2018/2019 COMP1037 Coursework 2 – Machine Learning

Report

1. Dataset Preparation

Implemented

2. NN model training

Implemented

3. NN Model Testing

Implemented

4.Experiments

a. I gradually increase the hidden layer from 15 to 50 to 150,the accuracy was increased a few.

If I added more hidden layers,the thing will happen is that errors will be reached,the accuracy will be reduced and the iteration will reach 1000 easily.

b. If you shrink the picture size to 40x40 or 100x100,the most conspicuous effect is that the training rate is dramatically improved while the picture becomes vague and accuracy doesn't change too much.On the contrary,if you enlarge the picture the process will slow down.

c. Change the ratio of training set is the most obvious effect,for instance,the training set is increased and the test set is reduced,this would cause the accuracy increase dramatically.

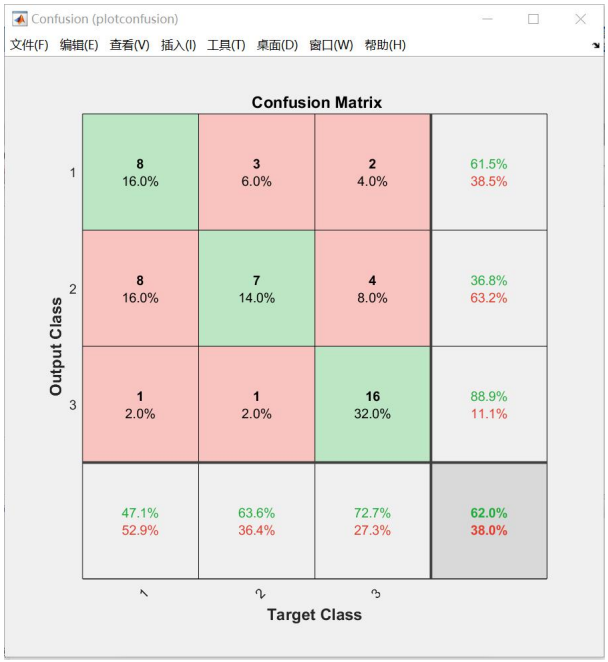
5.Reports

a.dataset structure

In build_dataset.m,X and y is built,in a easy sense,X is a matrix that stores all the pictures in the form of pixel intensity,y,on the other hand is a matrix that indicates which picture is dog,which is cat,which is panda.Every label is represented by either 100,010,or 001.

b.Confusion matrix

2018/2019 COMP1037 Coursework 2 – Machine Learning



(Confusion matrix)

Each column of the Confusion matrix represents the prediction category, and the total number of each column represents the number of data in that category. Each row represents the true category to which the data belongs, and the total number of data instances for each row represents the number of data instances for that class. The values in each column represent the number of classes to which the real data were predicted: as shown in the figure , the 110 in the first column of the first row indicates that 110 instances actually belonging to the first category were predicted to be in the first category; similarly, the 67 in the second column of the first row indicates that two instances actually belonging to the first category were wrongly predicted to be in the second category.

The green block means the right prediction while red means wrong prediction.On the right-bottom corner is the accuracy rate.

C. I increased the nodes inside the hidden layer,increase accuracy

```
% create a neural network
net = patternnet(150);
```

I increase the training set quantity,increased accuracy

```
% divided into training, validation and testing simulate
net.divideParam.trainRatio = 0.9;
net.divideParam.valRatio = 0;
net.divideParam.testRatio = 0.1;

rand_indices = randperm(size(X, 2));

trainData = X(:, rand_indices(1:2700));
trainLabels = y(:, rand_indices(1:2700));
testData = X(:, rand_indices(2701:end));
testLabels = y(:, rand_indices(2701:end));
```

2018/2019 COMP1037 Coursework 2 – Machine Learning

You can even choose the training function and other functions

Acronym	Algorithm	Description
LM	trainlm	Levenberg-Marquardt
BFG	trainbfg	BFGS Quasi-Newton
RP	trainrp	Resilient Backpropagation
SCG	trainscg	Scaled Conjugate Gradient
CGB	traincgb	Conjugate Gradient with Powell/Beale Restarts
CGF	traincgf	Fletcher-Powell Conjugate Gradient
CGP	traincgp	Polak-Ribiére Conjugate Gradient
OSS	trainoss	One Step Secant
GDX	traingdx	Variable Learning Rate Backpropagation

I change the fail exception to 50,a moderate level of 50.

```
% create a neural network
net =  patternnet(30);
net.trainParam.max_fail=50;
```

6.Haven't done

7.Haven't done