

There are two datasets, crx.csv and data.csv, please use them to complete the requirements for this assignment below:

1. The labels column in crx and data are "label" and "Diagnosis"
2. Please implement the Linear Classifier from scratch with the update rule in the slide. It means you cannot adopt any existing package like sklearn in this assignment.
3. When the " $J=WX+b$ " could be represented as the matrix form for the linear classifier, please find the solution by solving this equation using least-squared manner. Also, please implement it and make a comparison between this method and the previous one implemented in 2.
4. Please implement the voted perception and make the performance comparison between methods.
5. With minimizing the $\|w\|^2$, it should drive the marginal to be maximized as well. Please implement the linear classifier with the minimum $\|w\|^2$ property and verify whether the margin of this version is larger than that of the conventional linear classifier or not.
6. Based on 5, please add the slack variable term in the linear classifier and find the most effective weighting value C .
7. Use any existing SVM package (libSVM, sklearn, or Matlab SVM), please make a performance comparison between the built-in SVM and your implementation.

All the source code should be uploaded to Github with public access. The performance comparison should be organized into a formal report with the standard template indicated in Moodle.