

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**

**INFORMATIKOS FAKULTETAS**

**Laboratorinis darbas Nr. 2**

**Parengė:**

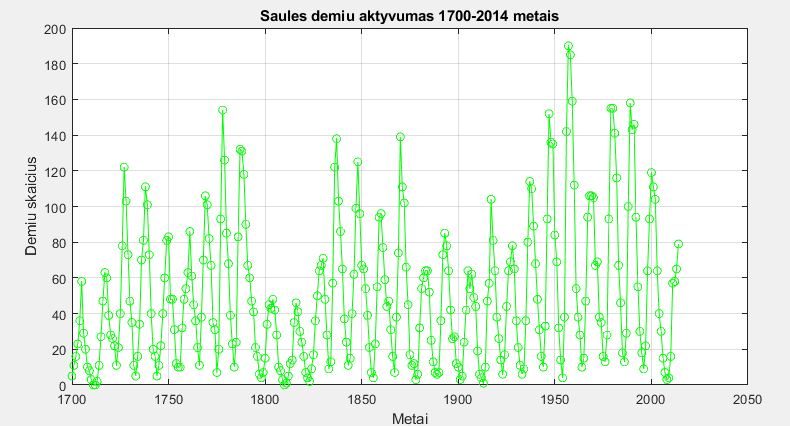
IFF-7/11 Martynas Jodeika

**Darbą priėmė:**

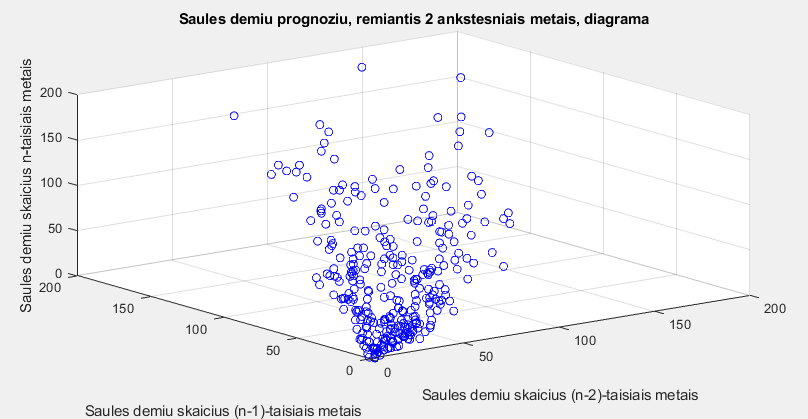
lekt. Budnikas Germanas

**KAUNAS, 2020**

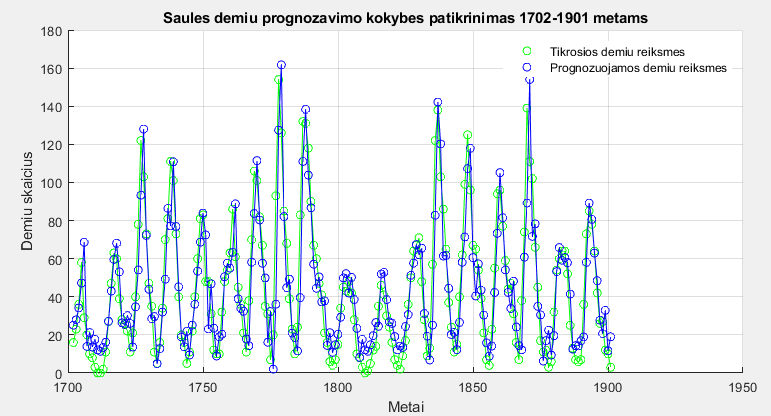
## Pirmos užduoties punktai 1-14

****

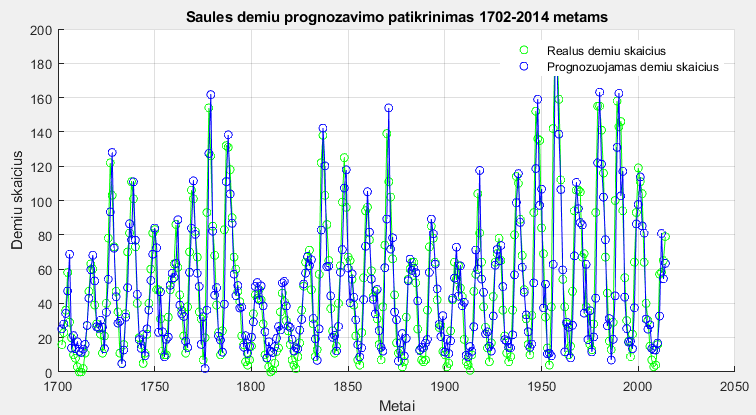
pav. 1 saulės dėmių aktyvumas 1700-2014 metais

****

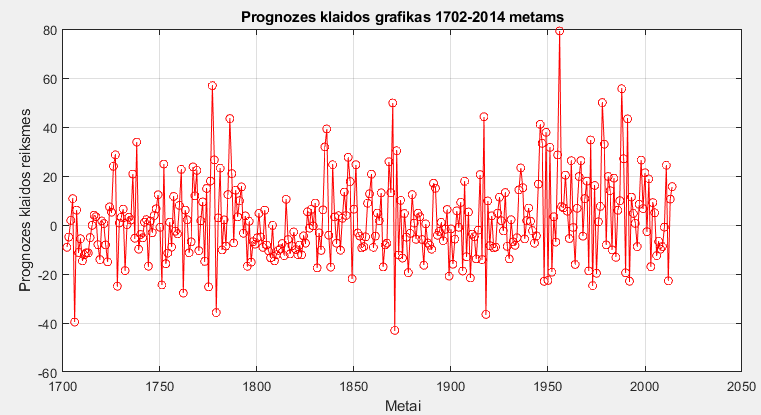
pav. 2 saulės dėmių prognozės gautos naudojantis dviejų ankstesnių metų duomenimis

****

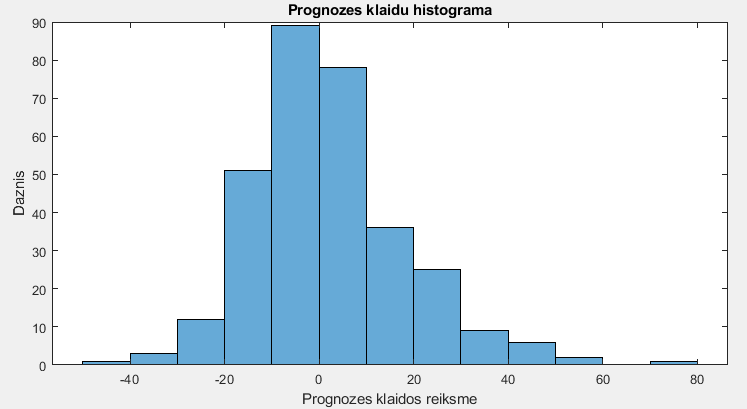
pav. 3 saulės dėmių prognozių koykbės patikrinimas 1702-1901 metais

****

pav. 4 saulės dėmių prognozių patikrinimas 1702-2014 metais

****

pav. 5 prognozės klaidų grafikas 1702-2014 metais

****

pav. 6 prognoziu klaidų histograma

Neuronų svorio koeficientai:

* -0.6761 1.3715

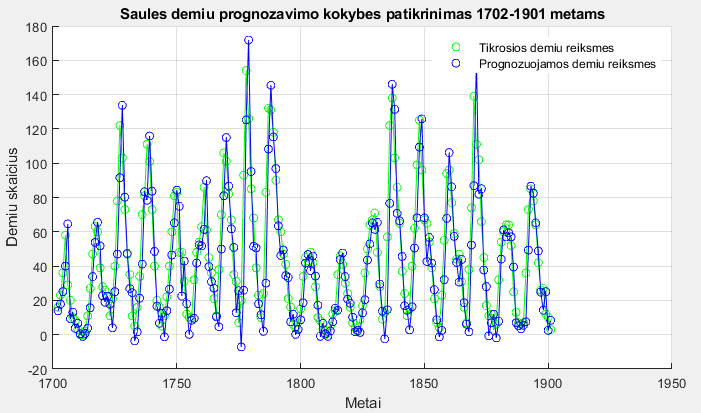
MSE (vidutinės kvadratinės prognozės klaidos reikšmė):

* 278.2687

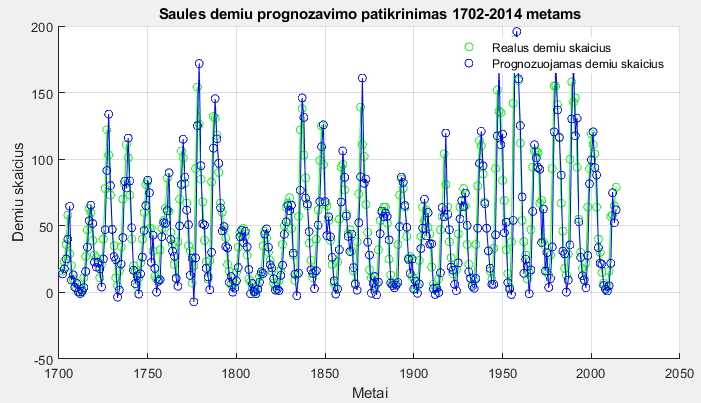
MAD (prognozės absoliutaus nuokrypio medianos reiškmė):

* 12.5616

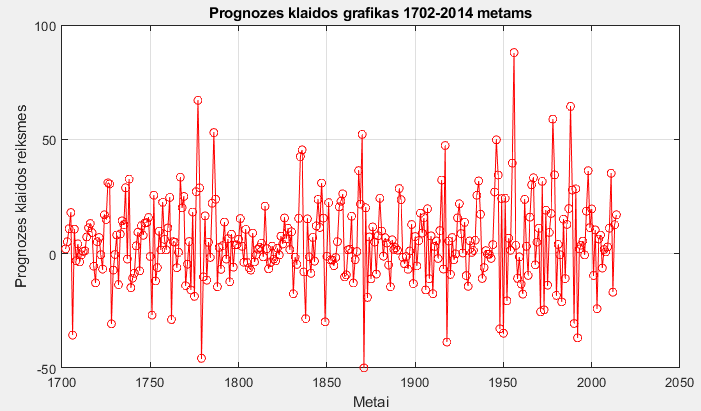
## Pirmos užduoties punktai 15-20



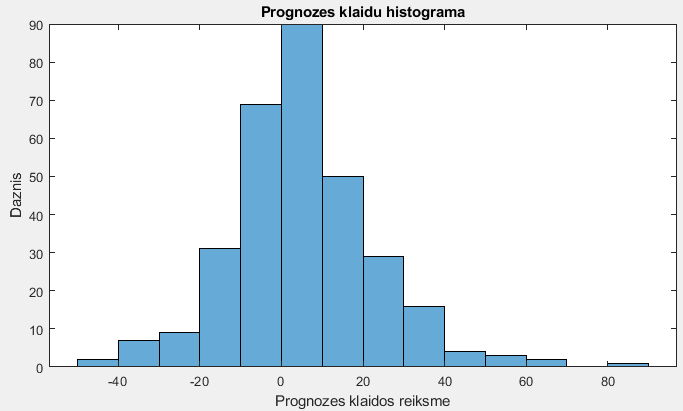
pav. 7 saulės dėmių prognozių koykbės patikrinimas 1702-1901 metams



pav. 8 saulės dėmių prognozių kokybės patikrinimas 1702-2014 metams



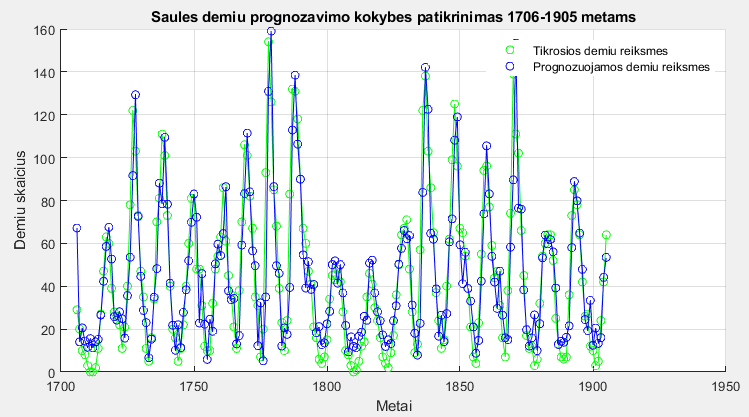
pav. 9 prognozės klaidų grafikas 1702-2014 metams



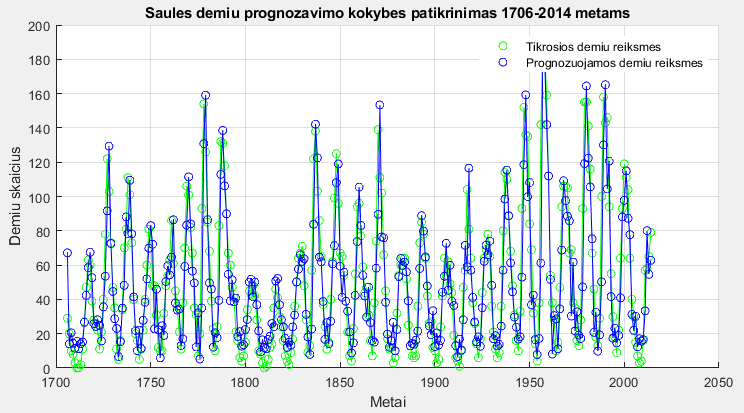
pav. 10 prognozės klaidų histograma

1. MSE (vidutinės kvadratinės prognozės klaidos reikšmė):  
   349.6499
2. MAD (prognozės absoliutaus nuokrypio medianos reiškmė):  
   13.0594
3. Ką pavaizduoja diagrama, kuri vaizduojama mokymosi proceso metu?  
     
   Atvaizduoja neurono treniravimo proceso efektyvumą ir tinklo struktūrą.
4. Ar mokymosi procesas yra konverguojantis? Jeigu ne, pamąstyti kas gali būti priežastimi ir pakeisti atitinkamą parametrą:  
     
   Procesas nėra konvervuogantis. Norint, kad mokymosi procesas konverguotų, reikia pakeisti tinklo atributo goal reikšmę.
5. Kokios yra naujos neurono svorių koeficientų reikšmės?  
     
   -0.5851 1.4638
6. Kokia yra neurono darbo kokybės įverčio – vidutinis kvadratinis nuokrypis – reikšmė?  
     
   349.6499
7. Kokia yra maksimali leistina mokymosi proceso greičio koeficiento lr reikšmė, kuri užtikrina proceso konvergenciją?  
     
   8.2087e-07

## Pirmos užduoties 21 punktas

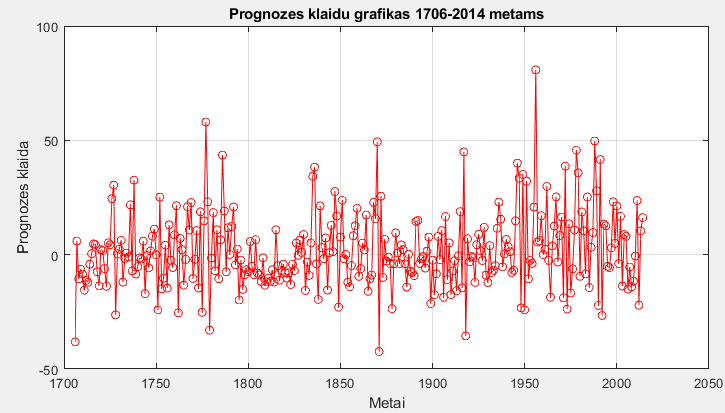


pav. 11 saulės dėmių prognozių koykbės patikrinimas 1706-1901 metais

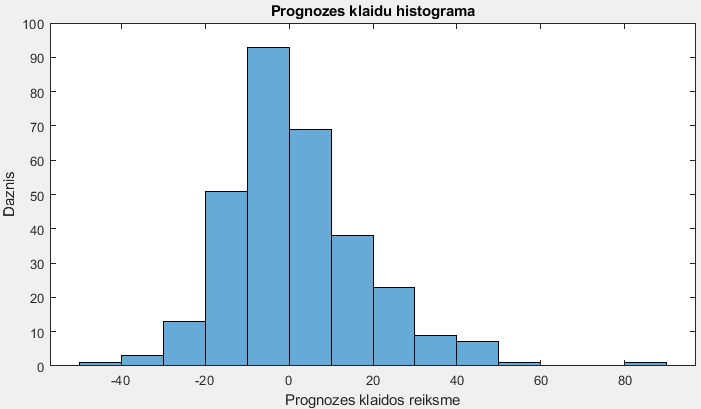


pav. 12 saulės dėmių prognozių koykbės patikrinimas 1706-2014 metais

Iš pirmų dviejų diagramų pastebime, kad prognozuojamos reikšmės yra daug arčiau tikrūjų reikšmių naudojant paskutines šešias vertes prognozei, nei naudojant paskutines dvi. Šis tinklas yra daug tikslesnis.



pav. 13 prognozės klaidų grafikas 1706-2014 metais



pav. 14 prognozės klaidų histograma

Iš klaidų diagramų galime pastebėti, kad dauguma klaidų vertė yra 0 arba arti 0, tai dar patvirtina tinklo tikslumą.

# Laboratorinio darbo antra dalis

## Pradinio duomenų rinkinio aprašymas

Pasirinktas duomenų rinkinys, kuris nusako informacija apie Viduro-Atlanto valstijose gyvenančius vyrus ir jų algą, turintis 3000 duomenų eilučių. Duomenų rinkinį sudaro:

4 tolydinio tipo atributai:

* Metai, kuriais surinkta informacija,
* darbininko amžius,
* atlyginimo logoritmas
* grynas darbuotojo atlyginimas

6 kategorinio tipo atributai:

* ar darbuotojas vedęs
* darbuotojo rasė (baltaodis, juodaodis, azijietis, kita)
* darbuotojo išsilavinimas (nebaigęs mokyklos, baigęs mokyklą ir t.t.)
* regionas (tik Vidurio-Atlanto)
* sveikata (gera arba blogiau, labai gera)
* ar turi sveikatos draudimą

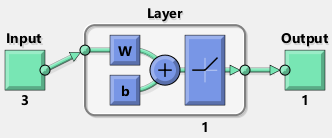
## Duomenų rinkinio pertvarkymo aprašas

Duomenų rinkiniu pertvarkyti buvo visos kategorinės reikšmės konvertuotos į skaitines, išmestos ekstremalios reikšmės ir „wage“ stulpelio sumažintas intervalas (nuo 1 iki 10). Taip pat pasirinkau nenaudoti tų stulpelių, kurie nekoreliavo su mano tikslo atributu (atlyginimas).

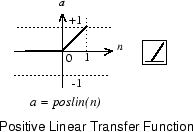
## DNT architektūra, įskaitant parametrų vertes



pav. 15 tinklo parametrų vertės



pav. 16 DNT architektūra



pav. 17 aktyvavimo funkcija poslin

## 10 intervalų kryžminės patikros eksperimentų rezultatai

Po tinklo treniravimo nadojantis 10 intervalų kryžmine patikra gautos tokie MSE ir MAD vidurkiai:

MSE:4.9341

MAD:1.7631

## Priemonių, kurių buvo imtasi siekiant pagerinti ANN veiklą, aprašymas.

Kad pagerinti gautus rezultatus ir sumažinti klaidas pakeičiau tinklo goal parametrą į 0.01 ir aktyvavimo funkciją į „purelin“. Duomenų pertvarkymą buvau atlikęs nuo pat pradžių todėl nieko nekeičiau. Gauti nauji vidurkiai:

MSE:4.6678

MAD:1.7161

## Išvados

Užduoties metu supratau kaip veikia dirbtinis neuroninis tinklas su tikrais duomenimias, kaip jis gali atlikti reikšmių prognozavimą. Antroje užduotyje išmokau panaudoti kryžminę patikrą dirbant su neuronų tinklu ir kaip tinklo parametrai daro įtaką prognozuojamoms reikšmėms.

# Realizacijos kodas:

Antra dalis:

close all;

clear all;

clc

data = csvread('wage.csv');

inputData = [data(1:600,1:3),data(1:600,4)];

j = 10;

mseSum = 0;

madSum = 0;

k1 = [inputData(1:30,:); inputData(300:330,:)];

k2 = [inputData(31:60,:); inputData(331:360,:)];

k3 = [inputData(61:90,:); inputData(361:390,:)];

k4 = [inputData(91:120,:); inputData(391:420,:)];

k5 = [inputData(121:150,:); inputData(421:450,:)];

k6 = [inputData(151:180,:); inputData(451:480,:)];

k7 = [inputData(181:210,:); inputData(481:510,:)];

k8 = [inputData(211:240,:); inputData(511:540,:)];

k9 = [inputData(241:270,:); inputData(541:570,:)];

k10 = [inputData(271:300,:); inputData(571:600,:)];

for i=1:j

if i == 1

trainMatrix = [k1;k2;k3;k4;k5;k6;k7;k8;k9];

testMatrix = k10;

end

if i == 2

trainMatrix = [k1;k2;k3;k4;k5;k6;k7;k8;k10];

testMatrix = k9;

end

if i == 3

trainMatrix = [k1;k2;k3;k4;k5;k6;k7;k9;k10];

testMatrix = k8;

end

if i == 4

trainMatrix = [k1;k2;k3;k4;k5;k6;k8;k9;k10];

testMatrix = k7;

end

if i == 5

trainMatrix = [k1;k2;k3;k4;k5;k7;k8;k9;k10];

testMatrix = k6;

end

if i == 6

trainMatrix = [k1;k2;k3;k4;k6;k7;k8;k9;k10];

testMatrix = k5;

end

if i == 7

trainMatrix = [k1;k2;k3;k5;k6;k7;k8;k9;k10];

testMatrix = k4;

end

if i == 8

trainMatrix = [k1;k2;k4;k5;k6;k7;k8;k9;k10];

testMatrix = k3;

end

if i == 9

trainMatrix = [k1;k3;k4;k5;k6;k7;k8;k9;k10];

testMatrix = k2;

end

if i == 10

trainMatrix = [k2;k3;k4;k5;k6;k7;k8;k9;k10];

testMatrix = k1;

end

inputTrain = trainMatrix(:,1:3);

disp('train matrix:')

disp(trainMatrix)

target = trainMatrix(:,4);

inputTrain = inputTrain';

outputTrain = target';

disp('inputtrain matrix:')

disp(inputTrain)

disp('outputtrain matrix:')

lr = maxlinlr(inputTrain);

net = newlin(inputTrain,outputTrain,0,lr);

disp('Neurono bias reiksme:');

disp(net.b{1});

disp('Neurono svorio koeficientai:');

disp(net.IW{1});

net.trainParam.goal = 0.01;

net.trainParam.epochs = 1000;

net.layers{1}.transferFcn = 'purelin';

net = train(net, inputTrain, outputTrain);

disp('Neurono bias reiksme:');

disp(net.b{1});

disp('Neurono svorio koeficientai:');

disp(net.IW{1});

inputTest = testMatrix(:,1:3);

targetTest = testMatrix(:,4);

inputTest = inputTest';

outputTest = targetTest';

result = sim(net, inputTest);

e = outputTest - result;

mseSum = mseSum + mse(e)

madSum = madSum + mad(e)

view(net)

disp(result)

disp(outputTest)

end

avgTest=avgTest/10;

mseAvg = mseSum / 10;

madAvg = madSum / 10;

disp(['Mean-Square-Error, MSE:', num2str(mseAvg)]);

disp(['Median absolute Deviation, MAD:', num2str(madAvg)]);

disp(['network hit average (kfold = 10):', num2str(avgTest)]);

clc, clear, close all;

load sunspot.txt;

%4:

figure(1);

plot(sunspot(:, 1), sunspot(:, 2), 'g-o');

grid on;

xlabel('Metai');

ylabel('Demiu skaicius');

title('Saules demiu aktyvumas 1700-2014 metais');

%5:

L = length(sunspot);

P = [sunspot(1:L-2, 2)'; sunspot(2:L-1, 2)'];

disp('P matrica:')

disp(P)

disp('P matricos dydis:')

size(P)

T = sunspot(3:L, 2)';

disp('T matrica:')

disp(T)

disp('T matricos dydis:')

size(T)

%6:

figure(2)

plot3(P(1,:), P(2,:), T, 'bo');

grid on;

xlabel('Saules demiu skaicius (n-2)-taisiais metais');

ylabel('Saules demiu skaicius (n-1)-taisiais metais');

zlabel('Saules demiu skaicius n-taisiais metais');

title('Saules demiu prognoziu, remiantis 2 ankstesniais metais, diagrama');

%7:

disp('Pu matrica:')

Pu = P(:, 1:200)

disp('Tu matrica:')

Tu = T(:, 1:200)

disp('Pu dydis:')

size(Pu)

disp('Tu dydis:')

size(Tu)

%8:

net = newlind(Pu, Tu);

%9:

disp('Neurono svorio koeficientai:');

disp(net.IW{1});

disp('Neurono bias reiksme:');

disp(net.b{1});

w1 = net.IW{1}(1);

w2 = net.IW{1}(2);

b = net.b{1};

%10:

Tsu = sim(net, Pu)

figure(3), hold on;

plot(sunspot(3:202, 1), Tu, 'g-o');

plot(sunspot(3:202, 1), Tsu, 'b-o');

xlabel('Metai');

ylabel('Demiu skaicius');

grid on;

legend('Tikrosios demiu reiksmes', ...

'Prognozuojamos demiu reiksmes');

title('Saules demiu prognozavimo kokybes patikrinimas 1702-1901 metams');

%11:

Ts = sim(net, P);

figure(4), hold on;

plot(sunspot(3:315, 1), T, 'g-o');

plot(sunspot(3:315, 1), Ts, 'b-o');

xlabel('Metai');

ylabel('Demiu skaicius');

grid on;

legend('Realus demiu skaicius', ...

'Prognozuojamas demiu skaicius');

title('Saules demiu prognozavimo patikrinimas 1702-2014 metams');

%12:

e = T - Ts

figure(5);

plot(sunspot(3:315), e, 'r-o');

grid on;

title('Prognozes klaidos grafikas 1702-2014 metams');

xlabel('Metai');

ylabel('Prognozes klaidos reiksmes');

%13:

figure(6);

histogram(e);

title('Prognozes klaidu histograma');

xlabel('Prognozes klaidos reiksme');

ylabel('Daznis');

%14:

mse\_reiksme = mse(e)

mad\_reiksme = mad(e)

clc, clear, close all;

load sunspot.txt;

figure(1);

plot(sunspot(:, 1), sunspot(:, 2), 'b-o');

grid on;

xlabel('Metai');

ylabel('Demiu skaicius');

title('Saules demiu aktyvumas 1700-1950 metais');

L = length(sunspot);

disp('P matrica (ávesties duomenys):')

P = [sunspot(1:L-2, 2)'; sunspot(2:L-1, 2)']

disp('P matricos dydis:')

size(P)

disp('T matrica (iðvesties duomenys):')

T = sunspot(3:L, 2)'

disp('T matricos dydis:')

size(T)

figure(2)

plot3(P(1,:), P(2,:), T, 'bo');

grid on;

xlabel('Saules demiu skaicius (n-2)-aisiais metais');

ylabel('Saules demiu skaicius (n-1)-taisiais metais');

zlabel('Saules demiu skaicius n-aisiais metais');

title('Saules demiu prognoziu, remiantis 2 ankstesniais metais, diagrama');

disp('Pu matrica (apmokymo duomenys):')

Pu = P(:, 1:200)

disp('Tu matrica (apmokymo rezultatai):')

Tu = T(:, 1:200)

disp('Pu matricos dydis:')

size(Pu)

disp('Tu matricos dydis:')

size(Tu)

fprintf('Maksimali leistina mokymosi proceso greièio koeficiento lr reikðmë: %d\n', maxlinlr(Pu));

net = newlin(Pu, Tu, 0, maxlinlr(Pu));

net.trainParam.goal = 200;

net.trainParam.epochs = 1000;

net = train(net, Pu, Tu);

disp('Neurono svorio koeficientai:');

disp(net.IW{1});

disp('Neurono bias reikðmë:');

disp(net.b{1});

w1 = net.IW{1}(1)

w2 = net.IW{1}(2)

b = net.b{1}

Tsu = net(Pu)

figure(3), hold on;

plot(sunspot(3:202, 1), Tu, 'r-o');

plot(sunspot(3:202, 1), Tsu, 'b-o');

xlabel('Metai');

ylabel('Demiu skaicius');

grid on;

legend('Tikrosios demiu reiksmes', ...

'Prognozuojamos demiu reiksmes');

title('Saules demiu prognozavimo kokybes patikrinimas 1702-1901 metams');

Ts = net(P);

figure(4), hold on;

plot(sunspot(3:315, 1), T, 'r-o');

plot(sunspot(3:315, 1), Ts, 'b-o');

xlabel('Metai');

ylabel('Demiu skaicius');

grid on;

legend('Tikrosios demiu reiksmes', ...

'Prognozuojamos demiu reiksmes');

title('Saules demiu prognozavimo kokybes patikrinimas 1702-2014 metams');

disp('Prognozës klaidø vektorius')

e = T - Ts

figure(5);

plot(sunspot(3:315), e, 'r-o');

grid on;

title('Prognozes klaidos grafikas 1702-2014 metams');

xlabel('Metai');

ylabel('Demiu skaiciaus skirtumas tarp tikrosios ir prognozuojamos reiksmiu');

figure(6);

hist(e);

title('Prognozes klaidu histograma');

xlabel('Prognozes klaidos reiksme');

ylabel('Daznis');

disp('Vidutinës kvadratinës prognozës klaidos reikðmë')

mse\_reiksme = mse(e)

clc, clear, close all;

load sunspot.txt;

n = 6

figure(1);

plot(sunspot(:, 1), sunspot(:, 2), 'b-o');

grid on;

xlabel('Metai');

ylabel('Demiu skaicius');

title('Saules demiu aktyvumas 1700-1950 metais');

L = length(sunspot);

P = [];

for i = 1:n

P = [P; sunspot(i:L-(n+1)+i, 2)'];

end

disp('P matrica (ávesties duomenys):'), P

disp('P matricos dydis:')

size(P)

disp('T matrica (iðvesties duomenys):')

T = sunspot((n+1):L, 2)'

disp('T matricos dydis:')

size(T)

disp('Pu matrica (apmokymo duomenys):')

Pu = P(:, 1:200)

disp('Tu matrica (apmokymo rezultatai):')

Tu = T(:, 1:200)

disp('Pu matricos dydis:')

size(Pu)

disp('Tu matricos dydis:')

size(Tu)

net = newlind(Pu, Tu);

disp('Neurono svorio koeficientai:');

disp(net.IW{1});

disp('Neurono bias reikðmë:');

disp(net.b{1});

sunspot(n+1)

sunspot(200+n)

Tsu = net(Pu)

figure(3), hold on;

plot(sunspot((n+1):200+n, 1), Tu, 'g-o');

plot(sunspot((n+1):200+n, 1), Tsu, 'b-o');

xlabel('Metai');

ylabel('Demiu skaicius');

grid on;

legend('Tikrosios demiu reiksmes', ...

'Prognozuojamos demiu reiksmes');

title(sprintf('Saules demiu prognozavimo kokybes patikrinimas %d-%d metams', 1700+n, 1900+n-1));

Ts = net(P);

figure(4), hold on;

plot(sunspot(n+1:315, 1), T, 'g-o');

plot(sunspot(n+1:315, 1), Ts, 'b-o');

xlabel('Metai');

ylabel('Demiu skaicius');

grid on;

legend('Tikrosios demiu reiksmes', ...

'Prognozuojamos demiu reiksmes');

title(sprintf('Saules demiu prognozavimo kokybes patikrinimas %d-2014 metams', 1700+n));

disp('Prognozës klaidø vektorius')

e = T - Ts

figure(5);

plot(sunspot(n+1:315), e, 'r-o');

grid on;

title(sprintf('Prognozes klaidu grafikas %d-2014 metams', 1700+n));

xlabel('Metai');

ylabel('Prognozes klaida');

figure(6);

histogram(e);

title('Prognozes klaidu histograma');

xlabel('Prognozes klaidos reiksme');

ylabel('Daznis');

disp('Vidutinës kvadratinës prognozës klaidos reikðmë')

mse\_reiksme = mse(e)