Student name: amin Kamranfar
Project: Iris Flower

- Link to github:

https://github.com/tyrex1368/amin1368

Iris flower project

- It was introduced by Fisher in 1936

- It is an example of linear discriminant analysis

- Morphologic variation of iris flower

- We have three related species

- Iris is one of the first for multivariate data analysis

Data set

- 50 sample from each of three species of iris

- Iris versicolor, iris virginica, iris setosa

- The length and width of sepals and petals have been measured

- Developing a linear discriminant model to detect the species from each model

Download dataset

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:**	3840496

Types of Iris Flower

Iris Versicolor

Iris Setosa

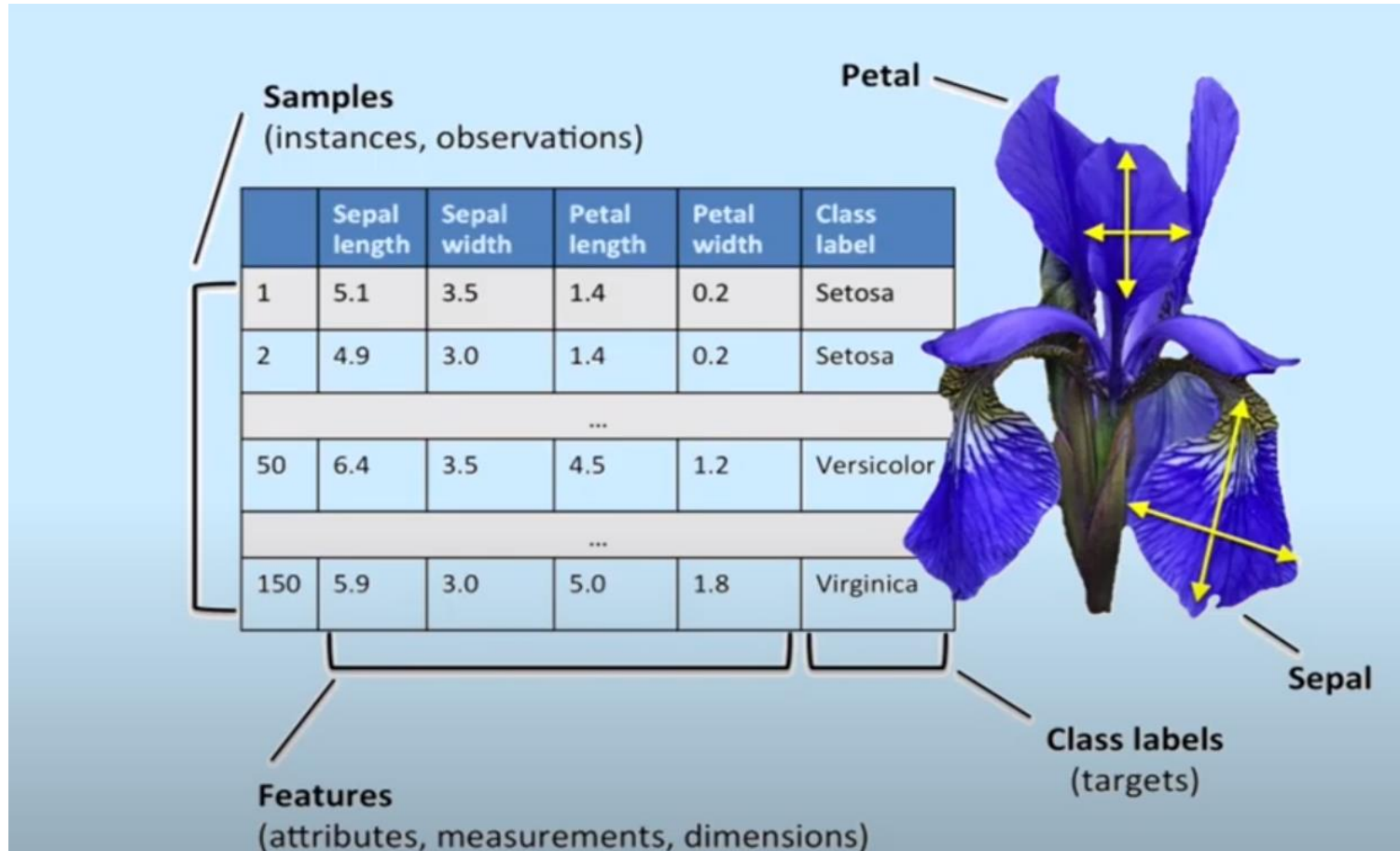**Iris Virginica**

Analyzing data set

- 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

Dataset Structure

Technologies

We can analyze data in four different aspect:

1. Data Visualization
2. Data analytics
3. Data Mining
4. Data storage

Each of these aspect of data have their own software-utility. For the project the dataset is included in R base and Python in the machine learning package Scikit-learn.

Dataset Exploration

- ## 'data.frame': 150 obs. of 5 variables:

- ## $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...

- ## $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...

- ## $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...

- ## $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...

- ## $ Species   : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1

1 1 1 1 1 ...

Goal

- The aim is to classify iris flowers among three species (setosa, versicolor, or virginica) from measurements of sepals and petals' length and width.

- Linear regression cases

- Logistic regression algorithms