Introduction to the Course

Niloy Ganguly

IITKGP, LUH

Introduction

Course Info

My Contact

- **Email:** ganguly@l3s.de
- **Webpage:** `https://cnerg-iitkgp.github.io/nlp/course/`

Teaching Assistants

- Gourab Patro, patro@l3s.de
- Souwmya Sundaram, sundaram@l3s.de
- Soumyadeep Roy, sroy@l3s.de

Class Timings

- Lectures: Monday 10.30-12.00 (online + hybrid)
- Excercises: Thursday 16.00-17.30 (online/hybrid)

Books and Materials

Reference Books

- Daniel Jurafsky and James H. Martin. 2009. Speech and Language Processing: *An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics*. 2nd edition. Prentice-Hall.
- Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT Press.

Lecture Material

- Lecture Slides
- IPython Notebooks

Lecture Delivery Plan

Online Part

- Initially at Least two lectures online
- Lecture uploaded by Friday morning
- Doubts to be posted by students latest by Sunday noon (Google form provided)

In-person Part

- Monday class in **Lange Laube 6, A112** in normal lecture time
- 50 Students can attend, Form will be posted to seek interest. First 50 students will be admitted
- Doubts will be cleared in the class
- At the end of the class assignment will be posted

Assignment

Assignment Content

Numerical and Programming Exercise

In-person Part

- Thursday practice class will be divided into two parts
  - Discussion of assignments given
  - New Exercise/programs to be solved in class
- Initially online, later plan is to have hybrid mode

Assignment Submission

- Each submission consists of
  - Exercise/programs done in class on Thursday (previous week)
  - Assignments posted on Monday
  - To be submitted by Thursday 3pm
- We will upload the solutions by Thursday EOD

Assignment 1

- In the first week (21st October), no assignment will be provided
- A tutorial will be provided (video) about the basics of python.
- On 21st, students can clear doubts regarding Python programming (online)

Rough Course Content

Entire content may not be covered due to want of time

- Basic Text Processing: Tokenization, Stemming, Spelling Correction
- Language Modeling: N-grams, smoothing
- Morphology, Parts of Speech Tagging
- Syntax: PCFGs, Dependency Parsing
- Lexical Semantics, Word Sense Disambiguation
- Distributional Semantics, Word Embeddings
- Topic Models
- Entity Linking and Information Extraction
- Text Summarization and Text Classification
- Sentiment Analysis and Opinion Mining

Course Evaluation Plan

- Final Test
- Three assignments - one bonus point
- At most three bonus points