

# Rozpoznawanie Obrazów

## Laboratorium 5

### Sprawozdanie

**Sławomir Dobrzański**

AGH Marzec 2007

<http://home.icslab.agh.edu.pl/~dobrzan/pr/index.html>

Celem laboratorium było zbadanie jak proste sieci neuronowe nadają się do klasyfikowania elementów zbioru wielowymiarowego.

#### Wstęp

Sieć neuronowa to ogólna nazwa struktur matematycznych i ich programowych lub sprzętowych modeli, realizujących obliczenia lub przetwarzanie sygnałów poprzez rzędy elementów, zwanych neuronami, wykonujących pewną podstawową operację na swoim wejściu. Oryginalną inspiracją takiej struktury była budowa naturalnych układów nerwowych, w szczególności mózgu. Takie właśnie struktury będziemy wykorzystywać w tym ćwiczeniu.

Bagging jest techniką poprawiającą klasyfikacje poprzez zwielokrotnienie predyktów, cały mechanizm można opisać w kilku krokach:

- zbiór uczący dzielimy na  $k$  podzbiorów
- trenujemy model na elementach z  $k$  podzbiorów (każdy podzbiór będzie miał około 63.2% unkalnych elementów – tzw. próbka bootstrap)
- wyniki symulacji są łączone z wykorzystaniem głosowania (dla problemu klasyfikacji)

#### Próbki danych

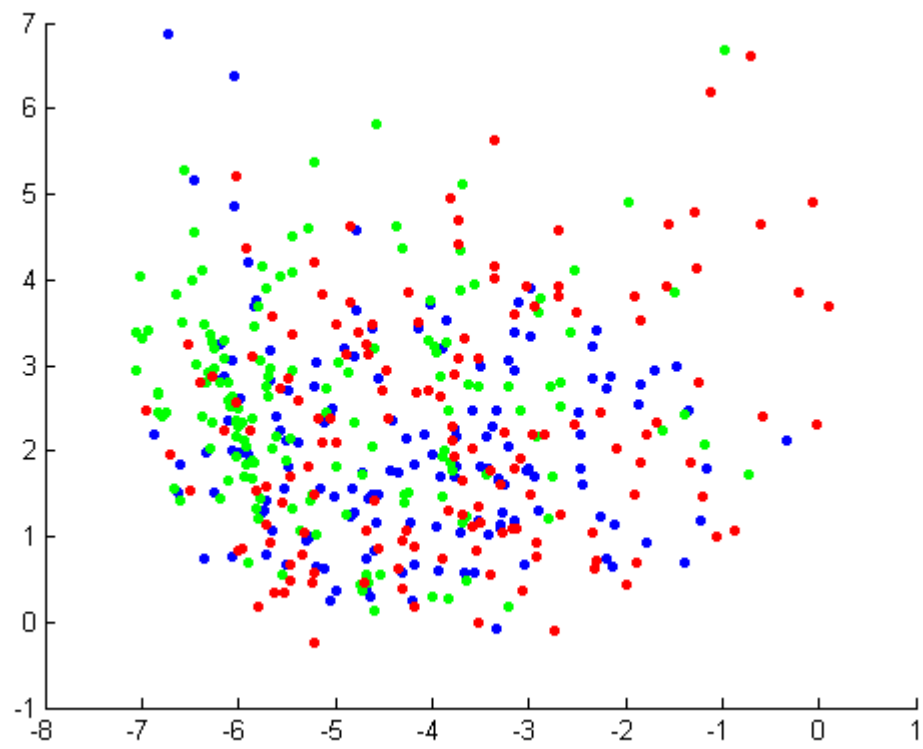
Do przeprowadzenia ćwiczenia użyłem następujących próbek danych:

- **cmc.data**: „*Contraceptive Method Choice*”. Zbiór ten zawiera dane statystyczne na temat rodzaju stosowanej metody antykoncepcyjnej stosowanej przez indonezyjskie kobiety w 1987 roku.
- **wine.data**: „*Wine recognition data*”. Zbiór ten zawiera dane na temat analizy chemicznej winogron używanych do produkcji wina trzech różnych rodzajów.
- **iris.data**: „*Iris Plants Database*”. Zbiór opisujący rodzaje Irysów (3 klasy)

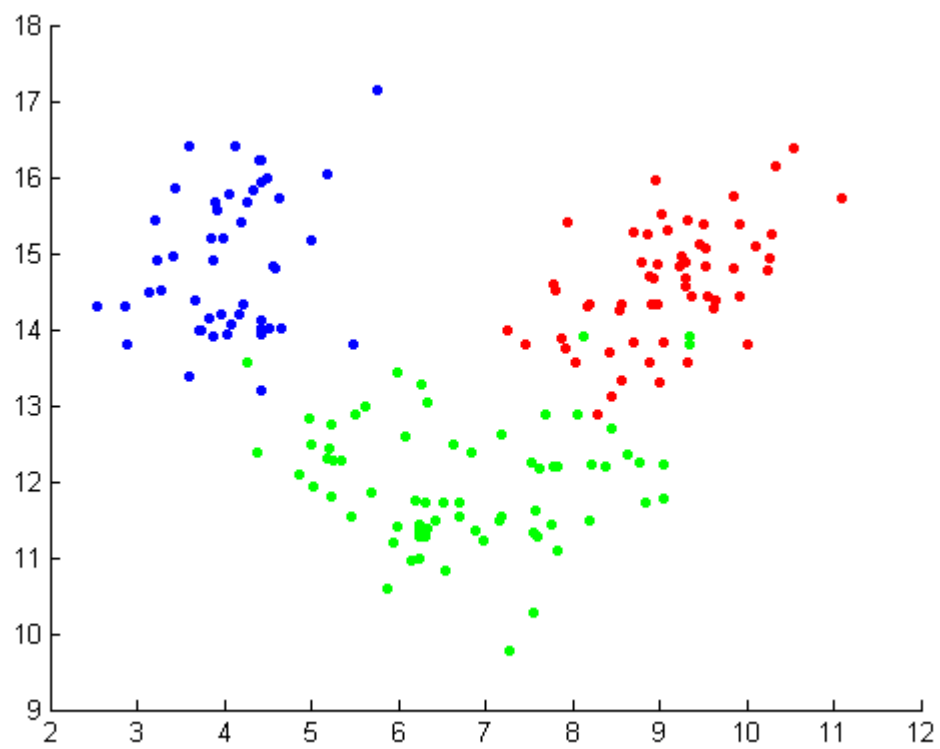
#### Wyniki

W celu wykonania ćwiczenia napisałem 3 skrypty (w 3 wersjach – po jednej dla każdego zbioru danych):

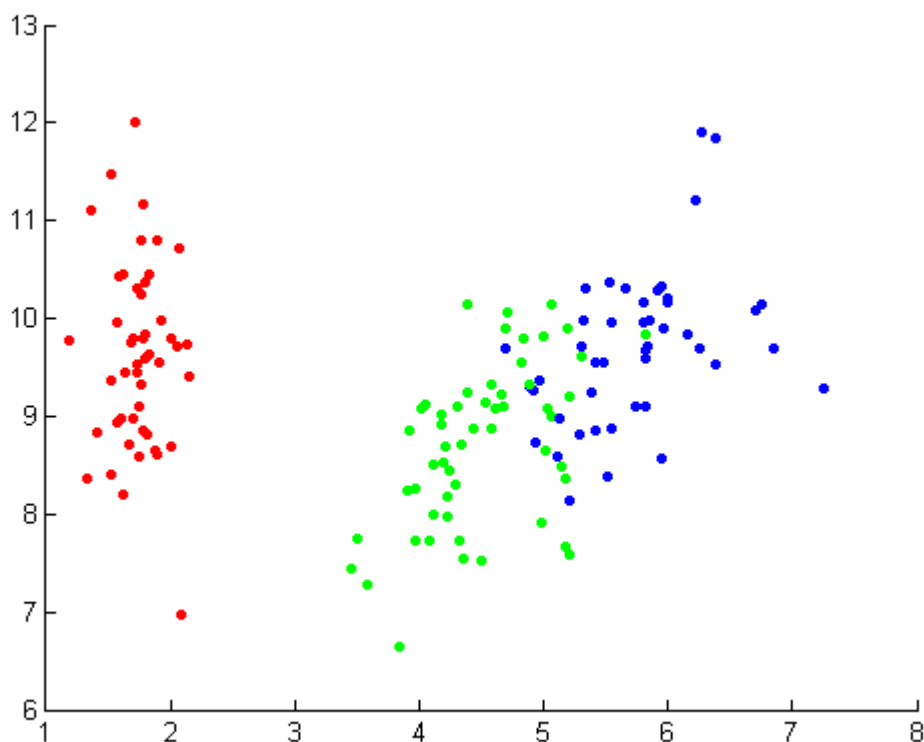
- **lab5\_zad2.m**: przekształcenie przy pomocy PCA do 2-D
- **lab5\_zad3.m**: przeprowadzić trening i symulacje na zbiorach danych z używając sieci neuronowych [5 5 1] oraz [10 10 5 5 1]
- **lab5\_zad5.m**: przeprowadzić trening i symulacje z wykorzystaniem klasyfikatora **bagging**



*lab5\_img1.png - cmc.data*



*lab5\_img2.png - wine.data*



lab5\_img3.png - iris.data

#### Wyniki klasyfikacji (lab5\_results.txt):

	Sieci neuronowa [5 5 1]		Typ sieci neuronowej [10 10 5 5 5 1]		Klasyfikator <i>bagging</i> [5 5 1]	
Zbiór danych	Średnia	Wariancja	Średnia	Wariancja	Średnia	Wariancja
cmc.data	37.4783	4.9527	38.3478	4.5747	54.0870	8.0907
wine.data	95.5056	6.3123	95.2809	2.1462	97.5281	4.6711
iris.data	95.4667	4.9778	95.2000	10.3111	97.0667	0.3556

#### Wnioski:

Przeprowadzone ćwiczenie pokazuje że w przypadku problemu klasyfikacji zwiększenie złożoności sieci neuronowej nie zawsze przynosi pozytywny rezultat lub jest on znikomy w porównaniu do zwiększenia złożoności (zmiana sieci neuronowej spowodowała minimalną zmianę błędy klasyfikacji).

Natomiast użycie klasyfikatora typu *bagging* z prostą siecią neuronową znacząco poprawia ilość poprawnie sklasyfikowanych elementów (17% poprawy dla najtrudniejszego zbioru cmc.data), dla zbiorów o dość dobrej charakterystyce poprawa wynosi 'tylko' 1.5% jednak znacznie zmniejszyła się wariancja (z 4.9 do 0.3) co oznacza że klasyfikacja jest znacznie stabilniejsza.

Tak więc na podstawie zaobserwowanych wyników można sformułować następujące wnioski:

- zwiększenie złożoności sieci neuronowej nie wpływa znacząco na polepszenie klasyfikacji (może ją pogarszać)
- klasyfikator typu *bagging* nie tylko poprawia wynik klasyfikacji ale również czyni ją stabilniejszą