|  |  |  |
| --- | --- | --- |
| UNIZNAK |  | FTNZNAK |

Projekat iz predmeta Algoritamske heuristike

Dušan Radišić E1 50-2022

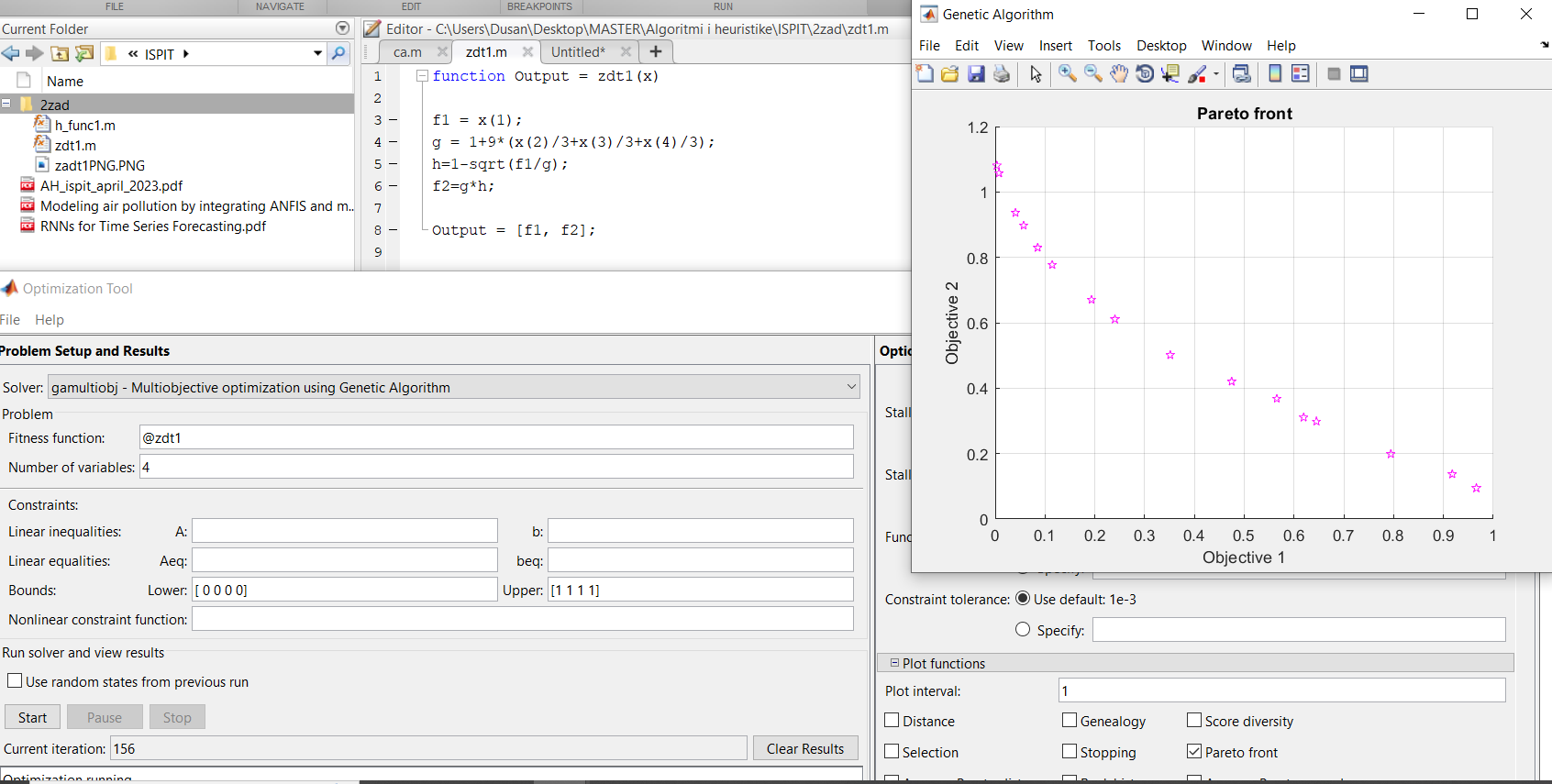
Sadržaj

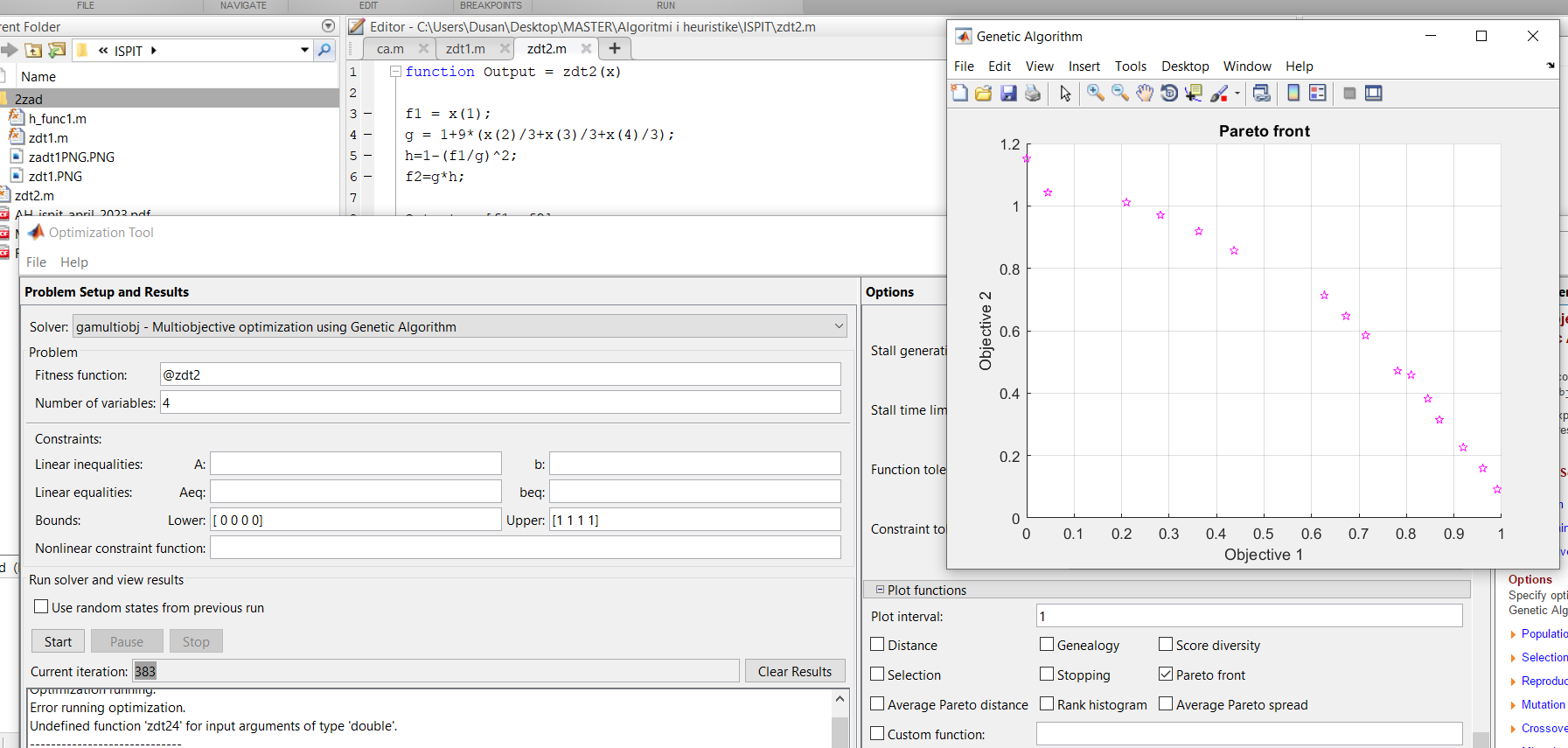
[1. Drugi zadatak - *Multiobj.* Optimizacije, Pareto front 1](#_Toc132319432)

[2. Prvi zadatak- ANFIS predikcija indeksa kvaliteta vazduha 4](#_Toc132319433)

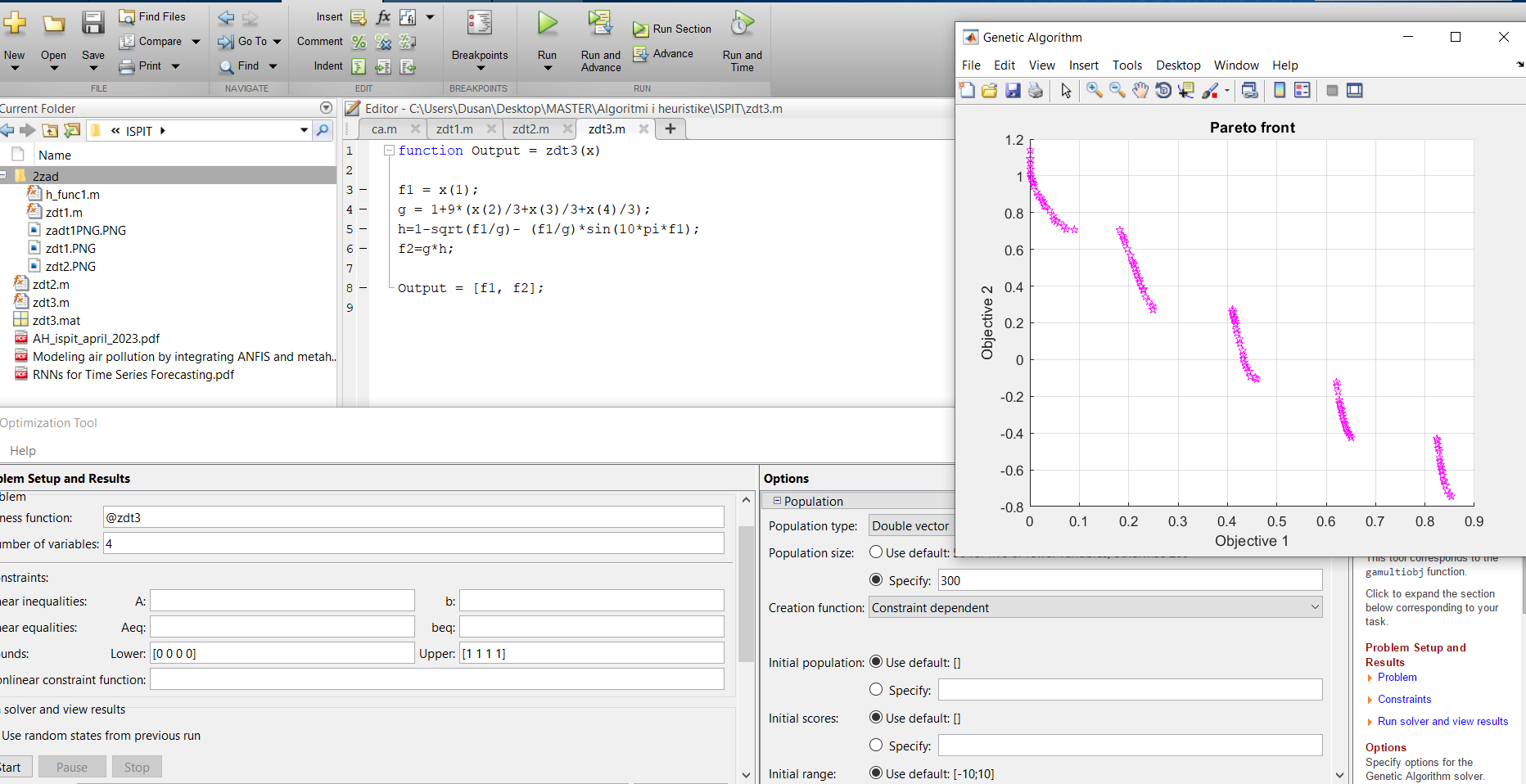
[3. Treći zadatak- Predikcija vremenskih serija korišćenjem Rekurentnih neuralnih mreša, RNN 8](#_Toc132319434)

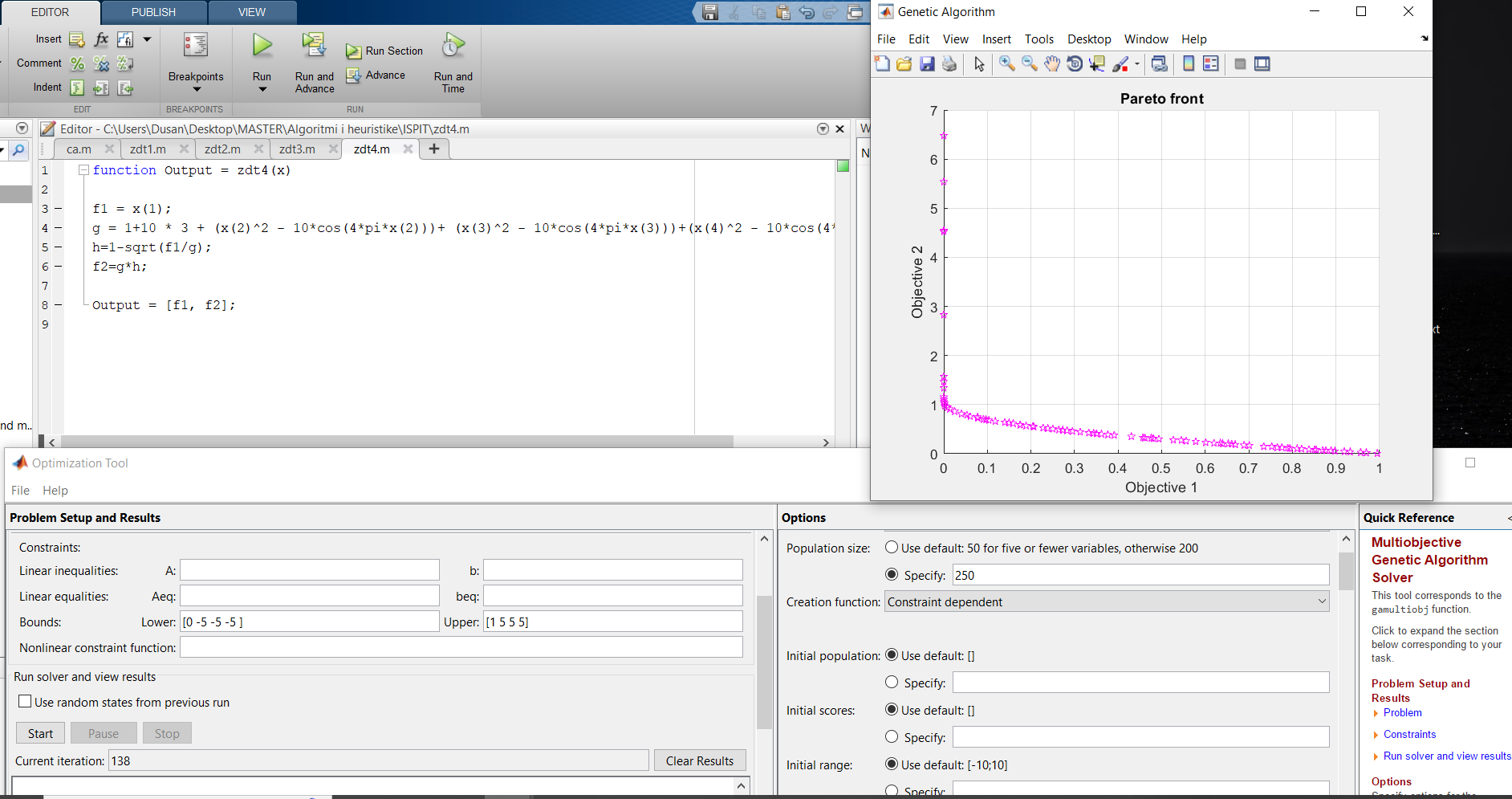
# 1. Drugi zadatak - *Multiobj.* optimizacije, Pareto front

Slede redom slike ekrana, po primerima u drugom zadatku. Na slikama je prikazan kod koji opisuje problem, konfiguracija GUI-ja i Pareto front.

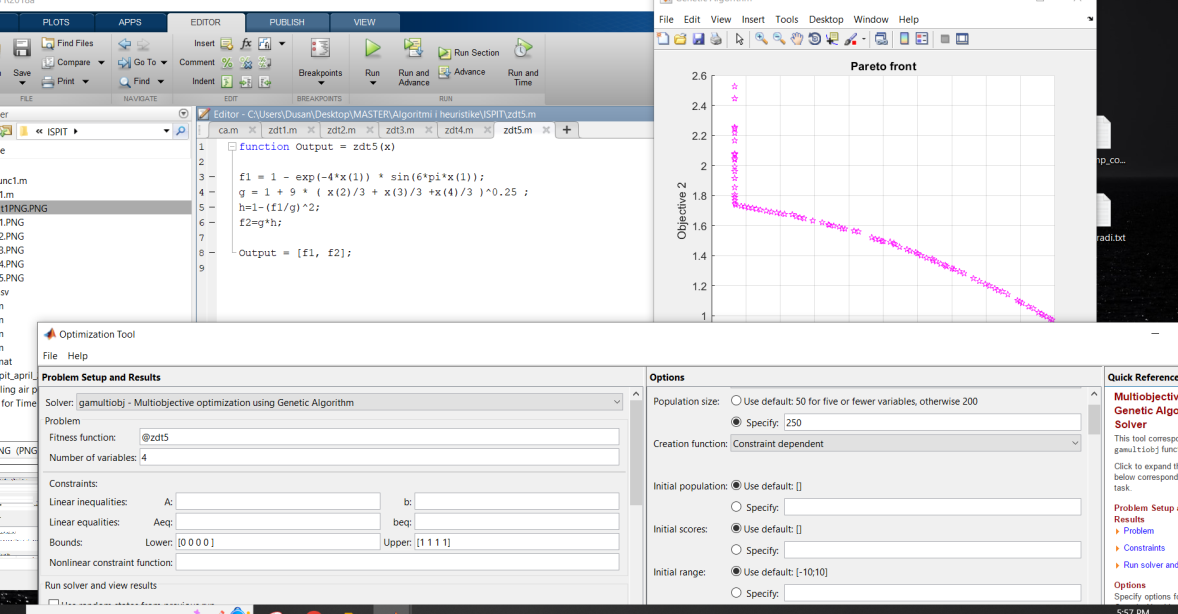
Slika 1: ZDT1

Slika 2: ZDT2



Slika 3: ZDT3

Slika 4: ZDT4



Slika 5: ZDT5

Koraci:

* Definisanje funkcija u MATLAB jeziku
* Otvaranje GUI-ja, konfigurisati ga da rešava *Multiobjective* optimizaciju koristeći genetskog algoritma, postavljanje fitnes funkcije, ograničenja, broja jedinki
* Prikaz Pareto fronta

**Šta je Pareto front?**

Zamislimo da imamo dve (ili više) međusobno zavisne funkcije koje treba da budu najbolje moguće (nekad želimo da maksimizujemo obe, nekad da ih minimizujemo, a nekad jednu da maksimizujemo, drugu da minimizujemo). Postavimo ih u dvodimenzijalnu ravan, tako da smatramo da, što je funkcija bolja, da je dalja koordinatnom početku.

Jedna tačka u ravni označava dvojku (f1,f2) i svojim položajem (koordinatama) pokazuje koliko su funkcije optimizovane, i koji je odnos izmedju njih. Uvek će se morati tražiti kompromis izmedju jedne i druge funkcije, ako zavise jedna od druge.

Tačke koje su bolji po obe funkcije se nazivaju Pareto poboljšanja. Ako pratimo poboljšanje (u Dekartovoj ravni ovo se prevodi kao udaljavanje od te tačke u smeru suprotnom koordinatnom početku), u jednom trenutku dolazimo do tačke od koje ne postoji bolja tačka po obe funkcije. Skup takvih tačaka sačinjavaju Pareto front.

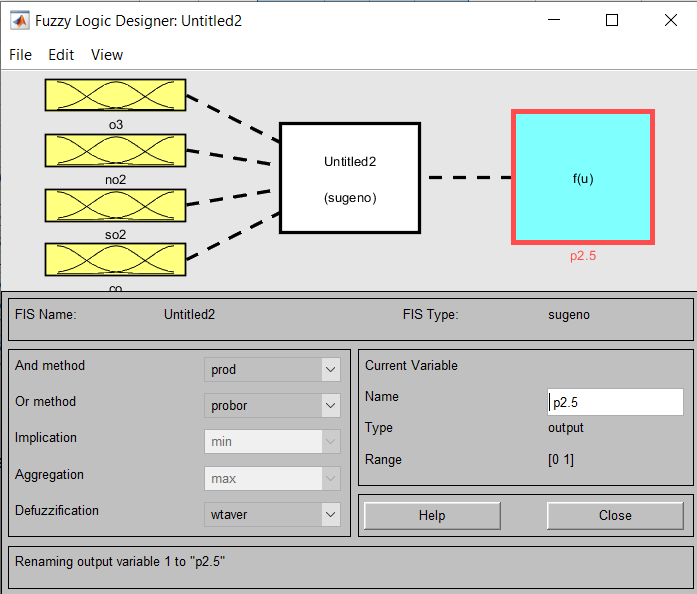
Ovo važi ne samo za dve, nego za proizvoljan broj funkcija koje se optimizuju.

Ukratko: Pareto front je skup tačaka od kojih ne postoji tačka koja je bolja po svim funkcijama koje se optimizuju.

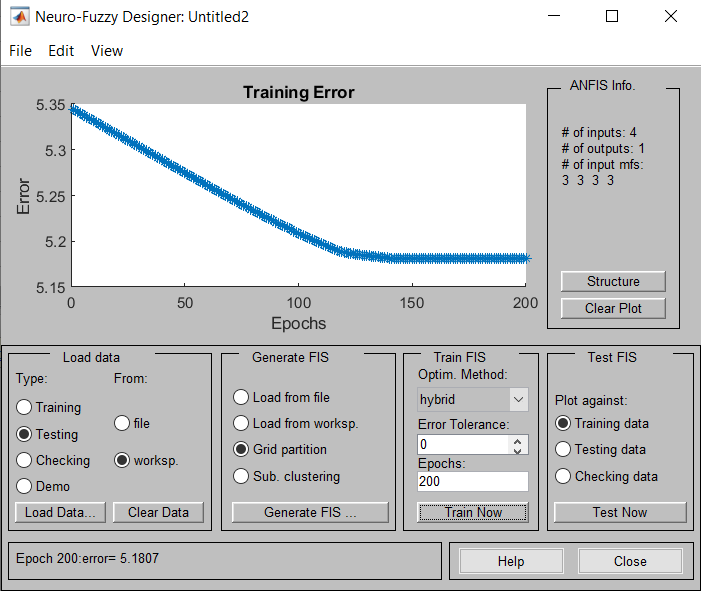
# 2. Prvi zadatak- ANFIS predikcija indeksa kvaliteta vazduha

U MATLAB je uneto 80 podataka za trening, 40 podataka za testiranje. Pod jednim podatkom se smatra uredjena četvorka (O3, NO2, SO2, CO) u kojoj svaki element određuje nivo zagađenosti vazduha datim jedinjenjem. Uzete su vrednosti od 1.2.2022. do 31.5.2022.

Uz pomoć MATLAB alata *fuzzy designer*-a i *neurofuzzy designera*, napravljena je fuzzy logika koja predviđa koncentraciju PM2.5 čestice. Na sledećim slikama su prikazani koraci tokom rada ovog zadatka.

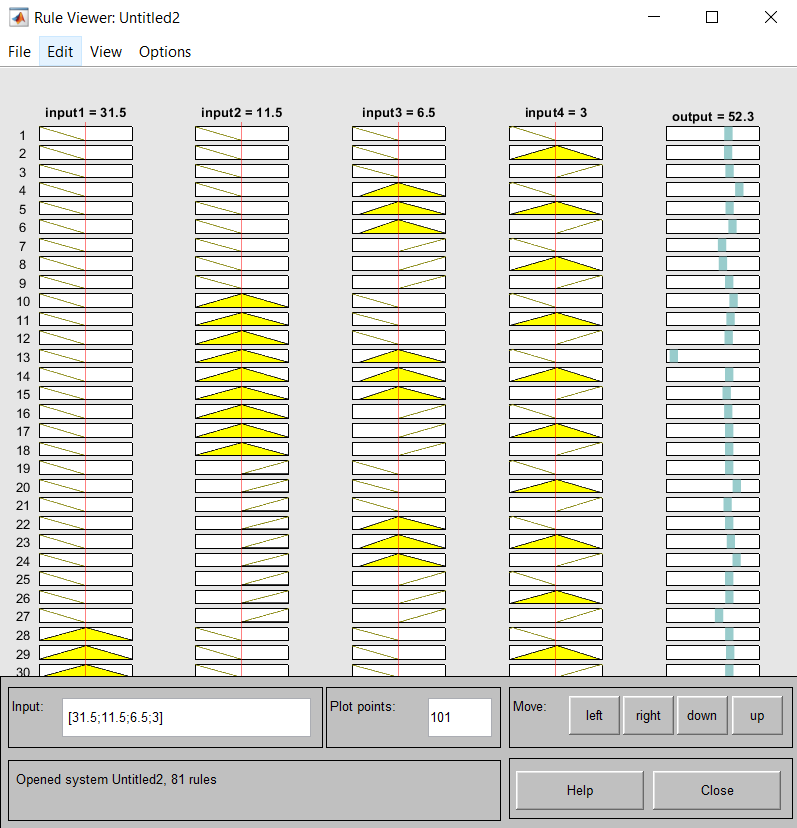


Slika : Definisanje ulaza i izlaza FUZZY logike



Slika : Treniranje FUZZY logike

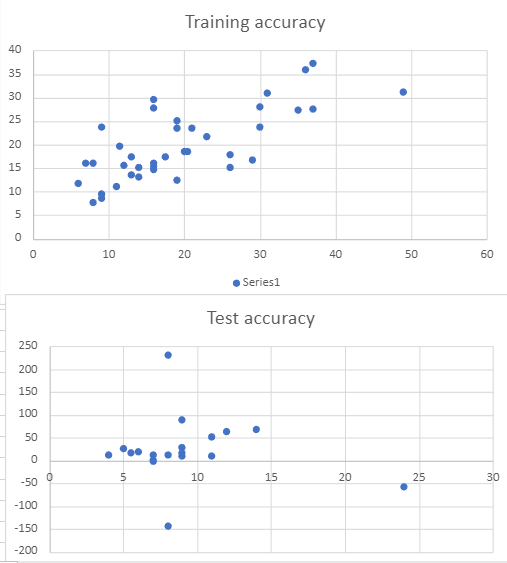
Slika 7 prikazuje pad greške tokom treniranja *fuzzy* logike. Greška je padala, dok na kraju treninga nije bila 5.1 . Ovo je velika greška, ali kao što se vidi na grafiku pred kraj treniranja funkcija greške se zaravnila, što znači da najverovatnije uz više epoha ne bi napredovala (imala značajno manju grešku).



Slika : *Rule viewer*

Iz *rule viewer*-a je ekstrakovana *fuzzy* funkcija, koja je u kombinaciji sa ulazima data izlaz: komanda *newOutput = evalfis(dataOutput,anfis\_rules);* U newOutput se nalazi predikcija koncentracije P2.5 .

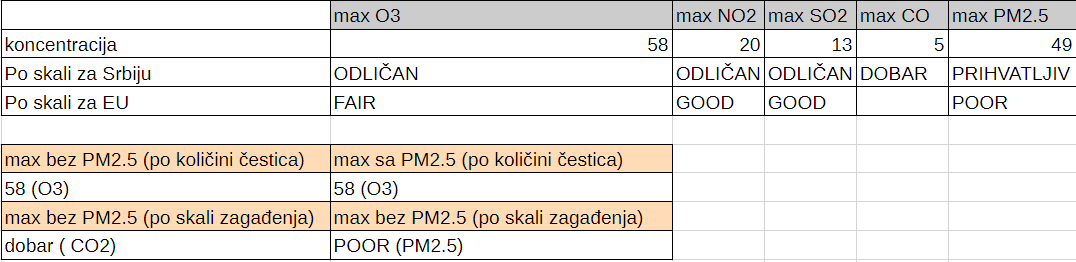
Na sledećim graficija su prikazane tačnosti podataka iz treninga, i testiranja.



Slika : Tačnost treninga i testiranja

Kriva koja pokazuje srednju vrednost greške nije bila dostupna u *online* *Excel*-u ni u *LibreOffice-u*. Ona bi bila jedan od pokazatelja kvaliteta predikcije.

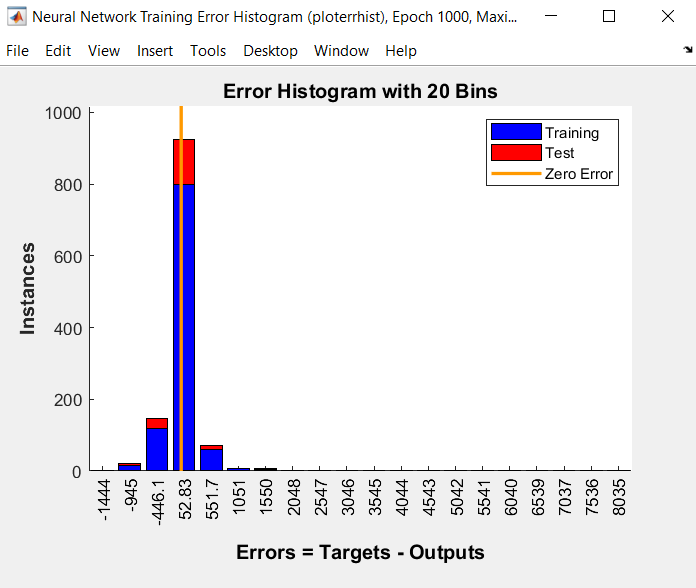
U sledećoj tabeli su predstavljene maksimalne koncentracije svake čestice, kao i ostala izračunavanja (uradjeno uz pomoć *EXCEL*-a).



Najviša koncentracija čestica je O3 (ozon), ali su pragovi zagađenja za njega manji nego za PM2.5. Ako posmatramo skale zagađenja(i srpsku i EU), najveći stepen zagađenja pravi P2.5.

# 3. Treći zadatak- Predikcija vremenskih serija korišćenjem Rekurentnih neuralnih mreša, RNN

Ulazni podaci su bili podaci zaraženih svaki dan od početka pandemije, do 4.4.2023. Pokretanjem *nnstart* , *time series app, NAR.* Izabrani algoritam je bio *Bayesian Regularization.* Razlog korišćenja ovog algoritma je to da je davao najmanje greške na *Error histogram*-u.

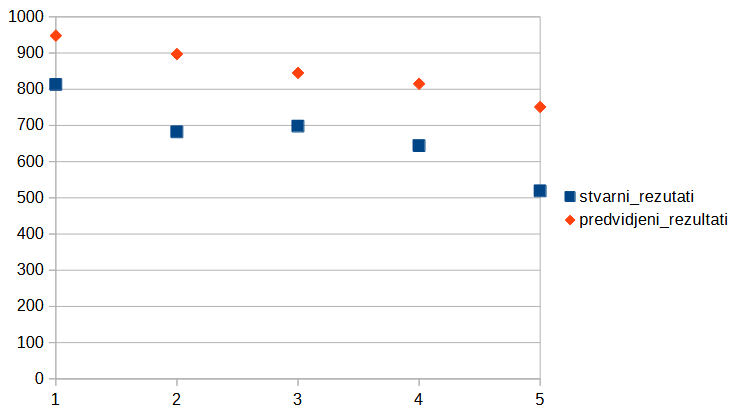


Slika : Error histogram

Na histogramu je prikazano da ni jedna apsolutna greška tokom treniranja ili testa nije bila veća od 1051. A velika većina podataka je nalazi u opsegu apsolutne greške 53.

Grenerisana je *Advanced script* iz *nntool* GUI-ja, i pokrenuta je. Od nje smo dobili predikciju za sledećih 5 dana. Na sledećoj tabeli je prikazan u prvoj vrsti stvaran broj obolelih od petog do devetog aprila ove godine, druga vrsta je predvidjena od strane mreže koju smo istrenirali.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 5.4.2023. | 6.4.2023. | 7.4.2023. | 8.4.2023. | 9.4.2023. |
| Stvaran broj | 813 | 682 | 698 | 644 | 519 |
| Izračunat broj | 948 | 897 | 845 | 815 | 751 |



Slika 11: Grafički prikaz stvarnih i predvidjenih broja obolelih u periodu od 5.4.2023-9.4.2022.