CSE5243 HW1

Xuyang Zhang

WF 11:10 AM

1) Clustering:

2) Noisy data:

Let’s say we want to build a classifier to classify Cats and Dogs. When training the classifier, the training dataset should only contain image of dogs and image of cats. If we don’t denoise our dataset before we feed the classification model, images that having dogs/cats and other animals on them will be quite confusing. The model may extract wrong features as classification criterion. The trained model may not maintain high accuracy when classify a new dataset.

Multimedia data:

multimedia data is usually unstructured. In the scenario of User Profiling, to analyze multimedia data, it is essential to transfer these unstructured in to well-structured and compatible data. If not, it is usually impossible to make full use of multimedia data and extract useful information.

3)

4) calculate density by using weight/volume

for X in unknow material lists:

calculate the absolute difference of the density and densities of given materials

find the minimum value D\_min among these absolute differences

find the material M corresponding to the minimum value D\_min

label X as material M