

2、训练时还随机地将输入句子的部分token替换为[MASK]，训练目标是输入一个带噪声的句子，输出原句子的一个相似句