

Cvičení 7. Lavinovité sítě (avalanche networks)

Program (**avalanch.exe**) a data (**Sequence.dat** a **Seq2.dat**) jsou k dispozici na IS MU (studijní materiály předmětu Hybridní systémy strojového učení). Program si uložte do svého pracovního adresáře.

Postup cvičení 7:

Simulátor vzhledem k omezení na max. 10 neuronů střední vrstvy je omezen na max. 10 vzorů v trénovací sekvenci. Měnit lze parametry A, T, F a G (podrobný popis lavinových sítí je v přednáškách a materiálech k nim). Pro toto cvičení je nutno nastavit aktivační útlum lavinové vrstvy na 0.5 a práh na 0.3. Doba přenosu τ se předpokládá 1.

a) Spusťte simulátor AVALANCH.EXE. V menu *File* nastavte vstup na SEQUENCE.DAT. V menu *Define* jsou rozměry vstupního a výstupního vzoru již nastaveny. Nastavte rozsah střední vrstvy na 10 neuronů (implicitní hodnota). V menu *Set* nastavte trénovací parametry pro jednokrokový trénovací proces tak, aby konstanty pro *learning* a *forgetting* (učící a zapomínací konstanty) byly $G=0.25$ a $F=0.05$. V menu *Run* nastartujte síť. Sledujte její chování během tréninku. Trénujte síť přibližně 10 cyklů a pak zastavte trénování (cyklem se rozumí kompletní reprodukce celé trénovací sekvence; protože v trénovacím souboru je 6 vzorů, jeden cyklus znamená 6 "tiknutí" hodin). Reprodukují prostřední vrstva sekvenci vzorů? Jaké jste pozorovali úrovně relativní aktivity v prostřední vrstvě během trénování?

b) Pokud je střední vrstva natrénovaná, nastavte mode na *decay only* (tj. pouze útlum) a spusťte síť na dobu jednoho cyklu. Nastavte mode na *one pass* (jeden průchod) a nastartujte síť. Reprodukují střední vrstva celou sekvenci?

c) Pokračujte v trénování jako dříve a občas zkontrolujte, zda se střední vrstva již natrénovala. Kolik trénovacích cyklů bylo zapotřebí k reprodukci sekvence vzorů?

d) Jakmile je střední vrstva natrénována, nastavte parametry sítě tak, aby *learning* a *forgetting* (učení a zapomínání) byly vypnuty ve střední vrstvě. (Pokračujte v trénování na nižší úrovni, cca 0.05.) Nastavte parametry tak, aby ve výstupní vrstvě bylo zvýšeno učení a zapomínání, přičemž poměr učení a zapomínání by měl číselně být asi 5:1. Pokračujte v tréninku. Protože se jedná o obyčejnou síť typu *outstar*, využijte trénovací techniky pro *outstar*. Kolik přídatných cyklů bylo zapotřebí k dokončení tréninku sítě?

e) Nastavte mode na *decay only* a nechte síť, aby jí poklesly aktivační úrovně po dobu jednoho cyklu. K čemu je to dobré? (Pozn.: na rozdíl od *outstar*, lavinová síť provádí některá významná přenastavení aktivity na počátku každého trénovacího cyklu.)

f) Vypněte trénování (*training off*) po dobu jednoho cyklu. Naučila se lavinová síť skutečně reprodukovat sekvenci vzorů? Jak silně reprodukuje výstupní vrstva tréninkové vzory?

g) Otevřete soubor typu report s názvem SEQ1.RPT. Spusťte síť po dobu jednoho cyklu s vypnutým trénováním.

h) Uložte síť do souboru SEQ1.NET.

i) Inicializujte síť a nastavte G a F na 0.25 resp. 0.05 pro lavinovou i výstupní vrstvu. Trénujte síť po stejnou dobu cyklů jako před tím (cca 15 cyklů). Uložte síť do SEQ2.NET. Všimněte si pořadí, v němž se lavinové neurony aktivovaly.

j) Vypněte trénování a otevřete soubor SEQ2.RPT. Spusťte síť nejméně po tři cykly. Bylo snadnější trénovat síť metodou jednoho nebo dvou kroků? Můžete z pouhého pozorování obrazovky říci, že se síť natrénovala přesněji s jednokrokovou nebo dvoukrokovou metodou v bodech 1-6? Ukončete simulátor a srovnajte záznamy SEQ1.RPT a SEQ2.RPT.

k) Nastartujte simulátor a otevřete tentýž trénovací soubor SEQUENCE.DAT. Nastavte úroveň šumu na 0.25 a zapněte šum (*noise on*) v menu *Set*. Použijte tytéž trénovací techniky jako v předchozích krocích (vyberte si jedno nebo dvoukrokovou metodu) pro natrénování sítě se šumem. U lavinových sítí se aplikuje šum na požadované výstupní vzory, nikoliv na vstupní podnět. Uložte natrénovanou síť do NOISY.NET. Trvalo trénování se šumem déle než bez šumu nebo bylo rychlejší? Vytvořte soubor NOISY.RPT, vypněte trénování a spusťte síť po dobu jednoho cyklu. Ukončete simulátor a porovnejte výstupní sekvence v NOISY.RPT s aktuálními výstupy v SEQ1.RPT a SEQ2.RPT. Jaké jsou kvalitativní rozdíly?

l) Použijte PRINTNET ke konverzi binárních souborů vah SEQ1.NET, SEQ2.NET a NOISY.NET do textových souborů SEQ1.WTS, SEQ2.WTS a NOISY.WTS. Porovnejte SEQ1.WTS a SEQ2.WTS. Jsou vidět nějaké kvalitativní rozdíly jiné než ty, které pocházejí z různého pořadí aktivace ve střední vrstvě? Nyní srovnajte váhy v NOISY.WTS s ostatními. Jsou patrné nějaké rozdíly?