

Sprach-KI zum Anfassen:

Docker und Jupyter Installation

für Windows

von Nicolas Coaca, 31.05.2024

Dieses Skript dient nur als Installationsanleitung für Docker und Jupyter. Es wurde auf Windows 10 getestet. Wenn ihr mehr über die beiden Programme erfahren wollt, könnt ihr in unserem Notebook “Sprach_KI_Anleitung.ipynb” in den Kapiteln “Überblick über das Lab/Notebook” und “Sprach-KI-auf dem eigenen Rechner” nachschlagen, nachdem ihr Jupyter installiert habt.

WICHTIG: Wenn ihr Code aus dem PDF kopiert solltet ihr darauf achten, das keine Zeilenumbrüche übernommen werden (das ist etwas, was PDFs machen)

(Hinweis: Wenn ihr eine Konsole (Terminal) öffnen wollt, drückt die Windows-Taste + R und gebt dort “cmd” ein)

1. Docker:

1.1 Kurze Einführung:

- Docker ist ein Programm zur Virtualisierung anderer Programme
- Sehr gut um Programme über verschiedene Systeme zu verwenden

1.2 Installation:

Ladet euch Docker Desktop für Windows hier herunter:

<https://www.docker.com/products/docker-desktop/>

Startet das eben heruntergeladene Programm mit Administratorberechtigung (Rechtsklick auf das Symbol und auf "Als Administrator ausführen" drücken) und wartet bis es fertig ist.

Es kann sein, dass ihr aufgefordert werdet den Computer neuzustarten, sollte das passieren führt einfach einen Neustart durch.

Um Docker dann zu starten müsst ihr Docker Desktop als Administrator ausführen.

Damit Docker funktioniert müsst ihr die Virtualisierung für euren Prozessor im BIOS aktivieren. Da es zuviele verschiedene Möglichkeiten gibt dies zu tun, rate ich euch danach im Internet zu suchen (wie ihr das BIOS öffnet und wie ihr die Virtualisierung aktiviert). Ihr könnt im Task Manager (ihr könnt diesen mit der Tastenkombination STRG + SHIFT + Esc öffnen) unter dem Punkt "Leistung" im Bereich "CPU" überprüfen ob die Virtualisierung bereits aktiviert ist.

Sollte es zu Problemen kommen oder Unklarheiten entstehen könnt ihr hier mehr Informationen bekommen: <https://docs.docker.com/desktop/install/windows-install/>

2. Jupyter:

2.1 Kurze Einführung:

- Jupyter ist ein Programm um Texte zu schreiben und diese mit funktionierendem und änderbarem Code auszustatten
- Es werden viele Sprachen unterstützt wie zum Beispiel C++, Java oder Ruby (Standardmäßig wird Python verwendet)
- Das Projekt ist Open-Source

2.2 Installation:

Hier könnt ihr entscheiden ob ihr Jupyter bei euch lokal installiert oder ob ihr dafür Docker verwenden wollt. Wenn ihr euch dazu entscheidet Jupyter lokal zu verwenden, dann müsst ihr auch Ollama (mehr Informationen später in dem Notebook "Sprach_KI_Anleitung.ipynb") lokal installieren.

2.2.1 Installation über Docker:

Diese Methode ist bevorzugt, da nichts zusätzlich auf euren Computer installiert wird sondern in einem Docker Container läuft. Es kann manchmal passieren, dass diese Installation nicht klappt (Ollama funktioniert nicht). Sollte das passieren, müsst ihr auf die lokale Installation wechseln.

Als erstes müsst ihr das Docker Image “ollama-jupyter” (die Datei heißt ollama-jupyter.tar) herunterladen, welches sich hier befindet:

<https://cloud.hs-augsburg.de/s/jcQanraHafR3rj8>

Öffnet als nächstes ein Terminal und verwendet diesen Befehl:

```
docker load -i <Dateipfad/zu/dem/Image>
```

Wobei ihr <Dateipfad/zu/dem/Image> gegen den Installationspfad austauschen müsst (Bsp.: docker load -i /home/benutzer/ollama-jupyter.tar). Der Befehl kann mehr Zeit in anspruch nehmen.

Anschließend könnt ihr entscheiden (wenn ihr eine Nvidia Grafikkarte habt), ob ihr eure Grafikkarte nutzen wollt oder nicht. Führt nur einen von den beiden Befehlen aus

Wenn ihr keine oder eine AMD Grafikkarte habt, nutzt diesen Befehl:

```
docker run -d -v ollama-jupyter:/root/.ollama -p 8888:8888 -p 11434:11434 --name ollama-jupyter ollama-jupyter
```

Wenn ihr eine Nvidia Grafikkarte habt könnt ihr diesen Befehl benutzen:

```
docker run -d --gpus=all -v ollama-jupyter:/root/.ollama -p 8888:8888 -p 11434:11434 --name ollama-jupyter ollama-jupyter
```

Sobald der von euch ausgeführte Befehl bearbeitet wurde, solltet ihr in der Lage sein Jupyter nutzen zu können. Ihr könnt das ausprobieren indem ihr auf diesen Link klickt: <http://localhost:8888/>

2.2.2 Lokale Installation:

Um Jupyter lokal zu installieren brauchen wir sowohl Python als auch ein Programm, welches Python-Pakete verwaltet. Dazu nutzen wir Miniforge3, da dieses beides tut. Über diesen Link könnt ihr das Installationsskript herunterladen: https://github.com/conda-forge/miniforge/releases/latest/download/Miniforge3-Windows-x86_64.exe

WICHTIG: Bitte nicht in den Namen des Installationsverzeichnis Leerzeichen, Umlaute oder ß verwenden

Zu guter Letzt müsst ihr noch ein paar Pakete installieren, die wir später brauchen werden. Öffnet dazu Miniforge (es ist eine Konsole) und verwendet diesen Befehl:

```
1. pip install jupyterlab==4.2.1 cython==3.0.10  
numpy==1.26.4 ollama==0.2.0 langchain==0.2.1  
langchain_community==0.2.3 tqdm==4.66.2  
chromadb==0.5.0 tiktoken==0.7.0 bs4==0.0.2  
gensim==4.3.2 scipy==1.10.1 deep-translator==1.11.4  
nltk==3.8.1 matplotlib==3.9.0 torch==2.3.0  
transformers==4.41.1
```

Um Jupyter auszuführen könnt ihr in der Miniforge Konsole diesen Befehl verwenden:

```
jupyter-lab
```

WICHTIG: Ihr dürft das Terminal nicht schließen!

3. Abschluss:

Jetzt da ihr die wichtigsten Programme installiert habt könnt ihr die Datei Sprach_KI_Anleitung.ipynb herunterladen (von hier: <https://cloud.hs-augsburg.de/s/jcQanraHafR3rj8>) und mit Jupyter öffnen. Dies macht ihr indem ihr in dem Dateiverzeichnis, welches euch in Jupyter auf der linken Seite angezeigt wird, die Datei sucht und selbige mit einem doppelklick öffnet.