1.

（1）

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| class | - | - | - | - | + | + | - | - | + | + |
|  | 0.08 | 0.15 | 0.35 | 0.44 | 0.45 | 0.47 | 0.55 | 0.67 | 0.69 | 0.73 |
| TP | 2 | | | | | | | | | |
| FP | 2 | | | | | | | | | |
| TN | 4 | | | | | | | | | |
| FN | 2 | | | | | | | | | |
| accuracy | 0.6 | | | | | | | | | |
| precision | 0.5 | | | | | | | | | |
| TPR(recall) | 0.5 | | | | | | | | | |
| FPR | 1/3 | | | | | | | | | |
| F-measure | 0.5 | | | | | | | | | |

（2）

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| class | + | + | - | - | + | - | - | - | + | - |
|  | 0.01 | 0.03 | 0.04 | 0.05 | 0.09 | 0.31 | 0.38 | 0.45 | 0.61 | 0.68 |
| TP | 1 | | | | | | | | | |
| FP | 1 | | | | | | | | | |
| TN | 5 | | | | | | | | | |
| FN | 3 | | | | | | | | | |
| accuracy | 0.6 | | | | | | | | | |
| precision | 0.5 | | | | | | | | | |
| TPR(recall) | 0.25 | | | | | | | | | |
| FPR | 1/6 | | | | | | | | | |
| F-measure | 1/3 | | | | | | | | | |

TPR,M1>M2,分类模型M1在这个测试集上表现得更好。

（3）

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| class | - | - | - | - | + | + | - | - | + | + |
|  | 0.08 | 0.15 | 0.35 | 0.44 | 0.45 | 0.47 | 0.55 | 0.67 | 0.69 | 0.73 |
| TP | 4 | | | | | | | | | |
| FP | 4 | | | | | | | | | |
| TN | 2 | | | | | | | | | |
| FN | 0 | | | | | | | | | |
| accuracy | 0.6 | | | | | | | | | |
| precision | 0.5 | | | | | | | | | |
| TPR(recall) | 1 | | | | | | | | | |
| FPR | 2/3 | | | | | | | | | |
| F-measure | 2/3 | | | | | | | | | |

TPR=1,阈值为0.2时结果更好。

（4）

对于模型M1，

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| class | - | - | - | - | + | + | - | - | + | + |  |
| Threshold>= | 0.08 | 0.15 | 0.35 | 0.44 | 0.45 | 0.47 | 0.55 | 0.67 | 0.69 | 0.73 | 1.0 |
| TP | 4 | 4 | 4 | 4 | 4 | 3 | 2 | 2 | 2 | 1 | 0 |
| FP | 6 | 5 | 4 | 3 | 2 | 2 | 2 | 1 | 0 | 0 | 0 |
| TN | 0 | 1 | 2 | 3 | 4 | 4 | 4 | 5 | 6 | 6 | 6 |
| FN | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 2 | 2 | 3 | 4 |
| accuracy | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.7 | 0.6 | 0.7 | 0.8 | 0.7 | 0.6 |
| precision |  |  |  |  |  |  |  |  |  |  |  |
| TPR | 1 | 1 | 1 | 1 | 1 | 0.75 | 0.5 | 0.5 | 0.5 | 0.25 | 0 |
| FPR | 1 | 5/6 | 2/3 | 0.5 | 1/3 | 1/3 | 1/3 | 1/6 | 0 | 0 | 0 |
| F-measure |  |  |  |  |  |  |  |  |  |  |  |

阈值取0.45时最优，此时accuracy = 0.8, TPR=1，FPR=1/3.

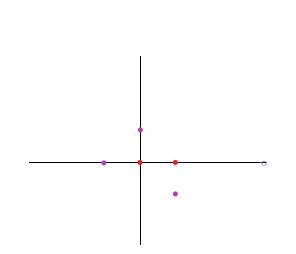
对于模型M2，

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| class | + | + | - | - | + | - | - | - | + | - |
| Threshold>= | 0.01 | 0.03 | 0.04 | 0.05 | 0.09 | 0.31 | 0.38 | 0.45 | 0.61 | 0.68 |
| TP | 4 | 3 | 2 | 2 | 2 | 1 | 1 | 1 | 1 | 0 |
| FP | 6 | 6 | 6 | 5 | 4 | 4 | 3 | 2 | 1 | 1 |
| TN | 0 | 0 | 0 | 1 | 2 | 2 | 3 | 4 | 5 | 5 |
| FN | 0 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 4 |
| accuracy | 0.4 | 0.3 | 0.2 | 0.3 | 0.4 | 0.3 | 0.4 | 0.5 | 0.6 | 0.5 |
| precision |  |  |  |  |  |  |  |  |  |  |
| TPR(recall) | 1 | 0.75 | 0.5 | 0.5 | 0.5 | 0.25 | 0.25 | 0.25 | 0.25 | 0 |
| FPR | 1 | 1 | 1 | 5/6 | 2/3 | 2/3 | 0.5 | 1/3 | 1/6 | 1/6 |
| F-measure |  |  |  |  |  |  |  |  |  |  |

阈值取0.61时最优，此时accuracy = 0.6, TPR = 0.25, FPR = 1/6.

2.

（1）

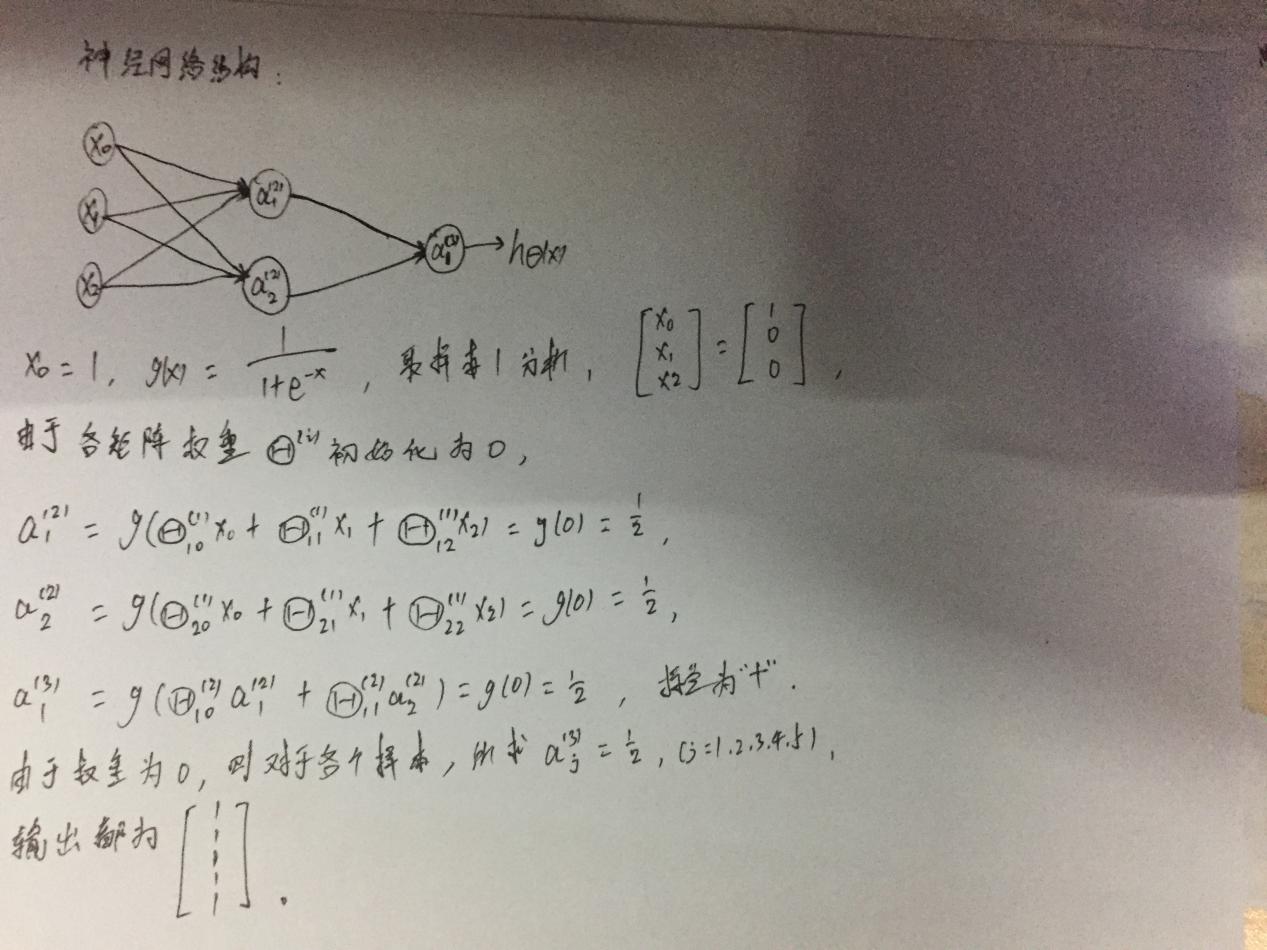


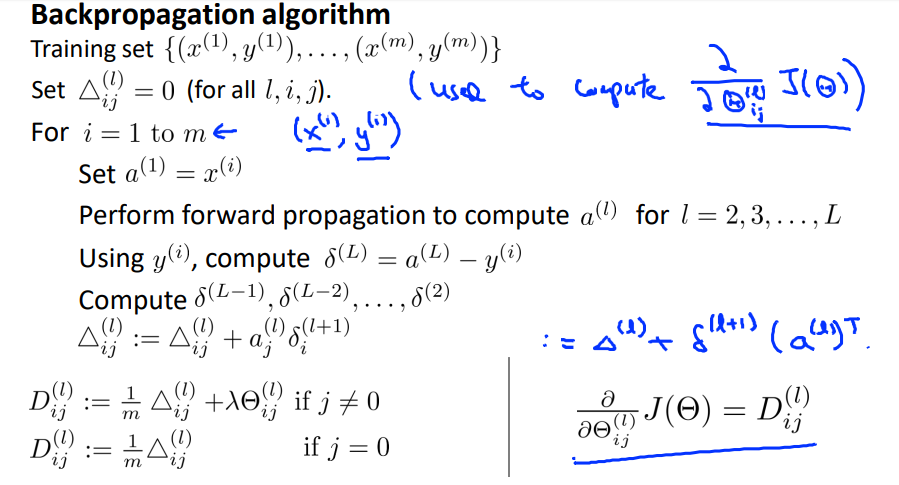
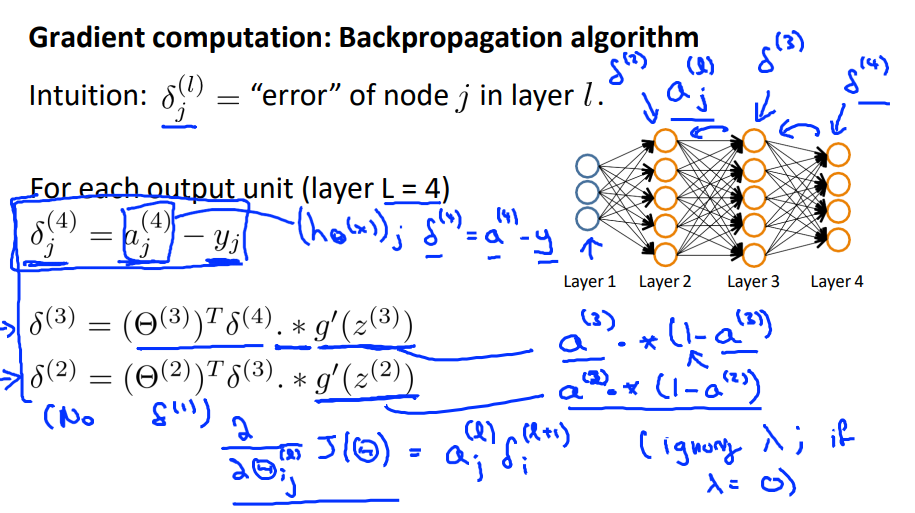
由图可以看出，不存在任意一条直线可以将样本正确分类，所以此训练样本线性不可分。

（2）

使用Sigmoid激活函数，给神经元引入了非线性因素，使得神经网络可以任意逼近任何非线性函数，这样神经网络就可以应用到众多的非线性模型中。而如果将Sigmoid函数换成线性函数，则隐藏层就失去了意义，每一层输出都是上层输入的线性函数，无论神经网络有多少层，输出都是输入的线性组合。

（3）

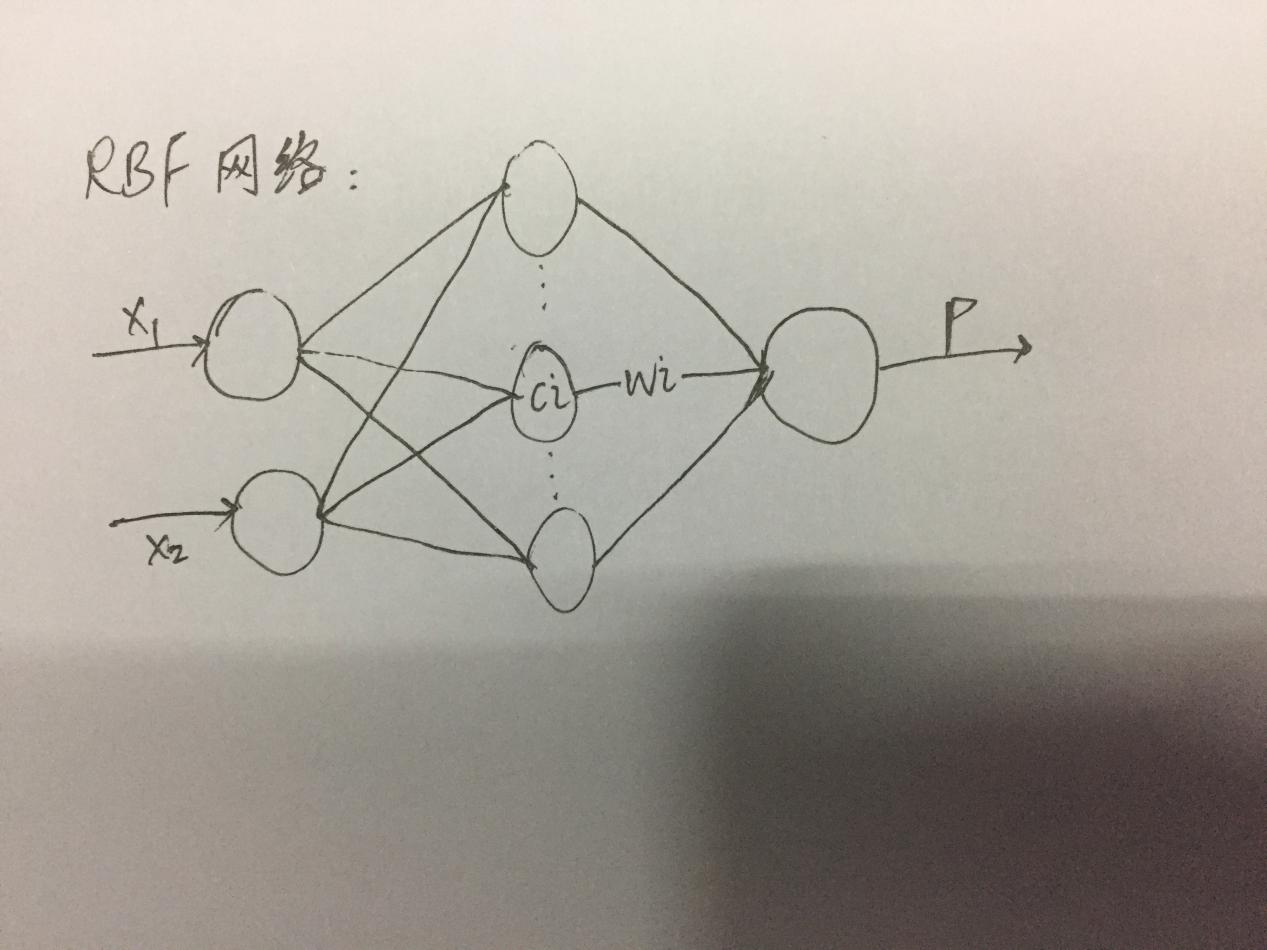


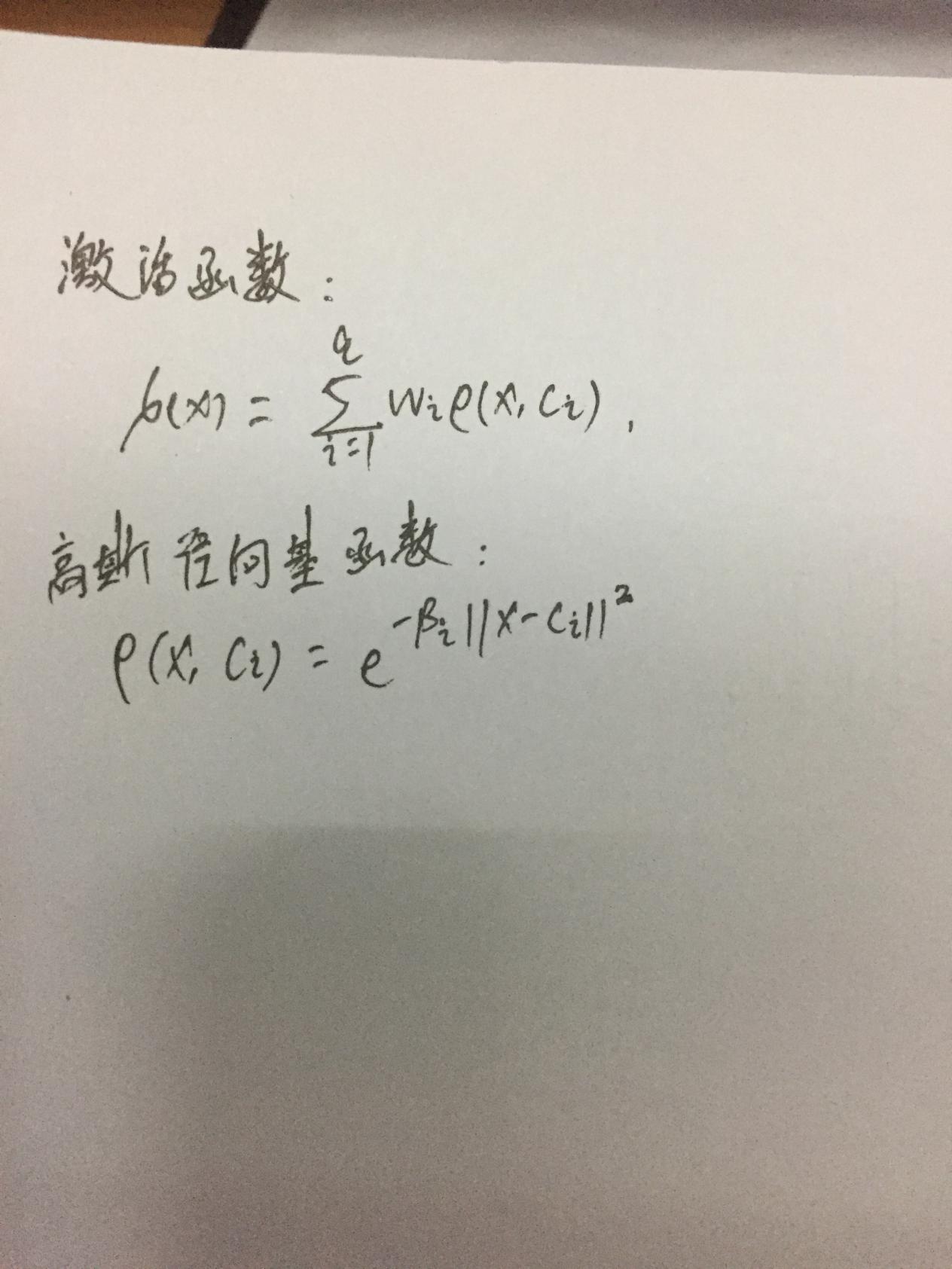


在前馈算法中，矩阵权重初始化为0导致计算的结果都为‘+’，分类不正确，神经网络没起到任何作用；在BP算法中，权重初始为0，将导致最终结果都为0，反向传播也失去了意义。

（4）

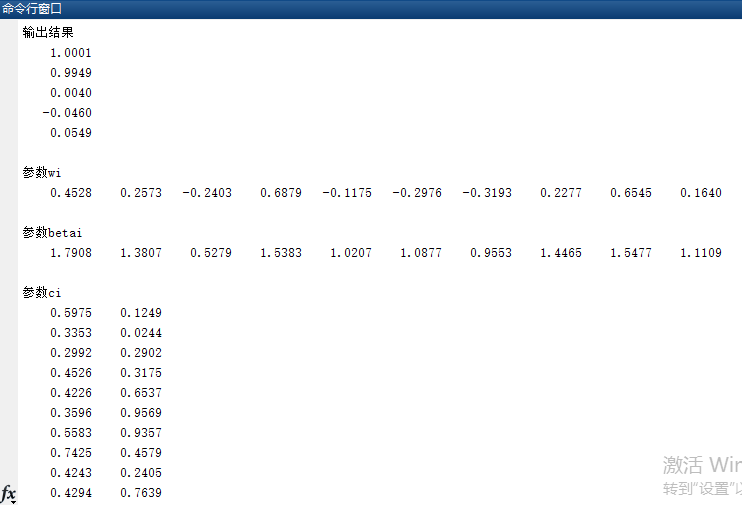
设计RBF神经网络如图，





通常采用两步过程来训练RBF网络，第一步：确定神经元中心Ci，常用的方式包括随机取样，聚类等；第二步，利用BP算法等来确定参数wi和βi。

经过代码测试，取隐藏层数为10，学习率为0.5，结果如图：



3.

（1）

根节点信息熵Ent(root) = -= -(4/10)-(6/10)=0.97

以属性A划分的信息熵Ent() = 0.985, Ent() = 0;

信息熵增益GAIN = Ent(D)-Ent() = 0.28;

以属性B划分的信息熵Ent() = 0.81, Ent() = 0.65;

信息熵增益GAIN = Ent(D)-Ent() = 0.26;

以A划分的信息熵增益更大，所以选择A划分。

（2）

根节点Gini系数Gini(root) = 0.48;

以属性A划分Gini(T) = 0.49, Gini(F) = 0;Gini增益为0.48-0.34 = 0.12;

以属性B划分Gini(T) = 0.38,Gini(F) = 0.28,Gini增益为0.48-0.32 = 0.16;

以属性B划分的Gini增益更大，所以选择B划分。

（3）

根节点的分类误差Error(root) = 0.4;

以属性A划分Error(T) = 3/7,Error(F) = 0,Error增益为0.1;

以属性B划分Error(T) = 0.25,Error(F) = 1/6,Error增益为0.2;

以属性B划分的分类误差增益更大，所以选择B划分。

（4）

信息熵增益偏好于属性A，Gini增益，分类误差增益更偏好于属性B。