NEURAL NETWORK

DATA DESCRIPTION

STUDENT: BUI TUAN ANH

STUDENT ID: 309051

1. Iris Plan Database

Source:

- Creator: R.A. Fisher
- Donor: Michael Marshall
- Created in July, 1988

This is perhaps the best-known database to be found in the pattern recognition literature. Fisher's paper is a classic in the field and is referenced frequently to this day. The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are not linearly separable from each other. Some important information:

- Number of Instances: 150 (50 in each of three classes)
- Number of Attributes: 4 numeric, predictive attributes and the class
- Attribute information:
  - Sepal length in cm
  - Sepal width in cm
  - Petal length in cm
  - Petal width in cm
  - Class: Iris Setosa, Iris Versicolour, Iris Virginica
- Missing attribute values: None
- Class Distribution: 33.3% for each of 3 classes

2. Wine Recognition Data

Source:

- Forina, M. et al, PARVUS – An Extendible Package for Data Exploration, Classification and Correlation. Institute of Pharmaceutical and Food Analysis and Technologies, Via Brigata Salerno, 16147 Genoa, Italy
- Setefan Aeberhard, email: stefan@coral.cs.jcu.edu.au
- Created in July 1991

These data are results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines. The initial data set had around 30 variables, but for some reason it has the 13 dimensions. Some important informations:

- The attributes are:
  - Alcohol
  - Malic acid
  - Ash
  - Alcalinity of ash
  - Magnesium
  - Total phenols
  - Flavanoids
  - Nonflavanoid phenols
  - Proanthocynanins
  - Colo intensity
  - Hue
  - OD280/OD315 of diluted wines
  - Proline
- Number of instances: class 1: 59; class 2: 71; class 3: 48.
- Number of Attributes: 13
- For each attribute: All attributes are continuous. No statistics available, but suggest to standardize variables for certain uses. 1st attribute is class identifier (1-3)
- Missing attribute values: none