Introduction to Machine Learning
Lab 10: Primal SVM and Data Augmentation for Fairness

Hongteng Xu

May 20, 2022

1 Motivation

- Implement the learning methods of SVM in the primal form.
- Develop your own data augmentation strategy to suppress the unfairness issue we found in Lab 9.
- Fairness is a challenging problem in ML, and there is no free lunch. Get some feelings about how to pursue fairness, and how to achieve a trade-off between performance and fairness.

2 Tasks

Please read Lecture 12 carefully before doing this lab work.

1. **Primal SVM:** Implement the training and the testing methods of SVM. The training method is the stochastic gradient descent (SGD) that optimizes the primal form of SVM. (Hint: This is the second non-smooth optimization problem we learned in this course. Be careful about the (sub-)gradient of the objective function.)
2. **Data augmentation:** Develop a data augmentation strategy as a pre-processing of training data, such that the gender unfairness issue of the trained SVM can be suppressed. Note that 1) this is an open problem, try any method you think makes sense; 2) think about what you sacrifice when pursuing fairness. (Hint: the data augmentation method is motivated via the criterion used in the fairness checker you developed in Lab 9.)