MACHINE LEARNING (IMC-4302C)

LAB 4: LOGISTIC REGRESSION (CONTINUATION)

SLIM BEN AMOR: SLIM.BEN-AMOR@INRIA.FR

**School Year
2017-2018**

16/2/2018

About Lab 4

- **Outline:**
  - Data pre-processing (cleaning, drop non relevant information, fill missing values…)
  - Train logistic classifier and predict outputs by using predefined functions in **sklearn Library**.
  - Use **Logistic Regression class** from **Linear Model module**.

→**Send your report and 2 labs code by mail before Sunday 03/04. Any submission after the correction will be penalized.**

→You may submit one report per group of 2-3 persons.

Assess classifier performance

- **Accuracy:**

$$accuracy = \frac{number\ of\ good\ prediction}{total\ number\ of\ sample}$$

- This measurement may inform about us how well the classifier prediction will be. But in general, accuracy is not very significant measurement .

- In the case of skewed data (all sample have y=0 or 1), accuracy could be high even with a naïve classifier that predict always 0.

→ Other assessments score will be studied in upcoming sessions

16/2/2018