

Installation

A number of Python libraries are required to be installed prior to using this software. We have made the program available in a Jupyter notebook which is one of the most popular methods of executing Python code and provides the flexibility to immediately see the results of running the Python code. It is an intermediate solution between a code editor and a full-fledged web application.

Selenium controls the browser through its web driver which communicates with a browser driver which is required to be installed on the system. We are using a chrome browser. Following steps explain how to install chrome's latest version and chrome driver.c

Installation

Vor der Verwendung dieser Software müssen eine Reihe von Python-Bibliotheken installiert werden. Wir haben das Programm in einem Jupyter-Notebook verfügbar gemacht, das eine der beliebtesten Methoden zum Ausführen von Python-Code ist und die Flexibilität bietet, die Ergebnisse der Ausführung des Python-Codes sofort anzuzeigen. Es ist eine Zwischenlösung zwischen einem Code-Editor und einer vollwertigen Webanwendung.

Selenium steuert den Browser über seinen Webtreiber, der mit einem Browsertreiber kommuniziert, der auf dem System installiert werden muss. Wir verwenden einen Chrome-Browser. Die folgenden Schritte erklären, wie Sie die neueste Version von Chrome und Chrome Driver.c installieren

Browser and Driver Installation

Browser

- Check your version of Chrome browser by clicking 3 dots in the upper right hand corner of the browser
=> help => About Google Chrome

From there you can check the version and also whether it is up to date or not.

If you require download the latest version of Chrome from the official website
(<https://www.google.com/chrome/>)

Browser- und Treiberinstallation

Browser

- Überprüfen Sie Ihre Version des Chrome-Browsers, indem Sie auf die drei Punkte in der oberen rechten Ecke des Browsers klicken => Hilfe => Über Google Chrome

Von dort aus können Sie die Version überprüfen und auch, ob sie aktuell ist oder nicht.

Wenn Sie die neueste Version von Chrome von der offiziellen Website (<https://www.google.com/chrome/>) herunterladen möchten

Driver

- Latest version of Google Chrome Driver will be required. To check your version:

- Visit the official Chrome driver download page at <https://sites.google.com/a/chromium.org/chromedriver/downloads>

Download the latest version of Chrome driver that corresponds to your operating system and Chrome browser version.

Extract the downloaded file to a suitable location on your machine.

Add the path to the extracted Chrome driver executable to your system's PATH environment variable so that it can be accessed from the command line (Type environmental variable in the windows search bottom left, and open edit environmental variables => system variables => path => edit => copy and paste the path to the driver executable where you had saved it.

Once you have installed the Chrome driver, you can use it with Selenium by specifying the path to the executable when you launch the driver, like this:

```
from selenium import webdriver
```

Set the path to the Chrome driver executable

```
chrome_driver_path = "/path/to/chromedriver"
```

Launch a new instance of Chrome driver

```
driver = webdriver.Chrome(chrome_driver_path)
```

Make sure to replace "/path/to/chromedriver" with the actual path to the Chrome driver executable on your machine.

Fahrer

- Die neueste Version des Google Chrome-Treibers ist erforderlich. So überprüfen Sie Ihre Version:

- Besuchen Sie die offizielle Chrome-Treiber-Downloadseite unter
<https://sites.google.com/a/chromium.org/chromedriver/downloads>

Laden Sie die neueste Version des Chrome-Treibers herunter, die Ihrem Betriebssystem und Ihrer Chrome-Browserversion entspricht.

Extrahieren Sie die heruntergeladene Datei an einen geeigneten Ort auf Ihrem Computer.

Fügen Sie den Pfad zur ausführbaren Datei des extrahierten Chrome-Treibers zur Umgebungsvariable PATH Ihres Systems hinzu, damit Sie über die Befehlszeile darauf zugreifen können (geben Sie die Umgebungsvariable in die Windows-Suche unten links ein und öffnen Sie Umgebungsvariablen bearbeiten => Systemvariablen => Pfad => Bearbeiten => kopieren und den Pfad zur ausführbaren Treiberdatei dort einfügen, wo Sie sie gespeichert haben.

Nachdem Sie den Chrome-Treiber installiert haben, können Sie ihn mit Selenium verwenden, indem Sie beim Starten des Treibers den Pfad zur ausführbaren Datei wie folgt angeben:

von Selen Import Webdriver

Legen Sie den Pfad zur ausführbaren Chrome-Treiberdatei fest

```
chrome_driver_path = "/pfad/zu/chromedriver"
```

Starten Sie eine neue Instanz des Chrome-Treibers

```
Treiber = webdriver.Chrome(chrome_driver_path)
```

Stellen Sie sicher, dass Sie „/path/to/chromedriver“ durch den tatsächlichen Pfad zur ausführbaren Chrome-Treiberdatei auf Ihrem Computer ersetzen.

Windows Miniconda

Go to the [Miniconda Downloads page](<https://docs.conda.io/en/latest/miniconda.html#windows-installers>). Download the appropriate (32- or 64-Bit) Python 3.X version of Miniconda.

- Double click on the .exe file and click Install.
- Read and agree to the licensing terms.
- Select if you want to install for 'Just Me' or 'All Users'. If you are installing for 'All Users', you must have Administrator privileges.
- You will be prompted to select the installation location. By default, Anaconda should try to install in your home directory. We recommend accepting this default. Click Install.
- You will be asked if you want to add Anaconda to your PATH environment variable. Do not add Anaconda to the PATH because it can interfere with other software.
- You will be asked if you want Anaconda to be your default version of Python. We recommend 'Yes'. There are some rare instances where you might not make Anaconda the default version, but they are beyond the scope of this project.
- Click the Install button.

First, let's verify that Miniconda was installed. We will use the Anaconda Prompt (miniconda3) Go to the Start Menu and search for the Anaconda Prompt. Anaconda Prompt is the Command Line Interface (CLI) we will use with Miniconda (more about the Command Line). In the shell, type:

```
conda --version
```

A number should be returned. This is the version number of conda that you have installed.

Windows-Miniconda

Rufen Sie die [Miniconda-Downloadseite] (<https://docs.conda.io/en/latest/miniconda.html#windows-installers>) auf. Laden Sie die entsprechende (32- oder 64-Bit) Python 3.X-Version von Miniconda herunter.

- Doppelklicken Sie auf die .exe-Datei und klicken Sie auf Installieren.
- Lesen und akzeptieren Sie die Lizenzbedingungen.
- Wählen Sie aus, ob Sie für „Nur mich“ oder „Alle Benutzer“ installieren möchten. Wenn Sie für „Alle Benutzer“ installieren, müssen Sie über Administratorrechte verfügen.
- Sie werden aufgefordert, den Installationsort auszuwählen. Standardmäßig sollte Anaconda versuchen, sich in Ihrem Home-Verzeichnis zu installieren. Wir empfehlen, diese Voreinstellung zu akzeptieren. Klicken Sie auf Installieren.

- Sie werden gefragt, ob Sie Anaconda zu Ihrer PATH-Umgebungsvariable hinzufügen möchten. Fügen Sie Anaconda nicht zum PATH hinzu, da es andere Software stören kann.
- Sie werden gefragt, ob Anaconda Ihre Standardversion von Python sein soll. Wir empfehlen „Ja“. Es gibt einige seltene Fälle, in denen Sie Anaconda möglicherweise nicht zur Standardversion machen, aber sie würden den Rahmen dieses Projekts sprengen.
- Klicken Sie auf die Schaltfläche Installieren.

Lassen Sie uns zunächst überprüfen, ob Miniconda installiert wurde. Wir werden die Anaconda-Eingabeaufforderung (miniconda3) verwenden. Gehen Sie zum Startmenü und suchen Sie nach der Anaconda-Eingabeaufforderung. Anaconda Prompt ist die Befehlszeilenschnittstelle (CLI), die wir mit Miniconda verwenden werden (mehr über die Befehlszeile). Geben Sie in der Shell Folgendes ein:

```
conda --version
```

Es soll eine Zahl zurückgegeben werden. Dies ist die Versionsnummer von Conda, die Sie installiert haben.

Install Packages

Now we need to install a few key packages. In the Anaconda Prompt, type

```
conda install jupyter
```

When prompted with 'Proceed ([y]/n)?' type y

Once the installation completes, in the Anaconda Prompt, type:

```
jupyter notebook
```

A Jupyter Notebook interface will appear in your default browser.

!image-5.png](attachment:image-5.png)

An Anaconda Prompt might also open and display a url. If it does, do not close it until you are done working with Jupyter Notebook. If it does not appear, don't worry — this has to do with your operating system and will not affect Jupyter Notebook's performance.

Let's install a few packages to get you started working with data in Jupyter Notebook. These are all Python packages, so we will use pip instead of conda to install them.

We will load:

- requests for communicating with remote servers and websites
- Numpy for working with arrays.
- pandas for working with DataFrames.
- scikit-learn for machine learning.
- Matplotlib for visualizations.
- seaborn for statistical visualizations.
- NLTK for text analysis.
- spaCy for more text analysis.
- Selenium for front-end testing and web scraping.
- BeautifulSoup 4 for web scraping.
- Scrappy for web crawling and web scraping.

Open a new Anaconda Prompt from start menu and type the following commands (hit 'Enter' after each row):

```
pip install requests
```

```
pip install numpy
```

```
pip install pandas
```

```
pip install scikit-learn
```

```
pip install matplotlib
```

```
pip install seaborn
```

```
pip install nltk
```

```
pip install beautifulsoup4
```

```
pip install scrapy
```

```
pip install selenium
```

You can check to be sure these are all installed by starting Python in your Anaconda Prompt.

On a new line, type python

The header will change to be >>>:

![[image-6.png]](attachment:image-6.png)

You can import each package on a separate line or take a shortcut and import them all at once. Type:

```
import numpy
```

```
import pandas
```

```
import requests
```

```
import sklearn
```

```
import matplotlib
```

```
import seaborn
```

```
import nltk
```

```
import bs4
```

```
import scrapy
```

```
import selenium
```

If nothing happened, everything is installed perfectly, and your system is ready

Pakete installieren

Jetzt müssen wir ein paar Schlüsselpakete installieren. Geben Sie in der Anaconda-Eingabeaufforderung Folgendes ein

Conda installiert Jupyter

Wenn Sie mit „Fortfahren ([y]/n)?“ aufgefordert werden, geben Sie y ein

Geben Sie nach Abschluss der Installation in der Anaconda-Eingabeaufforderung Folgendes ein:

Jupyter-Notizbuch

In Ihrem Standardbrowser wird eine Jupyter Notebook-Oberfläche angezeigt.

(Anlage:image-5.png)

Eine Anaconda-Eingabeaufforderung kann auch geöffnet werden und eine URL anzeigen. Wenn dies der Fall ist, schließen Sie es nicht, bis Sie die Arbeit mit Jupyter Notebook abgeschlossen haben. Wenn es nicht angezeigt wird, machen Sie sich keine Sorgen – dies hat mit Ihrem Betriebssystem zu tun und wirkt sich nicht auf die Leistung von Jupyter Notebook aus.

Lassen Sie uns ein paar Pakete installieren, um Ihnen den Einstieg in die Arbeit mit Daten in Jupyter Notebook zu erleichtern. Dies sind alles Python-Pakete, daher verwenden wir pip anstelle von conda, um sie zu installieren.

Wir werden laden:

- Anfragen zur Kommunikation mit entfernten Servern und Websites
- Numpy für die Arbeit mit Arrays.
- Pandas für die Arbeit mit DataFrames.
- scikit-learn für maschinelles Lernen.
- Matplotlib für Visualisierungen.
- Seaborn für statistische Visualisierungen.
- NLTK für die Textanalyse.
- spaCy für mehr Textanalyse.
- Selenium für Front-End-Tests und Web-Scraping.

- BeautifulSoup 4 für Web Scraping.
- Scrappy für Web-Crawling und Web-Scraping.

Öffnen Sie eine neue Anaconda-Eingabeaufforderung aus dem Startmenü und geben Sie die folgenden Befehle ein (drücken Sie nach jeder Zeile die Eingabetaste):

Pip-Installationsanfragen

pip installiere numpy

Pandas installieren

pip installiere scikit-learn

pip installiere matplotlib

Pip installieren Seaborn

pip installiere nltk

Pip installiere BeautifulSoup4

pip install kratzig

pip Selen installieren

Sie können überprüfen, ob diese alle installiert sind, indem Sie Python in Ihrer Anaconda-Eingabeaufforderung starten.

Geben Sie in einer neuen Zeile python ein

Die Kopfzeile ändert sich zu >>>:

![[image-6.png]](Anlage:image-6.png)

Sie können jedes Paket in einer separaten Zeile importieren oder eine Verknüpfung verwenden und alle auf einmal importieren. Typ:

importiere numpy

Pandas importieren

Anfragen importieren

sklearn importieren

matplotlib importieren

seegeboren importieren

nltk importieren

bs4 importieren

schrottig importieren

Selen importieren

Wenn nichts passiert ist, ist alles perfekt installiert und Ihr System ist bereit