Practical Group Exercise (3-4 in a group)

You will have one group assignment for my part of this course. In both options, **it is vital that you document who is in charge of which parts of the assignment**. That doesn't mean that this person has to do everything alone, but rather that they take responsibility from the beginning for the work, organization, and presentation of this part of the assignment. This should be agreed upon before starting the assignment and you should send me a document explaining this agreement (jeremy.barnes@ehu.eus). If you have any problems during the assignment, let me know as soon as possible to avoid misunderstandings later.

Option 1. Implement an HMM PoS tagger
- needs to be implemented in github
- the HMM should follow the model we showed in class
- frequent pushes – it should be clear that the code is updated bit by bit
- experiments with the HMM
    - in-domain experiments:
        - use 2 datasets from UniversalDependencies to train and test your model
        - report the results using accuracy and give a detailed error analysis
    - out-of-domain results
        - use UniversalDependencies models to tag informal data (social media, etc)
        - you may need to annotate some data, or simply do a human evaluation
        - you should determine how robust your model is to domain shift
- experiments, results, and discussion can be done in Colab

Grade will be based on:
- correct implementation of HMM in github (30%)
- correctness of experiments   (40%)
- analysis of results   (30%)

Option 2. Propose a language-specific tagset for a non-English language
- Should not be like UniversalDependencies
- Instead, the tagset should be designed to capture language-specific phenomena
- Each tag should be clearly explained and you should give examples
- Difficulties when deciding on which tags to include should be documented
- Further document any important assumptions you make

- You will need to double annotate a set of 100 sentences
- After annotation, calculate inter-annotator agreement (see here)
- You should document examples where the annotators don't agree and analyse these
- If needed, propose changes to the tagset and clearly define any limitations.
- The tagset, data, and a short report (3-5 pages) will need to be submitted.

Grade will be based on:
- documentation, explanation and reasoning behind the proposed tagset (30%)
- the double annotation amount (30%)
- inter-annotator agreement, analysis and further thoughts (40%)