**神经网络阶段考试**

神经网络、机器学习相关考试

# 选择题

1. 下面哪种结果不是利用机器学习算法从数据中得到的？ [单选题] \*

|  |
| --- |
| 1. 规则 |
| 1. 常识 |
| 1. 回归模型 |
| 1. 神经网络 |

2. 有关神经网络训练时使用的学习率参数说法错误的是？ [单选题] \*

|  |
| --- |
| 1. 学习率可以与其他网络参数一起训练，对降低代价函数是有利的。 |
| 1. 网络训练时刚开始学习率可以大一些，以便提高学习速度，随后应减少学习率，以免引起学习震荡。 |
| 1. 学习率可以随着训练误差动态调整效果更好。 |
| 1. 学习率过大更容易导致训练陷入局部极小值。 |

3. 下面哪个选项不是神经网络训练过程中过拟合的防止方法？ [单选题] \*

|  |
| --- |
| 1. 增加学习率 |
| 1. dropout |
| 1. L2正则化 |
| 1. 提前终止 |

4. 关于训练样本的描述中，正确的说法是哪个？ [单选题] \*

|  |
| --- |
| 1. 样本越少，模型的方差越大 |
| 1. 增加数据可以减少模型方差 |
| 1. 如果模型性能不佳，可增加样本多样性进行优化 |
| 1. 样本越多，模型训练越快，性能越好 |

5. 下面有关神经网络的说法，错误的是？ [单选题] \*

|  |
| --- |
| 1. 神经网络的训练主要是针对神经元之间的权重和神经元的偏置进行一定的调整，使得代价函数极小化 |
| 1. 神经网络不同层次的神经元可以使用不同的激活函数 |
| 1. 神经网络神经元的输出都是传给其他神经元，不能再反馈回来 |
| 1. 均方差损失函数是神经网络常用的一种代价函数（损失函数） |

6. 关BP神经网络的说法，错误的是哪个？ [单选题] \*

|  |
| --- |
| 1. 训练时新样本的加入对已经学习的结果没什么影响 |
| 1. 易陷入局部极小值 |
| 1. 训练次数多，学习效率低，收敛速度慢 |
| 1. 隐层节点的选取缺乏理论指导 |

7. 有关神经网络参数初始化的说法错误的是哪个？ [单选题] \*

|  |
| --- |
| 1. 权重和偏置都可以取全零初始化 |
| 1. 使用正态分布初始化，但权重取太大或太小都会影响网络的训练 |
| 1. Xavier初始化可以减少梯度消失 |
| 1. 合适的权重初始化可以加快神经网络的训练和效果 |

8. 有关BP网络的说法，哪个是错误的？ [单选题] \*

|  |
| --- |
| 1. 与批量梯度下降法相比，使用小批量梯度下降法可以提高训练速度，但达到全局最优解可能需要更多的迭代次数 |
| 1. 交叉熵也可以作为回归预测问题的损失函数 |
| 1. 在使用梯度下降时，加上冲量项会减少训练的速度，但可能会减少陷入局部极小值的可能 |
| 1. 神经元的激活函数选择影响神经网络的训练过程和最终性能 |

9. 梯度消失问题的认识哪个是正确的？ [单选题] \*

|  |
| --- |
| 1. 神经元处于Sigmoid等激活函数的饱和区工作 |
| 1. 隐藏层太多时，可能导致靠近输入层的权重的偏导数太小而得不到更新 |
| 1. 隐层的权重取值太小不容易导致梯度消失 |
| 1. 隐藏层神经元的个数太多导致 |

10. 有关数据预处理对神经网络的影响，以下哪个说法是错误的？ [单选题] \*

|  |
| --- |
| 1. 连续性属性可以采用最小最大值归一化，减少数值过大和不同属性量纲不同对网络的影响 |
| 1. 预处理与否对神经网络的训练速度和分类准确率影响比较大 |
| 1. BP神经网络的输入属性不需要筛选，因为网络本身有特征获取能力 |
| 1. 对于非数值型的属性，可以使用独热（onehot)编码或模糊化转换成数值 |

11. 下面关于单个神经元输入输出正确的是哪个？ [单选题] \*

|  |
| --- |
| 1. 一个神经元可以拥有多个输入但只有一个输出 |
| 1. 一个神经元可以有多个输入和多个输出 |
| 1. 一个神经元只能拥有一个输入和一个输出 |
| 1. 一个神经元只能拥有一个输入但可以有多个输出 |

12. 一个含有2个隐层的BP神经网络，神经元个数都为20，输入和输出节点分别有8和5个节点，这个网络的权重和偏置数分别是多少？ [单选题] \*

|  |
| --- |
| 1. 16000，48 |
| 1. 660，45 |
| 1. 3000，32 |
| 1. 3200，45 |

13. ‎以下不属于超参的是哪个因素？ [单选题] \*

|  |
| --- |
| 1. 输出编码形式 |
| 1. 激活函数 |
| 1. mini-batch的样本数 |
| 1. 学习步长（率）和冲量引子 |

14. 有关神经网络训练过程的说法，正确的是？ \*

|  |
| --- |
| 1. 神经网络权重的初始化大小会对网络的训练结果影响 |
| 1. 分析问题确定后，神经网络合适的结构就可以确定 |
| 1. 对神经网络训练的优化需要综合考虑激活函数、网络结构、权重更新方法等多种因素 |
| 1. 使用增加训练次数的方法不一定可以减少代价函数的取值 |

15. 激活函数通常具有以下哪些性质？ \*

|  |
| --- |
| 1. 非线性 |
| 1. 计算简单 |
| 1. 单调性 |
| 1. 可微性 |

16. 关于模型参数（权重值）的描述，正确的说法是哪些？ \*

|  |
| --- |
| 1. 在训练神经网络过程中，参数不断调整，其调整的依据是基于损失函数不断减少 |
| 1. 每一次Epoch都会对之前的参数进行调整，迭代次数越多，损失函数一般越小 |
| 1. 模型参数量越多越好，但没有固定的对应规则 |
| 1. 训练好的神经网络模型存储于一定结构的神经元之间的权重和神经元的偏置中 |

17. 减少神经网络过拟合的说法，以下哪些是正确的？ \*

|  |
| --- |
| 1. 通过增加数据扰动的数据增强增加了神经网络的过拟合 |
| 1. 通过正则化可以减少网络参数的取值或个数，一定程度可能减少过拟合 |
| 1. 在神经网络训练过程中类似dropout减少神经元或相关链接权的数量 |
| 1. 利用L1或L2正则化可以使权重衰减，从而一定程度上减少过拟合 |

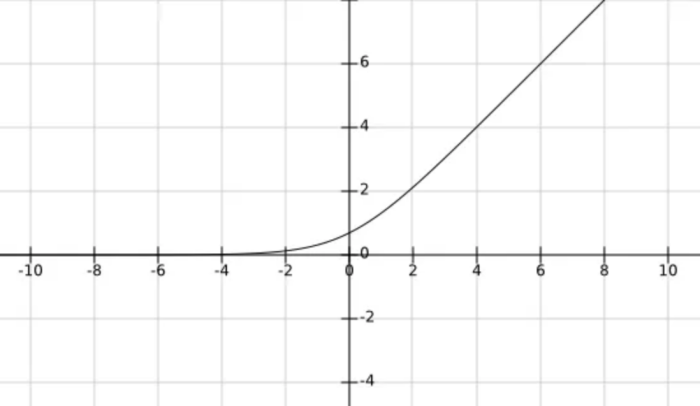
18. 有关数据增强的说法正确是哪些？ \*

|  |
| --- |
| 1. 数据增强带来了噪声，因此一般会降低神经网络模型的效果 |
| 1. 数据增强会增加样本的个数，因此可能会减少过拟合 |
| 1. 对于手写体的识别，对样本的反转、旋转、裁剪、变形和缩放等操作会提高神经网络的识别效果 |
| 1. 对于分类问题，数据增强是减少数据不平衡的一种方法。 |

19. 下列用学习率进行15次梯度下降迭代，每次迭代后计算损失。如果发现损失值下降缓慢，并且在15次迭代后仍在下降。基于此，以下哪个结论最可信？ [单选题] \*

|  |
| --- |
| 1. α = 0.3 是学习率的有效选择 |
| 1. 当前学习率设置偏小 |
| 1. 当前学习率设置偏大 |
| 1. 无法从当前现象对学习率进行判断 |

20. 下图是某一激活函数的图像，下列哪个选项可能是该激活函数的表达式？



[单选题] \*

|  |
| --- |
| 1. e \*\* (-x \*\* 2) |
| 1. ln(1 + e \*\* x) |
| 1. if x ＜ 0 : e\*\*x - 1 ；else：x |
| 1. 1/（1 + e\*\* x） |

# 简答题

1.什么是深度学习网络？

2.如何提升模型的泛化能力？

3. 在训练神经网络时，[损失函数](https://www.zhihu.com/search?q=%E6%8D%9F%E5%A4%B1%E5%87%BD%E6%95%B0&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22article%22%2C%22sourceId%22%3A505597862%7D" \t "_blank)(loss)在最初的几个epochs时没有下降，可能的原因是什么

4. 如果增加多层感知机（Multilayer Perceptron）的隐藏层层数，分类误差便会减小。这种陈述正确还是错误？

5.什么是dropout

6.训练神经网络时，模型损失迟迟不收敛的原因有哪些

7、简单介绍一下你使用的激活函数，以及使用场景

8、KL散度主要是做什么的

9、请解释pytorch中下述程序的含义：optimizer = optim.SGD(model.parameters(), lr=0.002, momentum=0.5)

10、已知输入层特征数量为n,一个隐藏层神经元数量为m，如果输出为l，所有偏置b均为0，求运算的总量

# 设计题

1.试设计一个前馈神经网络来解决 XOR 问题，要求该前馈神经网络具有两个隐藏神经元和一个输出神经元，并使用 ReLU 作为激活函数。



2.P为给定的二维平面整数点集。定义P中某点x，如果x满足P中任意点都不在x的右上方区域内（横纵坐标都大于x），则称为“最大值“，求出所有”最大值“的集合。（所有点的横坐标和纵坐标都不重复，坐标轴范围在（0，1e9）内）

如下图：实心点为满足条件的点的集合，请实现代码找到集合P中的所有“最大值“点的集合并输出。P点形状为（10,10, 2）

