

12. Definizione del problema

Corso di Python per il Calcolo Scientifico

Outline

- Determinare l'obiettivo
- Comprendere i dati
- Scegliere il modello

Determinare l'obiettivo

- Il primo step è capire **obiettivo** ed **output atteso**.

Applicazione	Obiettivo del problema	Output atteso
Previsioni meteo	Calcolare le precipitazioni orarie in una determinata zona	Predizione delle precipitazioni orarie
Spam detector	Individuare lo spam	Alert per un possibile spam
Prevenzione bancaria	Identificare transazioni fraudolente	Blocco transazioni sospette
Identificazione degli oggetti	Individuare il tipo di oggetto in una foto	Classe dell'oggetto rappresentato

- In alcuni casi, **il machine learning non è la soluzione migliore!**

Comprendere i dati

- I dati devono avere diverse caratteristiche:
 - **Abbondanza**
 - **Consistenza**
 - **Affidabilità**
 - **Disponibilità**
 - **Correttezza**
 - **Rappresentatività**
- Se i dati non le rispettano, potrebbe essere impossibile avere modelli con prestazioni adeguate!

Comprendere i dati

- Ecco dei possibili esempi di dati 'desiderabili' nelle applicazioni precedenti.

Applicazione	Dati desiderabili (un esempio)
Previsioni meteo	Serie storiche su temperatura, pressione, umidità e precipitazioni per ogni giorno nei cento anni precedenti
Spam detector	Numerosi esempi di mail ordinarie e di spam provenienti da mittenti eterogenei
Prevenzione bancaria	Numerosi esempi di transazioni ordinarie e fraudolente prese in contesti eterogenei e variegati
Identificazione degli oggetti	Numerose foto di oggetti differenti in diverse angolazioni, pose, condizioni di illuminazione ed effetti di occlusione o rumore

Scegliere il modello

- La scelta del modello è strettamente correlata agli step precedenti.

Applicazione	Output atteso	Modello
Previsioni meteo	Predizione delle precipitazioni orarie	Modello di regressione univariata
Spam detector	Alert per un possibile spam	Modello di classificazione binaria
Prevenzione bancaria	Blocco transazioni sospette	Modello di classificazione binaria
Identificazione degli oggetti	Classe dell'oggetto rappresentato	Modello di classificazione multiclasse

Domande?

42