

## Selece neuronů aktuální vrstvy

Je přirozené, že výše uvedeným algoritmem vypočítané koeficienty  $a$  až  $f$  dobře vyhovují těm vzorům, které jsme pro jejich výpočet použili, ale nemusí být nejvhodnější v případě vzorů jiných. Kromě toho by zcela jistě počet neuronů další vrstvy, který odpovídá počtu kombinací  $C\left(\begin{matrix} N \\ 2 \end{matrix}\right)$ , přesáhl únosnou míru. Z tohoto důvodu je zapotřebí vybrat mezi neurony aktuální vrstvy jen ty, které nejlépe aproximují výstup sítě pro všechny vstupní vektory, tedy provést seleci.

Seleci můžeme například provést následujícím postupem:

- Vypočítáme střední kvadratickou odchylku každého neuronu pro soubor tzv. testovacích dat. Ta jsme získali tak, že jsme soubor vstupních dat předem rozdělili na část použitou při učení sítě (učící data) a na část testovací, kterou zkoumáme, jak kvalitně naučená síť aproximuje. Výpočet této odchylky tedy probíhá se souborem jiných dat, než která byla použita k učení. Výpočet střední kvadratické odchylky může například probíhat podle tohoto vztahu:

$$r^2 = \frac{\sum_{i=1}^k (y_i - \hat{y}_i)^2}{\sum_{i=1}^k y_i^2}, \text{ kde } y_i \text{ je hodnota požadovaného výstupu, } \hat{y}_i \text{ skutečný výstup uvažova-}$$

ného neuronu a  $k$  počet vektorů testovací množiny.

- Neurony aktuální vrstvy seřadíme podle dosažené chyby. To znamená, že neurony s menší chybou budou před neurony s chybou větší.

- V síti ponecháme jen „vyhovující“ neurony. To znamená, že odstraníme neurony z konce setí-děné posloupnosti. Jsou to totiž neurony s největší chybou a nechceme tedy, aby jako takové měly na chování sítě v dalším vliv.

- Neurony, které v aktuální vrstvě zůstaly, „zmrázíme“, jejich koeficienty se při další tvorbě sítě už měnit nebudou a v dalším tyto neurony budou sloužit pouze k distribuci dat do vrstev následujících.

Pod formulaci „ponecháme vyhovující neurony“ si v praxi představme asi toto: zjistíme například, kolik nám zbývá volně paměti pro vytvoření nových neuronů v následující vrstvě. To ovlivní naši volbu, kolik neuronů v aktuální vrstvě ponecháme.

Po selekci a vyřazení je tvorba aktuální vrstvy ukončena a můžeme tedy pokračovat v přidávání dalších vrstev, dokud výstup sítě nespĺňuje naše požadavky. Praktické experimenty prokázaly, že odstranění neuronů s největší chybou nijak celkovou chybu sítě nezhorší.

## Ukončení učení sítě

Tvorba sítě je ukončena, jakmile síť splňuje naše požadavky na výstup. Postupujeme tak, že v každém učícím kroku (po přidání nové vrstvy, vyčíslení koeficientů a selekci) otestujeme, zda nejlepší neuron s požadovanou přesností neaproximuje všechny vstupní vektory testovací množiny. V kladném případě tvorba sítě skončila a právě tento neuron je (jediny) výstupním neuronem sítě. Obvyklý průběh střední kvadratické chyby v závislosti na počtu skrytých vrstev ukazuje následující Obr. 6.4.

V něm si můžeme všimnout faktu, že se s rostoucím počtem vrstev výsledná chyba sítě zmenšuje (uvažujeme střední kvadratickou chybu neuronu dosahujícího v dané vrstvě nejlepší výsledky). Jakmile se dosáhne minima, začíná se obvykle chyba sítě opět zvyšovat. Síť je přeučená. Snažíme se tedy najít globální minimum této chyby a tvorbu sítě po jejím dosažení ukončíme. Přesněji řečeno tvorbu sítě ukončíme,

- v okamžiku, kdy dosáhneme globálního minima, nebo

- v okamžiku, kdy síť dosáhne menší chyby, než požadujeme.