Installing Orange
Reading Data in Orange

Installing Orange 3.0
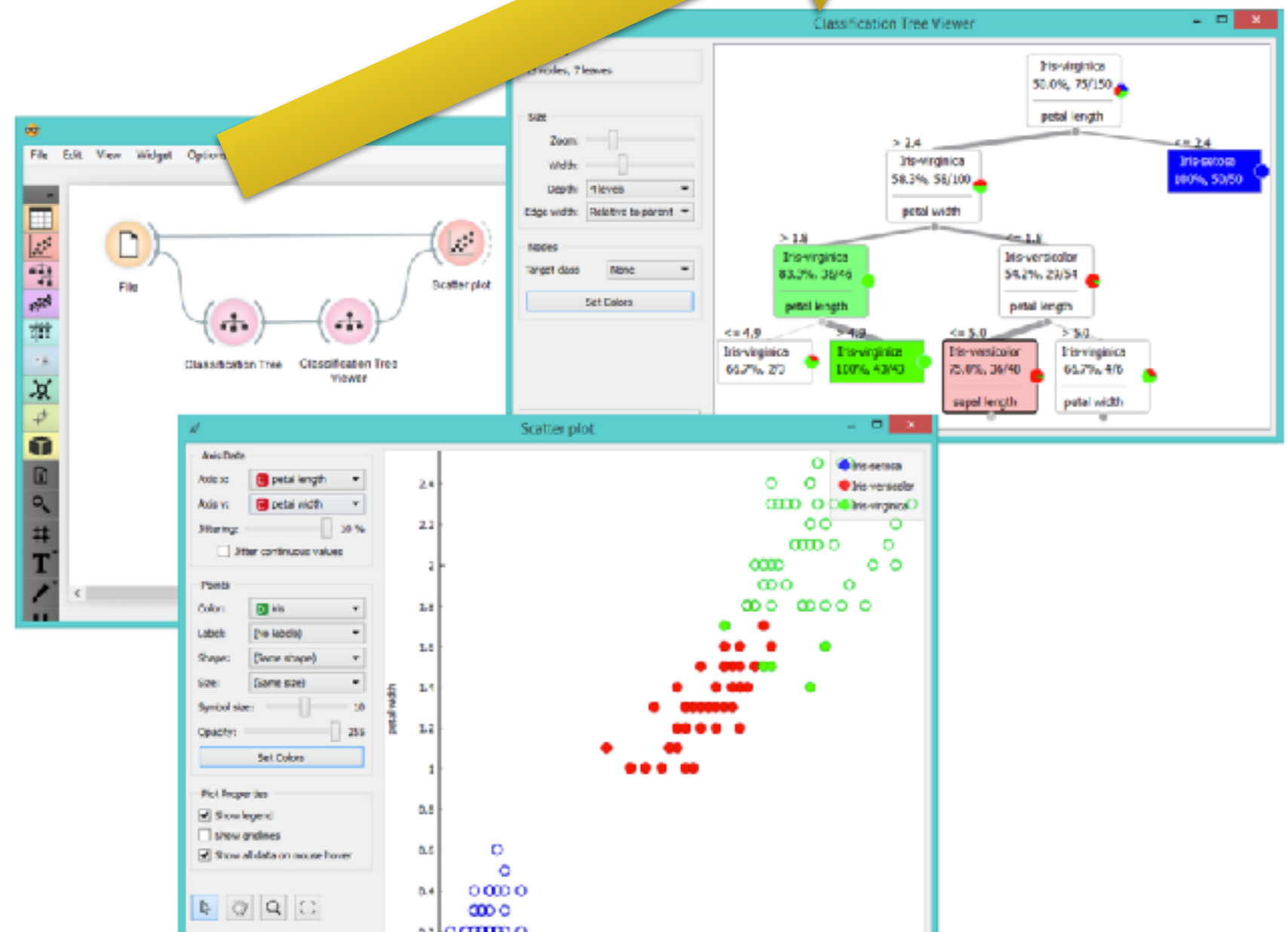
orange.biolab.si

Reading Data in Orange

Data Format

- Files in tab-delimited format ("orange.tab") <— DEFAULT

- Files in csv format ("orange.csv") <— SIMPLIFIED HEADER

- Files in excel format ("orange.xlsx")

Data Format (header)

3 header lines (rows) in tab format

ROW1: feature names
ROW2: feature types
- continuous (c), discrete (d), text (s), basket
ROW3: optional flags
- ignore (i), class (c), multiclass, meta (m)

```
sepal length      sepal width       petal length    petal width     iris
c         c       c         c       d
                                    class
5.1       3.5     1.4       0.2     Iris-setosa
4.9       3.0     1.4       0.2     Iris-setosa
4.7       3.2     1.3       0.2     Iris-setosa
4.6       3.1     1.5       0.2     Iris-setosa
```

iris.tab

Simplified header

- Instead of a three-line header a single-line header can be used with tags listed before feature names and separated from a feature name with a hash ("#") sign. Supported tags are:

  c for class feature
  i for feature to be ignored
  m for the meta attribute
  C for continuous-typed feature
  D for discrete feature
  S for string

```
C#sepal length  iC#sepal width  mC#petal length mC#petal width  cD#iris
5.1      3.5        1.4      0.2        Iris-setosa
4.9      3.0        1.4      0.2        Iris-setosa
4.7      3.2        1.3      0.2        Iris-setosa
```

iris.csv

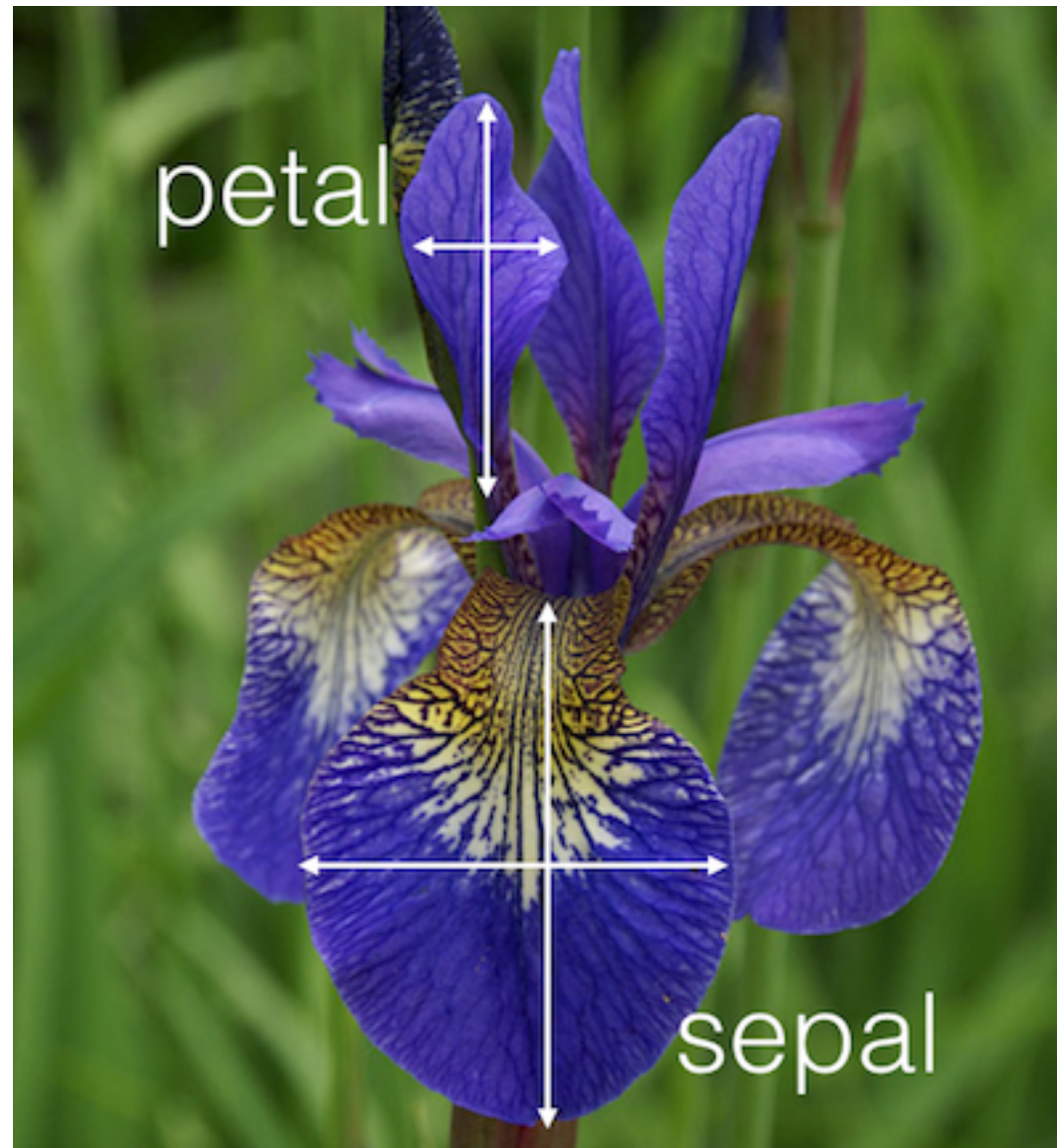# Iris Dataset
classification task (3 classes)

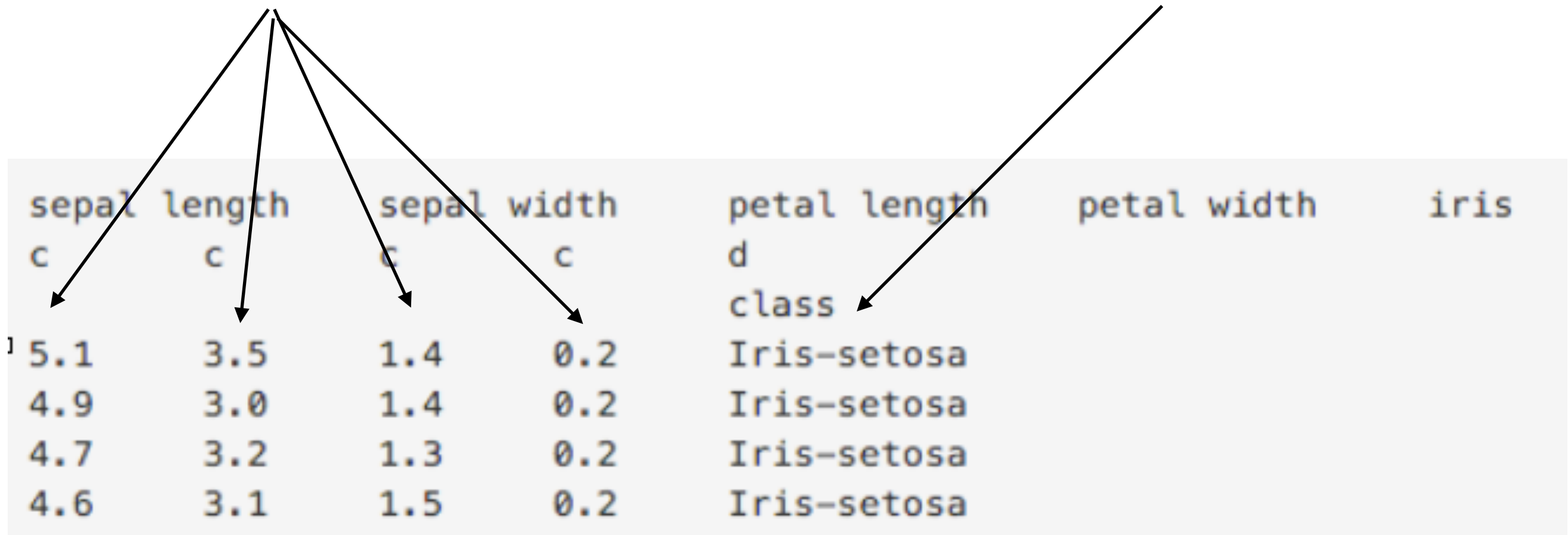Iris
Setosa

Iris
Versicolor

Iris
Virginica

Iris Dataset - 4 features

sepal length	sepal width	petal length	petal width	iris

4 continuous values per row + class label

4 continuous values per row + class label

sepal length	sepal width	petal length	petal width	iris
c	c	c	c	d
				class
5.1	3.5	1.4	0.2	Iris-setosa
4.9	3.0	1.4	0.2	Iris-setosa
4.7	3.2	1.3	0.2	Iris-setosa
4.6	3.1	1.5	0.2	Iris-setosa