FML HW2 Zhe HUANG 22302183 IASD

Ex 7

(1) • The VC-dim of linear classifier here is 3.

Because in $\mathbb{R}^2$, the maximum number of points that can be shattered is 3. A line can shatter any combination of 3 points.

• Sauer lemma : $\left[\forall n \geq d,\ \Pi_F(n) \leq \sum_{i=1}^{n} \binom{n}{i} \leq \left(\frac{en}{d}\right)^d\right]$ for $F \subseteq \{-1, +1\}^X$ such that $VCdim(F) \leq d < +\infty$

This lemma tells us that there is an upper bound on the number of different ways the hypothesis class can perfectly seperate the data points. Also, as the VC dim increases, the number of different sets of data points that can be shattered grows, but it is bounded by a polynomial function of d.

(2) • We can't apply Halving's algorithm here (without any modification) because the number of all possible linear classifiers is infinite, and Halving relies on a finite hypothesis class.

• The number of classifiers would depend on the specific distribution and arrangement of data points in $\mathbb{R}^2$.

• The dataset is perfectly linearly seperable.
Thus we would make at most $\log_2$(number of classifiers) mistakes.

Ex 8. Assuming that we have unlimited computational resources while trying to limit the number of prediction errors as best as possible, we adapt Halving algo for the scenario where $|H|$ is infinite.

1. Initialization
   - Start with a distribution D over the infinite hypothesis class H. We can initialize this distribution uniformly, giving equal weight to all hypotheses.

2. Iteration
   - For each round, we sample a hypothesis h from current distribution D.
   - Observe the outcome of the chosen hypothesis h.
   - Update the weights of the hypotheses in H based on their performance. Hypotheses that perform well are assigned higher weights, while those that perform poorly are assigned lower weights.
   - Re-normalize the distribution D to make it a probability distribution

3. Prediction
   - To make predictions, we can use the weighted majority vote from the hypotheses in H, with each hypothesis's weight reflecting its performance in the past rounds.

4. Repeat
   - Repeat the iteration process for as many rounds as needed.