

Lilien (利廉) - 76246003

Magalie (周诗瞳) – 16241018

Compte rendu de TP (discrimination linéaire)

1) Pour analyser les performances des deux discriminateurs pour une base linéairement séparable, on peut compter le nombre de points du mauvais côté de la frontière PI et de la frontière Hebb, et déterminer le meilleur algorithme. Sinon, dans le cas où les points sont trop nombreux, on peut aussi calculer l'écart relatif entre l'angle de w_{PI} (respectivement w_{Hebb}) et w_{vrai} . On constate que l'écart entre w_{PI} et w_{vrai} est plus faible que l'écart entre w_{Hebb} et w_{vrai} , cela est d'autant plus vrai lorsqu'on augmente la longueur de la base d'apprentissage. Pour une base non-linéairement séparable, on observe une intersection entre les points de la classe 0 et de la classe 1, les frontières PI et Hebb ont tendance à être identiques, d'autant plus lorsqu'on augmente la longueur de la base d'apprentissage, les frontières sont alors quasi-superposables.

Remarque : si, par curiosité, on choisit la base non-linéairement séparable avec « choix_base3 », alors la répartition des classes est aléatoire, on constate que l'algorithme PI et Hebb scindent le mélange en deux parties d'aire environ égale.

2) La représentation graphique de l'écart-type en fonction de la taille de la base de généralisation, en échelle logarithmique, est une droite décroissante. En réarrangeant cette expression, on peut retomber sur la relation reliant l'écart-type avec la moyenne. Si on fait un seul tirage, alors on peut se contenter de remplacer la valeur de μ par la valeur de τ_g . On aura alors des résultats moins précis, voire pas peu précis si la valeur de τ_g s'écarte trop de la moyenne. Faut-il alors répéter plusieurs fois l'expérience puis faire une moyenne (pour éviter d'effectuer l'expérience avec une valeur de τ_g trop éloignée de la moyenne) ?

3) On remarque que τ_{app} est égal à 1 pour une base de longueur inférieure à 42, puis diminue lorsque la longueur de la base dépasse 42. Cela est dû au fait que la base est de dimension 42 : il y a donc 42 degrés de liberté. Ainsi, lorsqu'on dépasse une longueur de 42, il y a plus d'inconnues que d'équations donc τ_{app} diminue. Sinon, la discrimination PI (entre 0,95 et 1) reste plus performante que la discrimination de Hebb (entre 0,85 et 0,95) pour τ_{app} . En ce qui concerne τ_g , on remarque que les discriminateurs PI et Hebb sont équivalents en termes de performances et que les points sont plus resserrés lorsqu'on augmente P_{gen} . Le bruit est-il dû au fait que lorsqu'on augmente P_{app} , il y a une nouvelle réalisation de la base d'apprentissage et que cela engendre deux sources d'erreurs (la base de généralisation et la base d'apprentissage) ?

4) Lorsqu'il n'y a pas de changement de base, on remarque que pour des σ croissants sur τ_{app} , R_A démarre comme PI puis diminue par paliers pour atteindre Hebb. De même, pour τ_g , R_A démarre comme PI et augmente très rapidement pour converger vers Hebb. En changeant de base de généralisation, on remarque la présence de bruit sur τ_g , ce qui paraît logique puisqu'on effectue des changements de base pour chaque σ . En revanche, lorsqu'on change de base d'apprentissage, on observe du brouillage à la fois sur τ_{app} et τ_g , ce qui nous laisse penser que τ_g ne dépend pas que de P_{gen} . Comment expliquer ce phénomène ?

Autre question : vous parlez, lors de vos recherches, de capteurs marins permettant de d'identifier précisément l'origine des sons diffusés dans l'océan, notamment d'explosions ou de baleines, mais y a-t-il réellement des baleines dans la mer Méditerranée ?