

Haberman's Survival prediction

Please read the dataset description to understand the background and dataset.

Dataset

3 features, see the dataset description.

1 output: 1 = the patient survived 5 years or longer (negative class)

2 = the patient died within 5 year (positive class)

Training data: first 85%: row 1 to row 260

Testing data: row 260 to row 306

Task:

1. Using Maximum Likelihood Estimation with gradient ascent for training. Report true positive, false positive, false negative, true negative, precision, recall, and F1 score for the testing data.
2. Normalize the feature, report the results again. Any improvement?

1. We run plagiarism check for submitted code. Please don't look for solutions online.
2. You will not receive any credit if you directly use off-the-shelf machine learning tools. Mathematic computation tools such as Numpy and other basic tools, e.g., matplotlib are allowed to use.
3. Submit your code (.py file) and your report (no more than 2 pages, no format requirement, .pdf file). Not following the submission requirement, e.g., code in txt/pdf file, exceeding report page limit, etc. will receive a credit penalty.