

SOM: VIZUALIZÁCIA ABSTRAKTNÝCH DÁT NEURÓNOVÉ SIETE – PROJEKT 3B

Peter Kováč

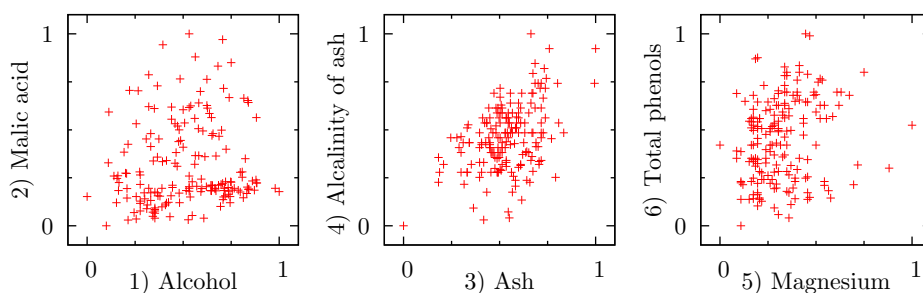
10. mája 2010

Implementoval som samorganizujúcu sa mapu a aplikoval ju na vizualizáciu 13 rozmerného priestoru.

1 Dáta

Dáta (wine set) som pred tréňovaním normalizoval – každý z 13 rozmerov podľa svojho minima a maxima. Celkove bolo v dátovej množine 178 prvkov v troch triedach s rozdelením 59:71:48.

Obrázok 1: Vizualizácia vybraných dvoch rozmerov normalizovaných dát



2 SOM

Na vizualizáciu som použil SOM v tvare mriežky 10×10 neurónov, takže každý neurón zodpovedal v priemere 1,78 vstupného prvku. Počiatočné váhy neurónov som inicializoval rovnomerne náhodne z intervalu $\langle 0, 1 \rangle$.

Celkový počet epoch som nastavil na 10 000. Rýchlosť učenia klesá exponenciálne, ale nikdy neklesne pod hodnotu 0.01 (obr. 2)

$$\alpha(t) = 1.02^{-t} + 0.01.$$

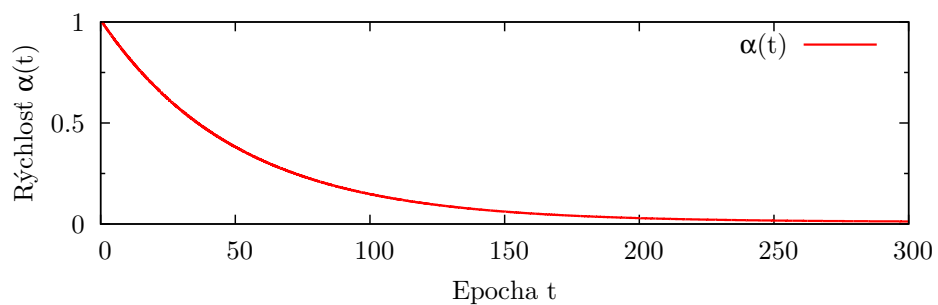
Pri tréňovaní som použil gaussovskú funkciu veľkosti okolia. Označme vstupný vektor x , vĺaza $i(x)$ a nejaký neurón j . Predpis funkcie okolia $h(i(x), j, t)$ je potom

$$h(i(x), j, t) = \exp \left(-\frac{d_e(i(x), j)^2}{2\lambda(t)^2} \right)$$

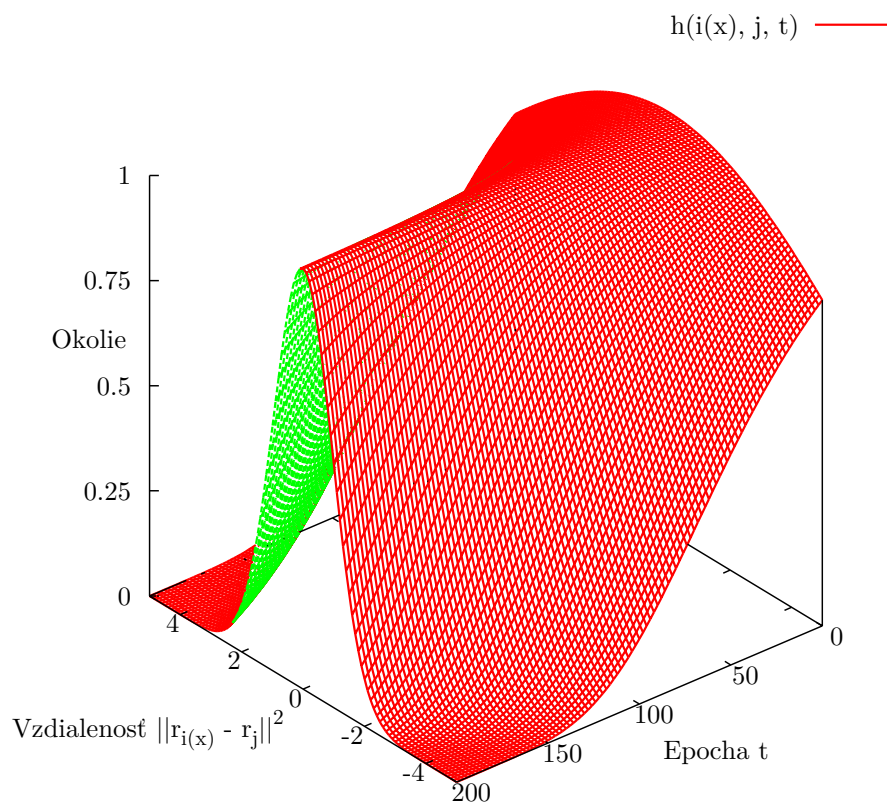
$$\lambda(t) = 7(1.01)^{-t}.$$

Táto funkcia spĺňa požiadavky na funkciu okolia (obr. 3).

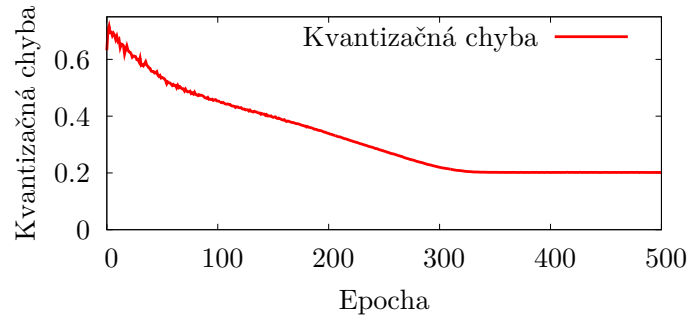
Obrázok 2: Rýchlosť učenia $\alpha(t)$



Obrázok 3: Veľkosť okolia $h(i(x), j, t)$



Obrázok 4: Vývoj kvantizačnej chyby



3 Výsledky

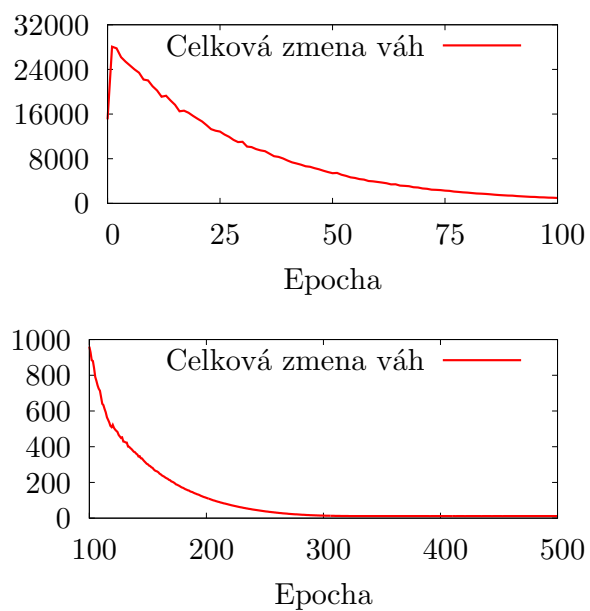
Z obrázkov vývoja kvantizačnej chyby (obr. 4) a celkovej zmeny váh (obr. 5) vidieť, že fáza usporiadavania trvala zhruba 300 epoch. Potom sa už váhové vektory neurónov príliš nemenili.

Výsledná mapa (obr. 6) približne odzrkadľuje pomerné zastúpenie tried v dátovej množine. Skutočný pomer tried je 33%:40%:27%, pomer na mape je 28%:45%:27%. Pri použití vyššieho počtu neurónov v mape by sa dalo dosiahnuť vyššie „rozlíšenie“ a teda väčšia presnosť.

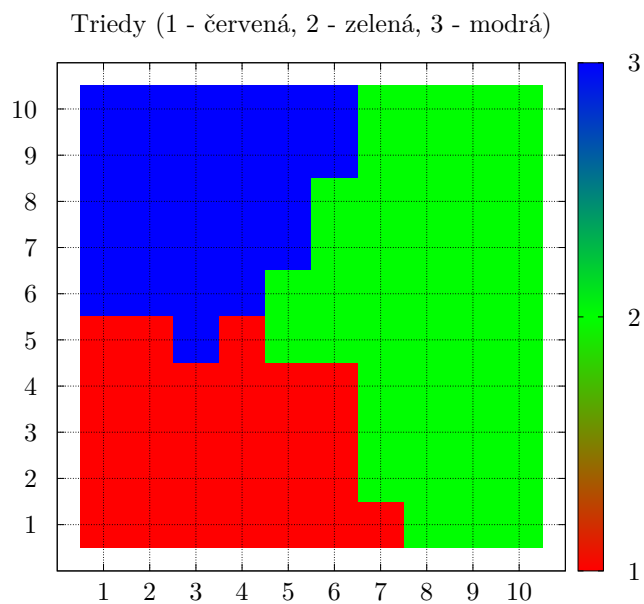
4 Magnifikačný faktor

Na otestovanie magnifikačného faktora mapy som upravil dátovú množinu – prvú a tretiu triedu som zmenšil na polovicu. Celkove teda bolo v dátovej množine 125 prvkov v troch triedach s rozdelením 30:71:24, v percentách 24%:57%:19%. Na výslednej mape (obr. 8) je pomer tried 22%:62%:16%. Aj v tomto prípade platí, že vyšší počet neurónov by mohol spresniť vizualizáciu.

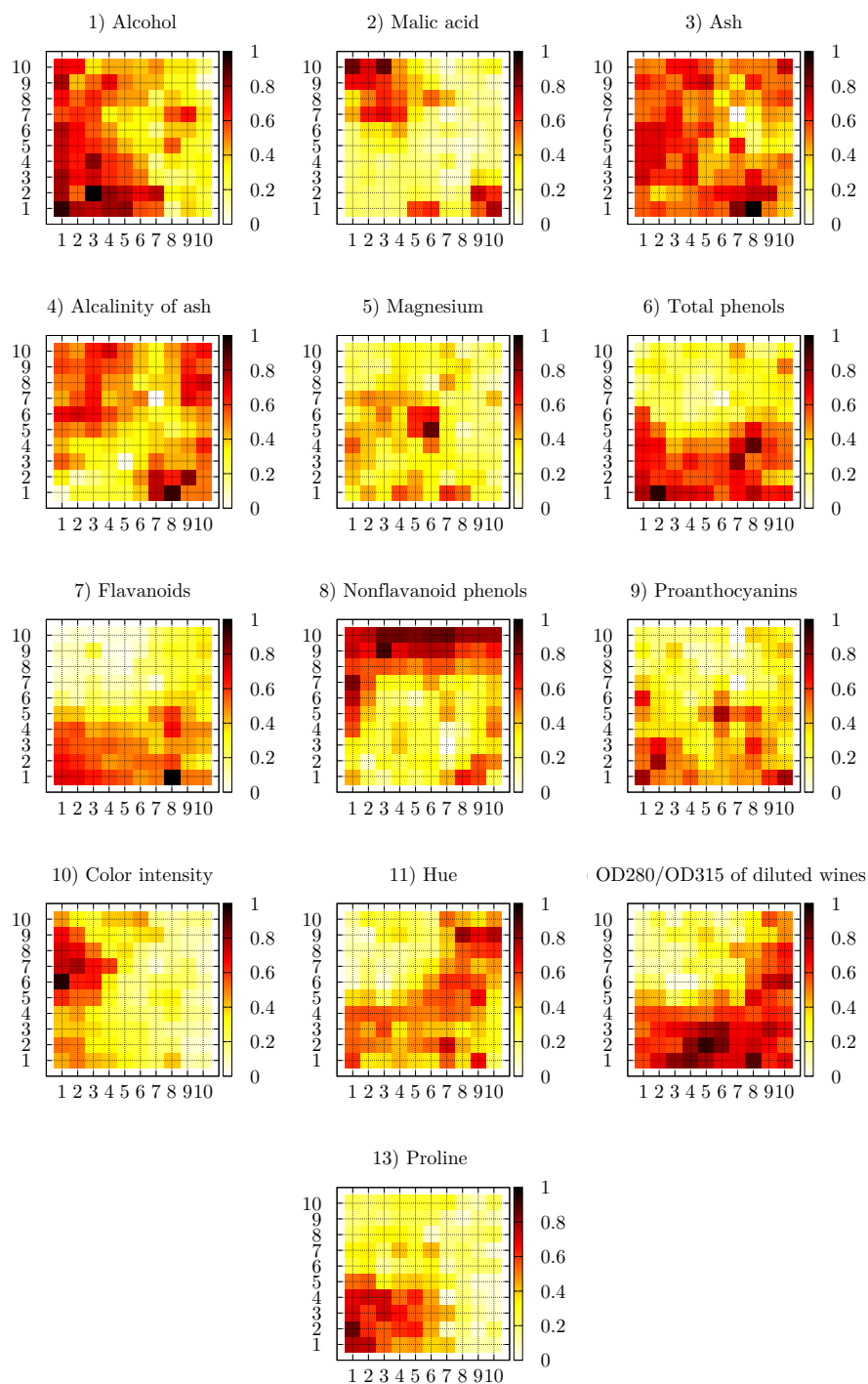
Obrázok 5: Vývoj celkovej zmeny váh



Obrázok 6: Triedy v mape



Obrázok 7: Atribúty v mape



Obrázok 8: Prejav magnifikačného faktora

