Programming Assignment 0: A Sample Classification Problem

Due September 4, 2023

In the repository you have access to tabular data with ten features $x_j \in \mathbb{R}$, for $j = 1, \ldots, 10$, and a binary label $y \in \{0, 1\}$, saved as .csv file under pa0/data/pa0_train.csv. You are tasked with constructing a model $\overline{y} : \mathbb{R}^{10} \to \{0, 1\}$ so that $\overline{y}(x) = y$ for as many pairs $(x, y) \in \mathbb{R}^{10} \times \{0, 1\}$ *as possible*. There is starter code in pa0/code/pa0_code.py which you may and should use to guide your development. In fact, your submission should be this code *as .py file* with all methods defined and the *main* function deleted, namely everything under

```
if __name__ == "__main__":
```

This function and subsequent function calls are there only for your guidance. You are encouraged to use a debugger and comment portions of the main function as you develop.

Your submission in canvas should report (accuracy, true positive rate, false positive rate, your prediction score (for class 1) average, and prediction score (for class 1) standard deviation) as an ordered tuple. You should also submit the .py file as

```
pa0_[lastname].py
```