Ejercicios puntuables de APR, curso 2018-2019, grupo 4CO21

La numeración de paginas corresponde a los materiales docentes
que hay en www.prhlt.upv.es/~evidal/students/apr/

La puntuación de cada ejercicio está en el rango [0,2]. La puntuación
total es el máximo entre 2.0 y la suma de puntuaciones de los
ejercicios entregados.

Tema 3:

3.1. (p.3.9, 0.5 puntos) Obtener los estimadores de máxima
verosimiltud del vector media de una gaussiana bi-variada, cuya matriz
de covarianza es fija y conocida partir de una muestra de de vectores
bi-dimensionales x_1,x_2,...,x_N.

3.2. (p.3.9, 1 punto) Obtener los estimadores de máxima verosimiltud
de los 5 parámetros de una gaussiana bi-variada, a partir de una
muestra de de vectores bi-dimensionales x_1,x_2,...,x_N.

3.3. (p.3.10, 0.25 puntos) Desarrollar y explicar con detalle todos los
pasos necesarios para obtener el estimador de máxima verosimiltud del
vector media de una de una gaussiana multivariada cuya matriz de
covarianza es fija y conocida, a partir de una muestra de vectores
x_1,x_2,...,x_N.

3.4. (p.3.16, 0.25 puntos) Ejercicio b) de la página indicada:
minimizar una función con una condición de igualdad.

3.5. (p.3.18, 0.5 puntos) Ejercicio al pie de la página indicada:
estimación por máxima verosimilitud de las probabilidades a priori de
un clasificador genérico en C clases.

3.6. (p.3.20, 0.25 puntos) Ejercicio al pie de la página indicada:
minimizar una función con una restricción de desigualdad haciendo
uso de la condición complementaria de KKT.

3.7. (p.3.26, 0.5 puntos) Mostrar la traza de tres iteraciones de
descenso por gradiente para minimizar la función indicada usando paso
decreciente con el número de iteraciones, k (por ejemplo \rho_k =
1/(2k)).

3.8. (p.3.26, 1 puntos) Implementar en el lenguaje de programación que se desee el algoritmo de descenso por gradiente para minimizar la función de Rosenbrock. Mostrar una traza de ejecución hasta convergencia. Ensayar paso fijo (valor pequeño fijo para \rho_k) y paso decreciente con el número de iteraciones, k (por ejemplo \rho_k = 1/(2k)). Esayar también distintas inicializaciones de los parámetros; al menos proporcionar traza(s) empezando en (-1, 1)^t.

3.9 (p.3.41, 0.25 puntos) Desarrollar con detalle los pasos necesarios para obtener la expresión de la esperanza de las variables latentes de una mezcla de gaussianas que se muestra en el Paso E la página 39.

3.10 (p.3.41, 1 puntos) Desarrollar con detalle los pasos necesarios para obtener las expresióones de actualización de parámetros (\pi_k, \mu_k, 1<=k<=K) de la mezcla de gaussianas que se muestran en el Paso M de la página 41. Sugerencia: Optimización analítica para las \mu_k (para las que no hay restricciones) y optimización lagrangiana (solo) para las \pi_k.

Tema 4:

4.1 (p.4.9, 0 puntos, pues está resuelto en el boletín de ejercicios) Detallar y completar los desarrollos indicados en esta página hasta obtener la Lagrangiana Dual en función de los multipicadores de Lagrange y de los productos escalares entre vectoresw de aprendizaje.

4.2. (p.4.14, 1 punto) Ejercicio 1 de la página indicada: Obtener de forma analítica los parámetros de una SVM a partir de 3 datos de entrenamiento en dos dimensiones.

4.3 (p.4.16-17, 0.75 puntos) Detallar y completar los desarrollos indicados en estas página s para SVM en caso de no seprabilidad lineal.

4.4 (p.4.18, 0.75 puntos) Una SVM con márgenes "blandos" asigna a cada vector de entrenamiento un valor de a=\alpha^* y de z=\zeta^*. Completar una tabla de tres filas y tres columnas como esta:

	a = 0	0 < a < C	a = C
z = 0	x	x	x
z <= 1	x	x	x
z > 1	x	x	x

donde C es la constante de tolerancia.

Cada casilla de la tabla, x, indica cómo quedaría clasificado un vector de entrenamiento para los correspondientes valores de a y z.

```
Esto debe indicarse mediante:

  "m"  : mal clasificado
  "b+" : bien clasificado y con margen mayor que el margen óptimo
  "b=" : bien clasificado justo con el margen óptimo
  "b-" : bien clasificado pero con margen menor que el margen óptimo
  "--" : imposible


Tema 5:
-------

5.1. (1 punto) Considerar el problema resuelto 5 del boletín de ejercicios
(www.prhlt.upv.es/~evidal/students/apr/EjercYexam/Ejercicios.pdf).

a) Calcular el error cuadrático, q_S(\Theta), de la red inicial para
   la muestra dada ((-2,1)^t, (0,1,0)^t) y el correspondiente error de
   la red reentrenada con dicha muestra.

b) Realizar una segunda iteración del algoritmo BackProp, con la misma
   muestra de entrenamiento y factor de aprendizaje (1.0).

c) Calcular el error cuadrático de la red obtenida en b) para la
   muestra dada.

5.2 (p.5.39, 0.5 puntos) Ejercicio propuesto al final de la página
indicada: Obtener las ecuaciones de actualización de un perceptrón
multicapa para la minimización de la entropía cruzada.


Tema 6:
-------

6.1. (p.6.9, 0.25 puntos) Ejercicio propuesto al final de la página
indicada: Calcular los 8 valores de P(A|L,C) y obtener la mejor
predicción para el estado del aspersor, sabiendo que llueve y el
cesped está mojado.

6.2. (p.6.10, 0.5 puntos) Ejercicio propuesto al final de la página
indicada: Determinar la probabilidad de que un paciente no fumador
*no* tenga cancer de pulmón, si la radiografía ha dado un resultado
negativo pero sufre de disnea.  (Ejercicio también propuesto en
Prácticas para ser resuelto de forma computacional).

6.3. (0.75 puntos): Realizar una traza del método de aprendizaje con
observaciones incompletas por "reestimación a la Viterbi" sugerido en
la P.6.40, para el ejemplo considerado en esa página con S = {(0,1),
```

(1,0), (1,1)} y p_0=0.4, p_1=0.6.

6.4. (0.75 puntos) Ejercicio propuesto al final de la sección de Aprendizaje EM: Realizar una traza del EM para el ejemplo desarrollado en las páginas 6.37-6.43, con S = {(0,1), (1,0), (1,1)} y p_0=0.4, p_1=0.6.