Predictive Models

Second practical project

Ivo Gonçalves
igoncalves@novaims.unl.pt

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Objectives

- Given a regression dataset, explore Genetic Programming (GP) strategies that generate individuals with competitive generalization

- Find ways to properly balance between learning the training data and generalizing to unseen data

Task

- You are provided with a dataset and are allowed to use it as you wish

- You can use any form of tree-based GP (e.g., standard, geometric semantic, etc)

- You can use methods from outside the GP area as long as the final individual is a GP individual

- In the end of the project, you are submitting a final individual to be evaluated on data that was withheld

The only restriction

- The final model must be a tree-based GP individual, defined by the same structure as provided in the baseline code

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Dataset

- Real-world regression dataset

- 12 attributes

- 210 instances provided to build your individuals

- 90 instances withhold to evaluate the final individual

Delivery

- Send me an email until 2016-06-10 23:59, with the following:

    - the proposed individual
    - the source code
    - a small report describing how the final individual was build

Evaluation

- 90% based on the performance of the proposed individual

- 10% for the report