Machine Learning

Assignment 9.4

Submitted By: Ranji Raj

January 6, 2021

a) Active Learning for CB

Active Learning

- Active Learning (subset of machine learning) is a kind of query-based learning where the user queries for specific information based on a certain region of interest.

- The fundamental belief behind the active learner algorithm concept is that an ML algorithm could potentially reach a higher level of accuracy while using a smaller number of training labels if it were allowed to choose the data it wants to learn from.

Building CB based on AL

- Instead of storing each instance in the memory the learner can exploit from the hypothesis postulated by AL that user has the liberty to choose the data to learn from specific to the learning algorithm.

- So in comparison to regular ML algorithms where we randomly picked up the instances we can employ AL to do it more smartly in an interactive approach.

b) TWO examples for AL

- **Speech recognition**: Accurate labeling of speech utterances is time consuming and requires expertise. Annotation at word level can take ten times longer than actual audio, and annotating phonemes can take 400 times as long.

- **Document classification and filtering**: Learning to classify documents or any kind of media requires that users label each document or media file with particular labels, like *relevant* or *irrelevant.*

c) AL vs. IB2

Active Learning	IB2
Instances are queried interactively	Instances near neighborhood of the boundary line.
Accuracy is higher	Accuracy is lower as noise level increases
Storage requirements are less	Stores only missclassified instances
Uses both labeled and unlabeled data	Uses only labeled data

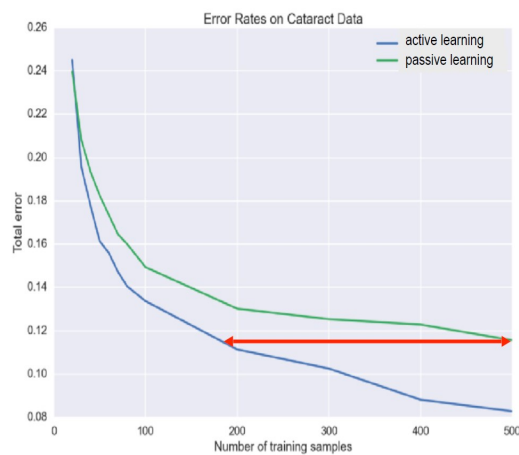Table 1: Differences of Active Learner with simple algorithms like IB2

Figure 1: Less than half as many labeled examples needed by active learning

References

[1] Burr Settles, *Active Learning Literature Survey*. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.