Machine Learning 2020 - Milestone 2 - Supervised Learning

Francisco P. Romero - University of Castilla La Mancha 26/10/2020

Your task is to predict the number of dengue cases each week (in each location) based on environmental variables describing changes in temperature, precipitation, vegetation, and more. These data are from a competition of the site DrivenData [1].

Deliverables

The students will make a report (strong limit 10 pages) describing the following: the faced, the used techniques and their parameterization and the results obtained including the DrivenData user.

Evaluation

- (1.0) Report Presentation
- (1.5) Baseline: test with basic regression techniques (kNN, Decision Trees,) with simple optimization.
- (1.0) Optimization: Use of GridSearch/ to optimize some
- (1.5) Novelty: Introduce variations and lines of improvements
- (2.0) Competition Result: from $[2 * (32 - result)]/10$.
- (3.0) Evaluation of the code and the github repository (presentation, code, reproducibility, etc.)

Template

Student submissions must include a short report following the ACM template.

`https://www.acm.org/publications/proceedings-template`

Deadline: 15 Jan

[1] `https://www.drivendata.org/competitions/44/dengai-predicting-disease-spread/`