Course software for decision trees / random forests

In the practicals, we will use only a small dataset and we will need only little computation power and memory ressources. You can therefore do the practicals on various computing platforms. However, since the participants may use various types of computers and softwares, we recommend to use the UNIL JupyterLab to do the practicals.

- JupyterLab: Working on the cloud is convenient because the installation of the Python and R packages is already done and you will be working with a Jupyter Notebook style even if you use R. Note, however, that the UNIL JupyterLab will only be active during the course and for one week following its completion, so in the long term you should use either your laptop or Curnagl. Access requires that you connect either via the eduroam Wi-Fi with your UNIL account or through the UNIL VPN. This point is especially crucial for researchers from the CHUV.

- Laptop: This is good if you want to work directly on your laptop, but you will need to install the required libraries on your laptop. Warning: We will give general instructions on how to install the libraries on your laptop but it is sometimes tricky to find the right library versions and we will not be able to help you with the installation. The installation should take about 15 minutes.

- Curnagl: This is efficient if you are used to work on a cluster or if you intend to use one in the future to work on large projects. If you have an account you can work on your /scratch folder or ask us to be part of the course project but please contact us at least a week before the course. If you do not have an account to access the UNIL cluster Curnagl, please contact us at least a week before the course so that we can give you a temporary account. The installation should take about 15 minutes.

If you choose to work on the UNIL JupyterLab, then you do not need to prepare anything since all the necessary libraries will already be installed on the UNIL JupyterLab. In all cases, you will have access to the UNIL JupyterLab.

Otherwise, if you prefer to work on your laptop or on Curnagl, please make sure you have a working installation before the day of the course as on the day we will be unable to provide any assistance with this.

If you have difficulties with the installation on Curnagl we can help you so please contact us before the course at helpdesk@unil.ch with subject: DCSR ML course.

On the other hand, if you are unable to install the libraries on your laptop, we will unfortunately not be able to help you (there are too many particular cases), so you will need to use the UNIL Jupyter Lab during the course.

Before the course, we will send you all the files that are needed to do the practicals.

JupyterLab for Python and R

Here are some instructions for using the UNIL JupyterLab to do the practicals.

Access requires that you connect either via the eduroam Wi-Fi with your UNIL account or through the UNIL VPN.

This point is especially crucial for researchers from the CHUV.

The webpage's link will be given during the course.

Enter your UNIL credentials (username and password).

The JupyterLab is provided by Open On Demand.

Fill in the form as shown in the lecture's slides.

Copy / paste the commands from the html practical file to the Jupyter Notebook.

To execute a command, click on "Run the selected cells and advance" (the right arrow), or SHIFT + RETURN.

When using TensorFlow, you may receive a warning

WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
I0000 00:00:1789027670.401347 159087 cudart_stub.cc:31] Could not find cuda drivers on your machine, GPU will not be used.
I0000 00:00:1789027671.309670 159087 cpu_feature_guard.cc:227] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
To enable the following instructions: AVX2 FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
I0000 00:00:1789027675.036749 159087 cudart_stub.cc:31] Could not find cuda drivers on your machine, GPU will not be used.

You should not worry. By default, TensorFlow is trying to use GPUs and since there are no GPUs, it writes a warning and decides to use CPUs (which is enough for our course).

When you have finished the practicals, simply click on Log out.

Laptop for Linux, Mac and Windows

You may need to install development tools including a C and Fortran compiler (e.g. Xcode on Mac, gcc and gfortran on Linux, Visual Studio on Windows).

Python installation

On Windows, if you do not have Python installed, you can use either Conda:

https://docs.conda.io/en/latest/miniconda.html or Python official installer:

https://www.python.org/downloads/windows/ . Also, you may need to use the commands "py -3.12 -m venv mlcourse" and "mlcourse\Scripts\activate" instead of the ones written below.

Here are some instructions for installing decision tree and random forest libraries on your laptop. You need Python >= 3.12.

We will use a terminal to install the libraries.

Let us create a virtual environment. Open your terminal and type:

```
python3.12 -m venv mlcourse

source mlcourse/bin/activate

pip install scikit-learn pandas matplotlib graphviz seaborn notebook
```

You can terminate the current session:

```
deactivate

exit
```

TO DO THE PRACTICALS (today or another day):

You can use any Python IDE (e.g. Jupyter Notebook or PyCharm), but you need to launch it after activating the virtual environment. For example, for Jupyter Notebook:

```
source mlcourse/bin/activate

jupyter notebook
```

Information: Use Control-C to stop this server.

R installation

Here are some instructions for installing decision tree and random forest libraries on your laptop.

You need R >= 4.0. Run R in your terminal or launch RStudio.

For Windows users, you can download R here: https://cran.r-project.org/bin/windows/base/

REMARK: The R libraries will be installed in your home directory. To allow it, you must answer yes to the questions:

Would you like to use a personal library instead? (yes/No/cancel) yes

Would you like to create a personal library to install packages into? (yes/No/cancel) yes

And select Switzerland for the CRAN mirror.

```
install.packages("rpart")

install.packages("rpart.plot")

install.packages("randomForest")

install.packages("tidyverse")
```

The installation of "tidyverse" may lead to some conflicts, but do not worry you should be able to do the practicals fine.

You can terminate the current R session:

```
q()
```

Save workspace image? [y/n/c]: n

TO DO THE PRACTICALS (today or another day):

Simply run R in your terminal or launch RStudio.

Curnagl

For the practicals, it will be convenient to be able to copy/paste text from a web page to the terminal on Curnagl. So please make sure you can do it before the course. You also need to make

sure that your terminal has a X server.

For Mac users, download and install XQuartz (X server): https://www.xquartz.org/

For Windows users, download and install MobaXterm terminal (which includes a X server). Click on the "Installer edition" button on the following webpage: https://mobaxterm.mobatek.net/download-home-edition.html

For Linux users, you do not need to install anything.

Python installation

Here are some instructions for installing decision tree and random forest libraries on the UNIL cluster called Curnagl. Open a terminal on your laptop and type (if you are located outside the UNIL you will need to activate the UNIL VPN):

```
ssh -Y < my unil username >@curnagl.dcsr.unil.ch
```

Here and in what follows we added the brackets < > to emphasize the username, but you should not write them in the command. Enter your UNIL password.

For Windows users with the MobaXterm terminal: Launch MobaXterm, click on Start local terminal and type the command ssh -Y < my unil username >@curnagl.dcsr.unil.ch. Enter your UNIL password. Then you should be on Curnagl. Alternatively, launch MobaXterm, click on the session icon and then click on the SSH icon. Fill in: remote host = curnagl.dcsr.unil.ch, specify username = < my unil username >. Finally, click ok, enter your password. If you have the question "do you want to save password ?" Say No if your are not sure. Then you should be on Curnagl.

See also the documentation: https://wiki.unil.ch/ci/books/high-performance-computing-hpc/page/ssh-connection-to-dcsr-cluster

```
cd /scratch/< my unil username >
```

or

```
cd /work/TRAINING/UNIL/CTR/rfabbret/cours_hpc/
mkdir < my unil username >
cd < my unil username >
```

For convenience, you will install the libraries from the frontal node to do the practicals. Note however that it is normally recommended to install libraries from the interactive partition by using (Sinteractive -m 4G -c 1).

```
module load python/3.12.12

python -m venv mlcourse

source mlcourse/bin/activate

pip install scikit-learn pandas matplotlib graphviz seaborn
```

You can terminate the current session:

```
deactivate

exit
```

TO DO THE PRACTICALS (today or another day):

```
ssh -Y < my unil username >@curnagl.dcsr.unil.ch
```

```
cd /scratch/< my unil username >
```

or

```
cd /work/TRAINING/UNIL/CTR/rfabbret/cours_hpc/< my unil username >
```

For convenience, you will work directly on the frontal node to do the practicals. Note however that it is normally not allowed to work directly on the frontal node, and you should use (Sinteractive -m 4G -c 1).

```
module load python/3.12.12

source mlcourse/bin/activate

python
```

R installation

Here are some instructions for installing decision tree and random forest libraries on the UNIL cluster called Curnagl. Open a terminal on your laptop and type (if you are located outside the UNIL you will need to activate the UNIL VPN):

```
ssh -Y < my unil username >@curnagl.dcsr.unil.ch
```

Here and in what follows we added the brackets < > to emphasize the username, but you should not write them in the command. Enter your UNIL password.

For Windows users with the MobaXterm terminal: Launch MobaXterm, click on Start local terminal and type the command ssh -Y < my unil username >@curnagl.dcsr.unil.ch. Enter your UNIL password. Then you should be on Curnagl. Alternatively, launch MobaXterm, click on the session icon and then click on the SSH icon. Fill in: remote host = curnagl.dcsr.unil.ch, specify username = < my unil username >. Finally, click ok, enter your password. If you have the question “do you want to save password ?” Say No if your are not sure. Then you should be on Curnagl.

See also the documentation: https://wiki.unil.ch/ci/books/high-performance-computing-hpc/page/ssh-connection-to-dcsr-cluster

```
cd /scratch/< my unil username >
```

or

```
cd /work/TRAINING/UNIL/CTR/rfabbret/cours_hpc/
mkdir < my unil username >
cd < my unil username >
```

For convenience, you will install the libraries from the frontal node to do the practicals. Note however that it is normally recommended to install libraries from the interactive partition by using (Sinteractive -m 4G -c 1).

```
module load r-light/4.5.2


R
```

REMARK: The R libraries will be installed in your home directory. To allow it, you must answer yes to the questions:

Would you like to use a personal library instead? (yes/No/cancel) yes

Would you like to create a personal library to install packages into? (yes/No/cancel) yes

And select Switzerland for the CRAN mirror.

```
install.packages("rpart")


install.packages("rpart.plot")


install.packages("randomForest")
```

```
install.packages("tidyverse")
```

The installation of "tidyverse" may lead to some conflicts, but do not worry you should be able to do the practicals fine.

You can terminate the current R session:

```
q()
```

Save workspace image? [y/n/c]: n

TO DO THE PRACTICALS (today or another day):

```
ssh -Y < my unil username >@curnagl.dcsr.unil.ch
```

```
cd /scratch/< my unil username >
```

or

```
cd /work/TRAINING/UNIL/CTR/rfabbret/cours_hpc/< my unil username >
```

For convenience, you will work directly on the frontal node to do the practicals. Note however that it is normally not allowed to work directly on the frontal node, and you should use (Sinteractive -m 4G -c 1).

```
module load r-light/4.5.2


R
```

Créé 2022-09-12 08:44:41 UTC par Ewan Roche
Mis à jour 2026-09-14 13:19:51 UTC par Philippe Jacquet