

Vytěžování dat, cvičení 12:

Neuronové sítě typu Back Propagation

Miroslav Čepek



Evropský sociální fond
Praha & EU: Investujeme do vaší budoucnosti

Fakulta elektrotechnická, ČVUT

- ▶ V software Rapidminer experimentujte s implementací neuronové sítě typu back-propagation.
- ▶ Touto sítí klasifikujte dodaná data (Glass identification dataset viz. <http://archive.ics.uci.edu/ml/datasets/Glass+Identification>).
- ▶ Experimentujte s počtem skrytých vrstev a počty neuronů v těchto vrstvách - najděte co možná nejpřesnější a nejméně přesnou konfiguraci neuronové sítě.
- ▶ Zkuste najít minimální počty neuronů ve skrytých vrstvách, pro které síť ještě dokáže data klasifikovat s rozumnou chybou.
- ▶ Pro odhad chyby vždy použijte 10 násobnou Cross-validaci.

Zadání domácího úkolu (2)

- ▶ Pro nejpreciznější a nejméně přesnou konfiguraci neuronové sítě vytvořte matici záměn (confusion matrix) a ilustруйте na nich rozdíly v klasifikaci.
- ▶ V datech změňte výstupní třídu tak, aby indikovala "Sklo typu 2 nebo 3" a "Všechna ostatní skla".
- ▶ Pro takto změněná data a několik neuronových sítí (viz další bod) zobrazte ROC křivku a okomentujte, co ROC křivka říká. Na jejich základě učiňte závěr, která konfigurace neuronové sítě dává nejlepší výsledky.
- ▶ Použijte NS s jedním neuronem v obou skrytých vrstvách a pak vaši nejlepší a nejhorší nalezenou konfiguraci.

- ▶ Obsahem zprávy bude:
 - ▶ popis výsledků pro jednotlivé konfigurace (počty skrytých vrstev a počty neuronů v nich) neuronové sítě a volitelně i zobrazení NS v jedné konfiguraci,
 - ▶ popis výsledků získaných z ROC křivek a jejich zobrazení,
 - ▶ váš komentář k předchozím bodům.
- ▶ V domácím úkolu můžete použít Rapidminer proudy, které naleznete na webu cvičení.

- ▶ Neuronová síť typu back-propagation, kterou máte zkoumat se skrývá v uzlu nazvaném `Neural Net`
- ▶ Uzly v programu `Rapid miner`, které se vám také mohou hodit:
 - ▶ `Apply Model`
 - ▶ `Performance`
 - ▶ `Validation`
 - ▶ `Read CSV`