

This is a guide to help you set up Python, Jupyter notebook, and other packages you will need on your system to engage in class lab activities. *You are NOT required to use Python for your assignments and projects.* However, because Python seems to be the most commonly used language for text mining/NLP these days, we will focus on it for practice.

Download and install Anaconda environment for Python 3.12

Download Anaconda: <https://www.anaconda.com/download/>

Anaconda environment will help you manage and deploy packages/libraries more efficiently.

To get started with conda, see Conda Getting Started guide.

Create new anaconda environment for this class called *textmining*

```
conda create --name textmining python=3.12
```

Activate environment

Run either

```
source activate textmining
```

or

```
conda activate textmining
```

Check version (should be 3.12)

```
python --version
```

Install packages

Be sure to install these specific versions so that everyone in the class is using the same environment and debugging is easier.

```
conda install spacy=3.7.2
conda install pandas=2.2.3
conda install nltk=3.9.1
conda install wordcloud=1.9.3
conda install scikit-learn=1.5.2
```

```
conda install gensim=4.3.3
conda install seaborn=0.13.2
conda install tabulate=0.9.0
conda install pytorch=2.5.1
conda install transformers=4.45.2
conda install unidecode=1.3.8
conda install xmltodict=0.13.0
pip install sklearn_crfsuite
```

Install spaCy English model

```
python -m spacy download en_core_web_sm
```

Add conda environment to jupyter notebook

```
conda install -c anaconda ipykernel
python -m ipykernel install --user --name=textmining
```

Use Jupyter notebooks

Most class exercises will be provided to you as Jupyter notebooks. To open a Jupyter Notebook in this setup, open up the terminal and navigate to the folder containing the notebook; then select “textmining” as the kernel in Jupyter Notebook to access these libraries:

```
jupyter notebook
```

If you’re new to Jupyter notebooks, you can check out the following tutorial:

- How to Use Jupyter Notebook: A Beginner’s Tutorial

If you receive Jupyter-related error messages while downloading packages, check out the following link:

- How To Use Conda Environment In a Jupyter Notebook

This guide has been adapted from <https://github.com/dbamman/anlp21/blob/main/0.setup>.