Challenge submission

Kai Puolamäki

22 November 2022

You should submit your answer to the term project challenge as a single file named `answers.csv`.

The first line of the file should contain only one real number, namely your estimate for binary accuracy.

The second line of the file should contain the following string:

```
class4,p
```

The remaining 965 lines of the file should contain two comma-separated values, namely your predicted class in `nonevent`, `Ia`, `Ib`, and `II`, and your predicted probability $P(\text{class2} = \text{event})$ for the respective lines in the `npf_test_hidden.csv`, where for each row in the data matrix `class2` is defined as follows:

$$\text{class2} = \begin{cases} \text{event} & , \quad \text{class4=Ia or class4=Ib or class4=II} \\ \text{nonevent} & , \quad \text{class4=nonevent} \end{cases} .$$

The attached file `dummy.csv` is an example of syntactically valid `answers.cvs` file. Here are the first ten lines of `dummy.csv`:

```
0.976932
class4,p
II,0.950482646119781
II,0.53541361738462
II,0.595561691676266
II,0.88846462406218
Ia,0.601566751836799
nonevent,0.201882255263627
Ia,0.723090314073488
Ia,0.80314773437567
```

We use R script `check.R` (or equivalent) to compute the performance metrics. Notice that you cannot yet run `check.R` because you need the file `npf_test.csv`, which contains the correct labels.

You must run the script `lint.R` on your `answers.csv` file before submitting it to the challenge. We will accept only files that pass the lint check!

```
% Rscript --vanilla lint.R dummy.csv
Everything seems to be ok.
```