直播预告 | 2020语言与智能技术竞赛冠军团队与你畅聊比赛经验

语言是人类信息传递最重要的媒介，让机器理解语言并进行交互是人工智能的基本挑战。为了推动语言理解与交互技术发展，中国中文信息学会、中国计算机学会和百度公司联合举办“2020语言与智能技术竞赛”。**比赛历时3个月，共吸引了来自海内外知名企业、高校及科研机构的5307支队伍报名，收到近1.5万份结果提交，五大任务的系统效果较基线提升平均约90%。**

在精彩纷呈的角逐中，云知声智能科技股份有限公司、国防科技大学、深圳追一科技有限公司、京东方智能物联首席技术官组织、深思考人工智能机器人科技（北京）有限公司和大连理工大学脱颖而出，获得了比赛的冠军。作为冠军队伍，他们的方法和思路，对于相关方向的研究者们有着借鉴意义。

**11月7日，五大任务冠军团队将亲临直播间做评测报告，为大家带来比赛经验的分享！**机器阅读理解、面向推荐的对话、语义解析、关系抽取和事件抽取五个核心任务全面覆盖，干货满满，精彩不容错过！

**直播时间**：11月7日（周六）10:00-12:00

**10:00-10:25 机器阅读理解任务**

针对过敏感和过稳定问题，分别从相似问句构建和相似篇章构建两个角度提出两种数据增强方式，针对领域内和未知领域问题使用增量和二阶段训练增强模型性能，有效增强了模型的鲁棒性。

**10:25-10:50 面向推荐的对话任务**

任务是在给定的知识库下，针对目标完成多轮对话。在该任务目标下，基于预训练模型，采用分桶的编码方式和多源信息融合的解码器，将知识按长度分桶、独立编码，大大减少了内存占用，构建了一个端到端的对话模型，在自动和人工评测均取得第一的好成绩。

**10:50-11:15 语义解析任务**

相较于Seq2Seq的传统方案，新方法将Seq2Seq与模板填充进行结合，首先使用Seq2Seq生成SQL模板，然后采用模板填充技术对模板中的细节部分进行预测填充，有效地解决了传统Seq2Seq所具有的SQL组件顺序敏感、计算资源依赖较高、SQL生成效率较低等问题，在最终测试集上取得了76.5%的成绩

**11:15-11:40 关系抽取任务**

相对于先抽取实体再关系分类的传统方法，把实体抽取任务分为主语抽取、宾语抽取、关系分类三个步骤，输出层采用机器阅读理解中的指针网络作为基本结构，有效解决了实体进行两两匹配带来的大量负样本，用引入虚拟节点的方式解决复杂宾语的问题，用PU Learning解决远监督召回率较低的问题。

**11:40-12:05 事件抽取任务**

针对限定领域的事件抽取任务，不同于先进行触发词抽取再进行事件角色抽取的传统方案，该方案采用基于BERT的多层标签指针网络，对事件角色进行端到端的标注，并有效解决了事件抽取任务中的角色重叠和元素重叠等问题。

如何观看直播

扫描长图中的二维码进入交流群获取直播链接，周六早10点，不见不散！

