|  |  |
| --- | --- |
| **Iris Data Set** *Download*: [Data Folder](https://archive.ics.uci.edu/ml/machine-learning-databases/iris/), [Data Set Description](https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.names)  **Abstract**: Famous database; from Fisher, 1936 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 150 | **Area:** | Life |
| **Attribute Characteristics:** | Real | **Number of Attributes:** | 4 | **Date Donated** | 1988-07-01 |
| **Associated Tasks:** | Classification | **Missing Values?** | No | **Number of Web Hits:** | 2841227 |

**Source:**

Creator:  
  
R.A. Fisher  
  
Donor:  
  
Michael Marshall (MARSHALL%PLU **'@'** io.arc.nasa.gov)

**Data Set Information:**

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. (See Duda & Hart, for example.) 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.  
  
Predicted attribute: class of iris plant.  
  
This is an exceedingly simple domain.  
  
This data differs from the data presented in Fishers article (identified by Steve Chadwick, spchadwick **'@'** espeedaz.net ). The 35th sample should be: 4.9,3.1,1.5,0.2,"Iris-setosa" where the error is in the fourth feature. The 38th sample: 4.9,3.6,1.4,0.1,"Iris-setosa" where the errors are in the second and third features.

**Attribute Information:**

1. sepal length in cm  
2. sepal width in cm  
3. petal length in cm  
4. petal width in cm  
5. class:  
-- Iris Setosa  
-- Iris Versicolour  
-- Iris Virginica