**中移在线首届人工智能大赛**

**—面向客服领域的用户意图识别**

**一、背景简介**

如今，移动互联网使生活发生了翻天覆地的变化，数字化服务的加速启动将我们的生活带入一个实时分享、永远在线的全新时代。传统的服务方式，已难满足客户对互联网的依赖和对“以指代步”的期待。中国移动服务8亿客户的模式，也正因此迎来前所未有的重大转变。曾经，我们等待客户；如今，客户在哪里，我们的服务就要出现在哪里。在这样的背景下，中移在线服务有限公司正式成立了。在线公司的服务宗旨是：

**1）移动全网集中服务的提供者**

通过互联网、电话、社交媒体、APP等服务界面与客户建立互动沟通关系，为客户构建线上服务体系

**2）移动全网渠道运营的集中支撑者**

为社会代理渠道、连锁渠道、电商渠道及虚拟运营商等提供一点对接服务

**3）移动全网业务的后台集中处理者**

为实名制验证、开户等业务提供集中审核与办理等服务

从2017年开始，全国计算语言学学术会议（CCL）计划举办评测活动。作为全国计算语言学学术会议(CCL)的系列评测，18年在线公司与CCL合作举办第一届“中移在线杯--客服领域用户意图识别评测” ，并将于2018年下半年与第十七届全国计算语言学学术会议（CCL2018，2018年10月19日~21日，湖南长沙）共同召开。本届客服领域用户意图识别评测由中国中文信息学会计算语言学专业委员会（CIPS-CL）主办，在线公司承办并冠名赞助，旨在进一步促进客服领域研究及发展并且为相关领域学者提供一个良好的沟通平台。在此，评测委员会诚邀各界人士参加本次评测活动！

**二、任务描述**

在线公司管理移动全网10086热线、10086APP互联网客服入口，拥有海量的客户来话录音、在线交互记录，但一直无法准确识别客户诉求及来电偏好，无法将海量录音的数据价值应用于热线运营、产品服务中。利用非结构化录音、转写文本的结果进行用户来电原因分析，明晰用户来电诉求，提高服务质量。

针对本次竞赛，分为两个任务：来电原因多分类、用户意图识别。

**1）来电原因分类**

在客服和用户对话的过程中，用户可能会有多种来电意图，相应地会触发多个业务分类，例如投诉（抱怨），办理、查询等。本任务要求参赛队伍基于给定的来电原因类别表，对用户与客服的对话录音转写数据完成用户来电原因分类。

**2）用户意图识别**

随着新技术的发展，在线公司积极探索新的服务形态，打造一个基于智能语音、自然语言理解等先进技术的多轮对话服务机器人。在对话中，用户通过与对话机器人的多轮交互，提出所需的信息和服务，通过多次陈述完善需求，在一次完整通话中，会设计到多个业务场景和业务节点。因此首先要抽取出用户的意图。本任务要求参赛队伍完成单句级别用户意图识别与槽位抽取。

**三、重要日期**

报名：

5月1日-6月15日：发布评测方案，报名（于6月15号零点报名截止）。

6月1日-8月31日：初赛

9月1日-10月1日：复赛

10月8日-10月21日：确认评测排名，揭晓获奖名单。技术评测论坛（颁奖+技术报告）

**四、比赛赛程**

此次评测分为初赛与复赛两个阶段，

**1）初赛：**

6月1日-6月7日：发布初赛训练数据。（训练数据集获取方式会邮件通知各参赛队伍）。

6月18日-8月27日：18日发布初赛TestA测试集，以周为时间单位，每个参赛队每周最多可提交3次结果发送至zyzx\_ccl\_2018@163.com邮箱，以最优成绩为准。下一周的周一上午10:00会将上周的排名、分数发送给个参赛队伍。

8月27日-8月31日：27日发布初赛TestB测试集，以天为单位，每个参赛队每天可提交1次结果发送至zyzx\_ccl\_2018@163.com邮箱，以最优成绩为准。第二天的上午10:00会将前一天的排名、分数发送给个参赛队伍。

8月31号24点初赛TestB测试集结果提交截止，以各参赛队提交TestB的最终结果为准进行排名，取前50名参加复赛！

**2）复赛:**

9月3日-9月24日：3日发布复赛训练集，复赛TestA测试集，以周为时间单位，每个参赛队每周最多可提交5次结果发送至zyzx\_ccl\_2018@163.com邮箱，以最优成绩为准。下一周的周一上午10:00会将上周的排名、分数发送给个参赛队伍。

9月24日-9月30日：24日发布复赛TestB测试集，以天为单位，每个参赛队每天可提交2次结果发送至zyzx\_ccl\_2018@163.com邮箱，以最优成绩为准。第二天的上午10:00会将前一天的排名、分数发送给个参赛队伍。

9月30号24点复赛TestB测试集结果提交截止，以各参赛队提交TestB的最终结果为准确定名次，每支队伍需在提交最终结果截止时间之前，提交评测集的最终结果，包括打包源码、算法运行、说明文档两部分内容。其中，打包源码要求是可运行的，需提供环境依赖和使用说明。

举办方将组织人力对参赛队伍提交的代码进行复验，确认评测排名，揭晓获奖名单。

**3）提交物**

针对两个测试任务，每次提交需要选手提交2个结果文件，文件须为UTF-8编码(无BOM)的文件。文件命名格式为：参赛队伍队长手机号\_classify\_result\_时间戳；参赛队伍队长手机号\_intent\_result\_时间戳。邮件命名方式为：队长手机号\_时间戳\_队名。

邮件命名样例：

188888888888\_1527755190\_XXX

提交结果文件命名样例：

188888888888\_classify\_result\_1527755178

188888888888\_intent\_result\_1527755191

**(1)来电原因分类提交文件格式**

输出文件格式：

ID + \t + 一级标签 + \t + 二级标签

ID + \t + 一级标签 + \t + 二级标签

测试输出结果样例：

20170714\_b18//170791050139.V3 查询 套餐流量

20170708\_b3//91960759252.V3 取消 套餐

**(2)用户意图识别提交文件格式**

输出文本格式：

ID+\t+意图识别json字符串

ID+\t+意图识别json字符串

测试结果输出文件样例格式:

1500912014-1574618 {"sentence": "月初扣费都扣了哪些", "intents": [{"action": {"value": "查询"}, "target": {"value": "月初扣费"}}], "slots": []}

1500911992-722173 {"sentence": "用银行卡为本机交话费", "intents": [{"action": {"value": "办理"}, "target": {"value": "手机充值"}}], "slots": [{"key": "附属标签", "value": "银行卡"}]}

**4）评价指标**

**（1）来电原因分类**

文本分类评测最终使用F值作为评价指标，指标定义如下：

准确率（P）＝正确预测的文本标签数量/ 预测的文本标签总量  
 召回率（R）＝正确预测的文本标签数量/ 测试数据中文本标签总量

，其中=1.2

**（2）意图识别分值计算**

意图识别分值计算时同时考虑意图识别和槽位识别，一个通话中用户意图及槽值识别均正确得1分，其中意图识别和槽值识别分别占50%的权重，如果对话中包含多个槽值信息，所有槽值信息均识别正确得0.5分，其余情况不得分。最终使用总分值作为评价指标，评价指标定义如下：

Iscore=(0.5\*意图识别正确通话数+0.5\*槽值识别正确通话数)/通话总数

**测试最终得分为文本分类与意图识别的加权得分，两者权重分别为0.5。计算公式如下：**

finalScore = 0.5 \* F + 0.5 \* Iscore

**(F值、Iscore精确到3位小数、4舍5入)**

**5）讨论区**

为方便答疑，欢迎参赛队伍加入微信去群,举办方技术人员将不定期进行答疑。



**五、奖励**

一等奖1名：20000元

二等奖2名：每名10000元

三等奖3名：每名5000元

优秀奖5名：每名3000元

一、二、三等获奖队伍将被邀请CCL-2018大会举办的技术评测论坛上进行口头报告。

获奖队伍可获得由中移在线颁发的成绩证书。

**六、注意事项**

1. 每支队伍需指定一名队长，每支队伍最多不超过5名队员；

2. 每名选手只能参加一支队伍，一旦发现某选手参加多支队伍，将取消所有相关队伍的参赛资格；

**3. 除主办方提供的数据集，不得使用任何外部未公开发布或者需要授权的数据；**

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.进入**复赛的队伍需要在9月30号前将**可执行程序、代码、代码说明及最优结果生成文档（包括模型和特征的说明以及本周内每次提升所作改进的说明），上传到主办方指定的Linux服务器，并在小规模样例上测试通过，经组委会评估后即确认获奖资格；如拒绝或未在规定时间内提供相应文档则取消获奖资格，奖品顺延；

6. 主办方自支付奖金之时起，拥有数据竞赛获奖参赛队伍所提交的作品的使用权；

7. 欢迎国内外在校生及社会在职人士参加。中移在线员工（含正式员工、派遣员工、实习生）不可参赛；

8. 如特殊情况需要，比赛主办方拥有更新比赛时间线的权利.

**附录：**

# 任务1：来电原因分类数据格式说明

**（1）训练数据格式说明：**

文件ID\t一级标签+\t+二级标签

角色标识A\t转写内容，分词的文本

角色标识B\t转写内容，分词的文本

角色标识A\t转写内容，分词的文本

角色标识B\t转写内容，分词的文本

。。。

文件ID\t一级标签+\t+二级标签

角色标识A\t转写内容，分词的文本

角色标识B\t转写内容，分词的文本

。。。

格式说明：

（a）角色A、角色B代表不同的说话人，一般情况为0为客服、1为客户。受限于语音引擎能力，转写精度与说话人识别精度均在88%左右（全文转写准确率87%左右，说话人角色识别85%左右）。

（b）本测试为单标签测试，每通对话一个分类标签。

**（2）训练数据样例格式:**

20170714\_b18//170791050139.V3 查询 套餐流量

0 您好很高兴为您服务

1 唉你好你能帮我查一下我手机流量套餐是多少然后帮我取消了吧

0 呃您现在这个手机的话用的是三十八元套餐除了三十八元之外的话您有一个手机上网的五十元包月和一个来电提醒这三个业务是收费的

1 噢等于说是八十八元套餐是吧

0 对八十八加两块总共是九十

1 行那个那个你能帮我看一下就是呃我一个月那个全国的流量是多少吧

0 好我看一下三十八套餐里面全国流量是五百兆然后五十块钱里面包含的全国流量是一个g一千五百兆

1 噢好谢谢啊

0 不客气还有别的需要帮助的吗

1 两

0 嗯好的

20170708\_b3//91960759252.V3 查询 套餐流量

0 您好很高兴为您服务

1 这个好的好为什么

0 唉你好我我

1 我移动公司了

0 嗯是的请问什么需要帮您

1 呃你能帮我看一下就是呃我一个月那个全国的流量是多少吧

**（3）未标注数据及测试数据格式说明：**

文件ID

角色标识A\t转写内容

角色标识B\t转写内容

角色标识A\t转写内容

角色标识B\t转写内容

。。。

文件ID

角色标识A\t转写内容

角色标识B\t转写内容

。。。

**（4）测试结果输出文件格式说明：**

**提交结果文件格式为：**

ID + \t + 一级标签 + \t + 二级标签

ID + \t + 一级标签 + \t + 二级标签

**提交结果样例：**

20170714\_b18//170791050139.V3 查询 套餐流量

20170708\_b3//91960759252.V3 取消 套餐

**（5）来电原因类别列表与说明**

**来电原因类别体系与说明详见如下文件**

****

# 任务2：用户意图识别数据格式说明

**（1）训练数据格式说明：**

意图识别标注数据中包含用户意图和槽位信息，每个用户通话数据保存为一个json格式字符串。格式的设定参考了国际通用的DSTC的标准。（详见：<https://www.microsoft.com/en-us/research/event/dialog-state-tracking-challenge/> 和 <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/Dialog20state20tracking20challenge20handbook20V21.pdf> ）

训练数据格式为：

ID+\t+json格式字符串

ID+\t+json格式字符串

其中json字符串中包括：

（a）sentence：用户通话内容。

（b）intents：用户意图信息，每个用户意图由操作（action）和目标（target）两部分构成。

（c）slots：槽值信息，一个通话可能包含多个槽值信息，每个槽值信息包括槽类型、槽值、槽值在文本中的起始、结束位置。

**(2)训练数据样例格式:**

1500912014-1574618 {"sentence": "我想办理流量业务", "intents": [{"action": {"value": "开通"}, "target": {"value": "流量"}}], "slots": []}

1500911992-722173 {"sentence": "融合宽带资费介绍", "intents": [{"action": {"value": "咨询"}, "target": {"value": "宽带"}}], "slots": [{"key": "宽带类型", "value": "4G融合宽带"}]}

**（3）测试数据格式说明：**

**测试数据格式说明：**

ID+\t+客户录音文本

ID+\t+客户录音文本

其中客户录音文本为未分词文本数据。

**测试数据样例文件格式：**

1500912014-1574618 我要查询1月份账单

1500911992-722173 我要查询1月份话费

**（4）测试结果输出文件格式说明：**

**提交结果文件格式与训练文本格式保持一致：**

ID+\t+意图识别json字符串

ID+\t+意图识别json字符串

**提交结果样例格式:**

1500912014-1574618 {"sentence": "月初扣费都扣了哪些", "intents": [{"action": {"value": "查询"}, "target": {"value": "月初扣费"}}], "slots": []}

1500911992-722173 {"sentence": "用银行卡为本机交话费", "intents": [{"action": {"value": "办理"}, "target": {"value": "手机充值"}}], "slots": [{"key": "附属标签", "value": "银行卡"}]}

**（5）意图识别范围说明：**

意图识别范围中包含训练数据和测试数据中可能的用户意图和槽位信息的取值范围，相关内容可参阅以下2个文件。

****