**决策树**

**一、实验目的**

1. 了解典型决策树算法。
2. 熟悉决策树算法的思路与步骤。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

1.编程实现对数据集做决策树分析，并画出最终的决策树。

数据集名称：鸢尾花卉Iris数据集

选择了部分数据集来区分Iris Setosa（山鸢尾）及Iris Versicolour（杂色鸢尾）两个种类。

%% 部分数据预处理代码

ID3算法数据预处理，把字符串转换为0,1编码

txt={ '序号' '花萼大小' '花瓣长度' '花瓣宽度' '类型'

'' '小' '长' '长' 'versicolor'

'' '小' '长' '长' 'versicolor'

'' '小' '长' '长' 'versicolor'

'' '小' '短' '长' 'versicolor'

'' '小' '长' '长' 'versicolor'

'' '小' '短' '长' 'versicolor'

'' '小' '长' '短' 'versicolor'

'' '大' '长' '长' 'versicolor'

'' '大' '长' '短' 'versicolor'

'' '大' '长' '长' 'versicolor'

'' '大' '长' '长' 'versicolor'

'' '大' '长' '长' 'versicolor'

'' '大' '长' '长' 'versicolor'

'' '小' '长' '长' 'setosa'

'' '大' '短' '长' 'versicolor'

'' '大' '短' '长' 'versicolor'

'' '大' '短' '长' 'versicolor'

'' '大' '短' '长' 'versicolor'

'' '大' '短' '短' 'versicolor'

'' '小' '短' '短' 'setosa'

'' '小' '短' '长' 'setosa'

'' '小' '短' '长' 'setosa'

'' '小' '短' '长' 'setosa'

'' '小' '短' '短' 'setosa'

'' '小' '长' '短' 'setosa'

'' '大' '短' '长' 'setosa'

'' '大' '短' '长' 'setosa'

'' '小' '短' '短' 'setosa'

'' '小' '短' '短' 'setosa'

'' '大' '短' '短' 'setosa'

'' '小' '长' '短' 'setosa'

'' '大' '短' '长' 'setosa'

'' '大' '短' '短' 'setosa'

'' '大' '短' '短' 'setosa'

}

attributes=txt(1,2:end);% attributes: 属性和Label；

activeAttributes = ones(1,length(attributes)-1);% activeAttributes : 属性向量，全1；

data = txt(2:end,2:end);

%% 针对每列数据进行转换

[rows,cols] = size(data);

matrix = zeros(rows,cols);% matrix： 转换后的0,1矩阵；

for j=1:cols

matrix(:,j) = cellfun(@trans2onezero,data(:,j));

end

end

**BP神经网络**

**一、实验目的**

掌握利用感知器和BP网进行模式识别的基本方法，通过实验加深对基本概念的理解。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

感知器实验：1、设计线性可分实验，要求训练样本10个以上。

2、奇异样本对网络训练的影响。

3、以线性不可分样本集训练分类器。

BP网实验：利用BP网对上述线性不可分样本集进行分类。

线性可分实验数据

P=[-5 -7 -4 -1 0 -5 4 2 -4 -2 1 4 4 3 1 -2;0 -5 4 2 -4 1 4 1 -1 -3 -1 7 -2 -3 5 -5];T=[1 0 1 1 0 1 1 0 1 0 0 1 0 0 1 0];

奇异样本实验数据

P=[-3 -1 -5 4 2 -4 -2 1 4 3 1 -3;5 2 1 4 1 -1 -3 -1 -2 -4 5 -6];T=[1 1 1 1 0 1 0 0 0 0 1 0];

BP神经网络实验数据

P=[-5 -7 -4 -1 0 -5 4 2 -4 -2 1 4 4 3 1 -2;0 -5 4 2 -4 1 4 1 -1 -3 -1 7 -2 -3 5 -5];T=[1 0 1 0 0 1 1 0 1 0 0 1 0 0 1 0];

其中P代表二维空间中的数据点，T代表数据点的类别标签。

**朴素贝叶斯分类器设计**

**一、实验目的**

1.了解朴素贝叶斯算法的基本原理；

2.能够使用朴素贝叶斯算法对数据进行分类；

 3.学会对于分类器的性能评估方法。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

1.编程实现对数据集做朴素贝叶斯分类算法分类。

数据集名称：鸢尾花卉Iris数据集，详情见iris.txt。对于iris数据集，它包含了150个实例（每个分类包含50个实例），共有sepal length、sepal width、petal length、petal width和class五种属性。期中前四种属性为数值类型，class属性为分类属性。该数据集中的全部实例共可分为三类：Iris Setosa、Iris Versicolour和Iris Virginica。

实验数据集中所有的数据都是实验所需的，因此不存在属性筛选的问题。在txt文件中第一列为分类属性，其他四列为数值类型属性。

2.将Iris数据集每类选取25个样本，总计75个样本作为训练样本，其余75个样本作为测试样本。通过训练样本训练模型，最后对测试样本进行分类。

**主成分分析**

**一、实验目的**

掌握主成分分析原理，能够利用编程实现主成分分析，并能够利用算法解决简单问题。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

（1）针对实例数据,计算样本协方差矩阵，求解数据第一主成分，并重建原始数据。

（3）针对实例数据,采用自相关矩阵作为产生矩阵，对其做二维特征提取。

**支持向量机**

**一、实验目的**

掌握支持向量机原理，能够利用编程实现支持向量机，并能够利用算法解决简单问题。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

对于以下文件里的数据，用SVM进行两分类。

x1(1,1)=5.1418; x1(1,2)=0.5950;

x1(2,1)=5.5519; x1(2,2)=3.5091;

x1(3,1)=5.3836; x1(3,2)=2.8033;

x1(4,1)=3.2419; x1(4,2)=3.7278;

x1(5,1)=4.4427; x1(5,2)=3.8981;

x1(6,1)=4.9111; x1(6,2)=2.8710;

x1(7,1)=2.9259; x1(7,2)=3.4879;

x1(8,1)=4.2018; x1(8,2)=2.4973;

x1(9,1)=4.7629; x1(9,2)=2.5163;

x1(10,1)=2.7118; x1(10,2)=2.4264;

x1(11,1)=3.0470; x1(11,2)=1.5699;

x1(12,1)=4.7782; x1(12,2)=3.3504;

x1(13,1)=3.9937; x1(13,2)=4.8529;

x1(14,1)=4.5245; x1(14,2)=2.1322;

x1(15,1)=5.3643; x1(15,2)=2.2477;

x1(16,1)=4.4820; x1(16,2)=4.0843;

x1(17,1)=3.2129; x1(17,2)=3.0592;

x1(18,1)=4.7520; x1(18,2)=5.3119;

x1(19,1)=3.8331; x1(19,2)=0.4484;

x1(20,1)=3.1838; x1(20,2)=1.4494;

x1(21,1)=6.0941; x1(21,2)=1.8544;

x1(22,1)=4.0802; x1(22,2)=6.2646;

x1(23,1)=3.0627; x1(23,2)=3.6474;

x1(24,1)=4.6357; x1(24,2)=2.3344;

x1(25,1)=5.6820; x1(25,2)=3.0450;

x1(26,1)=4.5936; x1(26,2)=2.5265;

x1(27,1)=4.7902; x1(27,2)=4.4668;

x1(28,1)=4.1053; x1(28,2)=3.0274;

x1(29,1)=3.8414; x1(29,2)=4.2269;

x1(30,1)=4.8709; x1(30,2)=4.0535;

x1(31,1)=3.8052; x1(31,2)=2.6531;

x1(32,1)=4.0755; x1(32,2)=2.8295;

x1(33,1)=3.4734; x1(33,2)=3.1919;

x1(34,1)=3.3145; x1(34,2)=1.8009;

x1(35,1)=3.7316; x1(35,2)=2.6421;

x1(36,1)=2.8117; x1(36,2)=2.8658;

x1(37,1)=4.2486; x1(37,2)=1.4651;

x1(38,1)=4.1025; x1(38,2)=4.4063;

x1(39,1)=3.9590; x1(39,2)=1.3024;

x1(40,1)=1.7524; x1(40,2)=1.9339;

x1(41,1)=3.4892; x1(41,2)=1.2457;

x1(42,1)=4.2492; x1(42,2)=4.5982;

x1(43,1)=4.3692; x1(43,2)=1.9794;

x1(44,1)=4.1792; x1(44,2)=0.4113;

x1(45,1)=3.9627; x1(45,2)=4.2198;

x2(1,1)=9.7302; x2(1,2)=5.5080;

x2(2,1)=8.8067; x2(2,2)=5.1319;

x2(3,1)=8.1664; x2(3,2)=5.2801;

x2(4,1)=6.9686; x2(4,2)=4.0172;

x2(5,1)=7.0973; x2(5,2)=4.0559;

x2(6,1)=9.4755; x2(6,2)=4.9869;

x2(7,1)=9.3809; x2(7,2)=5.3543;

x2(8,1)=7.2704; x2(8,2)=4.1053;

x2(9,1)=8.9674; x2(9,2)=5.8121;

x2(10,1)=8.2606; x2(10,2)=5.1095;

x2(11,1)=7.5518; x2(11,2)=7.7316;

x2(12,1)=7.0016; x2(12,2)=5.4111;

x2(13,1)=8.3442; x2(13,2)=3.6931;

x2(14,1)=5.8173; x2(14,2)=5.3838;

x2(15,1)=6.1123; x2(15,2)=5.4995;

x2(16,1)=10.4188; x2(16,2)=4.4892;

x2(17,1)=7.9136; x2(17,2)=5.2349;

x2(18,1)=11.1547; x2(18,2)=4.4022;

x2(19,1)=7.7080; x2(19,2)=5.0208;

x2(20,1)=8.2079; x2(20,2)=5.4194;

x2(21,1)=9.1078; x2(21,2)=6.1911;

x2(22,1)=7.7857; x2(22,2)=5.7712;

x2(23,1)=7.3740; x2(23,2)=2.3558;

x2(24,1)=9.7184; x2(24,2)=5.2854;

x2(25,1)=6.9559; x2(25,2)=5.8261;

x2(26,1)=8.9691; x2(26,2)=4.9919;

x2(27,1)=7.3872; x2(27,2)=5.8584;

x2(28,1)=8.8922; x2(28,2)=5.7748;

x2(29,1)=9.0175; x2(29,2)=6.3059;

x2(30,1)=7.0041; x2(30,2)=6.2315;

x2(31,1)=8.6396; x2(31,2)=5.9586;

x2(32,1)=9.2394; x2(32,2)=3.3455;

x2(33,1)=6.7376; x2(33,2)=4.0096;

x2(34,1)=8.4345; x2(34,2)=5.6852;

x2(35,1)=7.9559; x2(35,2)=4.0251;

x2(36,1)=6.5268; x2(36,2)=4.3933;

x2(37,1)=7.6699; x2(37,2)=5.6868;

x2(38,1)=7.8075; x2(38,2)=5.0200;

x2(39,1)=6.6997; x2(39,2)=6.0638;

x2(40,1)=5.6549; x2(40,2)=3.6590;

x2(41,1)=6.9086; x2(41,2)=5.4795;

x2(42,1)=7.9933; x2(42,2)=3.3660;

x2(43,1)=5.9318; x2(43,2)=3.5573;

x2(44,1)=9.5157; x2(44,2)=5.2938;

x2(45,1)=7.2795; x2(45,2)=4.8596;

x2(46,1)=5.5233; x2(46,2)=3.8697;

x2(47,1)=8.1331; x2(47,2)=4.7075;

x2(48,1)=9.7851; x2(48,2)=4.4175;

x2(49,1)=8.0636; x2(49,2)=4.1037;

x2(50,1)=8.1944; x2(50,2)=5.2486;

x2(51,1)=7.9677; x2(51,2)=3.5103;

x2(52,1)=8.2083; x2(52,2)=5.3135;

x2(53,1)=9.0586; x2(53,2)=2.9749;

x2(54,1)=8.2188; x2(54,2)=5.5290;

x2(55,1)=8.9064; x2(55,2)=5.3435;

**奇异值分解**

一、**实验目的**

1、掌握用奇异值分解的方法求A+ ；

2、学会使用pinv函数求广义矩阵；

3、理解inv函数与pinv函数的不同；

4、利用MATLAB编程求矩阵的奇异值A+；

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

1、利用奇异值分解的方法求A+ 。



2、给出一个具体应用案例，如：推荐系统或图像压缩等。

**K近邻算法**

**一、实验目的**

掌握K近邻算法原理，能够利用编程实现K近邻算法，并能够利用算法解决简单问题。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

1.编程实现对数据集做K近邻算法分类。

数据集名称：鸢尾花卉Iris数据集，详情见iris.txt。对于iris数据集，它包含了150个实例（每个分类包含50个实例），共有sepal length、sepal width、petal length、petal width和class五种属性。期中前四种属性为数值类型，class属性为分类属性。该数据集中的全部实例共可分为三类：Iris Setosa、Iris Versicolour和Iris Virginica。

实验数据集中所有的数据都是实验所需的，因此不存在属性筛选的问题。在txt文件中第一列为分类属性，其他四列为数值类型属性。

2.将Iris数据集每类选取25个样本，总计75个样本作为训练样本，其余75个样本作为测试样本。通过训练样本训练模型，最后对测试样本进行分类。

**EM算法**

**一、实验目的**

掌握EM算法原理，能够利用 编程实现EM算法，并能够利用算法解决简单问题。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

给定数据X，共有三类，每类都成正态分布；X = [randn(100,2)+ones(100,2);randn(100,2)-ones(100,2);randn(100,2)+[ones(100,1),-ones(100,1)]];使用GMM程序对这个样本集分类，用不同颜色标识聚类。最后画出聚类结果图。

**PageRank算法**

**一、实验目的**

掌握PageRank算法原理，能够利用编程实现PageRank算法，并能够利用算法解决简单问题。

**二、实验的硬件、软件平台**

硬件：计算机

软件：操作系统：WINDOWS 10

应用软件：Java、C、C++、MATLAB、Python

**三、实验内容**

1. 创建一个图形，用于说明每个节点如何将其 PageRank 得分赋予图形中的其他节点。效果图如下图所示。



1. 创建并绘制一个有向图，其中包含六个表示虚假网站的节点。计算该图形的 PageRank 中心度得分。使用 0.85 的点进概率（或称为阻尼因子）。查看每个网页的 PageRank 得分和级别信息。

s = [1 1 2 2 3 3 3 4 5];

t = [2 5 3 4 4 5 6 1 1];

names = {'http://www.example.com/alpha', 'http://www.example.com/beta', ...

'http://www.example.com/gamma', 'http://www.example.com/delta', ...

'http://www.example.com/epsilon', 'http://www.example.com/zeta'};

