Compte rendu

1. Dans la question 1.c), on trouve que les points présents sont non-linéairement séparables par une ligne. Est-ce que le *choix_base=2* signifie que les points sont linéairement séparables par un plan ou un hyperplan dans la dimension plus haute en utilisant la fonction du noyau?

2. Dans la question 2.d), on a appris que la moyenne du taux de généralisation peut prévoir l'écart type.

3. On a appris que quand la taille de la base d'apprentissage P_{app} est petite, le taux d'apprentissage τ_{app} est presque 1, le taux de généralisation τ_g est bas. Au contraire, quand P_{app} agrandit, τ_{app} va diminuer et τ_g va augmenter.

4. D'après la question 2.d), on a σ_{τ_g} et P_{gen} sont inversement proportionnelles. Mais on ne le trouve pas vrai dans le graphique de la question 3.b). Si j'ai bien entendu, vous nous expliquiez que τ_g dépend de P_{app} et P_{gen}, pas seulement P_{gen}. Mais si on fixe P_{app}, pourquoi σ_{τ_g} ne change pas en fonction de P_{gen} ?

5. On trouve que quand $\sigma \to 0$, *ridge approximation* est plus proche de PI.
   Quand $\sigma \to \infty$, *ridge approximation* est plus proche de Hebb.

6. On pense que c'est inconvénient d'obtenir la bonne valeur de σ à l'aide du taux de réussite. Puisque si on a P_{app} de taille 100000, chaque fois qu'on prend une nouvelle valeur de σ, il faut refaire des calculs pour calculer τ_g et τ_{app}. Et pour apprendre la bonne valeur de σ, il faut trouver la valeur maximum du τ_g. Ça coûte trop cher donc pas pratique.

Romain et Lou