

Tp4

1. En sachant que $P(w_i|w_j)$ est la proba de choisit w_i en sachant de classe w_j . Donc à partir du matirx, $P(w_i|w_1)=d_{i1}/(d_{11}+d_{21}+d_{31})$, et $R=\frac{\sum (\sum_{i,j} d_{ij})}{(d_{11}+d_{21}+d_{31})}$.

En calculant, $R=(4/66+9/67+10/67)/3=0.1349$

$1-R=0.8650 \approx \tau$ de généralisation=0.865

Dans le cas ou $a=0$ pour i,j sont meme ,cad vrai ,et $a=1$ sinon, le risque de bayes est egale à la probabilité d'erreur. La performance de 2 diecriminateurs sont la meme.

D'après la figure, $P(x|w_i)$ représente la densité de proba de x qui appartient à w_i .

$P(w_i|x)$ représente pour tous les x , la probabilité des x qui appartiennent à w_i dans tous les x .

Dans le cas ou $a_{12}=100$,cad que si la vrai classe est 2, mais on choit la classe 1 , le poits de l'erreur augemente , ce représente de ce erreur devient grave. Donc le système préfere considérer un x qui ressemble aux classe 1 et 2, appartient à la classe 2 pour éviter les erreurs. c'est pourquio $P(w_2|w_1)$ est bp plus grand et $P(w_1|w_2)$ est bp plus petit que les resultats précédent. En general , la performance de bayes est pire.

Dans le cas $P(w_2)=10P(w_1)$,le sys sait que la plupart des x sont en réalité appartient aux classe 2, donc il préfere consider que les x appartient à la classe 2. c'est la raison pour laquelle $P(w_2|w_2)$ est tres grand.de plus, $P(w_2|w_1) > P(w_3|w_1)$, $P(w_2|w_3) > P(w_1|w_3)$.

2. Si le corfficient de corrélation est plus grand, les images des chiffres sont plus claires(plus semblable au chiffre), donc on se trompe plus dans la discrimination pour les chiffres similaire.

Pour 2.2, d'après le figure, la performance de bayes est meilleur parce que connaitre tous les paramères,la risque sont tres petit, et les 2 autres peuvent arrive un petit rique seulement quant le Papp augemente. En comparant, le discriminateur bayes quadra a besion un plus grand nombre de Papp pour un petit risque. Parce que pour le discriminateur bayes quadratique, on ne connait pas les paramètres de $P(x|w_i)$, et les covariance ne sont pas les meme, donc il demande une grande base à apprendre pour atteindre une bonne performance. Pour lineaire, il a y les meme covariance, donc le seuil de Papp est petit que celui de quadratique.

Comme dans la question 1.3,1.4, le système qui a une tendance d'éviter d'erreur(sur cet classe, il y moins de nombre), a un cout(a) plus grave que les autres. De plus ,il a souvent une meilleur performance sur la classe qui a le grand prior.

Diffeculté :En regardant le nombre ,on peut déduire le prior $P(w_2)$ devient grand,masi sait pas le precis cout à cause de grand nombre de classe.

La performance de $RN\beta$ est meilleure en raison de la bas risque. Avec le chgmet de prior et cout ,le taux de généralisation devient meilleur. De plus, quand on realise plus de passe, la performance devient mieux. Si la taille d'apprentissage est assez grande,la performance de RN est meilleur que Bayes.

Pour ne pas avoir besion d'effectuer un nouvel apprentissage à chaque foi, on garde le w du RN.Mais on ne comprend pas les 2 formules de solutions signifient à quoi et lien entre le resultat.