Quiz 3: Supervised Learning and Unsupervised Learning

Nam Nguyen

September 2020

1 Statistical Learning

In this assignment, we will compare the differences of two main statistical learning methods: supervised and unsupervised. The main difference is that supervised learning requires labels in dataset, while unsupervised learning does not.
Supervised Learning: There two main tasks are regression and classification. The corresponding loss function for regression task is typically L_2 (MSE) and L_1 (MAE) loss. While in classification task, we usually use cross-entropy function.
Unsupervised Learning: We aim to cluster the data samples into groups. The loss function is usually a measure (distance) function. Typical methods are K-Nearest Neighboors (KNN), K-means, etc.
Self-supervised Learning: I would like to introduce my current research interest in PhD program. The underlying theory is that we pre-assign label for training task. There are many advancements in NLP, however, the SSL recently witnesses remarkable results in computer vision task also. For example, simCRL can achieve state-of-the-art result with only 1-10% of ImageNet.

References

[1] Trevor Hastie et al, "Statistical Learning"

[2] Chen et al, "A Simple Framework for Contrastive Learning of Visual Representations"