这是限选课的机器学习，期末只占40分，开卷，强烈建议带一本书(模式分类或西瓜书都行，我个人考试时基本只用到西瓜书)

名词解释

knn

独立同分布

间隔(margin)

激活函数

简答

什么是outlier，为什么k均值对outlier敏感

一道朴素贝叶斯，为应该分到哪类

一道决策树，根据信息增益判断选哪个属性

什么是线性可分，svm怎么解决线性不可分问题，简述思路

综合

什么是训练误差，什么是泛化误差

画图说明误差，指出过拟合和欠拟合区域

梯度下降为什么选择负梯度优化目标函数

怎么选择学习率，过大过小的影响?

综合

验证集的作用，怎么使模型学习到全部数据

adaboost，权重α的推导

试卷共六个大题，基本都在老师考前画的重点内，不过还是有些出乎意料的…可能因为往年开卷的原因，感觉复习资料特别少，简单整理一下供大家参考。

自己整理的复习资料，欢迎下载：添加链接描述

一、 概念题

1、机器学习过程，每个环节主要操作

2、集成学习的概念，BOOSTING和BAGGING 思想

3、交叉验证法

4、特征提取和特征选择的区别

5、隐马尔科夫模型定义形式，以及解决的三个问题？

6、贝叶斯决策过程

二、 svm

1、svm概念,其目的,什么是最优化分类面

2、最优拉格朗日形式

3、核函数为什么要引入函数

4、验证核函数，基本如下，只是x,y换成了x1,x2，映射函数是（x^2 x 1/2）T

三、 神经网络、

描述bp算法

前向传播表达式a1

输出层输出形式

（嗯，又一个原封不动的）

四、 深度学习

卷积层作用

激活函数以及形式

输出

给出下图左1，左二，问你卷积后输出几个，大小为多少，（输出即右一图，这页ppt我根本就没瞅啊，现在就是后悔！！！）

五、 决策树

1.决策树算法思想是什么，两个分类

2.给出一个表格

写出预处理数据集

决策树缺失项 以及依据

决策树分类规则

给数据判断结果

给出的表格决策树基本如下：

六、 DBSCAN算法应用计算（给距离矩阵求聚类结果）

一、综合概念

1.机器学习概念？机器学习过程？

2.泛化能力，过拟合，欠拟合及解决方法

3.贝叶斯决策一般过程

4.二分类：混淆矩阵、查准率、查全率、ROC曲线横纵坐标、AUC

5.Kmeans思想，K的初始化

6.有监督学习/无监督学习的学习机制、代表算法

二、决策树

1.计算题（基于信息增益决策）

2.处理过拟合：剪枝

三、神经网络

1.ML算法基本思想

2.第l层第i单元的激活值；

3.神经网络输出过程

四、SVM

1.支持向量？margin

2.SVM基本型

3.核函数的条件、作用

4.软间隔、硬间隔区别（给定软间隔，写出对应硬间隔）

一、名词解释（每个5分，共20分）

监督学习

KNN

线性可分

多层感知器

二、简答题（每个10分，共60分）

两个一模一样的碗，一号碗有30颗水果糖和10颗巧克力糖，二号碗有水果糖和巧克力糖各20颗。现在随机选择一个碗，从中摸出一颗糖，发现是水果糖。请问这颗水果糖来自一号碗的概率有多大？

简述一下k均值聚类的基本思想，并说明如何初始化k

简述一下SVM的基本思想，针对线性不可分问题，SVM有哪些方法

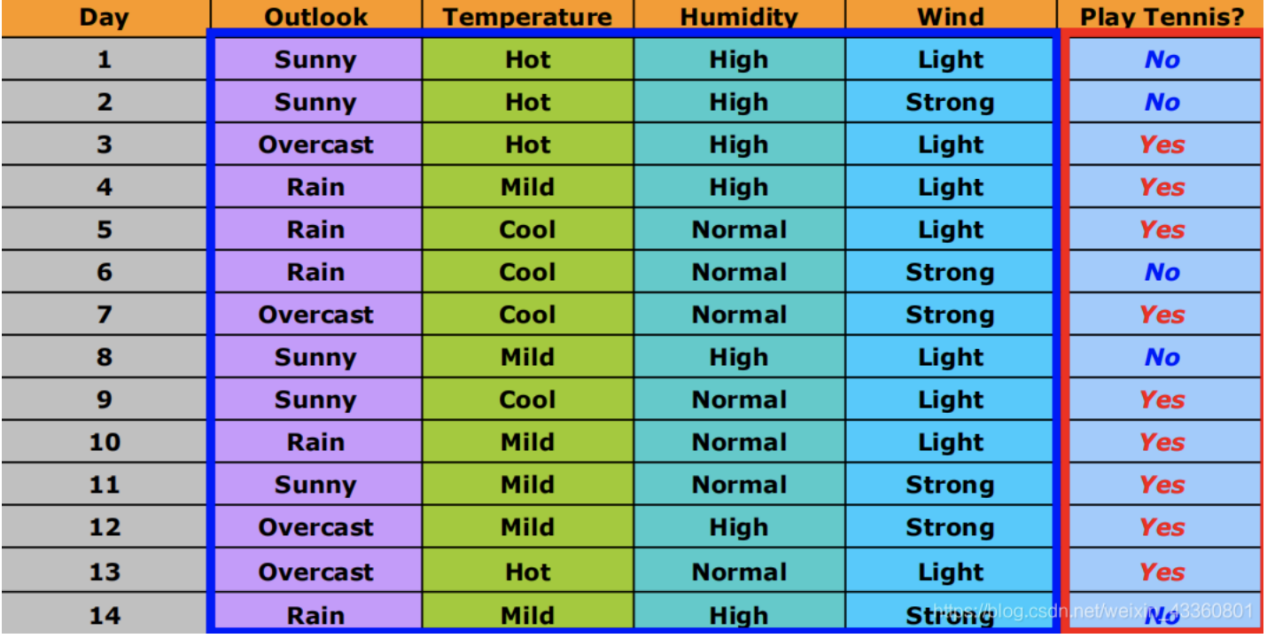
简述一下什么是bagging算法，并比较bagging与AdaBoost的异同

简述一下线性回归，并比较岭回归和lasso回归的区别

从Graph embedding、word embedding、Graph CNN 中选择一个你熟悉的并进行介绍

三、综合分析题（20分）

（个人说明：卷中是10个样本，这里只是一个例子）

（1）从信息增益的角度考虑，Outlook和Humidity哪个作为划分属性，给出具体计算过程

（个人说明：这里给出了可能用到的对数值）

1. 决策树存在过拟合问题，采用什么方法可以减少过拟合风险

一、名词解释（5\*4）

1、机器学习

2、主动学习

3、ID3

4、神经网络

二、简答题（10\*3）

1、简述Parzen窗方法的原理和过程。证明为什么可以用高斯概率密度函数作为窗函数。

2、简要说明梯度下降法和牛顿法的基本思想和区别。解释为什么梯度下降法能够保证一定是下降的？

3、什么是过拟合？为什么会出现过拟合？如何解决过拟合问题？

三、综合题（20+30）

1、证明AdaBoost算法的合理性。

1. 在SVM支持向量机中，简述为什么要使margin最大化？求下列函数的对偶函数（PPT上原公式）。对于非线性问题该如何解决又什么方法吗？

名词解释（5' \* 4）

机器学习

主动学习

ID3

神经网络

简答题（10' \* 3）

parzen窗简述。为什么可以选用高斯密度函数作为窗函数？

梯度下降算法与牛顿法的基本思想和区别。证明为什么梯度下降算法可以保证目标函数下降

什么是过拟合？模型为什么会出现过拟合？如何避免过拟合？

综合分析题

从期望损失角度解释adaboost，如分布和分类器权重更新的依据。（20'）

SVM。（1）从VC维和结构风险角度分析为什么margin要最大化。（2）推导优化函数的对偶形式。（3）简述SVM线性不可分的情况下如何求解(30')

* 线性回归和逻辑回归(梯度下降算法)
* 朴素贝叶斯
* SVM(原始问题和对偶问题)
* K-means算法流程
* PCA算法流程

