

## TIPI, CLASSIFICAZIONE

17:33

### Prima classificazione di machine learning

Supervised Learning  
 Unsupervised Learning  
 Semi-supervised Learning  
 Reinforcement Learning  
 Recommender Systems

L'apprendimento supervisionato o anche chiamato predittivo.

OBIETTIVO --> è apprendere una mappatura dagli input  $x$  agli output  $y$ , dato un set etichettato di coppie input-output  $D = \{X_i, y_i\}^N$

$$X_i = \text{Vettore } D\text{-dimensionale} \quad \leadsto D = \left\{ x_i \right\}_{i=1}^N$$

$X_i$  rappresenta anche ad esempio il peso di una persona o anche l'altezza

TALI  $x_i$  sono COVARIATES

SI PARLA DI REGRESSIONALE ORDINALE QUANDO AD ESEMPIO  $Y$  CIOE' L'OUTPUT E' NOTO CIOE' DELINEATO TIPO I VOTI DA 1 a 10

APPROCCIO DESCRITTIVO APROCCIO NON SUPERVISIONATO

TALE APPROCCIO DOVE ESSENZIALMENTE NOI ABBIAMO SOLO L'INPUT, DOVE L'OBIETTIVO E' TROVARE I MODELLI INTERESSANTI NEI DATI.

QUINDI ESSENZIALMENTE NON SONO ETICHETTATI, INVECE IN QUELLO SUPERVISIONATO ESSENZIALMENTE ESISTE UNA CORRELAZIONE TRA  $X$  E  $Y$ .

✗ APPREDIMENTO PER RINFORZO || REINFORCEMENT LEARNING

SUCCEDA QUANDO VENGONO DATI SEGNALI OCCASIONALI E COME REAGIRE A QUESTI DATI

IL LIBRO FA RIFERIMENTO A UN BAMBINO CHE CAMMINA.

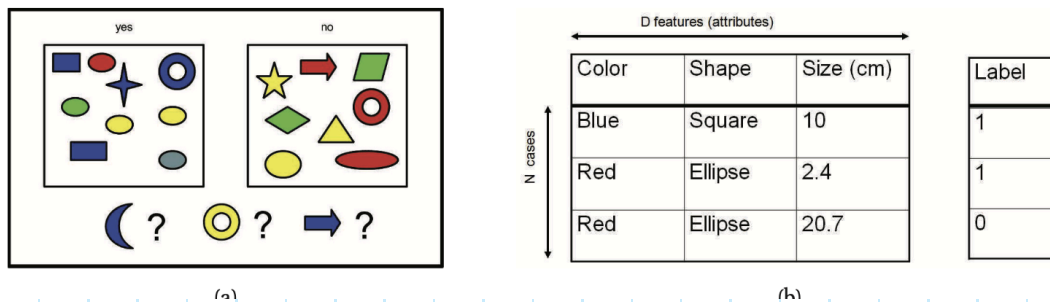
~~CLASSIFICAZIONE~~IPOTESI  $y = f(x)$ 

L'OBIETTIVO E' STIMARE DA DATO UN TRAINING SET ETICHETTATO

QUINDI FARE DELLE PREVISIONI SU

$$\hat{y} = \hat{f}(x)$$

IL NOSTRO OBIETTIVO E' QUELLO DI FARE PREVISIONE SUI NUOVI INPUT

 $D \xrightarrow{f} \rightarrow$  GENERALIZZAZIONE

QUI ABBIAMO UNA DISTIZIONE ANCHE SE NON MOLTO PRECISA IN QUANTO COSA SUCCEDDE ABBIAMO UNA SUDDIVISIONE 0 O 1 LA MEZZALUNA DOVE VA? IL CERCHIO DOVE VA? LA FRECCIA DOVE VA?

LO SI FA ATTRAVERSO LA PROBABILITA'.

COSA SIGNIFICA  $P(y | x, D)$  \*\*\*\*

\*\*\*\* descrive la probabilità che un modello assegni una certa etichetta  $y$  (ad esempio, una classe in un problema di classificazione) a un nuovo esempio con caratteristiche  $X$ , tenendo conto delle informazioni apprese dal dataset  $D$ .

DOVE  $y = 0, 1$ 

DOVE COMUNQUE IN GENERALE

$$P(y | x, D, M)$$

 $M =$  MODELLO

$$\hat{y} = \hat{f}(\mathbf{x}) = \underset{c=1}{\operatorname{argmax}} p(y = c | \mathbf{x}, \mathcal{D})$$

FORMULA MIGLIORE IPOTESI

$\hat{y}$  = PREVISIONE FINALE DEL MODELLO

$\hat{f}(\mathbf{x})$  = FUNZIONE DEL MODELLO ADDESTRATO

$\underset{c=1}{\operatorname{argmax}}$  = TROVARE QUELL'ARGOMENTO CHE MASSIMIZZA LA FUNZIONE

IL MODELLO CERCA QUALE CLASSE DOVE SI HA LA PROBABILITA' CONDIZIONATA PIU' ALTA

↳ T.C  $p(y = c | \mathbf{x}, \mathcal{D})$

IN DEFINITIVA

"LA PROBABILITA' CHE Y APPARTIENE A C TENENDO CONTO DI X DEL TRAINING SET D.

CARNE O PESCE ?

$$y = \text{CARNE} \quad y = \text{PESCE} \quad \text{se } p(y = \text{PESCE} | \mathbf{x}, \mathcal{D}) > p(y = \text{CARNE} | \mathbf{x}, \mathcal{D})$$

ALLORA il modello assegna  $y = \text{PESCE}$

IN DEFINITIVA

L'APPRENDIMENTO SUPERVISO, QUANDO IL dataset contiene esempi etichettati, ovvero ogni input  $X$  è associato a un'uscita corretta (o etichetta)  $Y$ . Per esempio, in un problema di classificazione di immagini, ogni immagine potrebbe essere associata a una categoria come "gatto" o "cane".

OBIETTIVO Il modello impara a mappare correttamente gli input alle uscite, cercando di minimizzare l'errore tra le sue previsioni e le etichette fornite.

## NELL'APPRENDIMENTO NON SUPERVISIONATO INVECE

IL dataset contiene solo input non etichettati, cioè non ci sono etichette  $y$  associate agli input  $X$ . Il modello cerca di trovare strutture nascoste o modelli nei dati.

Obiettivo: Il modello esplora i dati e cerca di rilevare pattern, strutture o gruppi nascosti (ad esempio, raggruppamento o riduzione della dimensionalità).

### CLASSIFICAZIONE 2

## DISCRIMINATIVE MODELS || MODELLI DISCRIMINATIVI

RITORNANDO A APPRENDIMENTO AUTOMATICO -> Dato un dataset di immagini di animali con etichette "gatto" o "cane", il modello impara a distinguere tra le due categorie sulla base degli esempi.

I modelli discriminativi sono una categoria di modelli di machine learning che si concentrano sulla stima della probabilità condizionata  $P(y|x)$ , ovvero la probabilità che un dato input  $X$  appartenga a una determinata classe  $Y$ .

- Esempio: Un modello discriminativo come la regressione logistica prende un insieme di caratteristiche  $\mathbf{x}$  e stima direttamente la probabilità  $p(y|\mathbf{x})$ , cioè "data questa immagine, qual è la probabilità che sia un gatto piuttosto che un cane?"
- Obiettivo: Separare o distinguere tra le classi nei dati.

QUINDI LA DIFFERENZA E' CHE SE AP APPRENDIMENTO AUTOMATICO || SUPERVISED LEARNING E' UN APPROCCIO DI APPRENDIMENTO IL MODELLO DISCRIMINATIVO E' APPUNTO UN MODELLO !!!

## GENERATIVE MODELS

## MODELLI GENERATIVI

I modelli generativi sono una classe di modelli di apprendimento automatico che cercano di modellare l'intera distribuzione dei dati, non solo la relazione tra input e output. In particolare, i modelli generativi stimano la probabilità congiunta  $P(x,y)$  dove:

- $X$  rappresenta l'input (i dati osservati),

- $y$  rappresenta l'etichetta (o la variabile di classe, quando si tratta di classificazione).

Quindi, i modelli generativi cercano di catturare il modo in cui i dati vengono generati. Una volta stimata la distribuzione congiunta  $p(\{x\}, y)$ , possono essere utilizzati per:

- Generare nuovi dati (come immagini o testo),
- Classificare facendo uso della regola di Bayes per calcolare, la probabilità che un input appartenga a una determinata classe.