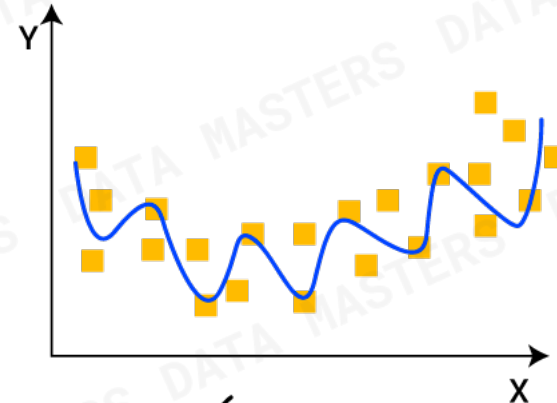
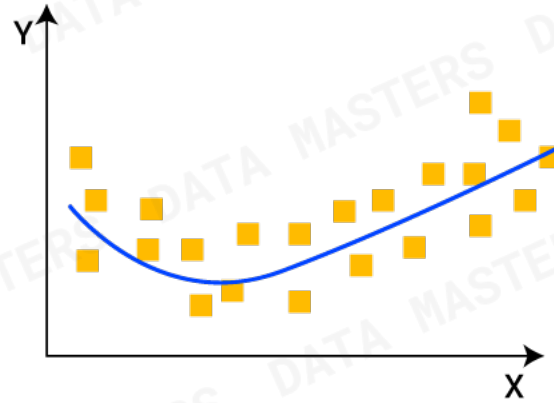
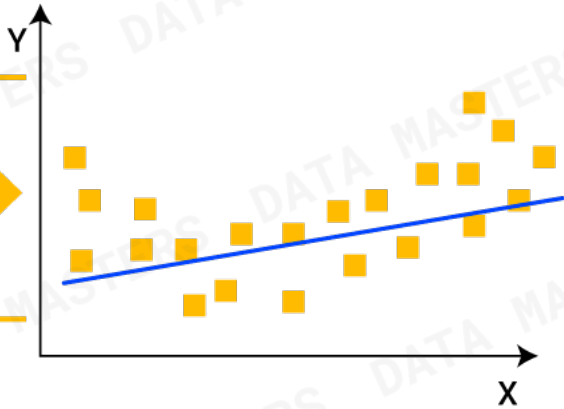
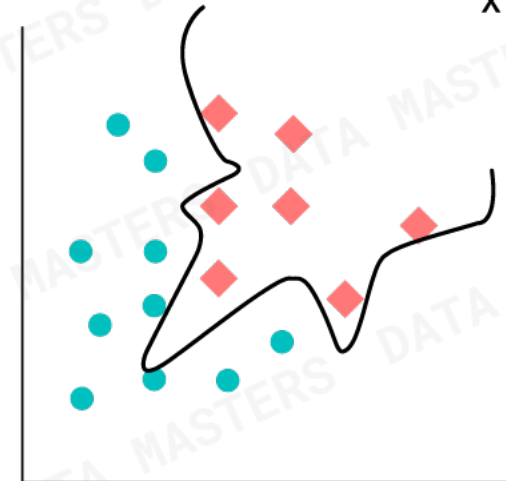
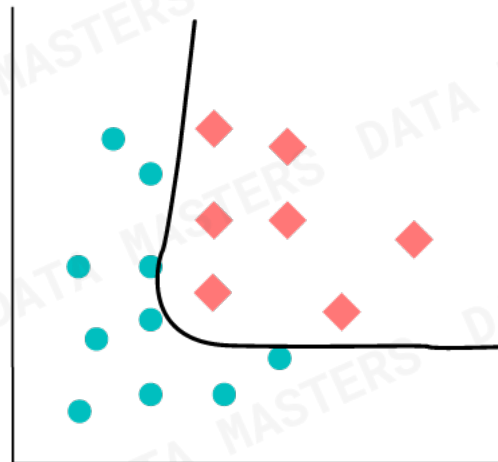
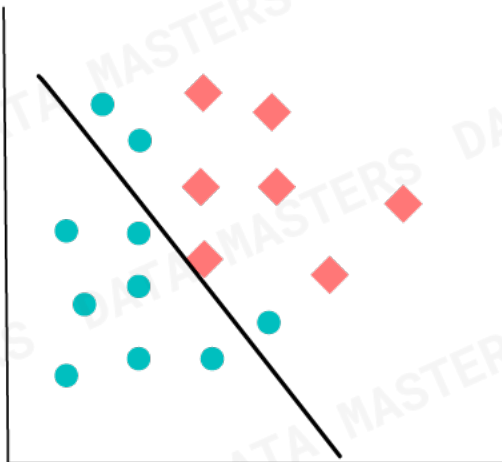


# Differenze di addestramento

REGRESSIONE



CLASSIFICAZIONE



UNDERFITTING

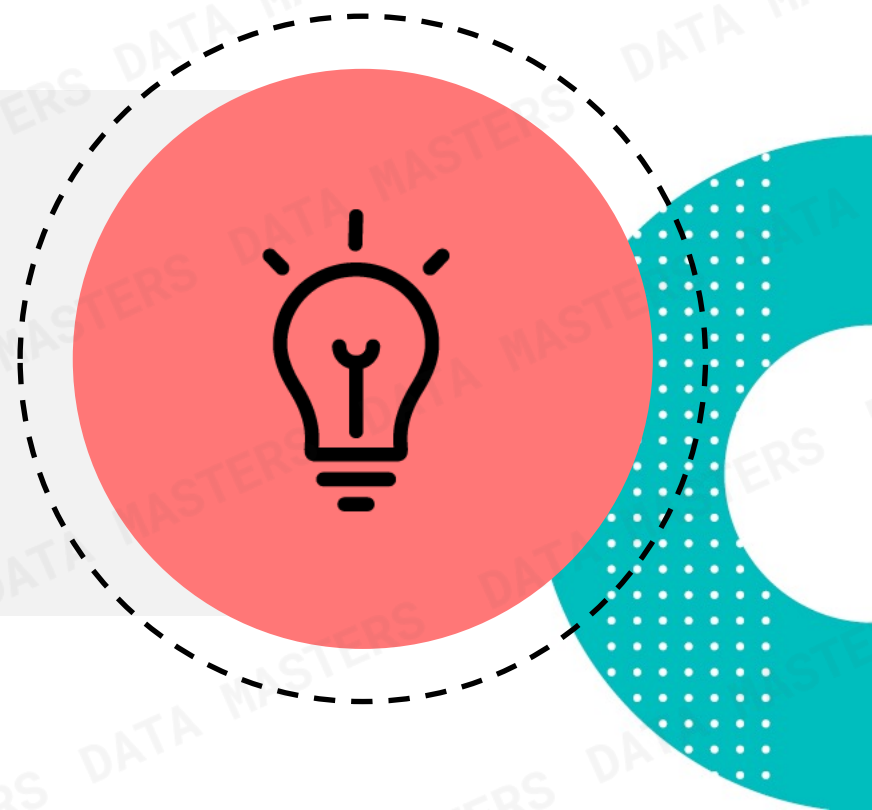
OVERFITTING

# Underfitting

- modello troppo semplice
- pochi dati a disposizione per addestrare il sistema
- poche features per addestrare il modello
- errore alto nel training-set e di valore simile a quello del test-set

## Soluzioni:

- features engineering
- regressioni polinomiali, di grado sempre più alto o comunque modelli più complessi
- numero maggiore di epoche di addestramento

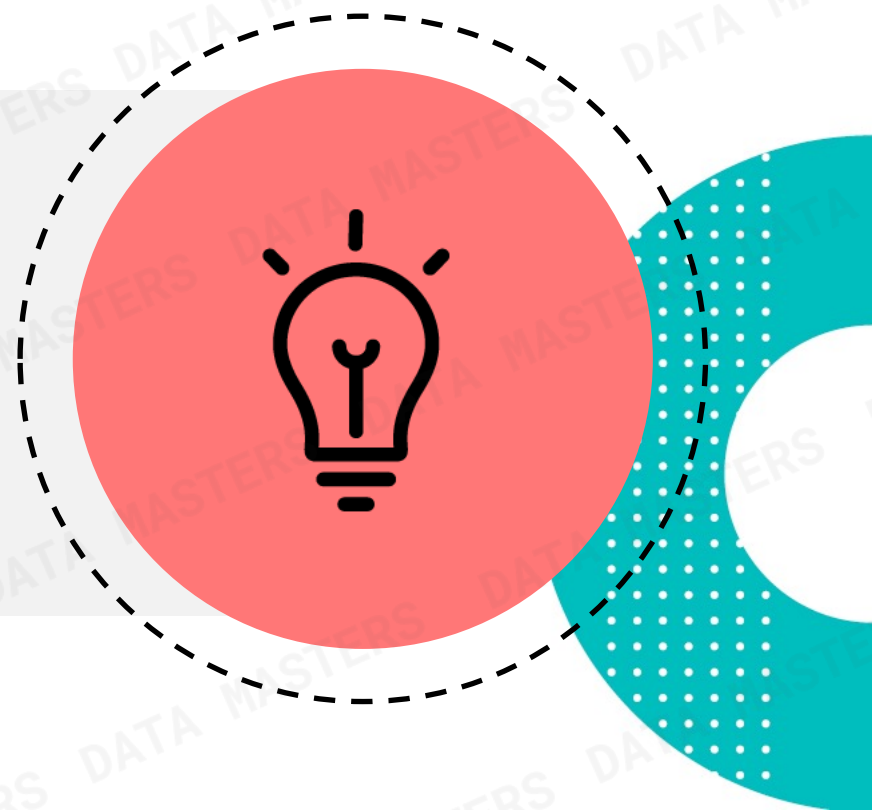


# Overfitting

- il modello non generalizza bene le predizioni relative ad input mai visti in precedenza
- errore nel test-set molto maggiore dell'errore nel training-set

## Cause (e soluzioni):

- modello troppo complesso per il problema analizzato  
(vengono generate troppe curve per seguire i dati di addestramento, tanto che il modello finisce per impararli a memoria)
- troppe features per addestrare il modello



# Regularizzazione

- E' una delle soluzioni al problema di overfitting
- Mantiene tutte le features ma ne riduce i relativi parametri  $\theta_j$

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$

Funzione di costo

$$\frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$

Inizia da 1 e  
non da 0

# Discesa del gradiente con Regularizzazione

Repeat {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_0^{(i)}$$

$$\theta_j := \theta_j - \alpha \left[ \left( \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \right) + \frac{\lambda}{m} \theta_j \right] \quad j \in \{1, 2 \dots n\}$$

}



# Esercitazione

Es03\_regressione\_logistica\_con\_regolarizzazione.ipynb

