

Klasifikácia pomocou neurónovej siete

Vytvorte program realizujúci algoritmus spätného šírenia chyby na doprednej neurónovej sieti. Na daných vstupných dátach (D1, D2 resp. D3) dokumentujte naučenie siete a jej schopnosť zovšeobecnenia. Sieť trénujte po cykloch, čo je jeden prechod tréningovými vzormi v náhodnom poradí. Na konci každého cyklu vyhodnocujte kumulatívnu chybu a výsledky klasifikácie.

Reprezentácie vstupov a výstupov: Kódovanie vstupov a požadovaných výstupov zvolte podľa danej množiny: pri D1 reálne čísla na vstupe a lokalistické reprezentácie (LR) na výstupe, pri D2 binárny kód na vstupe i výstupe, pri D3 binárny kód na vstupe, LR na výstupe. Pri klasifikácii, t.j. pri rozhodovaní či daný výstup sa zhoduje s požadovanou odozvou (nie však pri výpočte chyby), vyskúšajte dva varianty interpretácie výstupov siete v závislosti od prahu  $g$  (0.1 alebo 0.2). V oboch prípadoch považujte skutočný výstup neurónov  $out$  za 1 ( $out \leftarrow 1$ ), keď  $out > 1 - g$ , alebo  $out \leftarrow 0$ , keď  $out < g$  (sieť teda môže “nevedieť” odpovedať). Počiatkové váhy inicializujte ako náhodné malé čísla, napr. z intervalu (-0.5, 0.5).

Postup: Skúste nájsť optimálnu konfiguráciu siete, t.j. sieť s najlepším zovšeobecnením na testovacej množine (pomocou metódy skorého zastavenia učenia). Pod optimálnou konfiguráciou siete rozumieme (a) optimálnu architektúru siete, t.j. počet skrytých neurónov, prípadne pridanie druhej skrytej vrstvy s optimálnym počtom neurónov, a (b) optimálny variant algoritmu učenia a jeho parametre. Pri architektúre vyskúšajte aspoň 3 rôzne počty skrytých neurónov (použite metódu  $k$ -násobnej prekríženej validácie s  $k = 5$ ), a prípadne skúste nájsť optimálny počet neurónov v druhej skrytej vrstve. Spomedzi variantov učenia pomocou spätného šírenia chyby vyskúšajte štandardnú verziu algoritmu, s použitím momentu a metódu klesajúcich váh. Pri tréningu a testovaní náhodne rozdeľte dátovú množinu v pomere 4:1, t.j. 4/5 dát na tréning a zostávajúcu 1/5 dát na testovanie. V každej simulácii inak (náhodne) rozdeľte dátovú množinu na tréningové a testovacie dáta v pomere 4:1, nanovo inicializujte váhy, ale použite identické parametre učenia. Pri hľadaní optimálnej konfigurácie siete použite len jednu hodnotu prahu, a to  $g = 0.1$ . Pripojte stručný sprievodný komentár ku svojim krokom.

Výsledky: Pre každú vyskúšanú architektúru a zvolené parametre a  $g = 0.1$  v tabuľke uveďte finálne výsledky klasifikácie (správne klasifikácie, nesprávne klasifikácie, “nevedela”, všetko v %) pre tréningové i testovacie dáta. Výsledky musia odpovedať priemerným hodnotám z 5-ich simulácií (s uvedením aj štandardnej odchýlky od priemeru). Okrem toho, pre Vami nájdenú optimálnu konfiguráciu siete, uveďte tieto výsledky aj pre  $g = 0.2$  (pre porovnanie), a graficky znázorníte priebeh chyby na tréningových i testovacích dátach v jednej zvolenej simulácii s optimálnou konfiguráciou siete. Výsledky okomentujte. Víťaz v každej kategórii (D1,D2,D3 a  $g = 0.1$ ) získa bonus 5 bodov!

Dátovú množinu si vyberte podľa iniciály Vášho priezviska: A-J iris.dat (D1), K-O nasob.dat (D2), P-Z parita.dat (D3). Všetky 3 dátové súbory sú na [www.ii.fmph.uniba.sk/~farkas/Courses/meno\\_suboru.dat](http://www.ii.fmph.uniba.sk/~farkas/Courses/meno_suboru.dat). Všetky 3 súbory obsahujú dáta v tvare vstup | výstup.

Projekt musí byť vypracovaný samostatne, musí obsahovať požadované textové a grafické výstupy (pozor zmena: zdrojový kód programu netreba). Splnenie projektu predpokladá dva kroky: (1) prezentáciu modelu na cvičení (Pia, 2.4.), a (2) odovzdanie vypracovaného projektu (vo formáte PDF) emailom na [farkas@ii.fmph.uniba.sk](mailto:farkas@ii.fmph.uniba.sk) (napr. eleonora.patoprsta.proj2.pdf).