

## Cvičení 2. Trénování vícevrstvé dopředné sítě metodou zpětného šíření chyb

Program (**backprop.exe**) a data (**aceqstv.dat**) jsou k dispozici na IS MU (studijní materiály předmětu Hybridní systémy strojového učení). Data i software jsou stejné jako v minulém cvičení č. 1.

Program a data si pro cvičení uložte do svého pracovního adresáře.

### Postup cvičení 2:

a) Cílem je porozumět funkci **učící konstanty**  $\beta$  při trénování sítě. V menu *Set* vypněte šum (nastavte na *noise off*). V menu *Run* zvolte *Initialize*. Případně potvrďte, že chcete zrušit dosud natrénované váhy. V menu *Run* zapněte trénování (*training on*) a *mode* nastavte na *continuous* (průběžné učení). Uložte si tuto **nenatrénovanou** síť pod nějakým jménem, abyste její počáteční stav mohli použít vícekrát.

b) V menu *Set* zvolte *Set Parameters* a nastavte  $\beta$  na 0,1. Ujistěte se, že *momentum* je nastaveno na 0.

c) Trénujte síť stejně jako ve cvičení 1 (zadání pro minulé cvičení 1 je zopakováno níže pod čarou). Jakmile je síť natrénována, zaznamenejte si, kolik průchodů trénovacími příklady bylo zapotřebí k naučení sítě rozpoznávat vzory. Opakujte body **g), h) a i) uvedené níže v kopii zadání cvičení 1** a zaznamenejte si, jak dobře se síť naučila **pro danou hodnotu  $\beta$  vždy pro různé hladiny šumu (0%, 20%, 40%, 60%, 80%, 100%)**.

d) Opakujte kroky a), b) a c) a měňte postupně  $\beta$  na 0,2, 0,3, ... , 0,9. Pro každý experiment vždy načtěte inicializovanou (tj. nenatrénovanou) síť, kterou jste si uložili v kroku a), aby trénování vždy začínalo za stejných podmínek. Inicializovaná síť má váhy nastavené náhodně a její opakované spouštění zajistí porovnání výsledků jednotlivých experimentů za stejných výchozích podmínek.

Může se síť správně natrénovat pro  $\beta=1.0$ ? Proč ano či ne? Došlo někdy k tomu, že se síť zcela nenaučila?

e) Inicializujte znovu síť, nastavte  $\beta=0,5$ , natrénujte síť a zaznamenejte si počet trénovacích průchodů příklady. Je tento počet stejný jako před tím? Proč ano nebo ne?

f) Pro  $\beta=0,5$  inicializujte síť a nejméně 10x ji natrénujte, pokaždé si zaznamenejte potřebný počet průchodů příklady do natrénování. Jaký je min a max počet průchodů? Proč je či není počet průchodů pokaždé stejný?

### Kopie zadání cvičení 1:

a) V menu File zvolte New Network. Dále zvolte v menu File Set Input a vyberte datový soubor ACEQSTV.DAT. V menu Define vyberte Set Layer Sizes. Ověřte, zda vstupní vrstva má rozměry  $x=5$  a  $y=7$ . Výstupní vrstva musí mít  $x=1$  a  $y=8$ . Nastavte rozměry skryté 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 písmena 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.

Sledujte, zda po několika průchodech trénovacích dat chyba podstatně klesá.

e) Nechte síť dokončit trénování (síť se automaticky zastaví, jakmile je dosažen chybový práh  $\leq 0,10$ ). Stavový box udržuje záznamy o počtu průchodů tréningové 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 tomtéž 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?

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