APRENDIZAJE NO SUPERVISADO

Modelo de Hopfield

Sistemas de Inteligencia Artificial - 2024

TABLA DE CONTENIDOS

01. INTRODUCCIÓN

02. MODELO DE KOHONEN

03. MODELO DE HOPFIELD

04. AUTOVALORES Y AUTOVECTORES

05. COMPONENTES PRINCIPALES

06. REGLA DE OJA Y SANGER

03.1

MODELO DE HOPFIELD

Arquitectura de Red

MEMORIAS ASOCIATIVAS

El almacenamiento y recuperación de información por **asociación** con otras informaciones

Consiste en almacenar un conjunto de *p* patrones tal que

- cuando se presenta un nuevo patrón P_{in}, la red responda con el patrón almacenado P_{out} que más se parece a P_{in}

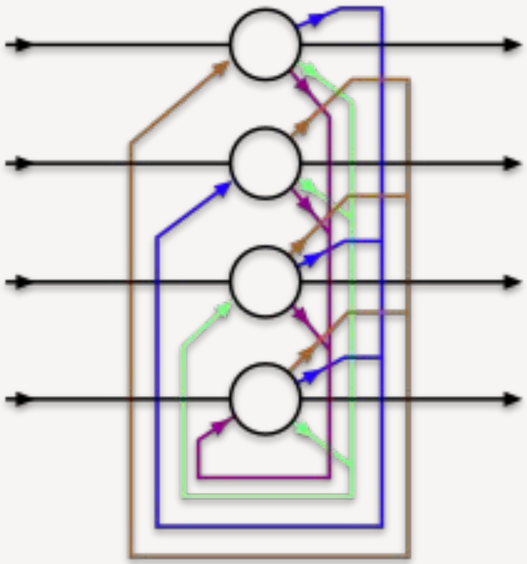
P_{in} = Estímulo

P_{out} = Respuesta

MODELO DE HOPFIELD

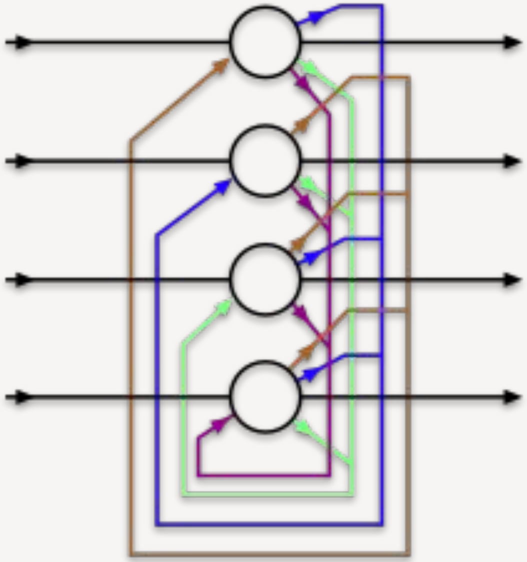
Propuesto en 1982 por JJ. Hopfield

Red Recurrente
Existe retroalimentación entre las neuronas, están todas conectadas entre sí
(a diferenca de las redes feed-forward)



MODELO DE HOPFIELD

- Las neuronas no están conectadas con sí mismas
- Conjunto de input y ouput permitido {-1, 1 } (activo o inactivo)
- Todas las neuronas están en una sola capa de entrada y salida



En esta materia se verá el modelo discreto

OBJETIVO

¿Para qué sirve?

Asociar un patrón de consulta binario (con perturbaciones) con alguno de los patrones almacenados

INPUT: patrón ruidoso
OUTPUT: el más similar

EJEMPLO

1	1
-1	-1

$\xi_1 = (1, 1, -1, -1)$

-1	-1
1	1

$\xi_2 = (-1, -1, 1, 1)$

¿Qué patrón se le asigna al siguiente vector de consulta? $\zeta = (1, -1, -1, -1)$

ESTADOS

Cada neurona tiene 2 estados

- S_i = +1 (activo)
- S_i = -1 (inactivo)

Para una red de N neuronas el estado queda representado por el vector de estados

$$S = [S_1 , S_2, \dots ,S_N]$$

PLANTEO DEL PROBLEMA

Dados los patrones almacenados ξ^{μ} , $\mu = 1, \ldots, p$.

Presentamos un nuevo patrón ζ

Queremos encontrar el patrón almacenado más cercano a ζ (usando redes neuronales)

PLANTEO DEL PROBLEMA

Buscamos los **pesos sinápticos**

$$w_{ij}\ ,\ i = 1, \ldots, N,\ j = 1, \ldots, N$$

tal que la red nos devuelva el patrón ξ^{μ} más cercano a ζ.

MODELO DE HOPFIELD

Es una red neuronal que cumple las siguientes características:

- Cada neurona i es un **perceptrón simple** con la función de activación **escalón** (1, -1).
- Cada par de neuronas (i, j) se conectan por el peso sináptico w_{ij} .

MODELO DE HOPFIELD

Comenzamos con una configuración $S_i = \zeta_i$, i = 1, . . . , n, (patrón de consulta)

Queremos ver si hay un conjunto de pesos w_{ij} que hagan a la red alcanzar el estado $S_i = \xi_i^{\mu 0}$, donde $\xi_i^{\mu 0}$ es un patrón almacenado.

Patrón

$\xi_i^{\mu 0}$, i = 1, . . . , n es el estado del patrón almacenado más parecido a ζ_i , i = 1, . . . , n.

MODELO DE HOPFIELD

Sea

$$h_i = \sum_{j=1}^{N} w_{ij} S_j \, , i \neq j$$

Donde:

- wij es el peso entre la neurona i y la neurona j
- S_j es el estado de la neurona j
- N es la cantidad de neuronas

MODELO DE HOPFIELD

La neurona i modifica su estado S_i de acuerdo a la regla:

$$S_i = \begin{cases} 1 & si \quad h_i > 0 \\ \\ -1 & si \quad h_i < 0 \end{cases}$$

$$S_i = sign(h_i), h_i \neq 0$$

Si h_i = 0 entonces la neurona i permanece en el estado previo.

ESTADO ESTABLE

La red evoluciona hasta que S_i no se modifique más $\forall i = 1, \ldots n$.

Quiere decir que la red alcanzó un **estado estable**

El estado al que evoluciona se lo denomina atractor

CONVERGENCIA

-1	-1
1	1

$\xi = (-1, -1, 1, 1)$

ξ_1 ξ_N

Tenemos N neuronas en la red y un **único patrón** almacenado ξ
Calculamos los w_{ij}

$$w_{ij} = \frac{1}{N}\xi_i\xi_j \qquad i, j = 1, \ldots N$$

$$S_i = sign(\sum_{j=1}^{N} w_{ij}\xi_j) = sign(\xi_i) = \xi_i \forall_i$$

debido a que $\xi_j^2 = 1 \ \forall j = 1, \ldots N$

Por lo tanto, se produce la convergencia: alcanza el patrón almacenado.

MÚLTIPLES PATRONES

1	1
-1	-1

$\xi^1 = (1, 1, -1, -1)$

-1	-1
1	1

$\xi^2 = (-1, -1, 1, 1)$

Sean *p* patrones almacenados.
Podemos elegir los pesos como:

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \quad , i \neq j$$

Esta ecuación incrementa los pesos entre dos neuronas i y j cuando ambos están activados o ambos desactivados.

Se toma también $w_{ii} = 0$ y $w_{ji} = w_{ij}$

03.2

ALGORITMO

ALGORITMO

1. **Almacenamiento:** Sean $\xi_1, \xi_2, \ldots, \xi_p$, patrones N-dimensionales binarios.
2. **Pesos Sinápticos:**

$$w_{ij} = \begin{cases} \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu & si \quad i \neq j \\ 0 & si \quad i = j \end{cases}$$

Donde:
- w_{ij} es el peso sináptico entre la neurona i y la j.
- $\xi_1^\mu \in \{-1, 1\}$

ALGORITMO

3. **Inicialización.** Sea ζ un vector de consulta N-dimensional desconocido, presentado a la red.

$$S_i(0) = \zeta_i \quad i = 1, . . . N$$

- $S_i(0)$ es el estado de la neurona i en el tiempo t = 0.
- ζ_i es el i-ésimo elemento del vector de consulta que queremos asociar con alguno de los patrones almacenados.

ALGORITMO

4. **Iteración hasta la convergencia**

Actualizar los elementos del vector de estado S(t) de acuerdo a la regla:

$$S_i(t+1) = sign(\sum_{j=1}^{N} w_{ij} S_j(t)) \ , i \neq j$$

Repetir la iteración hasta que el vector de estados S permanezca estable.

MÚLTIPLES PATRONES

5. **Output**

Sea S_{fijo} el estado fijo o estable calculado al final del paso 3.

S_{fijo} es el patrón asociado con ζ.

Observar que el vector de pesos permanece fijo y que en cada paso cambian los estados S_i.

03.3

EJEMPLO

EJEMPLO

Cantidad de patrones p = 2

$\xi^1 = (1, 1, -1, -1)$, $\xi^2 = (-1, -1, 1, 1)$

EJEMPLO

Cantidad de patrones p = 2

$\xi^1 = (1, 1, -1, -1)$, $\xi^2 = (-1, -1, 1, 1)$

Cantidad de neuronas N = 4 (porque cada patrón tiene 4 elementos)

¿Cuál es el valor de w_{12} y w_{13}?
Calcular la matriz de pesos W

EJEMPLO

$$w_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \ , i \neq j$$

Cantidad de patrones p = 2

$\xi^1 = (1, 1, -1, -1)$, $\xi^2 = (-1, -1, 1, 1)$

Cantidad de neuronas N = 4 (porque cada patrón tiene 4 elementos)

$$w_{12} = \frac{1}{4}(\xi_1^1\xi_2^1 + \xi_1^2\xi_2^2) = \frac{1}{4}(1+1) = 0.5$$

$$w_{13} = \frac{1}{4}(\xi_1^1\xi_3^1 + \xi_1^2\xi_3^2) = \frac{1}{4}(-1-1) = -0.5$$

EJEMPLO

Cantidad de patrones p = 2

$\xi^1 = (1, 1, -1, -1)$, $\xi^2 = (-1, -1, 1, 1)$

Cantidad de neuronas N = 4 (porque cada patrón tiene 4 elementos)

$$W = \begin{pmatrix} 0 & 0.5 & -0.5 & -0.5 \\ 0.5 & 0 & -0.5 & -0.5 \\ -0.5 & -0.5 & 0 & 0.5 \\ -0.5 & -0.5 & 0.5 & 0 \end{pmatrix}$$

EJEMPLO

Para calcular la matriz de pesos W

K = (ξ^1 ξ^2) (la matriz que se obtiene de poner los patrones en columna)

Entonces

$$W = \frac{1}{N} K K^t$$

$$S(t+1) = sign(WS(t))$$

Recordar que la diagonal debe ser 0

EJEMPLO

Pregunta: ¿Qué patrón se le asigna al vector de consulta $\zeta = (1, -1, -1, -1)$?

$$S_1 = sign(W\ \zeta) = sign \begin{pmatrix} 0 & 0.5 & -0.5 & -0.5 \\ 0.5 & 0 & -0.5 & -0.5 \\ -0.5 & -0.5 & 0 & 0.5 \\ -0.5 & -0.5 & 0.5 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

$$= sign \begin{pmatrix} 0.5 \\ 1.5 \\ -0.5 \\ -0.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

Llegó al patrón almacenado ξ^1

$\xi^1 = (1, 1, -1, -1)$, $\xi^2 = (-1, -1, 1, 1)$

EJEMPLO

$$S_2 = sign(WS_1) = sign\begin{pmatrix} 0 & 0.5 & -0.5 & -0.5 \\ 0.5 & 0 & -0.5 & -0.5 \\ -0.5 & -0.5 & 0 & 0.5 \\ -0.5 & -0.5 & 0.5 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$sign\begin{pmatrix} 1 \\ 0.5 \\ -1.5 \\ -1.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

Se mantiene **estable**

$S1 = \xi^1 = (1, 1, -1, -1)$

03.4

CONVERGENCIA

ESTABILIDAD

$$w_{ij} = \frac{1}{N}\sum_{\mu=1}^{p} \xi_i^{\mu}\xi_j^{\mu} \ , i \neq j$$

¿Se llega a un estado de **Estabilidad**?

Sea ξ_i^{ν} el elemento i del patrón de consulta ν.
La condición de estabilidad es:

$$sign(h_i^{\nu}) = \xi_i^{\nu}$$

Patrones almacenados

$$h_i^{\nu} = \sum_j w_{ij}\xi_j^{\nu} = \frac{1}{N}\sum_j \sum_{\mu} \xi_i^{\mu}\xi_j^{\mu}\xi_j^{\nu}$$

Patron de consulta

ESTABILIDAD

$$h_i^\nu = \frac{1}{N}\sum_j \sum_\mu \xi_i^\mu \xi_j^\mu \xi_j^\nu$$

$$h_i^\nu = \frac{1}{N}\sum_j \sum_{\mu\neq\nu} \xi_i^\mu \xi_j^\mu \xi_j^\nu + \frac{1}{N}\sum_j \xi_i^\nu \xi_j^\nu \xi_j^\nu$$

$$h_i^\nu = \frac{1}{N}\sum_j \sum_{\mu\neq\nu} \xi_i^\mu \xi_j^\mu \xi_j^\nu + \frac{1}{N} \xi_i^\nu$$

crosstalk = 0

$$sign(h_i^\nu) = \xi_i^\nu$$

El sistema es estable si:

- el término de crosstalk es cero.
- Es decir, cuando los **patrones son ortogonales**

  Esta es una de las limitaciones de esta red.

CONVERGENCIA

Hopfield demostró que su red está asociada a una **función de Energía**, dada por:

$$H(w) = -\frac{1}{2}\sum_{i,j} w_{ij} S_i S_j$$

Y que los mínimos locales de esta función son los patrones almacenados.

FUNCIÓN DE ENERGÍA

Propiedad central de una **Función de Energía**

- Siempre decrece (o permanece constante) cuando el sistema evoluciona.
- Como los pesos sinápticos son simétricos:

$$H(w) = -\sum_{j>i} w_{ij} S_i S_j$$

Porque $w_{ij} = w_{ji}$, entonces puedo sumar solo la parte superior de la matriz y multiplicar por 2 (parte superior e inferior de la matriz W) . Además w ii = 0

FUNCIÓN DE ENERGÍA

$$H(w) = -\sum_{j>i} w_{ij} S_i S_j$$

Sea S_i' el nuevo valor de S_i $\quad S_i' = sign \sum_{j \neq i} w_{ij} S_j$

Si $S_i' = S_i$ la energía no se modifica
Si $S_i' = -S_i$ entonces (cambió de signo):

$$\Delta H = H' - H = -\sum_{j>i} w_{ij} S_i' S_j + \sum_{j>i} w_{ij} S_i S_j$$

$$2S_i \sum_{j>i} w_{ij} S_j < 0$$

S_i'

Estamos considerando los S_i' con signo contrario a S_i entonces la energía decrece cada vez que un S_i cambia.

LIMITACIONES

- El número máximo de patrones que puede almacenar es igual al 15 % del número de neuronas de la red. Es decir,

$$p \leq 0{,}15 * N$$

  donde N es la dimensión de los patrones

- Los patrones deben ser “más o menos” ortogonales.

ESTADOS ESPURIOS

- Los patrones son atractores.

- La Función de Energía H puede tener otros mínimos locales que no son los patrones almacenados.

Estados Espurios

También son atractores
Puede desembocar en ciclos

BIBLIOGRAFÍA

[1] McKay D.J.C. *Hopfield Networks*. Information Theory, Inference and Learning Algorithms, Cambridge, 2003.

[2] Anders Krogh John Hertz and Richard Palmer. *Introduction to the Theory of Neural Computation*. Addison-Wesley, 1991.