

Domaći zadatak iz neuralnih mreža



Beograd, januar 2021.

Profesori:

Goran Kvaščev
Aleksandar Rakić
Marija Novičić

Studenti:

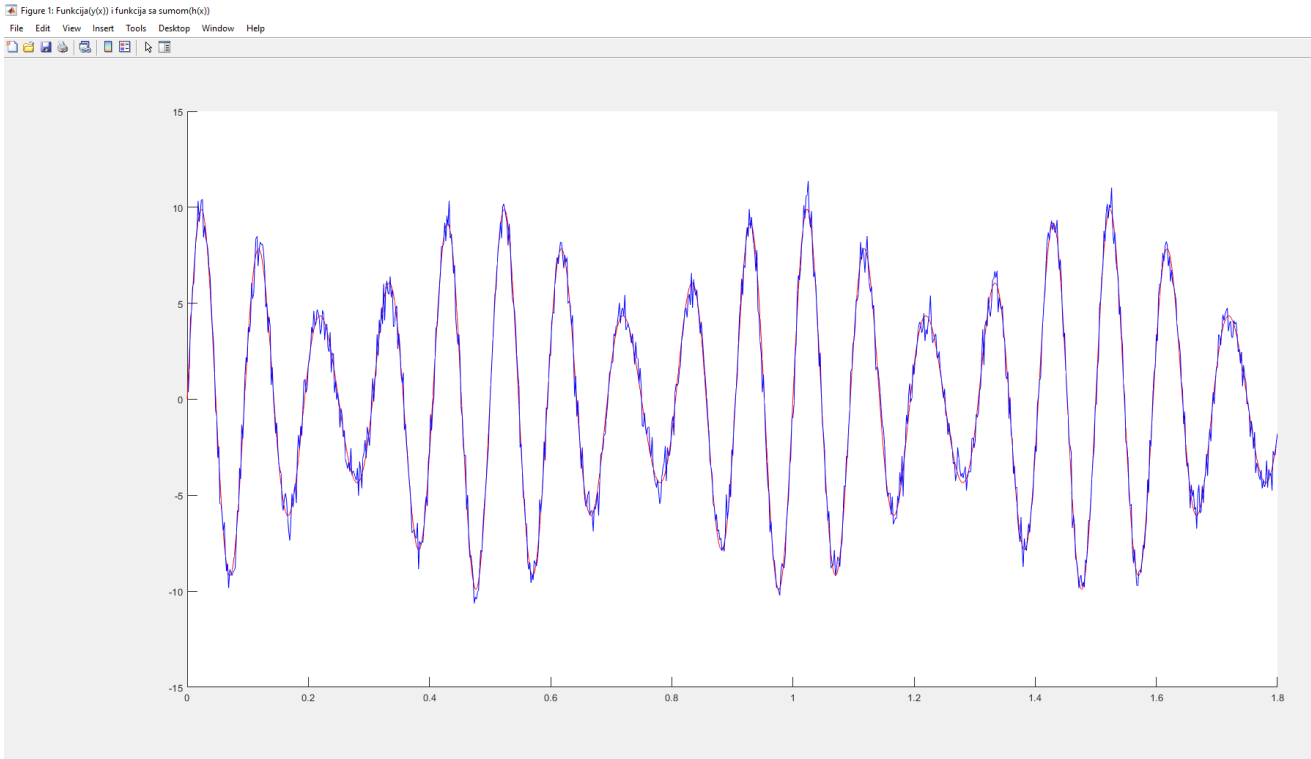
Luka Tomanović 0410/2018
Kosta Matijević 0034/2018

1.

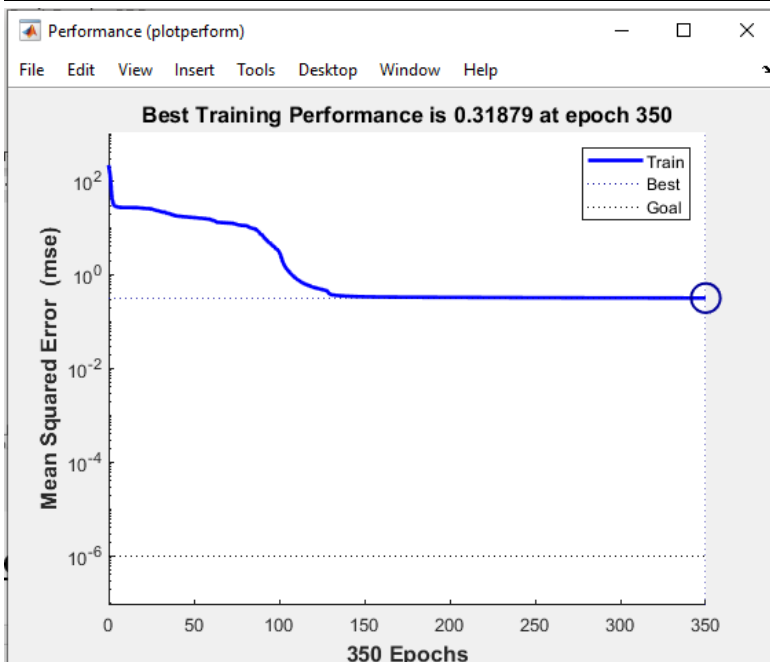
Funkcija $y(x)$ i funkcija sa sumom $h(x)$

Funkcija bez šuma na grafiku je označena crvenom bojom, i oblika je $h(x) = 7 * \sin(2\pi * 10 * x) + 3 * \sin(2\pi * 12 * x)$.

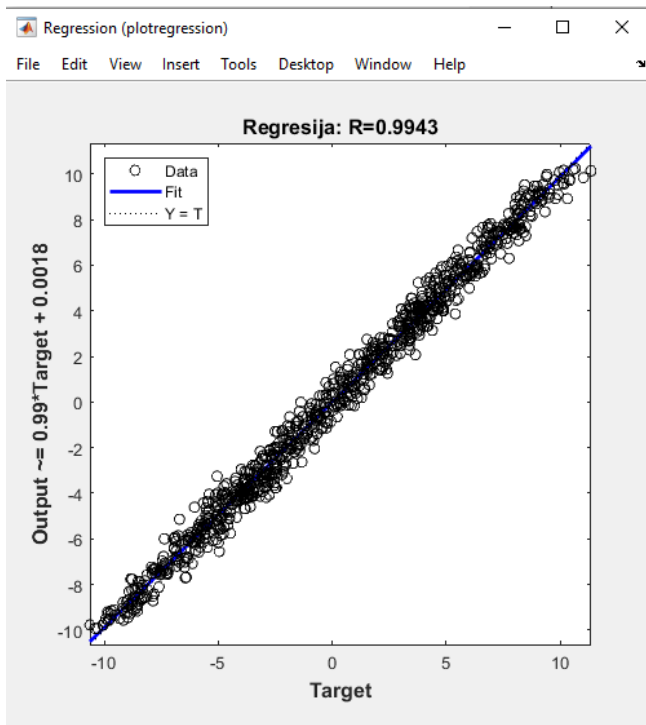
Funkcija sa slučajnim šumom, standardne devijacije $std = 0.6$, prikazana je plavom bojom, i oblika je $y(x) = h(x) + s(x)$.



Kriva performanse u zavisnosti od epohe treniranja

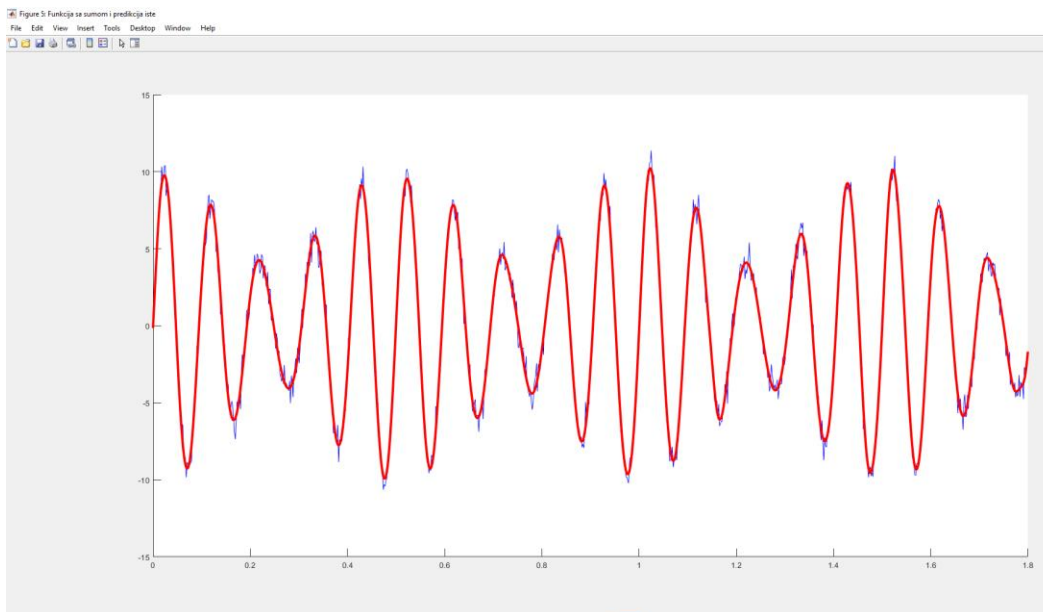


Regresiona kriva



Možemo videti da postoji skoro linearna veza između predikcije i željenog izlaza ($R=0.9943$). Predikcijom neuralne mreže dobijeni su skoro isti željeni izlazi.

Funkciju $y(x)$ i predikciju neuralne mreže na celom skupu podataka x



Crvena linija predstavlja predikciju mreže, dok je funkcija sa šumom predstavljena plavom bojom. Mreža je prilično lepo istrenirana što možemo videti sa slike.

2.

Vizuelizacija podataka



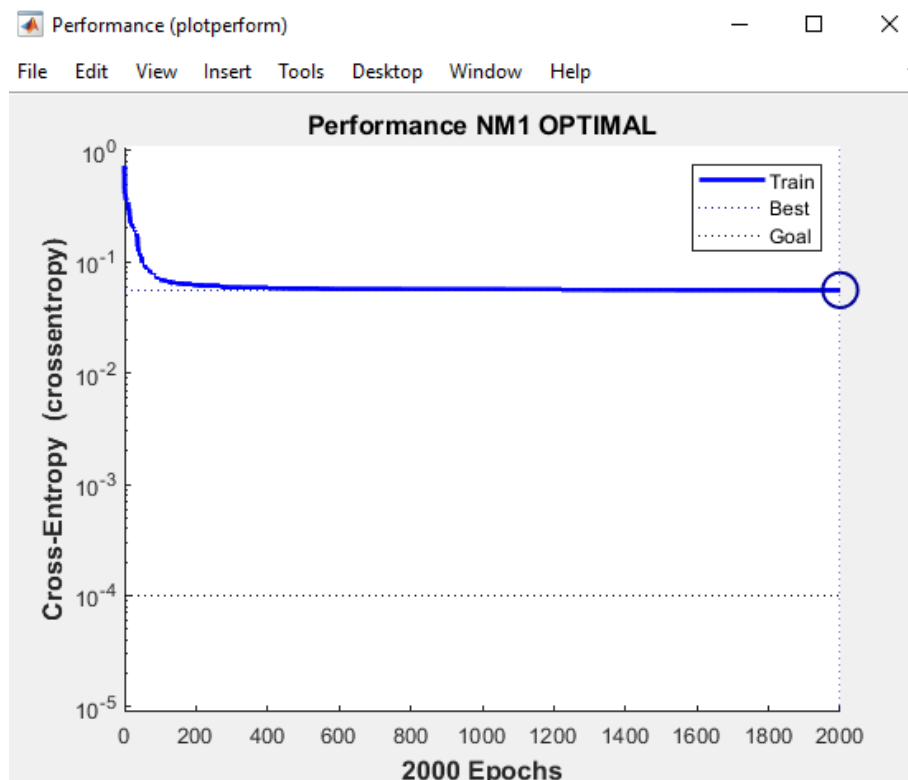
Podaci nisu idealni jer postoji preklapanje između klasa. U daljem tekstu odбирке obeležene žutom bojom ćemo smatrati klasom K3, crvenom bojom klasom K2 a plavom bojom klasom K1.

Podela podataka

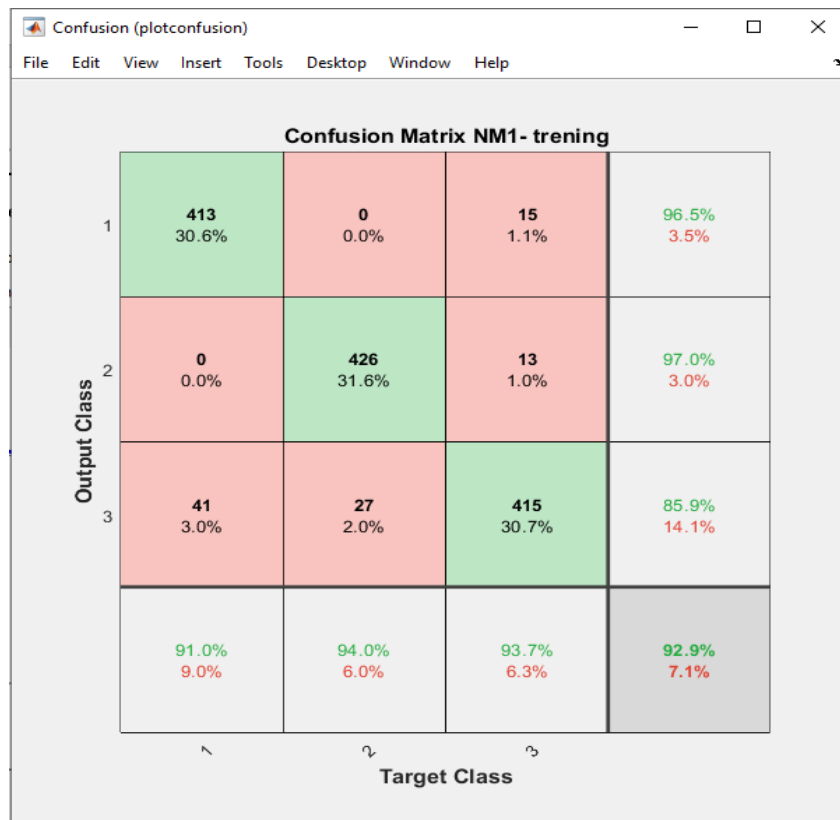
Podela podataka je izvršena tako da je 90% odбирaka uzeto za trening skup, a 10% za test skup. Podaci za validacioni skup nisu korišćeni jer se ne traži zaštita od preobučavanja. Podela je tako uređena da bismo imali i podatke koji se ne koriste u fazi treniranja (nepoznati su mreži) već se koriste samo za testiranje performansi mreže. Na taj način možemo videti da li se mreža naučila na podatke sa kojima je vršeno treniranje, odnosno da li je preobučena.

Analiza optimalne arhitekture

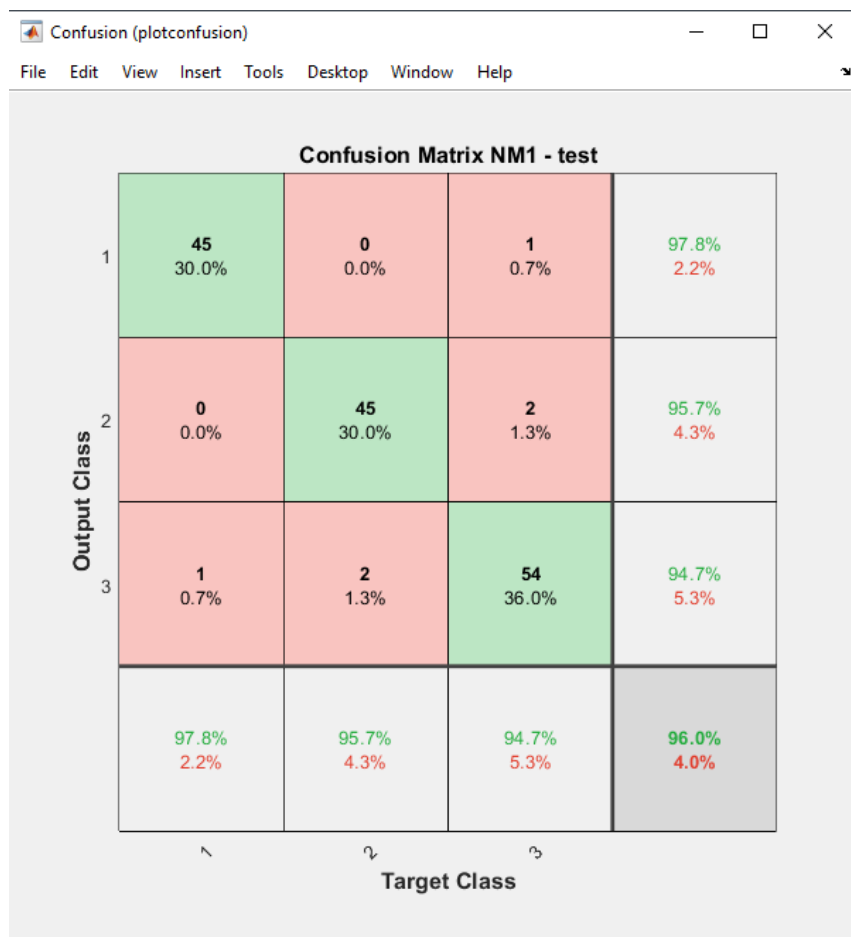
Kriva performanse u zavisnosti od epohe treniranja.



Konfuzionna matrica za trening skup:



Konfuzionna matrica za test skup:

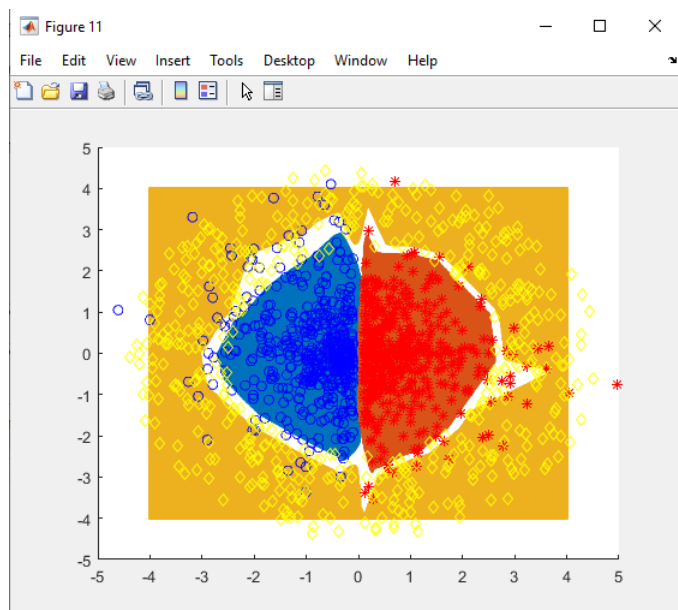


Precision i recall za klasu 1:

Precision: 97.83%

Recall: 97.83%

Granica odlučivanja

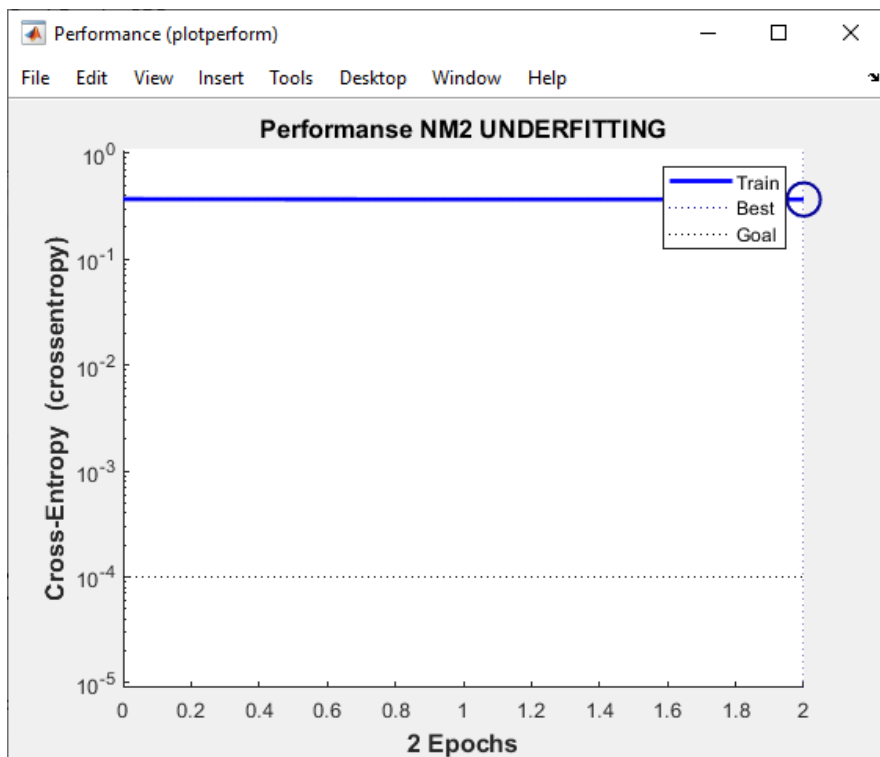


Granica je prilično jasna ali zbog preklapanja odbiraka klasa postoji određeno odstupanje. Podaci su klasifikovani u određenu klasu samo ako ih je mreža sa više od 70% sigurnosti u tu klasu i ubacila.

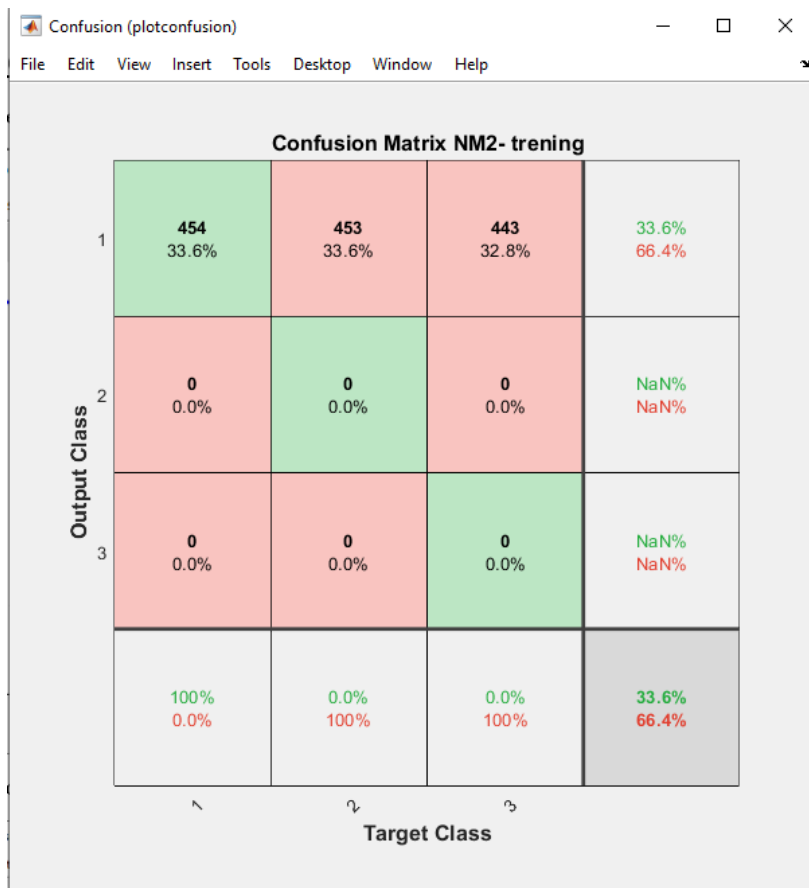
```
K1go = ulazGO(:, predGO(1, :) >= 0.7);
K2go = ulazGO(:, predGO(2, :) >= 0.7);
K3go = ulazGO(:, predGO(3, :) >= 0.7);
```

Analiza underfitting arhitekture

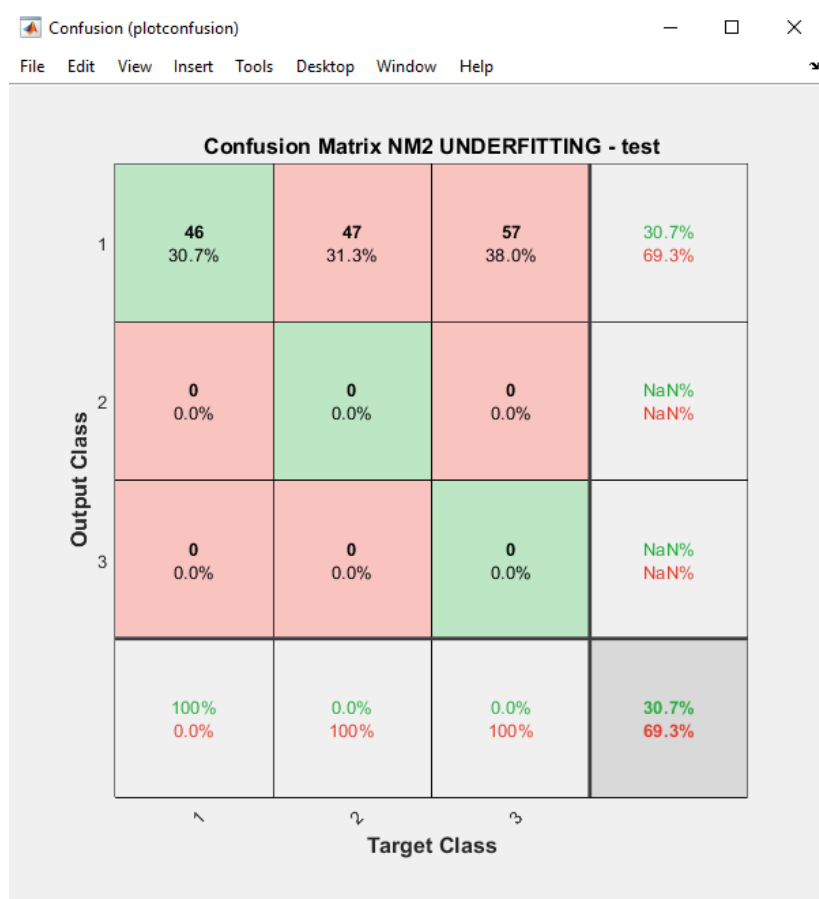
Kriva performanse u zavisnosti od epohe treniranja.



Konfuziona matrica za trening skup:



Konfuzionna matrica za test skup:

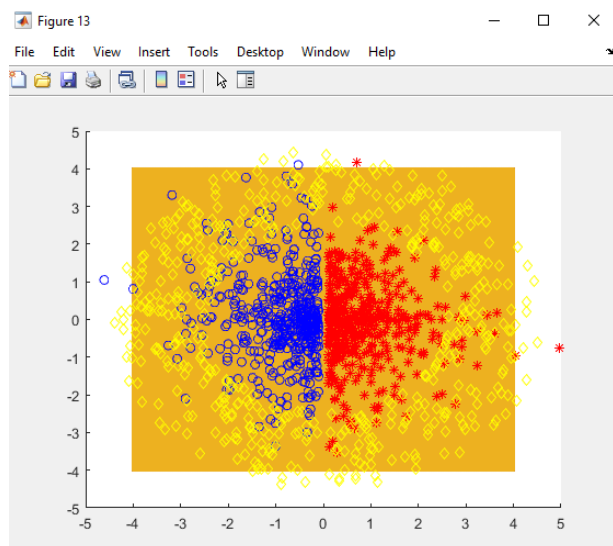


Precision i recall za klasu 1:

Precision: 30.67%

Recall: 100%

Granica odlučivanja

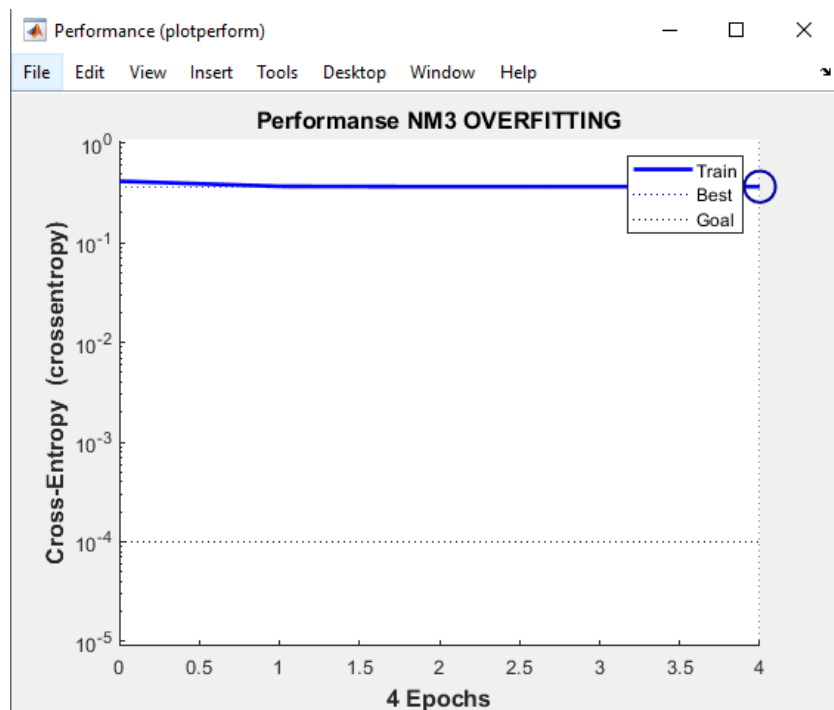


Kod ove mreže smo morali spustiti procenat sigurnosti mreže prilikom odredjivanja kojoj klasi odbirak pripada. Jer nijedan podatak nije klasifikovan sa sigurnošću od 70%. Zato smo ovde stavili da je potreban procenat sigurnosti od najmanje 30%, ali i tada dobijamo loše (očekivane rezultate) gde se vidi da naša mreža sve podatke klasifikuje u klasu 1. Ovo je u skladu sa prethodno prikazanom konfuzionom matricom.

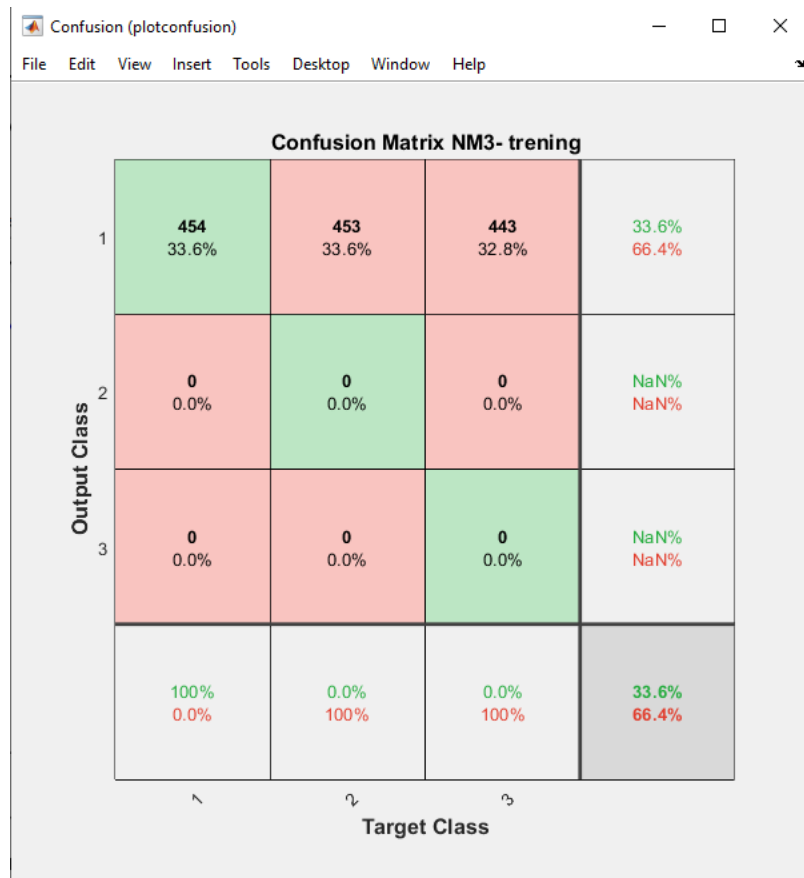
```
K1go = ulazGO(:, predGO(1, :) >= 0.3);
K2go = ulazGO(:, predGO(2, :) >= 0.3);
K3go = ulazGO(:, predGO(3, :) >= 0.3);
```


Analiza overfitting arhitekture

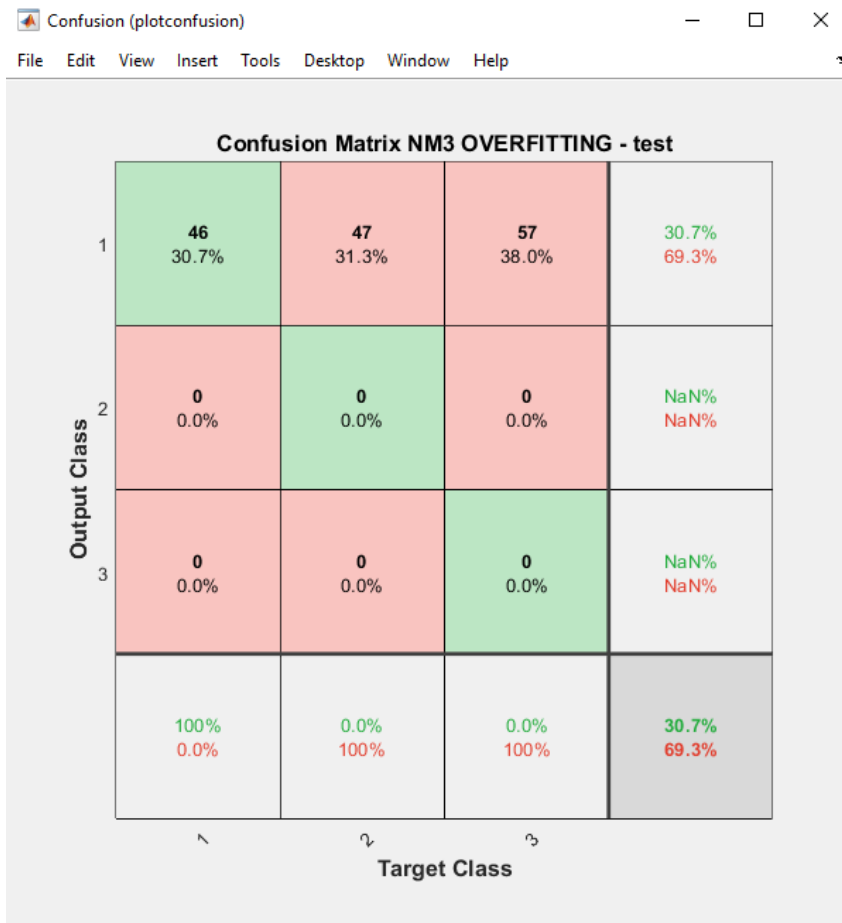
Kriva performanse u zavisnosti od epohe treniranja.



Konfuziona matrica za trening skup:



Konfuzionna matrica za test skup:

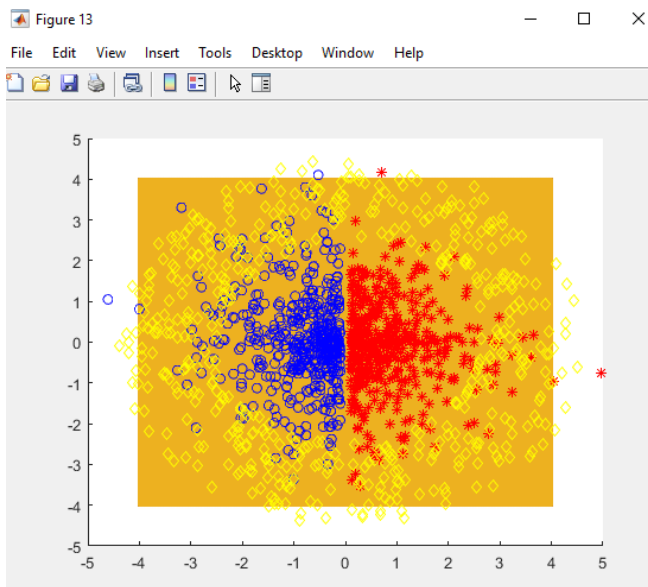


Precision i recall za klasu 1:

Precision: 30.67%

Recall: 100%

Granica odlučivanja



Kod ove mreže smo morali spustiti procenat sigurnosti mreže prilikom određivanja kojoj klasi odbirak pripada. Jer nijedan podatak nije klasifikovan sa sigurnošću od 70%. Zato smo ovde stavili da je potreban procenat sigurnosti od najmanje 30%, ali i tada dobijamo loše (očekivane rezultate) gde se vidi da naša mreža sve podatke klasifikuje u klasu 1. Ovo je u skladu sa prethodno prikazanom konfuzionom matricom.

```
K1go = ulazGO(:, predGO(1, :) >= 0.3);
K2go = ulazGO(:, predGO(2, :) >= 0.3);
K3go = ulazGO(:, predGO(3, :) >= 0.3);
```

3.

Opis problema

Rešava se problem predikcije (**regresije**) količine emitovanog ugljen-monoksida (CO) na osnovu ulaznih parametara koje koristi u radu gasna turbina.

Ulazni parametri su:

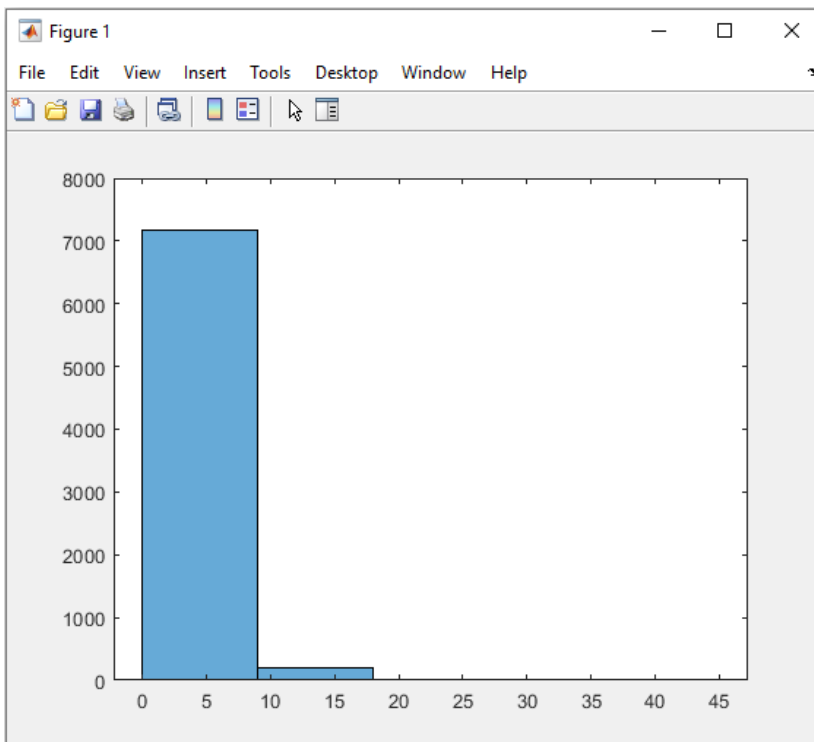
Ambient temperature (AT) C
Ambient pressure (AP) mbar
Ambient humidity (AH) (%)
Air filter difference pressure (AFDP) mbar
Gas turbine exhaust pressure (GTEP) mbar
Turbine inlet temperature (TIT) C
Turbine after temperature (TAT) C
Compressor discharge pressure (CDP) mbar
Turbine energy yield (TEY) MWH

Izlazni su:

Carbon monoxide (CO) mg/m³

Nitrogen oxides (NO_x) mg/m³ (nije od interesa)

Pregled podataka



Ulazni podaci su klasifikovani u pet klasa sa korakom od 9 mg/m³ na osnovu izlaznog parametra CO. Sa slike vidimo da su podaci loše balansirani i da skoro svi pripadaju prvoj klasi, i nešto malo drugoj. Podaci ostalih klasa su jako slabo zastupljeni pa se ni ne vide na histogramu.

Podela podataka(odbiraka)

Od svake od pet klasa 80% obdiraka uzima se za trening skup, 10% za validacioni i preostalih 10% za testiranje.

Hiperparametri (i njihove optimalne vrednosti)

Hiperparametri su struktura, funkcije aktivacije, koeficijent regularizacije, koeficijent težinskih faktora.

Struktura NM:

Povećanje broja slojeva iznad tri nije dalo povećanja u performansama NM, stoga smo zaključili da je minimalan optimalan broj slojeva tri. Strukture koje su uzimane u obzir su: $\{[10 \ 50 \ 10], [20 \ 40 \ 30], [10 \ 20 \ 10]\}$. Metodom kros-validacije došli smo do toga da je najbolja struktura: **[10 50 10]**.

Funkcija aktivacije:

Za svaki od slojeva pojedinačno je metodom kros-validacije određena funkcija aktivacije koja je dala najbolje rezultate. U obzir su uzimane sledeće funkcije aktivacije: tansig, logsig i poslin(ReLU).

Za prvi sloj kao najbolja se koristi **poslin** funkcija aktivacije.

Za drugi sloj kao najbolja se koristi **tansig** funkcija aktivacije.

Za treći sloj kao najbolja se koristi **logsig** funkcija aktivacije.

Koeficijent regularizacije:

Za koeficijent regularizacije uzimane su u obzir vrednosti 0-1. Metodom kros-validacije je utvrđeno da najbolja vrednost za koeficijent regularizacije iznosi **0.6**.

Koeficijent težinskih faktora:

Inicijalna matrica težina je definisana tako da su odbircima koji su slabije zastupljeni dodeljene veće težine, na sledeći način:

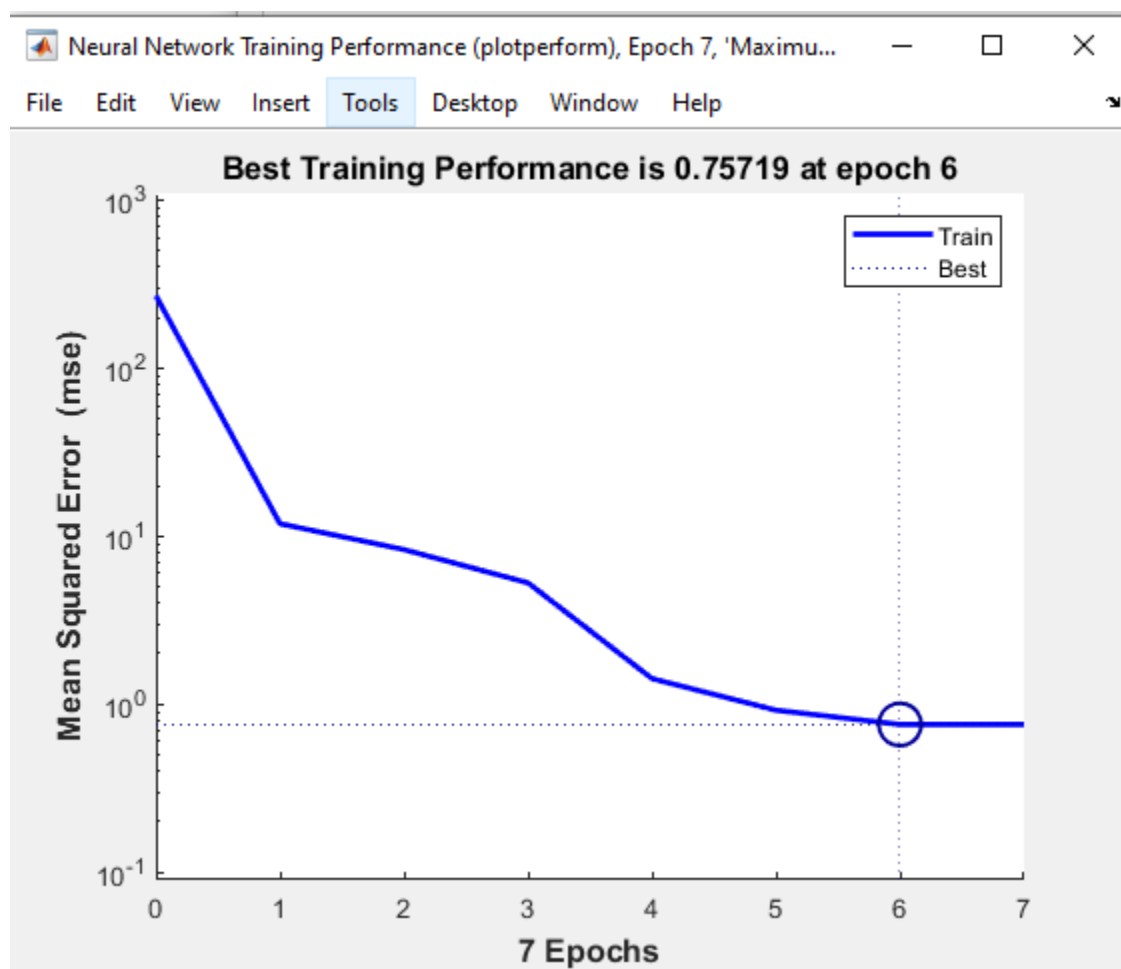
```
W = ones(1,length(izlazSve));  
W(izlazSve>10 & izlazSve<15) = weight;  
W(izlazSve>=15) = 2*weight;
```

Za parametar weight metodom kros-validacije dobijena je vrednost **weight=1**.

Mera performanse:

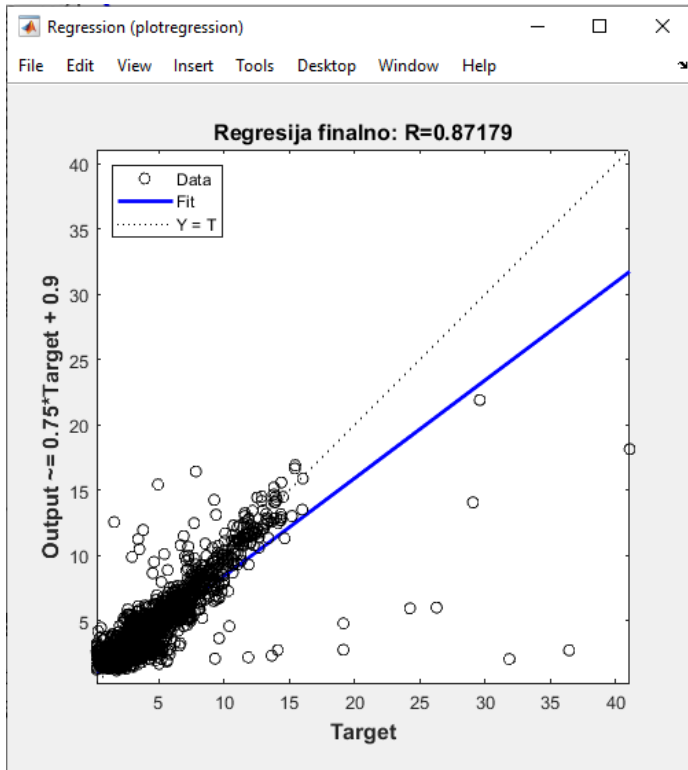
Kao mera performanse korišćena je vrednost **MSE** (srednja kvadratna greška). Minimalna dobijena vrednost srednje kvadratne greške tokom obučavanja metodom kros-validacije je **0.6373**.

Kriva performanse u zavisnosti od epoha treniranja



Regresiona kriva

-Na trening skupu(kada su određeni hiperparametri, za trening skup su korišćeni i podaci koji su ranije korišćeni za validaciju):



Zbog loše balansiranoosti podataka, mreža ne daje idealne rezultate(ranije analizirano), ali su ipak zadovoljavajući.

-Na test skupu

