Report on TCS NLP Winter School 2008

24 Dec 2007 - 7 Jan 2008
IIIT - Hyderabad

TCS NLP Winter School 2008 was held in IIIT-Hyderabad from 24 Dec, 2007 to 7 Jan, 2008. The objective of this winter school was to train students and faculty across the Indian subcontinent in advanced Natural Language Processing (NLP). There were three main areas of NLP: Morphology, Parsing and Machine Translation with focus on Indian Languages. The winter school was project based, in which groups of 2-3 participants were formed to work on a project. These projects were assigned from day 1. Apart from the projects, there were research lectures in various areas of NLP by well-known researchers from India and abroad.

Sixty participants from all over India, Nepal and Sri Lanka attended this two-week winter school. They worked on one of the sub-projects in the three focus areas of the winter school. The sub-projects had varying goals from unsupervised to semi-supervised machine learning with limited linguistics resources. We provided the participants with tools, language resources and text corpora needed for the projects. Each project had a faculty guide and a research student mentor from IIIT-H helping the participants with the projects. There were project meetings every day or once in two days to assess the progress of the projects and to try new ideas.

In parallel, each day, the students had the opportunity to listen to research talks from invited lecturers. There were lectures by eminent NLP researchers from both India and Abroad (The detailed list of invited lecturers is given at the end of this report). There were wide range of topics discussed in the winter school: Morphology, Machine Translation, Dependency Parsing, Grammar Development, Discourse Parsing and Prosody. The idea was for the students to get an exposure to advanced topics in Speech and Language Processing and interact with practitioners in the field.

An NLP Winter School Symposium was held on 7th January, 2008 at which each student group presented results of its project work over the winter school. The achievements were impressive in many projects. The two invited peakers at the symposium were Yuji Matsumoto (NAIST, Japan) and Junichi Tsuji (University of Tokyo, Japan).

Many students plan to follow up on the projects after returning to their respective institutions. We have encouraged them to write papers to upcoming conferences after some more results are obtained.

The room and boarding of all the students was funded, thanks to the sponsorship of Tata Consultancy Services who donated 2 Lakh Rupees for this school.

All in all, this was a great success and personally very gratifying to see students take away some positive enthusiasm for language processing. Managing all these events was a learning experience.

There will an advanced NLP Summer School along similar lines at IIIT, Hyderabad from 26 May, 2008 to 9 June, 2008 (See URL below).

Speakers:

 (i) From India : Prof. Rajeev Sangal (IIIT-Hyderabad), Samar Husain (IIIT-Hyderabad),   Dr. Dipti Misra Sharma (IIIT-Hyderabad), Prof. Sivaji Bandyopadhyay (Jadavpur University), Sudeshna Sarkar (IIT-Kharagpur), Sobha Latha (AU-KBC Research Lab),   Ganesh Ramakrishnan (IIT-Bombay & IBM Research) and Pushpak Bhattacharya (IIT-Bombay).

 (ii) From Abroad : Srinivas Bangalore (AT&T Research Labs, USA), Gerard Huet (Inria, France),   Gina Levow (University of Chicago, USA), Phillip Koehn (University of Edinburgh, UK), Sadao Kurohashi (Kyoto University, Japan), Fei Xia (University of Washington, USA),   Aravind Joshi, (University of Pennsylvania, USA), Michael Gamon (Microsoft Research, USA),Yuji Matsumoto (NAIST, Japan), Junichi Tsuji (University of Tokyo, Japan)

References:

 1. NLP Winter School 2008 - Full Schedule
              -- http://ltrc.iiit.ac.in/winterschool08/index.cgi?topic=schedule
 2. NLP Winter School 2008 - List of Projects
              -- http://ltrc.iiit.ac.in/winterschool08/index.cgi?topic=projects
 3. NLP Summer School 2008 - Announcement
              -- http://ltrc.iiit.ac.in/iasnlp2008/

Report Prepared by:

 Sriram Venkatapathy, Lecturer,
 Language Technologies Research Centre,
 IIIT-Hyderabad, Hyderabad, India.
 sriram@research.iiit.ac.in