Machine Learning in practice

Marta Sestelo & Nora M. Villanueva

GDG DevFest 2018

Santiago de Compostela, Spain. 17 November 2018

[+34] 986 120 430 | gradiant@gradiant.org | www.gradiant.org

Summary

Machine Learning framework will be introduced by showing both supervised and unsupervised learning methods. We will describe some of the most well-known and widely used techniques: (1) regression models, (2) classification models, (3) generalized additive models, (4) tree-based models and (5) clustering. We will focus on understanding some mathematical details of these models and how to carry out a practical use of them. We will illustrate these techniques analyzing two different case studies, using to this end a free interactive computing environment, a Jupyter notebook including the R kernel.

Installation Instructions

Docker installation

We are going to install Docker Community Edition (CE). It is ideal for developers and small teams looking to get started with Docker and experimenting with container-based apps. You can find instructions in the following links:

- Ubuntu: https://docs.docker.com/install/linux/docker-ce/ubuntu/
- Mac OS: https://store.docker.com/editions/community/docker-ce-desktop-mac
- Windows: https://store.docker.com/editions/community/docker-ce-desktop-windows

Once it has been installed, the following step will be download and run the *Gradiant Jupyter Notebook* image that contains the jupyter app with the IRKernel and the packages needed already installed.

Gradiant Jupyter R Image

- This image includes libraries for data analysis from the R community.

- Run the Gradiant Jupyter Notebook image in your Docker container:

```
mkdir ml_notebook
docker run -p 8888:8888 -v "$(pwd)/ml_notebook:/notebooks" -ti --rm gradiant/jupyter:5.7.0-gdg2018
```

- Browse to http://localhost:8888

- Download the *workshop_machine_learning.ipynb* file from https://github.com/sestelo/machine_learning_workshop_gdg2018