Range Spaces

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.

(so $\mathcal{R}$ is a family of subsets of X).

Range Spaces

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.
(so $\mathcal{R}$ is a family of subsets of X).

Examples:

$(\mathbb{R}, \mathcal{I})$, where $\mathcal{I}$ is the set of all closed intervals

$(\mathbb{R}^2, \mathcal{D})$, where $\mathcal{D}$ is the set of all closed disks

$(\mathbb{R}^2, \mathcal{T})$, where $\mathcal{T}$ is the set of all triangles

$(\mathbb{R}^2, \mathcal{AR})$, where $\mathcal{AR}$ is the set of all axis-aligned rectangles

$(\mathbb{R}^2, \mathcal{GR})$, where $\mathcal{GR}$ is the set of all general (i.e. arbitrarily oriented) rectangles

$(\mathbb{R}^2, \mathcal{H})$, where $\mathcal{H}$ is the set of all closed halfplanes

$(\mathbb{R}^2, \mathcal{C})$, where $\mathcal{C}$ is the set of all closed convex sets in the plane

Range Spaces

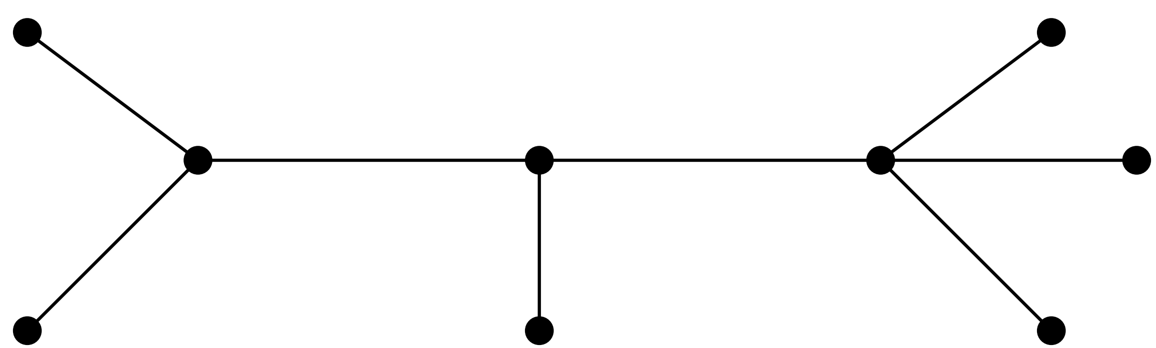
Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.

(so $\mathcal{R}$ is a family of subsets of X).

Examples:

Let T be a tree with vertex set V

$(V, \mathcal{S})$, where $\mathcal{S}$ comprises all sets that are vertex sets of subtrees (connected subgraphs) of T



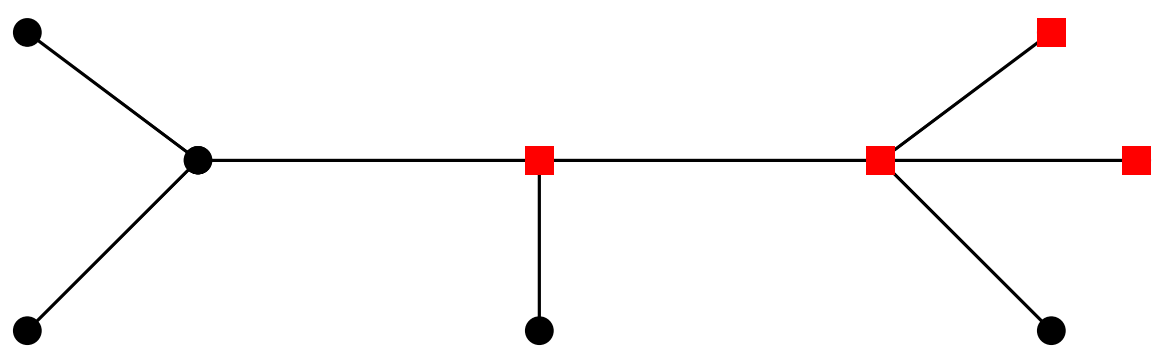
Range Spaces

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.
(so $\mathcal{R}$ is a family of subsets of X).

Examples:

Let T be a tree with vertex set V

$(V, \mathcal{S})$, where $\mathcal{S}$ comprises all sets that are vertex sets of subtrees (connected subgraphs) of T



Range Spaces

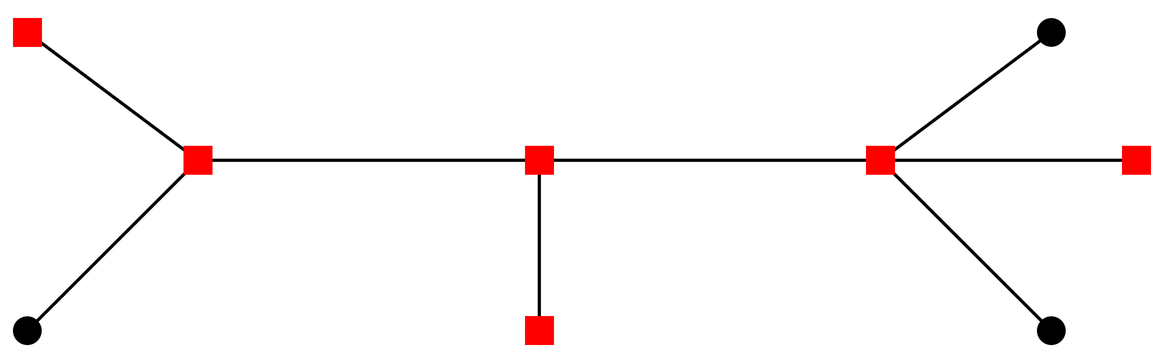
Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.

(so $\mathcal{R}$ is a family of subsets of X).

Examples:

Let T be a tree with vertex set V

$(V, \mathcal{S})$, where $\mathcal{S}$ comprises all sets that are vertex sets of subtrees (connected subgraphs) of T



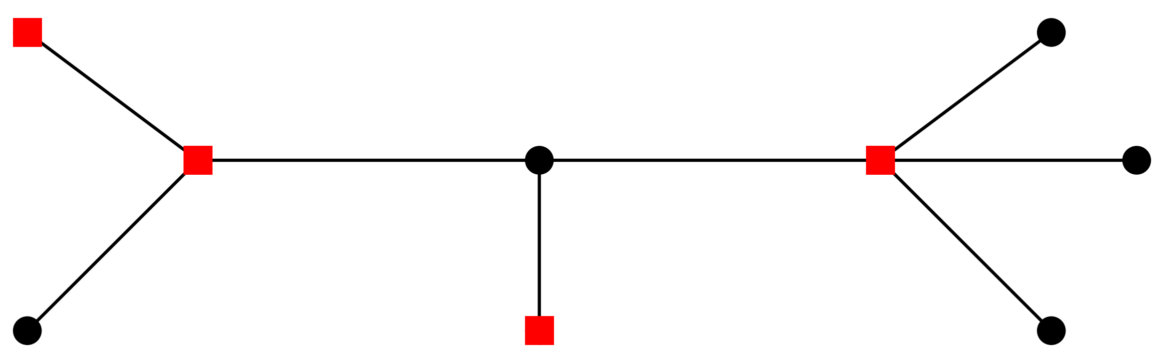
Range Spaces

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.
(so $\mathcal{R}$ is a family of subsets of X).

Examples:

Let T be a tree with vertex set V

$(V, \mathcal{S})$, where $\mathcal{S}$ comprises all sets that are vertex sets of subtrees (connected subgraphs) of T



not in $\mathcal{S}$

Range Spaces

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.
(so $\mathcal{R}$ is a family of subsets of X).

Examples:

$(\mathbb{R}^d, \mathcal{H}_d)$, where $\mathcal{H}_d$ is the set of all closed halfspaces in $\mathbb{R}^d$

$(\mathbb{R}^d, \mathcal{B}_d)$, where $\mathcal{B}_d$ is the set of all closed balls in $\mathbb{R}^d$

$(\mathbb{R}^d, \mathcal{S}_d)$, where $\mathcal{S}_d$ is the set of all closed simplices in $\mathbb{R}^d$

$\vdots$

Range Spaces

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.

(so $\mathcal{R}$ is a family of subsets of X).

$A \subset X$: $\mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$(A, \mathcal{R}_{|A})$ is the range space induced (projected) by $(X, \mathcal{R})$ on A

Range Spaces, Shattering

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.
(so $\mathcal{R}$ is a family of subsets of X).

$A \subset X$: $\mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$(A, \mathcal{R}_{|A})$ is the range space induced (projected) by $(X, \mathcal{R})$ on A

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-Dimension

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.

(so $\mathcal{R}$ is a family of subsets of X).

$A \subset X$: $\quad \mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$(A, \mathcal{R}_{|A})$ is the range space induced (projected) by $(X, \mathcal{R})$ on A

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

The *VC-dimension* of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

VC-Dimension

Range Space: Pair $(X, \mathcal{R})$, where X is a set and $\mathcal{R} \subset 2^X$.
(so $\mathcal{R}$ is a family of subsets of X).

$A \subset X$: $\quad \mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$(A, \mathcal{R}_{|A})$ is the range space induced (projected) by $(X, \mathcal{R})$ on A

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

The *VC-dimension* of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

"VC" ... Vapnik-Chervonenkis

VC-Dimension, Examples

$A \subset X$: $\mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

$(\mathbb{R}, \mathcal{I})$, where $\mathcal{I}$ is the set of all closed intervals

VC-Dimension, Examples

$A \subset X$: $\qquad \mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

Let T be a tree with vertex set V

$(V, \mathcal{S})$, where $\mathcal{S}$ comprises all sets that are vertex sets of subtrees (connected subgraphs) of T

VC-Dimension, Examples

$A \subset X$: $\mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

$(\mathbb{R}^2, \mathcal{D})$, where $\mathcal{D}$ is the set of all closed disks

VC-Dimension, Examples

$A \subset X$: $\quad \mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

$(\mathbb{R}^2, \mathcal{H})$, where $\mathcal{H}$ is the set of all closed halfplanes

VC-Dimension, Examples

$A \subset X$: $\mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

$(\mathbb{R}^2, \mathcal{C})$, where $\mathcal{C}$ is the set of all closed convex sets in the plane

VC-Dimension, Examples

$A \subset X$: $\quad \mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

$(\mathbb{R}^d, \mathcal{H}_d)$, where $\mathcal{H}_d$ is the set of all closed halfspaces in $\mathbb{R}^d$

VC-Dimension, Examples

$A \subset X$: $\mathcal{R}_{|A} = \{r \cap A | R \in \mathcal{R}\}$

$A \subset X$ is *shattered by* $\mathcal{R}$ iff $\mathcal{R}_{|A} = 2^A$

VC-dimension of $(X, \mathcal{R})$ is the cardinality of the largest $A \subset X$ that is shattered by $\mathcal{R}$

$(\mathbb{R}^d, \mathcal{H}_d)$, where $\mathcal{H}_d$ is the set of all closed halfspaces in $\mathbb{R}^d$

Radons Theorem Any set A of $d + 2$ points in $\mathbb{R}^d$ can be partitioned into two non-empty sets A' and A'' whose convex hulls intersect.

Sauer's Lemma

Sauer's Lemma For every range space $(X, \mathcal{R})$ of VC-dimension d and with $|X| = n$ we have $|\mathcal{R}| \leq \Phi_d(n) = \binom{n}{\leq d} = \sum_{0 \leq i \leq d} \binom{n}{i}$.

Sauer's Lemma

Sauer's Lemma For every range space $(X, \mathcal{R})$ of VC-dimension d and with $|X| = n$ we have $|\mathcal{R}| \leq \Phi_d(n) = \binom{n}{\leq d} = \sum_{0 \leq i \leq d} \binom{n}{i}$.

Immediate consequence of the following:

Lemma For every finite range space $(X, \mathcal{R})$ the number s of sets shattered by $\mathcal{R}$ is at least $|\mathcal{R}|$.

Sauer's Lemma

Sauer's Lemma For every range space $(X, \mathcal{R})$ of VC-dimension d and with $|X| = n$ we have $|\mathcal{R}| \leq \Phi_d(n) = \binom{n}{\leq d} = \sum_{0 \leq i \leq d} \binom{n}{i}$.

Immediate consequence of the following:

Lemma For every finite range space $(X, \mathcal{R})$ the number s of sets shattered by $\mathcal{R}$ is at least $|\mathcal{R}|$.

Since VC-dimension d means $s \leq \binom{n}{\leq d}$ and hence

$$|\mathcal{R}| \leq s \leq \binom{n}{\leq d}.$$

Sauer's Lemma

Sauer's Lemma For every range space $(X, \mathcal{R})$ of VC-dimension d and with $|X| = n$ we have $|\mathcal{R}| \leq \Phi_d(n) = \binom{n}{\leq d} = \sum_{0 \leq i \leq d} \binom{n}{i}$.

Immediate consequence of the following:

Lemma For every finite range space $(X, \mathcal{R})$ the number s of sets shattered by $\mathcal{R}$ is at least $|\mathcal{R}|$.
Proof by induction on $|\mathcal{R}|$

True for $|\mathcal{R}| = 1$ as the empty set is shattered.

For $|\mathcal{R}| > 1$ choose some x that is in some but not all ranges in $\mathcal{R}$ and split $\mathcal{R}$ into $\mathcal{R}^+$ (the ranges that contain x) and $\mathcal{R}^-$ (the ranges that do not contain x)

Shattering Function and Shattering Dimension

$S = (X, \mathcal{R})$ range space. Its shatter function

$$\pi_S(m) = \max_{B \in \binom{X}{m}} \left| \mathcal{R}_{|B} \right|$$

The **shattering dimension** of S is the smallest d such that $\pi_s(m) = O(m^d)$.

Shattering Function and Shattering Dimension

$S = (X, \mathcal{R})$ range space. Its shatter function

$$\pi_S(m) = \max_{B \in \binom{X}{m}} \left|\mathcal{R}_{|B}\right|$$

The **shattering dimension** of S is the smallest d such that $\pi_s(m) = O(m^d)$.

Lemma For any range space $S = (X, \mathcal{R})$ its shattering-dimension is at most as large as its VC-dimension.

Shattering Function and Shattering Dimension

$S = (X, \mathcal{R})$ range space. Its shatter function

$$\pi_S(m) = \max_{B \in \binom{X}{m}} \left|\mathcal{R}_{|B}\right|$$

The **shattering dimension** of S is the smallest d such that $\pi_s(m) = O(m^d)$.

Lemma For any range space $S = (X, \mathcal{R})$ its shattering-dimension is at most as large as its VC-dimension.

Observation The shattering dimension of a family of geometric shapes (e.g. disks) is bounded by the number of points necessary to determine the shape.

Dual Range Space and Dual Shatter Function

$S = (X, \mathcal{R})$ range space

For $p \in X$ define $R_p = \{r \in \mathcal{R} | p \in r\}$.

Dual range space for S is defined as

$$S^* = (\mathcal{R}, \{R_p | p \in X\})$$

Dual Range Space and Dual Shatter Function

$S = (X, \mathcal{R})$ range space

For $p \in X$ define $R_p = \{r \in \mathcal{R} | p \in r\}$.

Dual range space for S is defined as

$$S^* = (\mathcal{R}, \{R_p | p \in X\})$$

Dual shatter function $\pi^*_S(m) = \pi_{S^*}(m)$.

Dual shatter dimension of S is the shatter dimension of S^*.

Dual Range Space and Dual Shatter Function

$S = (X, \mathcal{R})$ range space

For $p \in X$ define $R_p = \{r \in \mathcal{R} | p \in r\}$.

Dual range space for S is defined as

$$S^* = (\mathcal{R}, \{R_p | p \in X\})$$

Dual shatter function $\pi_S^*(m) = \pi_{S^*}(m)$.

Dual shatter dimension of S is the shatter dimension of S^*.

Lemma If range space S has VC-dimension d then its dual space S^* has VC-dimension at most 2^d.

Composing Shapes

Lemma If range space $S = (X, \mathcal{R})$ has VC-dimension d then its complementary space $\overline{S} = (X, \overline{\mathcal{R}})$ also has VC-dimension d, where $\overline{\mathcal{R}} = \{\overline{r} | r \in \mathcal{R}\}$.

Composing Shapes

Lemma If range space $S = (X, \mathcal{R})$ has VC-dimension d then its complementary space $\overline{S} = (X, \overline{\mathcal{R}})$ also has VC-dimension d, where $\overline{\mathcal{R}} = \{\overline{r} | r \in \mathcal{R}\}$.

Lemma Let $S = (X, \mathcal{R})$ and $S' = (X, \mathcal{R}')$ be range spaces of VC-dimension $d > 1$ and $d' > 1$, respectively.
Then the space $(X, \widehat{\mathcal{R}})$ with $\widehat{\mathcal{R}} = \{r \cup r' | r \in \mathcal{R},\ r' \in \mathcal{R}'\}$ has VC-dimension $O((d + d') \log(d + d'))$.

Composing Shapes

Lemma If range space $S = (X, \mathcal{R})$ has VC-dimension d then its complementary space $\overline{S} = (X, \overline{\mathcal{R}})$ also has VC-dimension d, where $\overline{\mathcal{R}} = \{\overline{r} | r \in \mathcal{R}\}$.

Lemma Let $S = (X, \mathcal{R})$ and $S' = (X, \mathcal{R}')$ be range spaces of VC-dimension $d > 1$ and $d' > 1$, respectively.
Then the space $(X, \widehat{\mathcal{R}})$ with $\widehat{\mathcal{R}} = \{r \cup r' | r \in \mathcal{R},\ r' \in \mathcal{R}'\}$ has VC-dimension $O((d + d') \log(d + d'))$.

Lemma Let $S = (X, \mathcal{R})$ and $S' = (X, \mathcal{R}')$ be range spaces of VC-dimension $d > 1$ and $d' > 1$, respectively.
Then the space $(X, \widetilde{\mathcal{R}})$ with $\widetilde{\mathcal{R}} = \{r \cap r' | r \in \mathcal{R},\ r' \in \mathcal{R}'\}$ has VC-dimension $O((d + d') \log(d + d'))$.

Composing Shapes

Lemma If range space $S = (X, \mathcal{R})$ has VC-dimension d then its complementary space $\overline{S} = (X, \overline{\mathcal{R}})$ also has VC-dimension d, where $\overline{\mathcal{R}} = \{\overline{r} | r \in \mathcal{R}\}$.

Lemma Let $S = (X, \mathcal{R})$ and $S' = (X, \mathcal{R}')$ be range spaces of VC-dimension $d > 1$ and $d' > 1$, respectively.
Then the space $(X, \widehat{\mathcal{R}})$ with $\widehat{\mathcal{R}} = \{r \cup r' | r \in \mathcal{R},\ r' \in \mathcal{R}'\}$ has VC-dimension $O((d + d') \log(d + d'))$.

Lemma Let $S = (X, \mathcal{R})$ and $S' = (X, \mathcal{R}')$ be range spaces of VC-dimension $d > 1$ and $d' > 1$, respectively.
Then the space $(X, \widetilde{\mathcal{R}})$ with $\widetilde{\mathcal{R}} = \{r \cap r' | r \in \mathcal{R},\ r' \in \mathcal{R}'\}$ has VC-dimension $O((d + d') \log(d + d'))$.

Consequence Any finite sequence of combining range spaces of finite VC-dimension results in a range space of finite VC-dimension.

ε-Samples

$S = (X, \mathcal{R})$ is a range space; $B \subset X$ finite; $0 < \varepsilon \leq 1$.

A subset $C \subseteq B$ is an ε-**sample** for B iff for every range $r \in \mathcal{R}$ we have

$$\left| \frac{|B \cap r|}{|B|} - \frac{|C \cap r|}{|C|} \right| \leq \varepsilon$$

ε-Samples

$S = (X, \mathcal{R})$ is a range space; $B \subset X$ finite; $0 < \varepsilon \leq 1$.

A subset $C \subseteq B$ is an ε-**sample** for B iff for every range $r \in \mathcal{R}$ we have

$$\left| \frac{|B \cap r|}{|B|} - \frac{|C \cap r|}{|C|} \right| \leq \varepsilon$$

ε-**Sample Theorem** (Vapnik, Chervonenkis) There is a constant $c > 0$ so that if $(X, \mathcal{R})$ is a range space of VC-dimension at most d and if B is a finite subset of X, then for every $0 < \varepsilon, \delta < 1$ a randomly chosen subset B of s elements, where s is at least the minimum of $|B|$ and of

$$\frac{c}{\varepsilon^2}\left(d \log \frac{d}{\varepsilon} + \log \frac{1}{\delta}\right)$$

fails to be an ε-sample for B with probability at most δ.

ε-Nets

$S = (X, \mathcal{R})$ is a range space; $B \subset X$ finite; $0 < \varepsilon \leq 1$.

A subset $N \subseteq B$ is an ε-**net** for B iff every range $r \in \mathcal{R}$ with $|r \cap B| \geq \varepsilon |B|$ contains a point of N, i.e. $r \cap N \neq \emptyset$.

ε-Nets

$S = (X, \mathcal{R})$ is a range space; $B \subset X$ finite; $0 < \varepsilon \leq 1$.

A subset $N \subseteq B$ is an ε-**net** for B iff every range $r \in \mathcal{R}$ with $|r \cap B| \geq \varepsilon |B|$ contains a point of N, i.e. $r \cap N \neq \emptyset$.

ε-**Net Theorem** (Haussler, Welzl) Let $(X, \mathcal{R})$ be a range space of VC-dimension at most d and let B be a finite subset of X, and let $0 < \varepsilon, \delta < 1$. A set N is obtained by m random independent draws from B with

$$m \geq \max\left\{\frac{4}{\varepsilon}\log\frac{2}{\delta}, \frac{8d}{\varepsilon}\log\frac{8d}{\varepsilon}\right\}$$

fails to be an ε-net for B with probability at most δ.

Application of ε-Nets

Given a set S of n points in $\mathbb{R}^3$ and some $0 < \varepsilon < 1$ find a ball that fails to contain at most and ε fraction of the points of S.

Weak ε-Nets

$S = (X, \mathcal{R})$ is a range space; $B \subset X$ finite; $0 < \varepsilon \leq 1$.

A subset $N \subseteq B$ is an ε-**net** for B iff every range $r \in \mathcal{R}$ with $|r \cap B| \geq \varepsilon|B|$ contains a point of N, i.e. $r \cap N \neq \emptyset$.

$S = (X, \mathcal{R})$ is a range space; $B \subset X$ finite; $0 < \varepsilon \leq 1$.

A subset $N \subseteq X$ is a weak ε-**net** for B iff every range $r \in \mathcal{R}$ with $|r \cap B| \geq \varepsilon|B|$ contains a point of N, i.e. $r \cap N \neq \emptyset$.

Weak ε-Nets for Convex Sets in the Plane

$(\mathbb{R}^2, \mathcal{C})$, where $\mathcal{C}$ is the set of all closed convex sets in the plane

There are no constant size ε-nets for this space,
but there is a weak ε-net of size $O(1/\varepsilon^2)$.