CS 4248
Natural Language Processing

Professor NG Hwee Tou
Department of Computer Science
School of Computing
National University of Singapore
nght@comp.nus.edu.sg

Teaching Staff

- Lecturer: Ng Hwee Tou
- Teaching assistants:
  - Lin Ruixi
  - Muhammad Reza Qorib
  - Song Yang
  - Wang Qian

Lecture

- Face-to-face in-person lectures
- Fridays 9am – 11.30am in LT15
- Lectures are recorded

Tutorial

- Bi-weekly in-person tutorial classes
  - Weeks 3, 5, 7, 9, 11, 13
- Venue: COM1 02-07 (SR7)
- Split into 8 tutorial groups (max of 20 students per group)

Tutorial

- Students are to work thorough the tutorial questions (at least attempt to solve them) before coming to each tutorial class
- Students do not submit their tutorial solutions (not graded)
- Solutions to tutorial questions will be posted after all sessions end

Lecture Topics

- Introduction
- Regular expressions, finite state automata
- Words, edit distance
- N-grams
- Part-of-speech tagging

Lecture Topics

- Linear models
- Feed-forward neural networks
- Neural network training
- Neural language models
- Word embeddings

Lecture Topics

- Convolutional neural networks
- Recurrent neural networks
- Sequence-to-sequence models with attention
- Transformers
- Pre-trained language models and fine-tuning

Lecture Topics

- Formal grammars
- Parsing
- Semantics
- Discourse

Textbooks

- Speech and Language Processing, Daniel Jurafsky and James H. Martin, Prentice Hall, Second edition and third edition (draft) (SLP)
- Neural Network Methods for Natural Language Processing, Yoav Goldberg, Synthesis Lectures on Human Language Technologies (NNM4NLP)

Grading

- 5% Participation
- 25% Assignments
- 20% Project
- 50% Exam

Participation

- Active class participation during lectures and bi-weekly tutorial classes
- Attending tutorial classes, asking questions, answering questions, presenting tutorial solutions
- We will take attendance at tutorial classes
- Please attend the tutorial group that you register

Assignments

- All assignments are programming assignments, to be done individually
- Programming language: Python 3
- PyTorch
  - Open source Python package for neural networks

Project

- Group project (5 students per group)
- Form your own project group
- Hands-on programming and implementation
- Project topics
  - One pre-defined topic
  - Other self-proposed topics (subject to approval)

Project

- Project proposal due date: Friday 16 September 2022 9am
- Final project due date: Friday 11 November 2022 9am
- What to submit:
  - Project report (4 pages of content + 1 page of references), Times Roman 11-point font
  - Presentation slides
  - Source code and data (including README instructions on how to run your source code)
  - 15-minute oral presentation in a video recording

Project

- Marks breakdown
  - 50%: Content
  - 30%: Presentation
  - 20%: Novelty

Final Exam

- Date: Tuesday 29 November 2022, 9am – 10.30am (90 minutes)
- Closed book (but one double-sided A4 size sheet is allowed on which anything can be written)