DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

LAB 1: Linear Classification and Regression

Machine Learning 2021
Slides P. Zanuttigh

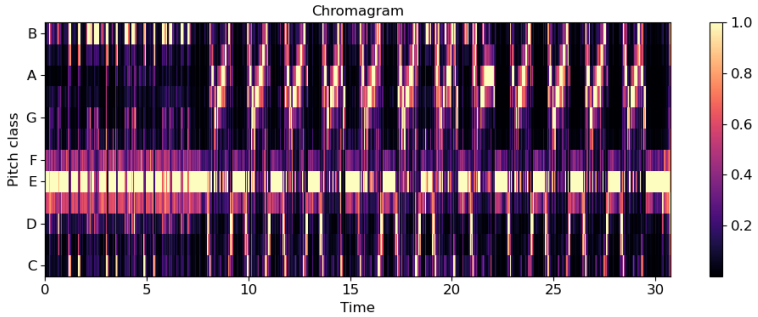
LAB 1: Linear Classification and Regression

There are 2 tasks:

	
Classification of music songs	*Regression on house prices data*

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Classification Task



For each song the training data contains 3 features

1. "*tempo*" (beats per minute)
2. "*chroma_stft*" (mean chromagram activation)
3. "*spectral_centroid*" (spectrum center of mass)

Task: classify into 2 classes, classical music and heavy metal

1. Load the data file, divide into train and test sets
2. Implement the Perceptron algorithm
3. Use Logistic Regression from Python libraries

Regression Task

For each house the training data contains several features (e.g., crime rate, amount of residential buildings, school stats…)
Task: Estimate the house price

1. Load the data file, divide into train and test sets
2. Use Least Square implementation in Python libraries

Complete the Notebook

- You have to complete the jupyter notebook:
  - classification problem (music songs, perceptron and logistic regression)
  - regression task (house prices, least square)
- FIRST THING TO DO: you need to put your name and ID number in the notebook
  - You can use the ID also as seed for random number generators
- The notebook has missing code: need to fill in what is missing
- You must write the answer to all the questions in the notebook
- You should also place some text/comments (to explain choices or describe results)
- Feel free to add cells with text if you need to explain or describe some "non-standard" decision!
  - But do not change the input data files, they will not be submitted

The Task

- Complete the jupyter notebook
  - i.e., write the code and answer to the questions
- Check that they run properly from the beginning with the provided data
  - use the `"restart kernel&run all"` command
- Save them as `surname_name_lab1.ipynb`
- Submit on elearning

Timeline

- Thu 28/10: Homework released
- Tue 2/11: Lab 1 (room De+Te+Ue+online)
- Mon 15/11: Delivery deadline (HARD deadline)
- The outcome is an on-off mark (i.e., +1 for the exam mark if the homework is reasonably done)