Dr. Sanjan Gupta
PhD in Applied AI| UW Madison, USA
IIT-Madras | B.Tech. Honors alum
tpsanjan@gmail.com

For DL and NLP demos

***ETI – Estimated Time to Install when tested on a 16GB RAM, i7 11th gen laptop with ~30 Mbps internet connection speed**

To install PyTorch (**ETI <15 min** - downloads about 2.7 GB of contents)

>> pip3 install torch torchvision torchaudio --index-url https://download.pytorch.org/whl/cu118

To create a new environment in conda:

>> conda create --name nlp_projects

>> conda activate nlp_projects

To install spaCy:

Option 1: Using conda **(ETI < 12 min)**

>> conda install -c conda-forge spacy

>> python -m spacy download en_core_web_sm

OR

Option 2: Using pip

>> pip install -U pip setuptools wheel

>> pip install -U spacy

To download models for using trained pipelines:

>> python -m spacy download en_core_web_sm

For more info: https://spacy.io/usage#installation

Troubleshooting guidelines:

(1) If installing packages through conda is taking a very long time i.e. stuck at solving environment step, it is very likely that there are too many packages in base environment. In general, it's always a best practice to have the base environment loaded with minimal number of packages and use miniconda as against conda.

Enter 'Ctrl + C' or 'Ctrl + Z' to forcefully exit and then try following to deinstall unnecessary packages from base environment:

```
>> conda install --rev 0
>> conda update conda
>> conda update conda-build          # Optional
```

If removing dependencies is not an option, one could try the following command:

```
>> conda install conda=22.9.0
```

Alternatively, one could use libmamba solver (a faster option by Anaconda):

```
# install mamba
>> conda install -n base conda-forge::mamba
>> conda config --set solver libmamba

# use mamba
>> mamba install <package-name>
```

If installing mamba is not an option (for corporate issued laptops that use Z-scaler), try:

```
>> conda config --set channel_priority flexible
```

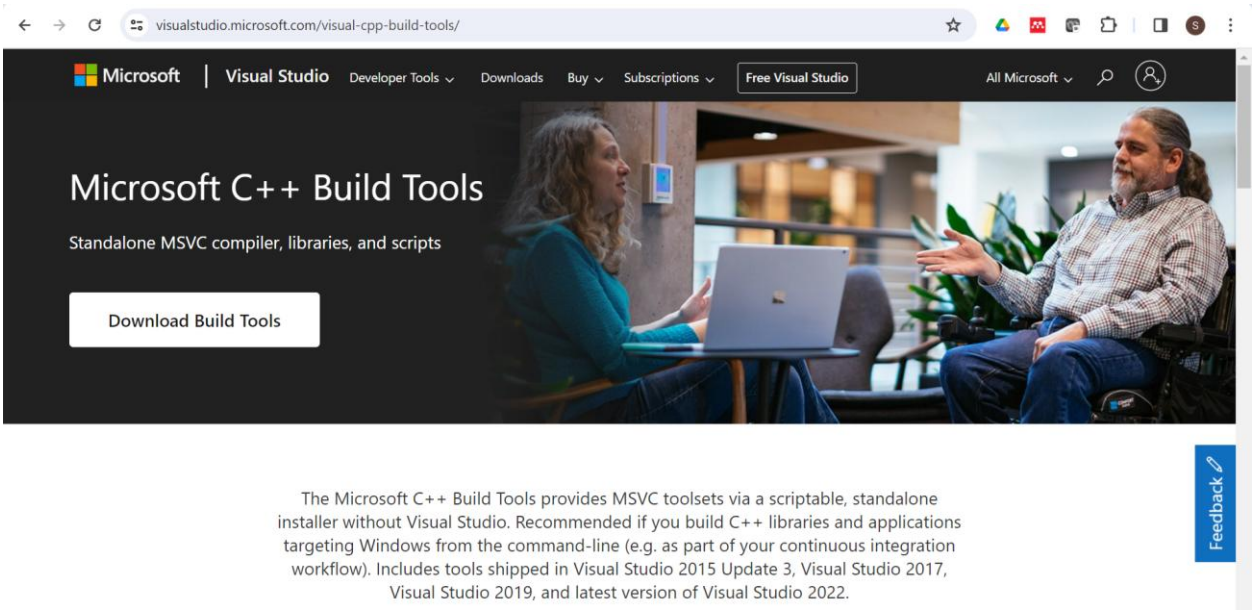
A faster alternative to mamba is to use micromamba that can be installed via homebrew on MacOS or using PowerShell on Windows OS:
https://mamba.readthedocs.io/en/latest/installation/micromamba-installation.html

(2) If you encounter Microsoft C++ build error during installation of scispaCy or medspaCy (error screenshot enclosed below):

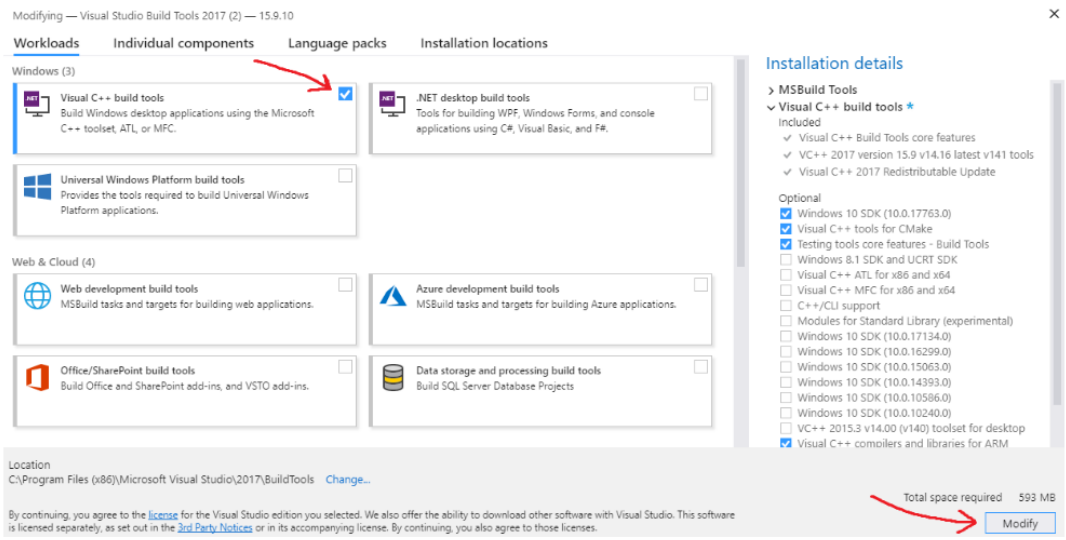
```
      error: Microsoft Visual C++ 14.0 or greater is required. Get it with "Microsoft C++ Build Tools": https://visuals
udio.microsoft.com/visual-cpp-build-tools/
      [end of output]
```

[ETI: < 15 min]

(a) Go to https://visualstudio.microsoft.com/visual-cpp-build-tools/ and click 'Download Build Tools'. Once the download is complete, run the .exe file by double clicking.



(b) Just choose the 'Visual C++ build tools' option by pressing the checkmark and then click 'Modify' or 'Install' at bottom right corner of the screen to begin installation.



(c) After the C++ tools finish installing, reboot the computer

(d) Now go to the anaconda prompt, activate the conda environment, and run the pip command again. It should work this time!