Compte rendu de TP4

Océane et Cécilia

1.

$P(wi|wj)$ est la probabilité de classe estimée w_i en sachant de classe vraie w_j, donc à partir de la matrix, par exemple, $P(w_i|w_1)=d_{i1}/(d_{11}+d_{21}+d_{31})$.

Pour figure 13, $P(x|w_i)$ représente la densité de probabilité de x appartient à w_i. $P(w_i|x)$ représente dans tous les x, la probabilité des x qui appartiennent à w_i dans tous les x.

- Dans le cas alpha = 0 pour i = j et alpha = 1 sinon,

$$R = (1/3) * (8/66+9/67+10/67) = 0.1349$$

$$1-R = 0.8650 = \text{tau de généralisation}$$

le risque de Bayes est bien égale à la probabilité d'erreur. La performance des deux discriminateurs sont la même.

- Dans le cas alpha = 100 pour i = 1 et j = 2, on analyse l'influence du choix de la fonction coût.

$$R = (1/3)*(8/66+10/67)+(9/67)*100*(1/3) = 0.2015$$

On amplifie le poids de l'erreur de considérer un x appartient à la classe 1 alors que celui-ci appartient à la classe 2. Cela veut dire que le résultat de ce erreur est plus grave. Donc le système préfère considérer qu'un x qui ressemble aux classe 1 et 2 appartient à la classe 2 pour éviter les erreurs. C'est pourquoi ici $P(w_2|w_1)$ est beaucoup plus grand et $P(w_1|w_2)$ est beaucoup plus petit que le cas précédent. En général, la performance de Bayes est pire.

- Dans le cas $P(w2)=10p(w1)$, on analyse l'influence de prior.

$$R=5/6*3/166+1/12*(4/17+1/17)=0.03957$$

Le système sait que la plupart des x sont de classe 2, docn il préfère considérer que les x appartient à la classe 2 pour éviter les erreurs. C'est pourquoi $P(w_2|w_2)$ est très grand. De plus, $P(w_2|w_1) > P(w_3|w_1)$ et $P(w_2|w_3) > p(w_1|w_3)$. En général, la performance de Bayes est mieux.

2.

Plus le coef de corrélation est grand, plus les images des chiffres sont similaires, donc plus on se trompe dans la discrimination pour les chiffres similaires.

Pour le discriminateur Bayes quadratique, on ne connaît pas les paramètres de $P(x|w_j)$ et les covariances ne sont pas le même, donc on demande une grande base à apprendre pour atteindre une bonne performance. Dans le cas linéaire, on a les même covariance, donc le seuil de Papp est beaucoup plus petit que celui de quadratique. En tout cas, le discriminateur Bayes a toujours une excellente performance en sachant tous les paramètres.

Comme dans la question 1.3&1.4, le système a une préférence d'éviter les erreurs qui ont des résultats plus grave que les autres. De plus, il a souvent une meilleure performance sur la classe qui a un plus grand prior. Parce que comme ça il peut minimiser le risque.

Mais comme on ne sais pas exactement le choix de la fonction coût, c'est difficile de recalculer la risque numérique sur une nouvelle base de généralisation.

3.

La performance de RN_β sur la classe 8 est évidemment pire que RN_k et RN. Mais on ne comprend pas la différence de la formule de β&k.

Pour ne pas avoir besoin d'effectuer un nouvel apprentissage à chaque fois, on peut sauvegarder le w du réseau de neurones.