

ML_循环神经网络

填空题

1. RNN的基本结构包括三个主要组件，其中_____接收输入序列，_____负责记忆和传递信息，_____产生最终输出。
2. _____可能导致数值溢出，_____可能导致长时依赖问题。应对策略可能包括梯度截断、使用LSTM或GRU等。
3. 在LSTM中，_____用于控制是否遗忘之前的记忆细胞的信息。
4. RNN的“隐藏状态”在网络中的_____层。
5. 描述LSTM中的三个门包括遗忘门、输入门和输出门。_____决定遗忘多少之前的记忆，_____决定存储多少新的信息，_____决定记忆细胞的哪部分被输出。

选择题

1. 为什么在某些任务中选择使用循环神经网络（RNN）而不是其他神经网络结构？（选择题）
 - A. 处理序列数据
 - B. 具有记忆能力
 - C. 适用于时序关系建模
 - D. 所有选项都正确
2. LSTM和GRU是RNN的两种常见变体，它们的主要区别是什么？（选择题）
 - A. LSTM有三个门，而GRU只有两个门
 - B. LSTM具有记忆细胞和输入门
 - C. GRU相对于LSTM参数更少
 - D. 所有选项都正确
3. 在训练中，如果梯度爆炸发生，可能会导致什么问题？（选择题）
 - A. 数值溢出
 - B. 模型参数不稳定
 - C. 训练无法收敛
 - D. 所有选项都正确
4. 对于长序列的处理，LSTM和GRU相对于传统RNN具有哪些优势？（选择题）
 - A. 更好地捕捉长时依赖关系
 - B. 减缓梯度消失问题
 - C. 更适用于时间序列建模
 - D. 所有选项都正确

判断题

1. RNN中的“沿时间反向传播”用于解决梯度传播中的长时依赖问题。

2. 在RNN训练中，正向传播用于计算模型的预测输出，反向传播用于计算梯度以更新模型参数。

简答题

1. 当训练长序列时，梯度爆炸和梯度消失问题可能变得更加显著。请简要解释为什么会出现这些问题。
2. 在选择LSTM或GRU时，你会在什么情况下更倾向于选择其中之一？

答案

填空题

1. 输入层；隐藏层；输出层；
2. 梯度爆炸；梯度消失；
3. 遗忘门
4. 隐藏
5. 遗忘门；输入门；输出门

选择题

1. D. 所有选项都正确
2. D. 所有选项都正确
3. D. 所有选项都正确
4. D. 所有选项都正确

判断题

1. 对
2. 对

简答题

1. 答案：在长序列中，梯度在反向传播时可能多次相乘，导致指数级增长或减小，从而引发梯度爆炸或梯度消失。
2. 答案：倾向选择LSTM当需要更复杂的建模能力时，倾向选择GRU当计算成本和参数数量更为重要时。