Statistical NLP Logistics

Tim Rocktäschel, Sebastian Riedel

Module Details

- Lecturer: Tim Rocktäschel, Sebastian Riedel, Pasquale Minervini
- Teaching Assistants: Yuxiang Wu, Zhengyao Jiang, Yihong Chen, Mikayel Samvelyan, Fanghua Ye, Robert Kirk
- Lectures:
  - Tuesday 3:30PM - 5PM, Online
  - Thursday 2:30PM - 4PM, Online
- Office Hours, Wednesday 5PM - 6 PM

Prerequisites 1

- **Python**: including classes and inheritance, difference between list and tuples, variable arguments, dictionaries etc.
- **Probabilities**: understand Bayes rule; what does marginalisation mean? What is conditioning? How to sample from a categorical distribution
- **Math**: basic linear algebra, multi-dimensional calculus (differentiation); understand what this means: $\arg\max_x f(x)$

Prerequisites 2

- **Jupyter and Colab Notebooks**: know how to create and run notebooks
- **Command line tools**: know how to use a command line tools, ssh, cd, mv, rm etc.
- **Git**: know how to use git (clone, merge, pull etc.)
- **Installation**: be able to install docker on your machine, or be comfortable running a VM on the cloud (Azure)

Enrolment

- Moodle: https://moodle.ucl.ac.uk/course/view.php?id=1441
- Enrolment key: ???

Managing Your Expectations

- We aren't *linguists*
  - We probably don't know your favourite linguistic framework or terminology
- We aren't *cognitive scientists*
  - We don't know how humans process language
- This course has no *exams*!
  - You should learn how to construct stat NLP systems

Lecture Preparation

- Go through lecture notes
  - Updated as we go
- Read Background Material

Coursework

- **Group Project**, 4 students per group, **100%**
  - Report (**40%**)
  - Video and Code (**30%**)
  - *Release*: Now
  - *Deadline*: Wed 8/04/2022
  - **FIND GROUPS NOW**
- *Top K projects can be presented at London's top NLP Meetup*

MSc Machine Learning Paper Receives International Acclaim

8 October 2019

MSc Machine Learning students achieved double success with published research paper; "SentiMATE: Learning to play Chess through Natural Language Processing", presented at AIIDE-19. It also received a write-up in MIT Technology Review.

Python

- Lectures and assignments focus on Python (version 3.6)
- Python is a leading language for data science, machine learning etc., with many relevant libraries
- We expect you to know Python, or be willing to learn it on your own
- Labs and assignments focus on development within Jupyter/Colab notebooks

Material

- Our Slides (available on Syllabus)
- Goldberg, A Primer on Neural Network Models for Natural Language Processing
  - Completely free online
- Jurafsky & Martin, Speech and Language Processing
  - Many parts free online

Discussion Forum

- Our moodle page has a **discussion forum**
- Please post questions there (instead of private emails)
- Our teaching assistants are actively monitoring the discussion forum
- We give **low priority to questions already answered** in previous lectures, tutorials and posts,
  - and to **pure programming related issues**
- We expect you to **online-search** for answers before.
- You are highly **encouraged to participate and help each other on the forum**.

Contacting TAs and Lecturers

- In case where you want to avoid other students to see your messages to the TAs or Lecturers, use the “Discuss with TA/Lecturer” forum on Moodle.

- Please do not send us direct emails

UCL NLP & DARK

- Research groups at UCL working on NLP, Reinforcement Learning, and their intersection
- Webpage: http://nlp.cs.ucl.ac.uk/, https://ucl-dark.github.io
- Twitter:
  - @ucl_nlp https://twitter.com/ucl_nlp
  - @ucl_dark https://twitter.com/ucl_dark
- Always looking for strong MSc and PhD students!

MSc Projects @ UCL NLP and DARK

- We plan to supervise a range of MSc projects
- Project Ideas are here and here
- Get in touch with the contacts mentioned on the slides