机器学习考试重点

1 机器学习导论

> 考试题型

5道判断题（5*2分）
10道单选（10*3分）
5道名词解释（5*3分）
3道简答（3*5分）
目标函数是什么 求导
3道综合（3*10分）
损失函数 如何求导 算法

1.1 绪论

1. **背景 范围 例子**

背景：人工智能，现实世界对智能的需求强烈，无法量化，有不确定性的 潜在规律的。人工智能就是从数据中去挖掘这种规律和确定性。
范围：必须要求潜在的统计规律，必须被数据反映，数据可以隐藏规律但是不可以没有规律。
例子：自动驾驶、大预言模型.....

实际上还是数据依赖，并不具备”智能“

2. **定义（软件、研究）**

CMU机器学习系主任，ICML创始人给出的定义：
从经验数据中学习的一段程序，和任务相关，有评价指标，提升性能指标。

人工智能和机器学习是什么关系?
推理类属于人工智能导论，判断步骤顺序，防止枚举爆炸，减小搜索空间

现在的认为人工智能包括机器学习，机器学习包括深度学习。

现在的定义都是小范围达到共识。

1.2 线性模型（基本概念，易出题）

使用吴恩达老师的课程，从示例入手

1. **线性回归、Logistic回归、线性判别分析**

分类和回归：区别主要在输出上，分类的输出是离散的

Logistics起初做一些非线性的回归分析，不过基于曲线的特性，绝大多数点都分布在01范围，因此二分类问题

Logistics的输出实际不是概率，只是单纯取值在01之间，表示不出来0和1

线性回归和logistic回归的代价函数和梯度的计算

LDA只需知道思路：在分割面的法线的投影上类间距越大，类内间距越小

2. 模型表示、代价函数、优化方法（梯度下降法）

模型表示：用到logistic回归上是很典型的，变换（sigmoid）时为了向01变换，最终的判别还是使用线性模型的判别

优化方法：梯度下降法比直接求解更通用。

3. 机器学习的基本概念（训练、预测、线上线下、独立同分布）

什么是训练，什么是预测，训练为了预测，预测为了预测真实情况
泛化性能、过拟合

训练是耗资源的

线上线下：训练是线下，预测是线上，指立刻。

训练的预测数据必须和预测数据独立同分布

4. 数据、模型、算法

机器学习的灵魂，数据质量很关键。

1.3 神经网络

此前称为ANN（人工神经网络），近些年，前面称为pre-train，后面称为多层感知机。实际上neuron、perception、mlp等等都是一个意思

1. 感知器（XOR）

logistics回归分类器

代价函数有从交叉熵，有均方误差

2. 多层神经网络

BP算法（链式求导法）：隐层参数梯度计算

代价函数：交叉熵以及均方误差（考试时，不用加正则项）

3. 深度学习

语音识别、计算机视觉、自然语言处理

名词解释：什么是卷积神经网络、GPT的缩写、是干什么的

任务：目标检测、跟踪算法

4. 深度学习平台

1.4 支持向量机

1. 代价函数推导
   使用周志华老师教材，$y \in \{1, -1\}$
   吴恩达老师教程中，$y \in \{1, 0\}$
2. 小测的公式
   优化方法：如果一个原始问题解决不了，可以找到其对偶问题解决（只求参数的时候）
   只有alpha大于0的点，才叫支持向量
3. 度量间隔，训练误差

组成，重要元素
核函数

1.5 机器学习模型评估

工程最重要的是量化

1. 指标

混淆矩阵、TP、FP、TN、FN、Precision、Recall、F1、AUC、ROC

2. 方法

交叉验证，充分利用标注数据

1.6 贝叶斯分类

1. 朴素贝叶斯

几个独立：

训练和预测：训练数据集和测试数据集独立同分布

朴素贝叶斯：特征独立

最大似然：训练样本从训练数据中独立抽样出来

2. 生成式、判别式

此处的生成式主要指概率方法

3. 最大似然估计、EM（不考）

1.7 网络机器学习（非重点）

1. RWR（Restrart）

主要思想
随机游走的损失函数推导

2. 周登勇那篇论文

Local（Regulaization）&Global（Training）

3. 图卷积神经网络

利用邻居关系表征节点，深度神经网络的野望

1.8 集成学习

1. AdaBoost

可计算学习理论，弱学习器和强学习器等价

2. Bagging 随机森林
3. Stacking

1.9 PCA

1. 空间、距离（聚类里也有）
2. PCA特征压缩（最大可分性、正交、最大方差）

不考代价函数的优化

1.10 聚类

1. 方法

KNN DBSCAN

2. 性能评价指标的思路

类间越远、类内越近
已知标注、未知标注

1.11 统计理论

可计算学习理论、SVM、偏差方差平衡理论

岭回归（L2范式）、Lassu（L1范式）、最大间隔（SVM）、

1.12 变成基本操作

numpy的基本操作