

**INF-578 Máquinas de Aprendizaje**  
**Cuestionario Control 3 II 2017**

Los ejercicios marcados con (★) son solo para alumnos de postgrado.

1. Explique cuál es la estrategia implementada en una red neuronal (ANN) para abordar problemas de clasificación que no son linealmente separables.
2. ¿Si consideramos un gran número de iteraciones. Adaboost dará error de training cero independiente de los clasificadores que se estén usando?
3. (★) ¿Es cierto o es falso que el algoritmo de backpropagation genera una capacidad de generalización de un MLP que es independiente del número de capas que éste tenga? Justifique seriamente.
4. (★) Explique brevemente la descomposición sesgo-varianza del error de un ensamblado.
5. En una ANN, discuta brevemente sobre los posibles problemas de usar un *learning rate* demasiado alto, o demasiado bajo. Mencione alguna alternativa para abordar este problema.
6. ¿Cuál es el objetivo del *kernel trick* que usan las SVMs?
7. ¿Qué propiedades debe tener una matriz de Kernel?
8. ([1] 14.5) Considere un ensamblado agregado por una combinación lineal no-uniforme

$$F(\mathbf{x}) = \sum_{n=1}^N \alpha_n f_n(\mathbf{x}).$$

Supongamos que queremos limitar la salida del ensamblado a

$$f_{\min}(\mathbf{x}) \leq F(\mathbf{x}) \leq f_{\max}(\mathbf{x}).$$

Muestre que es condición necesaria y suficiente que los coeficientes satisfagan  $\alpha_n \geq 0$  y

$$\sum_{n=1}^N \alpha_n = 1.$$

9. Si  $K_1(x, z)$  y  $K_2(x, z)$  son funciones de Kernel.
  - (a) Demuestre que  $K(x, z) = K_1(x, z) + K_2(x, z)$  es una función de Kernel.
  - (b) Demuestre que  $K(x, z) = K_1(x, z)K_2(x, z)$  es una función de Kernel.
  - (c) (★) Demuestre que el Kernel gaussiano

$$K(x, z) = e^{-\frac{\|x-z\|^2}{\sigma^2}},$$

donde  $\sigma^2 > 0$  es una constante positiva, es un kernel válido. [Ayuda: Considere que  $\|x - z\|^2 = \|x\|^2 - 2x^T z + \|z\|^2$ .]

10. Sea  $F(\mathbf{x}) = \sum_{n=1}^N w_n f_n(\mathbf{x})$ . demuestre que:

$$(y - F(\mathbf{x}))^2 = \sum_{n=1}^N w_n (y - f_n(\mathbf{x}))^2 - \sum_{n=1}^N w_n (f_n(\mathbf{x}) - F(\mathbf{x}))^2.$$

11. Derive las ecuaciones de backpropagation usando regularización L2 para los pesos de la capa esondida y de salida.

12. Derive Las ecuaciones de forward y backward para la red neuronal feed-forward de tres capas entrenada con la función de pérdida *cross-entropy* y con función de activación softmax en la capa de salida.

## References

- [1] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.