Short Story Assignment Proposal

Topic:

Image Classification with Deep Learning in the Presence of Noisy Labels: A Survey

Source:

https://arxiv.org/pdf/1912.05170.pdf

Abstract:

Advancement of deep neural networks has placed a major importance in Image Classification. But, they require huge amounts of labeled data to train the model. This is a very expensive process and now it has become a major problem and is practically very challenging. So, Label noise has become a common problem in many datasets and this paper suggests numerous methods to train deep networks with label noise. Data features, true label of data, and the labeler characteristics are the main factors which affect the label noise. According to these factors it is again classified into three subclasses namely Random noise, Y - Dependent noise and XY - Dependent noise which are discussed in the paper.

Source of label noise has been discussed. Firstly, there is a large amount of data available in the web. But this data may have noisy labels as they are given by the users. Secondly, multiple experts can label the data, but each expert has different experience levels which again leads to noisy labels. Sometimes, data is too complicated and even experts also will not be able to label them correctly. For example, Medical Imaging. Sometimes, the noisy labels are injected on purpose. In the presence of noise labels, we need to find the best estimator for hidden distribution which can be used to reverse the effect of noisy samples. Here, we use algorithms to try to find the noise structure and train the base classifier with estimated noise parameters. This survey paper throws light on those algorithms which are categorized into two subgroups. They are noise model based and noise model free methods.

Finally, this paper mainly concentrates on different types of noise model based and noise model free methods.