

中国科学技术大学
2018 -- 2019 学年第二学期考试试卷

考试科目: 人工智能原理 得分: _____

学生所在系: _____ 姓名: _____ 学号: _____

要求: (1) 所有答题必须写在试卷上; (2) 每页试卷都必须写上学号和姓名。

题号	1-1	1-2	1-3	2-1	2-2	2-3	2-4	3-1	3-2	总分
得分										

一、基础概念题

1、如何判断“过学习”和“欠学习”、如何处理“过学习”和“欠学习”。(8分)

2、请简述特征选择的作用,什么是过滤式特征选择?什么是包裹式特征选择,什么是嵌入式特征选择?(8分)

3、请简述构建集成分类器的几种方法，并说明集成分类器能够改善分类器性能的原因。Bagging 是否可以提升朴素贝叶斯分类器的性能，请阐明原因。（6分）

二、计算推导题

1、设 $X = \{x_1, x_2, \dots, x_N\}$ 为来自指数分布的样本集，
即： $x \geq 0, p(x, \theta) = \theta \exp(-\theta x); \quad x < 0, p(x, \theta) = 0;$
试求参数 θ 的最大似然估计。（15分）

2、设有一个三层的前馈网络，包括一个含有 d 个节点的输入层、一个含有 n_H 个节点的隐层和一个含有 c 个节点的输出层，层间全连接，层内无连接。

1) 请画出该神经网络。(3 分)

2) 如果该网络训练的目标函数如下；

$$J = \frac{1}{4} \sum_{k=1}^c (y_k - \hat{y}_k)^2$$

请推导输入层和隐层之间的连接权值以及隐层和输出层之间的连接权值的迭代公式。(15 分)

3、设有如下训练数据：

x_1	1	1	0	1	1	0	0	0
x_2	0	1	1	1	0	0	0	0
x_3	1	1	1	0	1	0	0	1
y	1	1	0	0	0	1	1	0

其中 x_1 , x_2 和 x_3 是特征， y 为标签。如果采用朴素贝叶斯分类器，请计算：

$$P(y = 1 | x_1 = 1, x_2 = 1, x_3 = 0)$$

$$P(y = 0 | x_1 = 1, x_2 = 1)。(10 分)$$

4、设有外星人，其体重、眼睛的颜色和眼睛的个数可以用来判断该外星人是勤劳还是懒惰。请根据如下训练数据，构造决策树。（15分）

体重	眼睛的颜色	眼睛的个数	类型
正常	琥珀色	2	懒惰
正常	紫罗兰	2	懒惰
正常	紫罗兰	2	懒惰
较瘦	紫罗兰	3	懒惰
较瘦	紫罗兰	3	懒惰
较瘦	琥珀色	4	勤劳
正常	琥珀色	4	勤劳
正常	紫罗兰	4	勤劳
较瘦	琥珀色	3	勤劳
较瘦	琥珀色	3	勤劳

三、思考题

1、请思考一个新的半监督学习方法，请给出该半监督学习方法的前提假设以及目标函数，并解释目标函数中各项的作用以及如何利用无标签数据的（15 分）。

2、请思考一下本课程的优缺点，并给出建设性的建议（5 分）。