NOTA SOBRE EL TEOREMA DE CONVERGENCIA DEL PERCEPTRON

por Andrés Cristobal Lorente.

INTRODUCCION

Por R denotamos un conjunto arbitrario denominado retina. Cualquier subconjunto de R se llama figura o modelo.

Definición 1. Decimos que un predicado P es lineal umbral respecto un conjunto finito (P_i) $1 \leq i \leq n$, de predicados definidos en R, si existe un conjunto de $n+1$ números, (a_{p_i}), $1 \leq i \leq n$, θ, tal que, para todo $X \subset R$

$$P(X) = 1 \quad \text{sí y sólo si} \quad \sum a_{p_i} p_i(X) > \theta$$

La expresión anterior se escribira más brevemente como

$$P(X) = \lceil \sum a_{p_i} p_i(X) > \theta \rceil ,$$

es decir, el valor de P(X) es el valor de verdad de la expresión encerrada por los ángulos $\lceil, \rceil$

Por otra parte, a cada figura $X \subset R$, la hacemos corresponder un vector, cuyas coordenadas sean los correspondientes valores de cada $p_i(X)$. Es decir,

$$X \longrightarrow (p_i(X), \ldots p_n(X))$$

El vector, $(p_i(X), \ldots p_n(X))$ lo denotaremos por $\emptyset(X)$, o más brevemente por $\emptyset$.

Con la letra F indicamos, indistintamente, un conjunto de figuras o el conjunto de vectores correspondientes.

El conjunto (a_{p_i}), de coeficientes, lo representaremos como un vector de un espacio de dimensión la del cardinal del conjunto de predicados. Usaremos la letra A.

El Perceptron es un mecanismo de proceso paralelo para el reconocimiento de modelos geométricos usando propiedades "locales" las cuales combinadas, permiten una caracterización "global" del modelo.

D e f i n i c i ó n 2 . Un Perceptron es un recurso capaz de computar todos los predicados lineal umbral respecto de algún conjunto básico de predicados.

Teorema de Convergencia

El teorema de convergencia asegura la habilidad del perceptron en la tarea de clasificación.

El proceso de clasificación mediante el perceptron puede describirse de la forma siguiente:

Sea H un espacio real de Hilbert, $F_o , F_1 \subset H$, $F_o \cap F_1 = 0$. Sea una sucesión de elementos de $F = F_o \cup F_1$, que denotamos por S.

$$S = \emptyset_1, \emptyset_2 , \ldots , \emptyset_n , \ldots$$

tal que

1) - Cada $\emptyset_i$ de S pertenece a F

2) - Cada elemento de F puede estar repetido en S

El proceso de clasificación presupone la existencia de un vector solución $A^* \in H$ y un escalar $\delta > 0$, tal que

$$\text{Si } \emptyset \in F_1 \qquad (A \mid \emptyset) > \delta$$

$$\text{Si } \emptyset \in F_o \qquad (A \quad \emptyset) < \delta$$

Para encontrar la solución se genera una sucesión S'

$$S' = A_1 , A_2, \ldots A_n , \ldots$$

tal que para algún n

$$A_n = A_{n+1} = A_{n+2} = \ldots$$

Es decir, el vector inicial $A_1 \in H$, cambiará un número finito de veces hasta alcanzar un cierto A_n - no tiene por que coincidir con A^* - para el cual $(A_n | \emptyset)$ tiene signo propio, o sea, el predicado

$$\psi = \lceil (A_n \mid \emptyset) > 0 \rceil$$

tendrá el siguiente comportamiento

$$\text{Si} \quad X \in F_1 \qquad \Psi(X) = 1$$

$$\text{Si} \quad X \in F_o \qquad \Psi(X) = 0$$

En otras palabras, se dice que Ψ separa F en los conjuntos F_o y F_1

La sucesión S' se genera de la S según las reglas siguientes:

a) - Si el elemento $\emptyset_i$ está correctamente clasificado usando el elemento A_i, entonces el elemento $i+1$ de la sucesión S' es el mismo que el anterior. Es decir

$$A_{i+1} = A_i \quad \text{si } (A_i \mid \emptyset_i) < 0 \quad y \quad \emptyset_i \in F_o$$

$$A_{i\ \ 1} = A_i \quad \text{si } (A_i \mid \emptyset_i) > 0 \quad y \quad \emptyset_i \in F_1$$

b) - En caso contrario, el elemento $i+1$ de S' es dado por

$$A_{i+1} = A_i - C_i \emptyset_i \quad \text{si} \quad (A_i \mid \emptyset_i) \geqslant 0 \quad y \quad \emptyset_i \in F_o$$

$$A_{i+1} = A_i + C_i \emptyset_i \quad \text{si} \quad (A_i \quad \emptyset_i) \leq 0 \quad y \quad \emptyset_i \in F_1$$

donde $C_1, C_2, \ldots C_n$ es una sucesión de escalares $C_i > 0$, para todo i

Obsérvese, que si ponemos $F'_o = \{-\emptyset_i \mid \emptyset_i \in F_o\}$ entonces a, b, son equivalentes a las siguientes

a') $A_{i+1} = A_i \quad \text{si } (A_i \mid \emptyset_i) > 0 \quad y \quad \emptyset_i \in F' = F'_o \cup F_1$

b') $A_{i+1} = A_i + C_i \emptyset_i \quad \text{si } (A_i \mid \emptyset_i) \leqslant 0 \ y \quad \emptyset_i \in F' = F'_o \cup F_1$

También, para demostrar que la sucesión S' se estabiliza, basta considerar una subsucesión de S', en la que se excluyen los elementos A_i tales que $(A_i \mid \emptyset) > 0$

Teorema . Sea H un espacio real de Hilbert. Suponemos la existencia de un $A^* \in H$, tal que $(A \mid \emptyset) > \delta$, $\delta > 0$, $\emptyset \in F' \subset H$.

También, suponemos que los elementos de F están acotados, o sea, existe un K tal que $0 < \|\emptyset\| \leq K$, para todo $\emptyset \in F'$. Entonces la sucesión S' se estabiliza, en el sentido de que si, c, C, son tales que para todo i, $c \leq C_i \leq C$, entonces

$$n \leq \frac{C^2}{c^2} \frac{K^2}{\delta^2} \|A^*\|^2$$

RECAPITULACION

En la versión más generalizada de los perceptrones se consideran útiles en los procesos de clasificación de ciertas propiedades geométricas y como sistemas paramétricos de aprendizaje. Sin embargo, son quizás más interesantes algunas limitaciones teóricas del perceptron para el reconocimiento de modelos, por lo cual nosotros pretendemos, esencialmente con este trabajo divulgar y acentuar el estudio de los perceptrones como agradecimiento al Prof. P. Dou que me interesó con su interés y ayuda por estos temas cuyo desarrollo se cree que será importante en los próximos años. También mi agradecimiento al Dr. F. Briones que me ha ayudado con la minuciosa lectura de este trabajo.

BIBLIOGRAFIA

. Marvin Minsky - Seymoor Papert. - Perceptrons. - MIT Press - 1969.