VC Dimension

Martin Sewell
Department of Computer Science
University College London

September 2008

VC dimension (for Vapnik Chervonenkis dimension) measures the capacity of a hypothesis space. *Capacity* is a measure of complexity and measures the expressive power, richness or flexibility of a set of functions by assessing how wiggly its members can be. The definitions below are taken from Vapnik (1999).

Definition 1 THE VC DIMENSION OF A SET OF INDICATOR FUNCTIONS *(Vapnik and Chervonekis 1968; Vapnik and Chervonenkis 1971)*
The VC dimension of a set of indicator functions $Q(z, \alpha), \alpha \in \Lambda$, *is the maximum number h of vectors $z_1, \ldots, z_h$ that can be separated into two classes in all 2^h possible ways using functions of the set*[1] *(i.e., the maximum number of vectors that can be* shattered *by the set of functions). If for any n there exists a set of n vectors that can be shattered by the set $Q(z, \alpha), \alpha \in \Lambda$, then the VC dimension is equal to infinity.*

Definition 2 THE VC DIMENSION OF A SET OF REAL FUNCTIONS *(Vapnik 1979)*
Let $A \leq Q(z, \alpha) \leq B, \alpha \in \Lambda$, be a set of real functions bounded by constants A and B (A can be $-\infty$ and B can be ∞).

The indicator *of level β for the function $Q(z, \alpha)$ shows for which z the function $Q(z, \alpha)$ exceeds β and for which it does not. The function $Q(z, \alpha)$ can be described by the set of all its indicators.*

Let us consider along with the set of real functions $Q(z, \alpha)$, $\alpha \in \Lambda$, the set of indicators

$$I(z, \alpha, \beta) = \theta\{Q(z, \alpha) - \beta\}, \alpha \in \Lambda, \beta \in (A, B), \tag{1}$$

where $\theta(z)$ is the step function

$$\theta(z) = \begin{cases} 0 & \text{if } z < 0, \\ 1 & \text{if } z \geq 0. \end{cases} \tag{2}$$

[1] Any indicator function separates a given set of vectors into two subsets: the subset of vectors for which this indicator function takes the value 0 and the subset of vectors for which this indicator function takes the value 1.

The VC dimension of a set of real functions $Q(z, \alpha)$, $\alpha \in \Lambda$, *is defined to be the VC dimension of the set of corresponding indicators* (1) *with parameters* $\alpha \in \Lambda$ *and* $\beta \in (A, B)$.

Figure 1 below gives a simple example of how to calculate the VC dimension.

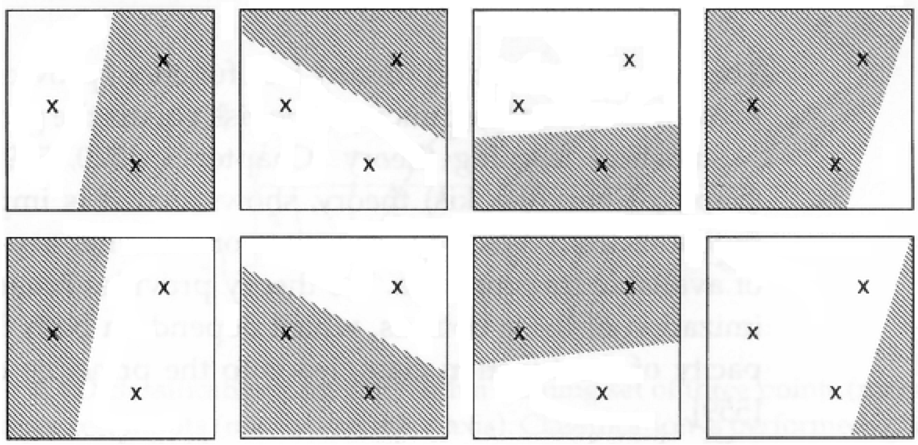

Figure 1: A simple VC dimension example. There are $2^3 = 8$ ways of assigning 3 points to two classes. For the displayed points in $\mathbb{R}^2$, all 8 possibilities can be realized using separating hyperplanes, in other words, the function class can shatter 3 points. This would not work if one was given 4 points, no matter how they were placed. Therefore, the VC dimension of the class of separating hyperplanes in $\mathbb{R}^2$ is 3. Schölkopf and Smola (2001)

References

SCHÖLKOPF, Bernhard, and Alexander J. SMOLA, 2001. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Adaptive Computation and Machine Learning. Cambridge, MA: The MIT Press.

VAPNIK, V., 1979. *Estimation of Dependences Based on Empirical Data [in Russian]*. Moscow: Nauka.

VAPNIK, Vladimir N., 1999. *The Nature of Statistical Learning Theory*. Second ed. Statistics for Engineering and Information Science. New York: Springer-Verlag.

VAPNIK, V. N., and A. Ja. CHERVONEKIS, 1968. On the uniform convergence of relative frequencies of events to their probabilities. *Doklady Akademii Nauk USSR*, **181**(4).

VAPNIK, V. N., and A. Ya. CHERVONENKIS, 1971. On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. *Theory of Probability and Its Applications*, **16**(2), 264–280. Translated by B. Seckler.