

Neuronové sítě

L. Horký*, K. Brinda**

Fakulta jaderná a fyzikálně inženýrská, Břehová 7, 115 19 Praha 1

*horkyladislav@seznam.cz, **brinda@fjfi.cvut.cz

Abstrakt

Cílem našeho příspěvku je získat uživatelský vhled do problematiky umělých neuronových sítí (UNS) včetně možnosti použití některých softwarových implementací.

1 Úvod

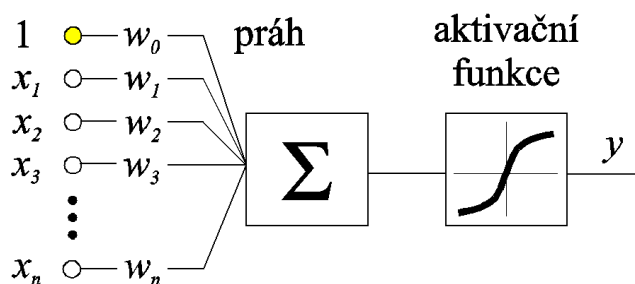
UNS jsou výpočetní modely inspirované biologickými soustavami. Jejich základní jednotka neuron je inspirován biologickým neuronem. Síť se pak skládá z mnoha takových, vzájemně propojených neuronů. UNS se uplatňují v oblastech, jako je predikce, klasifikace, aproximace, dynamické systémy, komprese dat, rozpoznávání obrazu, robotika a mnohé další.

2 Historie

Prvopočátky neuronových sítí sahají do čtyřicátých let dvacátého století, kdy byl poprvé vytvořen matematický model neuronu a navržena základní pravidla učení. První neuropočítače však byly zajímavé spíše z hlediska technického řešení, funkčně příliš nového nepřinášely. Znatelnější výsledky se dostavily na konci padesátých let po vynalezení perceptronové sítě. Byl např. sestaven neuropočítač na rozpoznávání znaků (vstupem fotodiody). V šedesátých letech došlo k velkému boomu UNS - objevily se první společnosti prodávající komerční produkty na bázi UNS, o problematiku se zajímalo mnoho vědců a vývoj pokračoval rychlým tempem. Nekritické nadšení z počátečních výsledků nakonec pomohlo celou oblast zdiskreditovat. Objevovala se tvrzení typu "do několika let budeme mít umělý mozek", nicméně něčeho takového se dostat nepodařilo. V důsledku došlo k odlivu finančních prostředků do jiných oblastí umělé inteligence. UNS se dostaly do popředí znovu až v osmdesátých letech a od té doby výzkum opět intenzivně pokračuje.

3 Modely neuronu

Nejstarším a nejrozšířenějším modelem je McCulloch-Pittsův perceptron (viz obrázek 1). Jeho předobrazem je biologický neuron - má jeden výstup a několik vstupů, z nichž každý je opatřen vahou. Po zadání vstupních hodnot se udělá ze všech vstupů vážený průměr a výsledek se použije jako argument aktivační funkce. Vrácená hodnota je pak přímo výstupem neuronu. Pokud je úroveň výstupu vysoká, říkáme, že je neuron aktivní. Jeden ze vstupů bývá tzv. práh - ten je připojen na nejvyšší úroveň signálu, jeho hodnota je dána jen jeho vahou a udává implicitní potenciál ("napětí") neuronu. Aktivačních funkcí je celá řada, mohou být spojité i binární (výstup jen 0, nebo 1) a podle nich se neurony dělí na spojité, nebo binární (spojité pak dále na lineární a nelineární).



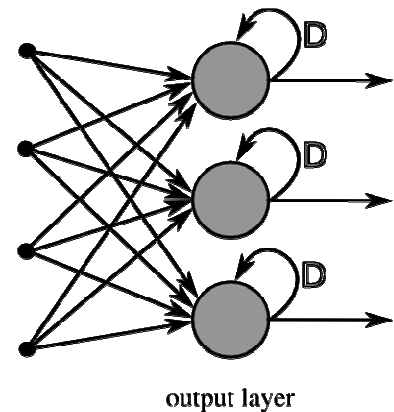
Obrázek 1 - McCulloch-Pittsův perceptron

Volba vhodné aktivační funkce závisí nejvíce na povaze problému/modelu a podrobně se jí zabývá teorie UNS. U speciálních sítí dokonce nemusí být ani u všech neuronů stejná.

Speciálním typem neuronu je RBF neuron (Radial Basis Function), který využívá toho, že se na vstupy a na váhy můžeme dívat jako na dva vektory a zkoumat jejich vzdálenost (x_i a w_i jsou složky vektorů). Místo váženého průměru se použije kvadrát euklidovské vzdálenosti x a w a aktivační funkce bývá gaussovská (tzn. neuron je aktivní, jen pokud jsou x a w dostatečně blízko). Tyto neurony se nejčastěji využívají při klasifikaci vzorů (naučený vzor je uložen ve vektoru vah) a aproximaci.

4 Topologie

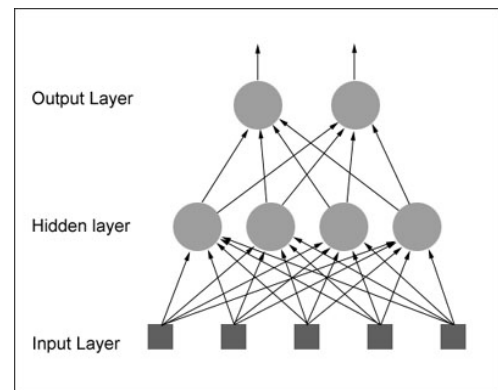
Topologie sítě určuje způsob vzájemného propojení neuronů. Většina sítí má neurony organizovány ve vrstvách, které jsou mezi sebou propojeny systémem každý s každým. Každá síť obsahuje vstupní a výstupní vrstvu (mohou být i totožné - jednovrstvé sítě) a může obsahovat několik skrytých vrstev. Podle topologie můžeme sítě v základě rozdělit na přímé (dopředné, obousměrné) a rekurentní (viz obr. 2,3). Vstupní resp. výstupní data označujeme jako vstupní resp. výstupní vektor, dimenze vektorů odpovídají počtu vstupů resp. výstupů sítě v určitých případech se mohou shodovat (Hopfieldova síť).



Obrázek 2 – rekurentní topologie

5 Učení

Učení sítě obstarává tzv. učící algoritmus (iterativní), který během učení mění váhy jednotlivých spojů. Teorie učení UNS je postavená kolem Hebbova zákona, publikovaného ve čtyřicátých letech, který je inspirován způsobem učení mozku: "Pokud jsou dva neurony v jednom okamžiku aktivní (jeden vybudil druhý), zesil jejich vazbu (příslušná úprava vah), v opačném případě ji zeslab." Tento zákon implementuje v obměnách většina učících algoritmů. Naučení sítě je tedy uloženo ve vahách spojů mezi neurony, samotná topologie se při učení nemění (lze realizovat pouze přerušení spojení neuronů nastavením příslušné váhy na 0). U učení rozlišujeme dva druhy: učení s učitelem a bez učitele. V prvním případě máme k dispozici páry vstupní a k němu příslušející výstupní vektor z reálně naměřených dat (tzv. trénovací množina). Algoritmus porovná požadovaný výstup se skutečným výstupem sítě a upraví váhy spojů tak, aby se výstup přiblížil požadované hodnotě. To provede s každým prvkem trénovací množiny. Při učení bez učitele má síť k dispozici pouze vstupní data a hledá v nich shluky - podobná data (shluková analýza). Algoritmus pak posouvá jednotlivé neurony (reprezentované vektorem vah) ve stavovém prostoru. Proces aplikace naučených poznatků se nazývá vybavování.



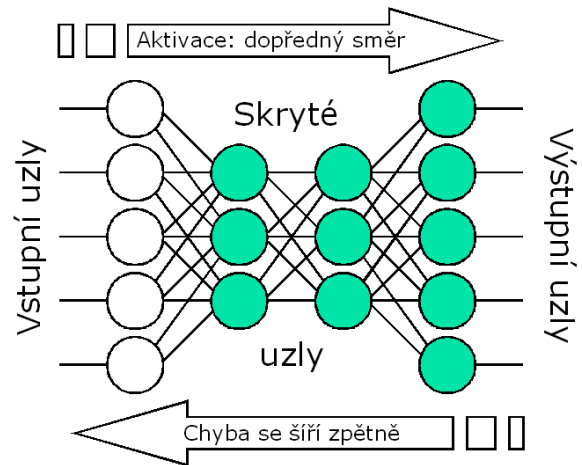
Obrázek 3 - přímá topologie

S učením je také spojen jeden z problémů UNS, konkrétně přeučení - síť sice perfektně odpovídá na vstupy z trénovací množiny, ale není schopna generalizace a odpovědi na reálná data jsou nesmyslné.

6 Některé topologie sítí

1 Dopředná síť (feedforward)

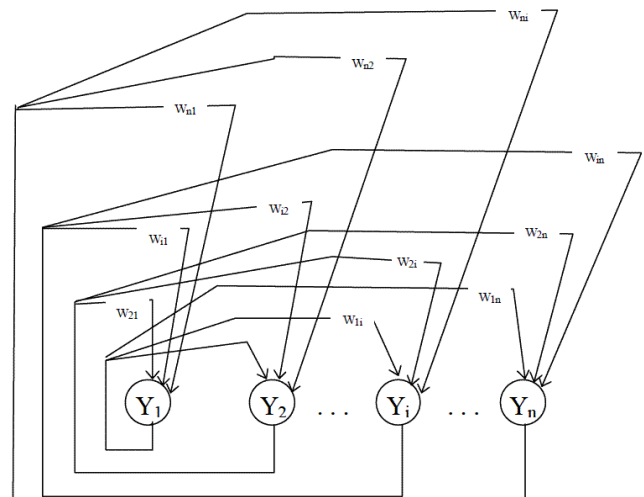
Jedná se o přímou topologii sítě s minimálně jednou skrytou vrstvou a sigmoidální aktivační funkcí (obr. 4). Jako učící algoritmus používá zpětné šíření chyby (backpropagation), kdy se po porovnání skutečného výstupu s očekávaným upravují nejdříve a nejvíce váhy v poslední vrstvě, pak méně v předposlední atd.. Mezi nejčastější použití patří aproximace funkcí. Jako příklad jsme aproximovali funkci $\sin(x)$ v intervalu $0-\pi$ s pomocí dvouvrstvé sítě (3 neurony vstup, 1 neuron výstup), k jejímu naučení jsme použili 10 párů hodnot $(x, \sin(x))$.



Obrázek 4 - obecná dopředná síť

2 Hopfieldova síť

Jedná se o jednovrstvou, plně rekurentní síť, která funguje jako autoasociativní paměť. Zjednodušeně řečeno to znamená, že pokud bude mít síť na vstupu A , má odpovědět A (heteroasociativní by mohla být naučena odpovědět B). Síť se používá při odstraňování šumu ze signálu, případně jeho rekonstrukci. Například v případě obrazu je vstupní vektor tvořen obrázkem rozsekaným do rastru. Vybavování v síti funguje tak, že se přiloží vstupní vektor a neurony se začnou opakovaně vybuzovat, dokud se výstup neustálí - ten je pak brán jako odpověď sítě.

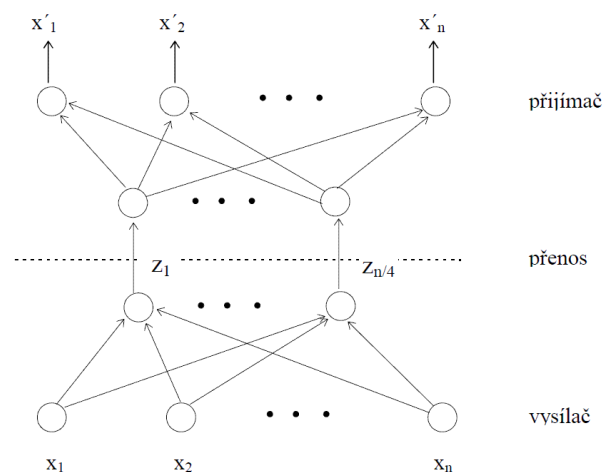


Obrázek 5 - Hopfieldova síť

7 Využití neuronových sítí

1 Komprese signálu

Komprese signálu pomocí UNS se využívá např. u přenosu televizního signálu. Síť (obr. 6. - prostřední dvě vrstvy ve skutečnosti splývají) se naučí tak, aby její odpověď byla stejná jako vstupní vektor. Poté se roztrhne v místě prostřední vrstvy (která má méně neuronů, než vstupní a výstupní - tudíž méně dat) a použije se podle obrázku. Stačí nám tedy přenášet menší objemy dat, ale zato se více projeví případné zkreslení.



Obrázek 6 - příklad neuronové sítě sloužící ke kompresi a dekompresi

2 Robotika

Zde je potenciál zřejmý - prakticky každý řídicí systém v robotice lze nahradit, nebo už je nahrazen UNS. Ať jde o rozpoznávání obzru, rozhodování jednoho, případně skupiny robotů, nebo jakoukoliv jinou automatizaci. Síť se může

učit doslova "koukáním". Můžeme si představit, že člověk řídí auto na dálkové ovládání s pár senzory kolem překážek a síť sleduje jeho reakce na dané situace. Navíc, pokud jí k tomu dáme (výpočetní prostředky), bude schopna síť reagovat správně i za stížených podmínek, větší rychlosti, protože se dostane za hranici lidského reflexu (případně jiných omezujících faktorů). (něco přidat? - asi daisy, když už je v referencích....)

3 Chemické provozy

Místo najímání zkušených a drahých pracovníků může chemický závod (potažmo nakrátko najmutý pracovník) naučit neuronovou síť řídit chemickou reakci (proces, apod.) stejně spolehlivě, jako by to dělal onen pracovník.

8 Některé softwarové implementace

Jelikož programovat neuronovou síť bývá náročné, byly vytvořeny knihovny implementující některé modely neuronových sítí. Mezi nejzdařilejší počiny patří knihovna NeuralNetworks pro matematické prostředí Wolfram Mathematica, více informací viz [2]. Alternativou může být dále např. Neural Networks Toolbox pro prostředí Matlab.

9 Závěr

Ve svém příspěvku jsme vysvětlili, co to je neuronová síť, jak funguje a k čemu se používá. Pro detailní seznámení s problematikou doporučujeme knihy a články uvedené v sekci Reference.

10 Poděkování

Chtěli bychom poděkovat Ing. Zdeňku Bukovi z katedry počítačů Fakulty elektrotechnické Českého vysokého učení technického v Praze za konzultace, rady a laskavé umožnění přístupu k softwaru Wolfram Mathematica s knihovnou NeuralNetworks.

Reference

- [1] J. Šíma, R. Neruda: *Teoretické otázky neuronových sítí*,
<http://www2.cs.cas.cz/~sima/kniha.pdf>
- [2] J. Sjöberg: *Mathematica - Neural Networks*,
<http://www.wolfram.com/products/applications/neuralnetworks/fullmanual.pdf>
- [3] P. Chlumský, *Demonstrační aplikace pro podporu kurzu neuronových sítí*, 2008
- [4] O. Drábek, P. Seidl, I. Taufer, *Umělé neuronové sítě - Základy teorie a aplikace (1)*,
http://www.chemagazin.cz/Texty/CHXV_4_cl6.pdf
- [5] O. Drábek, P. Seidl, I. Taufer, *Umělé neuronové sítě - Základy teorie a aplikace (2)*,
http://www.chemagazin.cz/Texty/CHXV_6_cl2.pdf
- [6] O. Drábek, P. Seidl, I. Taufer, *Umělé neuronové sítě - Základy teorie a aplikace (3)*,
http://www.chemagazin.cz/Texty/CHXVI_1_cl3.pdf
- [7] O. Drábek, P. Seidl, I. Taufer, *Umělé neuronové sítě - Základy teorie a aplikace (4)*,
http://www.chemagazin.cz/Texty/CHXVI_2_cl9.pdf
- [8] I. Taufer, O. Drábek, P. Seidl, *Umělé neuronové sítě - Základy teorie a aplikace (5)*,
http://www.chemagazin.cz/Texty/CHXVI_5_cl6.pdf
- [9] I. Taufer, O. Drábek, P. Seidl, *Umělé neuronové sítě - Základy teorie a aplikace (6)*,
http://www.chemagazin.cz/Texty/CHXVI_6_cl9.pdf

- [10] I. Taufer, O. Drábek, P. Seidl, *Umělé neuronové sítě - Základy teorie a aplikace (7)*,
http://www.chemagazin.cz/Texty/CHXVII_3_c110.pdf
- [11] V. Kvasnička a kol., *Úvod do teórie neurónových sietí*
- [12] E. Volná, *Neuronové sítě 1*
- [13] Y. Zhou, D. Wilkins, R. P. Cook, *Neural Network Control for A Fire-Fighting Robot*,
<http://www.cs.wustl.edu/~zy/Robot.pdf>