



# MACHINE LEARNING CON T-SQL

# Iris Dataset

**Obiettivo:** costruire un algoritmo da addestrare su uno storico di dati già classificati, al fine di costruire una regola matematica per predire con buona accuratezza la classe degli iris futuri a partire dalle variabili *sepal\_length*, *sepal\_width*, *petal\_length*, *petal\_width*

acknowledgements

Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science.

*sepal\_length*; *sepal\_width*; *petal\_length*; *petal\_width*; *class*

4;3.5;1.4;0.2;Iris setosa

4.6;3.1;1.5;0.2;Iris setosa

5.0;3.6;1.4;0.2;Iris setosa

7.0;3.2;4.7;1.4;Iris versicolor

6.4;3.2;4.5;1.5;Iris versicolor

6.9;3.1;4.9;1.5;Iris versicolor

6.3;3.3;6.0;2.5;Iris virginica

5.8;2.7;5.1;1.9;Iris virginica

# Pre-processing

- Aggiungere una colonna per gestire le caratteristiche non riportate nel Dataset
- Aggiungere una colonna per indicare se si tratta di dati di Training o di Test
- Sostituire tutti i NULL con le medie delle colonne di Training
- Standardizzare tutti i valori con le medie e le deviazioni standard delle colonne di Training:

$$x_{\text{new}} = (x - \text{media}(X)) / \text{Dev.St}(X)$$

- Trascodificare in -1 e 1 la colonne target

# Perceptron: training

$X$  = matrice di features di training

$X_i$  =  $i$ -esima riga di  $X$

$y$  = colonna di target di training

$y_i$  =  $i$ -esima target

$n$  = numero di righe di  $X$

$w = (0, 0, 0, 0, 0)$

**FOR**  $k = 1$  : numero epoche

permutiamo le righe di  $(X, y)$

**FOR**  $j = 1$  :  $n$

**IF**  $(X_i * w \geq 0)$  **THEN**  $\text{predizione}_i = 1$  **ELSE**  $\text{predizione}_i = -1$  **END IF**

$w = w + \eta * (y_i - \text{predizione}_i) * X_i$

**END FOR**

**END FOR**

**RETURN**  $w$

# Perceptron: predizione

$X$  = features di test

$y$  = target di test

$w$  = vettore trovato nella fase di training

**IF** ( $x_i * w \geq 0$ )

**THEN** predizione<sub>i</sub> = 1

**ELSE** predizione<sub>i</sub> = -1

**END**