

Shared Tasks in NLPCC 2014

Call for Participation

The conference on Natural Language Processing and Chinese Computing (NLPCC) is the annual conference of CCF TCCI (Technical Committee of Chinese Information, China Computer Federation). The first and the second NLPCC conferences were successfully held in Beijing in 2012 and Chongqing in 2013. This year's NLPCC conference will be held in Shenzhen on December 5-9, 2014.

NLPCC2014 will follow the NLPCC tradition of holding several shared tasks in natural language processing and Chinese computing. This year's shared tasks focus on several hot research topics, such as sentiment analysis, information linking and organization, and question answering.

Participants from both academia and industry are welcomed. Each group can participate in one or multiple tasks and members in each group can attend the NLPCC conference to present their techniques and results. A few selected groups will be invited to submit papers to the main conference and the accepted papers will appear in the conference proceedings published by Springer LNCS.

1. Overview of the Shared Tasks

There are six shared tasks in this year's NLPCC conference and the details of each task can be found in the document of task guidelines. Here we give a very brief overview of each task.

➤ **Emotion Analysis in Chinese Weibo Texts**

This task aims to evaluate the techniques for analyzing the emotions in Weibo (Chinese twitter) text. The bakeoff includes three sub-tasks, namely Weibo emotion classification (required), emotion sentence identification and classification (required), and emotion expression extraction (optional).

➤ **Deep Learning for Sentiment Classification**

This task aims to evaluate the deep learning techniques for sentiment classification. The data sets are document-level reviews in both Chinese and English. The reviews are collected from product review web sites. One basic requirement is to

employ deep learning techniques in the sentiment classification method.

➤ **Chinese Entity Linking**

This task aims to evaluate the current advance of techniques in aligning named entities from various text resources to entities in a reference Chinese knowledge base.

➤ **Cross-Lingual Knowledge Linking**

Given the datasets of English Wikipedia and Chinese Wikipedia, a set of known cross-lingual links, the task is to discover the equivalent English articles for a set of specified Chinese articles.

➤ **Large-Scale English Question Answering**

This task aims to evaluate the techniques for answering open-domain questions, namely Large-Scale Question Answering over Open-domain Data. Given a natural language question, the purpose of the task is to automatically find 1) whether the question has answers; 2) if the question has answer(s), find the correct answer(s) for it. The questions and answers are limited within English.

➤ **Large-Scale Chinese News Categorization**

This task aims to prompt research in and to evaluate the current development of techniques in automatically categorizing text documents into a predefined taxonomy. In this Text Categorization task, given a news document and a predefined hierarchy of categories with a depth of 2, a system is required to provide the IDs of the categories which this document belongs to.

2. How to Participate

Please fill out the registration form and send it to the coordinator Xinjie Zhou by email (xinjiezhou@pku.edu.cn) before April 15, 2014.

If you have any question about the shared tasks, please do not hesitate to contact us by email (xinjiezhou@pku.edu.cn).

3. Important Dates

2014/3/20: announcement of shared tasks and call for participation;

2014/4/15: registration deadline;

2014/4/30: sample data release;

2014/5/30: test data release;

2014/6/06: systems' results submission deadline;

2014/6/20: evaluation results release and call for conference papers;

2014/7/31: conference paper submission deadline;

2014/8/31: conference paper accept/reject notification and call for system reports;

2014/9/15: camera-ready paper submission deadline;

2014/9/30: notification of system report presentation format;

2014/12/5~9: NLPCC2014 conference;

4. Shared Task Organizers

Yigang Diao, Xinhua News Agency

Yansong Feng, Peking University

Shoushan Li, Soochow University

Xiaojun Wan, Peking University

Zhichun Wang, Beijing Normal University

Furu Wei, Microsoft Research Asia

Ruifeng Xu, Harbin Institute of Technology Shenzhen Graduate School