

TP1: NLP basics - Manipulation and Preprocessing

Les exercices de ce TP, les instructions et le code sont entièrement contenus dans le Notebook `tp_intro.ipynb`, que vous allez trouver dans ce répertoire. Il y a deux méthodes pour l'ouvrir et l'exécuter, dont la première est vivement conseillée parce que ça vous évite d'installer du software sur votre machine locale. En outre, l'exécution du code sera plus rapide.

Méthode 1: Google Colab

- Inscrivez-vous à Google Colab (si vous n'avez pas déjà de compte Google) et demandez l'accès.
- Vous pouvez ouvrir un nouveau fichier Colab depuis [votre compte Google Drive](#) en cliquant sur `+ New->More->Google Colaboratory`
- Une fois que vous avez obtenu l'accès à un fichier Colab, vous pouvez télécharger un Notebook en utilisant `fichier->Upload Notebook`. Téléchargez le fichier `tp_intro.ipynb`.
- Autrement, en tapant `Ctrl + O` (ou dans les options `File->Open`) vous pouvez télécharger une copie du Notebook present sur le répertoire GitHub.
- Pour activer l'utilisation de GPU pour votre notebook: `Runtime->Change runtime type->Hardware Accelerator->GPU`.

Méthode 2: conda environment

Installez [Miniconda3](#) (Anaconda3 marche aussi), ensuite, placez-vous dans le répertoire de ce TP (avec `cd home/username/mypath`) et exécutez les commandes suivantes dans l'ordre:

```
conda create --name nlpenv python
conda activate nlpenv
conda install -c conda-forge notebook
conda install scikit-learn
conda install -c conda-forge ipywidgets
```

Maintenant, vous pouvez ouvrir `tp_intro.ipynb`, il suffit de lancer la commande `jupyter notebook` dans le répertoire de travail.

À Rendre

Chaque exercice de ce TP prévoit une réponse sous forme textuelle ou de code. Toute réponse doit être écrite dans `tp_intro.ipynb`, dans une ou plusieurs cellules sous-jacentes l'énoncé de chaque exercice.

Vous rendrez un répertoire compressé `tp_intro_nom1_nom2.zip` avec le contenu du répertoire `tp_intro.zip` dont le fichier `tp_intro.ipynb` aura été mis à jour avec vos réponses.