CECS 551 Statistical Learning Theory Exercises

Exercises

1. A **decision stump** is a decision tree that has a single query node. Let $\mathcal{F}$ denote the family of all decision stumps that classify points in $\mathcal{R}^2$, where a query is limited to either $x \geq a$? or $y \geq a$?, for arbitrary real number a. Determine VC($\mathcal{F}$).

2. A binary tree T is said to be **full** iff every internal node of T has two children. Prove that the number of leaves of a full tree is one more than the number of internal nodes.

3. Consider the family $\mathcal{F}_n$ of binary decision trees having n query nodes, and that classify one-dimensional data points. Each query node of a tree has the form $x \geq a$?, where a is a real number. Determine the VC($\mathcal{F}_n$). Defend your answer. How does this result carry over to higher dimensions?

4. A two-dimensional binary ellipse classifier e consists of a two-dimensional ellipse E. Moreover e classifies a two-dimensional vector $\overline{x}$ as $+1$ iff either $\overline{x}$ lies on the ellipse boundary or inside the ellipse. Otherwise it is classified as -1. If $\mathcal{F}$ denotes the family of two-dimensional binary ellipse classifiers, then determine VC($\mathcal{F}$). Defend your answer.

5. Prove that the set of hyperplane classifiers in $\mathcal{R}^n$ have a VC-dimension of at least $n + 1$. Hint: shatter the basis vectors of $\mathcal{R}^n$, plus one additional point.

6. Let $\mathcal{F}_n$ denote the family of all neural networks with n total weight/bias parameters, and that classify points in $\mathcal{R}^2$. Provide a good lower-bound for VC($\mathcal{F}_n$). Assume each neuron has a discrete activation function.

7. Assume model $\hat{f}$ has a training accuracy of 0.12 on a data set having size 10000. Moreover, $\hat{f}$ was selected from a family having a VC dimension that is at most 15. Provide an upper bound for the expected risk of $\hat{f}$ that is accurate with 90% certainty.