附件1

机器翻译技能竞赛

当前机器翻译技术研究与系统开发基本都是以英语为核心，而“一带一路”国家战略的推动和实施迫切需要以汉语为核心的高质量机器翻译技术。本项目通过机器翻译技能大赛，一方面检验当前以汉语为核心的机器翻译技术水平，另一方面发现科学院内从事汉语和一带一路国家语言之间机器翻译技术研发的高水平团队。

**一、项目设置**

本项目设置汉语和西班牙语、葡萄牙语以及印地语之间的双向翻译任务，即设计的机器翻译模型不仅可以将西班牙语、葡萄牙语和印地语翻译为汉语，还需要将汉语翻译为西班牙语、葡萄牙语和印地语。

**二、参赛要求**

本项目允许参赛队伍参加一种语言的比赛或参加多种语言的比赛。例如，可以参加西班牙语和汉语的机器翻译，也可以参加所有语言与汉语之间的机器翻译。参加多个语言与汉语的互译比赛时既可以为每种语言设计一个翻译模型，也可以设计统一的多语言机器翻译模型处理所有语言。所有参赛队伍必须仅使用组织方提供的训练数据，不允许使用额外数据，也不允许使用开源的预训练模型。各参赛队伍使用自己的机器训练、调参和测试，翻译速度建议在一块V100显卡上进行测试。

**三、参赛方式**

组委会将为参赛队伍统一提供训练集、开发集和测试集。

训练集：西班牙语、葡萄牙语和印地语与汉语之间分别100万平行句对；

开发集：西班牙语、葡萄牙语和印地语与汉语之间分别1000个句子；

测试集：西班牙语、葡萄牙语和印地语与汉语之间分别1000个句子。

参赛队伍需提交测试集的翻译结果、测试集按句顺序翻译的时间（说明测试环境）、以及模型规模的截图文件。

**四、测试方法**

本项目采用客观评价指标评价参赛队伍的成绩，客观评价指标包括三类：

**译文质量**：本项目采用BLEU作为自动评价指标衡量机器翻译系统的译文质量，西班牙语、葡萄牙语和印地语到汉语的翻译系统采用汉字级别的BLEU-5对比翻译系统结果与参考答案，汉语到西班牙语、葡萄牙语和印地语的翻译系统采用词语级别的BLEU-4对比翻译系统结果与参考答案。

**翻译系统速度**：在同等翻译质量的前提下，翻译系统的速度越快，成绩越高。

**翻译模型规模**：在同等翻译质量和速度的前提下，翻译模型的参数规模越小，成绩越高。

奖项排名仅考虑译文质量排名前50%的参赛队伍。在译文质量排名前50%的参赛队伍中，按照翻译质量、翻译速度和模型规模三个指标分别排名，然后进行得分的归一化操作，最后按照0.5，0.25和0.25的比列对翻译质量、翻译速度和模型规模的排名进行加权求和，获得最终的排名。如果三种语言采用统一的多语言翻译模型，得分\*1.2。