

Cognitive Systems. 2nd assignment

Goal

The goal of the second assignment is to put in practice the theory of data science from a model creation and deployment perspective. Tomorrow, if a client asks you to create a model to be used in a real environment you will need to:

- Investigate your data (by applying CRISP-DM methodology)
- Create your models
- Evaluate your models
- Deploy your models

Description

Mushrooms4all is an organization that aims to promote the collection and use of mushrooms for feeding purposes. However, this organization is aware that there are many mushrooms that are not suitable for human consumption, being dangerous to eat them since they can cause poisoning. Although they have a quite complete database of dangerous/non-dangerous mushrooms, they aim to create a machine learning model that, based on the mushroom characteristics, allows to classify a not known mushroom species as safe or not for human consumption.

Mushrooms4all provides their current dataset of mushrooms that are dangerous and non-dangerous and asked you to create a model that allows to classify a new mushroom as that can cause poisoning or not. They also require having an “application”, or at least, an endpoint/API to query the model. Hence, the deployment of the model is required.

Deliverables

1. Report: A report that contains at least the following information:
 - a. Data exploration, description, preparation and analysis of the dataset variables (distribution of the variables, outlier detection, correlation analyses, ...).
 - b. Model creation: explanation of the models created, the reason behind choosing one algorithm or another, as well as the parameters used.
 - c. Evaluation of the models, and explanation of the results.
 - d. Deployment: explanation of how the model has been deployed and how can be used in a real environment.
2. Code: any code created/used.

Students will work on groups of 2. The cover page will contain the name of both persons together with contact details.

Deadline

December 7th, 2018 (23:55)

Submission

Mail to alejandro.rg@upm.es with Subject [Cognitive Systems: Project 2nd delivery]. All the components of the group should be in copy in the email

Doubts

Any doubt regarding the assignment can be asked Alejandro Rodríguez by email always on an email containing in the subject: [Cognitive Systems]

In Moodle we can discuss on the forum in the channel opened for this purpose.

Presentation

The members of each group will do a presentation to show the obtained results.