KUHN Théotime
MOSSERI Julia

Compte rendu TP 2

Le but de l'étude est d'analyser les performances de trois discriminateurs donnés basés sur des connaissances et estimations différentes de la moyenne et de la matrice de covariance

En générant différents examples de bases de données pour des cas de moyennes et matrices de covariances égales ou différentes parmi les classes, on se rend compte de la difficulté dans de nombreux de tracer un critère pouvant être représenté par une frontière physique.

Lorsque la moyenne et la matrice de covariance de chaque classes sont connues, le discriminateur correspondant permet de minimiser la probabilité d'erreur.

Cependant, les moyennes et covariances ne sont pas toujours connues. On utilise pour cela le discriminateur linéaire ou le discriminateur quadratique.

On mesure le taux de généralisation pour chaque discriminateur à partir de la même base d'apprentissage. En choisissant comme barre d'erreur le paramètre $\sigma_{\tau g} = \sqrt{\frac{\mu_{\tau g}(1-\mu_{\tau g})}{P_{gen}}}$, et en augmentant la taille de la base de généralisation pour plus de précision, on obtient que le discriminateur linéaire est considérablement moins efficace que le discriminateur quadratique.

Le même test est réalisé sur une base d'apprentissage dans laquelle chaque classe est plus clairement définie, donc plus simplement séparables. Les résultats donnent une nette amélioration des performances du discriminateur linéaire, sans toutefois dégrader les performances du discriminateur quadratique.

En faisant varier la taille de la base d'apprentissage, on peut appercevoir une chute de performance du discriminateur linéaire autour de P_{app} = N la dimension de l'espace, ainsi qu'un chute similaire de performance du discriminateur quadratique autour de P_{app} = C*N, avec C le nombre de classes. Ceci est dû au terme d'inversion de la matrice de covariance dans l'expression de la fonction de distribution de la loi normale.

En effet, la matrice de covariance Γ est de taille N*N. Si P_{app} = N, la probabilité qu'il existe un vecteur dans la base d'apprentissage qui soit généré par d'autres vecteurs de cette base est forte. On a alors $\lambda_N \rightarrow 0$ sans l'atteindre pour tout $\hat{}\Gamma_c$. On a donc aussi $\lambda_N \rightarrow 0$ sans l'atteindre pour $\hat{}\Gamma_{lin}$. Or, l'inverse d'une matrice avec un lambda $\rightarrow 0$ est très bruitée.

Le discriminateur linéaire est donc peu performant pour P_{app}=N.

Le discriminateur quadratique, quand à lui, considère une matrice de covariance différente pour chacune des classes. Pour P_{app}=N et C!=1, on a au moins un $\hat{}\Gamma_c$ qui n'est pas inversible. On utilise donc la méthode de la pseudo-inverse, ce qui nous donne un résultat non bruité. Toutefois, lorque P_{app} = C * N, on s'approche de la configuration précedente, pour chaque classe c. Il est donc propable que l'inverse de chaque $\hat{}\Gamma_c$ soit bruitée, diminuant ainsi la performance du discriminateur