

项目一：“标准问”分类

项目背景

实现一个简单的问答系统的一种方法是：收到用户提出的问题后，找到用户提问在已知回答的标准集中对应的标准问，再利用标准问及其对应的回答答复用户。

你的任务是：给定一个提问，例如：“预约取钱的有效时间是多久”，预测它的标准问是什么，例如，这个问题对应的标准问是：“预约取款的有效期限”。（测试数据中，所有可能的候选标准问是已知的，即都在训练数据中出现）

项目参与方式

具体的数据集及项目介绍，已经发布在Kaggle竞赛平台上，你需要从以下链接进入该项目，**注册账号并加入竞赛**。（必须从下面的邀请链接中进入）

项目邀请链接：<https://www.kaggle.com/t/5155976b21e94378b136763c8b2d15c9>

利用train.csv文件中的数据训练你的模型，然后对test.csv文件中的数据进行预测，以submission.csv格式上传到平台。然后该平台会自动对你提交的结果进行评分，并显示出你当前的排名。

务必将Kaggle排行榜的名字，改为 SUFE_你的名字_你的学号，否则视为无效提交。

评分标准

1. 有效提交结果：我们仅接受正确率不低于0.6的提交。若正确率 < 0.6 ，视为无效提交。
2. 成绩排名得分：若你的总排名是 a （从0开始计算），共有 b 位同学，那么你的得分为 $\text{基础分值} * (b - a) / b$ 。其中，基础分值待定。
3. 早鸟加分：如果你在 2023/12/11 23:59 分前提交了有效的结果（正确率高于0.6），不论排名，课程总成绩+1分。