INTRODUCCION A LOS MODELOS COMPUTACIONALES 25 enero 2016

Alumno/a D...

Cuestiones.-

1) (2 puntos) En un modelo de análisis discriminante lineal tenemos la siguiente regla de decisión.

$$Si\ (\mathbf{m_2} - \mathbf{m_1})^{\mathbf{T}}\mathbf{\Sigma^{-1}x} + 1/2(\mathbf{m_1}^{\mathbf{T}}\mathbf{\Sigma^{-1}m_1} - \mathbf{m_2}^{\mathbf{T}}\mathbf{\Sigma^{-1}m_2}) - \ln\frac{P(C_1)}{P(C_2)} > 0$$

$$entonces\ \mathbf{x} \in C_1$$

Indique brevemente como hemos llegado a esta regla de decisión. ¿Qué significado tienen P(C_1) y P(C_2)? Cuál sería la decisión en el caso de que el vector asociado a un patrón sea $\mathbf{x^T}$= (-1, -1), $\mathbf{m}_1^{\mathrm{T}}$= (2,2) $\mathbf{m}_2^{\mathrm{T}}$=(-2,-2) $\mathbf{\Sigma^{-1}} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, sabiendo además que tenemos 50 patrones de la clase 1 y 50 de la clase 2 de la muestra de entrenamiento.

Solución.- Por una parte

$$((-2,-2)-(2,2))\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \end{pmatrix} = 0$$

y por otra

$$(2,2)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = 0$$

$$(-2,-2)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ -2 \end{pmatrix} = 0$$

como además

$$-\ln\frac{P(C_1)}{P(C_2)} = -\ln\frac{0,5}{0,5} = 0$$

Tenemos que la ecuación es 0 y por tanto **x** está en la función de decisión y no podemos decidir a qué clase pertenece.

2) (1 punto) ¿Qué es el margen en la metodología SVM? ¿Cuáles son las ecuaciones de los dos hiperplanos del margen H^+ y H^-?

Solución

El margen es la distancia entre los dos hiperplanos del margen H + y H-.

Sus ecuaciones son:

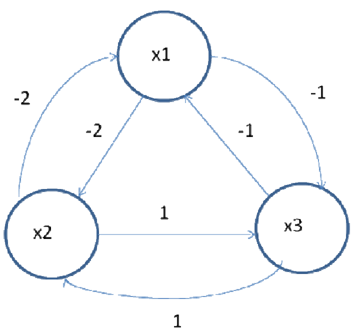
H +: $\langle w \bullet x^+ \rangle + b = 1$

H^-: $\langle w \bullet x^- \rangle + b = -1$

donde x^+ y x^- , son los puntos de datos que están más cerca del hiperplano $\langle \mathbf{w} \cdot \mathbf{x} \rangle + b = 0$

3) (1 punto) ¿Qué diferencias existen entre las redes MLP y la redes RBF? Escriba dos modelos sencillos de dichas redes. Cite y explique alguno de los algoritmos de entrenamiento de estas redes.

Ejercicio 1.- (2 puntos) Para la red de Hopfield ilustrada a continuación, dibuje el diagrama de transiciones síncronas posibles y luego determine los puntos fijos, o estados estables de la red.



Solución.-
Para la unidad elegida, se calcula la suma de los pesos de las conexiones sólo a los vecinos activos, si los hay. Si la suma es > 0, entonces la unidad elegida se convierte en activa, de lo contrario, se vuelve inactiva. Si suponemos que los tres nodos están activos tenemos de inicio X = [1 1 1], para los nodos 1, 2 y 3. La suma de las conexiones para x_1 es -2+(-1)=-3, por lo que se hace inactiva. la suma de las conexiones para x_2 tiene ahora como -2+1=-1, luego se hace inactiva. La suma de las conexiones para x_3 es -1+1 luego sigue activa. De esta forma el estado estable es X = [-1 -1 1].

La matriz de transición de estados es $W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$ y matricialmente tenemos

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

luego se deduce que (1,1,1) $\Leftrightarrow$ (-1,-1,1)

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -3 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

luego se deduce que (1,1,-1) $\Leftrightarrow$ (-1,-1,-1)

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ 2 \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

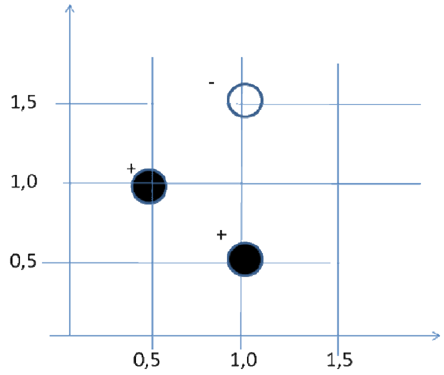
luego se deduce que (-1,1,-1) $\Rightarrow$ (-1,1,1) y que (-1,1,1) es un estado estable

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

$$W = \begin{pmatrix} 0 & -2 & -1 \\ -2 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \\ -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

luego se deduce que (1,-1,1) $\Rightarrow$ (1,-1,-1) y que (1,-1,-1) es un estado estable

Ejercicio 1.- (2 puntos) Considere los tres vectores bidimensionales linealmente separables de la siguiente figura. Encuentre el SVM lineal que separa de manera óptima las clases al maximizar el margen.



Solución Todos los puntos son vectores soporte el hiperplano de margen H^+ es la línea que pasa por los dos puntos positivos. El hiperplano de margen H^- es la recta que pasa por el punto negativo y es paralela a H^+. La función de decisión es la recta que está entre H^+ y H^-. Esta recta tiene por ecuación $-x_1 + 2 = x_2$.

Tenemos la clase positiva formada por los vectores soporte (0.5 ,1) y (1, 0.5) o también (0.5 ,1, 1) y (1, 0.5, 1)
La clase negativa está formada por el vector soporte (1, 1.5) o también (1, 1.5 ,1)
La ecuación del hiperplano es $\mathbf{w}^T.\mathbf{x} + w_0 = 0$, o lo que es lo mismo $w_1x_1+w_2x_2+w_0=0$

A) Primal
Si sustituimos los vectores soporte en los hiperplanos positivo y negativo
H^+: $\langle w \bullet x^+ \rangle + w_0 = 1$
H^-: $\langle w \bullet x^- \rangle + w_0 = -1$
tenemos las ecuaciones
$0.5w_1+w_2+w_0=1$
$w_1+0.5w_2+w_0=1$
$w_1+1.5w_2+w_0=-1$
y operando adecuadamente $w_1=-2$; $w_2=-2$ y $w_0= 4$, por lo que la ecuación del hiperplano separador es $x_1+x_2-2=0$

B) Dual

$$\max \sum_{i=1} \alpha_i - \frac{1}{2} \sum_{i,j=1} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i^T \mathbf{x}_j) = \max \sum_{i=1} \alpha_i - \frac{1}{2} \boldsymbol{\alpha}^{\mathbf{T}} \mathbf{Q} \boldsymbol{\alpha}$$

$$s.a. \ \sum_{i=1} \alpha_i y_i = 0, \quad \alpha_i \geq 0, \ \text{i=1,...,n}$$

$$\hat{\mathbf{w}} = \sum_{i=1} \hat{\alpha}_i y_i \mathbf{x_i} = \sum_{i \in Sop} \hat{\alpha}_i y_i \mathbf{x_i}$$

$$\hat{w}_{\mathbf{0}} = 1 - \sum_{j \in Sop} \hat{\alpha}_j y_j (\mathbf{x_j^T x_i}), \quad \text{con } \mathbf{x_i} \in \omega_1 \ y \ \ \hat{\boldsymbol{\alpha}}_{\mathbf{i}} > 0$$

Ahora la función a maximizar es

$$\mathrm{L}(\alpha_1, \alpha_2, \alpha_3, \lambda; \ \mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}) =$$

$$= \alpha_1 + \alpha_2 + \alpha_3 - \ \frac{1}{2} \left(\begin{array}{l} \alpha_1^2 (\mathbf{x_1^T}.\mathbf{x_1}) + \alpha_1 \alpha_2 (\mathbf{x_1^T}.\mathbf{x_2}) - \alpha_1 \alpha_3 (\mathbf{x_1^T}.\mathbf{x_3}) + \alpha_2 \alpha_1 (\mathbf{x_2^T}.\mathbf{x_1}) + \alpha_2^2 (\mathbf{x_2^T}.\mathbf{x_2}) - \alpha_2 \alpha_3 (\mathbf{x_2^T}.\mathbf{x_3}) + \\ -\alpha_3 \alpha_1 (\mathbf{x_3^T}.\mathbf{x_1}) - \alpha_3 \alpha_2 (\mathbf{x_3^T}.\mathbf{x_2}) + \alpha_3^2 (\mathbf{x_3^T}.\mathbf{x_3}) - \lambda (\alpha_1 + \alpha_2 - \alpha_3) \end{array} \right)$$

sustituyendo los valores tenemos

$$\mathrm{L}(.) = \alpha_1 + \alpha_2 + \alpha_3 - \frac{1}{2}\left(\begin{array}{l} \alpha_1^2(\frac{5}{4}) + \alpha_1\alpha_2(1) - \alpha_1\alpha_3(2) + \alpha_2\alpha_1(1) + \alpha_2^2(\frac{5}{4}) - \alpha_2\alpha_3(\frac{7}{4}) + \\ -\alpha_3\alpha_1(2) - \alpha_3\alpha_2(\frac{7}{4}) + \alpha_3^2(\frac{13}{4}) - \lambda(\alpha_1 + \alpha_2 - \alpha_3) \end{array}\right)$$

Derivando con respecto a los α y λ tenemos

$$\frac{\partial L}{\partial \alpha_1} = 0, \text{ esto es}$$

$$1 - \frac{5}{4}\alpha_1 - \alpha_2 + 2\alpha_3 - \lambda = 0$$

$$\frac{\partial L}{\partial \alpha_2} = 0, \text{ esto es}$$

$$1 - \alpha_1 - \frac{5}{4}\alpha_2 + \frac{7}{4}\alpha_3 - \lambda = 0$$

$$\frac{\partial L}{\partial \alpha_3} = 0, \text{ esto es}$$

$$1 + 2\alpha_1 + \frac{7}{4}\alpha_2 - \frac{13}{4}\alpha_3 + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = -\alpha_1 - \alpha_2 + \alpha_3 = 0, \text{ luego}$$

$$\alpha_3 = \alpha_1 + \alpha_2$$

Resolviendo el sistema de 4 ecuaciones con 4 incognitas tenemos

$$\alpha_1 = 4;\ \alpha_2 = 0 \text{ y } \alpha_3 = 4$$

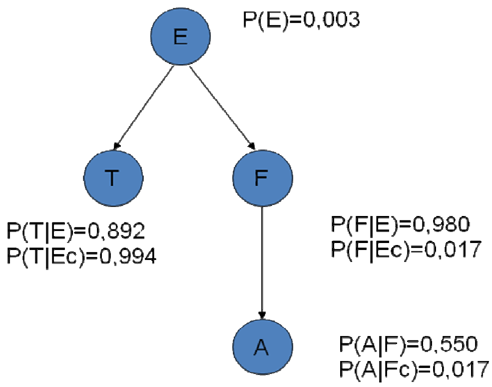
$$\hat{\mathbf{w}} = \sum_{i \in Sop} \hat{\alpha}_i y_i \mathbf{x_i} = (-2,-2)$$

$$\hat{w}_{\mathbf{0}} = 1 - \sum_{j \in Sop} \hat{\alpha}_j y_j (\mathbf{x_j^T x_i}), \quad \text{con } \mathbf{x_i} \in \omega_1 \ y \ \ \hat{\alpha}_i > 0$$

$$\hat{w}_{\mathbf{0}} = 4$$

La ecuación es $-2x_1+2-x_2+4=0$

Ejercicio.- (2 puntos)
Supongamos que para detectar cierta enfermedad, hacemos un test. Definimos cuatro variables:
E = presencia de la enfermedad, que toma los valores Si, No
T = resultado del test, que toma los valores. Positivo, Negativo
F = presencia de fiebre en el enfermo, que toma los valores. Si, No
A= amígdalas inflamadas. Si, No
Entre las variables establecemos una relación de influencia causal o red bayesiana.



Se pide, teniendo en cuenta la independencia condicional y la independencia de los nodos no conectados
a) ¿Cuál es la probabilidad de que una persona cuyo test de positivo y tenga fiebre, padezca la enfermedad?
b) ¿Cuál es la probabilidad de que una persona cuyo test de positivo, padezca la enfermedad?
c) ¿Cuál es la probabilidad de que una persona con amígdalas inflamadas y que no tenga fiebre padezca la enfermedad?

Solución.-a)

$$\mathrm{P}\left(E|\ \mathrm{T}\cap\mathrm{F}\right)=\frac{P(E\cap T\cap F)}{P(T\cap F)}=\frac{P(E)P(T\cap F\,|\,E)}{P(T)P(F)}=$$

$$\frac{P(E)P(T\,|\,E)P(F\,|\,E)}{P(T)P(F)}=\frac{0,003*0,892*0,980}{0,9937*0,0199}=0,1326$$

$$P(T)=P(T\cap(E\cup E^c))=P(E)P(T/E)+P(E^c)P(T/E^c)$$

$$=0,003*0,892+0,997*0,994=0,9937$$

$$P(F)=P(F\cap(E\cup E^c))=P(E)P(F/E)+P(E^c)P(F/E^c)$$

$$=0,003*0,980+0,997*0,017=0,01988$$

b)

$$\mathrm{P}\left(E|\ \mathrm{T}\right)=\frac{P(E\cap T)}{P(T)}=\frac{P(E)P(T\,|\,E)}{P(E)P(T/E)+P(E^c)P(T/E^c)}$$

$$=\frac{0,003*0,892}{0,003*0,892+0,997*0,994}=\frac{0,00267}{0,9937}=0,00269$$

c)

$$\mathrm{P}\left(E \mid \mathrm{A} \cap \mathrm{F^c}\right) = \frac{P(E \cap A \cap F^c)}{P(A \cap F^c)} = \frac{P(E \cap F^c)\,\mathrm{P(A/\,E \cap F^c)}}{P(F^c)P(A/F^c)} = \frac{P(E \cap F^c)P(A/F^c)}{P(F^c)P(A/F^c)}$$

$$= \frac{P(E)P(F^c \mid E)}{P(F^c)} = \frac{0,003*0,020}{(0,003*0,020+0,997*0,983)} = \frac{0,00006}{0,980111} = 6{,}1*10^{-5}$$

$$P(F^c) = P(F^c \cap (E \cup E^c)) = P(E)P(F^c / E) + P(E^c)P(F^c / E^c).$$

$$= 0,003*0,020 + 0,997*0,983 = 0,98$$

Tambien $\mathrm{P(F^c)=1\text{-}P(F)=1\text{-}0{,}01988} \simeq 0,98$