1. 混淆矩阵的召回率（Recall）公式为？（A）
   A. TP/(TP+FN)
   B. FN/(TP+FN)
   C. TN/(TN+FP)
   D. FP/(FP+TN)
2. 在构建决策树时，需要计算每个用来划分数据特征的得分，选择分数最高的特征，以下可以作为得分的是？（D）
   A. 信息熵
   B. 基尼系数（Gini）
   C. 训练误差
   D. 以上都是
3. 下列描述无监督学习错误的是？（C）。
   A. 无标签
   B. 多用于聚类
   C. 多用于分类
   D. 多用于降维
4. 强化学习算法可以分为有模型学习(model-based)算法和免模型(model-free)学习算法，以下属于有模型学习算法的是？（D）。
   A. Policy Gradient
   B. Deep Deterministic Policy Gradient (DDPG)
   C. Deep Q Network (DQN)
   D. AlphaZero
5. 感知器可以解决一下那些问题？（ABCD）
   A. 逻辑“与”
   B. 逻辑“或”
   C. 逻辑“与非”
   D. 逻辑“或非”

6. 如果一个“线性回归”模型完美地拟合了训练样本，也就是训练样本误差为零，则下面哪个说法是正确的？（C）

A. 测试样本误差始终为零

B. 测试样本误差不可能为零

C. 以上答案都不对

7. 在一个线性回归问题中，我们使用 R 平方（R-Squared）来判断拟合度。此时，如果增加一个特征，模型不变，则下面说法正确的是？（C）

A. 如果 R-Squared 增加，则这个特征有意义
B. 如果 R-Squared 减小，则这个特征没有意义
C. 仅看 R-Squared 单一变量，无法确定这个特征是否有意义。
D. 以上说法都不对

8. 下列哪些假设是我们推导线性回归参数时遵循的（多选）？

A. X 与 Y 有线性关系（多项式关系）

B. 模型误差在统计学上是独立的

C. 误差一般服从 0 均值和固定标准差的正态分布

D. X 是非随机且测量没有误差的

答案：ABCD

9. 一般来说，下列哪种方法常用来预测连续输出变量？

A. 线性回归

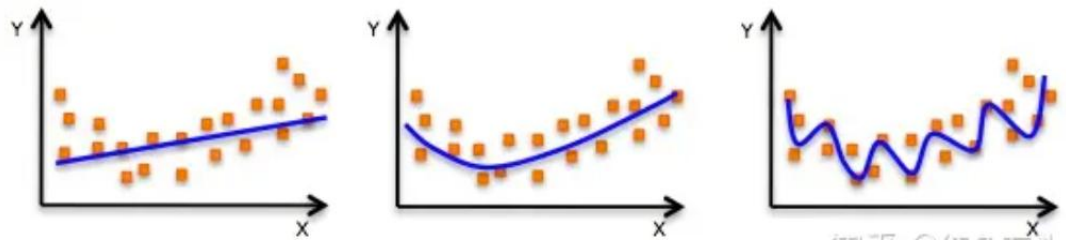
B. 逻辑回顾

C. 线性回归和逻辑回归都行

D. 以上说法都不对

答案：A

10. 下面三张图展示了对同一训练样本，使用不同的模型拟合的效果（蓝色曲线）。那么，我们可以得出哪些结论（多选）？



A. 第 1 个模型的训练误差大于第 2 个、第 3 个模型

B. 最好的模型是第 3 个，因为它的训练误差最小

C. 第 2 个模型最为“健壮”，因为它对未知样本的拟合效果最好

D. 第 3 个模型发生了过拟合

E. 所有模型的表现都一样，因为我们并没有看到测试数据

答案：ACD

11. 两个变量相关，它们的相关系数 r 可能为 0。这句话是否正确？

A. 正确

B. 错误

答案：A

12. 逻辑回归将输出概率限定在 [0,1] 之间。下列哪个函数起到这样的作用？

A. Sigmoid 函数

B. tanh 函数

C. ReLU 函数

D. Leaky ReLU 函数

答案：A

13. 关于 k 折交叉验证，下列说法正确的是？

A. k 值并不是越大越好，k 值过大，会降低运算速度

B. 选择更大的 k 值，会让偏差更小，因为 k 值越大，训练集越接近整个训练样本

C. 选择合适的 k 值，能减小验方差

D. 以上说法都正确

答案： D

14. 我们知道二元分类的输出是概率值。一般设定输出概率大于或等于 0.5，则预测为正类；若输出概率小于 0.5，则预测为负类。那么，如果将阈值 0.5 提高，例如 0.6，大于或等于 0.6 的才预测为正类。则准确率（Precision）和召回率（Recall）会发生什么变化（多选）？

A. 准确率（Precision）增加或者不变

B. 准确率（Precision）减小

C. 召回率（Recall）减小或者不变

D. 召回率（Recall）增大

答案：AC

15. 关于神经网络，下列说法正确的是？

A. 增加网络层数，可能会增加测试集分类错误率

B. 增加网络层数，一定会增加训练集分类错误率

C. 减少网络层数，可能会减少测试集分类错误率

D. 减少网络层数，一定会减少训练集分类错误率

答案：AC

16. 下面哪句话是正确的？

A. 机器学习模型的精准度越高，则模型的性能越好

B. 增加模型的复杂度，总能减小测试样本误差

C. 增加模型的复杂度，总能减小训练样本误差

D. 以上说法都不对

答案：C

17. 如果一个经过训练的机器学习模型在测试集上达到 100% 的准确率，这是否意味着该模型将在另外一个新的测试集上也能得到 100% 的准确率呢？

A. 是的，因为这个模型泛化能力已经很好了，可以应用于任何数据

B. 不行，因为还有一些模型不确定的东西，例如噪声

答案：B

18. 下面有关分类算法的准确率，召回率，F1 值的描述，错误的是？

A.准确率是检索出相关文档数与检索出的文档总数的比率，衡量的是检索系统的查准率

B.召回率是指检索出的相关文档数和文档库中所有的相关文档数的比率，衡量的是检索系统的查全率

C.正确率、召回率和 F 值取值都在 0 和 1 之间，数值越接近 0，查准率或查全率就越高

D.为了解决准确率和召回率冲突问题，引入了 F1 分数

答案：C

19. “增加卷积核的尺寸，一定能提高卷积神经网络的性能。” 这句话是否正确？

A. 正确

B. 错误

答案：B

20. 假如现在有个神经网络，激活函数是 ReLU，若使用线性激活函数代替 ReLU，那么该神经网络还能表征异或（XNOR）函数吗？

A. 可以

B. 不可以

答案：B

21. 下列哪种方法可以用来减小过拟合？（多选）

A. 更多的训练数据

B. L1 正则化

C. L2 正则化

D. 减小模型的复杂度

答案：ABCD

22. 下列说法错误的是？

A. 当目标函数是凸函数时，梯度下降算法的解一般就是全局最优解

B. 进行 PCA 降维时，需要计算协方差矩阵

C. 沿负梯度的方向一定是最优的方向

D. 利用拉格朗日函数能解带约束的优化问题

答案：C

23. 关于 L1、L2 正则化下列说法正确的是？

A. L2 正则化能防止过拟合，提升模型的泛化能力，但 L1 做不到这点

B. L2 正则化技术又称为 Lasso Regularization

C. L1 正则化得到的解更加稀疏

D. L2 正则化得到的解更加稀疏

答案：C

24. 假定你在神经网络中的隐藏层中使用激活函数 X。在特定神经元给定任意输入，你会得到输出 -0.01。X 可能是以下哪一个激活函数？

A. ReLU

B. tanh

C. Sigmoid

D. 以上都有可能

答案：B

25. 以下哪些方法不可以直接来对文本分类？

A. K-Means

B. 决策树

C. 支持向量机

D. kNN

答案：A

（1） 什么是机器学习模型的过拟合和欠拟合？导致模型过拟合的原因有哪些？请结合决策树和神经网络模型进一步阐述解决模型过拟合的方法。

参考答案：
过拟合：模型在训练集上错误率很低，但是在未知数据上错误率很高。
欠拟合：模型不能很好地拟合训练数据，在训练集和测试集上的错误率都比较高。
过拟合的原因：过拟合问题往往是由于训练数据少和噪声以及模型复杂度过高等原因造成的。
解决过拟合的方法：（1）数据层面：增加训练数据、清除数据噪声；（2）模型层面：在经验风险最小化的基础上引入参数的正则化、模型训练提前迭代终止远侧、模型剪枝原则等。例如，决策树模型可以通过先剪枝操作来控制决策树的生长或通过后剪枝操作对决策树进行修剪。神经网络模型可以通过加入 L1 正则化或者 L2 正则化、Dropout、early stopping 等
（2）关于机器学习算法评估，分别阐述适用于回归算法和分类算法的评估指标有哪些？并阐述各种评估指标的优缺点。

参考答案：
回归算法评估指标：

a) 平均绝对误差（Mean Absolute Error）
b) 均方误差（Mean Squared Error）
c) 均方根误差（Root Mean Squared Error）
d) 决定系数（Coefficient of determination）

分类算法评估指标：

a) 精度 Accuracy
b) 混淆矩阵 Confusion Matrix
c) 准确率（查准率） Precision
d) 召回率（查全率）Recall
e) Fβ Score
f) AUC Area Under Curve
g) KS Kolmogorov-Smirnov

回归算法评估指标的优缺点举例：

a) MAE 虽能较好衡量回归模型的好坏，但是绝对值的存在导致函数不光滑，在某些点上不能求导，可以考虑将绝对值改为残差的平方；
b) MSE 与目标变量的量纲不一致；
c) RMSE 可以保证量纲一致性；
d) 以上基于误差的均值对进行评估的指标，均值对异常点（outliers）较敏感，如果样本中有一些异常值出现，会对以上指标的值有较大影响。

分类算法评估指标的优缺点举例：

a) 对于有倾向性的问题，往往不能用精度指标来衡量。
b) 对于样本类别数量严重不均衡的情况，也不能用精度指标来衡量。
c) AUC 是一种模型分类指标，且仅仅是二分类模型的评价指标。
d) AUC 对样本类别是否均衡并不敏感；