VC-dimension

Zdeněk Dvořák

November 30, 2023

Definition 1. *Let $\mathcal{F}$ be a set system. For a set X, we define $X \cap \mathcal{F} = \{F \cap X : F \in \mathcal{F}\}$. We say X is* broken *by $\mathcal{F}$ if $X \cap \mathcal{F}$ consists of all subsets of X.* Vapnik-Chervonenkis dimension *(*VC-dimension*) of the system $\mathcal{F}$ is the maximum integer k such that some set of size k is broken by $\mathcal{F}$.*

Examples:

- The system of all half-planes in $\mathbb{R}^2$ has VC-dimension 3. On one hand, there exists a 3-element broken set (in fact, any set of 3 points in general position is broken). On the other hand, consider any 4-element set X. If some $x \in X$ is contained in the convex hull of $X \setminus \{x\}$, then no half-plane intersects X in exactly $X \setminus \{x\}$. Otherwise, the points of X are in convex position, say in cyclic order x_1, x_2, x_3, x_4 in the boundary of the convex hull of X. Then no half-plane intersects X in exactly $\{x_1, x_3\}$.

- The system of all axis-aligned rectangles in the plane has VC-dimension at most 5: Consider any 6-element set X, and let x_1 be a leftmost point of X. Without loss of generality, at least three of the remaining five points in X have the y-coordinate at least as large as x_1; let X' be the set of these points. Let x_2 be a point in X' with the largest y-coordinate and let x_3 be the rightmost point of X'. Then every axis-aligned rectangle containing $\{x_1, x_2, x_3\}$ contains the whole X', and thus X is not broken.

- The system of all convex polygons in the plane has an infinite VC-dimension, since every set of points in convex position is broken.

Lemma 1. *Every set system $\mathcal{F}$ breaks at least $|\mathcal{F}|$ subsets of $\bigcup \mathcal{F}$.*

Proof. By induction on $|\mathcal{F}|$. The claim is trivial if $|\mathcal{F}| = 0$, and $\mathcal{F}$ breaks the empty set in $\mathcal{F} = 1$. Hence, we can assume $|\mathcal{F}| \geq 2$, and thus there exists an element $c \in \bigcup \mathcal{F}$ that does not belong to all sets of $\mathcal{F}$. Let $\mathcal{F}_1 = \{F \in \mathcal{F} : c \in F\}$ and $\mathcal{F}_2 = \{F \in \mathcal{F} : c \notin F\}$. By the choice of c we have $|\mathcal{F}_1| < |\mathcal{F}|$ and $|\mathcal{F}_2| < |\mathcal{F}|$. Observe that for any set X, if $c \in X$, then neither $\mathcal{F}_1$ nor $\mathcal{F}_2$ breaks X.

For $i = 1, 2$ let $\mathcal{R}_i$ be the system of subsets of $\bigcup \mathcal{F}_i$ that are broken by $\mathcal{F}_i$; by the induction hypothesis, we have $|\mathcal{R}_i| \geq |\mathcal{F}_i|$. Let $\mathcal{R}_3 = \{X \cup \{c\} : X \in \mathcal{R}_1 \cap \mathcal{R}_2\}$. By the observation at the end of the last paragraph, the sets in $\mathcal{R}_3$ are broken neither by $\mathcal{F}_1$ nor by $\mathcal{F}_2$, and thus $\mathcal{R}_3 \cap \mathcal{R}_1 = \emptyset$, $\mathcal{R}_3 \cap \mathcal{R}_2 = \emptyset$, and $|\mathcal{R}_3| = |\mathcal{R}_1 \cap \mathcal{R}_2|$. Furthermore, for each $X' \in \mathcal{R}_3$, the system $X' \cap \mathcal{F}_1$ consists of all subsets of X' containing c and $X' \cap \mathcal{F}_2$ consists of all subsets of X' not containing c, and thus X' is broken by $\mathcal{F}$. Every set broken by $\mathcal{F}_1$ or $\mathcal{F}_2$ is also broken by $\mathcal{F}$, and thus the number of sets broken by $\mathcal{F}$ is at least $|\mathcal{R}_1 \cup \mathcal{R}_2| + |\mathcal{R}_3| = (|\mathcal{R}_1| + |\mathcal{R}_2| - |\mathcal{R}_1 \cap \mathcal{R}_2|) + |\mathcal{R}_1 \cap \mathcal{R}_2| \geq |\mathcal{F}_1| + |\mathcal{F}_2| = |\mathcal{F}|$. $\square$

Corollary 2. *Let $\mathcal{F}$ be a set system of VC-dimension at most k. Every set X satisfies*

$$|X \cap \mathcal{F}| \leq \sum_{i=0}^{k} \binom{|X|}{i},$$

and thus if $k, |X| \geq 2$, then $|X \cap \mathcal{F}| \leq |X|^k$.

Proof. The VC-dimension of the system $X \cap \mathcal{F}$ is at most as large as the VC-dimension of $\mathcal{F}$, and thus it is at most k. So $X \cap \mathcal{F}$ can only break the subsets of X of size at most k, and there are $\sum_{i=0}^{k} \binom{|X|}{i}$ of them. Lemma 1 gives the required bound on $|X \cap \mathcal{F}|$. $\square$

Definition 2. *Let μ be a measure, let Y be a measurable set with $\mu(F)$ finite, and let $\mathcal{F}$ be a system of measurable sets. Let $\varepsilon > 0$ be a real number. Then $N \subseteq Y$ is an* ε-net *if every set $F \in \mathcal{F}$ such that $\mu(F \cap Y) \geq \varepsilon\mu(Y)$ satisfies $F \cap N \neq \emptyset$.*

Example: Let Y be an axis-aligned rectangle with sides of length 1 and let $\mathcal{F}$ be the system of all axis-aligned rectangles contained in Y. A set $N \subseteq Y$ is an ε-net iff N intersects every axis-aligned rectangle $D \subseteq Y$ of area at least ε. Note that every such rectangle has both sides of length at least ε. Hence, N can be chosen as the regular ε-spaced grid of points; then $|N| = \varepsilon^{-2}$. As we will see next, there exist asymptotically smaller ε-nets.

Theorem 3. *Let μ be a measure, let Y be a measurable set with $\mu(F)$ finite, and let $\mathcal{F}$ be a system of measurable sets of VC-dimension $k \geq 2$. Let $0 < \varepsilon \leq 1$ be a real number such that $k/\varepsilon \geq 15000$. Let N be a set of $\lceil 3\frac{k}{\varepsilon} \log \frac{k}{\varepsilon} \rceil$ points chosen independently from the probability distribution π on Y defined by $\pi(X) = \mu(X)/\mu(Y)$ for every measurable $X \subseteq Y$. Then N is an ε-net with a non-zero probability.*

Proof. Omitted; uses Corollary 2 in a tricky way. □

For a set system $\mathcal{F}$, let $\tau(\mathcal{F})$ denote the minumum size of a set $X \subseteq \bigcup \mathcal{F}$ intersecting all sets in $\mathcal{F}$ (for example, if the sets in $\mathcal{F}$ are edges of a graph, then $\tau(\mathcal{F})$ is the minimum size of a vertex cover). Let $\tau^\star(\mathcal{F})$ denote the fractional relaxation defined as the minimum of

$$\sum_{v \in \bigcup \mathcal{F}} x_v$$

subject to $x_v \geq 0$ for all $v \in \bigcup \mathcal{F}$ and

$$\sum_{v \in F} x_v \geq 1$$

for every $F \in \mathcal{F}$.

Corollary 4. *Let $\mathcal{F}$ be a set system with a finite union. If $\mathcal{F}$ has VC-dimension at most $k \geq 2$, then $\tau(\mathcal{F}) = O(k\tau^\star(\mathcal{F}) \log(k\tau^\star(\mathcal{F})))$.*

Proof. Let $Y = \bigcup \mathcal{F}$ and $\tau^\star = \tau^\star(\mathcal{F})$. Let x_v for $v \in Y$ are the values in an optimal solution to the linear program defining $\tau^\star(\mathcal{F})$. For $X \subseteq Y$, we let $\mu(X) = \sum_{v \in X} x_v$. Then μ is a measure on Y and $\mu(Y) = \tau^\star$. Furthermore, $\mu(F) \geq 1$ for every $F \in \mathcal{F}$. By Theorem 3, there exists a $1/\tau^\star$-net N of size $O(k\tau^\star \log(k\tau^\star))$; then N intersects all sets from $\mathcal{F}$, and $\tau(F) \leq |N|$. □

Example: For a given set $\mathcal{F}$ of axis-aligned rectangles in the plane and a finite set Y of points intersecting all of them, we want to find a smallest subset $S_{\text{opt}} \subseteq Y$ that intersects all rectangles in $\mathcal{F}$. By solving the linear program and using Corollary 4 (for the system $Y \cap \mathcal{F})$), we can at least find such a subset of size $O(|S_{\text{opt}}| \log |S_{\text{opt}}|)$.

Let G be a graph. For a vertex $v \in V(G)$ and a non-negative integer r, let $B_G(v, r)$ denote the set of vertices of G at distance at most r from v, and let $\mathcal{B}_G = \{B_G(v, r) : v \in V(G), 0 \leq r \leq |V(G)|\}$.

Lemma 5. *If G does not contain K_t as a minor, then $\mathcal{B}_G$ has VC-dimension at most $t - 1$.*

Proof. Suppose for a contradiction that some set $X = \{v_1, \ldots, v_t\} \subseteq V(G)$ is broken by $\mathcal{B}_G$. Hence, for $1 \leq i < j \leq t$ there exist vertices $v_{ij} \in V(G)$ and integers $r_{ij} \in \{1, \ldots, |V(G)|\}$ such that. $B_G(v_{ij}, r_{ij}) \cap X = \{v_i, v_j\}$. Choose such vertices v_{ij} so that the integers r_{ij} are minimum. Note that G contains shortest paths $P_{ij,1}$ and $P_{ij,2}$ from v_{ij} to v_i and v_j, and both of them have length at most r_{ij}. The minimality of r_{ij} implies that $P_{ij,1}$ and $P_{ij,2}$ intersect only in v_{ij}; their union P_{ij} consequently is a path from v_i to v_j containing v_{ij}. For $i > j$, let us set $P_{ij} = P_{ji}$, $v_{ij} = v_{ji}$, and $r_{ij} = r_{ji}$.

Suppose $y \neq v_{ij}$ appears on P_{ij} between v_i and v_{ij}. We claim that $d(v_i, y) < d(v_j, y)$: otherwise, if $d(v_j, y) \leq d(v_i, y) = r$, then note that $r < r_{ij}$, since the path $P_{ij,1}$ from v_{ij} to v_i has length at most r_{ij}. But $B_G(y, r) \cap X = \{v_i, v_j\}$, contradicting the minimality of r_{ij}. It follows that

($\star$) all vertices of P_{ij} before v_{ij} are strictly closer to v_i than to v_j, and symmetrically, all those after v_{ij} are closer to v_j than to v_i.

Suppose x is a vertex belonging to the intersection of paths $P_{i_1j_1}$ a $P_{i_2j_2}$ for some $\{i_1, j_1\} \neq \{i_2, j_2\}$. By symmetry, we can assume that $d(x, v_{i_s}) \leq d(x, v_{j_s})$ for $s \in \{1, 2\}$, and ($\star$) implies that x lies on a shortest path from $v_{i_sj_s}$ to v_{i_s}; therefore, $d(v_{i_sj_s}, x) + d(x, v_{i_s}) = d(v_{i_sj_s}, v_{i_s})$. By symmetry, we can also assume that $d(x, v_{i_1}) \leq d(x, v_{i_2})$. Using the triangle inequality, we have

$$d(v_{i_2j_2}, v_{i_1}) \leq d(v_{i_2j_2}, x)+d(x, v_{i_1}) \leq d(v_{i_2j_2}, x)+d(x, v_{i_2}) = d(v_{i_2j_2}, v_{i_2}) \leq r_{i_2j_2}.$$

Therefore $v_{i_1} \in B_G(v_{i_2j_2}, r_{i_2j_2}) \cap X = \{v_{i_2}, v_{j_2}\}$. If $d(x, v_{i_2}) < d(x, v_{j_2})$, then $d(x, v_{i_1}) < d(x, v_{j_2})$, and thus $i_1 \neq j_2$; it follows that $i_1 = i_2$. If $d(x, v_{i_2}) = d(x, v_{j_2})$, then we can assume $i_1 = i_2$ by symmetry. Let $i = i_1 = i_2$. If $d(x, v_i) = d(x, v_{j_s})$ for some $s \in \{1, 2\}$, then

$$d(v_{ij_{3-s}}, v_{j_s}) \leq d(v_{ij_{3-s}}, x)+d(x, v_{j_s}) = d(v_{ij_{3-s}}, x)+d(x, v_i) = d(v_{ij_{3-s}}, v_i) \leq r_{ij_{3-s}},$$

implying that $v_{j_s} \in B_G(v_{ij_{3-s}}, r_{ij_{3-s}}) \cap X$, which is a contradiction. Therefore,

($\star\star$) if $x \in V(P_{i_1j_1}) \cap V(P_{i_2j_2})$, then the labels can be chosen so that $i_1 = i_2$ and $d(x, v_{i_s}) < d(x, v_{j_s})$ for $s \in \{1, 2\}$.

For $i = 1, \ldots, t$, let

$$X_i = \{x \in V(P_{ij}) : j \in [t] \backslash \{i\}, d(x, v_i) \leq d(x, v_j), \text{and if } d(x, v_i) = d(x, v_j), \text{ then } i < j\}.$$

By ($\star$), the sets X_i induce connected subgraphs of G, and by ($\star\star$) the sets X_i are pairwise disjoint. By contracting them we obtain a minor of K_t in G, which is a contradiction. □

Let us consider the following generalization of the dominating set. For any function $r : V(G) \to \mathbf{Z}_0^+$, let $\mathrm{dom}_r(G)$ denote the minimum size of a set $X \subseteq V(G)$ such that $d(v, X) \leq r(v)$ for every $v \in V(G)$.

Corollary 6. *For every positive integer t, there exists a polynomial-time algorithm that for every K_t-minor-free graph G and every function $r : V(G) \to \mathbf{Z}_0^+$ returns a real number d satisfying $d \leq \mathrm{dom}_r(G) = O(td \log(td))$.*

Proof. Consider the set system $\mathcal{F} = \{B_G(v, r(v)) : v \in V(G)\}$. Since $\mathcal{F} \subseteq \mathcal{B}_G$, Lemma 5 implies that $\mathcal{F}$ has VC-dimension at most $t-1$. Note also that $\mathrm{dom}_r(G) = \tau(\mathcal{F})$. By Corollary 4 we can let $d = \tau^\star(\mathcal{F})$ (the number d can be determined in polynomial time by solving the linear program). $\square$