

● 新题目

课题13：基于评论特征的“爆款”言论预测

数据获取：通过网络爬虫等技术手段获取小米SU7的热门微博评论数据。

数据网址： https://github.com/fyuo863/weibo_content_SU7

要求：

- 1. 将评论的“点赞数”或“回复数”作为标签，构建预测模型。可以定义为回归问题（预测具体数值）或分类问题（预测是否为“高赞热门评论”）。
- 2. **特征工程：**

文本特征：评论长度、情感强度、是否包含特定关键词（如“雷军”、“保时捷”）、标点符号使用（如感叹号数量）。

时间特征：发布时间段（凌晨/高峰期）、距离原博发布的时间差。

- 3. 分析哪个特征对评论成为“爆款”贡献最大（例如：是骂得越狠赞越多？还是字数越多赞越多？）。

参考分析手段：

- 自定

论文撰写：完成论文撰写，并提交数据、程序、以及论文，完成PPT答辩。