



全部赛题 / 金融大脑-金融智能NLP服务 进行中 参赛人数 2,632 总奖金(RMB) 600,000 举办方: 

举办方：



报名

2018-04-18



第一赛季

2018-04-21



第二赛季

2018-07-1



4 决赛

2018-09-01

已报名

赛题信息

赛题简介

赛题与数据

常见问题

排行榜

我的信息

我的队伍

提交结果

我的成绩

第二赛季

第一赛季

文件名称	文件格式	操作
atec_nlp_sim_train	csv	下载
atec_nlp_sim_train_add	csv	下载

<

1

>

1、赛题任务描述

问题相似度计算，即给定客服里用户描述的两句话，用算法来判断是否表示了相同的语义。

示例：

1. “花呗如何还款”--“花呗怎么还款”：同义问句
2. “花呗如何还款”--“我怎么还我的花呗呢”：同义问句
3. “花呗分期后逾期了如何还款”--“花呗分期后逾期了哪里还款”：非同义问句

对于例子a，比较简单的方法就可以判定同义；对于例子b，包含了错别字、同义词、词序变换等问题，两个句子乍一看并不类似，想正确判断比较有挑战；对于例子c，两句话很类似，仅仅有一处细微的差别“如何”和“哪里”，就导致语义不一致。

2、数据

本次大赛所有数据均来自蚂蚁金服金融大脑的实际应用场景，赛制分初赛和复赛两个阶段：

初赛阶段

我们提供10万对的标注数据（分批次更新，已更新完毕），作为训练数据，包括同义对和不同义对，可下载。数据集中每一行就是一条样例。格式如下：

行号\t句1\t句2\t标注，举例：1 花呗如何还款 花呗怎么还款 1

- 行号指当前问题对在训练集中的第几行；
- 句1和句2分别表示问题句对的两个句子；
- 标注指当前问题对的同义或不同义标注，同义为1，不同义为0。

评测数据集总共1万条。为保证大赛的公平公正、避免恶意的刷榜行为，该数据集不公开。大家通过提交评测代码和模型的方法完成预测、获取相应的排名。格式如下：

行号\t句1\t句2

初赛阶段，评测数据集会在评测系统一个特定的路径下面，由官方的平台系统调用选手提交的评测工具执行。

复赛阶段

我们将训练数据集的量级会增加到海量。该阶段的数据不提供下载，会以数据表的形式在蚂蚁金服的数巢平台上供选手使用。和初赛阶段类似，数据集包含四个字段，分别是行号、句1、句2和标注。

评测数据集还是1万条，同样以数据表的形式在数巢平台上。该数据集包含三个字段，分别是行号、句1、句2。

3、评测及评估指标

初赛阶段，比赛选手在本地完成模型的训练调优，将评测代码和模型打包后，提交官方测评系统完成预测和排名更新。测评系统为标准Linux环境，内存8G，CPU4核，无网络访问权限。安装有python 2.7、java 8、tensorflow 1.5、jieba 0.39、pytorch 0.4.0、keras 2.1.6、gensim 3.4.0、pandas 0.22.0、sklearn 0.19.1、xgboost 0.71、lightgbm 2.1.1。提交压缩包解压后，主目录下需包含脚本文件run.sh，该脚本以评测文件作为输入，评测结果作为输出（输

出结果只有0和1)，输出文件每行格式为“行号\t预测结果”，命令超时时间为30分钟，执行命令如下：

```
bash run.sh INPUT_PATH OUTPUT_PATH
```

预测结果为空或总行数不对，评测结果直接判为0。

复赛阶段，选手的模型训练、调优和预测都是在蚂蚁金服的机器学习平台上完成，后台定时运行选手保存的模型。评测以问题对的两句话作为输入，相似度预测结果（0或1）作为输出，同样输出为空则终止评估，评测结果为0。

本赛题评分以F1-score为准，得分相同时，参照accuracy排序。选手预测结果和真实标签进行比对，几个数值的定义先明确一下：

True Positive（TP）意思表示做出同义的判定，而且判定是正确的，TP的数值表示正确的同义判定的个数；

同理，False Positive（FP）数值表示错误的同义判定的个数；

依此，True Negative（TN）数值表示正确的不同义判定个数；

False Negative（FN）数值表示错误的不同义判定个数。

基于此，我们就可以计算出准确率（precision rate）、召回率（recall rate）和accuracy、F1-score：

$$\text{precision rate} = TP / (TP + FP)$$

$$\text{recall rate} = TP / (TP + FN)$$

$$\text{accuracy} = (TP + TN) / (TP + FP + TN + FN)$$

$$\text{F1-score} = 2 * \text{precision rate} * \text{recall rate} / (\text{precision rate} + \text{recall rate})$$

平台合作

支付宝

蚂蚁金融云

数巢平台

机器学习平台

数据集成开发平台

相关协议

平台协议

联系方式

大赛咨询邮箱：

atec_dc@service.alipay.com



蚂蚁金服技术支持



大赛钉钉答疑群



金融科技号

沪ICP备15027489号