Project 2 CMLI & CDS courses: Classification

The objective of this project is to create a model for predicting forest cover Class_Type=7. You will continue with the activity started in the Computational Machine Learning I class "Forest cover classification". This project will be evaluated by both courses (Computational Machine learning and Computing for Data Science) independently. In order to deliver the project, create a Github repository for this project (make it private for the Kaggle competition) and share it with both professors (jejewson, Icedarr).

Evaluation:

Computing for Data Science:

Your coding skills will be evaluated for this course. In this project, you should try to apply the principles taught during the classes: DRY, the use of functions, loops, Functional Programming and Classes with OOP. That does not mean you need to use everything, use what it makes sense to you based on the lectures but try to be DRY and use functions and/or Classes. I will evaluate how clean the code is and how well is following the principles and best practices learnt during the classes.

For the final delivery, try to write clean and reusable code and organise it in .py files to be imported in the final notebook. It is not compulsory to create a library for this project (with the setup.py and all) - although it is advised to do so-, but the .py files should be importable from the notebook. When making Git commits, follow the best practices taught in class (clear comments, imperative statements, atomic commits).

15% of the Computing for Data Science course final grade is based on this project.

Deadline: Sunday 28th November 11:59pm