

# 宁波大学 2018 年博士研究生招生考试初试试题(B 卷)

(答案必须写在考点提供的答题纸上)

科目代码： 3809 科目名称： 智能系统

1. 机器学习通常可分为指导性学习和非指导性学习, 学习模型则可分为生成模型和判别模型二类。假定现在要完成以下的任务:

(a) 函数拟合与插值; (b) 聚类分析; (c) 模式识别,

请简要解释以下学习模型各自属于生成模型还是判别模型, 可以胜任哪种任务:

(1) 深度学习网络, (2)RBF网络, (3)Boltzmann机, (4) SOM网络。

(10%)

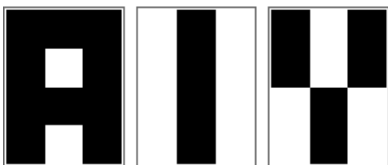
2. 如果将异或 (XOR) 问题推广为 3 输入的 parity 问题, 要求输入和输出实现如下真值表:

输入 1	输入 2	输入 3	输出
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	1

请设计一个多层网络实现 3-parity, 并给出相应的学习算法。 (12%)

3. Hopfield 网络是全相连的反馈型网络, 可以用于实现“吸引子”联想记忆, 如果要使一个 Hopfield 网络同时记住“A”、“I”、“Y”三个字母(3x4 点阵), 试为其设计一个合理的结构, 并给出学习算法和具体的网络权值。

(12%)



# 宁波大学 2018 年博士研究生招生考试初试试题 (B 卷)

(答案必须写在考点提供的答题纸上)

科目代码： 3809 科目名称： 智能系统

4. Adaboost 算法通过适当地组合多个弱分类器（如感知器）形成一个强分类器，它是一种迭代算法，试给出 Adaboost 算法的详细步骤，你认为 Adaboost 算法中最为关键的是哪一步？这些弱分类器必须具备什么条件才能通过该算法形成一个强分类器？Adaboost 算法的优势与弱点各是什么？ (13%)

5. 在 BP 网络中，最常用的激励函数是 Sigmoid 函数， (14%)

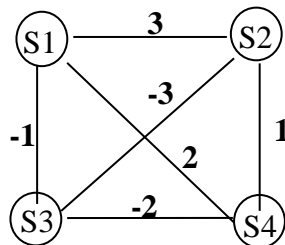
$$\phi_j(v_j(n)) = \frac{1}{1 + \exp(-v_j(n))}, \quad -\infty < v_j(n) < \infty$$

如果将激励函数改为如下的 S 型函数：

$$\varphi_j(v_j(n)) = 1 - \frac{1}{1 + |v_j(n)|}, \quad -\infty < v_j(n) < \infty$$

请问，我们是否还能利用梯度下降法，如果可以，请设计一个相应的 BP 算法。

6. 对于下图所示的 Boltzmann 机，假设每个结点的阈值都为 0，连接权值如图所示，如何采用模拟退火 (Simulated Annealing) 技术，使网络尽快达到系统平衡态？如果达到了平衡态，请计算系统处在状态  $\{S1=-1, S2=-1, S3=1, S4=-1\}$  的概率是多少？要求给出具体的计算方式。 (13%)



7. 深度学习研究近来在生成对抗建模方面取得了突破，生成对抗建模通过一个生成模型与判别模型的竞争学习来逐步完成对输入数据的建模。生成模型 Gen 通过输入随机数  $z$  生成数据，由判别模型 Disc 来判断真假。要求判别模型在输入真实数据时输出 1，在输入由生成模型产生的数据时输出 0，即极大化以下函数：

# 宁波大学 2018 年博士研究生招生考试初试试题 (B 卷)

(答案必须写在考点提供的答题纸上)

科目代码： 3809 科目名称： 智能系统

$$\log D(x^{(i)}) + \log (1 - D(G(z^{(i)})))$$

同时生成模型则以产生出能够“愚弄”判别模型使其以为真的数据为目标，即极小化以下函数：

$$\log (1 - D(G(z^{(i)})))$$

试为这种生成对抗学习建立一个合理的网络结构，并设计合理算法来训练上述生成模型和判别模型。 (14%)

8. 设有海量数据分布在网络上的 100 台计算机中，数据的总量远远超过单台计算机的内存和外存大小，请给出办法高效统计出这批数据的 TOP10（即前 10 个最大数据或前 10 个最小数据）。 (12%)