1. (简答题)

梯度学习的改进优化算法有哪些？

◼ 动态学习率

➢ Adagrad

➢ RMSProp

◼ 处理梯度为0的情况

➢ Momentum

◼ 综合方法

➢ Adam (RMSProp+Momentum)

1. (简答题)卷积神经网络中，为什么可以使用ReLU激活函数？

relu激活函数的优势：

（1）反向传播时，可以避免梯度消失。

（2）Relu会使一部分神经元的输出为0，这样就造成了网络的稀疏性，并且减少了参数的相互依存关系，缓解了过拟合问题的发生（以及一些人的生物解释balabala）。

（3）相对于sigmoid激活函数，tanh激活函数，求导简单。采用sigmoid等函数，算激活函数时（指数运算），计算量大，反向传播求误差梯度时，求导涉及除法，计算量相对大，而采用Relu激活函数，整个过程的计算量节省很多。

缺点：

（1）左侧神经元为0，导致神经元死亡，不再更新。

（2）输出非负，仍然存在zigzag现象。

1. (简答题)RBF Network与常用的Neural Network相比，隐藏层（Hidden Layer）有什么区别？

- \*\*Inner-product tanh\*\*：使用内积和tanh激活函数的隐藏层，具有全局特性，适合经典神经网络。

- \*\*Distance Gaussian\*\*：使用距离和高斯激活函数的隐藏层，具有局部特性，适合径向基函数网络（RBFN）。

1. (简答题)自组织特征映射神经网络中权重W存储了什么信息？

SOM的权重存储了竞争层每个节点的特征

**SVM和SVR分别用于什么任务**

支持向量机（Support Vector Machine, SVM）主要用于分类任务。它的目标是找到一个超平面，将不同类别的数据点分隔开来，并最大化两个类别之间的间隔（margin）。SVM 在解决线性和非线性分类问题时都表现出色，通过使用核函数（kernel）将数据映射到高维空间，从而实现复杂的分类。

支持向量回归（Support Vector Regression, SVR）主要用于回归任务。与SVM类似，SVR的目标是找到一个超平面，但不同的是，SVR的目标是拟合一个函数，使得它能够预测连续的目标变量。SVR试图在一定误差范围内（ε-不敏感区间）最小化预测误差，并最大化预测函数的间隔。

**为什么会产生过拟合？**

按照统计学习理论的解释，过拟合的产生是由于待拟合/学习的函数类的复杂性太高，该函数类的表征能力太强，通过参数的调整可以找到函数类中的某个“完美”函数。这个完美函数恰好穿过全部的训练数据点，以至于使得训练误差为0。但是对于未知的样本，这个所谓的完美函数却误差很大。因此，产生了过拟合。

[判断题] 给定一个欧式空间和一个核函数K，则只能得到唯一的高维特征空间H以及从欧式空间到特征空间H的映射f，满足K(x,y) = f(x)·f(y),其中·为内积操作

【答案：错】

1. [单选题] RBF神经网络中隐藏层节点数

A. 小于样本数

B. 等于样本数

C. 大于样本数

D. 都有可能

【答案：B】

2. [单选题] 广义RBF神经网络中隐藏层节点数

A. 远远小于样本数

B. 等于样本数

C. 远远大于样本数

D. 都有可能

【答案：A】

1. [判断题] 对于线性可分的数据，感知器方法得到的分类超平面是不惟一的。

正确答案：对

2. [判断题] 对于线性可分的数据，线性可分支持向量机得到的分类超平面是惟一的。

正确答案：对

1. [单选题] PCA是

A. 有监督学习

B. 无监督学习

正确答案：B

2. [单选题] 使用PCA方法对某高维数据W进行降维处理，已知数据W在正交的u1,u2,u3三个方向上的投影数据的方差分别为30，25，51。请问应该先将数据W投影到哪个方向

A. u1

B. u2

C. u3

D. 哪个都可以

正确答案：C

**变分自编码器和自编码器有什么优势**

生成能力：VAE 能生成新数据样本，而 AE 主要重构输入数据。

潜在空间：VAE 的潜在空间连续、平滑，生成数据更一致。

概率解释：VAE 具有概率解释，利用贝叶斯推断和变分推断。

正则化效果：VAE 天然有正则化效果，避免过拟合。

特征表示：VAE 能捕捉数据的复杂分布，学习到更有用的特征表示。

推导局部加权线性回归参数学习公式

推导逻辑回归学习公式

在多分类问题中，one vs rest需要训练 K 个分类器，one vs one 需要训练 K(K-1)/2 个分类器。