

VC dimension of linear threshold networks

SORRY: I will not have time to type up this lecture.

Here is a summary.

1. We spent a lot of time constructing and then discussing a table with various complexity bounds for neural networks. In addition to the ones from last time, we also discussed the following bounds in the Anthony-Bartlett text: Theorem 3.4 (linear separators), Theorem 6.2 (linear threshold network), Theorem 7.1 (pathological activation function with infinite VC), Theorem 8.8 (networks with piecewise polynomial activations), Theorem 8.13 (networks with sigmoid activations).
2. We then gave a careful proof of Theorem 6.2. Main difference with bartlett version was maintaining a family of sign matrices with all outputs of all nodes for all inputs, and also making the induced partition on the parameter space explicit.

References