

Neuronové sítě RBM a DBM

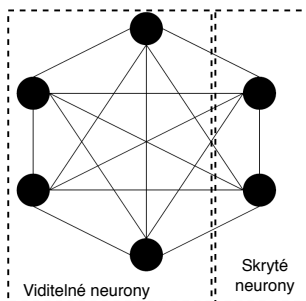
Patrik Lyčka (LYC0017)

Vysoká škola báňská - Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky

24. 3. 2020

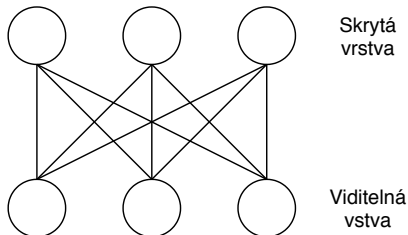
Boltzmann machine

- Síť má 2 druhy neuronů - skryté a viditelné
- Neurony jsou spojeny každý s každým
- Využívá principu simulovaného žíhání
- Velmi pomalý proces učení



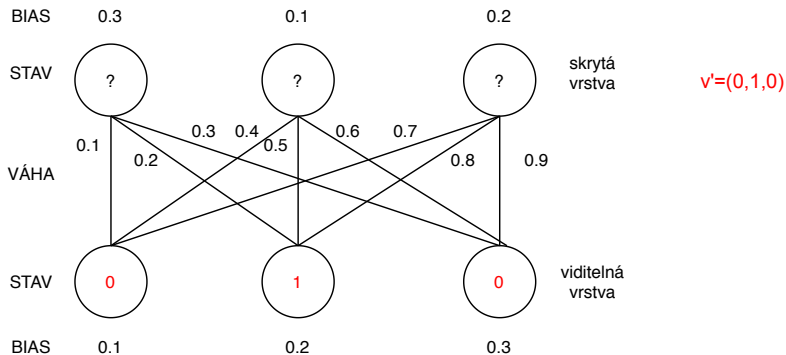
Rerstricted Boltzmann machine

- Došlo k omezení vazeb mezi neurony
- Síť má 2 vrstvy - skrytou a viditelnou
- Neurony v rámci stejné vrstvy mezi sebou propojeny nejsou
- Učení probíhá pomocí algoritmu Contrastive divergence
- Síť rekonstruuje vzory stejně jako Hopfieldovy sítě



RBM - Contrastive divergence

- Zapamatujeme si vstupy namapované na viditelné neurony jako v'_i .



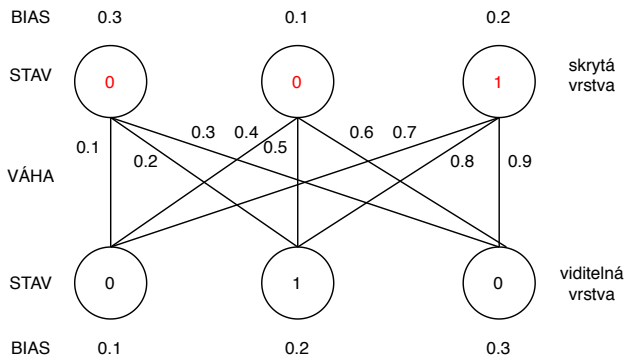
RBM - Contrastive divergence

- Aktualizujeme stavy skrytých neuronů. σ určuje pravděpodobnost s jakou se neuron přepne do stavu 1. Jako vstup x se použije aktivační energie neuronu h_j .
- Každý skrytý neuron si svůj stav zapamatuje jako h'_j

$$h_j = b_j + \sum_{i=1}^N w_{ij} v_i \quad (1)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

RBM - Contrastive divergence



$$v'=(0,1,0)$$

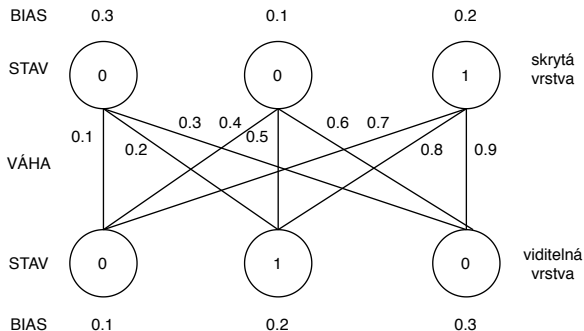
$$1) h=(0.5,0.6,1.0)$$

$$2) \sigma=(0.62,0.65,0.73)$$

$$3) h'=(0,0,1)$$

RBM - Contrastive divergence

- Pro každou vazbu w_{ij} vypočteme hodnotu $Positive(w_{ij}) = v'_i * h'_j$



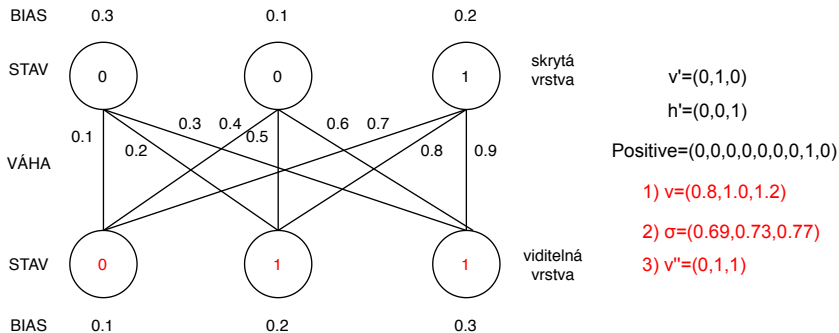
$v'=(0,1,0)$

$h'=(0,0,1)$

$Positive=(0,0,0,0,0,0,0,0,1,0)$

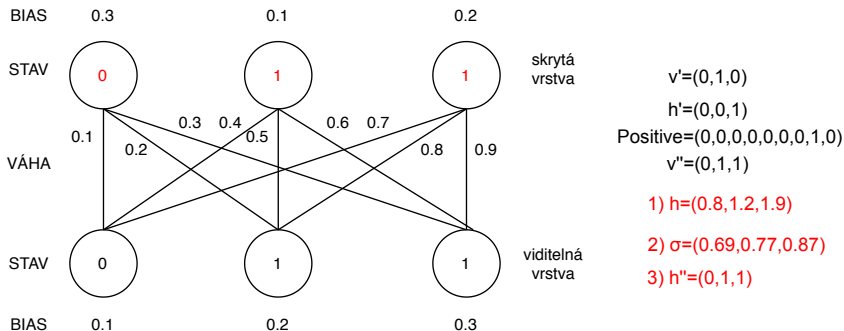
RBM - Contrastive divergence

- Analogickým způsobem aktualizujeme stavy všech viditelných neuronů.
- Každý viditelný neuron si svůj stav zapamatuje jako v_i'' .



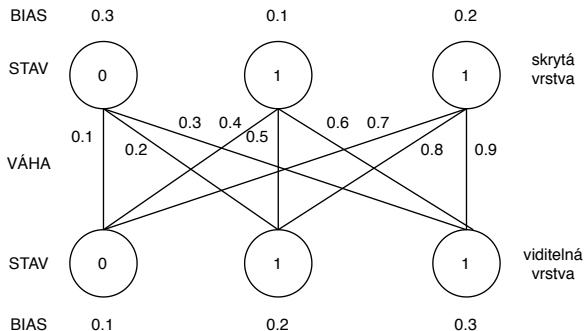
RBM - Contrastive divergence

- Opět aktualizujeme stavy všech skrytých neuronů.
- Každý skrytý neuron si svůj stav zapamatuje jako h'_i .



RBM - Contrastive divergence

- Pro každou vazbu w_{ij} vypočteme hodnotu
 $Negative(w_{ij}) = v_i'' * h_j''$



$v'=(0,1,0)$

$h'=(0,0,1)$

Positive=(0,0,0,0,0,0,0,1,0)

$v''=(0,1,1)$

$h''=(0,1,1)$

Negative=(0,0,0,0,1,1,0,1,1)

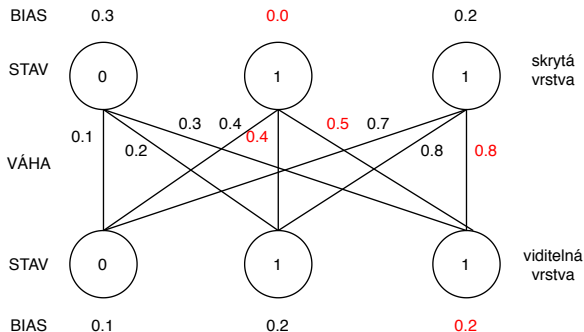
RBM - Contrastive divergence

- Aktualizujeme váhy všech vazeb a biasů.

$$w_{ij} = w_{ij} + \alpha(\text{Positive}(w_{ij}) - \text{Negative}(w_{ij})) \quad (3)$$

$$a_i = a_i + \alpha(v_i' - v_i'') \quad b_j = b_j + \alpha(h_j' - h_j'') \quad (4)$$

RBM - Contrastive divergence



$$v'=(0,1,0)$$

$$v''=(0,1,1)$$

$$h'=(0,0,1)$$

$$h''=(0,1,1)$$

$$\text{Positive}=(0,0,0,0,0,0,0,1,0)$$

$$\text{Negative}=(0,0,0,0,1,1,0,1,1)$$

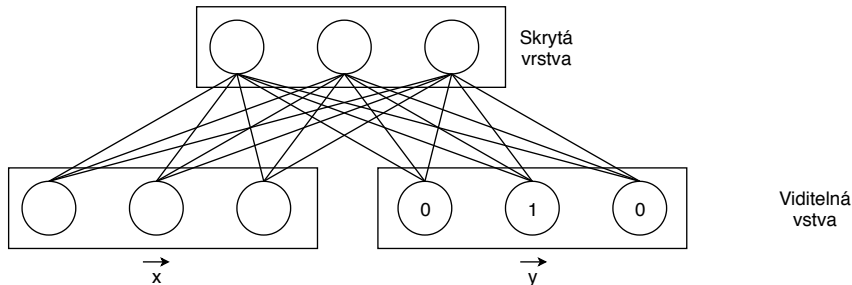
$$\alpha=0.1$$

RBM - Contrastive divergence

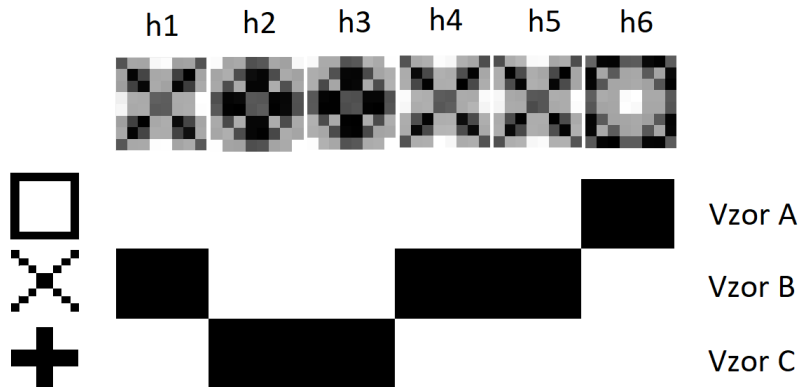
- Opakujeme pro všechny vzory trénovací množiny v několika epochách.
- Při vybavení děláme pouze průchod od viditelné vrstvy do skryté a zpět.

RBM - Klasifikace

- RBM lze použít prostým rozšířením vstupu pro klasifikaci (x - vzor, y - vektor tříd).
- Vektor tříd má jedničku na pozici třídy do které spadá daný vstupní vektor.
- Tento rozšířený vstup lze následně použít pro klasifikaci.

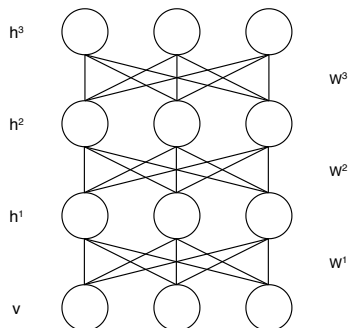


Vizualizace RBM



Deep Boltzmann Machine

- Jedná se o vícevrstvou neuronovou síť jejíž základem jsou RBM.
- Algoritmus učení je značně komplikovanější oproti RBM. Skládá se ze dvou fází - předtrénování a trénování.

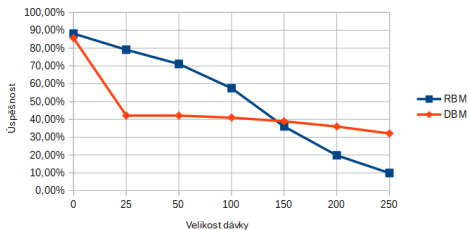


Datový paralelismus

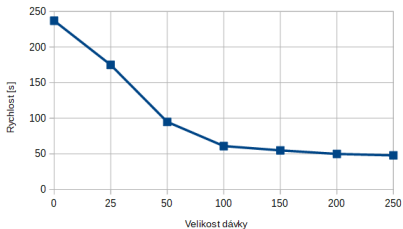
- Trénovací množinu rozdělíme na několik dávek, kde každá dávka obsahuje N vstupů.
- Výpočet probíhá paralelně. Každému uzlu je přiřazena dávka. Po zpracování dávky dochází k synchronizaci vah.
- Klíčové je nastavení správné velikosti dávky.
- Experimenty zobrazující vliv velikosti dávky na úspěšnost a čas jsou provedeny na 9 uzlech.
- Experimenty zobrazující vliv počtu uzlů na úspěšnost a čas jsou provedeny při velikosti dávky 50.

Datový paralelismus

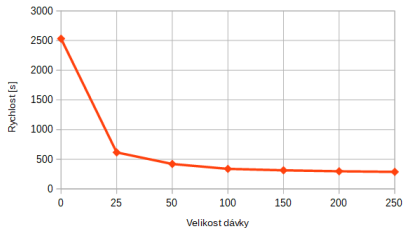
Vliv velikosti dávky na úspěšnost



Vliv velikosti dávky na rychlost (RBM)

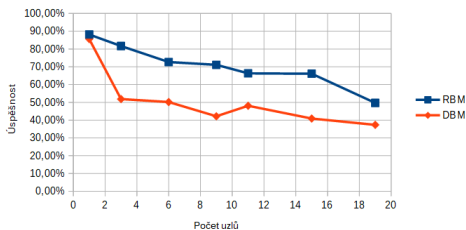


Vliv velikosti dávky na rychlost (DBM)

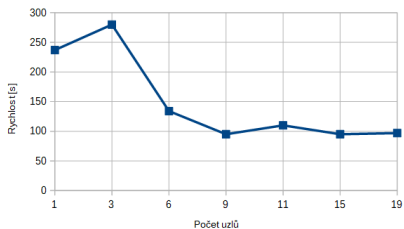


Datový paralelismus

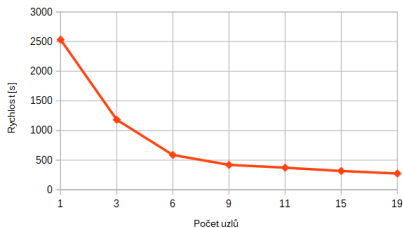
Vliv počtu uzlů na úspěšnost



Vliv počtu uzlů na rychlost (RBM)



Vliv počtu uzlů na rychlost (DBM)



Literatura



Tal Ben-Nun and Torsten Hoefler, *Demystifying parallel and distributed deep learning: An in-depth concurrency analysis*, ACM Computing Surveys (CSUR) **52** (2019), no. 4, 1–43.



Geoffrey E Hinton, *A practical guide to training restricted boltzmann machines*, Neural networks: Tricks of the trade, Springer, 2012, pp. 599–619.



Hugo Larochelle and Yoshua Bengio, *Classification using discriminative restricted boltzmann machines*, Proceedings of the 25th international conference on Machine learning, ACM, 2008, pp. 536–543.



Ruslan Salakhutdinov and Hugo Larochelle, *Efficient learning of deep boltzmann machines*, Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010, pp. 693–700.



Jason Yosinski and Hod Lipson, *Visually debugging restricted boltzmann machine training with a 3d example*, Representation Learning Workshop, 29th International Conference on Machine Learning, 2012.