



Cos'è il Machine Learning?

Il machine learning sta trasformando il modo in cui i computer risolvono problemi e prendono decisioni. Questa sessione ti introdurrà ai concetti fondamentali del machine learning, esplorerà quando applicarlo e ti fornirà le basi per capire come gli algoritmi apprendono dai dati senza programmazione esplicita.

La Visione di Arthur Samuel

Ciò che dà ai Computer la capacità di Imparare

Arthur Samuel, un pioniere nel campo, ha definito l'apprendimento automatico come "**il campo di studio che dà ai computer la capacità di imparare senza essere esplicitamente programmati**". Questa definizione, risalente agli anni '50, rimane straordinariamente rilevante ancora oggi.

L'approccio rivoluzionario di Samuel sfidò la saggezza convenzionale secondo cui i computer potevano fare solo esattamente ciò che veniva loro detto. Invece, dimostrò che le macchine potevano migliorare le loro prestazioni attraverso l'esperienza, proprio come fanno gli esseri umani.



Il leggendario programma di dama



01

La sfida

Arthur Samuel non era un giocatore di dama esperto, eppure voleva creare un programma in grado di padroneggiare il gioco.

02

L'innovazione

Ha programmato il computer per giocare decine di migliaia di partite contro se stesso, imparando da ogni incontro.

03

Il processo di apprendimento

Analizzando quali posizioni sulla scacchiera portavano a vittorie e quali a sconfitte, il programma ha identificato schemi di successo.

04

Il risultato

Grazie a questa esperienza di auto-gioco, il computer è diventato più bravo di Samuel stesso a giocare a dama.

Il genio dell'approccio di Samuel risiedeva nello sfruttare la pazienza e la potenza di calcolo del computer. Mentre un giocatore umano potrebbe giocare centinaia di partite in una vita, il computer poteva giocarne decine di migliaia in poche settimane.

Questo enorme volume di esperienza ha permesso al programma di riconoscere schemi sottili che persino i giocatori esperti potrebbero non notare. Il computer ha imparato a cercare posizioni vantaggiose e ad evitare quelle svantaggiose, affinando progressivamente la sua strategia ad ogni partita giocata.

I Due Tipi Principali di Machine Learning

Apprendimento Supervisionato

Gli algoritmi imparano da esempi etichettati, con risposte corrette fornite durante l'addestramento. Questo è il tipo più utilizzato nelle applicazioni del mondo reale.

- Prevede risultati da dati passati
- Richiede esempi di addestramento etichettati
- Include regressione e classificazione

Apprendimento Non Supervisionato

Gli algoritmi trovano schemi nascosti nei dati senza risposte etichettate, scoprendo la struttura in modo indipendente.

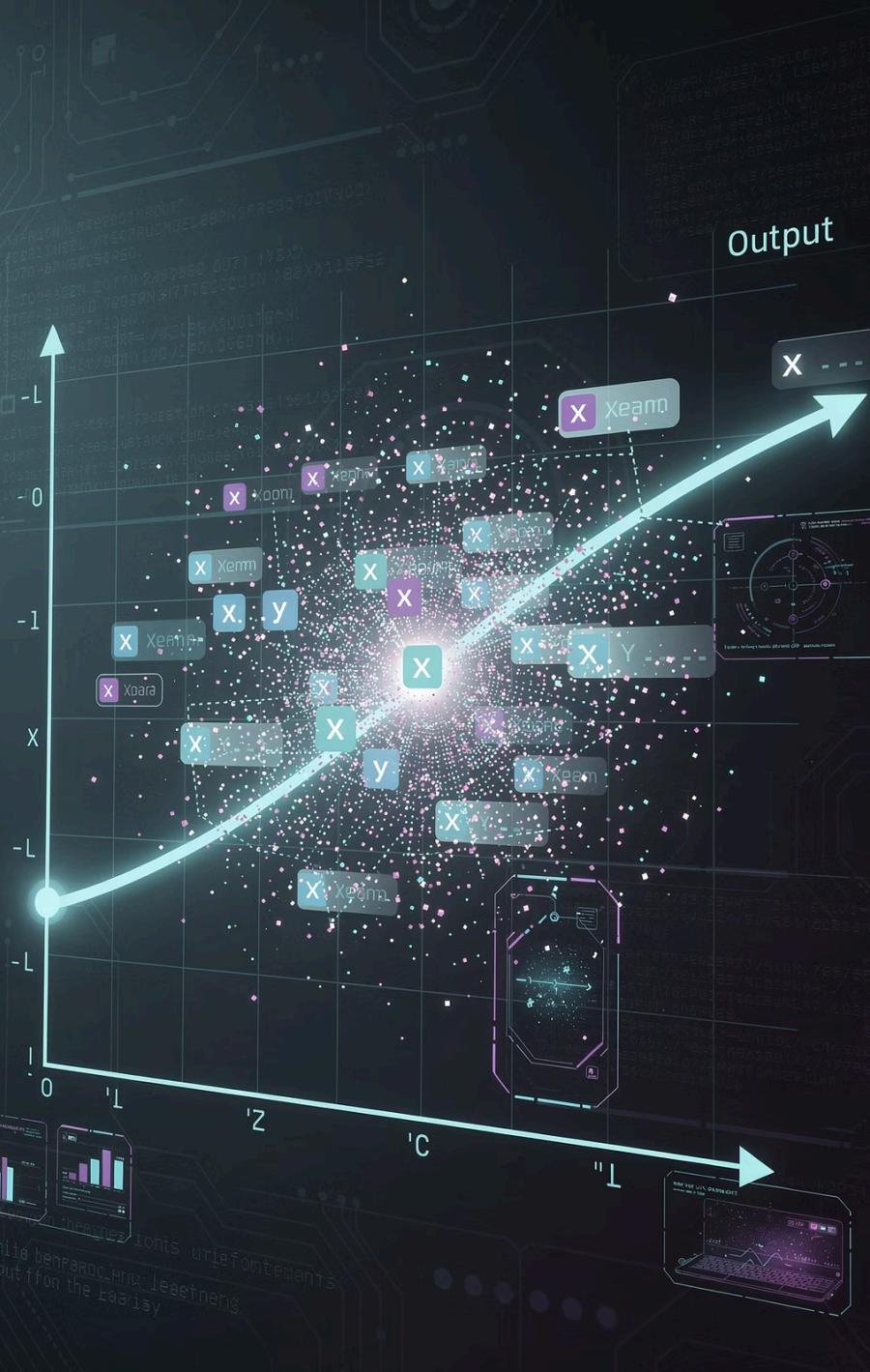
- Identifica schemi autonomamente
- Funziona con dati non etichettati
- Scopre relazioni nascoste

Apprendimento Supervisionato

Il machine learning supervisionato, o più comunemente apprendimento supervisionato, si riferisce ad algoritmi che imparano mappature da X a Y, o da input a output. Essenzialmente, questi algoritmi sono addestrati a prevedere un output (Y) basato su un dato input (X).

- **La Caratteristica Chiave:** Fornisci al tuo algoritmo di apprendimento esempi che includono le risposte corrette. Queste coppie corrette di input (X) e etichetta di output desiderata (Y) consentono all'algoritmo di apprendere schemi, in modo che possa eventualmente fare previsioni accurate su nuovi input non visti.

Pensalo come uno studente che impara con un insegnante: l'insegnante fornisce esempi con risposte corrette e lo studente impara a derivare le risposte da solo. Questo approccio strutturato rende l'apprendimento supervisionato altamente efficace per molte attività di previsione nel mondo reale.



Applicazioni nel Mondo Reale dell'Apprendimento Supervisionato

L'apprendimento supervisionato è il fulcro di molte tecnologie che utilizziamo quotidianamente, consentendo ai sistemi di effettuare previsioni e classificazioni intelligenti basate su dati passati.



Rilevamento dello Spam

Identifica le email indesiderate imparando dai messaggi precedentemente contrassegnati, migliorando continuamente la sua precisione con il feedback degli utenti.



Riconoscimento di Immagini

Classifica oggetti, volti e scene nelle immagini, alimentando applicazioni dall'organizzazione di foto ai veicoli autonomi.



Diagnosi Medica

Assiste i professionisti sanitari nella diagnosi di malattie analizzando i dati dei pazienti e le scansioni mediche per individuare schemi specifici.



Rilevamento delle Frodi

Monitora le transazioni finanziarie per rilevare attività insolite, segnalando acquisti potenzialmente fraudolenti per proteggere gli account.

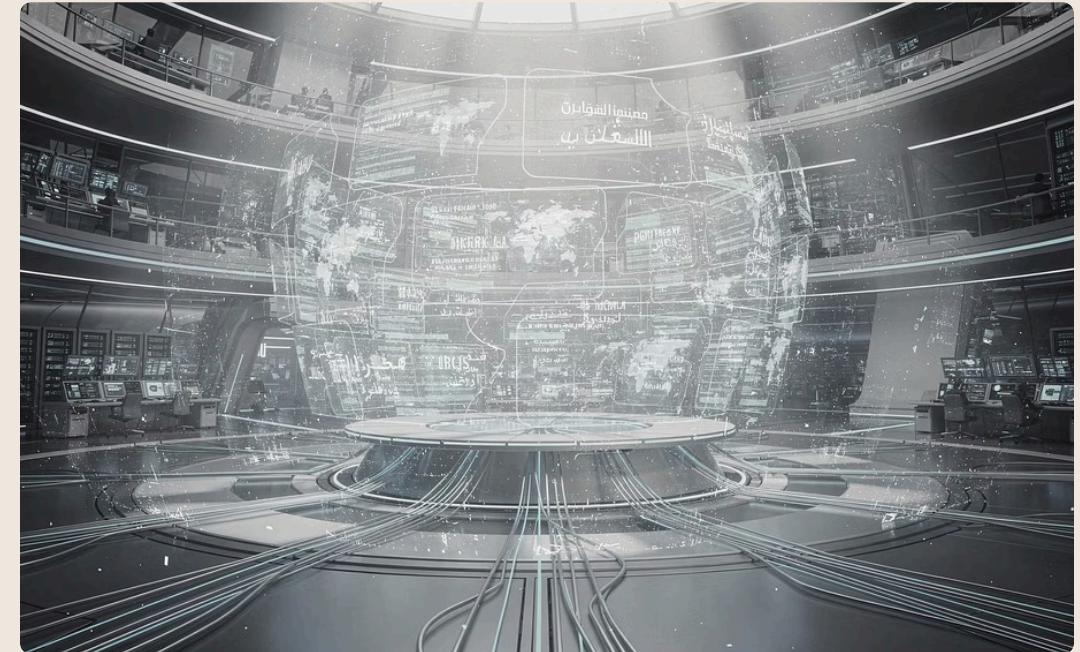
Altre applicazioni reali dell'apprendimento supervisionato



Riconoscimento vocale

Input: Clip audio

Output: Trascrizione testuale



Traduzione automatica

Input: Testo inglese

Output: Spagnolo, Arabo, Hindi, Cinese o altre lingue

L'Economia del Machine Learning



La pubblicità online si distingue come una delle applicazioni più lucrative dell'apprendimento supervisionato oggi. Le grandi piattaforme a livello globale utilizzano algoritmi di apprendimento sofisticati per analizzare vaste quantità di dati sia sugli annunci che sugli utenti.

Questi algoritmi prevedono la probabilità che un utente clicchi su uno specifico annuncio, ottimizzando efficacemente il posizionamento e il targeting degli annunci per massimizzare l'engagement.

- **Perché è importante:** Ogni singolo clic su un annuncio si traduce direttamente in entrate. Abbinando con precisione gli utenti agli annunci con cui è più probabile che interagiscano, queste piattaforme generano un valore economico sostanziale, consolidando la pubblicità online come un'applicazione fondamentale della redditività dell'apprendimento supervisionato.

Altre Applicazioni Che Trasformano le Industrie



Auto a Guida Autonoma

Input: Immagini e dati da sensori (radar, lidar)

Output: Posizioni di altre auto e ostacoli

Scopo: Navigare in sicurezza attorno ad altri veicoli



Ispezione Visiva

Input: Foto del prodotto fabbricato

Output: Rilevamento difetti (graffi, ammaccature)

Scopo: Ridurre i difetti e migliorare il controllo qualità



Come funziona l'apprendimento supervisionato



Fase di Addestramento

Per prima cosa, addestra il tuo modello con esempi di input X e le risposte corrette (etichette Y)

Apprendimento dei Modelli

L'algoritmo identifica modelli e relazioni tra input e output

Effettuare Predizioni

Il modello prende input completamente nuovi che non ha mai visto e produce output appropriati

Algoritmi di apprendimento supervisionato

Gli algoritmi di apprendimento supervisionato sono incredibilmente versatili e trovano applicazioni in una vasta gamma di settori. Permettono di mappare input specifici a output desiderati, fornendo soluzioni automatizzate per problemi complessi.

Input (X)	Output (Y)	Applicazione
e-mail	spam? (0/1)	filtro antispam
audio	trascrizioni di testo	riconoscimento vocale
testo in inglese	testo in spagnolo	traduzione automatica
annuncio, info utente	clic? (0/1)	pubblicità online
immagine, info radar	posizione di altre auto	auto a guida autonoma
immagine del telefono	difetto? (0/1)	ispezione visiva

Dal miglioramento della comunicazione al potenziamento della qualità di produzione e all'abilitazione dei sistemi autonomi, l'apprendimento supervisionato continua a guidare l'innovazione e l'efficienza in innumerevoli settori.



Caso di studio

Previsione dei prezzi delle case

Previsione del Prezzo delle Case: I Dati

Immagina un set di dati in cui ogni voce include le dimensioni di una casa (in metri quadrati) e il suo corrispondente prezzo di vendita (in migliaia di sterline). Se visualizzato, la dimensione della casa forma tipicamente l'asse orizzontale, con il prezzo sull'asse verticale, creando un grafico a dispersione di punti dati.

- ❑ **La Sfida:** Un amico cerca una stima affidabile del prezzo per la sua casa di 150 metri quadrati. Come può un algoritmo di machine learning sfruttare questi dati storici per fornire una previsione accurata?

Due approcci per adattare i dati

Quando si modellano le relazioni nei dati, possono essere utilizzati diversi approcci matematici per fare previsioni, ognuno con diversi livelli di complessità e assunzioni sui modelli sottostanti.

Modello Lineare

Un approccio comune è adattare una linea retta ai dati.

Questo modello semplice assume una relazione costante tra la dimensione della casa e il prezzo.

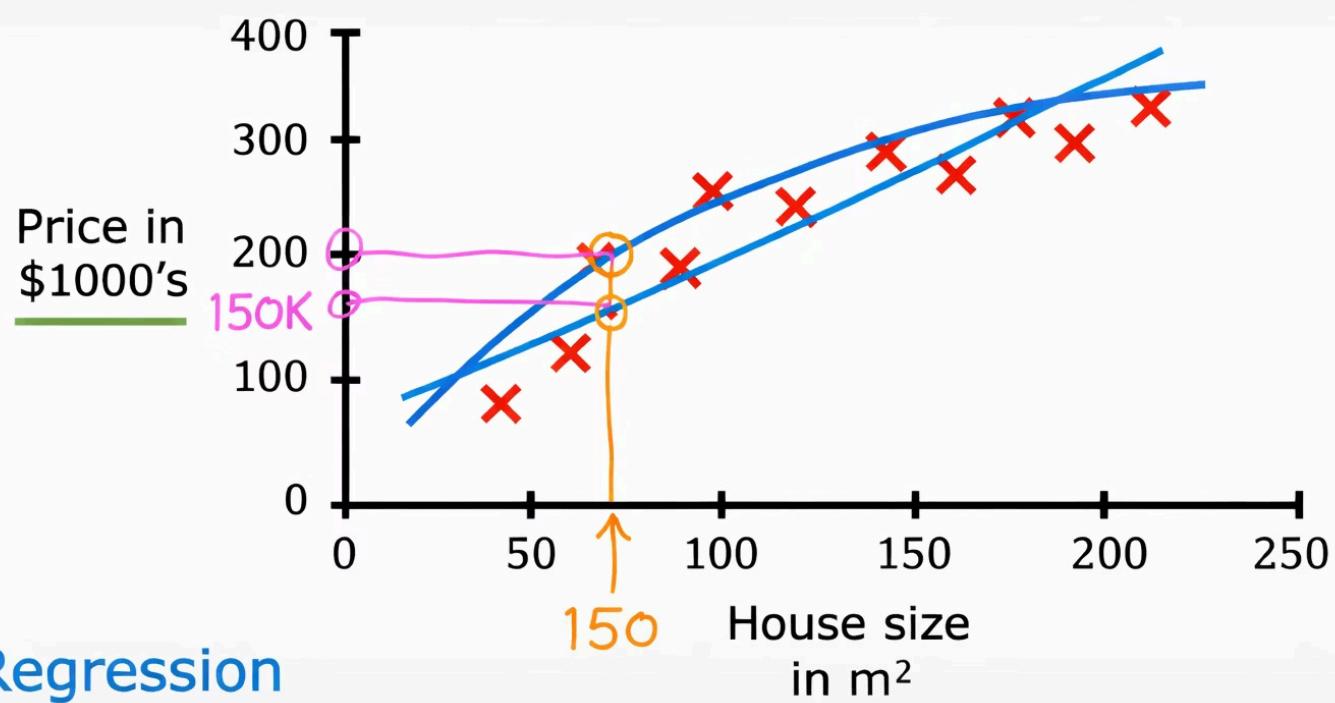
Leggendo da questa linea, la casa di 150 metri quadrati del tuo amico potrebbe essere venduta per circa £150.000.

Modello Curvo

In alternativa, potresti adattare una curva—una funzione leggermente più complessa. Questo approccio può catturare relazioni più sfumate e non lineari nei dati.

Utilizzando un modello curvo, la previsione per la stessa casa di 150 metri quadrati potrebbe essere più vicina a £200.000.

Regression: Housing price prediction



Regression

Predict a **number**

infinitely many possible outputs

Introduzione alla Regressione

Questa previsione del prezzo delle case è un tipo particolare di apprendimento supervisionato chiamato **regressione**.

Cos'è la Regressione?

La regressione significa prevedere un numero tra infiniti numeri possibili. Nel nostro esempio, i prezzi delle case potrebbero essere €150.000, €183.000 o qualsiasi altro numero intermedio, non solo poche opzioni discrete.

A differenza della classificazione (che esploreremo in seguito), la regressione produce previsioni numeriche continue anziché categorie distinte. Questo la rende ideale per prevedere quantità come prezzi, temperature o distanze, dove l'output può assumere qualsiasi valore all'interno di un intervallo.

Pausa Capitolo

Problemi di Classificazione

Previsione di categorie, non di numeri

Introduzione alla Classificazione

La classificazione è il secondo tipo principale di algoritmo di apprendimento supervisionato. A differenza della regressione, che predice valori numerici, la classificazione predice categorie discrete.

Esploriamo questo concetto attraverso un'applicazione medica cruciale: la rilevazione del cancro al seno. La diagnosi precoce può migliorare significativamente i risultati per il paziente, rendendo i sistemi di classificazione accurati strumenti diagnostici inestimabili.



L'Obiettivo: Costruire un modello in grado di distinguere accuratamente tra cellule benigne (non cancerose) e maligne (cancerose) basandosi su varie caratteristiche derivanti dall'imaging medico.

Esempio di Rilevamento del Cancro al Seno

Utilizzando i dati medici di un paziente, il tuo sistema di machine learning tenta di determinare se un tumore (un nodulo) è:

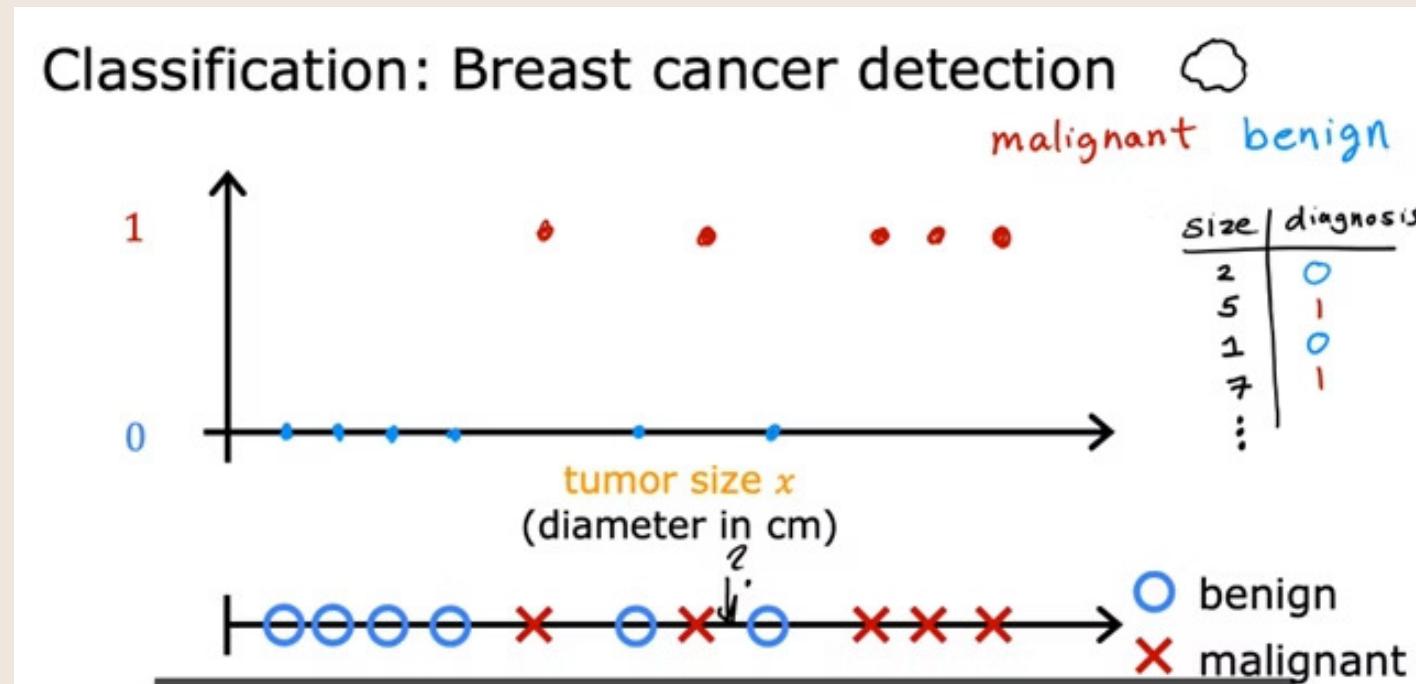
Benigno (0)

Non canceroso; un nodulo non pericoloso

Maligno (1)

Canceroso e potenzialmente pericoloso

Il dataset traccia tumori di varie dimensioni, etichettati come benigni (0) o maligni (1). Questo è fondamentalmente diverso dalla regressione perché stiamo prevedendo solo due possibili risultati, non un intervallo infinito di numeri.

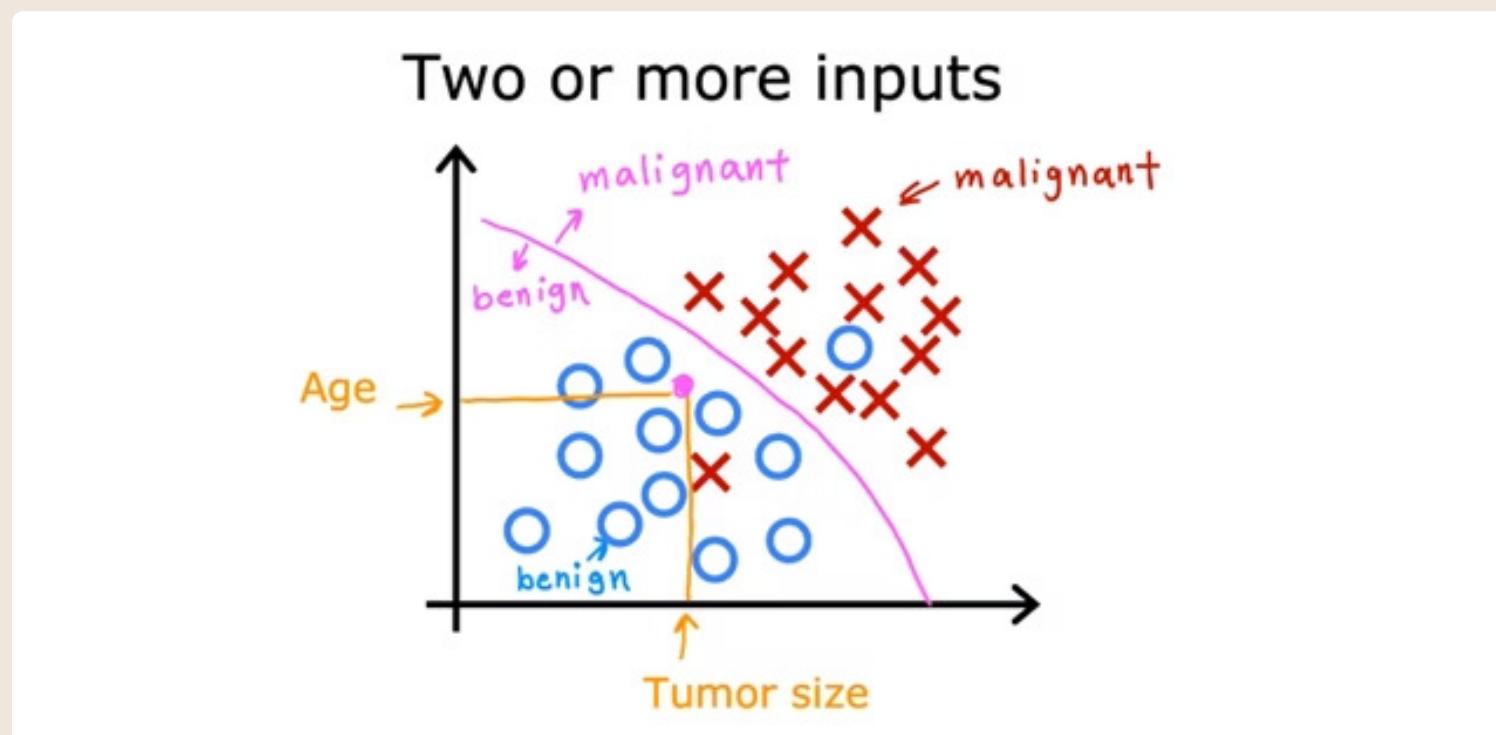


Oltre il Binario: Categorie Multiple

I problemi di classificazione possono estendersi oltre due semplici scelte. Il tuo algoritmo di apprendimento potrebbe dover distinguere tra diversi esiti, portando a uno scenario di classificazione multi-classe.

1	2	3
Benigno Nessun cancro rilevato	Cancro di Tipo 1 Primo tipo di cancro identificato	Cancro di Tipo 2 Secondo tipo di cancro identificato

- **Terminologia:** Nella classificazione, i termini "classi di output" e "categorie di output" sono usati in modo intercambiabile—significano la stessa cosa e si riferiscono agli esiti discreti che il modello predice.



Utilizzo di Funzionalità di Input Multiple

Invece di affidarci solo alla dimensione del tumore, possiamo utilizzare input multipli per migliorare le previsioni. In questo set di dati migliorato, abbiamo due input: età (in anni) e dimensione del tumore (in piedi quadrati).

I cerchi rappresentano pazienti con tumori benigni, mentre le croci indicano casi maligni. Quando un nuovo paziente arriva, il medico misura la dimensione del tumore e registra l'età del paziente.

Il compito dell'algoritmo: Trovare un confine che separi i tumori maligni da quelli benigni, aiutando i medici a fare diagnosi più informate.



Complessità del Mondo Reale: Molteplici Funzionalità di Input

I ricercatori che lavorano sulla rilevazione del cancro al seno utilizzano molti input aggiuntivi oltre all'età e alla dimensione del tumore:

Spessore del Grumo Tumorale

Misura la densità e la profondità della massa tumorale

Uniformità della Dimensione Cellulare

Valuta la coerenza nelle dimensioni delle cellule

Uniformità della Forma Cellulare

Valuta la regolarità della struttura cellulare

Caratteristiche Cliniche Aggiuntive

Inclusa l'adesione cellulare, le caratteristiche nucleari e altro ancora

Incorporando molteplici funzionalità, i sistemi di apprendimento automatico possono effettuare previsioni più accurate e sfumate.

Riepilogo: Punti chiave



Apprendimento Supervisionato

Mappa l'input X all'output Y imparando dalle "risposte corrette" nei dati di addestramento etichettati



Regressione

Prevede numeri da infiniti output possibili (es. prezzi delle case)



Classificazione

Prevede categorie da un piccolo, finito insieme di output possibili (es. benigno o maligno)

Ora comprendi l'apprendimento supervisionato, inclusi i problemi di regressione e classificazione. Nella prossima lezione, esploreremo l'apprendimento non supervisionato, il secondo tipo principale di machine learning, per completare le tue conoscenze fondamentali.