**任务：CCKS2018 -微众银行智能客服问句匹配大赛**

# 任务描述

微众银行智能客服问句匹配大赛是由微众银行提供语料支持，哈尔滨工业大学（深圳）智能计算研究中心负责组织实施的真实场景语句意图匹配任务。

语句匹配是自然语言处理的最基本任务之一，是自动问答、聊天机器人、信息检索、机器翻译等各种自然语言处理任务基础。语句匹配问题的复杂性在于，匹配的要求不同，对匹配的定义也不尽相同，比如经典的语句复述判别问题，需要判断两句话是否仅仅是表述方式不同，但意义相同，而在Quora问句匹配语料发布后，大量在该语料库上开展的语句匹配研究工作都沿袭语料发布者的定义，称为语义等价判别(semantic equivalence identification)，和复述判别相比，语义等价判别主要从语句（主要是问句）所蕴含的意图来判断两个语句是否等价，而不直接判断两个语句是否表达相同的语义，所以其核心是语句的意图匹配。由于来源于真实问答语料库，该任务更加接近于智能客服等自然语言处理任务的实际需求。

与基于Quora的语义等价判别相同，本次评测任务的主要目标是针对中文的真实客服语料，进行问句意图匹配。集给定两个语句，要求判定两者意图是否相同或者相近。所有语料来自原始的银行领域智能客服日志，并经过了筛选和人工的意图匹配标注。

**输入：**一个语句对

**输出：**表明该语句对是否表达相同或者相似意图的二值标签（0或1）

**示例：**

*样例1*

输入：一般几天能通过审核\t一般审核通过要多久

输出：1

*样例2*

输入：一般会在什么时候来电话\t一直在等待电话通知

输出：0

# 数据描述

本次数据是主要来自金融领域的真实文本，训练集、验证集及测试集的说明如下：

**训练集&验证集：**

在训练及验证数据发布阶段，我们会发布10w对左右的标注训练问句对数据集和1w对左右的验证问句对数据集。

**测试集：**

在测试数据发布阶段，我们将会再发布1w对左右的问句对数据集，不含标注结果，作为测试。

# 评价指标

本次任务采用精确率（Precision, P）、召回率（Recall, R）、F1值（F1-measure, F1）和准确率(Accuracy, Acc)作为评价指标。

$$语义匹配精确率=正例标注正确的语句对数/标注为正例的语句对总数$$

$$语义匹配召回率=正例标注正确的语句对数/正例语句对总数$$

$$语义匹配F1值=\frac{语义匹配准确率\*语义匹配召回率}{语义匹配准确率+语义匹配召回率}$$

$$语义匹配准确率=标注正确的语句对数/语句对总数$$

# 任务提交

* 1. **评测平台**

本次评测将依托biendata平台（https://biendata.com/）展开，请有意向的参赛队伍关注平台上的竞赛列表。

* 1. **评测形式**

本次任务将采取刷榜的方式，验证集发布后，允许参赛队伍多次向平台提交结果，文件命名为“参赛队名称\_valid\_result.txt”，格式与结果文件result.txt相同，每周一更新一次排名。参赛队伍可在评测集发布之前随时上传验证集的计算结果（每日最多可上传1次），管理系统会及时更新各队伍的最新排名情况；

测试集发布后，允许参赛队伍多次提交结果文件，结果文件提交格式如下。

**结果文件格式说明：**

选手将结果保存为result.txt，以utf-8编码格式保存，每行3列，以“\t”分隔，第一列为句子1，第二列为句子2，第三列为语义匹配结果0或1，即“句子1\t句子2\t匹配结果”。

最终提交文件要求：

每一个参赛队需提交的材料如下：

1. 语义匹配运行结果文件
2. 代码及说明
3. 方法描述文档（非评测论文，评测论文撰写要求见CCKS 2018官网）

以上三个文件需在任务提交截止日期前按照评测网站的要求提交结果。

结果文件用result.txt命名，采用utf-8格式存储，文件格式需要与训练数据中的标注结果文件完全一样(每行格式为“句子1\t句子2\t类标”)。

代码及其文档需打包成一个文件（tar，zip，gzip，rar等均可），用code.xxx命名，要求提交所有的程序代码及相关的配置说明，确保程序能够正确运行，且所得结果与result.txt相符。

方法描述文档用SM.pdf命名，包含算法描述及参数设置，需用pdf格式存储，页数不超过5页。

# 任务交流平台

本任务的交流平台为：

QQ群【CCKS2018评测\_Task3】：700736096

所有报名参赛的参赛队至少要有一名成员加入到该群中，验证消息请发送参赛队伍名称，后续所有有关本任务评测及数据的相关说明和通知也将会在讨论组中发布和交流，请所有参赛队务必在报名成功之后加入该群。

# 时间安排

评测任务发布：4月1日

报名时间：4月1日—7月15日

训练及验证数据发布：4月20日

测试数据发布：7月15日

提交测试结果：7月20日

评测论文提交：8月5日

# 规则

1. 所有参赛队伍都必须提交报名信息（参赛队名，队长姓名，成员姓名，联系邮箱，手机号码，参赛单位名称）至**chenjing@stu.hit.edu.cn**；

2. 每支队伍需指定一名队长，队伍名称不超过15个字符；

3. 每名选手只能参加一支队伍，一旦发现某选手以注册多个账号的方式参加多支队伍，将取消所有相关队伍的参赛资格；

4. 允许使用开源代码或工具，但不允许使用任何未公开发布或需要授权的代码或工具；

5. 本次比赛允许参赛队伍只用主办方提供的数据集，或者在主办方提供的数据集之外增加Quora数据集，如果引入了Quora数据集，需要在提交结果的时候加以明确说明。不允许使用除主办方提供的数据集和Quora数据集之外的任何外部数据。

6. 欢迎国内外在校生及社会在职人士参加。比赛组织方成员不可参赛。

**温馨提示：**如果选手没有在收件箱收到相应的通知邮件，请在垃圾邮箱里查找

# 组织者

哈尔滨工业大学（深圳），陈清财

哈尔滨工业大学（深圳），汤步洲

哈尔滨工业大学（深圳），陈静

深圳前海微众银行股份有限公司，杨海军

深圳前海微众银行股份有限公司，姜华

关于本评测有任何问题，请联系：**chenjing@stu.hit.edu.cn**