

CONTRIBUTION OF PRETRAINING BY AUTOENCODERS IN DEEP NEURAL NETWORKS

Michal Lohniský

Bachelor Degree Programme (3), FIT BUT

E-mail: xlohni01@stud.fit.vutbr.cz

Supervised by: Michal Hradiš

E-mail: ihradis@fit.vutbr.cz

Abstract: This paper evaluates benefit of pretraining by autoencoders in learning of deep neural networks. The pretraining approach was tested on MNIST hand-written digit dataset and CALTECH 101 dataset. Results showed positive contribution of autoencoder pretraining, mainly in combination with Fine tuning.

Keywords: autoencoder, pretraining, deep neural network, fine tuning, backpropagation

1 ÚVOD

Cílem této práce je seznámit čtenáře s výsledky experimentu na hlubokých neuronových sítích. Experiment se zabývá sestavením takovéto sítě a jejím trénováním na dvou datasetech. Cílem experimentu je zjistit, jak velký přínos má předtrénování autoenkodéry na celkové úspěšnosti hluboké neuronové sítě. Předtrénování je porovnáváno s algoritmem Backpropagation.

2 AUTOENKODÉR

Autoenkodér [2] je dopředná třívrstvá neuronová síť (viz. obr. 1), která se dokáže učit bez učitele. Každý spoj mezi dvěma neurony má svoji váhu, kterou je výstup (aktivace) neuronu z nižší vrstvy vynásoben.

Cílem autoenkodéru je na nejvyšší, výstupní vrstvě \hat{x} zrekonstruovat vstupní hodnoty autoenkodéru z nejnižší, vstupní vrstvy x .

Jako aktivační funkce neuronů se u autoenkodéru nejčastěji používá sigmoida $f(z) = \frac{1}{1+\exp(-z)}$.

Při trénování se používá optimalizační algoritmus (jako je například conjugate gradient method) pro minimalizaci objektivní funkce

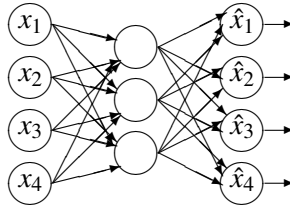
$$J = \frac{1}{m} \sum \frac{1}{2} |h(x) - y|^2, \quad (1)$$

kde m je počet trénovacích vzorků, $h(x)$ jsou dosažené hodnoty výstupu a y jsou výstupní hodnoty autoenkodéru, kterých chceme dosáhnout. Tento algoritmus však potřebuje znát gradienty vah, které optimalizujeme, a k tomu se využívá algoritmus zpětné propagace chyb (Backpropagation).

Složitější neuronové sítě je možné vytvářet vrstvením jednotlivých autoenkodérů na sebe tak, že odstraníme vždy nejvyšší vrstvu nižšího autoenkodéru a odezvy jeho skryté vrstvy použijeme jako vstupní hodnoty nového vyššího autoenkodéru.

3 SOFTMAX KLASIFIKÁTOR

Softmax regrese zobecňuje metodu logistické regrese a používá se ke klasifikaci problémů, ve kterých potřebujeme, aby výstup mohl dosáhnout více než jen dvou hodnot.



Obrázek 1: Neurony (kolečka) v jedné vrstvě autoenkodéru spolu navzájem nekomunikují (nejsou navzájem propojeny). Každý neuron přijímá výstupy neuronů z nižší vrstvy a nabízí své výstupy do vrstvy vyšší.

Softmax klasifikátor je jedna vrstva neuronů, která přijímá vstupy z nižších vrstev sítě. Výstupem je vektor pravděpodobností s Boltzmannovým rozložením pravděpodobnosti. Výstupem je tedy pro vstupní vzorek x pravděpodobnost $p(y = j|x)$ pro každou hodnotu $j = 1, \dots, k$.

Jako aktivační funkce neuronů se používá exponenciální funkce $f(z) = \exp(z)$.

Trénování probíhá s učitelem. Objektivní funkcí je

$$J = \frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y_i = j\} \log \frac{\exp(\Theta_j^T x_i)}{\sum_{l=1}^k \exp(\Theta_l^T x_i)} \right], \quad (2)$$

kde m je počet trénovacích vzorků, k je počet možných výstupních hodnot, Θ jsou váhy softmax vrstvy, x vstupní a y výstupní hodnoty a kde je výraz $1\{\cdot\}$ roven jedné, pokud obsahuje pravdivý výraz a v opačném případě je roven nule. Objektivní funkci se, stejně jako u autoenkodéru, snažíme minimalizovat pomocí optimalizačního algoritmu.

4 TRÉNOVÁNÍ NEURONOVÉ SÍTĚ

Algoritmus Backpropagation je již relativně dlouho používaný algoritmus na trénování hlubokých neuronových sítí. Jeho nevýhoda však spočívá v tom, že špatně optimalizuje nejnižší vrstvy hluboké sítě a proto se provádí předtrénování autoenkodéru.

Při předtrénování se postupuje od nejnižšího k nejvyššímu autoenkodéru v síti. Vždy po dokončení předtrénování nižšího autoenkodéru se odezvy jeho skryté vrstvy použijí jako vstup autoenkodéru nad ním. Jako poslední se trénuje softmax vrstva.

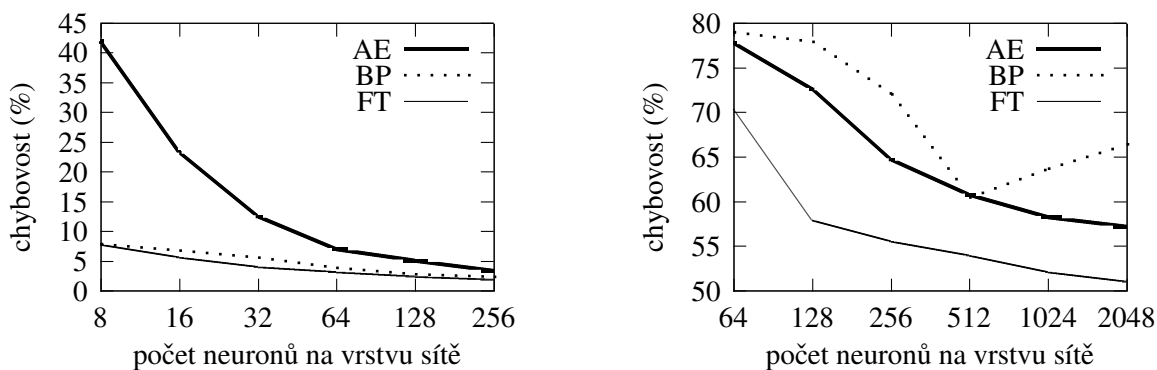
Samotným předtrénováním se neuronová síť učí velmi obecné znaky o vstupním datasetu a pro konkrétní úlohu tedy většinou nedává velmi dobré výsledky. Zlepšení předtrénované sítě přináší aplikace algoritmu Backpropagation, který pracuje se sítí jako s celkem a v jedné své iteraci optimalizuje všechny vrstvy najednou. Aplikaci algoritmu Backpropagation na předtrénovanou neuronovou síť nazýváme Fine tuning.

5 EXPERIMENTY A VÝSLEDKY

Experiment probíhal na dvou datasetech, v prostředí Matlab.

Prvním z datasetů byl dataset MNIST [3], obsahující předzpracované obrázky s ručně psanými číslicemi. Trénování probíhalo na síti se dvěma skrytými vrstvami, vždy na 60 000 vzorcích. Vstupem byly pixely předzpracovaného obrázku o velikosti 28x28 pixelů.

Druhým datasetem byl Caltech 101 [4], obsahující různě velké obrázky, rozřazené do 101 kategorií.



Obrázek 2: Výsledky pro dataset MNIST jsou vlevo, pro dataset Caltech 101 vpravo. V obou případech jsem použil neuronovou síť se dvěma skrytými vrstvami. AE značí předtrénování sítě autoenkodéry, BP algoritmus Backpropagation a FT Fine tuning.

Dataset byl před experimentem předzpracován [1]. Vstupem byl Bag of visual Words, vektor o velikosti 4096. Trénování probíhalo na síti se dvěma skrytými vrstvami a s 1515 vzorky (15 z každé kategorie).

Grafy na obrázku 2 zobrazují chybovost natrénované sítě v procentech při použití různých přístupů ke trénování. V obou případech nejlepších výsledků dosáhla síť předtrénovaná autoenkodéry s následným Fine tuningem.

Z grafů je vidět, že použití ještě větších sítí by mohlo přinést lepší výsledky. Experiment však byl limitován výpočetní technikou.

6 ZÁVĚR

Experiment na dvou různých datasetech ukázal přínos předtrénování autoenkodéry při trénování hlubokých neuronových sítí. Jako nejvhodnější se ukázalo jednotlivé vrstvy sítě předtrénovat autoenkodéry a následně vylepšit síť Fine tuningem. Dále experiment ukázal že přínos předtrénování je větší na sítích s větším počtem neuronů.

Do budoucni by bylo například vhodné experiment rozšířit o výsledky z hlubších a širších neuronových sítí.

REFERENCE

- [1] Hradiš, M., Řezníček, I., Behúň, K.: Brno University of Technology at MediaEval 2011 Genre Tagging Task, In: Working Notes Proceedings of the MediaEval 2011 Workshop, Pisa, Italy, IT, 2011, s. 2, ISSN 1613-0073
- [2] Hinton, G. E., Salakhutdinov, R. R.: Reducing the dimensionality of data with neural networks. Science. 2006, č. 313, s. 504-507. Dostupné z: <http://www.cs.toronto.edu/hinton/science.pdf>
- [3] MNIST handwritten digit database, Yann LeCun and Corinna Cortes. [online]. [cit. 2012-03-26]. Dostupné z: <http://yann.lecun.com/exdb/mnist/>
- [4] Caltech101. [online]. [cit. 2012-03-26]. Dostupné z: http://www.vision.caltech.edu/Image_Datasets/Caltech101/