

## Cvičení 1. Trénování vícevrstvé dopředné sítě metodou zpětného šíření chyb (back-propagation)

Program (**backprop.exe**) a data (**aceqstv.dat**) jsou k dispozici na URL:

<http://www.fi.muni.cz/~zizka/> -> [Odkazy](#) -> [Hybridní systémy strojového učení: Cvičení 1 \(program a data\)](#)

Program a data si pro cvičení uložte do svého pracovního adresáře. (Pozn.: Program a data jsou převzaty z přílohy ke knížce od dvojice autorů *Maureen Caudill a Charles Butler: Understanding Neural Networks: Computer Explorations. The MIT Press.* Publikace a software jsou určeny výhradně k výukovým účelům. Původní implementace programu je v grafice MS-DOS, ale bez problémů funguje i pod MS Windows.)

### Postup cvičení:

a) v menu File zvolte New Network. Dále zvolte v menu File položku Set Input a vyberte datový soubor ACEQSTV.DAT. V menu Define vyberte Set Layer Sizes. Ověřte, zda vstupní vrstva má rozměry  $x=5$  a  $y=7$  (5 sloupců a 7 řádků). Výstupní vrstva musí mít  $x=1$  a  $y=8$  (jeden sloupec s osmi pozicemi). Nastavte rozměry tzv. skryté (prostřední) vrstvy (middle layer) tak, že  $x=2$  a  $y=5$ . Po potvrzení se objeví zadané vrstvy na obrazovce. Ověřte správnost rozměrů.

b) V menu Set zvolte Set Parameters. Nastavte Alpha (momentum)=0.0 a Beta (learning—učící konstanta)=0.5.

c) V menu Run ověřte, zda je položka Mode nastavena na Continuous, Training na On, Noise na Off.

d) Soubor ACEQSTV.DAT obsahuje matici s body, znázorňující vybraná velká písmena z anglické (ACEQSTV) abecedy. Výstupním vzorem je pro každé písmeno ASCII kód. V menu Run zvolte Run k zahájení tréningu. V pravém dolním rohu se objeví dva diagramy. Vrchní diagram sleduje celkovou chybu pro každý průchod trénovacích dat a je aktualizován vždy po zpracování celé trénovací dávky. Spodní diagram zaznamenává chybu pro každý vzor dat a je aktualizován po zpracování každého příkladu.

Na obrazovce je (zleva doprava) vidět INPUT PATTERN (vrstva vstupu vzorků dat), MIDDLE LAYER (nelineární transformace vstupu do jiného prostoru), OUTPUT LAYER (výstup—transformace zpět do původního prostoru, de facto 8 bitů ASCII kódu písmene), a nakonec DESIRED OUTPUT (požadovaný výstup, tj. při tréninku je známo, co by mělo ze sítě vystupovat po přiložení vstupního vzoru; rozdíl mezi DESIRED OUTPUT a OUTPUT LAYER tvoří chybu, na jejímž základě dochází k úpravě hodnot vah propojení neuronů mezi vrstvami—cílem je minimalizovat celkovou chybu pro všechny trénovací vzory, tj. najít optimální hodnoty všech vah; počet vah dává počet dimenzí v transformovaném prostoru, kde se hledá řešení). Čím sytější je fialové zbarvení neuronu, tím menší chybu dělá (a naopak), takže lze do jisté míry sledovat opticky změny chyb a kvalitu dosaženého učení.

Sledujte, zda po několika průchodech trénovacích dat chyba podstatně klesá a jak rychle k poklesu dochází.

Pozn. k reprezentaci dat: Vybraná písmena (A, C, E, Q, S, T, D, V) jsou reprezentována formou rastru (jehož velikost je definována prvními dvěma čísly 5 a 7 v prvním řádku souboru ACEQSTV.DAT), kde jsou jedničky a nuly (data jsou v ASCII, lze si je bez potíží prohlédnout). Pod každým rastrem je řádek z osmi čísel (jedničky a nuly), která reprezentují ASCII kód písmene v rastru (např. pod prvním rastrem je 01000001, což odpovídá  $41_{16} = 65_{10} = A$ ). V prvním řádku (po 5 a 7) je 1 a 8, což je definice výstupní vrstvy neuronové sítě. Číslo 7 na konci prvního řádku udává počet tříd (tj. písmen), do nichž se síť učí klasifikovat, čili zařazovat vstupní datové vzory reprezentované uvedenými rastry.

e) Nechte síť dokončit trénování (síť se automaticky zastaví, jakmile je dosažen chybový práh  $\leq 0.10$ ). Stavový box udržuje záznamy o počtu průchodů trénovací množiny. Kolikrát musela síť “vidět” data před naučením?

f) Jakmile se trénování zastavilo, nastavte Mode na One Pattern (jeden vzor). Simulátor sítě bude nadále zpracovávat vzory po jednom. V menu Run nastavte Training Off (vypnuté trénování—váhy již nebudou nadále měněny). Dále volte opakovaně v menu Run položku Start a zpracujte kompletně celý soubor dat, vzor za vzorem. Rozeznává síť všechny vzory?

g) V menu Set nastavte úroveň šumu (Noise Level) na 20% (přidá se tím náhodné číslo v intervalu  $\pm 0.2$  ke každému vzoru). V tomtož menu nastavte Noise On (aby se aplikoval šum). Zvolte opakovaně Start a sledujte, zda natrénovaná síť korektně rozeznává zašuměné vzory. Existují některé vzory, které síť nepoznává správně? Které (a proč asi)?

h) Proveďte testování datových vzorů 2x nebo 3x a sledujte, jak konsistentně síť vzory rozeznává (šum je generován tak, že pokaždé jsou data mírně modifikována jinak).

i) Opakujte body g) a h) za současného zvyšování úrovně šumu na 40%, 60%, 80%, 100%. Zaznamenejte si počet průchodů dat a celkový počet nerozeznávaných vzorů.

j) Shrňte dosažené výsledky. Lze o síti říci, že je z hlediska šumu robustní (tj. určité zkeslení vstupu nemění skokově výstup)?

k) Zájemci mohou experimenty rozšířit dle vlastní vynalézavosti, včetně úpravy a rozšíření datového souboru, který je snadno editovatelný v ASCII kódu.