ML 4375 – Intro to Machine Learning

Homework: Machine Learning with Python

Worth 100 points

Instructions:

- Select one of your data sets from the R project. Create a Python notebook (Jupyter or CoLab) for this assignment with appropriate headings.
- Rework as much of your code as possible in Python for:
    - data cleaning
    - data exploration through code
    - data exploration through graphs
    - running at least 3 algorithms (try to use the same algorithms as your R project)
- Write a paragraph or more of your results analysis, comparing the algorithms in Python and their performance.
- Write a paragraph or more of your personal impressions of machine learning in R versus Python. Since this is opinion, the more opinion you share the higher your grade.

Turn in:
- Print-to-pdf of your notebook

Grading Rubric:

| Step | Points |
|---|---|
| 1. Perform similar data cleaning as you did in R to set up the data | 10 |
| 2. Perform similar data exploration functions | 30 |
| 3. Create some graphs that are descriptive of the data (at least 3) | 30 |
| 4. Run as many algorithms as you can (at least 3) | 40 |
| 5. Results analysis | 30 |
| 6. Comparison of Python and R for machine learning | 40 |
| 7. Project depth | 20 |
| Total points possible | 100 |