# 模式识别实验报告

## 实验三 线性分类器

学院：计算学部

姓名： 张亚博

学号：1170300521

1. **实验内容**
2. 使用Python或Matlab编程实现感知器算法和最小平方误差算法；
3. 分别使用感知器算法学习区分下列两类样本的线性分类器：



1. MNIST数据集测试：使用TrainSamples中的30000个17维特征手写数字样本训练线性分类器区分10个类别，TrainLabels中包含训练样本的标签；测试线性分类器对TestSamples中10000个样本的识别正确率。
2. **程序代码**

（感知器算法和最小平方误差算法，矩阵乘法和求逆可以调用其他函数库中的程序）

**感知机算法**：

**class** **Perceptron**(object):

**def** \_\_init\_\_(self):

self**.**learning\_step **=** 0.00001

self**.**max\_iteration **=** 5000

**def** **predict\_**(self, x):

wx **=** sum([self**.**w[j] **\*** x[j] **for** j **in** xrange(len(self**.**w))])

**return** int(wx **>** 0)

**def** **train**(self, features, labels):

self**.**w **=** [0.0] **\*** (len(features[0]) **+** 1) *#权重+偏置*

correct\_count **=** 0

time **=** 0

**while** time **<** self**.**max\_iteration:

index **=** np**.**random**.**randint(0, len(labels) **-** 1) *#随机找个样本*

x **=** list(features[index])

x**.**append(1.0) *#偏置*

y **=** 2 **\*** labels[index] **-** 1

wx **=** sum([self**.**w[j] **\*** x[j] **for** j **in** range(len(self**.**w))])

**if** wx **\*** y **>** 0:

correct\_count **+=** 1

**if** correct\_count **>** self**.**max\_iteration:

**break**

**continue**

*#错误就更新参数*

**for** i **in** range(len(self**.**w)):

self**.**w[i] **+=** self**.**learning\_step **\*** (y **\*** x[i])

**def** **predict**(self,features):

labels **=** []

**for** feature **in** features:

x **=** list(feature)

x**.**append(1)

labels**.**append(self**.**predict\_(x))

**return** labels

**最小平方误差算法：**

class LMSE:

def \_\_init\_\_(self, x1, x2):

self.y = []

for x in x1:

self.y.append(np.append(1, x1))

for x in x2:

self.y.append(np.append(-1, -x2))

self.y = np.array(self.y)

self.n, self.dim = self.y.shape

self.b = np.ones(self.n)

self.W – np.ones(self.dim) + np.random.rand(self.dim) – 0.5

def train(self, learn\_rate=0.001):

loss = 10

while loss > 1e-5:

sum\_y = np.zeros(self.dim)

for I in range(self.n):

sum\_y += (self.b[i] – self.W.dot(self.y[i])) \* self.y[i]

sum\_y\*= learn\_rate

self.W += sum\_y

loss = np.linalg.norm(sum\_y)

return self.W

1. **实验结果**
2. 仿真数据实验结果：分别给出使用感知器算法和最小平方误差算法得到的线性判别函数。

设判别式为y = WT \* X + b

感知机算法中：W=(-1.9952, -0.9952)T, b=2

最小平方误差算法中：W=(-1.9908, -0.9950)T, b=2

1. MNIST数据集实验结果：（多类别解决方案及分类正确率）

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 算法 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Total |
| 感知机 | 0.91 | 0.95 | 0.73 | 0.80 | 0.63 | 0.44 | 0.88 | 0.84 | 0.68 | 0.64 | 0.76 |
| MSE | 0.92 | 0.95 | 0.74 | 0.72 | 0.73 | 0.46 | 0.86 | 0.89 | 0.73 | 0.48 | 0.75 |