

Classes Planning

PRIVATE – ONLY FOR COURSE INSTRUCTORS

© Nuno António 2023 – v 1.01 (2023-02-12)

WEEK 0 – BEFORE CLASSES: SELF-ASSESSMENT SURVEY + PYTHON COURSE + BRANDWATCH CONFIGURATION

To be done one week before the first class.

1. Configure students' access to:
 - a. Datacamp
2. Send email to students:
 - a. Reinforce the requirement of completing the Datacamp's Python courses (Introduction to Python, Intermediate Python, and Data manipulation with pandas)
3. Publish CUF and update course website page
4. Publish Moodle content

WEEK 1 - THEORY: COURSE OVERVIEW + INTRODUCTION TO ML + INTRODUCTION TO CRISP-DM

1. Present slides 0 – Course overview
2. Reinforce:
 - a. Need to finish Datacamp courses in two weeks
 - b. Need to enroll into groups for the three group projects
 - c. Need to install Python packages and Anaconda
3. Mention the files available in Moodle
4. Present slides 1 – Introduction to Machine Learning
5. Present slides 2 – Introduction to CRISP-DM

WEEK 2 – THEORY: DATA UNDERSTANDING + DATA PREPARATION + MODELING

1. Present slides 3 – Data understanding
2. Present slides 4 – Data preparation
3. Present Project 1 – Regression:
 - a. Present the notebook
 - b. Discuss the deadline

- c. Discuss the need to start work on it ASAP
 - d. Discuss the practical classes
- 4. Present slides 5 – Regression (model validation part)
 - a. Do not show the notebooks of the application exercises – it is for the practical classes

WEEK 3 – THEORY: REGRESSION

- 1. Present slides 5 – Regression
 - a. Do not show the notebooks of the application exercises – it is for the practical classes

WEEK 4 – PRACTICE: DATA UNDERSTANDING + DATA PREPARATION + REGRESSION

- 1. Since some students do not have previous experience with Python and data analysis, at the same type, you should explain notebooks and the Python code. Always relate the findings to the business objective. If possible, leave some time for students to tweak some parameters and see how they impact results.
- 2. Show and detail the notebooks:
 - a. PredictMedicalExpenses_Modeling.ipynb
 - b. PredictMedicalExpenses_DeploymentExample.ipynb
 - c. PredictMedicalExpenses_Modeling_DT.ipynb
 - d. PredictCustomerCLV.ipynb
 - e. PredictCustomerCLV_GS.ipynb
 - f. PredictCustomerCLV_RS.ipynb
 - g. PredictCustomerCLV_Optuna.ipynb

WEEK 5 – PRACTICE: REGRESSION (CONT.)

- 1. Continue the work of the previous class
- 2. Make yourself available to answer any questions on project 1 and pass thru the different groups to check on their progress

WEEK 6 – THEORY: CLASSIFICATION

- 1. Present Project 2 – Classification:
 - a. Present the notebook
 - b. Discuss the deadline
 - c. Discuss the need to start work on it ASAP
 - d. Discuss the practical classes
- 2. Present slides 6 – Classification
 - a. Do not show the notebooks of the application exercises – it is for the practical classes

WEEK 7 - THEORY: CLASSIFICATION (CONT.)

1. Continue the previous class

WEEK 8 – PRACTICE: CLASSIFICATION

1. Since some students do not have previous experience with Python and data analysis, at the same type, you should explain notebooks and the Python code. Always relate the findings to the business objective. If possible, leave some time for students to tweak some parameters and see how they impact results.
2. Show and detail the notebooks:
 - a. PredictBankTelemarketingSuccess_LR.ipynb
 - b. PredictBankTelemarketingSuccess_SVM.ipynb
 - c. PredictBankTelemarketingSuccess_KNN.ipynb
 - d. PredictBankChurn_DT.ipynb
 - e. PredictBankChurn_NB.ipynb
 - f. PredictBankChurn_NN.ipynb

WEEK 9 – THEORY: ENSEMBLES OF MODELS AND MODELS’

INTERPRETABILITY

1. Present slides 7 – Ensembles of models, including the notebooks associated to it
2. Present slides 8 – Models’ interpretation, including the notebooks associated to it

WEEK 10 – THEORY: CLUSTERING + EVALUATION & DEPLOYMENT

1. Present Project 3 – Clustering:
 - a. Present the notebook
 - b. Discuss the deadline
 - c. Discuss the need to start work on it ASAP
 - d. Discuss the practical classes
2. Present slides 9 – Clustering
3. Present slides 10 – Evaluation and deployment

WEEK 11 – PRACTICE: CLASSIFICATION

1. Continue the presentation of the notebooks
2. Make yourself available to answer any questions on project 2 and pass thru the different groups to check on their progress

WEEK 12 – PRACTICE: CLUSTERING

1. Since some students do not have previous experience with Python and data analysis, at the same type, you should explain notebooks and the Python code. Always relate the findings to the business objective. If possible, leave some time for students to tweak some parameters and see how they impact results.
2. Show and detail the notebooks:
 - a. ClusteringWithPCA.ipynb
 - b. HierarchicalClustering.ipynb
3. Make yourself available to answer any questions on project 3 and pass thru the different groups to check on their progress

WEEK 13 – THEORY: CLUSTERING (CONT.) + EVALUATION AND DEPLOYMENT

1. Since some students do not have previous experience with Python and data analysis, at the same type, you should explain notebooks and the Python code. Always relate the findings to the business objective. If possible, leave some time for students to tweak some parameters and see how they impact results.
2. Continue to show and detail the notebooks:
 - a. ClusteringWithPCA.ipynb
 - b. HierarchicalClustering.ipynb
3. Make yourself available to answer any questions on project 3 and pass thru the different groups to check on their progress

WEEK 14 – PRACTICE: PROJECTS DISCUSSION

1. Taking into consideration at least project 1 and project 2 evaluation, give qualitative global and per group feedback on the positive points and the points where there is space for improvement. The objective is that understand what they could have done better and to relate the practice of the projects with the theory.

Python Bootcamp planning

GENERIC GUIDELINES

1. The objective of the Bootcamp is to allow students to learn by practicing, as such all students are required to have a computer and practice
2. Instructors should present the slides with the “theory” and at the end of each “theory” section of the slides, ask students to do the exercises in empty notebooks:
 - a. Never give notebooks to students before the exercise is finished (at the end of the classes, instructors can make available to students the notebooks they created during the class). If necessary, in longer texts, to avoid students from taking long time to type, copy & paste the code of the cell to the Zoom chat
 - b. Use also an empty notebook to do the exercises so that students do not see the expected result before you write and explain it
 - c. Always comment code similarly to what is the notebook so that students get used to commenting on their own code (explain the difference between # and ”)
3. Instructors are free to change or complement the examples on the notebooks to answer questions or doubts posed by students

DAY 1

1. At the end of slides B.1:
 - a. Conduct an Anaconda Overview, including how to install packages
 - b. Open a Jupyter notebook and conduct a quick overview of how to use Jupyter and markdown (mention the Cheatsheets in the Moodle folder)
 - c. Create a “hello world” cell, save the notebook, close it, and open it again
 - i. Show how to consult or look for documentation (e.g., pandas: <https://pandas.pydata.org/docs/>)
 - d. Show Visual Code and Google Colab (explain the differences between them and why Colab can be a good collaboration tool)
2. Intercalate slides B.2 with the respective exercises of the notebook DataStructures.ipynb
3. Intercalate slides B.3 with the respective exercises of the notebook FlowControl.ipynb

DAY 2

1. Intercalate slides B.4 with the respective exercises of the notebook Functions.ipynb
2. Intercalate slides B.5 with the respective exercises of the notebook Libraries.ipynb
3. Intercalate slides B.6 with the respective exercises of the notebook IntroductionToNumpy.ipynb

Intercalate slides B.2 with the respective exercises of the notebook IntroductionToPandas.ipynb