

1. 请简述 BERT 的基本原理和结构，并说明它与传统语言模型（如 N-gram 模型）相比有哪些优势。
2. 请介绍 Top-K, Top-P 和 Temperature 等。
3. 现在让你解决一个新闻分类任务，类别有 6 个类别（体育、科技、娱乐、商业、健康、政治），请给出你的解决方案。
4. 调用大语言模型的时候，发现性能不理想。请问有哪些策略可以提升大模型的性能。