**Evaluación de conocimientos - Semana 7 - Clasificación supervisada (DT, RNA, SVM)**

**NOMBRE: Frédéric Roux** **Fecha: 14/12/2018**

**BLOQUE 1: SELECCIONA LAS RESPUESTA(S) CORRECTA(S)**

**1.-** Respuesta(s) correcta(s): b

**2.-** Respuesta(s) correcta(s): c

**3.-** Respuesta(s) correcta(s): d

**4.-** Respuesta(s) correcta(s): c

**5.-** Respuesta(s) correcta(s): b

**6.-** Respuesta(s) correcta(s): c

**7.-** Respuesta(s) correcta(s): a

**BLOQUE 2: INDICA SI LAS SIGUIENTES AFIRMACIONES SON VERDADERAS O FALSAS. Justifica brevemente tus respuestas**

**8.-** Respuesta:

Verdadero. Los SVMs quedan mas robustos frente al curse of dimensionaliy. Es porque no hay muchos parámetros que deben ser tuneados en el caso del SVM para estimar el margin. El rendimiento del SVM es dicho de ser independiente del numero de features, porque depende del margin y del parámetro c. Por lo tanto, es menos probable de "caer en la trampa" del overfitting que otros algoritmos en cuales el numero de parámetros crece con el numero de variables input.

**9.-** Respuesta:

Verdadero. Mas nodos hoja tiene un árbol de clasificación, mas complejo resulta.

**10.-** Respuesta:

Falso. La complejidad dependen sobre todo del numero de nodos.

**11.-** Respuesta:

Falso. La complejidad de una red neuronal artificial depende del numero de capas intermedias.

**12.-** Respuesta:

Falso. Los problemas non-lineales se pueden resolver usando el soft margin (parámetro c) y el kernel trick.

**BLOQUE 3: CONTESTA BREVEMENTE A LAS SIGUIENTES CUESTIONES**
**Razona tus respuestas**

**13.-** Respuesta

Los diferentes criterios de división en un árbol de clasificación buscan a partir los datos con respecto a las clases (variable dependente) a cuales pertenecen los datos. Es común por ejemplo partir los datos de manera a obtener una partición muy homogena (entropia baja) de las clases en cada nodo.

**14.-** Respuesta

El árbol sera muy complejo, es decir que se ajustara hasta que quedan todo los casos clasificados, aunque los nodos resultantes no sean muy representativos para el conjunto de entrenamiento (overfitting). Sin pre- o post-poda, es muy probable que el árbol no tendrá una rendimiento muy alto en la fase de cross-validación con datos del conjunto test.

**15.-** Respuesta

La aplicación de un kernel trick a los datos de entrada en el marco de un clasificador de tipo SVM permite de proyectar los datos a un espacio en cual están separables linealmente. Consiste a transformar los datos con una función (el dicho kernel) de tal manera para que una vez los datos de entrada están transformados se pueden separar linealmente. Un ejemplo clásico es el kernel RBF que permite de proyectar datos de un espacio con dos dimensiones x e y, a un espacio con tres dimensiones x,y e z, en cual z refleja la similitud o distancia geométrica entre los casos definidos por las cordenadas x e y.

**BLOQUE 4: EJERCICIO**

**16.-** Respuesta:

Presencia gotas: el atributo es discreto y genera dos particiones (NO y SI)

NO: 3x0, 1x1, 0x2
SI: 0x1, 4x1,4x2

Entropia presencia gotas NO: H = -( (3/4)*log(3/4) + (1/4)*log(1/4) + (0/4)*log(0/4)) = 0.56
Entropia presencia gotas SI: H = -( (0/9)*log(0/9) + (4/9)*log(4/9) + (4/9)*log(4/9)) = 0.72

**Entropia mediana ponderada presencia gotas:** Hm = ((4/13)*0.56+(8/13)*0.72) = 0.61

Velocidad del vehículo: el atributo es discreto y genera tres particiones ( baja, mediana, alta)

baja: 2x0,2x1,0x2
mediana: 0x0, 2x1, 1x2
alta: 2x0,1x1,3x2

Entropia velocidad baja: H = -( (2/4)*log(2/4) + (2/4)*log(2/4) + (0/4)*log(0/4)) = 0.69
Entropia velocidad mediana: H = -( (0/3)*log(0/3) + (2/3)*log(2/3) + (1/3)*log(1/3)) = 0.63
Entropia velocidad alta: H = -( (2/6)*log(2/6) + (1/6)*log(1/6) + (3/6)*log(3/6)) = 1.01

**Entropia mediana ponderada velocidad del vehículo:** Hm = ((4/13)*0.69+(3/13)*0.63+(6/13)*1.01) = 0.82

Porque a mayor entropía, peor es la partición, el mejor split se realizara con la variable **presencia gotas.**