1.Keras是一个怎样的深度学习框架（ A ）

A.基于TensorFlow的高级API B.基于PyTorch的高级API

C.一种独立的深度学习框架 D.一个集成了多种深度学习框架的工具包

2.在一个神经网络里，知道每一个神经元的权重和偏差是最重要的一步。如果以某种方法知道了神经元准确的权重和偏差，就可以近似任何函数。实现这个最佳的方法是（ C ）

A.随机赋值，期望它们是正确的

B.搜索所有权重和偏差的组合，直到得到最佳值

C.赋予一个初始值，通过检查跟最佳值的差值，然后迭代更新权重和偏差

D.以上都不正确

3.将训练数据进行分割，设置批块大小的好处有（ D ）

A.有利于并行计算 B.有利于内存装载

C.配合小批量随机梯度算法的需要 D.以上均是

4.以下关于代码和参数的解释，不正确的是（ B ）

dataset = dataset.shuffle(1000).batch(32).repeat(20)

A.shuffle方法实现对数据的随机打散，避免模型陷入过拟合

B.batch方法设置批块大小，每一批块的样本数量一定是32个 可能小于

C.repeat能实现多个轮次的训练

D.调用dataset提供的工具函数会返回新的dataset对象，会覆盖旧的dataset对象

5.下列哪个神经网络结构会发生权重共享（ D ）

A.卷积神经网络 B.循环神经网络 C.连接神经网络 D.选项A和B

6.批规范化（BN）是有帮助的，因为（A ）

A.在将所有的输入传递到下一层之前对其进行归一化

B.它将返回归一化的权重平均值和标准差

C.它是一种非常有效的反向传播技术

D.它能增加模型的复杂度

7.在神经网络中，以下那些技术可以用于解决过拟合问题（ D ）

A.Dropout B.正则化 C.批规范化 D.以上均是

8.对于图像识别问题（比如识别照片中的猫），哪种神经网络模型结构更适合解决这类问题（ B ）

A.多层感知器 B.卷积神经网络 C.循环神经网络 D.感知器

9.小明在训练深度学习模型时，发现训练集误差不断减少，测试集误差不断增大，以下解决方法正确的是（ D ）

A.数据增广 B.提前停止训练 C.添加Dropout D.以上均是

10.以下不是keras中的一部分的是（ D ）

A.keras.datasets B.keras.models C.keras.layers D.keras.matplotlib

11.输入图片大小为200×200，依次经过一层卷积（kernel size 5×5，padding 1，stride 2），pooling（kernel size 3×3，padding 0，stride 1），又一层卷积（kernel size 3×3，padding 1，stride 1）之后，输出特征图大小为（ C ）

A.95 B.96 C.97 D.98

12.关于深度学习的描述正确的是（B ）

A.深度学习需要输入人工提取的特征

B.深度学习模型可用来分类

C.目前基于深度学习的图像分类准确率低于人类

D.深度学习功能可与人脑媲美

13.分析以下代码片段，MaxPooling2D 层的作用是：( C )

from tensorflow.keras.layers import MaxPooling2D

from tensorflow.keras.models import Sequential

model = Sequential()

model.add(MaxPooling2D(pool\_size=(2, 2), strides=(2, 2), padding='valid'))

A.对输入特征图进行平均池化，池化窗口大小为2x2，步长为2，不进行填充  
B.对输入特征图进行最大池化，池化窗口大小为2x2，步长为2，进行填充以保持输出尺寸与输入相同  
C.对输入特征图进行最大池化，池化窗口大小为2x2，步长为2，不进行填充  
D.对输入特征图进行卷积操作，卷积核大小为2x2，步长为2，不进行填充

14.下列关于卷积神经网络，说法错误的是（ D ）

A.卷积神经网络可以有多个卷积核，可以不同大小

B.卷积层的核心作用是通过滑动窗口提取局部特征

C.卷积核权重共享机制使得网络参数数量减少并提升泛化能力

D.最大池化（Max Pooling）能防止梯度消失问题

15.关于常用的激活函数及其特性，描述错误的是（ D ）

A.sigmoid函数可以将实数映射到（0，1）范围内，容易产生梯度消失现象

B.tanh函数可以将实数映射到（-1，1）范围内，和sigmoid函数存在线性关系

C.relu函数在零点处不可导，可以实现单侧抑制，缓解梯度消失问题

D.激活函数的目标是为了提高模型训练的速度 一定程度

16.为了适配模型，单个测试样本通常需要增加一个维度，对于样本A的操作正确的是（ A ）

A.tf.expand\_dims(A,axis= -1)

B.tf.reshape(A,axis= -1)

C.tf.squeeze(A,axis= -1)

D.tf.stack(A,axis= -1)

17.梯度下降算法的目标是？( B )

A.最大化损失函数

B.最小化损失函数

C.使损失函数保持不变

D.随机改变损失函数的值

18.下面这段代码构建的卷积神经网络中，最后一层Dense层通常用于什么任务：(A)

from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense

from tensorflow.keras.models import Sequential

model = Sequential()

model.add(Conv2D(filters=32, kernel\_size=(3, 3), activation='relu', input\_shape=(28, 28, 1)))

model.add(MaxPooling2D(pool\_size=(2, 2)))

model.add(Flatten())

model.add(Dense(units=10, activation='softmax'))

A. 图像的特征提取

B. 图像的降维

C. 图像的分类，输出10个类别的概率分布

D. 图像的上采样，恢复图像分辨率

19.当使用一个 2x2 的池化窗口，步长为 2 对一个4x4的特征图进行最大池化操作时，输出特征图的尺寸是？ (A)

A.2x2

B.3x3

C.4x4

D.1x1

20.以下代码中 Flatten 层的功能是： (C)

from tensorflow.keras.layers import Flatten

from tensorflow.keras.models import Sequential

model = Sequential()

model.add(Flatten())

A.将输入的多维特征图进行卷积操作，使其变为一维向量

B.将输入的多维特征图进行池化操作，减少特征图的尺寸

C.将输入的多维特征图展平成一维向量，以便后续连接全连接层

D.对输入的一维向量进行归一化处理

1.在回归任务中，均方误差（MSE）越小表示模型预测越准确（ 对 ）

2.模型的参数量越多，其性能一定越好（ 错 ）

3.在监督学习任务中，输出层中的神经元的数量要和类的数量（其中类的数量大于2）保持一致（ 对 ）

4.如果模型训练时使用的学习率太大，可能会导致网络将无法收敛 （ 对 ）

5.卷积神经网络具有稀疏连接、参数共享和平移不变性三个特性（ 对 ）

6.早停（Early Stopping）是一种正则化技术，可防止模型在训练过程中过拟合（ 对 ）

7.tf.linalg.matmul()函数可以实现矩阵乘法操作，等价于符号@（ 对 ）

8.数据增强（如旋转、缩放图像）会增加训练数据的多样性，缓解过拟合（ 对）

9.已知input = [0.01,0.01,0.02,0.087,0.02,0.02,0.02,0.01,0.01,0.01],

则tf.math.argmax(input)的结果为4（ 错 ）

10.在随机梯度下降算法中，每次更新参数使用的数据是一个批块的样本（ 错 ）

1.以下神经网络有 3 层，有 2 层隐藏层，输入层有 3 个神经元，输出层有 1 个神经元，属于 全连接 神经网络（神经网络类型）。（每空2分，共10分）

