

# 湖南科技大学考试试题纸（ A 卷）

（2020 – 2021 学年度第一学期）

课程名称: 机器学习概论 开课单位: 数学学院 命题教师: \_\_\_\_\_

授课对象: 数学与计算科学 学院 2017 年级 信息与计算科学 1-4 班

考试时量: 100 分钟 考核方式: 考查 考试方式: 开卷

审核人: \_\_\_\_\_ 审核时间: \_\_\_\_\_ 年 \_\_\_\_\_ 月 \_\_\_\_\_ 日

## 一、计算题（本大题共 62 分，第 1 题 12 分，第 2 题 24 分，第 3 题 26 分）

1. 下图 1(a) 是含有两个输入神经元的两层感知机示意图，其激活函数为(b)图所示的阶跃函数：

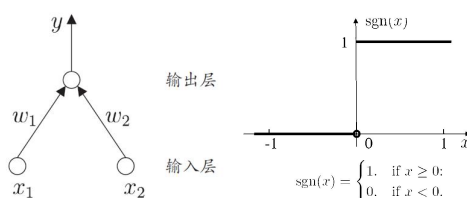


图 1 (a) 感知机网络结构图

(b) 阶跃函数

- ① 令  $w_1=w_2=1$ , 输出层神经元阈值取 0.5, 当  $x_1=0$ 、 $x_2=1$  时, 求输出的  $y$  值; (4 分)
  - ② 令  $w_1=w_2=1$ , 输出层神经元阈值取 2, 当  $x_1=0$ 、 $x_2=1$  时, 求输出的  $y$  值; (4 分)
  - ③ 令  $w_1=0$ ,  $w_2=-0.8$ , 输出层神经元阈值取 -0.5, 当  $x_1=0$ 、 $x_2=1$  时, 求输出的  $y$  值. (4 分)
2. 某瓜农拉来一车西瓜, 共 1000 个 (其中好瓜 950 个, 坏瓜 50 个), 我们用训练好的二分类模型进行判别, 分类结果的混淆矩阵(confusion matrix)如下表 1 所示:

表 1

真实情况	预测结果	
	好瓜	坏瓜
好瓜 (950)	900	50
坏瓜 (50)	20	30

- ① 计算该分类器的错误率、精度; (8 分)
- ② 计算该分类器的查准率、查全率 (保留到小数点后三位); (8 分)
- ③ 计算当前分类状态下的 F1 度量 (保留到小数点后三位); (4 分)
- ④ 给定该二分类任务的代价矩阵 (cost matrix), 求表 1 预测结果的代价敏感错误率。代价矩阵如下表 2 所示. (4 分)

表 2

真实类别	预测类别	
	好瓜	坏瓜
好瓜	0	0.5
坏瓜	1.5	0

3. 根据样本集 D(P76 表 4.1)上的属性“触感”数据，
- ① 写出该属性的所有可能取值 (2 分);
  - ② 写出利用该属性划分后的含编号子集; (6 分)
  - ③ 计算利用该属性划分后得到的子集的信息熵 (保留到小数点后三位); (6 分)
  - ④ 计算该属性的固有值 (intrinsic value) (需要写出详细的计算过程, 只写结果不  
计分, 保留到小数点后三位); (6 分)
  - ⑤ 计算该属性的信息增益率 (保留到小数点后三位)。 (6 分)

## 二、证明题 (本大题共 20 分, 第 1 题 10 分, 第 2 题 10 分)

- 1. 证明: 对于 Sigmoid 函数, 有  $f'(x) = f(x)(1 - f(x))$ .
- 2. 对于图 5.7, 试推导出 BP 算法中的更新公式 (5.13)

## 三、论述题 (本大题共 18 分, 第 1 题 18 分)

- 1. 试述机器学习能在互联网搜索的哪些环节起什么作用.