

Orlane Audrey

1.1/Les trois figures au-dessus représentent la densité de probabilité de x sachant que $x \in \omega_c$, ($c = 1,2,3$). Il représente la position de x la plus concentré dans chaque ω . Les trois figures au-dessous représentent la probabilité de $x \in \omega_c$. Il représente la frontière des ω .

1.2/Pour une matrice de confusion $M = (m_{k,l}), P(\omega_i|\omega_j) = \frac{m_{i,j}}{\sum_{k=1}^{k=3} m_{k,j}}$. Ici, les deux τ_g sont égaux, les deux barre sont égaux, les deux risques sont égaux. Puisque ici $\alpha_{i,j} = 0$ si $i=j$, $\alpha_{i,j} = 1$ sinon, les critères des deux discriminateurs sont mêmes, donc, τ_g , barre et risque sont même. Dans ce cas là, $P_{err} = R$.

1.3/Quand on change la valeur de $\alpha_{i,j}$, le calcul de d_{Bayes} change, donc la frontière de d_{Bayes} varie. τ_g et R de d_{Bayes} sont plus petit que lesquels de d_{connus} . C'est à dire que au point de vue de P_{err} , la performance de d_{Bayes} diminue car dans le prior uniforme, il n'est plus optimal. Mais au point de vue de risque, d_{Bayes} a une meilleur performance. Le choix de coût $\alpha_{i,j}$ est un choix de valeur, qui représente l'erreur qu'on préoccupe, donc R est un critère plus important.

1.4/On observe que τ_g de d_{Bayes} est plus grand et R de d_{Bayes} est plus petit. Car $\alpha_{1,2}$ est très grand, on veut diminuer $P(\omega_1|\omega_2)$. Donc on le réalise en augmentant largement $P(\omega_2)$. Ici, le prior n'est plus uniforme. Par ce méthode, on améliore τ_g et R en même temps.

2.1/ Dans le figure, les nombres dans la ligne diagonale $m_{k,k}$ représente le fois quand le classe vrai égale au classe estimé pour le numéro k . Donc, si $m_{k,k}$ est plus grand, la performance pour k est meilleur. Donc, la performance correspondant à 7 est le meilleur. La corrélation représente le ressemblage des deux chiffres, donc la performance. D'après la contraste de la figure, la performance correspondant à 7 est le meilleur.

2.2/ Les R de d_{Bayes} et $d_{Bayes,l}$ sont proches, qui sont moins que R de $d_{Bayes,q}$. Donc, la performance de d_{Bayes} et $d_{Bayes,l}$ est meilleure. D'après la figure, lorsque P_{app} se situe dans cette intervalle, la performance de $d_{Bayes,q}$ est toujours la mauvaise. Bien que $d_{Bayes,q}$ a une applicabilité plus large, mais il faut avoir une taille d'apprentissage énorme, donc ici, sa performance est mauvaise.

2.3/Oui, il est possible. D'après l'équation (65), quand on change $\alpha_{i,j}$ et $P(\omega_j)$, \hat{P} ne change pas, donc il n'est pas nécessaire de recalculer le risque.

Par cette fonction, on peut calculer R par $\alpha_{i,j}$ et $P(\omega_j)$ différent. Quand on modifie le choix cout, si on modifie aussi le prior, la performance est meilleur. C'est le même résultat que 1.4.

3.1 Question : pourquoi on choix ces deux forme de conditions à estimer P ?

D'après l'histogramme, on peut voir la distribution de R . $R_{RN} \in [2,3]$, $R_{RN,\beta} \in [0.2,0.6]$, $R_{RN,k} \in [0.7,1.4]$, donc les performances : $R_N, \beta > R_N, k > R_N$.

Quand on augmente P_{app} , $R_{RN,\beta}$ et $R_{RN,k}$ diminuent, mais $R_{RN,k}$ diminue plus largement. C'est à dire que la performance de R_N, k est plus sensible à P_{app} .

3.2/Quand on change le fonction coût mais pas le prior, les trois R sont grands, et les distributions dans l'histogramme sont dispersées. Et puis on change le prior =2, les trois R diminuent et deviennent plus concentrés. Donc la performance s'améliore, comme question 1.4. on a les performances : $R_N, \beta > R_N, k > R_N$. Quand le prior =3, dans l'histogramme, la concentration développe, et les courbes des trois R se superposent.

3.3/D'après l'équation 62, après l'apprentissage, on a $P(\omega_i|x)$, si on change la valeur de $\alpha_{i,j}$, il doit seulement calculer ρ sans changer $P(\omega_i|x)$.

Question : d'après l'équation 62, il n'y a pas de $P(\omega_j)$, donc, comment calculer sans effectuer un nouvel apprentissage quand on modifie $P(\omega_j)$?