Programming Assignment 0: A Sample Classification Problem

Due September 1, 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 each method fully 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. If you would like guidance on code, you may use Real Python's tutorial on binary classification for reference.

Your submission in canvas will contain **two** components:

1. You will report

```
(accuracy, tpr, fpr, mean(y_tilda), and stdev(y_tilda))
```

as an ordered tuple in the comments box on your submission page, where tpr denotes 'true positive rate', fpr 'false positive rate', and 'y_tilda' the prediction score for the positive class, and

2. You will also submit the .py file as

`pa0_[lastname].py`

Again, remember to delete the main function at the bottom of your code.