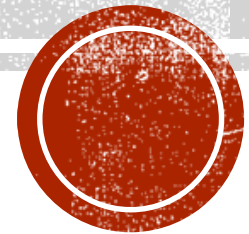


PRACTICAL ASSIGNMENT

Dealing with imbalanced datasets



EMPIRICAL STUDY ON IMBALANCED DATASETS

- Individual assignment
- Use Google Colab for your practical work.
- Datasets:
 - Boys or Girls
 - Another imbalance dataset of your own choice: Binary or multi-class
- Choose 2+ classifiers
- Methods for dealing with imbalanced datasets
- Original datasets vs “Processed” datasets
- **Deliverables: Shared link to your Google Colab and an ipynb file.**
- **Deadline: April the 3rd, 23:59**



RESEARCH QUESTIONS?

- What classifiers work the best with imbalanced datasets “off the shelf”?
- What methods best “offset” the effect of imbalanced datasets?
- Binary class vs Multi-class problems?
- Classification vs Regression?
- ... any questions that you want to investigate.



ORGANIZE YOUR WORK

- Experiment overview
 - What, Why and How
- Research Questions
- Flow of experiments
- Write your Google Colab as a tutorial for others.
- Cite references when appropriate.

