

Cílem je uchovat množinu vzorů $\{\vec{x}_k \mid k = 1, \dots, p\}$ tak, aby platilo následující:

Po předložení nového vstupu \vec{x} , který je „blízko“ některému \vec{x}_k bude výstup sítě roven \vec{x}_k .

Zejména by síť měla mít schopnost *reprodukce*:
Pro vstup \vec{x}_k by měla dát výstup \vec{x}_k .

Hopfieldova síť

- ▶ Definice
- ▶ Energetická funkce
- ▶ Reprodukce
- ▶ Asociace

Autoasociativní síť.

Organizační dynamika:

- ▶ úplná topologie, tj. každý neuron je spojen s každým
- ▶ všechny neurony jsou současně vstupní i výstupní
- ▶ označme ξ_1, \dots, ξ_n vnitřní potenciály a y_1, \dots, y_n výstupy (stavy) jednotlivých neuronů
- ▶ označme w_{ji} celočíselnou váhu spoje od neuronu $i \in \{1, \dots, n\}$ k neuronu $j \in \{1, \dots, n\}$.
- ▶ žádný neuron nemá bias a předpokládáme $w_{jj} = 0$ pro každé $j = 1, \dots, n$.

Adaptivní dynamika: Dána tréninková množina

$$\mathcal{T} = \{\vec{x}_k \mid \vec{x}_k = (x_{k1}, \dots, x_{kn}) \in \{-1, 1\}^n, k = 1, \dots, p\}$$

Adaptace probíhá podle Hebbova zákona (stejně jako u LAS).
Výsledná konfigurace je

$$w_{ji} = \sum_{k=1}^p x_{kj} x_{ki} \quad 1 \leq j \neq i \leq n$$

Všimněte si, že $w_{ji} = w_{ij}$, tedy matice vah je symetrická.

Adaptaci lze vidět jako hlasování vzorů o vazbách neuronů:

$w_{ji} = w_{ij}$ se rovná rozdílu mezi počtem souhlasných stavů $x_{kj} = x_{ki}$ neuronů i a j a počtem rozdílných stavů $x_{kj} \neq x_{ki}$.

Hopfieldova síť

Aktivní dynamika: Iniciálně jsou neurony nastaveny na vstup $\vec{x} = (x_1, \dots, x_n)$ sítě, tedy $y_j^{(0)} = x_j$ pro každé $j = 1, \dots, n$.

V t -tém kroku aktualizujeme neuron j , který splňuje $t = \tau \cdot (n - 1) + j - 1$ takto:

nejprve vypočteme vnitřní potenciál

$$\xi_j^{(t-1)} = \sum_{i=1}^n w_{ji} y_i^{(t-1)}$$

a poté

$$y_j^{(t)} = \begin{cases} 1 & \xi_j^{(t-1)} > 0 \\ y_j^{(t-1)} & \xi_j^{(t-1)} = 0 \\ -1 & \xi_j^{(t-1)} < 0 \end{cases}$$

Pozn. τ je počet period v nichž se aktualizovaly všechny neurony.

Hopfieldova síť - aktivní dynamika

Výpočet končí v kroku t^* pokud se síť nachází (poprvé) ve *stabilním* stavu, tj.

$$y_j^{(t^*+n)} = y_j^{(t^*)} \quad (j = 1, \dots, n)$$

Věta

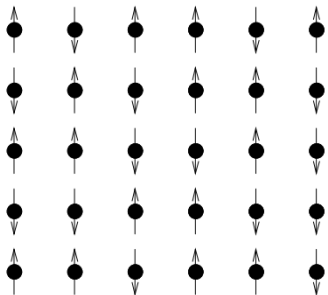
Za předpokladu symetrie vah, výpočet Hopfieldovy sítě skončí pro každý vstup.

Z toho plyne, že Hopfieldova síť počítá funkci z $\{-1, 1\}^n$ do $\{-1, 1\}^n$ (která závisí na hodnotách vah neuronů).

Označme $\vec{y}(W, \vec{x}) = (y_1^{(t^*)}, \dots, y_n^{(t^*)})$ hodnotu funkce sítě pro vstup \vec{x} a matici vah W . Dále označme $y_j(W, \vec{x}) = y_j^{(t^*)}$ složku hodnoty funkce sítě, která odpovídá neuronu j .

Pokud bude W jasné z kontextu, budu psát jen $y(\vec{x})$ a $y_j(\vec{x})$

Jednoduché modely magnetických materiálů připomínají Hopfieldovu síť.



- ▶ atomické magnety poskládané do mřížky
- ▶ každý magnet může mít pouze jednu ze dvou orientací (v Hopfieldově síti $+1$ a -1)
- ▶ orientaci každého magnetu ovlivňuje jednak vnější magnetické pole (vstup sítě), jednak magnetické pole ostatních magnetů (závisí na jejich orientaci)
- ▶ synaptické váhy modelují vzájemnou interakci magnetů

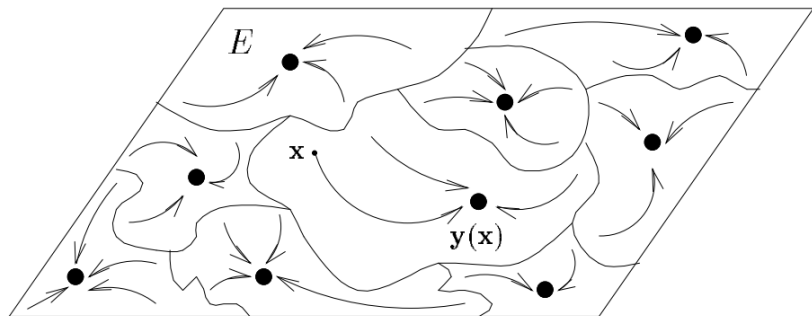
Energetická funkce

Energetická funkce E přiřazuje každému stavu sítě $\vec{y} \in \{-1, 1\}^n$ potenciální energii danou

$$E(\vec{y}) = -\frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n w_{ji} y_j y_i$$

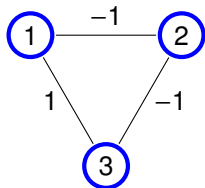
- ▶ velké (kladné) $w_{ji} y_j y_i$ je stabilní a malé (záporné) $w_{ji} y_j y_i$ nestabilní
- ▶ stavy s nízkou energií jsou stabilní (málo neuronů „chce“ změnit svůj stav), stavy s vysokou energií jsou nestabilní

V průběhu výpočtu se energie snižuje: $E(\vec{y}^{(t)}) \geq E(\vec{y}^{(t+1)})$, stav $\vec{y}^{(t^*)}$ odpovídá lokálnímu minimu funkce E .



Obr. 3.4: Energetická plocha.

Hopfield - příklad



y_1	y_2	y_3	E
1	1	1	1
1	1	-1	1
1	-1	1	-3
1	-1	-1	1
-1	1	1	1
-1	1	-1	-3
-1	-1	1	1
-1	-1	-1	1

- ▶ Hopfieldova síť se třemi neurony
- ▶ naučili jsme ji jeden vzor $(1, -1, 1)$ pomocí Hebbova učení (síť se automaticky „naučila“ i vzor $(-1, 1, -1)$)

Pomocí pojmu energie lze snadno dokázat, že výpočet Hopfieldovy sítě vždy zastaví:

- ▶ v průběhu výpočtu se energie nezvyšuje:
 $E(\vec{y}^{(t-1)}) \geq E(\vec{y}^{(t)})$
- ▶ pokud dojde v kroku t ke změně stavu, pak
 $E(\vec{y}^{(t-1)}) > E(\vec{y}^{(t)})$
- ▶ existuje pouze konečně mnoho stavů sítě: výpočet dosáhne lokálního minima funkce E , ze kterého už se nedostane

Hopfieldova síť - odučování

Při učení podle Hebbova zákona mohou vznikat lokální minima funkce E , tzv. nepravé vzory (*fantomy*), které neodpovídají tréninkovým vzorům.

Fantomy je možné odučovat např. pomocí následujícího pravidla: Mějme fantom $(x_1, \dots, x_n) \in \{-1, 1\}^n$ a váhy w_{ji} , pak nové váhy w'_{ji} spočítáme pomocí

$$w'_{ji} = w_{ji} - x_i x_j$$

(tj. podobně jako při adaptaci podle Hebbova zákona, ale s opačným znaménkem)

Hopfieldova síť - reprodukce

Matice všech vah v síti W je dána $W = \sum_{k=1}^p \vec{x}_k \vec{x}_k^T$

Označme $\vec{y}^{(t)} = \begin{pmatrix} y_1^{(t)} \\ \vdots \\ y_n^{(t)} \end{pmatrix}$

Pak $\vec{y}^{(t)} = \text{sgn}(W\vec{y}^{(t-1)})$

Pro jednoduchost předpokládáme, že všechny složky $W\vec{y}^{(t-1)}$ jsou nenulové.

Daný vstup \vec{x} je **stabilní** pokud $\vec{x} = \text{sgn}(W\vec{x})$

Pro daný tréninkový vzor \vec{x}_r máme

$$\text{sgn}(W\vec{x}_r) = \text{sgn}\left(\vec{x}_r + \sum_{k \neq r} \vec{x}_k \frac{(\vec{x}_k^T \vec{x}_r)}{(\vec{x}_r^T \vec{x}_r)}\right)$$

Tedy \vec{x}_r je stabilní pokud $\left| \sum_{k \neq r} \vec{x}_k \frac{(\vec{x}_k^T \vec{x}_r)}{(\vec{x}_r^T \vec{x}_r)} \right| < 1$

Kapacita Hopfieldovy paměti je dána poměrem p/n .

Zde n je počet neuronů a p je počet vzorů.

Předpokládejme, že tréninkové vzory jsou voleny náhodně takto: při volbě \vec{x}_k volím postupně (nezávisle) jednotlivé složky (1 s pravd. $1/2$ a -1 s pravd. $1/2$).

Uvažme konfiguraci W , kterou obdržíme Hebbovským učením na zvolených vzorech.

Označme

$$\beta = \mathbf{P} \left[\vec{x}_k = \vec{y}(W, \vec{x}_k) \text{ pro } k = 1, \dots, p \right]$$

Pak pro $n \rightarrow \infty$ a $p \leq n/(4 \log n)$ dostaneme $\beta \rightarrow 1$.

Tj. maximální počet vzorů, které lze uložit do Hopfieldovy paměti je úměrný $n/(4 \log n)$.

Hopfieldova síť - asociace

Problém:

- ▶ příliš mnoho vzorů implikuje existenci lokálních minim funkce E , která neodpovídají vzorům (tzv. *fantomy*)
- ▶ lokální minima pro vzory mohou dokonce zanikat

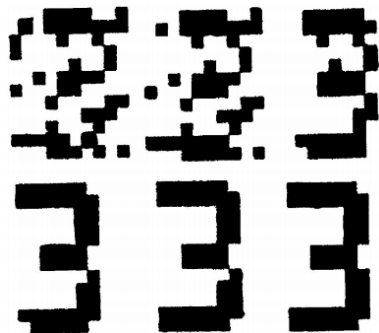
Podrobná analýza ukazuje následující

- ▶ Pro $p \leq 0.138n$ tréninkové vzory odpovídají lokálním minimům funkce E
- ▶ Pro $p > 0.138n$ lokální minima příslušející vzorům zanikají
- ▶ Pro $p < 0.05n$ tréninkové vzory odpovídají globálním minimům E a fantomy mají ostře větší energii

Tj. pro dobré zapamatování 10 vzorů je potřeba 200 neuronů a tedy 40000 spojů ohodnocených celočíselnými váhami
(autorovi knihy z roku 1996 to připadalo nepříliš praktické :-)

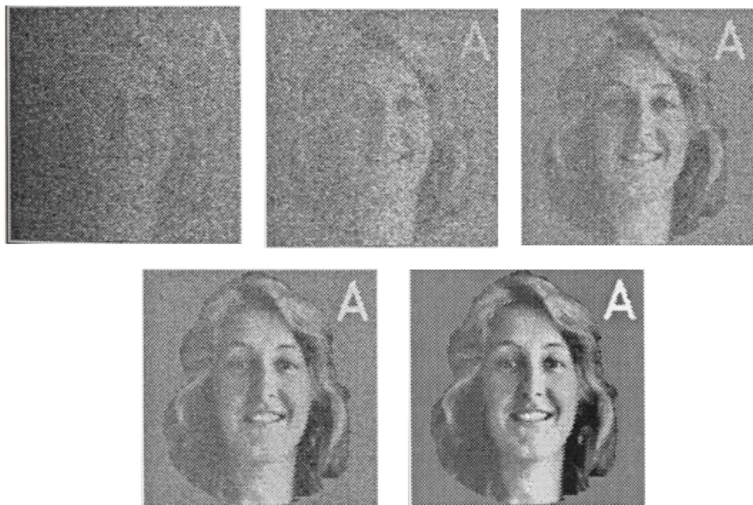
Pozn. Nevýhodou Hopfieldovy sítě je deterministický výpočet, který může skončit v mělkém lokálním minimu E bez možnosti uniknout. Tento problém částečně vyřeší Boltzmannův stroj - stochastické rozšíření Hopfieldovy sítě.

Hopfieldova síť - příklad kódování

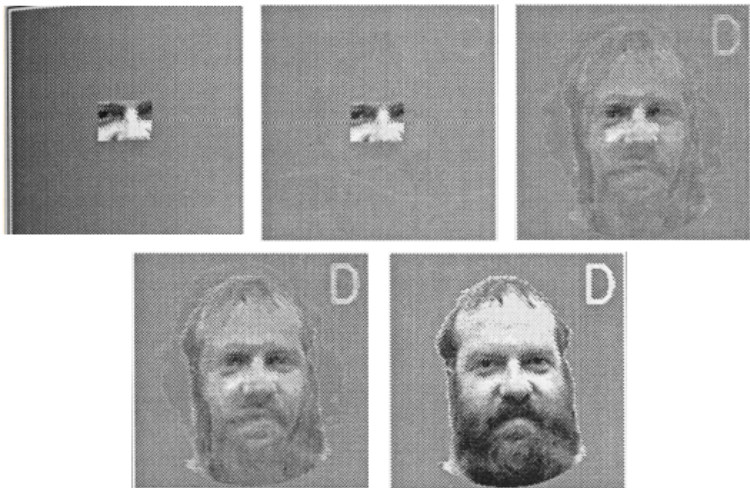


- ▶ číslice 12×10 bodů
(120 neuronů, -1 je bílá a 1 je černá)
- ▶ naučeno 8 číslic
- ▶ vstup vygenerován ze vzoru 25% šumem
- ▶ obrázek ukazuje postup výpočtu Hopfieldovy sítě

Hopfieldova síť - příklad obnovení vzoru



Hopfieldova síť - příklad rekonstrukce vzoru



Hopfieldova síť a optimalizační úlohy

Optimalizační úloha je zadána množinou přípustných řešení a účelovou funkcí. Cílem je nalézt přípustné řešení, které minimalizuje účelovou funkci.

Pro mnoho optimalizačních úloh lze nalézt Hopfieldovu síť (obvykle je nutné přidat biasy) takovou, že

- ▶ energetická funkce $E =$ účelová funkce
- ▶ minima $E \approx$ přípustná řešení

Cílem je nalézt globální minimum funkce E .

Problém: není jasné, v jakém stavu začít, abychom dosáhli globálního minima. Síť může skončit v mělkém minimu.

Řešení: V každém stavu umožníme s malou pravděpodobností přechod do stavů s vyšší energií. Tuto pravděpodobnost budeme postupně snižovat. Tohoto je schopen Boltzmannův stroj, viz. dále ...

Organizační dynamika:

- ▶ Cyklická síť se symetrickými spoji (tj. libovolný neorientovaný graf)
- ▶ Množinu všech neuronů značíme N
- ▶ označme ξ_j vnitřní potenciál a y_j výstup (stav) neuronu j
- ▶ stav stroje: $\vec{y} \in \{-1, 1\}^{|N|}$.
- ▶ označme w_{ji} reálnou váhu spoje od neuronu i k neuronu j .
- ▶ žádný neuron nemá bias a předpokládáme $w_{jj} = 0$ pro $j \in N$.

Botzmannův stroj

Aktivní dynamika: Stavby neuronů jsou iniciálně nastaveny na hodnoty z množiny $\{-1, 1\}$, tj. $y_j^{(0)} \in \{-1, 1\}$ pro $j \in N$.

V t -tém kroku aktualizujeme náhodně vybraný neuron $j \in N$ takto: nejprve vypočteme vnitřní potenciál

$$\xi_j^{(t-1)} = \sum_{i \in j_{\leftarrow}}^n w_{ji} y_i^{(t-1)}$$

a poté náhodně zvolíme hodnotu $y_j^{(t)} \in \{-1, 1\}$ tak, že $\mathbf{P}[y_j^{(t)} = 1] = \sigma(\xi_j^{(t-1)})$ kde

$$\sigma(\xi) = \frac{1}{1 + e^{-2\xi/T(t)}}$$

Parametr $T(t)$ se nazývá **teplota** v čase t .

Boltzmannův stroj - teplota a energie

- ▶ Velmi vysoká teplota $T(t)$ znamená, že $\mathbf{P}[y_j^{(t)} = 1] \approx \frac{1}{2}$ a stroj se chová téměř náhodně.
- ▶ Velmi nízká teplota $T(t)$ znamená, že buď $\mathbf{P}[y_j^{(t)} = 1] \approx 1$ nebo $\mathbf{P}[y_j^{(t)} = 1] \approx 0$ v závislosti na tom, jestli $\xi_j^{(t)} > 0$ nebo $\xi_j^{(t)} < 0$. Potom se stroj chová téměř deterministicky (tj. jako Hopfieldova síť).

Podobně jako u Hopfieldovy sítě definujeme energetickou funkci E

$$E(\vec{y}) = -\frac{1}{2} \sum_{j \in N} \sum_{i \in j_{\leftarrow}} w_{ji} y_j y_i$$

Boltzmannův stroj funguje jako Hopfieldova síť rozšířená o náhodný šum na přechodech, tj. energie se může občas zvýšit.

Pravděpodobnost přechodů do vyšší energetické hladiny se exponenciálně zmenšuje s velikostí energetického skoku.

Simulované žihání

Následujícím postupem lze dosáhnout globálního minima funkce E :

- ▶ Na začátku výpočtu nastavíme vysokou teplotu $T(t)$
- ▶ Teplotu postupně snižujeme, např takto:
 - ▶ $T(t) = \eta^t \cdot T(0)$ kde $\eta < 1$ je blízko 1
 - ▶ nebo $T(t) = T(0) / \log(1 + t)$

Lze dokázat, že při vhodném postupu chlazení dosáhneme globálního minima.

Pozn:

- ▶ Tento proces je analogií žihání používané při výrobě tvrdých kovových materiálů s krystalickou strukturou: materiál se nejprve zahřeje, čímž se poruší vazby mezi atomy, v průběhu následného pomalého chlazení se materiál „usadí“ do stavu s minimální vnitřní energií a s pravidelnou vnitřní strukturou.
- ▶ Jedná se také o rozšíření fyzikální motivace Hopfieldovy sítě: orientace magnetů jsou ovlivněny nejen vnitřním a vnějším magnetickým polem, ale také termálními fluktuacemi.