ML Based Shear Modulus Prediction

Introduction

1)Artificial Neural Networks (ANN)

CODE

Live Editor - C:\Users\Dell\Downloads\ShearModulus.mlx *

ShearModulus.mlx *

```
data=readmatrix('C:/Users/Dell/Downloads/Shear_Modulus.csv');
x=data(:,1:6);
y=data(:,7);
m=length(y);
```

Visualization of the data

```
histogram(y,10);
```

Normalize the features

```
y2=log(1+y);
for i=1:6
    x2(:,i)=(x(:,i)-min(x(:,i)))/max(x(:,i))-min(x(:,i));
end
histogram(y2,10);
plot(x2,y2,'o')
```

Training an ANN model

```
xt=x2';
yt=y2';
hiddenLayerSize=6;
net=fitnet(hiddenLayerSize);
net.divideParam.trainRatio=70/100;
net.divideParam.valRatio=30/100;
net.divideParam.testRatio=0/100;
[net,tr]=train(net,xt,yt);
```

Performance of the ANN network

```
yTrain=exp(net(xt(:,tr.trainInd)))-1;
yTrainTrue=exp(yt(tr.trainInd))-1;
sqrt(mean((yTrain-yTrainTrue).^2))
plot(yTrainTrue,yTrain,'x')
```

Activation Functions

Hidden Layer : tansig

Output Layer : purelin

Command Window

```
>> net.layers{1}

ans =

    Neural Network Layer

              name: 'Hidden'
        dimensions: 6
       distanceFcn: (none)
     distanceParam: (none)
         distances: []
           initFcn: 'initnw'
       netInputFcn: 'netsum'
     netInputParam: (none)
         positions: []
             range: [6x2 double]
              size: 6
       topologyFcn: (none)
       transferFcn: 'tansig'
     transferParam: (none)
          userdata: (your custom info)
```

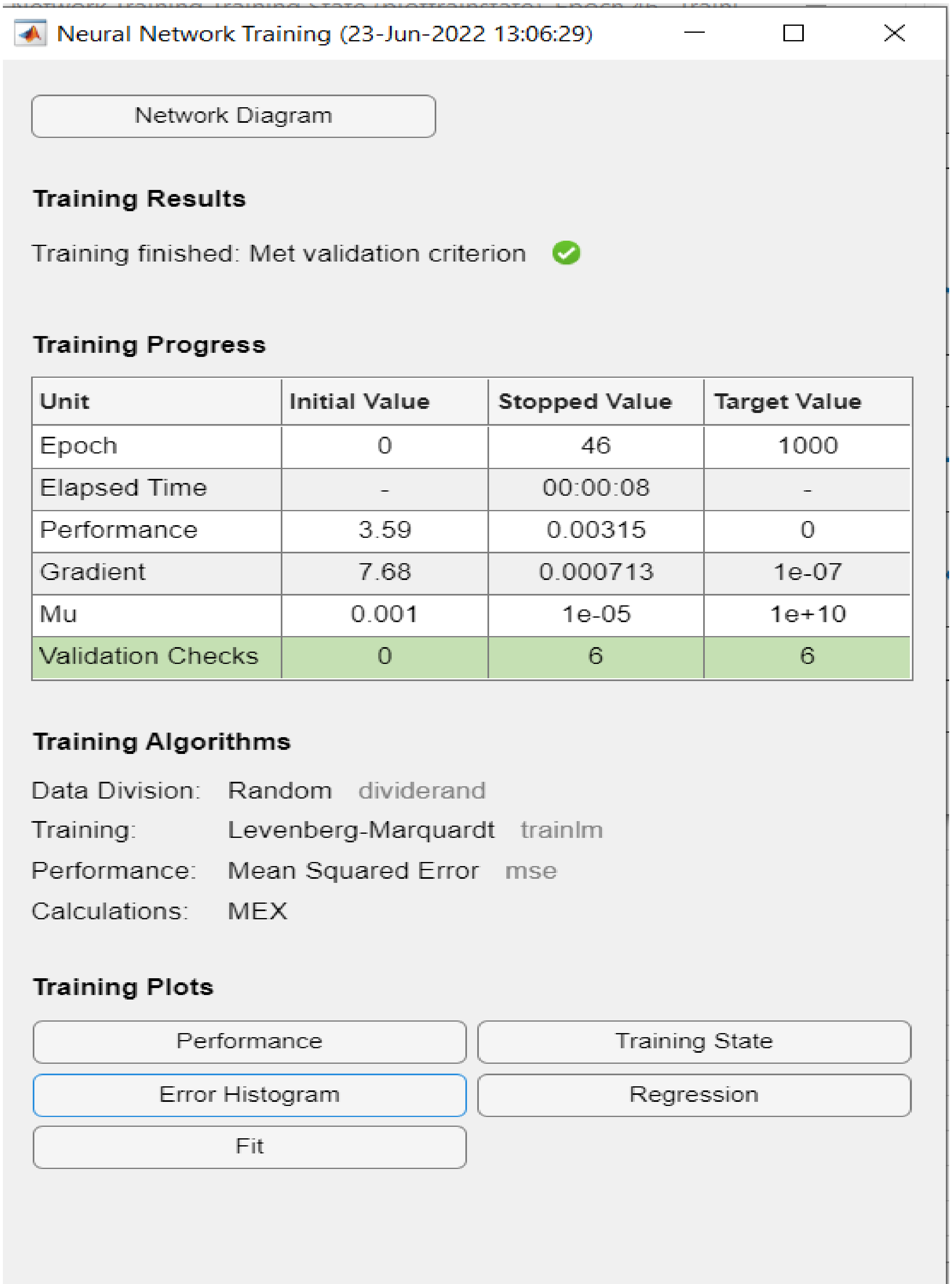
```
>> net.layers{2}

ans =

    Neural Network Layer

              name: 'Output'
        dimensions: 1
       distanceFcn: (none)
     distanceParam: (none)
         distances: []
           initFcn: 'initnw'
       netInputFcn: 'netsum'
     netInputParam: (none)
         positions: []
             range: [1x2 double]
              size: 1
       topologyFcn: (none)
       transferFcn: 'purelin'
     transferParam: (none)
          userdata: (your custom info)

fx >>
```



Neural Network Training (23-Jun-2022 13:06:29)

Network Diagram

Training Results

Training finished: Met validation criterion

Training Progress

Unit	Initial Value	Stopped Value	Target Value
Epoch	0	46	1000
Elapsed Time	-	00:00:08	-
Performance	3.59	0.00315	0
Gradient	7.68	0.000713	1e-07
Mu	0.001	1e-05	1e+10
Validation Checks	0	6	6

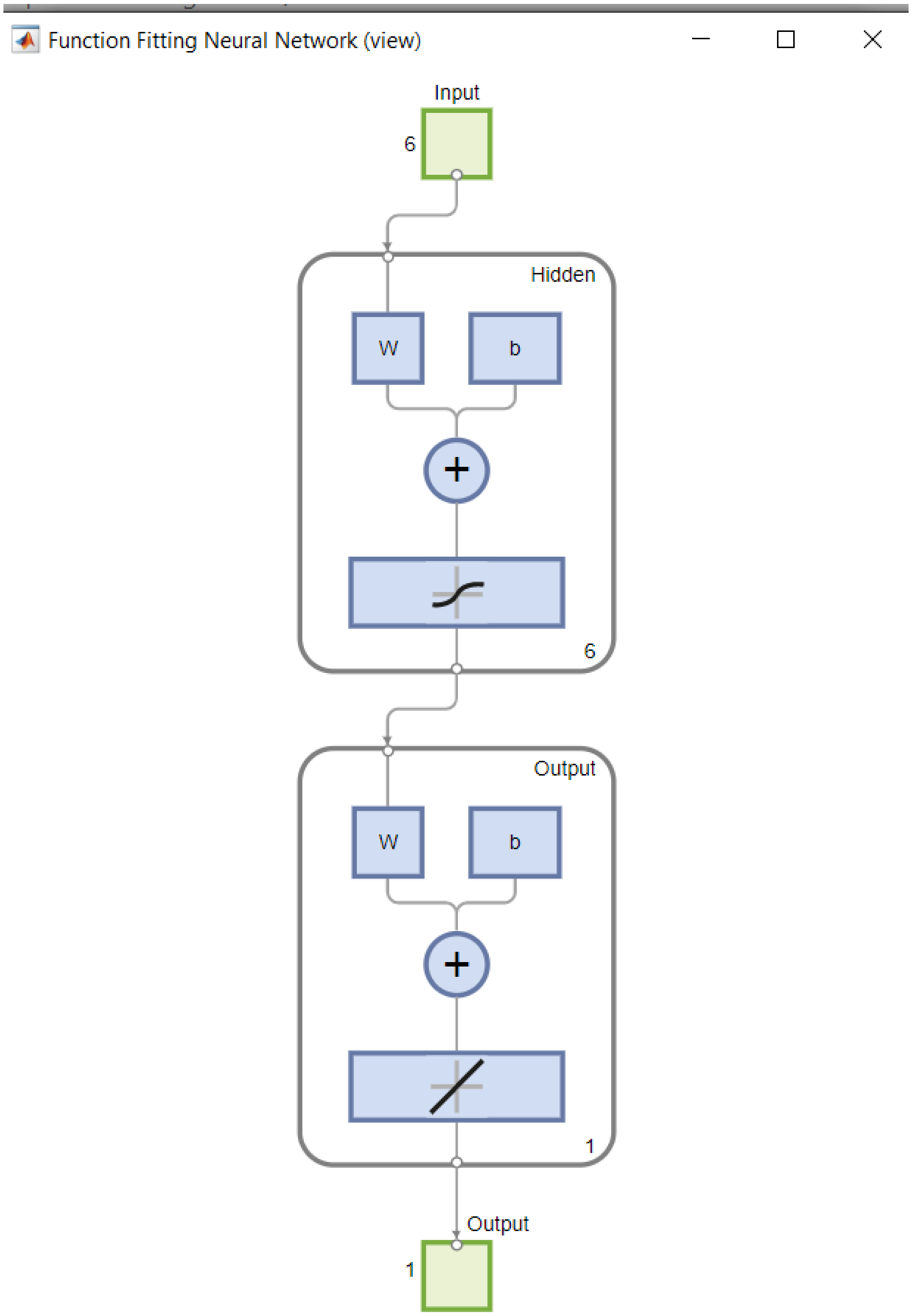
Training Algorithms

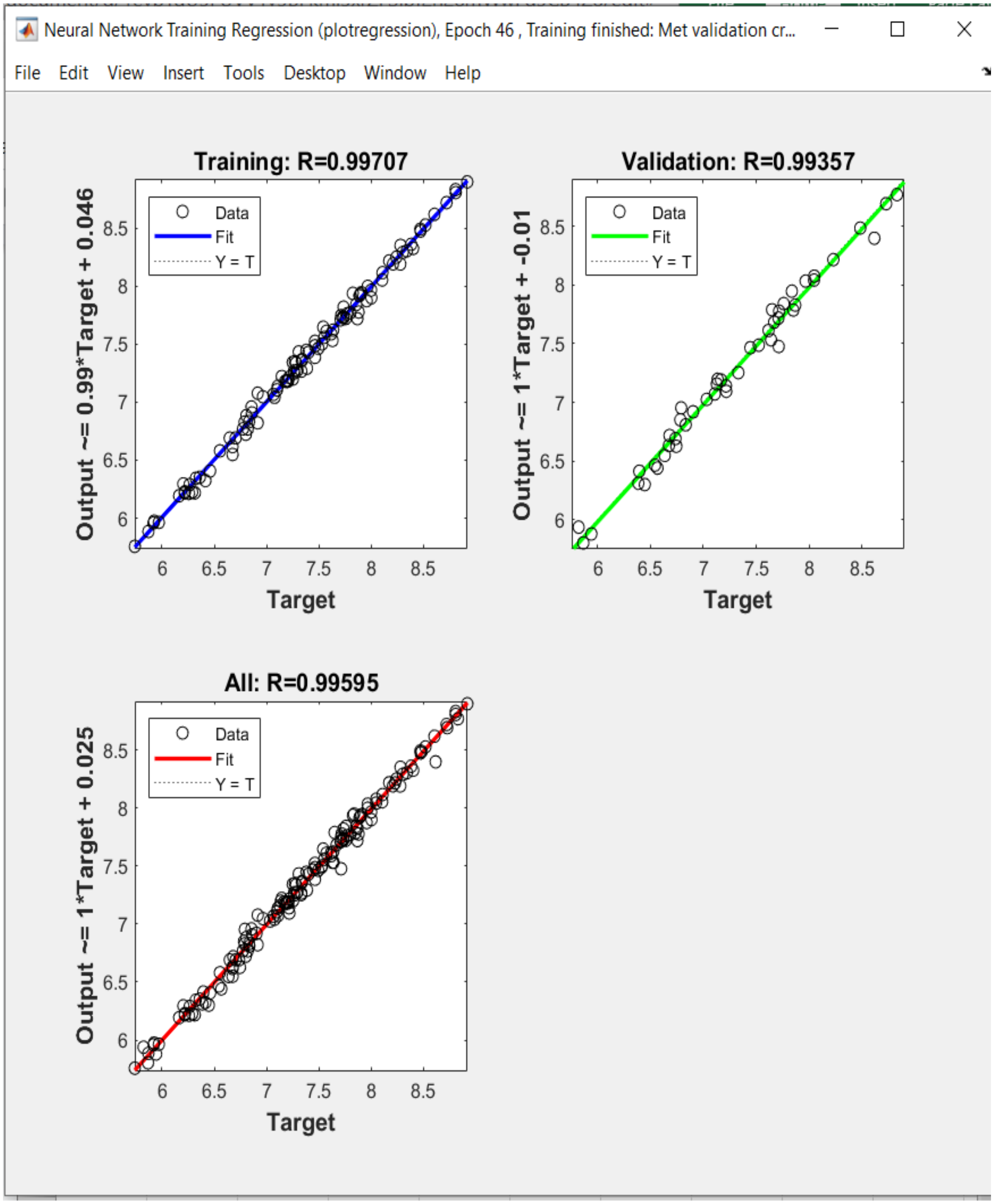
Data Division: Random dividerand
Training: Levenberg-Marquardt trainlm
Performance: Mean Squared Error mse
Calculations: MEX

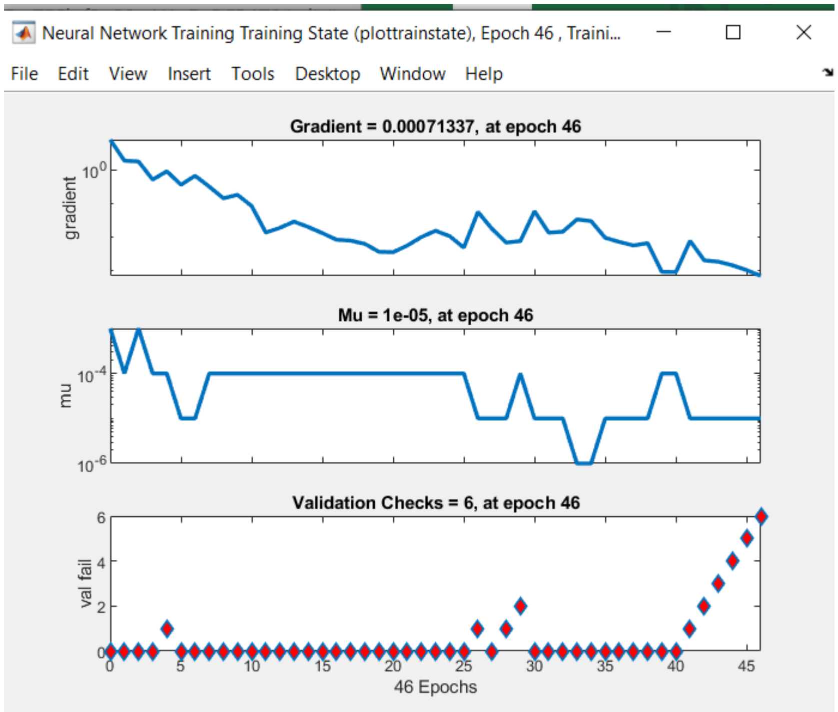
Training Plots

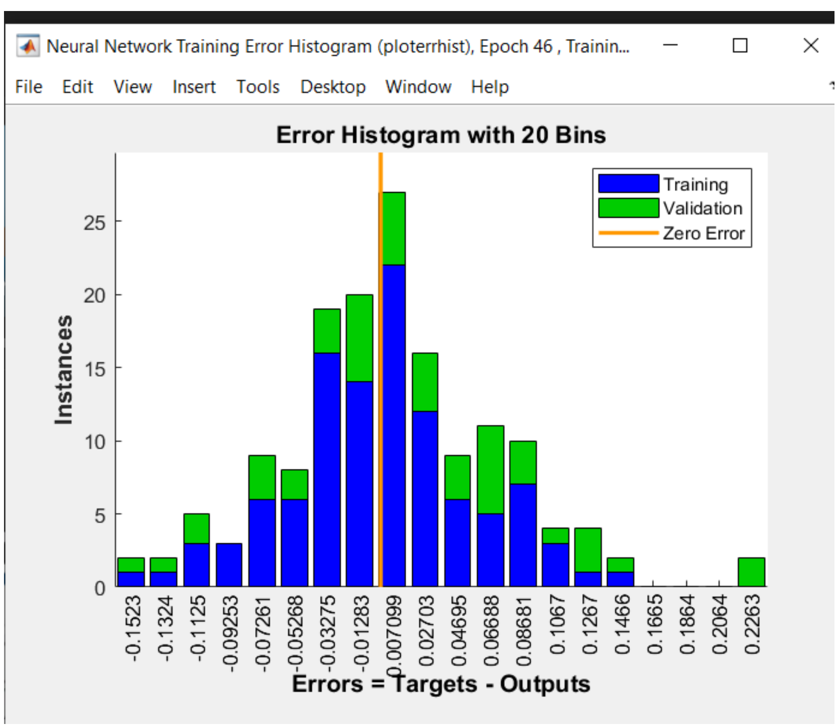
Performance | Training State
Error Histogram | Regression
Fit

Architecture

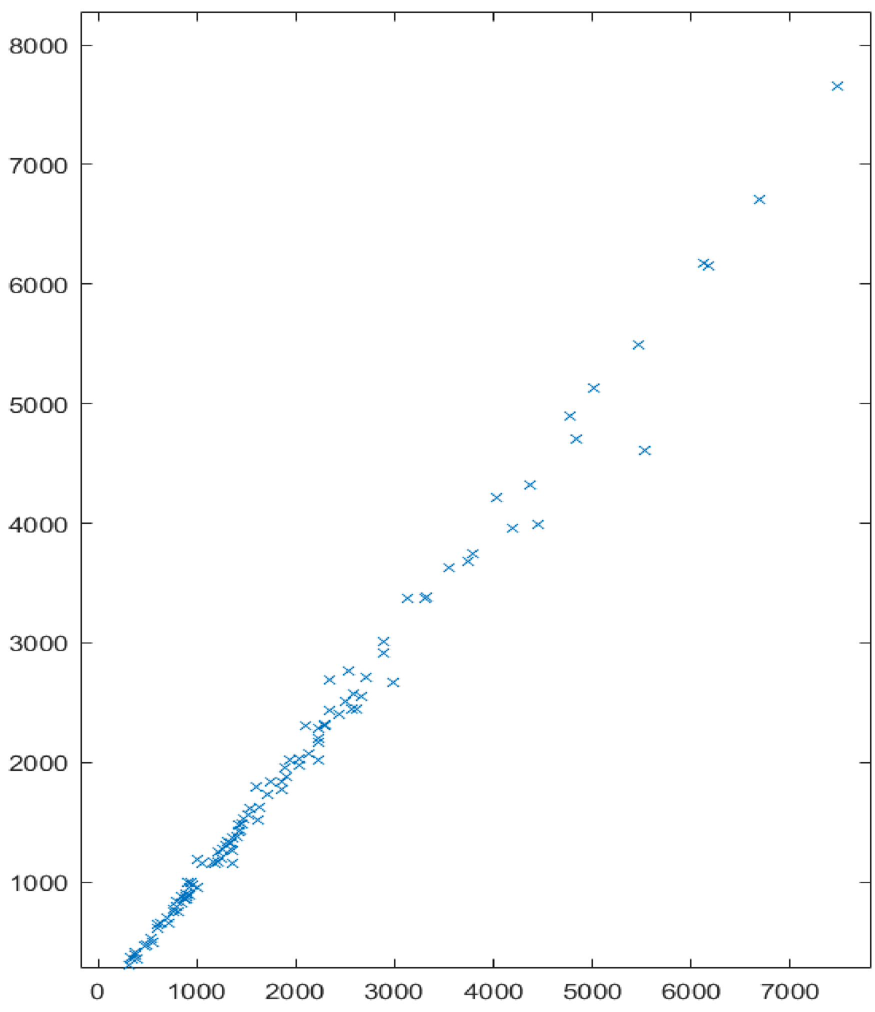



Neural Network Training Regression (plotregression), Epoch 46 , Training finished: Met validation cr...
File Edit View Insert Tools Desktop Window Help
Training: R=0.99707
Output ~= 0.99*Target + 0.046
Data
Fit
Y = T
Target
Validation: R=0.99357
Output ~= 1*Target + -0.01
Data
Fit
Y = T
Target
All: R=0.99595
Output ~= 1*Target + 0.025
Data
Fit
Y = T
Target


Neural Network Training Training State (plottrainstate), Epoch 46 , Traini...
File Edit View Insert Tools Desktop Window Help
Gradient = 0.00071337, at epoch 46
gradient
Mu = 1e-05, at epoch 46
mu
Validation Checks = 6, at epoch 46
val fail
46 Epochs


Neural Network Training Error Histogram (ploterrhist), Epoch 46 , Trainin...
File Edit View Insert Tools Desktop Window Help
Error Histogram with 20 Bins
Training
Validation
Zero Error
Instances
-0.1523
-0.1324
-0.1125
-0.09253
-0.07261
-0.05268
-0.03275
-0.01283
0.007099
0.02703
0.04695
0.06688
0.08681
0.1067
0.1267
0.1466
0.1665
0.1864
0.2064
0.2263
Errors = Targets - Outputs

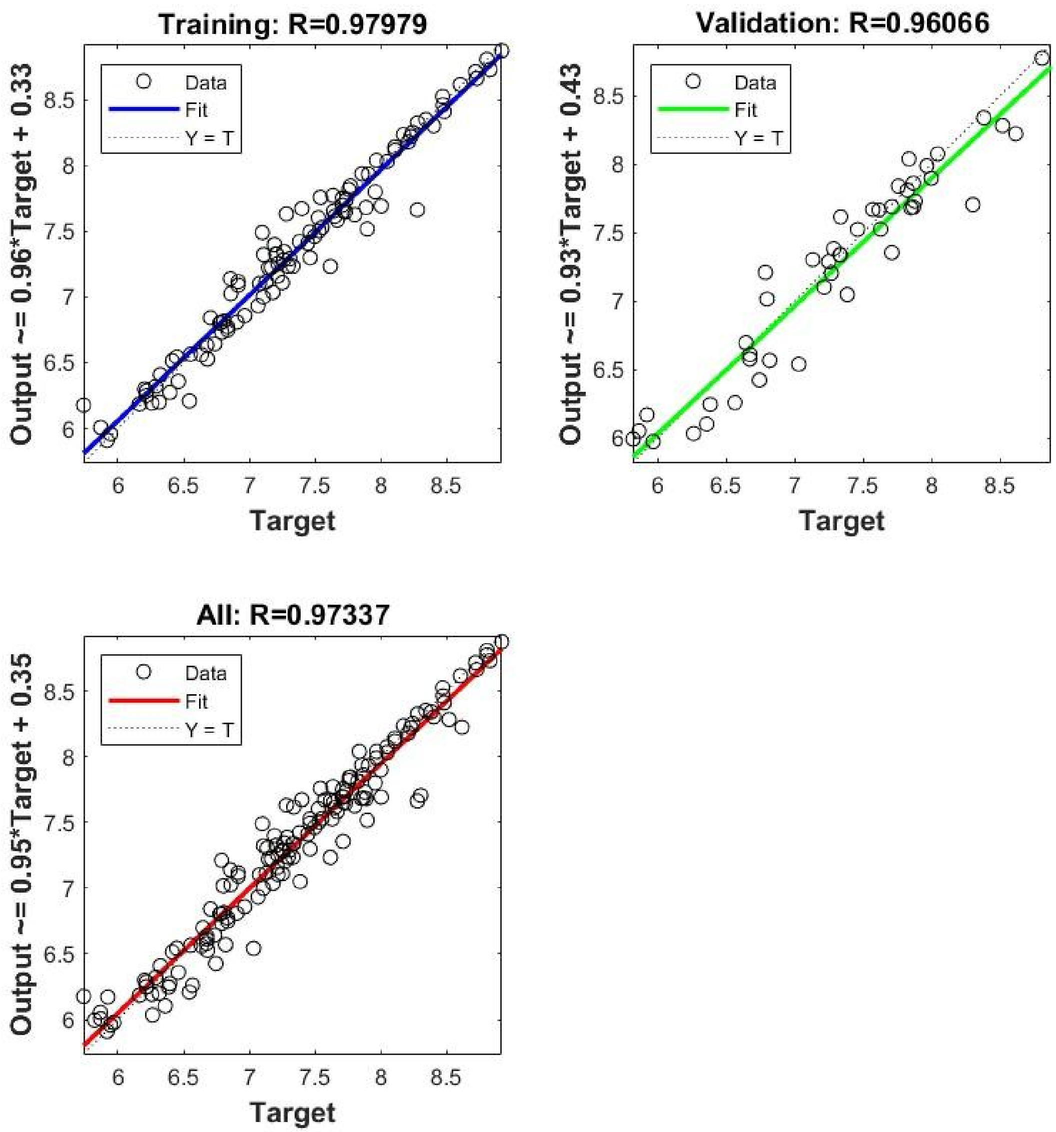
Predicted Output Vs Actual Output



Activation Functions

Hidden Layer: Logsig

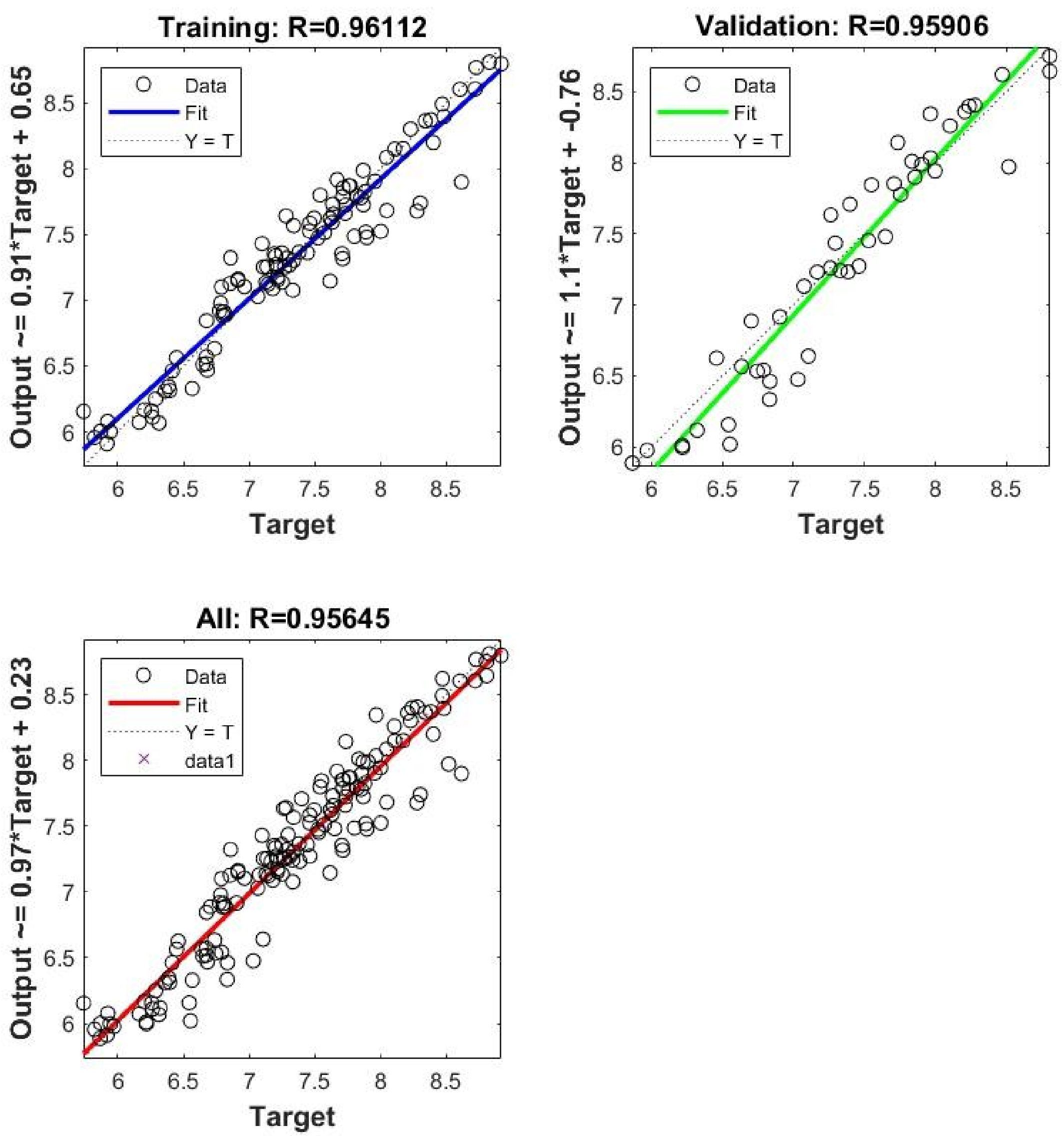
Output Layer: Purelin



Activation Functions

Hidden Layer: radbas

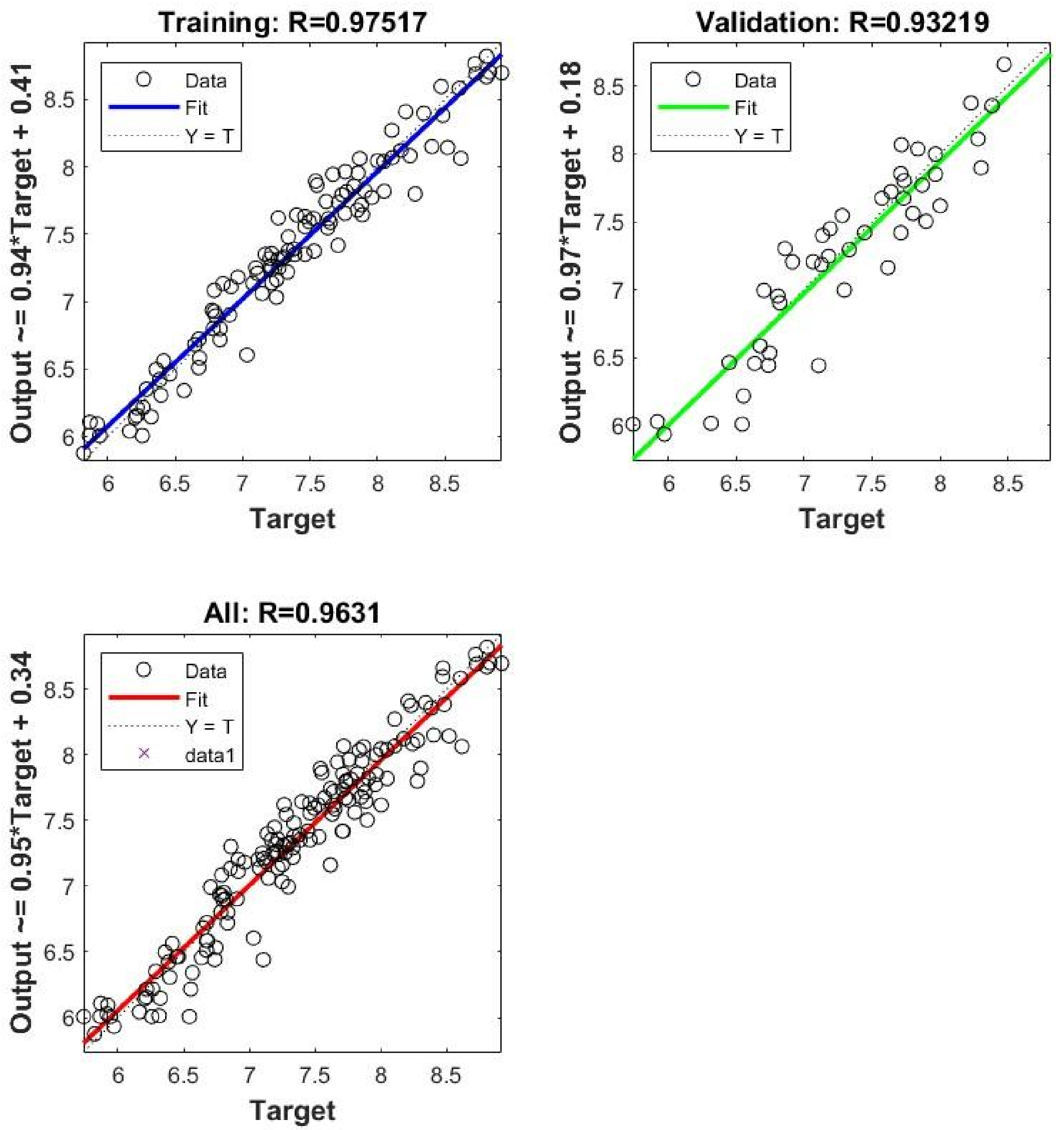
Output Layer: purelin



Activation Functions

Hidden Layer: satlin

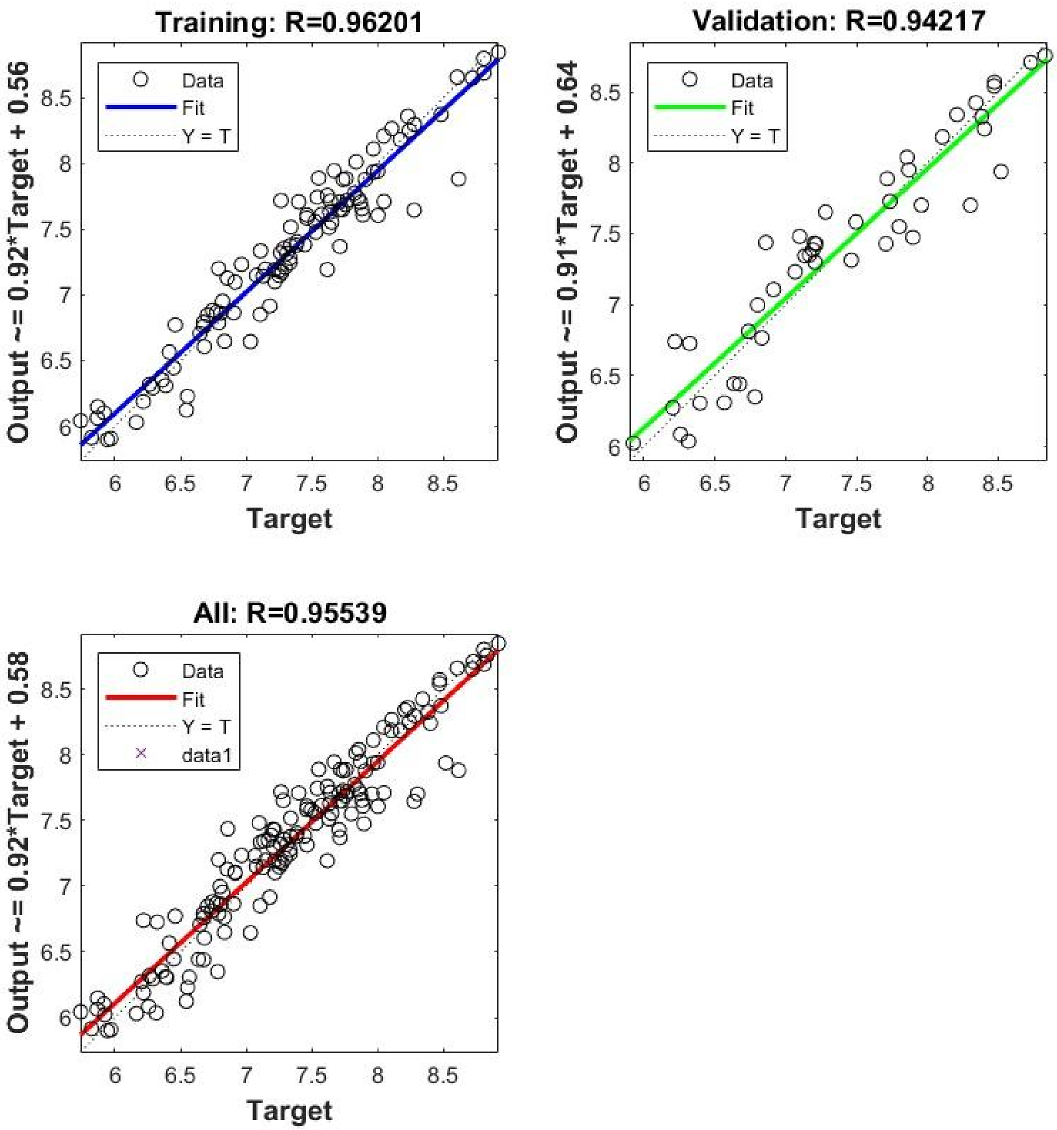
Output Layer: purelin



Activation Functions

Hidden Layer: poslin

Output Layer: purelin



2) Multivariate Adaptive Regression Splines (MARS)

CODE

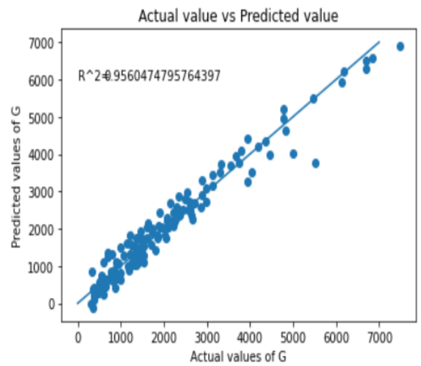
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
Shear_Modulus.csv

Disk 69.33 GB available

```
[131] y_pred_mars = MARS_model_fitted.predict(x)
      x=r2_score(y, y_pred_mars)
      plt.scatter(y, y_pred_mars)
      plt.plot(range(7000), range(7000))

      plt.xlabel("Actual values of G")
      plt.ylabel("Predicted values of G")
      plt.title("Actual value vs Predicted value")
      plt.text(0,6000,'R^2=')
      plt.text(600,6000,x)
```

```
Text(600, 6000, '0.9560474795764397')
```



✓ 0s completed at 1:37 PM

sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
Shear_Modulus.csv

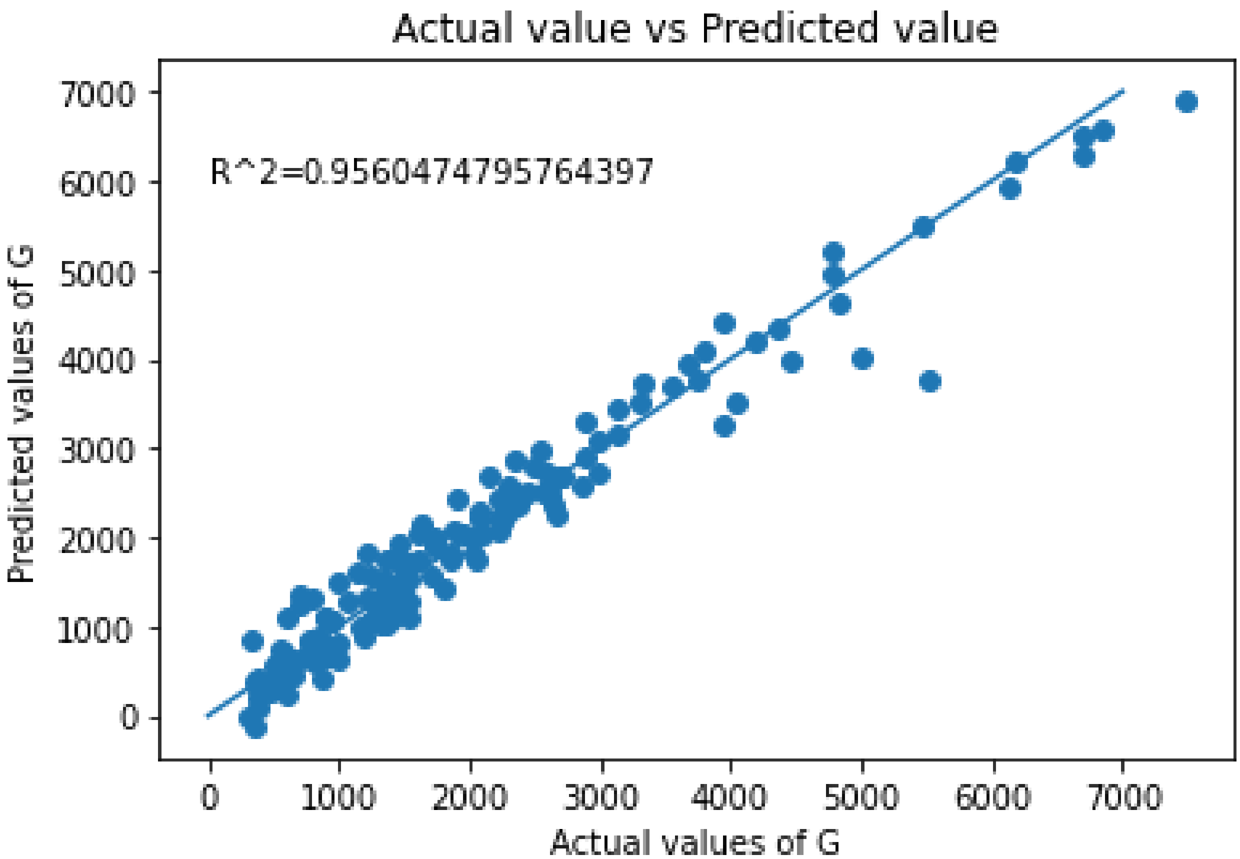
Disk 69.33 GB available

```
[132] mean_squared_error(y, y_pred_mars)
      print(np.sqrt(mean_squared_error(y, y_pred_mars)))
```

```
314.8546808105388
```

```
[133] print(MARS_model_fitted.summary())
```

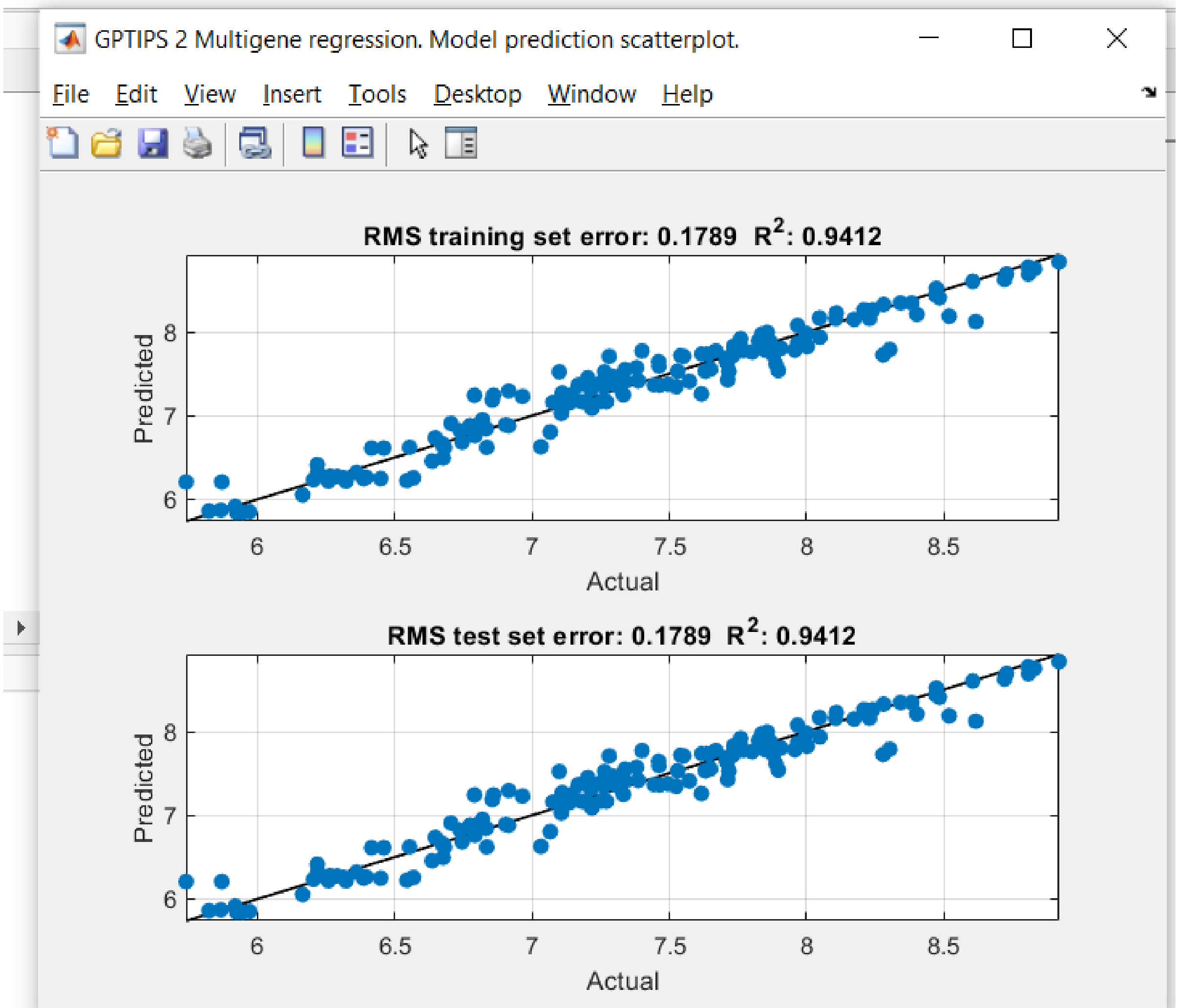
```
Earth Model
-------------------------------------
Basis Function  Pruned  Coefficient
-------------------------------------
(Intercept)     No      -2740.81
ShS             Yes     None
Age             No      19229.2
ShS*ShS         No      1155.25
Age*ShS         No      -78.4531
UW              No      412.407
CP              No      9.22808
ShS*ShS*ShS     No      -278.907
F*UW            No      65.4434
Age*Age         No      -1206.38
CP*ShS          No      -9.42502
Age*ShS*ShS     No      18.9459
UW*ShS          No      -235.495
CP*Age          No      21.8072
UW*Age*Age      No      1.28671
F*UW*ShS        No      -21.7568
ShS*UW*Age*Age  No      -0.231792
POP*CP          Yes     None
POP*POP*CP      No      3810.92
POP*POP*POP*CP  No      -21795.5
ShS*ShS*ShS*ShS Yes     None
POP*CP*ShS      Yes     None
ShS*F*UW*ShS    Yes     None
```



Actual value vs Predicted value
R^2=0.9560474795764397
Predicted values of G
Actual values of G
0
1000
2000
3000
4000
5000
6000
7000

3) Multi-Gene Genetic Programming (MGGP)

CODE



COMPARISON OF THE ABOVE THREE PREDICTIVE MODELS

Predictive Models	R^2
ANN	0.9970
MARS	0.9560
MGGP	0.9412

MODEL VALIDITY

MAD = 1/n($\sum_{i=1}^{n}$ (Pi-Median(P))

where P stands for the predicted values while n shows the length of the predicted data.

Then, the model's uncertainty can be obtained using the calculated MAD as follows

Uncertainty = (MAD/Median(P))*100

COMPARISON

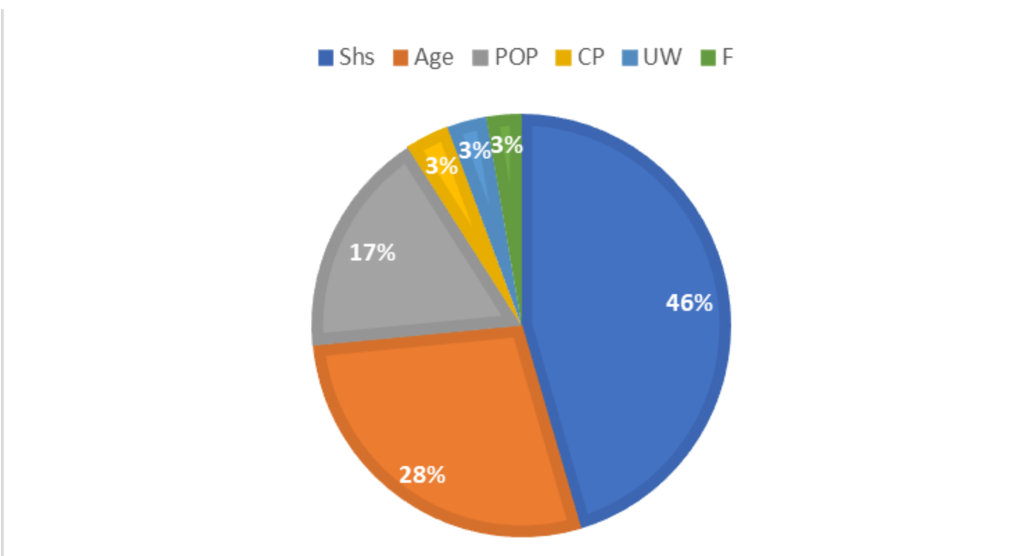
Predictive Models	Uncertainty %
ANN	18.5
MARS	26.8
MGGP	31.2

PARAMETRIC STUDY & SENSITIVITY ANALYSIS

A parametric study is an investigation of the effects of each input variables on the output. In this study the influence of changes in different parameters (i.e., ShS, Age, POP, UW, CP, and F) on the shear modulus was investigated using the ANN's network prediction function **"predict = net(input)".**

Input	MAPE	MAPE (10% reduction in Input)
Shs	6%	120%
Age	6%	74%
POP	6%	46%
CP	6%	9%
UW	6%	8%
F	6%	7%

Effect of each input parameter on the output -



Input	**-10%**	**-8%**	**-5%**	**-3%**	**0%**	**+3%**	**+5%**	**+8%**	**+10%**
Shs	150%	140%	120%	60%	6%	45%	65%	80%	100%
Age	60%	58%	40%	20%	6%	18%	20%	22%	25%
POP	20%	19%	18%	16%	6%	15%	15%	30%	40%
CP	8%	7.8%	8%	5.7%	6%	9%	9%	12%	15%
UW	9%	8.5%	8%	5.6%	6%	8%	8%	10%	8%
F	7%	8.6%	8%	5.6%	6%	7%	7%	8%	9%

Results on a new dataset with the following inputs as parameters -

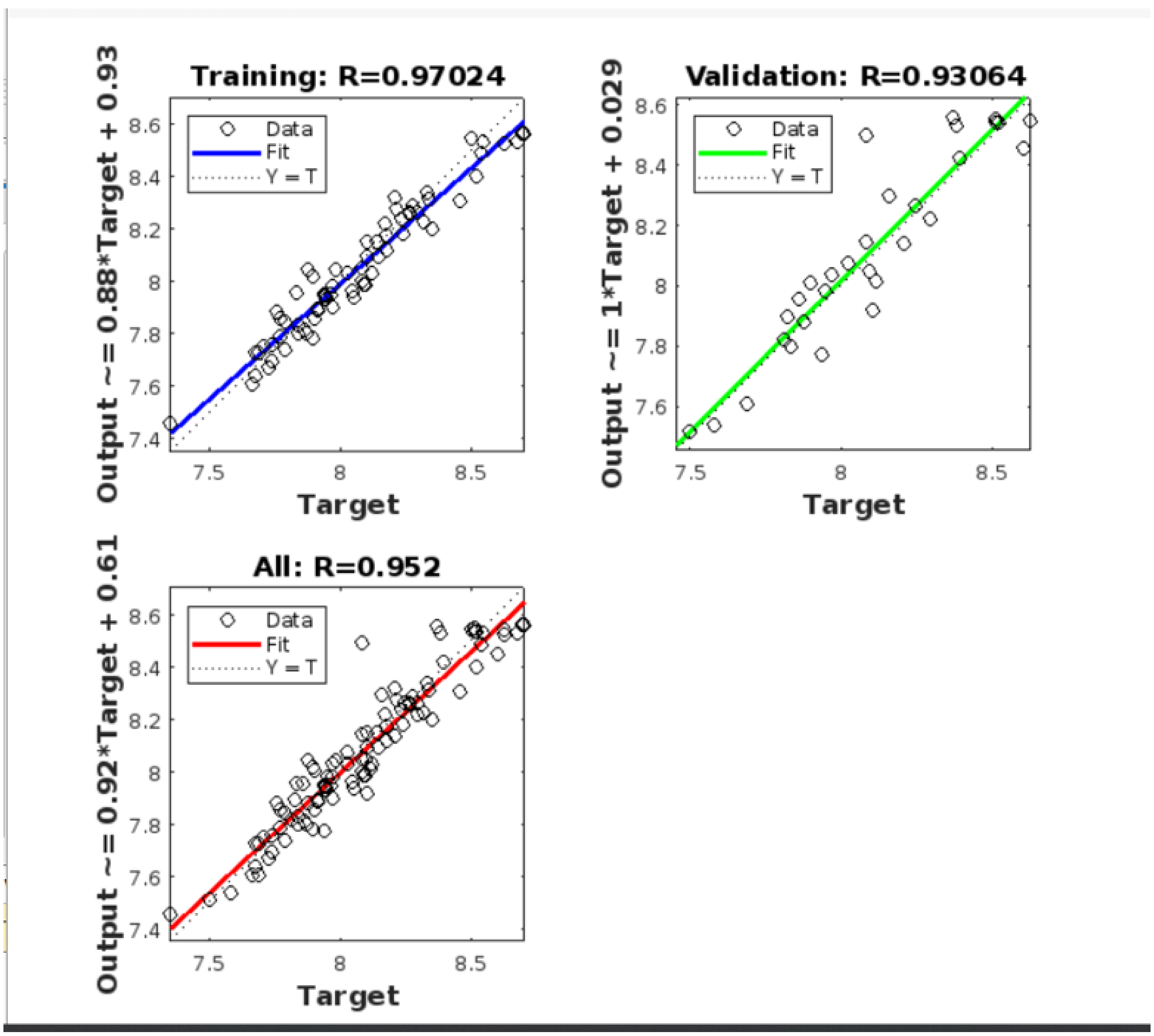
1)ShS

2)CP

3)F

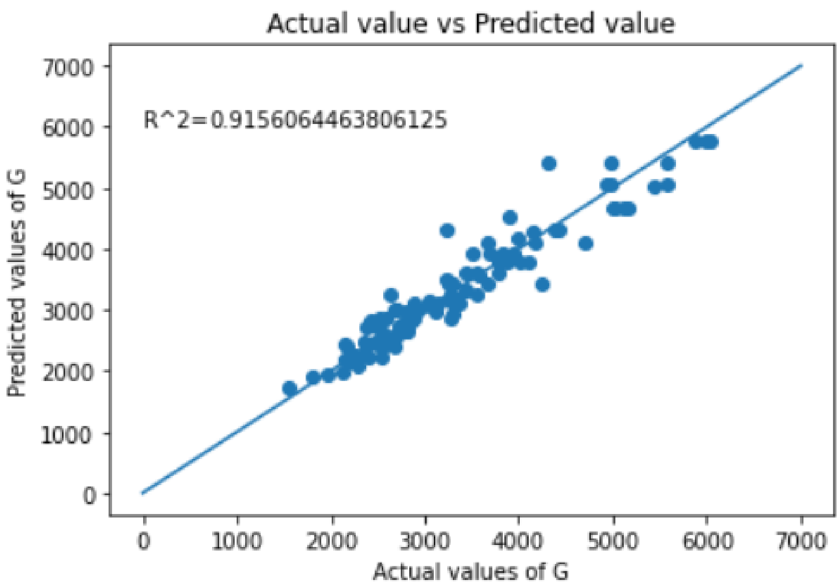
4)DD

1)ANN



Training: R=0.97024
Output ~= 0.88*Target + 0.93
Data
Fit
Y = T
Target
Validation: R=0.93064
Output ~= 1*Target + 0.029
All: R=0.952
Output ~= 0.92*Target + 0.61

2)MARS


Actual value vs Predicted value
R^2=0.9156064463806125
Predicted values of G
Actual values of G
0
1000
2000
3000
4000
5000
6000
7000

Parametric Study and Sensitivity Analysis (ANN model)

Input	MAPE	MAPE (10% reduction in Input)
Shs	12%	86%
CP	12%	32%
DD	12%	28%
F	12%	22%

Effect of each input parameter on the output