## 商品情感项目：

**项目要求：**

使用深度学习算法，实现输入一段文本，输出文本中的物品位置和物品对应的情感分类（好评/差评）

**项目步骤：**

首先对数据进行处理，训练样本数据中的数据都是竖着成列排列，第一列是文本，第二列是BIO标记，第三列是情感。首先将列转为一行一个样本；定义了一些工具函数，例如按BIO的标记来读取出实体位置，得到cdm和cdw的函数；接着定义dataset类，为了扩充样本将相邻的两句话进行拼接，拼接后按手动方法进行bert词嵌入表示（因为本数据集比较特殊，是纵列排序，若出现英文，则可能每个字母占一个位置，有一个情感和BIO，若用bert自带的编码则会出现一个unk，则和BIO标记无法对应）。进行批量读取，对一批句子中的每一句话都随机的选取一个实体和这个实体对应的情感，保存到batch\_pairs变量中。

工作中出现的问题和思考：

1、bert在冻结权重后效果确实会变差，微调后效果会好。并且训练会快。

2、数据比训练的网络更重要。

3、