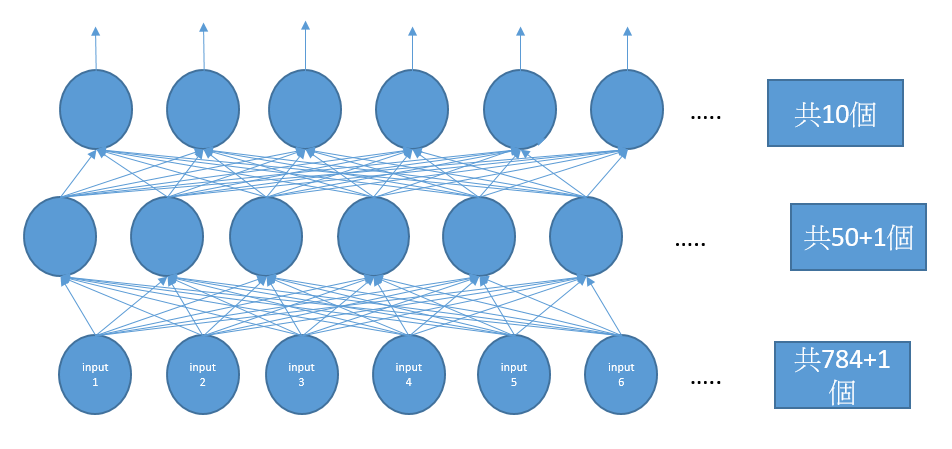
這次設計1層hidden layer，50個hidden nodes，架構如下



學習率採用0.1

本次測試了兩種情況

分別為

error>0.03 && iteration<100

error>0.01 && iteration<100

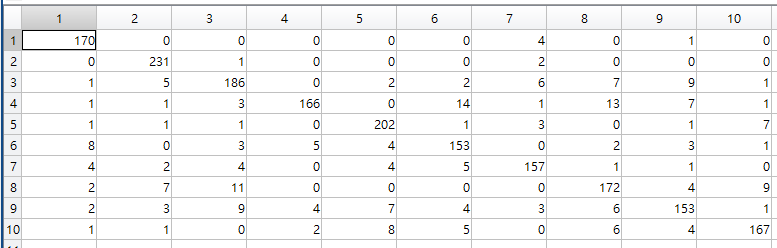
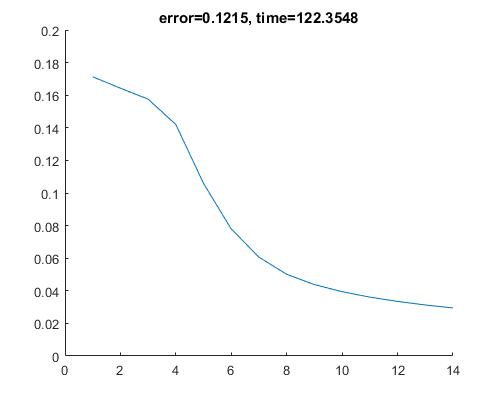
在error>0.03 && iteration<100為停止條件，收斂速度滿快的，時常要跑到15iteration以下就能收斂，時間大概120秒，iteration設為100避免無法收斂。

在error>0.01 && iteration<100為停止條件，收斂速度相對較慢，大概要跑到50iteration左右才能達到，時間大概480秒。

本次訓練使用mnist前5000筆training data來訓練，測試使用mnist前2000筆testing data測試，錯誤率大概在12%左右。training error<0.01的情況，有時候錯誤率可以達到11%，但大部分的情況並沒有明顯的好處，training時間卻要從120秒左右提升到接近500秒。

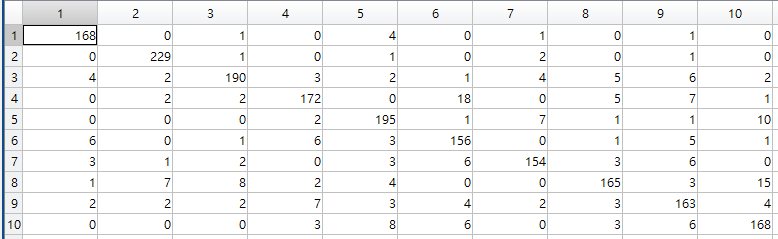
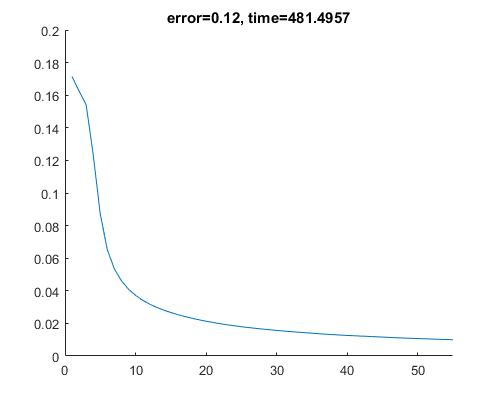
以下為結果

Error>0.03

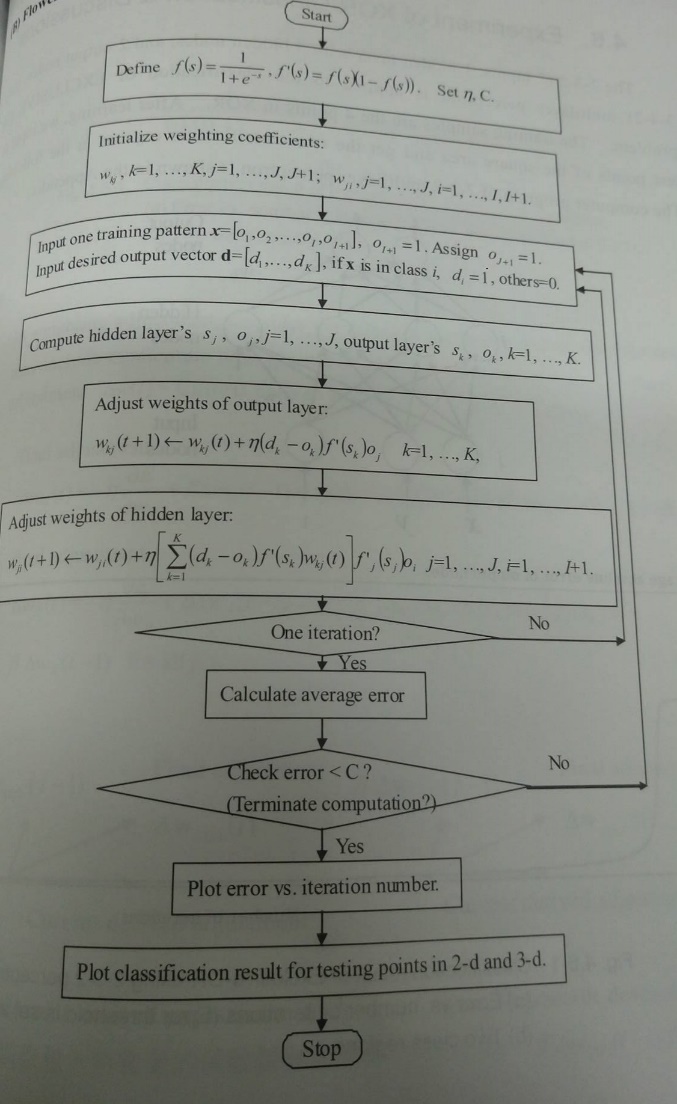


此表1對應到0，2對應到1，以此類推，row為原本的label col為被判斷的

Error>0.01



Flow chart



Matlab code

hFig = figure(1);

set(hFig,'name','1 hidden layers , 50 hidden nodes,learning rate=0.1, error>0.03 && iteration<100','Position', [200 500 500 400]);

alpha=0.1;

%read

fp = fopen('train-images.idx3-ubyte', 'rb');

magic = fread(fp, 1, 'int32', 0, 'ieee-be');

numImages = fread(fp, 1, 'int32', 0, 'ieee-be');

numRows = fread(fp, 1, 'int32', 0, 'ieee-be');

numCols = fread(fp, 1, 'int32', 0, 'ieee-be');

images = fread(fp, inf, 'unsigned char');

images = reshape(images, numCols, numRows, numImages);

images = permute(images,[2 1 3]);

fclose(fp);

images = reshape(images, size(images, 1) \* size(images, 2), size(images, 3));

images = double(images) / 255;

%

fp = fopen('train-labels.idx1-ubyte', 'rb');

magic = fread(fp, 1, 'int32', 0, 'ieee-be');

numLabels = fread(fp, 1, 'int32', 0, 'ieee-be');

labels = fread(fp, inf, 'unsigned char');

fclose(fp);

%read

WL1=0.1\*rand(50,785);

WL2=0.1\*rand(10,51);

error=1;

it=0;

d=[0 0 0 0 0 0 0 0 0 0];

errorstop=0.03;

time=tic;

while(error>errorstop && it<100)

it=it+1;

error=0;

for i=1:5000

for j=1:10

d(j)=0;

end

d(labels(i)+1)=1;

for j=1:50

SL1(j)=0;

for k=1:784

SL1(j)=SL1(j)+WL1(j,k)\*images(k,i);

end

SL1(j)=SL1(j)+WL1(j,785);

OL1(j)=1.0/(1.0+exp(-SL1(j)));

end

for j=1:10

SL2(j)=0;

for k=1:50

SL2(j)=SL2(j)+WL2(j,k)\*OL1(k);

end

SL2(j)=SL2(j)+WL2(j,51);

OL2(j)=1.0/(1.0+exp(-SL2(j)));

end

temperror=0;

for j=1:10

temperror=temperror+sqrt((d(j)-OL2(j))\*(d(j)-OL2(j)));

end

error=error+temperror/10;

%WL1

for j=1:50

temp=0;

for l=1:10

temp=temp+(d(l)-OL2(l))\*OL2(l)\*(1-OL2(l))\*WL2(l,j);

end

for k=1:784

WL1(j,k)= WL1(j,k)+alpha\*temp\*OL1(j)\*(1-OL1(j))\*images(k,i);

end

WL1(j,785)= WL1(j,41)+alpha\*temp\*OL1(j)\*(1-OL1(j))\*1;

end

%WL2

for j=1:50

for k=1:10

WL2(k,j)= WL2(k,j)+alpha\*(d(k)-OL2(k))\*OL2(k)\*(1-OL2(k))\*OL1(j);

end

end

for k=1:10

WL2(k,51)= WL2(k,51)+alpha\*(d(k)-OL2(k))\*OL2(k)\*(1-OL2(k))\*1;

end

end

error=error/5000;

EIT(1,it)=it;

EIT(2,it)=error;

end

timetoc=toc(time);

%read

fp = fopen('t10k-images.idx3-ubyte', 'rb');

magic = fread(fp, 1, 'int32', 0, 'ieee-be');

numImages = fread(fp, 1, 'int32', 0, 'ieee-be');

numRows = fread(fp, 1, 'int32', 0, 'ieee-be');

numCols = fread(fp, 1, 'int32', 0, 'ieee-be');

images = fread(fp, inf, 'unsigned char');

images = reshape(images, numCols, numRows, numImages);

images = permute(images,[2 1 3]);

fclose(fp);

images = reshape(images, size(images, 1) \* size(images, 2), size(images, 3));

images = double(images) / 255;

%

fp = fopen('t10k-labels.idx1-ubyte', 'rb');

magic = fread(fp, 1, 'int32', 0, 'ieee-be');

numLabels = fread(fp, 1, 'int32', 0, 'ieee-be');

labels = fread(fp, inf, 'unsigned char');

fclose(fp);

%read

finalmatch=zeros(10,10);

error=0;

for i=1:2000

for j=1:10

d(j)=0;

end

d(labels(i)+1)=1;

for j=1:50

SL1(j)=0;

for k=1:784

SL1(j)=SL1(j)+WL1(j,k)\*images(k,i);

end

SL1(j)=SL1(j)+WL1(j,785);

OL1(j)=1.0/(1.0+exp(-SL1(j)));

end

for j=1:10

SL2(j)=0;

for k=1:50

SL2(j)=SL2(j)+WL2(j,k)\*OL1(k);

end

SL2(j)=SL2(j)+WL2(j,51);

OL2(j)=1.0/(1.0+exp(-SL2(j)));

end

tempol=OL2(1);

tempj=1;

for j=1:10

if tempol<OL2(j)

tempol=OL2(j);

tempj=j;

end

end

if tempj~=labels(i)+1

error=error+1;

end

finalmatch(labels(i)+1,tempj)= finalmatch(labels(i)+1,tempj)+1;

end

error=error/2000

line(EIT(1,:),EIT(2,:));

axis([0 it 0 0.2]);

str = ['error=', num2str(error), ', time=', num2str(timetoc)];

title(str);

參考mnist官網最下面的表格，決定layer數，上網搜尋決定50個layer，參考<https://github.com/davidstutz/matlab-mnist-two-layer-perceptron>

與mnist官網讀入mnist檔案。發現train時計算的error越小，不一定代表測試出來的結果會越好，有發生過training error<0.01，testing error卻大於0.13(13%)的情況。