Exercice 1 Soit Π la classe des concepts *rectangle axe-parallele* de R^2. En cours on a demontré que VC-dim(Π) $\leq$ 4.

1. Montrer par un exemple à 4 points que VC-dim(Π) $\geq$ 4.
2. Refaire l'analyse vue en cours pour borner le nombre N d'exemples a tirer pour retournerun rectangle hypothèse h à la place du concept cible c, tel que $P[\text{Aire}(c\Delta h) \leq \epsilon] \geq 1 - \delta$.

Correction :

Pour montrer que VC-dim(Π) ≥ 4, nous pouvons donner un exemple de 4 points dans le plan qui peut être brisée par un concept rectangle axe-parallèle de R2, mais pas par un concept rectangle axe-parallèle plus petit.

Considérons les 4 points suivants dans le plan :

A = (0, 0)

B = (1, 0)

C = (0, 1)

D = (1, 1)

Nous pouvons voir que ces 4 points peuvent être brisés par un concept rectangle axe-parallèle de R2. Par exemple, si nous choisissons un rectangle hypothèse h qui couvre les points A, B, C et D, alors h peut représenter un concept rectangle axe-parallèle qui les englobe tous.

Ainsi, nous avons montré un exemple d'un ensemble de 4 points qui peut être brisé par un concept rectangle axe-parallèle de R2, ce qui implique que VC-dim(Π) ≥ 4.

Pour borner le nombre N d'exemples à tirer pour retourner un rectangle hypothèse h à la place du concept cible c, nous pouvons utiliser la borne de Hoeffding, comme vu en cours.
La borne de Hoeffding est donnée par l'inégalité suivante :
P[|ε(c, h) – ε(c, h)| >   ] ≤ 2e^(-2Nε^2)

où ε(c, h) est l'erreur de généralisation entre le concept cible c et le rectangle hypothèse h, ε est une petite constante positive qui mesure la marge d'erreur tolérée, et N est le nombre d'exemples d'apprentissage tirés.

Nous pouvons utiliser cette inégalité pour déterminer le nombre N d'exemples à tirer pour garantir que P[Aire(cΔh) ≤   ] ≥ 1 – δ, où δ est la probabilité d'erreur tolérée.

En substituant Aire(cΔh) pour ε(c, h) dans l'inégalité de Hoeffding, nous pouvons obtenir une borne sur le nombre d'exemples N nécessaire pour atteindre la probabilité de succès souhaitée. Le calcul précis de cette borne dépendra des valeurs spécifiques de ε, ε, et δ, ainsi que des propriétés du concept cible c et du rectangle hypothèse h.

Exercice 2 Soit $\mathcal{B}$ la classe de concepts *boite axe-parallele* de R^d et soit $\mathcal{I}_k$ la classes de concepts *union de k intervals de la droire réelle*.

1. Calculer VC-dim(B).
2. Calculer VC-dim(I_1) et VC-dim(I_2), ensuite calculer VC-dim(I_k) pour tout k.
3. Etant données N examples positives et negatives, comment calculer un concept-hypotèse de I_k (respectivement, de B) compatible avec les examples ?

Correction :

Pour calculer VC-dim(B), nous devons déterminer le nombre maximum de points que la classe de concepts boîte axe-parallèle B peut shatter, c'est-à-dire séparer en toutes les manières possibles. La VC-dimension de B est égale à la dimension de l'espace plus un, c'est-à-dire VC-dim(B) = d + 1, où d est la dimension de l'espace des points (dans ce cas, R^d). Donc, VC-dim(B) = d + 1.

Pour calculer VC-dim(I1) et VC-dim(I2), nous devons déterminer le nombre maximum de points que les classes de concepts union d'un intervalle (I1) et union de deux intervalles (I2) peuvent shatter.

VC-dim(I1): Pour la classe de concepts I1, qui est l'union d'un seul intervalle, il est possible de shatter au plus 3 points. Par exemple, on peut shatter les points {-1, 0, 1} en attribuant un intervalle [-1, 1] pour les points positifs et un intervalle [0, 0] pour les points négatifs. Donc, VC-dim(I1) = 3.

VC-dim(I2): Pour la classe de concepts I2, qui est l'union de deux intervalles, il est possible de shatter au plus 4 points. Par exemple, on peut shatter les points {-1, -0.5, 0.5, 1} en attribuant un intervalle [-1, 1] pour les points positifs et deux intervalles [-0.5, -0.5] et [0.5, 0.5] pour les points négatifs. Donc, VC-dim(I2) = 4.

Pour VC-dim(Ik) pour tout k, on peut utiliser le même raisonnement. VC-dim(Ik) serait égal à k + 2, car il serait possible de shatter k + 2 points en attribuant un intervalle à k points positifs et k + 1 intervalles à k + 1 points négatifs.

Pour calculer un concept-hypothèse compatible avec les exemples positifs et négatifs pour la classe de concepts Ik (respectivement, B), nous pouvons utiliser divers algorithmes d'apprentissage supervisé, tels que le Perceptron, le SVM, ou le Decision Tree, adaptés pour traiter des intervalles d'union comme des exemples de données. Ces algorithmes peuvent être entraînés sur les exemples positifs et négatifs pour apprendre un concept-hypothèse qui sépare les données selon les intervalles spécifiés dans la classe de concepts Ik ou B. Une fois que le concept-hypothèse est appris, il peut être utilisé pour effectuer des prédictions sur de nouvelles données en vérifiant si elles sont compatibles avec les intervalles spécifiés dans le concept-hypothèse.

Exercice 3 Soit C la classe de concepts *ensemble convexe* de R^2. Montrer que VC-dim(C) = ∞.

Correction :

La classe de concepts C, qui représente les ensembles convexes dans R^2, a une VC-dimension infinie, c'est-à-dire VC-dim(C) = ∞.

Pour le montrer, nous devons démontrer qu'il existe une séquence infinie de points dans R^2 que la classe de concepts C peut shatter, c'est-à-dire qu'elle peut les séparer en toutes les manières possibles.

Considérons la séquence infinie de points {(1, 0), (2, 0), (3, 0), ...} sur l'axe des abscisses dans R^2. Pour n'importe quel entier positif n, il est possible de trouver un ensemble convexe dans C qui sépare les n premiers points des autres points dans la séquence.

En particulier, pour n = 1, on peut choisir un cercle de rayon 1 centré en (1, 0) qui englobe uniquement le premier point (1, 0) et laisse les autres points en dehors de l'ensemble convexe.

Pour n = 2, on peut choisir un ensemble convexe qui englobe les deux premiers points (1, 0) et (2, 0), mais laisse les autres points en dehors de l'ensemble convexe. Par exemple, on peut choisir un triangle avec les sommets aux points (1, 0), (2, 0) et (1.5, 1), qui englobe les deux premiers points mais laisse les autres points en dehors de l'ensemble convexe.

De manière générale, pour n = k, on peut choisir un ensemble convexe avec k+1 sommets aux points (1, 0), (2, 0), ..., (k, 0) qui englobe les k premiers points mais laisse les autres points en dehors de l'ensemble convexe.

Ainsi, on peut continuer cette construction pour des valeurs de n arbitrairement grandes, ce qui montre que la classe de concepts C peut shatter une séquence infinie de points dans R^2. Par conséquent, VC-dim(C) = ∞.