INTRODUCTION

Natural Language Processing (NLP) deals with an important shortcoming of the revolution of information explosion. Most of the information on internet and at other sources appears in the form of human language. From a conventional computer professional's perspective, it consists of unstructured data that cannot be processed by tools for structured data. NLP bridges the gap between (seemingly) unstructured text and computer professional's discrete tools. It enables us to design spell checkers, grammar checkers, machine translators, question answering system, summarizers and many other applications. Natural Language Processing is the base of many killer applications of today that includes Google Translate, Google News, Summly and SRI etc.

OBJECTIVES

• To provide basic knowledge about the techniques of processing human language through computers
• To introduce the emerging and dominating field of Statistical Natural Language Processing
• To discuss development of NLP applications for Pakistani languages

EXPECTED

At the end of the course, the participants would be able to create simple NLP applications. A student with these skills can go to the industry or further studies. In industry giants e.g. Google, Yahoo and Microsoft, there is a huge demand of the people working on language processing application.

CONTENTS

Day 1: Spell Checking
Day 2: Language Models
Day 3: POS Tagging
Day 4: Named Entity Recognition and Chunking
Day 5: Issues in Natural Language Parsing
Day 6: Parsers for Natural Language
Day 7: Statistical Parsing
Day 8: Information Retrieval
Day 9: Text Mining
Day 10: Killer Applications of Today

The students are advised to also get enrolled for the sister course Linguistic Resources Creation which focuses on the preparation of data for computational processing and machine learning. (The NLP course focuses on how-to-process; the other course focuses on what-to-process.)

TARGET AUDIENCE

The course is primarily for BS-final year or MS/PhD student of Computer Science.

DURATION

The course will be conducted for ten days (5 days a week) from 24-June-2013 to 5-July-2013 at DHA Suffa University, Karachi. The time duration will be from 9:00 AM to 11:30 AM (2.5 hours per day).
APPLICATION
Candidates will be awarded with University approved certificate upon successful completion of the programme.

APPLICATION PROCEDURE
The application form should be obtained from tafseer@dsu.edu.pk. Completed application with course fee should be forwarded on or before 13th June, 2013 to Director CCE.

INTRODUCTORY SEMINAR
A free introductory seminar will be held on 10th June, 2013 at DHA Suffa University, Karachi.

COURSE FEES
Rs. 8,000 Rs. 4,000 per candidate

INSTRUCTOR
Dr. Tafseer Ahmed received PhD degree from University of Konstanz, Germany. He worked on polysemy of spatial markers and non-spatial usages of spatial markers for 10 South Asian languages (including 6 Pakistani languages). In his Post-Doc research, he worked on computational grammar of Urdu using LFG framework. Earlier, he has worked on English to Urdu machine translation and text mining for Urdu.

FURTHER DETAILS
Dr. Asim I. Wagan – Director, CCE, DSU Department of Electrical Engineering, DHA Suffa University, Karachi.
Tel.: 021 5556671~7 ext (274) Fax: 021 2245762