2020年1月机器学习期末考试回忆

——2017级人工智能实验班

题目类型：判断题26分 简答题 46分（ 9道 ）大题28分 (3道)

老师的话：重点我发现也没什么重点 每一章都有考题

判断题：

VC维，正则化，逻辑回归是回归问题还是分类问题，GBDT,ID3,C4.5提出者，SVM Gauss kernel 中lambda大小 ……13个，记不清

简答题：

1. 逻辑回归损失函数
2. 什么是监督学习，什么是无监督学习，并举例
3. 朴素贝叶斯推导，包括二项分布的解释，极大似然估计
4. 除了square loss 之外的两种loss function.
5. 贝叶斯要做拉普拉斯平滑，为什么要平滑
6. 什么是超参数和模型参数，怎样选择超参数
7. 使用SVD分解，SVD矩阵的一般格式及解释，PCA为什么经常使用SVD分解。
8. ……（ 还有两个记不起来了 ）

大题

1. LDA 计算类间距离矩阵 ，类内距离矩阵，两个类别，各给定三个点。
2. SVM
    1. 普通优化方程，对偶问题方程
    2. C=100 ，优化方程及对偶问题
    3. 课件原题：

> - Suppose we have 5 1-D data points
>   - $x_1 = 1, x_2 = 2, x_3 = 4, x_4 = 5, x_5 = 6$
>   - Labels $y_1 = 1, y_2 = 1, y_3 = -1, y_4 = -1, y_5 = 1$
> - Which kernel do you want to use?

- Suppose we have 5 1-D data points
  - $x_1 = 1, x_2 = 2, x_3 = 4, x_4 = 5, x_5 = 6$
  - Labels $y_1 = 1, y_2 = 1, y_3 = -1, y_4 = -1, y_5 = 1$
- We use the polynomial kernel of degree 2
  - $K(x_i, x_j) = (x_i x_j + 1)^2$
  - $C = 100$ (penalty parameter)

- We first find $\alpha_i (i = 1, \dots, 5)$ by

$$\max_{\alpha} \sum_{i=1}^{5} \alpha_i - \frac{1}{2} \sum_{i=1}^{5} \sum_{j=1}^{5} \alpha_i \alpha_j y_i y_j \underbrace{(x_i x_j + 1)^2}_{K(x_i, x_j)}$$

$$s.t.\ 0 \le \alpha_i \le 100,\ i = 1, \dots, 5$$

$$\sum_{i=1}^{5} \alpha_i y_i = 0$$

- Using a QP solver/SMO, we get
  - $\alpha_1 = 0, \alpha_2 = 2.5, \alpha_3 = 0, \alpha_4 = 7.333, \alpha_5 = 4.833$
  - Note that the constraints are indeed satisfied
  - The support vectors are $\{x_2 = 2, x_4 = 5, x_5 = 6\}$

- The discriminant function: $f = \boldsymbol{w}^T \phi(\boldsymbol{x}) + b = \sum_{i \in S} \alpha_i y_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b$

$$f = 2.5 * 1 * (2x + 1)^2 + 7.333 * (-1) * (5x + 1)^2 + \underbrace{4.833}_{\alpha_5} * \underbrace{1}_{y_5} * \underbrace{(6x + 1)^2}_{K(x, x_5)} + b$$

- b is recovered by solving $f(2) = 1$ or $f(4) = -1$ or by $f(5) = 1$.
  - As x_2, x_4, x_5 lie on the line $\phi(\boldsymbol{w})^T \phi(\boldsymbol{x}) + b = \pm 1$(support vectors).
- All three give $b = 9$

❖ $\boldsymbol{f(x) = 0.6667x^2 - 5.333x + 9}$

3. SGD，BGD

1. 设计经验误差函数
2. SGD，BGD 伪代码，给定样本x1,.....xm，以及区别
3. 添加正则化项L2norm，修正后的伪代码。

（机器学习是考的最难得一门课，哭了）

——2020年1月8日

20:30书