ML4NLP1 - Intro Tutorial
Fall Semester 2025

When, Where, Who

Every Monday at 16:15, RAK-E-8. After the lecture, in the same room. 45-60 mins

Teaching team:

Andrianos Michail
Doctoral Candidate
Computational Linguistics
X-Lingual Semantic Search

Michelle Wastl
Doctoral Candidate
Computational Linguistics
Identifying Semantic Diffs.

Konstantinos
Kalyfommatos
University of Athens
MSc Thesis @ CL

Special thanks to the previous members of the teaching teams:

They have studied and carefully created part of the materials we teach:

- Dr. Tannon Kew
- Shivam Adarsh
- Patrick Haller
- Janis Goldzycher
- Ahmet Yavuz Uluslu
- Phillip Ströbel

What

We bring you on track with modern practical Natural Language Processing. You can see these tutorials content as a hands on experience with:

- Building your own NLP solutions (without always relying on APIs!)
- Reading and understanding weak points of existing NLP solutions.
- Preparation for NLP focused ML Engineer / Data Scientist job assessments

Footnotes: 1. Content of the tutorial MIGHT be recorded. 2. Attendance is not mandatory, but it is wise to hear the clues :D

How

Five Practical Coding Assignments (Done in Teams of 3*) - Introduced biweekly

Exercise 1:

- Released Monday September 29th at 16:00
- Submission Deadline: Sunday 12th of October at 23:59
- Anonymous Peer Review* Deadline: Monday 20th of October at 23:59

....

– Other Practical Content (Can be part of the final exam)

We aim to cover a broad range of topics:

- NLP through classic ML methods
- Building your own Neural Networks in PyTorch
- Word Embeddings, Representation Learning, Embedding (Retrieval) models
- Understanding, prompting and fine-tuning Large (Small) Language Models
- Topic and Neural Topic Modelling

Anonymous Peer Review on OLAT

- After submitting the assignment, each person must complete an anonymous peer review by answering pre-determined questions for one assignment (therefore x3 peer reviews per team)

- Goal of the peer review is to have a guided process to help your peer improve their work and know how to do it better next time.

- The peer review includes a preliminary grade. Final grades are decided by Tutors and don't have to be at all related to the preliminary grade.

Forming Teams of 3 (mandatory)

Please speak with your classmates and form **TEAMS OF 3** that you will work together in all the assignments.

To form your teams, please write us in the forum the three names in the following OLAT thread (add link):

Team Forming Deadline: Friday 3rd of October 2025

How to access files from the tutorials

We put all the tutorial content within the GitHub Repo. In the repo you can find:

- All the content taught within the tutorials (also last year's)
- The given code samples for the exercises.
- Supplementary code that we provide for your ease of access.

The repository will be updated weekly. Content from last year's classes is there. Many times it will be the same content however the tutorial content **MIGHT BE UPDATED** before the tutorial.

Github: https://github.com/Andrian0s/ML4NLP1-2025-Tutorial-Notebooks

Compute for Exercises

The first exercise is executable locally on CPU

For all further exercises, you will need GPU access:

- Some of you might have a GPU in your laptop (e.g. MacBook M1)
- Google Colab GPU (T4) for a limited amount of time
- For esp. compute heavy exercises use Renku
  - Create an account using your SwitchID as described [here](#)
  - Add your Renku @username to [this sheet](#)
  - We will then add you to this course's group and give you instructions on how to use Renku once we introduce exercise 2

In both Google Colab and Renku, the compute will be limited, so make sure to develop your code locally first and then run it on GPU once you know it works to avoid unnecessary debugging time on the GPU instance!

Preliminary Assignment Deadlines (subject to change)

Assignment 01:

Release - 29th September

Submission - 12th October

Peer Review - 20th October

Assignment 02:

Release - 13th October

Submission - 26th October

Peer Review - 3rd November

Assignment 03:

Non practical Assignment handled by Simon. Details to come later but submission date is typically ~20th Jan

Assignment 04:

Release - 3rd November

Submission - 16th November

Peer Review - 24th November

Assignment 05:

Release - 17th November

Submission - 30th November

Peer Review - 7th of December

Assignment 06:

Release - 1st December

Submission - 14th December

Peer Review - 22nd December

FAQs:

Q: Can we also work in a pair or be a solo team?

A: We cannot accommodate pairs or solo teams. As part of also your future job, you will have to collaborate so it's good to get acquainted to the situation early. Importantly, we don't have enough compute/grading capacity for everyone to work in pairs/solo teams.

Q: What's the submission format for Assignments?

A: This year we plan to implement a uniform single file submission where there is an integrated report as part of the jupyter notebooks.

Further FAQs: