

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。