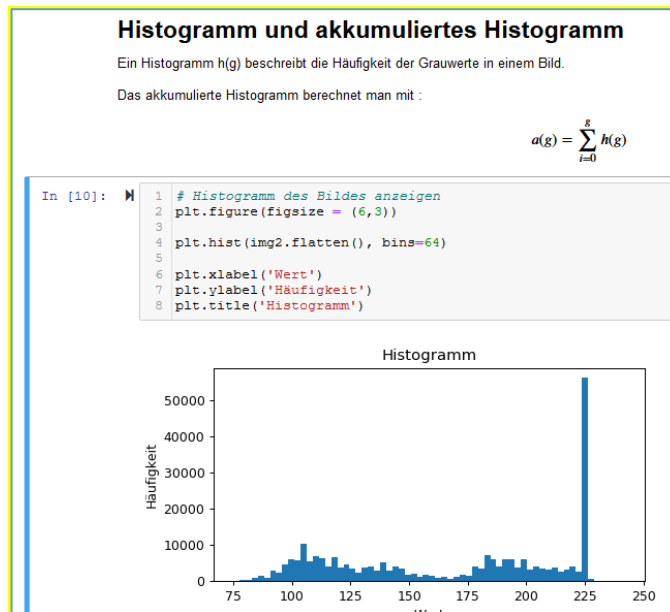


Jupyter Notebooks

Jupyter-Notebooks sind eine einfache und elegante Möglichkeit, Python-Programme im Webbrowser zu entwickeln und laufen zu lassen. Man ist damit komplett plattformunabhängig. Programm und Dokumentation bilden eine Einheit.

Da der größte Teil der Machine-Learning-Frameworks (Tensorflow, Keras, PyTorch,) mit Python arbeiten, sind Jupyter-Notebooks wie gemacht dafür.



Verwendete Versionen

Wir verwenden Python 3.7 und Tensorflow 2.x.

Tensorflow 2.x beinhaltet Keras, im Gegensatz zu den Vorgängerversionen.

Achtung: Tensorflow 1.x (und Keras) – ML-Programme sind mit Tensorflow 2.x (und Keras)-ML-Programmen nicht ohne Modifikationen kompatibel.

Einrichten einer Entwicklungsumgebung

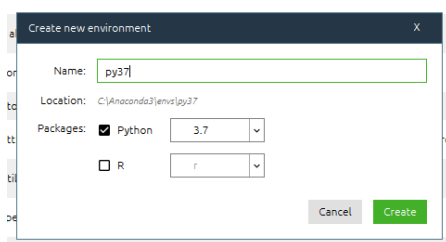
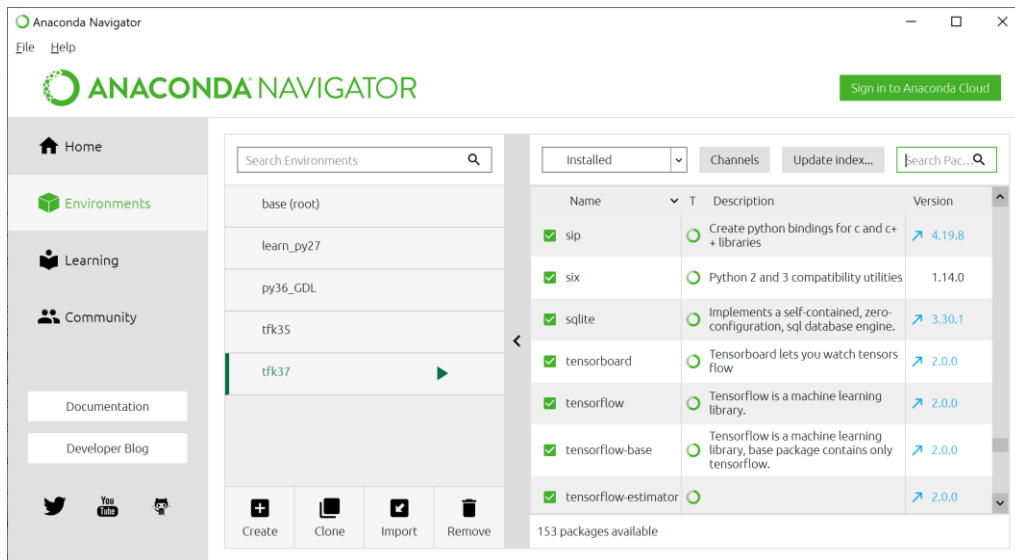
Es gibt es mehrere Möglichkeiten.

Möglichkeit 1: Installation auf dem eigenen Rechner

Anm.: Ich beschreibe die Installation unter Windows. Auf Mac und Linux gibt es vergleichbare Wege.

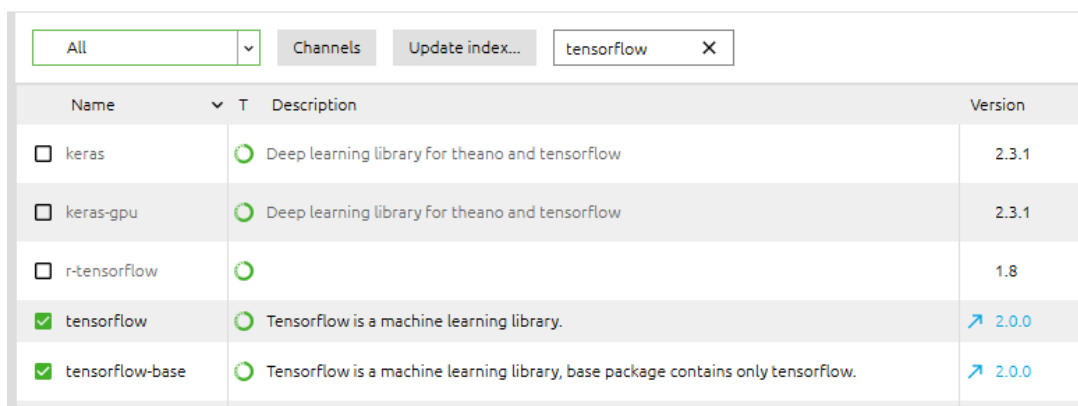
1. Anaconda 3 installieren (= Python-Environment-Manager. Unbedingt verwenden, wenn man sich nicht den Rechner zumüllen will.)

- Unter „Environments“ (s.Bild) ein neues Environment erzeugen (Create), z.B. mit dem Namen py37 o.ä..



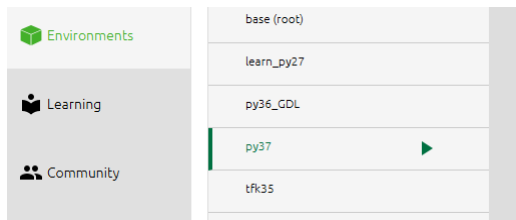
- Folgende Bibliotheken zusätzlich installieren (durch Anwählen und „Apply“, s.Bild):

tensorflow, tensorflow-base, scipy, scikit-learn, scikit-image, pillow
 opencv, pyopencv, libopencv, graphviz, pydot, numpy, matplotlib
 bokeh, h5py, pandas, cudatoolkit, cudnn, seaborn

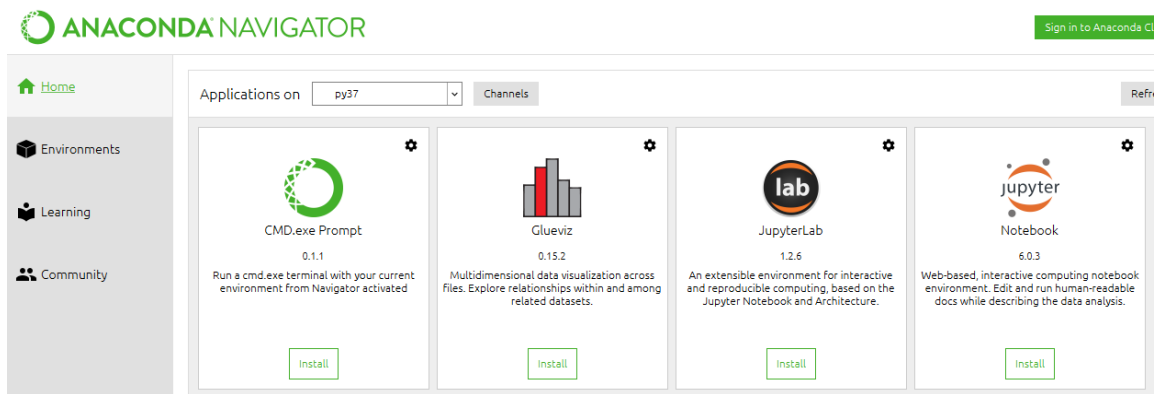


Sollte später noch was fehlen, kann man jederzeit Bibliotheken nachinstallieren.

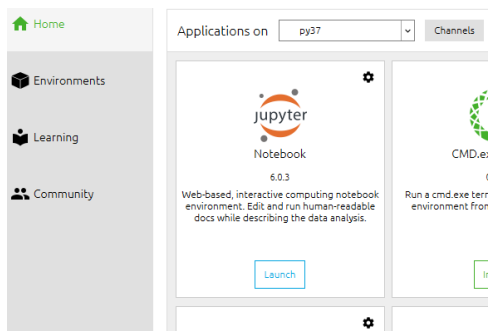
4. Environment anwählen und nach Home wechseln ...



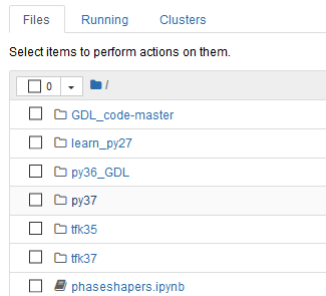
5. Jupyter-Notebook installieren (nur bei der ersten Verwendung des Environments)



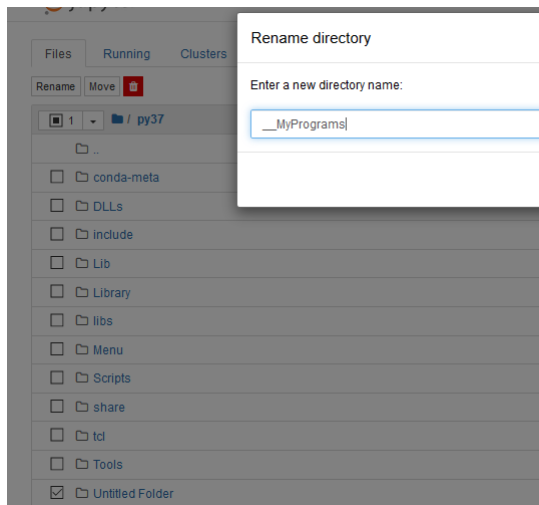
6. Notebook starten („Launch“)



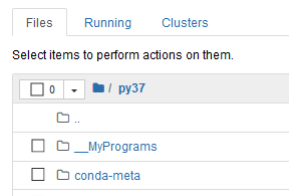
→ Webbrowser startet und zeigt die Verzeichnisstruktur der vorhandenen Environments an (nach der Erstinstallation gibt es natürlich nur ein Environment, z.B. py37).



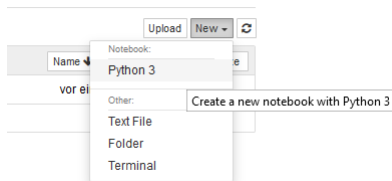
7. Environment-Verzeichnis (z.B. py37) öffnen und ein Verzeichnis „__MyPrograms“ erzeugen („New“)



I
V



8. Da reingehen und entweder ein leeres Notebook erzeugen („New“) oder ein vorhandenes laden („Upload“).



9. Jetzt kann's losgehen.

Anm.: Ich habe z.B. Anaconda3 direkt auf Laufwerk c: installiert.

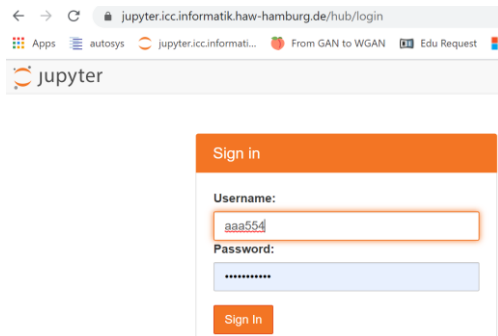
Im Dateibrowser ist der Pfad dann z.B.:

Local Disk (C:) > Anaconda3 > envs > py37 > __MyPrograms

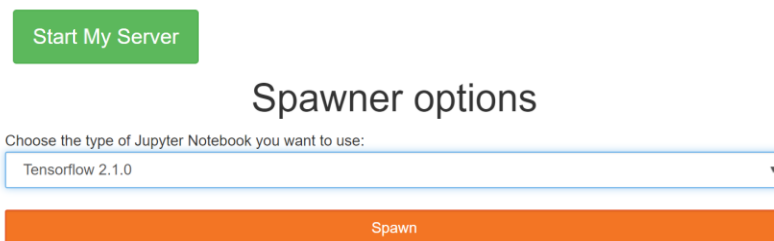
Möglichkeit 2: jupyter.icc.informatik.haw-hamburg.de

Anm.: Auch in der HAW-Computer-Cloud kann man Notebooks editieren und laufen lassen, ohne etwas installieren zu müssen.

1. Einloggen mit HAW-Account bei <https://jupyter.icc.informatik.haw-hamburg.de/>



2. Start Server

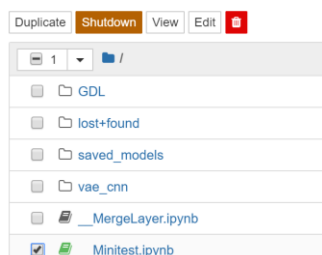


Es gibt auch die GPU-Variante, die sollte man aber erst dann wählen, wenn man sehr rechenintensive Machine-Learning-Jobs startet. Achtung: Vorher unbedingt den vorherigen Server stoppen (sonst kommt man sehr lange nicht mehr rein sehr dumm im Praktikum). Insgesamt gibt es z.Zt. 8 High-Performance-GPU's (TESLA). Man kann sich vorstellen, dass das eine gefragte Ressource ist. Also nicht unnötig belegen. Für die ersten Versuche (reine Bildverarbeitung) braucht man die GPUs gar nicht.

3. Notebook anlegen oder uploaden Und ab geht's

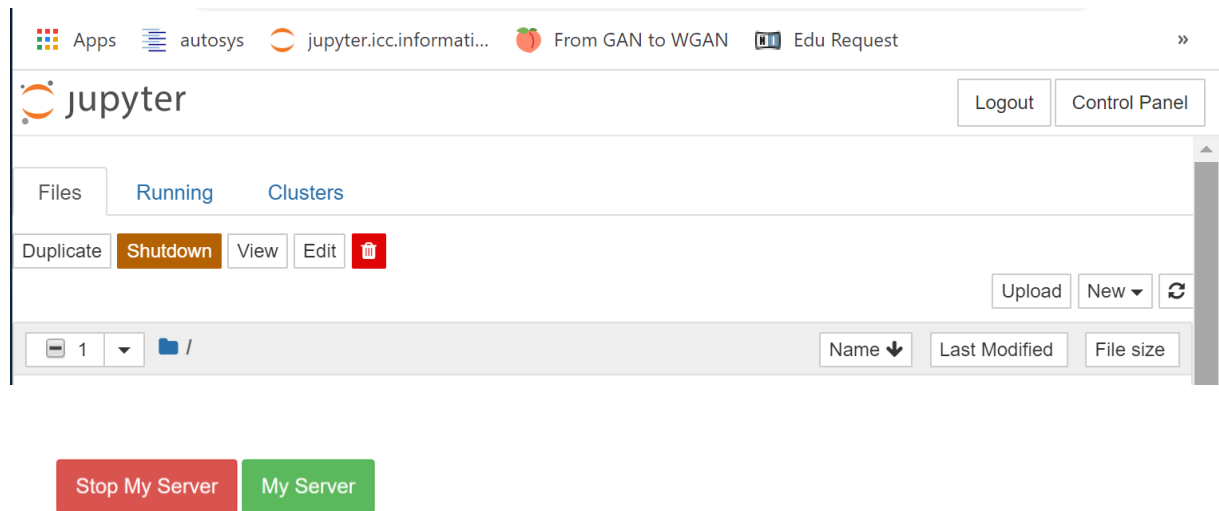
Anm. 1: Auf fremden Servern sollte man immer eine Kopie der Software auf dem eigenen Rechner haben. Man kann nie sicher sein, dass die Speicherung auf dem Server von Dauer ist.

Anm. 2: Bevor man ein neues Notebook startet, müssen alle anderen Notebooks jeweils gestoppt werden. Speziell bei Machine-Learning-Programmen kann es sonst zu Resource-Overload-Error kommen.



Anm. 3: Bevor man sich ausloggt, **unbedingt (!!!) vorher den Server stoppen** („Control Panel“) sonst kann man sich längere Zeit nicht mehr einloggen.

Hintergrund: Der Server bleibt belegt und wird erst nach längerer Untätigkeit autom. gestoppt.



Möglichkeit 3: Microsoft-Azure (nicht geeignet für Tensorflow 2.x-Programme)

Anm.: Alternativ geht auch Google-Colab, das beschreibe ich hier aber nicht. Das Design ist auch etwas anders.

1. Folgenden Link aufrufen: <https://notebooks.azure.com/>
2. Kostenfreien Account einrichten.
Wer ein hotmail-Konto hat, kann sich auch damit einen Account einrichten.
3. Einloggen
4. My Projects → New Project → New Notebook (im Moment nur Python 3.6)
Alternativ: Notebook (*.ipynb) uploaden

Anm. 1: Auf fremden Servern sollte man immer eine Kopie der Software auf dem eigenen Rechner haben. Man kann nie sicher sein, dass die Speicherung auf dem Server von Dauer ist.

Anm. 2: Leider wird z.Zt. nur bis Python 3.6 angeboten.

Die ML-Bibliotheken tensorflow und keras müssen deswegen anders importiert werden. Das spielt aber frühestens ab Versuch 3 eine Rolle. Die vorgefertigten Praktikums-Programmhüllen müssen dann geringfügig angepasst werden.