**比赛简介**

随着互联网的快速发展，在线新闻服务已成为主要的网络信息载体。越来越多的民众通过在线新闻了解时事动态，以新闻评论的形式表达观点、抒发情感。了解新闻评论中的用户情感，掌握网民态度和情绪变化，可辅助在线新闻服务商理解用户的偏好，为用户提供更好的个性化服务；同时，有助于政府进行民意问询和公共管理决策，有效实现舆情分析和监控。本次新闻评论情感极性分析比赛邀请参赛者设计算法和模型，对大量用户的新闻评论所属的情感极性进行预测，组委会将通过参赛者提交的预测结果与真实标注作比较，并结合答辩表现，评估参赛者的算法模型。

**数据介绍**

数据集分为训练集和测试集两个部分，训练集包括4.8万条用户新闻评论和已标注的情感极性标签，测试集包括1.2万没有标注情感极性标签的用户新闻评论，共涉及3种情感极性：positive（正向）、negative（负向）和neutral（中性）。

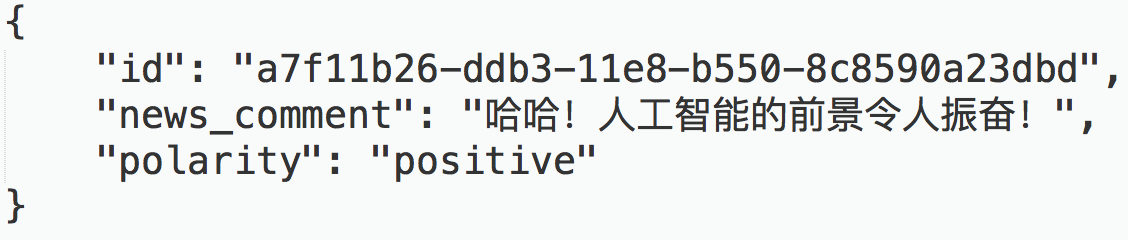
数据利用json格式储存，每一行为一条数据，每条数据均为一个字典。

id: 新闻评论id

news\_comment: 新闻评论

polarity: 情感极性

这里是简单的一条数据展示:



**评价标准**

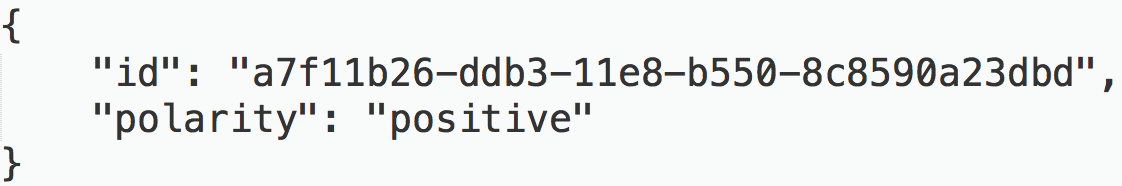
本次比赛将采用分类任务中的微平均F1值（Micro-F1）作为评价指标。比赛的所有数据将分为两个阶段发放给参赛选手，每个阶段以所有参赛队伍的历史最优成绩进行排名。

总排名以0.3 \* 阶段一成绩 + 0.7 \* 阶段二成绩，按从高到低排序

**结果提交说明**

选手返回的结果应存为JSON文件，提交结果应包含新闻评论的id与情感极性，每一行为一条数据，每条数据均为一个字典。

格式



结果在提交时，应以附件的形式附在邮件中，并且该附件命名为prediction.json，参赛队伍在提交结果时，邮件标题需采用“组别+队名+队长学号+队长姓名”的格式。（如：专业组+菜鸡起飞+20191818181818+张三）。如不按格式提交，不予测评。