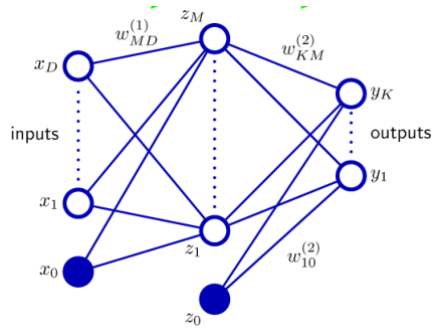


# 2023秋 机器学习导论 期末试题回忆版

共有九道大题, 前八道考察算法应用, 公式推导和模型原理, 第九道是用来捞人的开放题.

---

1. (10分) 题干给出了先验概率和决策损失表
  - (1) 根据最小错误率贝叶斯决策进行分类
  - (2) 根据最小风险贝叶斯决策进行决策(参考PPT例题, 做法一样)
2. (10分) 简述支持向量机 (SVM) 的原理, 遇到噪声和线性不可分时应该如何解决?
3. (15分) 题干给出了一张特征表
  - (1) 使用ID3算法构造一个决策树. 使用构造的决策树判断一个实例的分类结果.
  - (2) 如何解决决策树的过拟合问题? 简述一种解决办法的原理, 以及该方法的缺点.(参考PPT例题, 做法一样)
4. (10分) 题干给出了 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{a}_1, \mathbf{b}_1$ 的值, 且 $\mathbf{x}_1, \mathbf{x}_2 \in A$ ,  $\mathbf{x}_3, \mathbf{x}_4 \in B$   
使用梯度下降法求感知机准则函数 ( $\mathbf{a}, \mathbf{b}$ 的最终值).  
(参考PPT例题, 做法一样)
5. (10分) 简述 PCA 和 LDA (Fisher准则) 的原理和区别.
6. (10分) 简述 K-means 算法原理. K-means 算法有什么缺点, 给出一种改进方法.
7. (10分) 简述集成学习的基本思想, 简述Boosting和Bagging的原理和区别
8. (15分) 如下图所示, 推导神经网络中 $w_{md}$ 和 $w_{km}$ 的更新公式.



(参考PPT例题, 做法一样)

9. (10分) 现如今购物app的推荐算法总是会根据用户的浏览记录, 向用户推荐感兴趣的商品, 使用户陷入"信息茧房"中.

(1) 如何使用机器学习优化产品业务?

(2) 从算法的角度, 简述如何规避"信息茧房".

## 其它往年题

[2021年秋期末](#)

研究生也有一门叫《机器学习》的课, 比本科的《机器学习导论》考的要多和难一些, 下面两个是研究生的往年题.

[2020年秋研究生期末](#)

[2019年秋研究生期末](#)