| | |
|---|---|
| **Machine Learning: Foundations** | **Spring Semester, 1996/7** |

Homework 3: April 9

*Lecturer: Yishay Mansour*

# Homework number 3.

**Theory question:**
Compute the VC dimension of the following concept classes.

1. The class of convex polygons in the plane with $d$ edges.
2. The class of all circles in the plane.
3. The class of decision trees over $\{0,1\}^d$ with at most $m$ nodes. (Try to get the tightest bound.)

**Programming assignment:**
Write a program to implement the decision tree learning algorithm. Use as a hypothesis at the nodes the predicates $x_i \geq \alpha$, where $x_i$ is an attribute and $\alpha$ a real value. Implement the three splitting criteria (have it as an external input, which one you are using). Have another parameter which tells you what is the minimal number of inputs in a leaf (once you reach this threshold you will not continue splitting). Run the program on **all** the data sets in `http://www.math.tau.ac.il/~mansour/ml-course`.

Answer the following questions:

1. Compare the three criteria.
2. How do the result change with the changing the threshold for the number of input in a leaf.
3. What happens when you increase the training set.

**The homework is due in two weeks**