# Data Stream Learning and Application
# TP N° 1

## Online Machine Learning Model using River

### Submission date : 05/12/2022 à 12h00

The purpose of this lab is to practice data stream processing using the well known framework River. You will then perform different experiments on different classifiers using River and compare their performance giving guidance for the best choice.

## What to do ?

This is a two parts lab. First, you will discover the framework River. Second, you will compare classifiers performance using River.

1. Discover River by practice
   Download the zip code of Professor Albert Bifet River Lab (Source) and follow the instructions described in the Jupyter Notebook.

2. Classifiers performance comparison
   - ➢ Download Covertype dataset available here
   - ➢ Run several experiments with at least 5 classifiers, to find the classifier with better performance for this dataset
   - ➢ Write a short report presenting
     - classifiers selected for the experiments
     - the results of the experiments
     - a discussion about the results
     - which classifier do you recommend to use with the Covertype dataset

## What to submit ?

One compressed file (`.zip`, `.rar`, `tar.gz`...) which will contain :

1. The two Jupyter Notebooks for the two parts. On top of each Jupyter Notebook, put a cell with the team members name
2. 'html' version of notebook (go to: `File/Download as/HTML`), with visible outputs of your code
3. The report discusses the results of part two (PDF).

## How to submit ? Do the 2 steps

1. **First, on the Moodle of Polytechnique** : Upload the compressed file and name it as follow: ‘Student1LastName_Student2LastName_lab1.zip’
2. **Then, By email** : Send the two **HTML version of the notebook** + PDFreport to mariam.barry@polytechnique.edu and maurras.togbe@isep.fr ⇒ EMAIL OBJECT SHOULD BE - IP Paris - M2DS - Lab 1 - River - FirstStudent FirstName_LASTName + Second_student_FirstName_LASTName