

2.1 Machine Learning 简介

2.2 本单元小测1

2.3 本单元小测2

3 Neural Network

3.1 Neural Network课件

3.2 SEARCHING FOR ACTIVATION FUNCTIONS

3.3 本单元小测

4 About Training

4.1 About Training

4.2 本单元小测

5 Deep NN

5.1 Deep NN

5.2 拓展资料1: On the Origin of Deep Learning

5.3 拓展资料2: Efficient Processing of Deep Neural ...

5.4 拓展资料3: Attention is all you need

5.5 参考: Tensorflow和Pytorch环境安装

5.6 本单元小测

6 Problem Solving by Searching

章节测验

About Training章节小测

题量: 8 满分: 60.0

第1次作答

一. 简答题 (共3题)

1【简答题】请谈谈你对机器学习中的偏差(bias)和方差(variance)的认识。(15.0分)

我的答案:

在机器学习和统计学中, 偏差和方差是指两种不同类型的误差, 可以帮助解释模型泛化能力。

偏差是指模型的预测值与真实值之间的差, 他衡量了模型在不同训练数据集值之间的平均误差。高偏差的模型可能会错过真实数据内在的规律, 即使在训练集上表现很好, 也可能无法很好地预测新的未见数据。

方差是指模型在不同训练数据集上预测值的变化程度, 它衡量了模型对训练数据中的小波动产生的变化。高方差的模型可能在不同训练数据集上表现出对训练数据的过度拟合。

正确答案:

要点:

学习算法的泛化误差可以分解为三个部分: 偏差(bias), 方差(variance) 和噪声(noise)。我们主要关注偏差与方差。

偏差: 预测值与真实值之间的差别。偏差通常来自于对算法模型的错误假设, 在训练集上表现就不好。通常的解决方法: 增加更多的特征, 或使用更强的模型。

方差: 预测值的差异浮动, 反映模型的稳定性。在训练集上表现很好, 但在测试集上表现不好。解决方法: 特征降维; 增加数据; 降低模型复杂度; 加入正则化, early stop等。

2【简答题】举例说明机器学习中Hyperparameter和Parameter的区别。(10.0分)

我的答案:

在支持向量机SVM算法中, 选择核函数类型(如线性核、多项式核或高斯核)和核系数, 需要在训练之前设置。而学习到的分离超平面的系数和截距即为模型的参数。这些参数是在训练过程中学习到的, 并用于进行预测。

正确答案:

要点:

Hyperparameter: 预设的参数, 定义关于模型更高层次的概念(模型复杂度、学习率等)。

Parameter: 通过训练学习到的模型参数, 如神经网络的连接权重、偏置等。

3【简答题】试简述L1正则化和L2正则化的异同。(10.0分)

我的答案:

他们都是为了控制模型的复杂度并防止过拟合, 都是在损失函数中添加一个正则化项用于控制正则化对模型的影响程度。

而主要区别在于正则化的计算方式和效果。L1正则化通过加上参数权重的绝对值之和来控制模型的复杂度, L1正则化项为参数权重的绝对值之和。适用于具有大量特征的数据集, 稀疏性并进行特征选择。L2正则化通过加上参数权重的平方和来惩罚模型的复杂度, 防止过拟合, 平滑地调整权重, 并适用于缺乏足够数据的情况。

正确答案:

要点:

相同之处: 可用于缓解模型的过拟合; 都会使不重要的参数趋向于0。
不同: L1通过不断地加(减)一个值使参数往0的方向变化, 而L2则是通过不断地平方使参数往0的方向变化。

2.1 Machine Learning 简介

2.2 本单元小测1

2.3 本单元小测2

3 Neural Network

3.1 Neural Network课件

3.2 SEARCHING FOR ACTIVATION FUNCTIONS

3.3 本单元小测

4 About Training

4.1 About Training

4.2 本单元小测

5 Deep NN

5.1 Deep NN

5.2 拓展资料1: On the Origin of Deep Learning

5.3 拓展资料2: Efficient Processing of Deep Neural ...

5.4 拓展资料3: Attention is all you need

5.5 参考: Tensorflow和Pytorch环境安装

5.6 本单元小测

6 Problem Solving by Searching

二. 单选题 (共3题)

4【单选题】一般应该通过 () 来选择模型的超参数。(5.0分)

- A、 训练数据集
- B、 验证数据集
- C、 测试数据集
- D、 整个数据集

我的答案: B

正确答案: B

5【单选题】在测试阶段, batch normalization中所用的均值和方差来自于 (

- A、 训练过程中得到的均值和方差
- B、 多个测试数据
- C、 一个测试数据
- D、 /

我的答案: A

正确答案: A

6【单选题】若某个模型训练过程中使用0.3的dropout rate, 得到的参数为W, 则 (5.0分)

- A、 W
- B、 0.3W
- C、 0.5W
- D、 0.7W

我的答案: B

正确答案: D

三. 多选题 (共2题)

7【多选题】关于dropout的描述正确的是 () (5.0分)

- A、 每更新一次参数时, 需要对神经元重新进行dropout
- B、 对每一个mini-batch, 都需要对神经元重新进行dropout
- C、 模型训练过程中, 只需要对神经元进行一次dropout
- D、 使用dropout技术时, 测试时直接使用训练过程所得的模型参数。

我的答案: AB

正确答案: AB

8【多选题】Batch Normalization过程中的参数 () 是通过学习得到的。(5.0分)

- A、 μ

- 2.1 Machine Learning
- 2.2 本单元小测1
- 2.3 本单元小测2
- 3 Neural Network
- 3.1 Neural Network课件
- 3.2 SEARCHING FOR ACTIVATION FUNCTIONS
- 3.3 本单元小测
- 4 About Training
- 4.1 About Training
- 4.2 本单元小测
- 5 Deep NN
- 5.1 Deep NN
- 5.2 拓展资料1: On the Origin of Deep Learning
- 5.3 拓展资料2: Efficient Processing of Deep Neural ...
- 5.4 拓展资料3: Attention is all you need
- 5.5 参考: Tensorflow和Pytorch环境安装
- 5.6 本单元小测
- 6 Problem Solving for Searching

C、 γ

D、 β

我的答案: CD

正确答案: CD

上一节

下一节

