SL08. VC Dimensions

Introduction:

- Haussler Theorem states that number of examples *m*:

$$m \geq \frac{1}{\varepsilon}(\ln|H| + \ln\frac{1}{\delta})$$

- The problem is, Haussler Theorem breaks with infinite hypotheses spaces.
- If you have an infinite set of hypotheses, the above equation will be infinite.
- Linear separators, artificial neural networks and decision trees with continuous inputs all have infinite hypotheses spaces.
- Note that there is a notion of syntactic hypothesis space, which is anything that can be written as the hypothesis. However, there is also semantic hypothesis or only the meaningfully different hypothesis space.

Power of a Hypothesis Space:

- VC Dimension is the largest set of inputs that the hypothesis class can shatter (label in all possible ways).
- Interval training:
  - To prove that a specific VC dimension is possible, you need to prove that there's some set of points you can shatter.
  - To prove that a specific VC dimension is not possible, you need to prove that there's no example that can be shattered.
- For any *d*-dimensional hyperplane hypothesis class, the VC dimension will be *d+1*.
- If our hypothesis is points inside convex polygon, VC dimension is infinite.

Sample Complexity & VC Dimension:

- Adding the VC Dimension to Haussler Theorem:

$$m \geq \frac{1}{\varepsilon}(8.\mathrm{Vc(H)}.\log_2\frac{13}{\varepsilon} + 4.\log_2\frac{2}{\delta})$$

VC Dimension of Finite H:

- *H* is PAC-learnable if, and only if, VC dimension is finite.