

Lekcja 7: Sieci neuronowe - Zastosowanie

S. Hoa Nguyen

1 Zadania podstawowe

Zadanie 1 (Klasyfikacja gatunków *Irisa*) W systemie Weka zaprojektować sieć neuronową umożliwiającą klasyfikację kwiatów *Iris*. Zbiór danych jest umieszczony w katalogu *data*. Przeprowadź uczenie i przetestuj sieć. Zastosuj model testu "5-kroswalidacja".

- a) Zaprojektować dwu-warstwową sieć neuronową
- b) Ustaw *współczynnik uczenia* na 0.1, 0.25, 0.5, 0.75 i 1 oraz *liczba epok* na 500. Obserwuj poziom błędów i wybieraj optymalny współczynnik uczenia.
- c) Dobierać optymalną liczbę neuronów w warstwie ukrytej.
- d) Skomentować otrzymane wyniki

Zadanie 2 (Klasyfikacja wniosków o karty kredytowe) Zaprojektować sieć neuronową umożliwiającą klasyfikację wniosku o kartę kredytową. Opisy wnioskodawców są zapisane w zbiorze danych *data/australian.dat*. Zadaniem jest przewidywanie, czy wniosek będzie pozytywnie rozpatrywany? Przeprowadź uczenie i przetestuj sieć. Zastosuj model testu "10-kroswalidacja" do weryfikacji jakości modelu.

- a) Przeprowadzić zbiór danych uczących i testowych do stosownej postaci w systemie WEKA.
- b) Ustaw *współczynnik uczenia* na 0.1, 0.5, 1 oraz *liczba epok* na 1000. Obserwuj poziom błędów i wybieraj optymalny współczynnik uczenia.
- c) Sprawdź jakość klasyfikacji sieci *dwu-warstwowej*. Dobierać optymalną liczbę neuronów na warstwie ukrytej.
- d) Porównaj jakości sieci z i bez momentum
- e) Podać dokładność klasyfikacji
- f) Obserwując macierz błędów podaj w krócej klasie sieć popełnia najwięcej błędów?
- g) Powtórz kroki **a** - **e** dla sieci *trzy-warstwowej* i porównywać jakość klasyfikacji tej sieci z jakością sieci *dwu-warstwowej*.

Zadanie 3 (Wyznaczenie parzystobitowości) Zaprojektować sieć, która potrafi wyznaczyć parzystobitowość liczb trzybitowych.

- Liczby przedstaw graficznie, po cztery w każdej klasie (patrz Rysunek 1)
- Udowodnij, że sieć jedowarstwowa podczas nauki nie potrafi rozpoznać klasy liczb.
- Przygotuj odpowiedni zbiór danych dla systemu Weka.
- W systemie Weka utwórz sieć wielowarstwową, uczącą się metodą propagacji wstecznej błędu. Znajdź optymalne parametry: współczynnik uczenia, liczba epok, liczba warstw, ilość neuronów na każdej warstwie.



Rysunek 1: Klasyfikacja liczb parzystobitowych

Zadanie 4 (Rozpoznawanie cyfr ręcznie pisanych) Zaprojektować sieć, która potrafi rozpoznawać cyferki ręcznie pisane. Opis zbioru danych jest umieszczony w pliku *optdigit.doc* w katalogu *data/OpticalDigits*

- Przeprowadzić zbiór danych uczących i testowych do stosownej postaci w systemie WEKA (2 punkty).
- Zaprojektować sieć i dobierać odpowiednie parametry: współczynnik uczenia, liczba epok, uczenie z momentum czy nie
- Sporządzić raport: jakość klasyfikacji każdej z nauczonych sieci. Komentować otrzymane wyniki. Która sieć jest optymalna dla podanego problemu?
- Obserwując macierz pomyłek podaj w krócej klasie sieć popełnia najwięcej błędów?

Weka Home page: <http://www.cs.waikato.ac.nz/ml/weka>

Weka Tutorial: <http://www.cs.utexas.edu/users/ml/tutorials/Weka-tut/index.htm>