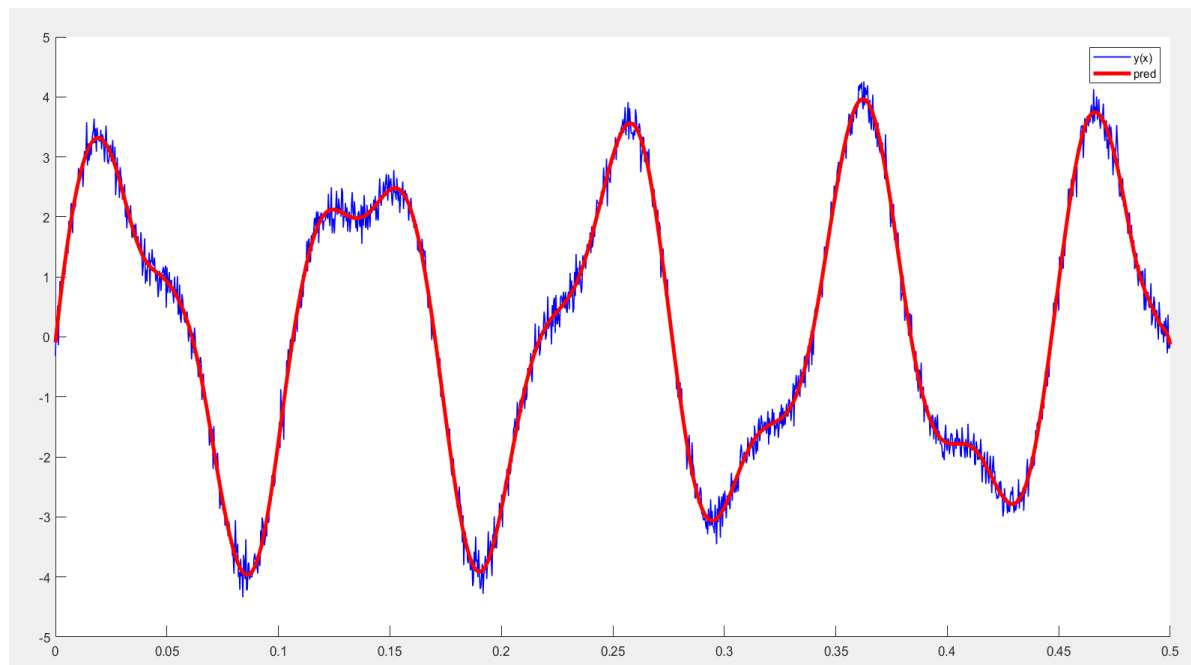
Kriva performanse:



Regresiona kriva:



Grafik funkcija  $y(x)$  i predikcije neuralne mreže:



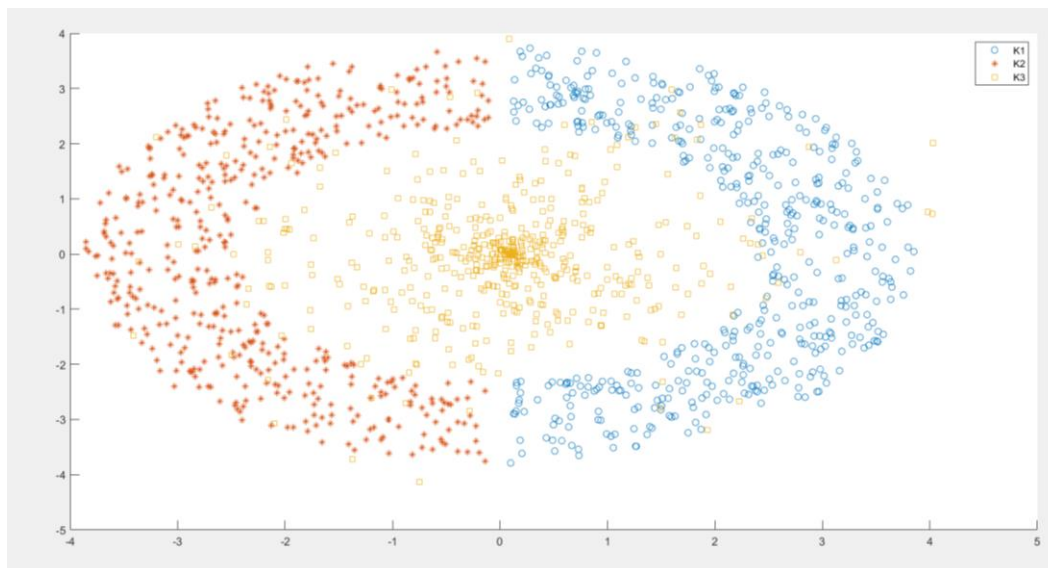
Mreža je dobro istrenirana, što pokazuje i velika vrednost regresije,

pa izlaz predikcije skoro u potpunosti odgovara izlazu originalne funkcije.

## 2.Zadatak

Radi se sa dataset2.mat datotekom.Postoje 3 klase obeležene brojevima 1,2,3.

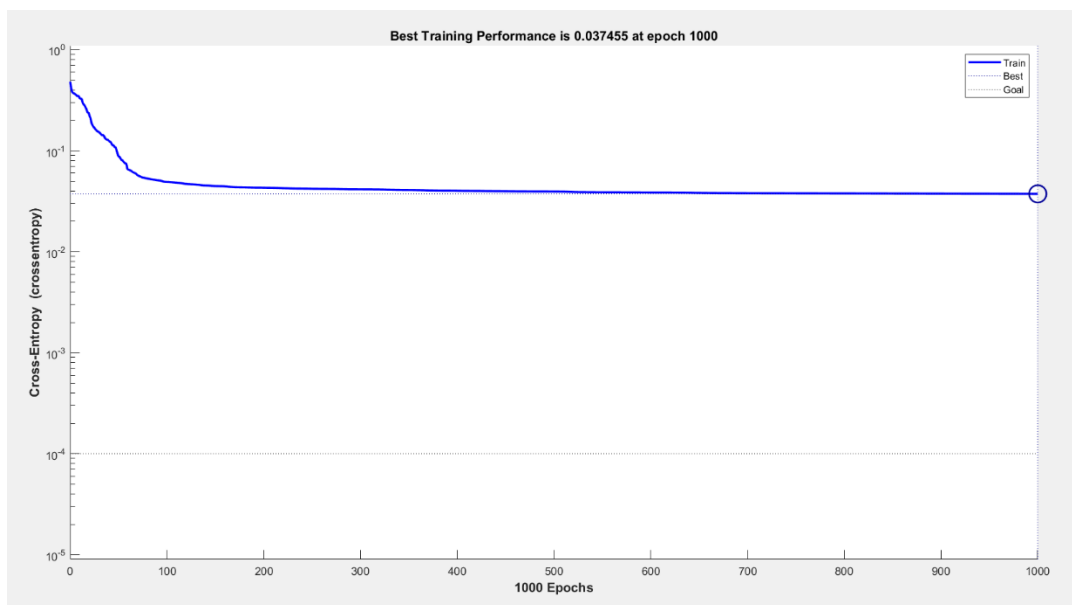
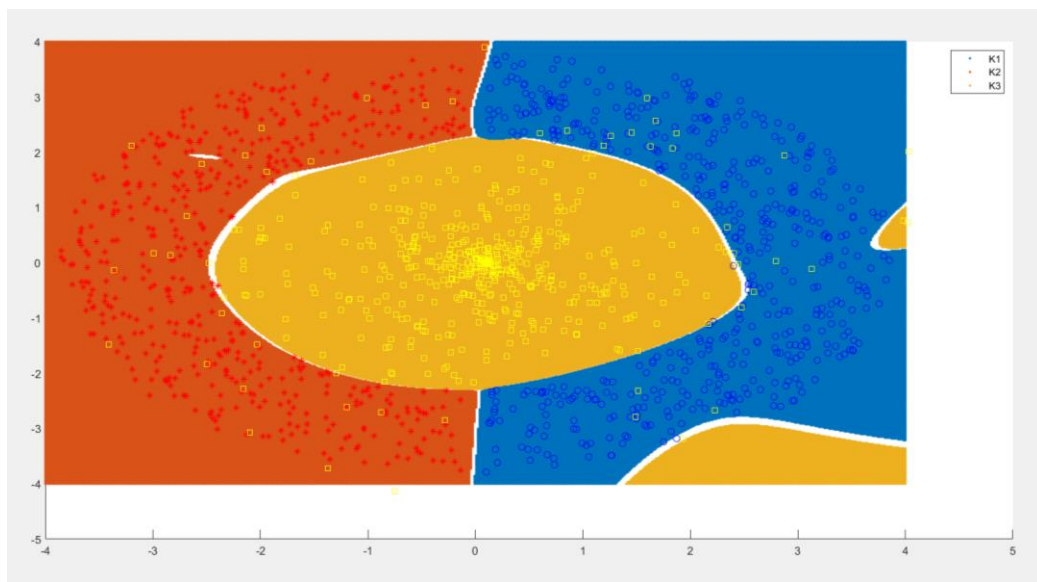
Prikaz podataka:



Posto se podaci dele na 3 grupe, izlazi su kodirani sa 3 neurona tako da za klasu 1 izlaz izgleda 100, za klasu 2 010, a za klasu 3 001 (one-hot encoding). Funkcije aktivacije koje su korišćene su tansig u skrivenim i softmax u izlaznom sloju. Podaci koji se dobiju simulacijom su razdvajani tako da ukoliko je izlaz veći od 0.7 aproksimira se na 1 a ukoliko je manji aproksimira se na 0. Na slikama su prikazani grafici klasifikacije, konfuzione matrice i grafici performansi treniranja.

Podela podataka je bitna jer želimo da mrežu naučimo da prepoznaje primerke sve 3 klase. Ne želimo da mreža uči samo na primercima jedne klase tako da podelu vršimo na slučajan način. 90% podataka je odvojeno za treniranje mreže, a 10% za njeno testiranje.

Arhitektura 1: mreza sa rasporedom neurona u skrivenim slojevima 3, 6, 6:



Confusion Matrix				
Output Class	1	2	3	
	453 33.6%	0 0.0%	19 1.4%	96.0% 4.0%
	0 0.0%	448 33.2%	23 1.7%	95.1% 4.9%
	2 0.1%	1 0.1%	404 29.9%	99.3% 0.7%
				99.6% 0.4%
				99.8% 0.2%
				90.6% 9.4%
				96.7% 3.3%
				Target Class
				1
				2
				3

Parametri na trening skupu:

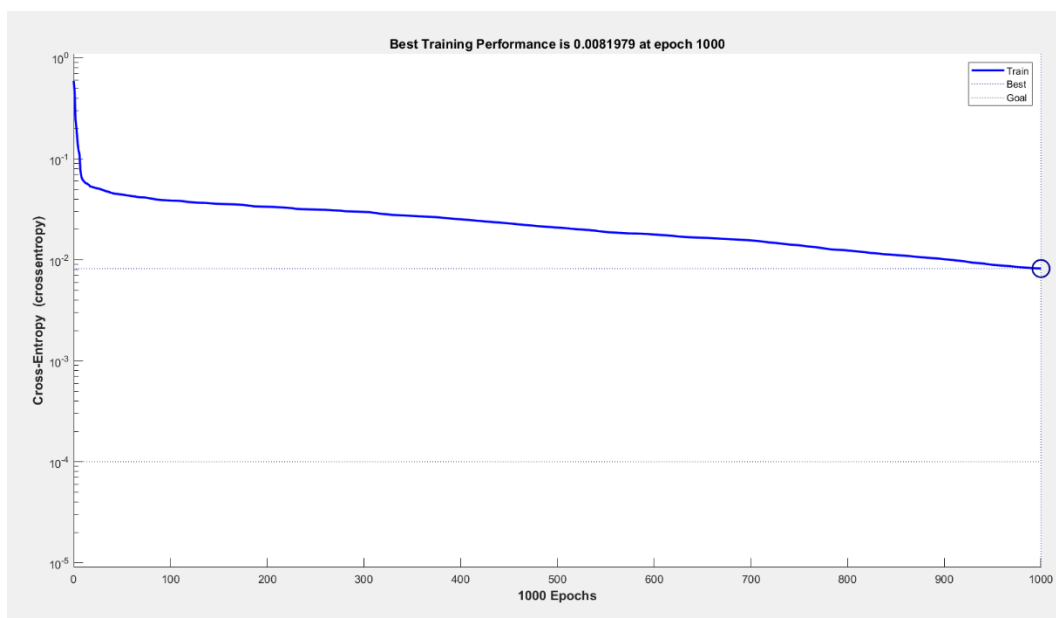
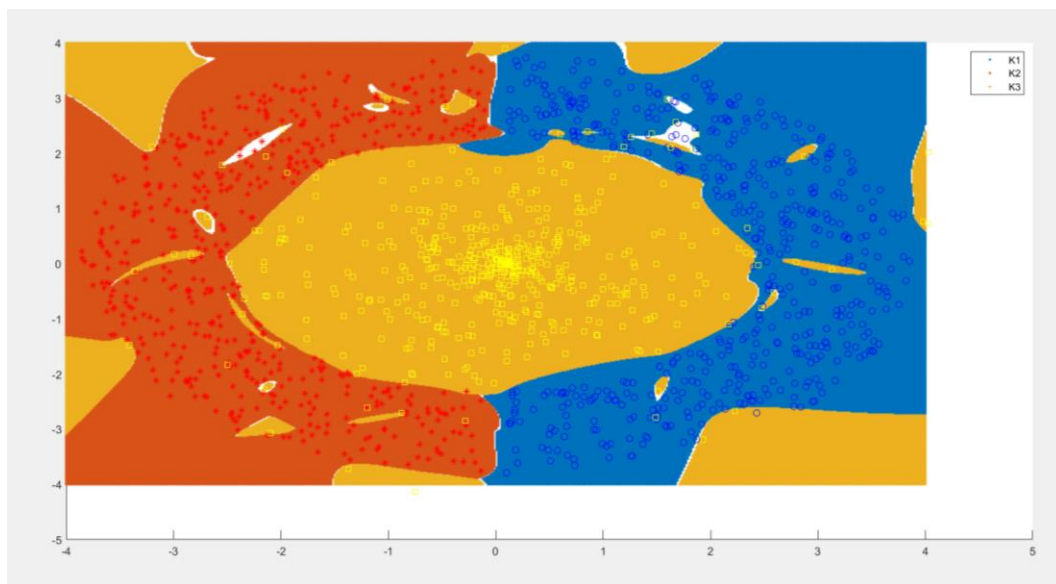
Precision = 0.9597, Recall = 0.9956

Confusion Matrix				
Output Class	1	2	3	
	45 30.0%	0 0.0%	3 2.0%	93.8% 6.3%
	0 0.0%	51 34.0%	4 2.7%	92.7% 7.3%
	0 0.0%	0 0.0%	47 31.3%	100% 0.0%
				100% 0.0%
				100% 0.0%
				87.0% 13.0%
				95.3% 4.7%
				Target Class
				1
				2
				3

Parametri na test skupu:

Precision = 0.9375, Recall = 1

Arhitektura 2: mreza sa rasporedom neurona u skrivenim slojevima 15, 15, 15:





Confusion Matrix				
Output Class	1	2	3	
	444 32.9%	0 0.0%	7 0.5%	98.4% 1.6%
	0 0.0%	445 33.0%	4 0.3%	99.1% 0.9%
	1 0.1%	1 0.1%	448 33.2%	99.6% 0.4%
	99.8% 0.2%	99.8% 0.2%	97.6% 2.4%	99.0% 1.0%
	1	2	3	Target Class

Parametri na trening skupu:

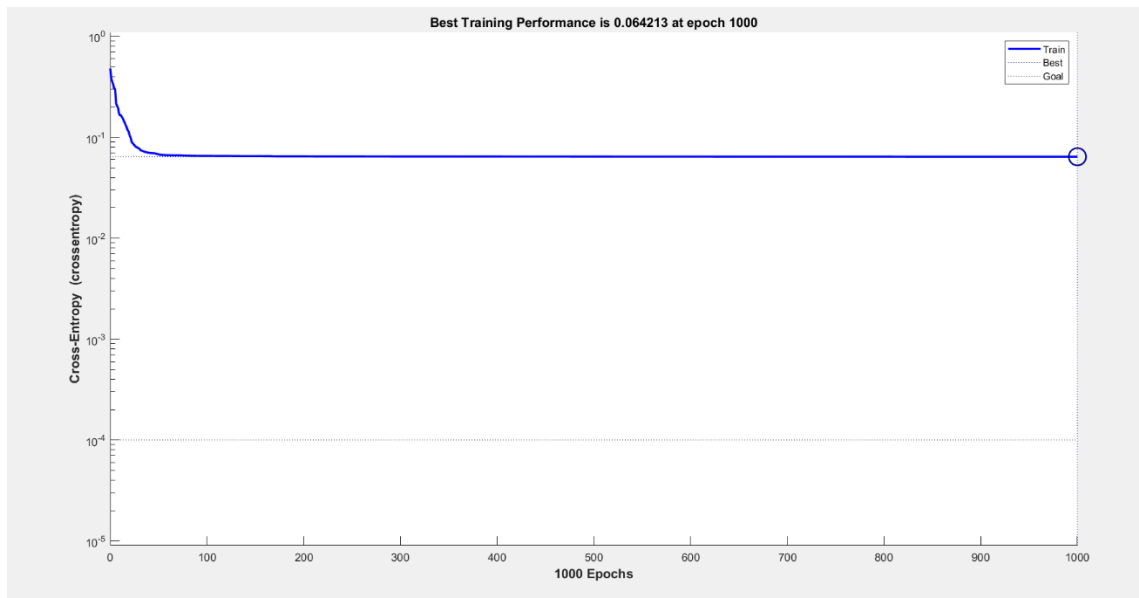
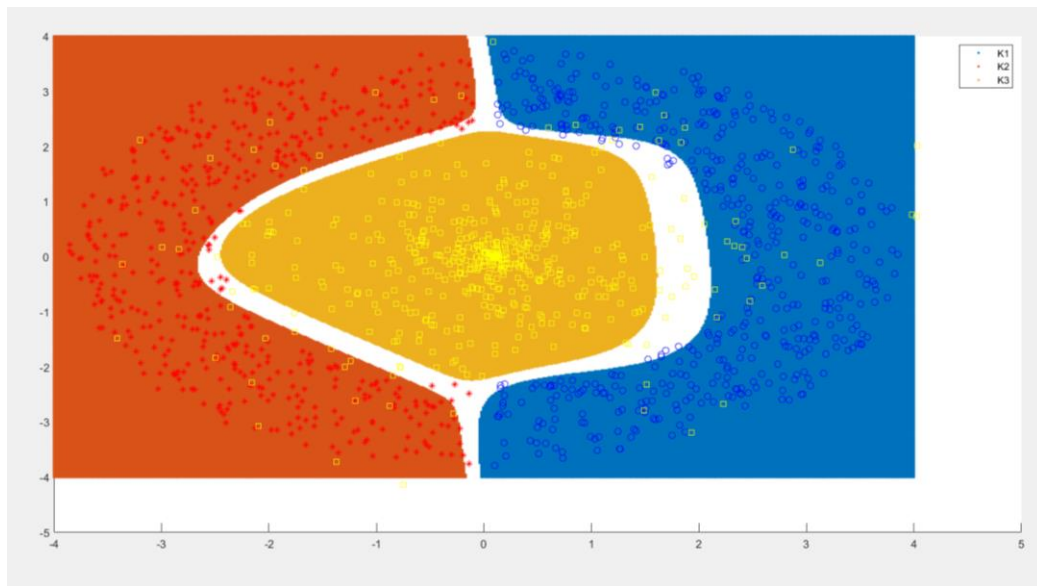
Precision = 0.9845, Recall = 0.9978

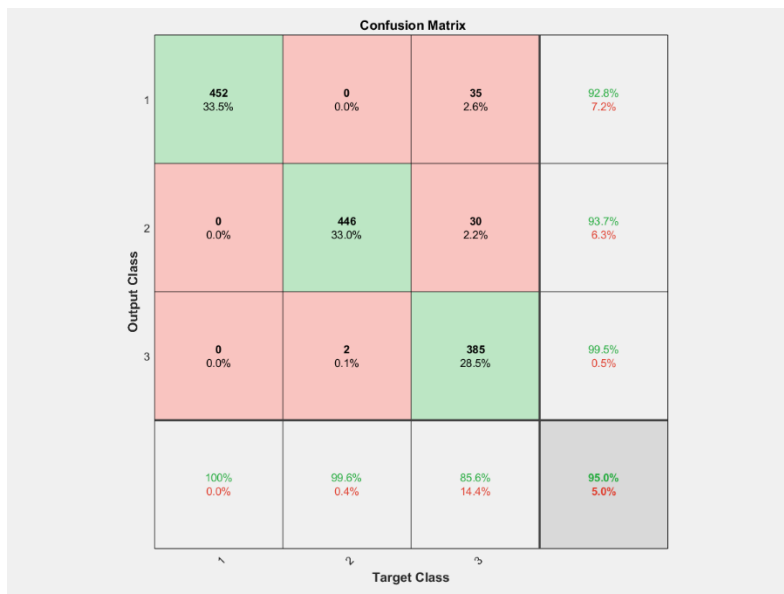
Confusion Matrix				
Output Class	1	2	3	
	54 36.0%	0 0.0%	3 2.0%	94.7% 5.3%
	0 0.0%	51 34.0%	3 2.0%	94.4% 5.6%
	1 0.7%	3 2.0%	35 23.3%	89.7% 10.3%
				Target Class
				1
				2
				3

Parametri na test skupu:

Precision = 0.9474, Recall = 0.9818

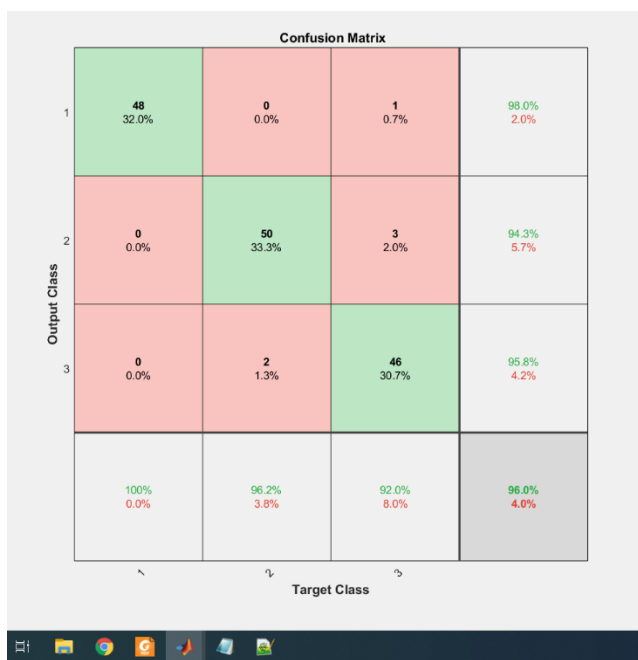
### Arhitektura 3: mreza sa rasporedom neurona u jednom skrivenom sloju 3:





Parametri na trening skupu:

Precision = 0.9281, Recall = 1



Parametri na test skupu:

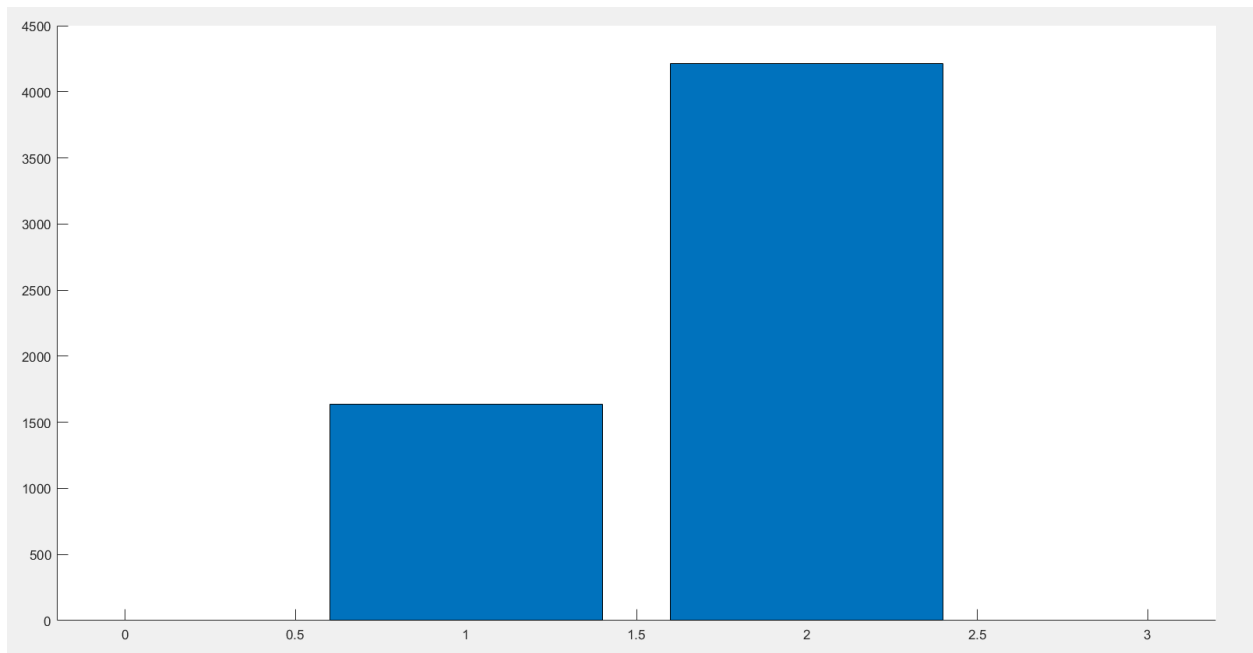
Precision = 0.9796, Recall = 1

U trecoj arhitekturi imamo приметно vise belih podrucja na slici, jer usled nedostatka neurona mreza ne moze sa zadatom verovatnocom da izvrši klasifikaciju.

### **3.Zadatak**

Na osnovu brojeva indeksa radi se sa Star datotekom. Problem koji se resava je klasifikacija pulsara. Pulsari su retka vrsta neutronske zvezde koji emituju zracenje. Svaki kandidat za pulsar je opisan sa 9 atributa, od kojih je poslednji indikator da li je kandidat pulsar (0 vrednost ako nije, 1 ako jeste).

Podaci nisu balansirani. Klasa 0 ima mnogo vise primeraka od klase 1 (4213 za K0, 1639 za K1).



Izvršena je podela na trening, validacioni i test skup u odnosu 80:10:10.

Posto je skup nebalansiran, kao metriku za određivanje tačnosti neuralne mreže koristimo harmonijsku sredinu parametara precision i recall, F1. Za hiperparametar struktura mreže uzete su višeslojne mreže sa različitim brojem neurona u njima. Za hiperparametar funkcija aktivacije uzete su funkcije tansig i logsig. Za hiperparametar regularizacija uzete su vrednosti između 0.2 i 0.8, posto on može da uzme vrednosti i bliže 0 i bliže 1. Za hiperparametar koeficijent težinskih faktora uzete su vrednosti veće od 1, da bi se otežale vrednosti za klasu 1 koja ima manje vrednosti.

Vrednosti optimalnih parametara:

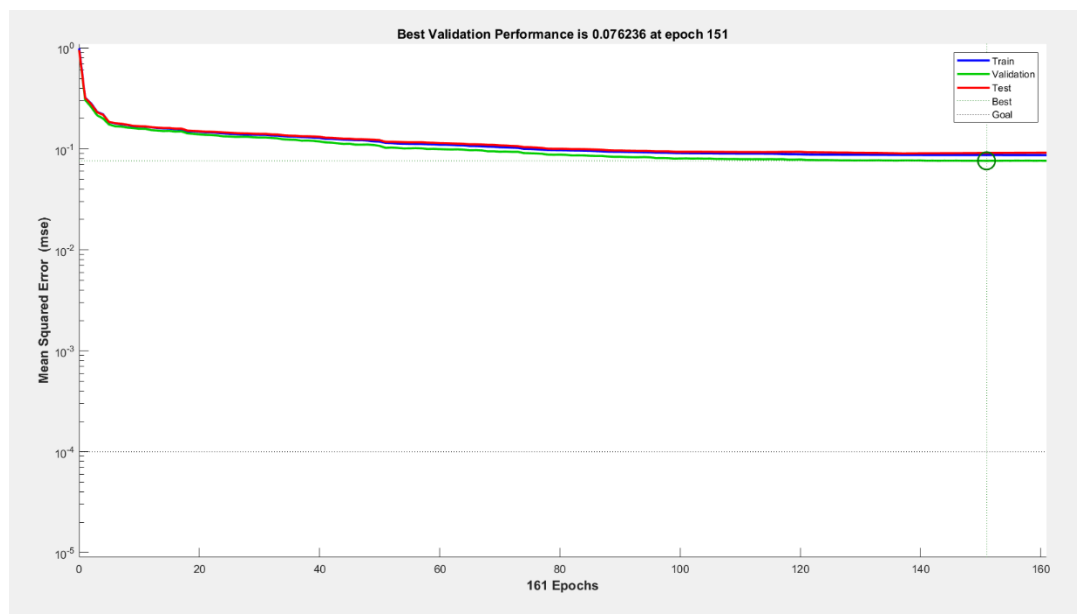
struktura mreže:20 15

funkcija aktivacije:tansig

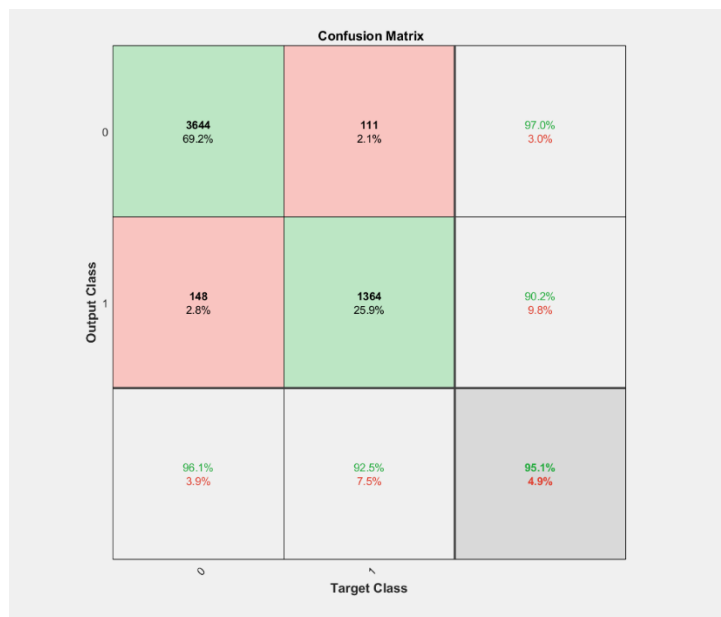
koeficijent regularizacije:0.2

koeficijent tezinskih faktora:4

Kriva performanse neuralne mreže nakon obucavanja optimalnim parametrima:



Matrica konfuzije za trening skup:

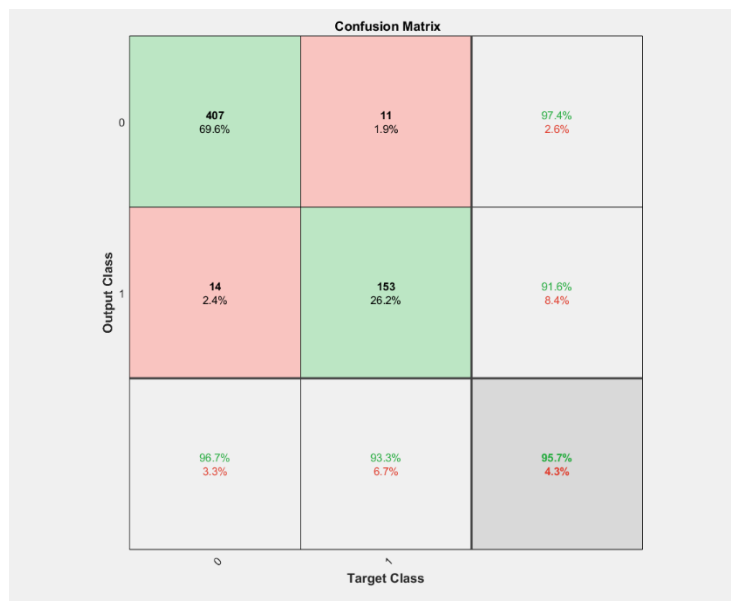


Dobili smo visoke vrednosti za parametre precision i recall, pa je i harmonijska sredina tih parametara visoka.

Parametri na trening skupu za K1:

Precision = 0.9021, Recall = 0.9247, F1 = 0.9133

Matrica konfuzije za test skup:



Ponovo smo dobili visoke vrednosti za parametre precision i recall, pa je i harmonijska sredina tih parametara visoka. Mreza dobro prepoznaje primerke klase.

Parametri na test skupu za K1:

Precision = 0.9162, Recall = 0.9329, F1 = 0.9245