

Il corso di Tuidi sul Machine Learning

# **INTRODUZIONE ALL'IA NELLA FILIERA ALIMENTARE**



**TUIDI**

# Questo corso ha l'obiettivo di introdurre il concetto di **Intelligenza Artificiale (IA)**, rispondendo a tre domande fondamentali:

- Cosa è l'IA?
- Da quali tecnologie è governata?
- Cosa può fare l'IA per noi?

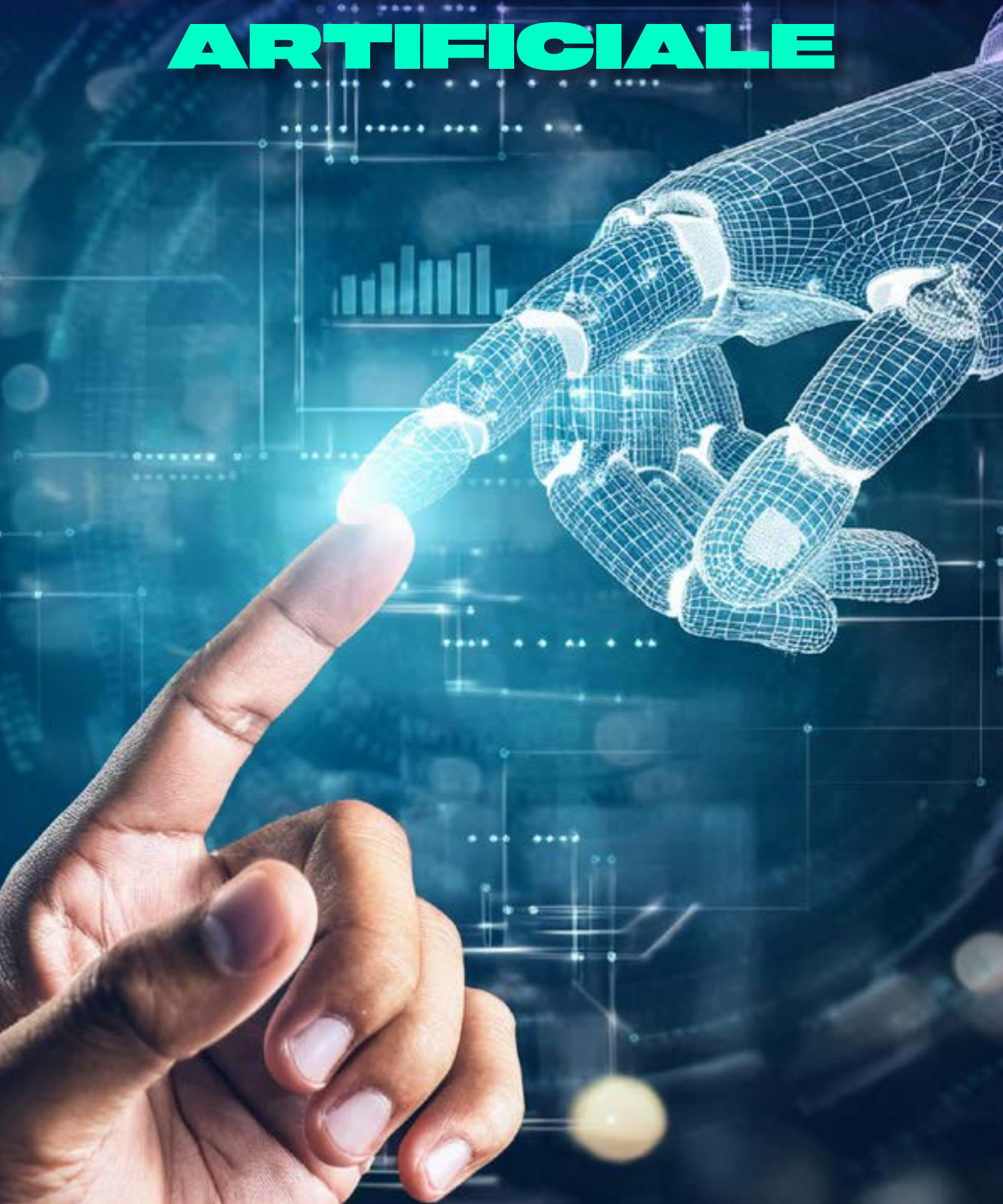
Lo scopo è offrire tutte le nozioni e gli strumenti necessari che aiutino nella comprensione di cosa si intenda per intelligenza artificiale, quale sia il processo di elaborazione dei dati per sviluppare dei modelli predittivi - attraverso il machine learning - e gli esempi applicativi contestualizzati nel settore del retail e della produzione.

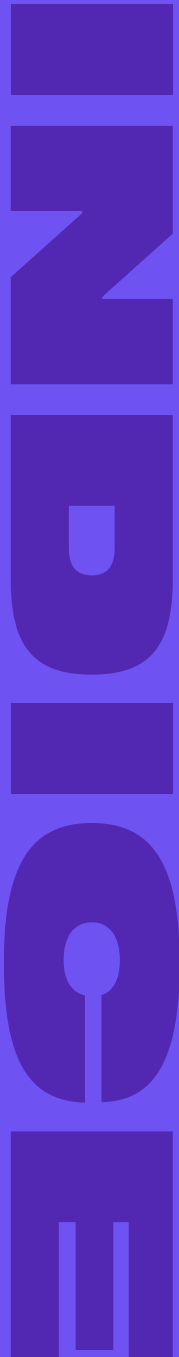
In particolare, il corso è composto da quattro lezioni:

- **Lezione 1** - L'intelligenza Artificiale: Cosa è questa disciplina, da cosa è composta, le sue applicazioni nel mondo del retail e le principali differenze con i modelli statistici tradizionali utilizzati fino ad ora.
- **Lezione 2** - Data Mining: In questa lezione verrà descritto il processo di estrazione dati, sottolineando l'importanza e la necessità di avere informazioni accurate al fine di raggiungere gli obiettivi preimpostati.
- **Lezione 3** - Machine Learning: L'apprendimento automatico e le tecniche di Classificazione, Regressione e Clustering.
- **Lezione 4** - Differenza tra analisi predittiva e prescrittiva: verrà spiegata la differenza tra questi due modelli, andando nel dettaglio di questi approcci e spiegandone i vantaggi e gli svantaggi dell'una o l'altra.

Lezione 1

# L'INTELLIGENZA ARTIFICIALE





- 1.1 Introduzione alla lezione
- 1.2 Che cosa è l'Intelligenza Artificiale
- 1.3 Deep Learning
- 1.4 Una macchina dotata di Intelligenza Artificiale è cosciente?
- 1.5 L'utilizzo dell'IA nella filiera alimentare
- 1.6 Perché investire nell'Intelligenza Artificiale?
- 1.7 Come identifico un'applicazione di Intelligenza Artificiale?
- 1.8 Differenza con i classici modelli statistici

## 1.1 Introduzione alla lezione

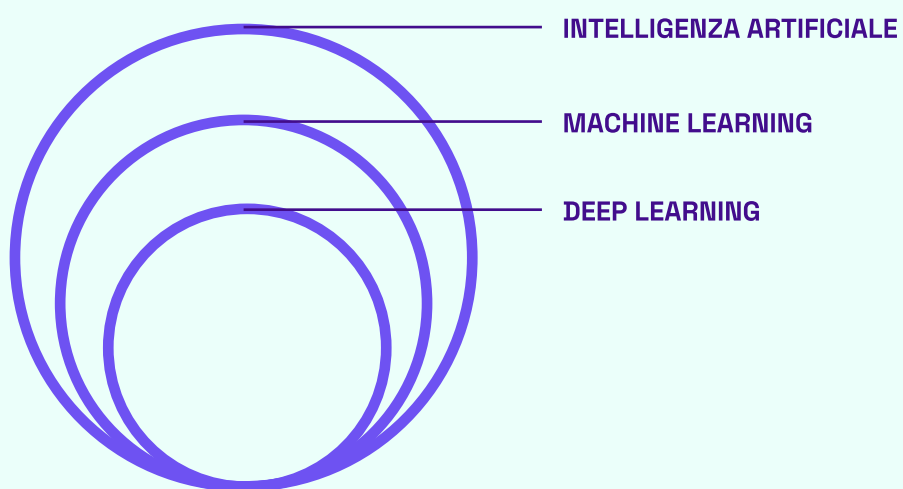
Durante la prima lezione si prenderà dimestichezza con l'intelligenza Artificiale definendo il concetto in senso lato ed entrando nello specifico attraverso alcuni esempi. Quello dell'IA è un tema molto vasto ed in continua evoluzione, caratterizzato da una moltitudine di possibili applicazioni. Per alcuni essa riguarda forme di vita artificiale in grado di superare l'intelligenza umana, altri invece qualificano l'IA come un insieme di tecnologie per l'elaborazione di dati.

Con questa lezione spiegheremo che cosa è l'Intelligenza Artificiale, quali sono i suoi utilizzi, con un focus particolare sul settore alimentare, e quali sono i vantaggi rispetto ai sistemi statistici tradizionali.

## 1.2 Che cosa è l'Intelligenza Artificiale

Si definisce Intelligenza Artificiale l'abilità di una macchina di mostrare capacità umane quali l'apprendimento, la pianificazione e la creatività. Ciò che rende "intelligente" questa tecnologia è proprio la capacità di riuscire ad apprendere dai dati che le vengono forniti manualmente ed utilizzare le informazioni acquisite per il raggiungimento di obiettivi preimpostati.

Spesso il concetto di IA viene confuso e sostituito con i termini "Machine Learning" (ML) e "Deep Learning" (DL) quasi come se fossero parole intercambiabili. Al fine di chiarire la differenza tra questi termini può essere utile rappresentarli come cerchi concentrici.



Infatti, ML e DL non sono sinonimi ma due sottocategorie dell'Intelligenza Artificiale. Nel caso del Machine Learning, o apprendimento automatico, è un sottoinsieme dell'IA che si basa sull'apprendimento dei dati, senza necessitare di regole definite manualmente. Tutti i sistemi di Machine Learning sono Intelligenze Artificiali, ma non tutte le Intelligenze Artificiali si basano sul Machine Learning. Il processo di elaborazione delle decisioni, quindi, può essere diviso

essenzialmente in tre fasi:

1. La macchina analizza e apprende da tutti i dati che le vengono forniti;
2. La macchina prende decisioni in maniera autonoma che si basano su tutti i dati che è riuscita ad analizzare;
3. Infine, la macchina riesce ad ottimizzare le decisioni prese in precedenza apprendendo automaticamente dai propri errori.

Il ML per operare in maniera efficace e per adattarsi alle varie situazioni che si verificano ha bisogno di grandi moli di dati. Un'applicazione di ML all'interno dell'e-commerce è il "Dynamic Pricing" ovvero la pratica di fissare un prezzo per un prodotto in base alle attuali condizioni di mercato.

Il ML è in grado di analizzare i dati in tempo reale e suggerire prezzi differenti a seconda dei giorni della settimana o delle fasce orarie, con il fine di massimizzare il profitto. Per esempio, il sistema può prevedere che alzare il prezzo di vendita delle bottiglie di vino di venerdì sera e abbassarlo di martedì a mezzogiorno può portare ad una massimizzazione del profitto (in questo caso, la nostra funzione obiettivo) derivante dalla vendita di tale alcolico. Come è possibile? Il Machine Learning studia i dati della clientela, come le caratteristiche demografiche, le scelte di acquisto, l'affluenza sul sito, proponendo un prezzo che consideri tutte le variabili che impattano la massimizzazione della funzione descritta.

## 1.3 Deep Learning

Quando si parla di Deep Learning (DL) invece si fa riferimento all'apprendimento autonomo (N.B. non più automatico). Oltre ad interpretare le variabili che l'uomo introduce all'interno dell'input del sistema (come nel Machine Learning), questi algoritmi riescono ad identificare autonomamente caratteristiche distintive senza categorizzazioni dall'esterno, ovvero senza alcun intervento umano. Il DL per questo motivo ha bisogno di molti più dati rispetto al ML e richiede maggiori risorse informatiche.

Un esempio di Deep Learning è il "Natural Language Processing" (NLP) ovvero "elaborazione del linguaggio naturale", si tratta di algoritmi in grado di analizzare, rappresentare e quindi comprendere testi intesi come sequenze di parole che in una lingua esprimono uno o più messaggi. L'NLP combina quindi tecniche linguistiche con tecniche informatiche per permettere alla macchina di avere abilità simili a quelle degli esseri umani come quella di leggere e capire i significati delle parole.

Questa tecnica, può essere utilizzata nella distribuzione alimentare per quantificare l'effetto delle cannibalizzazioni, in quanto il NLP è in grado di assegnare automaticamente dei valori matematici ad ogni prodotto analizzando le informazioni anagrafiche, e, in seguito, stabilendo quanto i prodotti sono contestualmente simili o dissimili. Un prodotto con descrizione "tonno all'olio d'oliva" sarà contestualmente simile ad un prodotto con descrizione "tonno al naturale", meno simile di un prodotto con descrizione "salsa tonnata", totalmente dissimile con un prodotto con descrizione "crema spalmabile al cioccolato". Identificando scientificamente tali relazioni, è possibile stabilire l'impatto positivo, negativo o neutrale che una promozione su una referenza può avere sulle altre.

## 1.4 Una macchina dotata di Intelligenza Artificiale è cosciente?

Per rispondere a questa domanda per prima cosa è opportuno descrivere il concetto di coscienza, essa può essere definita come "la consapevolezza che il soggetto ha di sé stesso e del mondo esterno con cui è in rapporto, della propria identità e del complesso delle proprie attività interiori". Quindi quando si intende una macchina dotata di coscienza non si intende una macchina solamente intelligente ma che sia anche consapevole di sé. Risulta così fondamentale differenziare la conoscenza dalla coscienza.

Il quesito è dunque capire cosa rende il cervello umano, composto da una rete fittissima di neuroni, consapevole che la spia della propria automobile sia accesa e cosa rende l'automobile, sofisticato pezzo di elettronica e ingegneria, ignara.

L'argomento è sicuramente molto spinoso e dibattuto perché se prendiamo ad esempio Deep Blue, la macchina che vinse a scacchi contro il campione mondiale Kasparov, c'è chi direbbe che nonostante abbia conoscenza delle regole degli scacchi e di quando fare la mossa giusta al momento giusto, essa non sia cosciente. Infatti, ogni sua mossa è solamente l'applicazione di regole e algoritmi che le sono stati forniti dai programmatori.

La macchina non è cosciente poiché ad esempio non sa in modo autonomo stabilire che cosa sia giusto o che cosa sia sbagliato, è quindi possibile vedere la coscienza sotto una logica di etica. L'idea di giusto o sbagliato che può avere deriva da come sia stata programmata, e non è quindi di un pensiero sviluppato autonomamente.

## 1.5 L'utilizzo dell'IA nella filiera alimentare

Il mercato dell'Intelligenza Artificiale sta crescendo ogni anno con un tasso vicino al 20% e si sta espandendo sempre in più settori, tutti ad esempio conoscono le auto a guida autonoma o gli assistenti vocali come Siri di Apple o Alexa di Amazon, ma si potrebbero citare molti esempi meno noti.

Uno tra questi è quello del mondo del retail e della produzione, l'IA infatti è una enorme opportunità per le aziende in questo ambito grazie ai tanti utilizzi che ne derivano.

Un'applicazione fondamentale è quella nell'ottimizzazione della gestione della supply chain per:

- Analizzare dati dai clienti e fornitori;
- Ottimizzazione della gestione del magazzino;
- Previsione della domanda.

In relazione all'ultimo punto citato, questa tecnologia è in grado di tenere in considerazione tantissimi fattori, tra cui variabili esterne come le condizioni meteorologiche, le festività e i cambiamenti delle offerte dei competitor, le quali possono avere un impatto significativo sulle variazioni della domanda. Può essere intuitivo il fatto che, le previsioni meteorologiche possano incidere sulle vendite di un determinato prodotto, ciò che però è meno intuitivo è il quantificare analiticamente l'impatto che queste variabili esterne hanno sull'intero assortimento di un rivenditore e sulla programmazione degli ordini per un produttore. A questo scopo vengono in aiuto gli algoritmi di Machine Learning che possono individuare tali relazioni: ad esempio, è possibile sapere quanta birra verrà venduta se la temperatura percepita è di 30 gradi e quanta se la temperatura è di 10 gradi. Inoltre, mantenendo lo stesso esempio precedente, risulterebbe possibile anche sapere quali sono le conseguenze nella vendita di una specifica birra se un competitor ha messo una promozione sulla stessa marca di birra oppure su una marca differente.

Retailer e grossisti e produttori hanno infatti a disposizione ingenti quantità di dati riguardo le transazioni e le interazioni con i clienti, ottenuti sia attraverso canali offline che online.

Gli algoritmi permettono di sfruttare le informazioni contenute in questi dati, aiutando le aziende a prendere veloci e accurate decisioni. Questo rappresenta un fattore chiave in un settore in cui devono essere gestiti e controllati milioni di flussi di informazioni sui prodotti, cercando di far combaciare il più possibile domanda e offerta. Una piccola azienda, in media, si trova a dover interagire più di 400 milioni di transazioni giornaliere, provenienti da fonti diverse e con diverse tipologie di interpretazione, rendendo impossibile per l'uomo il compito di analisi e studio scientifico del dato.



Secondo un report di McKinsey, l'Intelligenza Artificiale, e più precisamente il Machine Learning, nelle previsioni di vendita permette alle aziende di ridurre gli errori dal 30% al 50% lungo la catena di approvvigionamento. Le previsioni, infatti, sono rese più precise dalla combinazione di centinaia di variabili che influenzano la vendita di un singolo prodotto. Chiaramente la qualità del processo di approvvigionamento ha un impatto diretto sulle entrate e, di conseguenza, sui profitti.

Questa tecnologia può portare dei benefici sia per il retail sia per le aziende di produzione. Lato retail, grazie ad una accurata previsione della domanda è possibile:

- Aumento dei ricavi grazie ad una migliore disponibilità sugli scaffali dei prodotti, dunque diminuzione delle rotture di stock di prodotti in continuativo e in promozionale;
- Ottimizzazione della gestione di magazzino, consentendo una riduzione dei costi di gestione delle merci nei centri di distribuzione e nei negozi;
- Ottimizzazione delle scorte di sicurezza per i prodotti freschi, tipicamente molto difficili da gestire, riuscendo a ridurre lo spreco alimentare e massimizzando quindi le vendite;
- Ottimizzazione degli spazi nel magazzino e negli scaffali, il che permette di introdurre nuovi prodotti in assortimento, andando ad eliminare i “basso-vendenti” precedentemente utilizzati come prodotti sostitutivi;
- Trasferire prodotti tra i punti vendita, risolvendo il problema dell'acquisto in colli e della vendita in pezzi;
- Riconciliazione dei documenti, allineando le bolle ricevute con gli ordini effettuati per verificare inefficienze, come gli inevasi.

Mentre, per le aziende di produzione, implementare un sistema di previsione della domanda significa:

- Aumento dei ricavi: consentendo di ridurre gli ordini inevasi e l'overstock nel breve e nel lungo periodo
- Ridurre aleatorietà nelle richieste dei clienti: analizzando l'andamento storico dei prodotti per capire cosa impatta le vendite e prevedendo picchi di vendita improvvisi

- Reperibilità materie prime: suggerendo la quantità di ingredienti necessaria per soddisfare gli ordini
- Calendarizzare in anticipo la preparazione della merce: considerando vincoli commerciali, logistici e operativi
- Ottimizzazione della produzione: massimizzando l'efficienza e organizzando al meglio sia linee che forza lavoro
- Ottimizzazione della gestione di magazzino: gestendo le scorte per ottimizzare le consegne
- Miglior rotazione del magazzino: anticipando i picchi e riducendo i costi di immobilizzazione merci

Un esempio di come l'IA nel retail può essere un fattore determinante lo si può riscontrare nel caso di Walmart e il suo primo Intelligent Retail Lab (IRL), un “negoziante intelligente” impostato per raccogliere informazioni su ciò che accade al suo interno tramite molteplici sensori, telecamere e processori.

All'interno dell'IRL, l'Intelligenza Artificiale, combinata con il Machine Learning e Deep Learning, ha introdotto una nuova modalità di gestione dell'inventario, grazie ad una rintracciabilità dei prodotti che, in tempo reale, assicura la presenza di tutti gli articoli sugli scaffali. La mancanza di scorte infatti è uno dei problemi principali con cui i retailer si confrontano e di conseguenza provoca una perdita in termini di mancate vendite che può portare, oltre che ad una drastica riduzione del fatturato annuo, anche ad una diminuzione di clienti abituali. Per riuscire nel suo obiettivo l'IRL di Walmart dispone di 1500 telecamere posizionate al soffitto dei punti vendita che, tramite image recognition, garantiscono che i prodotti siano sempre disponibili.

Per Image Recognition, o riconoscimento di immagini, si intende una tecnica di Deep Learning che permette alla macchina di riconoscere specifiche forme tramite l'analisi di immagini. Per rendere il tutto possibile il sistema viene allenato con grandi set di immagini raffiguranti prodotti presenti nel supermercato, raffiguranti, ad esempio, scatolette di tonno. La macchina proverà a classificare il prodotto visualizzato e dopo numerosi errori sarà in grado di riconoscere con un grado di accuratezza estremamente elevato che cosa è una scatoletta di tonno e che cosa invece non lo è. Così, Walmart ed il suo IRL riescono a verificare la presenza dei prodotti sugli scaffali.

## 1.6 Perché investire nell'Intelligenza Artificiale?

L'enorme progresso dell'Intelligenza Artificiale apre nuovi orizzonti verso tematiche fino a qualche tempo fa considerate futuristiche. Questo grazie anche a:

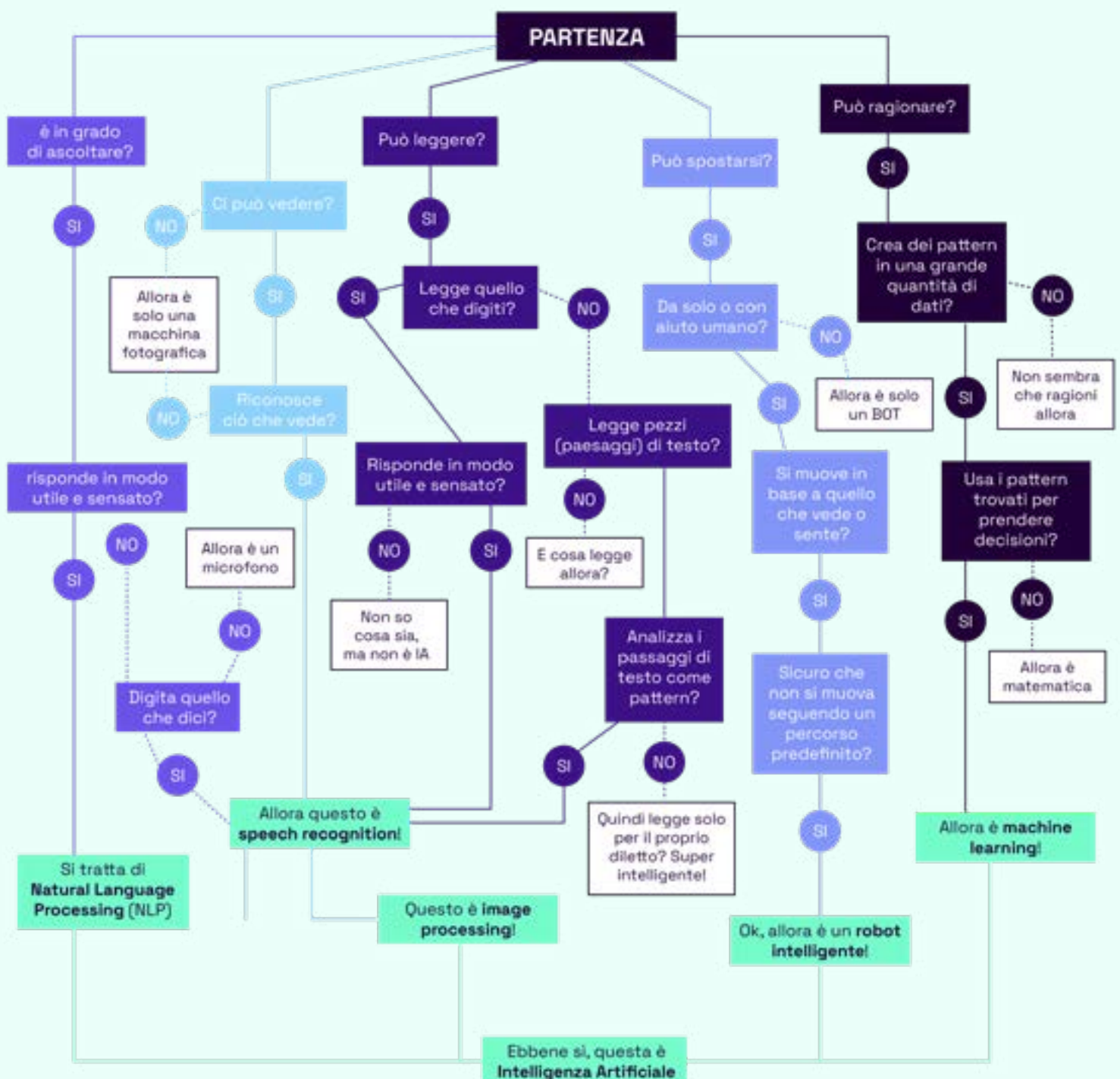
- Potenza di calcolo maggiore rispetto agli anni precedenti
- Possibilità di accesso ad enormi banche dati, che costituiscono un elemento fondamentale per lo sviluppo dei sistemi.

L'IA è in grado di fare analisi che sarebbero quasi impossibili da fare manualmente, tramite la gestione di grandi moli di dati essa riesce ad esplorarli, trovare correlazioni ed estrarne informazioni utili sulle quali basarsi per fare previsioni future.

Grazie ai numerosi benefici che ne derivano dall'utilizzo, gli investimenti delle imprese in soluzioni di Intelligenza Artificiale stanno crescendo in modo ingente. L'International Data Corporation (IDC) ha dettagliato una previsione sulla spesa globale per l'IA nel suo rapporto "Worldwide AI and Generative AI Spending Guide". Secondo questo studio: la spesa mondiale per l'intelligenza artificiale raggiungerà effettivamente \$632 miliardi entro il 2028. Questo valore include investimenti in applicazioni abilitate all'IA, infrastrutture e servizi IT e aziendali correlati.

## 1.7 Come identifico un'applicazione di Intelligenza Artificiale?

In senso lato l'IA si riferisce a macchine capaci di apprendere, ragionare e agire autonomamente. Allo stato attuale la maggioranza delle applicazioni di IA di cui si sente parlare appartiene alla categoria del Machine Learning. Quindi la domanda corretta non è "Come identificare un'applicazione di IA" (che è la capacità di un computer di eseguire compiti comunemente associati ad esseri intelligenti), piuttosto "come identificare un'applicazione di Machine Learning".



## 1.8 Differenza con i classici modelli statistici

È importante sottolineare le differenze che ci sono con i tradizionali modelli statistici. Entrambe le tecnologie aiutano a comprendere meglio i dati ma hanno due approcci differenti.

I modelli statistici tradizionali esistono da anni e precedono i computer. Si basano su piccoli campioni di dati con i quali fanno assunzioni a priori; ad esempio, la previsione della domanda si basa su uno storico di vendite passate. In questo contesto si assume a priori che la domanda passata sia un buon indicatore di previsione della domanda futura.

Negli ultimi anni però si sono aggiunti una serie di fattori in più da tenere in considerazione quando si effettua, ad esempio, una previsione della domanda futura di prodotti.

Tra questi: la presenza di nuovi dati da poter consultare, come le festività, dati provenienti dall'e-commerce, dati dei concorrenti, etc. La disponibilità di una potenza di calcolo più accessibile economicamente ed efficiente (con riferimento ai servizi cloud di noti provider con Microsoft, Amazon e Google) permette di controllare, attraverso previsioni di vendita sempre aggiornate, la crescente volatilità della domanda causata per esempio dall'introduzione di nuovi prodotti e/o di nuove promozioni commerciali.

Ecco perché il Machine Learning può giocare un ruolo fondamentale nella pianificazione della supply chain. In un contesto in cui la volatilità della domanda è il problema principale, può essere estremamente difficile creare una buona previsione. Il ML può prevedere la domanda non solo sulla base dei dati storici di vendita, ma anche tenendo conto di tanti altri parametri, tra cui variabili esterne come condizioni meteorologiche, stagionalità e festività.

Abbiamo quindi capito fino ad adesso che ciò che sta alle fondamenta di una previsione della domanda accurata sono i dati. È importante quindi che essi siano precisi ed estratti correttamente per poter essere studiati nel migliore dei modi attraverso un insieme di tecniche che fanno riferimento ad un nome comune: il Data Mining.

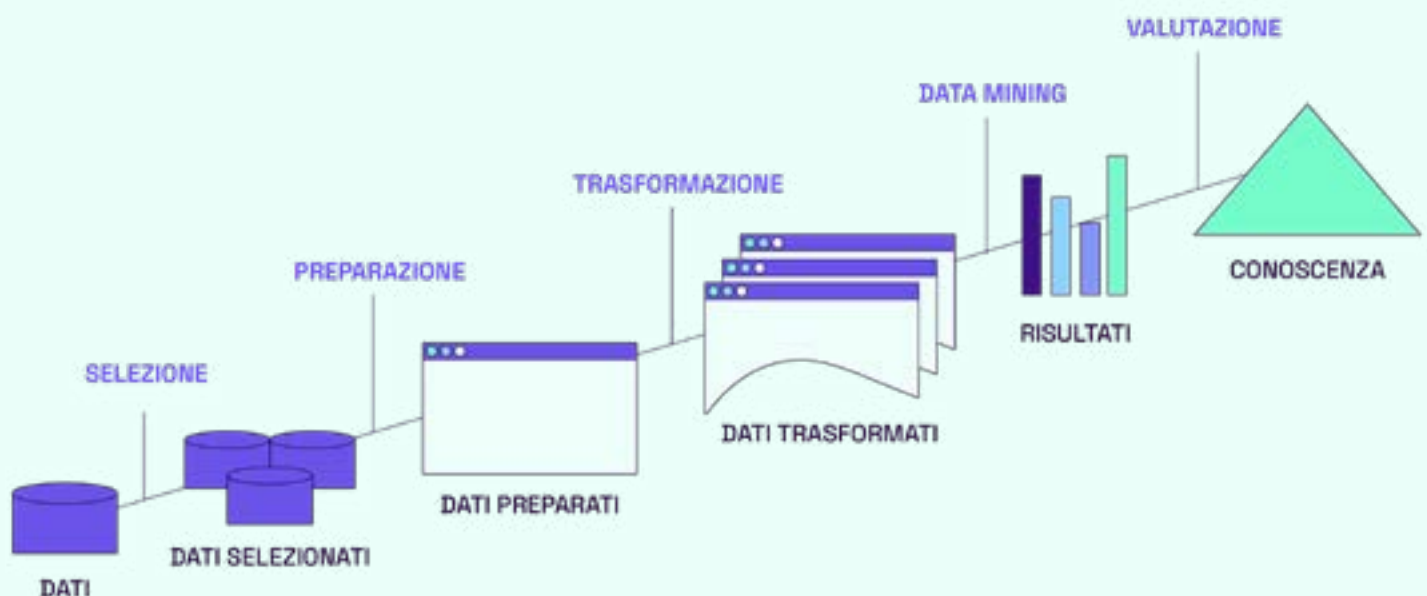
## 2.1 Data Mining: che cosa è e perché è importante

Per Data Mining si intende l'insieme di tecniche e metodologie che permettono l'estrazione di informazioni utili da grandi banche dati. I dati estratti ed elaborati tramite questo processo sono poi utilizzati per raggiungere obiettivi preimpostati come la massimizzazione o la minimizzazione di una funzione obiettivo.

Nel settore del retail, questo processo può aiutare i proprietari dei supermercati a riconoscere le scelte dei clienti. La cronologia degli acquisti di uno o più consumatori, le caratteristiche demografiche, l'orario di acquisto di certi prodotti, sono solo alcune fonti di dati tramite le quali è possibile riuscire a trovare delle correlazioni e capire le preferenze di acquisto dei clienti.

Nel settore della produzione alimentare, questo processo può aiutare i produttori a migliorare l'efficienza dei loro prodotti. Ad esempio, è possibile personalizzare le linee di produzione in base alle preferenze dei clienti, identificando i prodotti più popolari e adattando l'offerta. Se i dati mostrano che un certo tipo di pane è particolarmente apprezzato in una determinata stagione, si può aumentare la produzione di quel prodotto durante quel periodo.

In questo modo, è possibile anche raggruppare i clienti che presentano caratteristiche simili in gruppi, chiamati cluster, a cui verranno indirizzate promozioni specifiche. Per esempio, al cliente che acquista di rado ma in grande quantità verrà rivolta una promozione diversa rispetto al cliente che acquista molto più frequentemente ma con importi inferiori. Un processo di data mining che porta alla "estrazione della conoscenza" può essere identificato come in figura:



La selezione, preparazione, trasformazione e valutazione dei dati sono fasi imprescindibili per avere massima conoscenza ed essere in grado di optare per una soluzione ottimale di un problema. Come nell'esempio precedente una soluzione al problema "promozioni più efficaci" può essere quella di capire quali promozioni mirate effettuare ad un certo cluster di clienti al fine di aumentare la loro fidelizzazione, che nell'iter precedente rappresenta l'obiettivo da raggiungere.

Quando si parla di "dati grezzi" si intendono informazioni che non sono state elaborate e possono essere quasi senza significato se considerate come tali. Il codice binario può esserne un esempio, preso solo come una successione di 0 e 1 ha ben poco significato per l'utente del computer, ma se esso viene elaborato da una macchina può fornire informazioni più comprensibili. Ad esempio, questo codice binario "01000001" per noi non ha nessun significato ma per la macchina corrisponde alla lettera maiuscola "A".

I dati grezzi hanno quindi bisogno di una rielaborazione, anche perché talvolta vi sono informazioni mancanti che vanno integrate, oppure dati errati che hanno bisogno di correzioni. Lungo il percorso vi è dunque anche una notevole riduzione del volume di informazioni, ma un aumento del valore delle stesse in quanto più accurate e meno soggette ad errori. Infatti, se all'inizio bisogna selezionare i dati grezzi da enormi dataset, spesso anche generici, dopo la fase di Data Mining si ottengono pattern di informazioni utili per arrivare alla decisione finale, che senza questo processo sarebbero state impossibili da ottenere.

## 2.2 La metodologia CRISP-DM

Abbiamo visto che il Data Mining è una fase del processo di estrazione di conoscenza, e quando deve essere attuata vi sono diversi modi per farlo. Il "Cross Industry Standard

Process for Data Mining" (CRISP-DM) è il modello analitico più utilizzato a questo scopo.

La metodologia CRISP-DM è caratterizzata da questi sei passaggi principali, che verranno spiegati accompagnati da esempi relativi alla previsione della domanda nel retail:

**Business Understanding:** questa fase iniziale si concentra principalmente sul definire il problema di business o individuare il tipo di informazioni necessarie. L'obiettivo è quello di comprendere il problema e come la risoluzione andrà a impattare sulle decisioni di business.

Si risponde alla domanda "Qual è il problema che voglio risolvere?" Ad esempio: prevedere la vendita dei prodotti in promozione in modo più accurato tramite algoritmi di previsione della domanda.

**Data Understanding:** fase necessaria alla raccolta, comprensione e valutazione della qualità dei dati. In particolare, viene analizzata la bontà delle basi di dati identificando anomalie (note anche come outliers) e incompletezze che necessitano integrazioni da fonti esterne.

In questa fase vengono anche identificate le relazioni e le tendenze tra le variabili chiave (ad esempio il rapporto che c'è tra la variabile prezzo e la quantità venduta).

Si risponde alla domanda "Di quali dati abbiamo bisogno?"

Nel caso di una catena di supermercati dovremmo andare a considerare i dati provenienti dagli scontrini di acquisto dove vengono riportati i prodotti acquistati, l'orario e i dati del piano promozionale presente e passato.

Per un'azienda di produzione alimentare, i dati necessari potrebbero includere le quantità di materie prime acquistate, i tempi di produzione, i feedback dei clienti. Questi dati permettono di analizzare l'efficienza produttiva e migliorare il controllo qualità.

**Data Preparation:** fase che comprende tutte le attività relative alla costruzione del dataset finale tra cui figurano la selezione dei dati, la pulizia degli stessi, la costruzione di nuove informazioni ed eventuali trasformazioni. Questa fase è la più dispendiosa a livello di tempo all'interno del processo perché le operazioni di preparazione non sono immediate, richiedono uno studio accurato ed una navigazione approfondita dei dati. Le variabili primitive presenti all'interno della base di dati iniziale potrebbero non essere sufficienti.

Tramite una fase di data engineering esse vengono combinate e/o manipolate al fine di crearne di nuove più elaborate che aggiungano e diano informazioni aggiuntive al modello di intelligenza artificiale.

-Risponde alla domanda "| dati sono sufficienti, consistenti e accurati?". È necessario preparare i dati prima di fornirli al modello che porterà alla soluzione desiderata. Se i dati non sono sufficienti sarà necessario crearne di nuovi come, ad esempio, capire qual è la fascia oraria in cui solitamente si vende un certo prodotto, oppure in che giorno o in base a quali condizioni meteorologiche viene effettuato un certo acquisto, si possono anche creare cluster di clienti con caratteristiche simili oppure studiare la correlazione tra prodotti, ad esempio, se generalmente la maggior parte dei clienti che compra le fragole compra anche la panna spray.

In questo modo si può iniziare a preparare i dati per fare una previsione della domanda.

**Modeling:** fase di scelta ed applicazione dei modelli di machine learning che permettono di raggiungere un obiettivo preimpostato. Una volta individuata



la tipologia di modelli che si vuole utilizzare, al fine di massimizzare i risultati, è necessario effettuare una fase di settaggio (tuning) dei parametri con cui il modello viene allenato.

Durante questa fase il modello viene allenato con diversi parametri ottenendo performance più o meno soddisfacenti a seconda delle configurazioni. Qualora modificando i parametri, i risultati non dovessero essere soddisfacenti, sarà necessario tornare alla fase di ricerca e creazione di nuove variabili da dare in pasto al modello nella fase di addestramento.

- Risponde alla domanda “Quale tecnica si adatta meglio ai dati a disposizione e porterà ai risultati che sono stati prefissati?”

In ambito retail e produzione le variabili e i dati della fase precedente verranno utilizzati come input per un algoritmo di previsione della domanda che si baserà sui dati storici ma anche variabili esterne come condizioni metereologiche, festività e stagionalità.

**Evaluation:** durante la fase di valutazione, i modelli vengono valutati e confrontati al fine di selezionare quello che risponde maggiormente agli obiettivi di business e determinare se esistono eventuali ragioni per le quali il modello possa essere considerato debole.

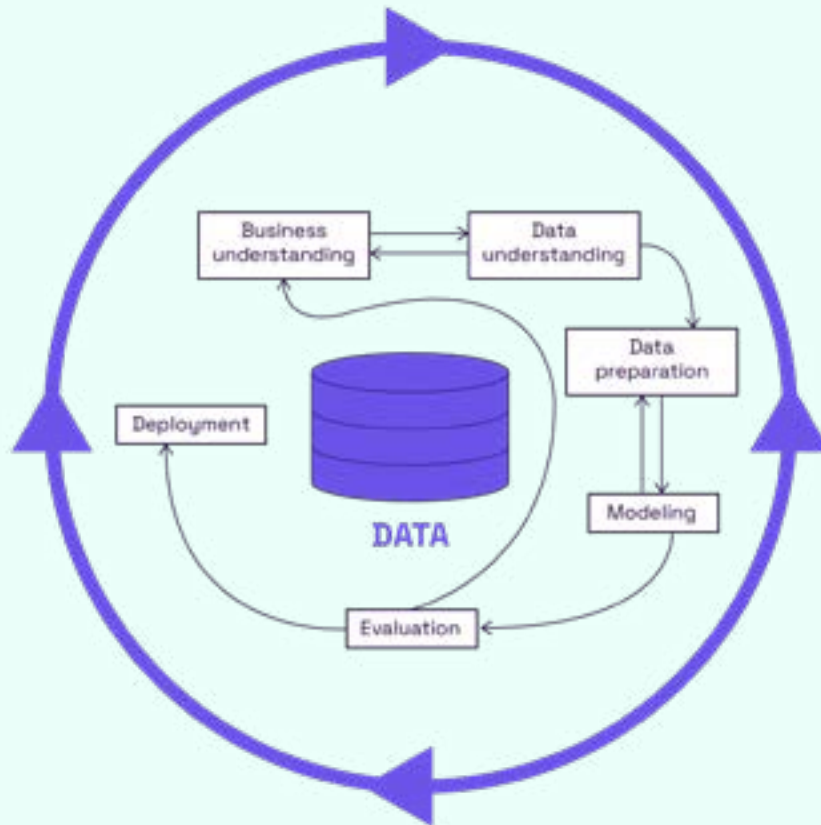
-La domanda da porsi in questa fase è “Ottengo una risposta significativa al mio problema iniziale?”. Per vedere se il modello è adatto, in questo caso, è necessario quindi analizzare la corrispondenza dei risultati del data mining ai criteri di successo aziendali.

**Deployment:** la corretta implementazione del modello è fondamentale per garantirne la sua effettiva integrazione nei processi aziendali.

-Risponde alla domanda “Come verrà messo in produzione il modello?”. In questa fase, quindi, avviene prima un’attività di pianificazione e monitoraggio della diffusione dei risultati e poi l’elaborazione di un rapporto finale per la revisione del progetto.

## LEZ. 2: DATA MINING

La differenza che c'è tra questo modello e altri processi tipicamente lineari e unidirezionali è che le fasi che caratterizzano il CRISP-DM sono pensate per essere ripetute ciclicamente.



In questo modo ad esempio se ci si rende conto che nella fase di “Modeling” non vi sono modelli validi è possibile tornare indietro alla fase di “Data Preparation” e verificare se è necessario selezionare un altro sottoinsieme di dati o se vi sono informazioni mancanti. In un modello lineare o unidirezionale, invece, le fasi si muovono linearmente verso un’unica direzione. Ad esempio, la costruzione di un palazzo prevede degli step ben precisi. Non è possibile costruire l’edificio e poi rifare le fondamenta. I progetti di mining sono per natura incerti, poiché strettamente dipendenti dalla bontà dei dati forniti.

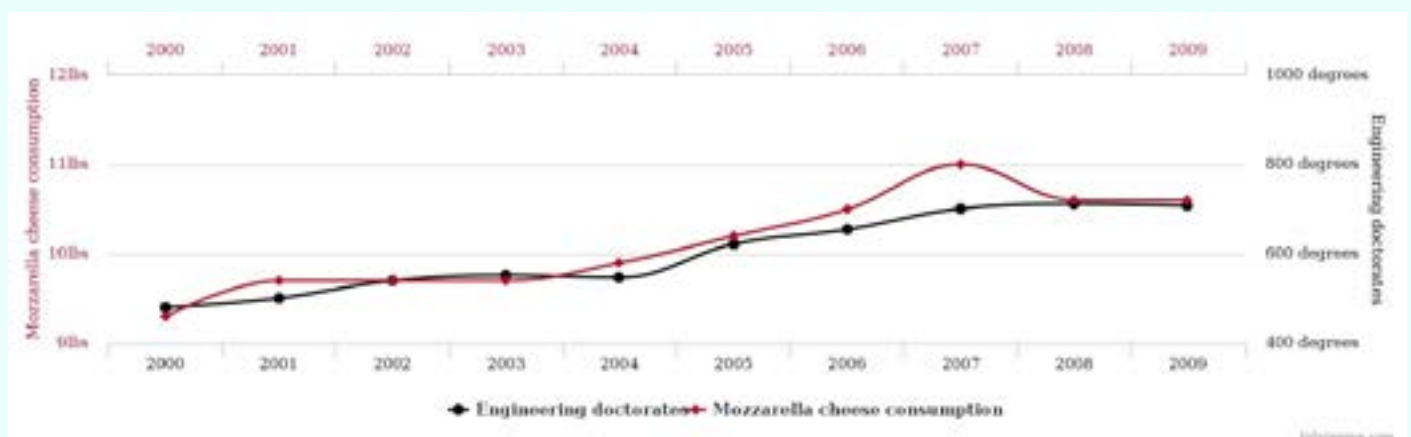
Tornare indietro e integrare sono elementi fondamentali di un processo di mining, diametralmente opposto ad un processo ingegneristico a cascata.

## 2.3 Data Mining e Intelligenza Artificiale

È necessario sottolineare l'importanza ricoperta dal Data Mining quando si parla di IA. Infatti, i dati ottenuti con questo processo fungono da input alla macchina per il raggiungimento di obiettivi preimpostati.

È fondamentale quindi che i dati selezionati siano il più accurati possibile, poiché da dati sbagliati non possono che derivare dei risultati sbagliati. Questo fenomeno da evitare viene definito come uno scenario "garbage in, garbage out" in cui i dati in entrata errati (garbage in) producono a loro volta risultati errati (garbage out).

Altre volte invece è necessario distinguere da causalità e casualità. Non sempre un'informazione che sembra correlata al problema in esame ha senso che lo sia. Un esempio di questo fenomeno sono le correlazioni spurie, ovvero situazioni in cui la macchina, basandosi su dati non accurati, individua correlazioni dove in realtà non ci sono facendo riferimento ad informazioni prive di senso. Guardiamo l'esempio in figura:



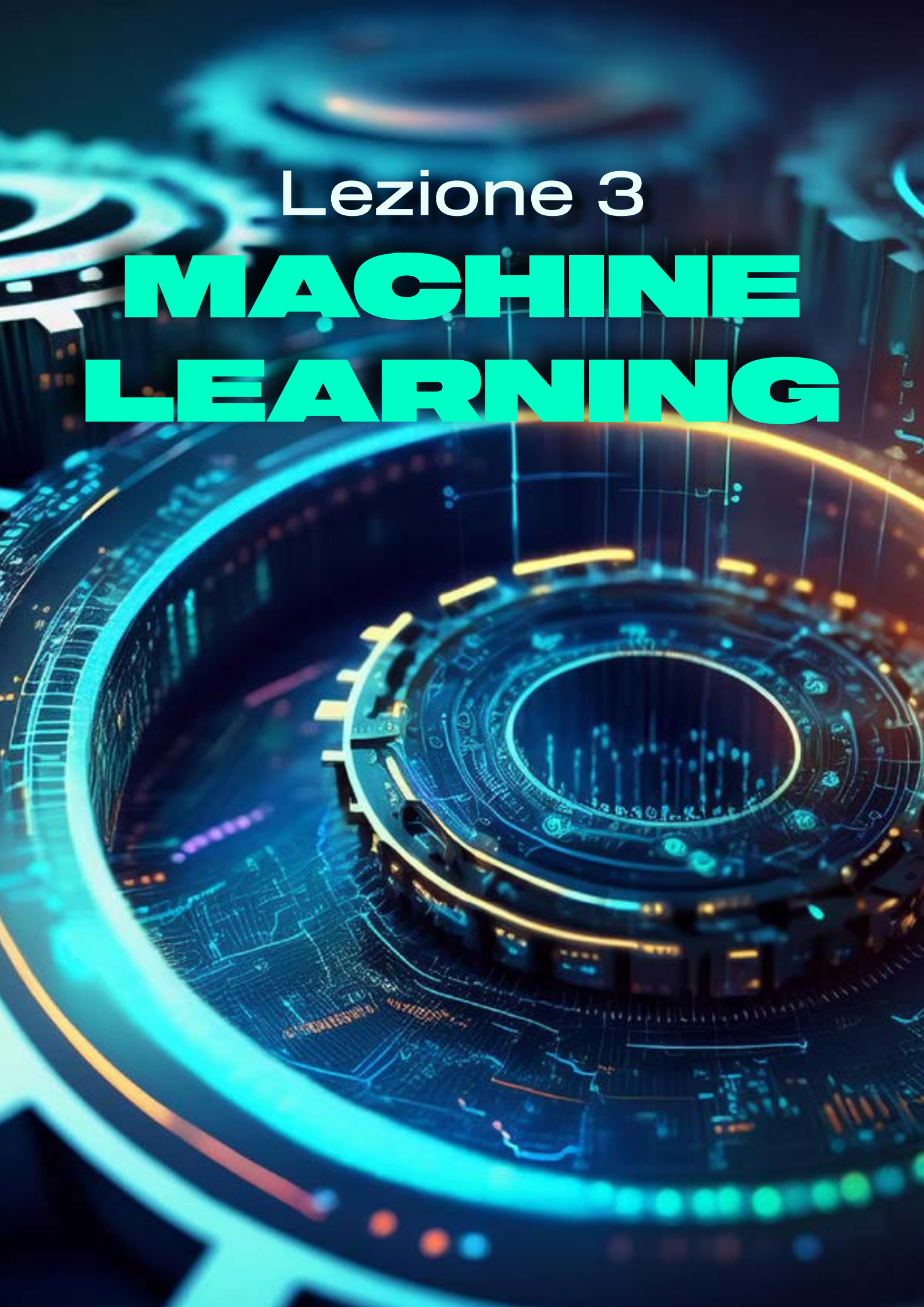
In questo grafico viene individuata una correlazione secondo cui chi ha un dottorato in ingegneria compra più mozzarelle.

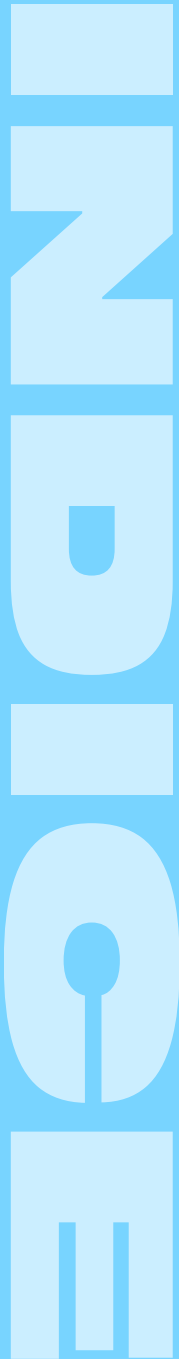
È evidente che questa è una associazione scorretta che esalta l'importanza di avere dati accurati, poiché se l'obiettivo è quello di prevedere la domanda futura delle vendite, è fondamentale non basarsi su correlazioni erronee come queste che porterebbero al riordino di determinati prodotti, in questo caso le mozzarelle, basandosi su informazioni prive di fondamento.

Abbiamo quindi visto che per parlare di Intelligenza Artificiale applicata a modelli predittivi è necessario avere dati accurati (qualità batte quantità), tramite un corretto utilizzo del processo di Data Mining. Abbiamo anche visto nella prima lezione che nel retail e nelle aziende di produzione per eseguire una corretta previsione della domanda è necessario il Machine Learning, ma cos'è nello specifico?

Lezione 3

# **MACHINE LEARNING**





3.1 Machine Learning

3.1 Classificazione, Regressione e Clustering

## 3.1 Machine Learning

Il Machine Learning (ML), è un sottoinsieme dell'IA che si occupa di creare sistemi che apprendono o migliorano le proprie performance in base ai dati che utilizzano. Ognuno di noi fa uso del Machine Learning quotidianamente, una sua applicazione ad esempio è quella del riconoscimento vocale di cui sono dotati molti smartphone che permettono di attivare comandi tramite la propria voce. Un altro utilizzo che può essere fatto è quello che permette alle aziende di realizzare pubblicità targettizzate tramite la profilazione degli utenti grazie alla quale, è possibile raccogliere interessi e necessità che ciascun utente manifesta nelle proprie ricerche su internet e mentre visita le pagine web. Se un utente passa molto tempo su una pagina web che parla di calcio, è altamente probabile che le prossime pubblicità che visualizzerà sui social saranno pubblicità di scarpe da calcio o articoli sportivi in generale.

Data Mining e Machine Learning sono strettamente correlati, poiché il primo è fondamentale per l'estrazione di informazioni e dati accurati che serviranno da input al secondo per ottenere una conoscenza più profonda e delle previsioni più accurate.

A seconda del tipo di algoritmo utilizzato per permettere l'apprendimento alla macchina, ossia a seconda delle modalità con cui la macchina impara ed accumula dati e informazioni, si possono suddividere differenti sistemi di ML.

Due dei principali tipi di algoritmi attualmente utilizzati sono quelli caratterizzati da un apprendimento supervisionato da una parte e quelli non supervisionati dall'altra.

- **ML supervisionato:** L'apprendimento supervisionato è una tecnica che genera delle previsioni grazie ai dei modelli che si allenano su dei dati "supervisionati" dall'uomo, ovvero in cui ad ogni input è associato un output ideale che il modello cercherà di replicare. Infatti, l'algoritmo apprende da un set di dati già etichettato e con un output predefinito. Quando si parla di dati etichettati si fa riferimento a dati a cui è stata apposta una specie di etichetta virtuale con scritte alcune proprietà o caratteristiche di quei dati. Parliamo quindi di ML supervisionato quando ad esempio è necessario addestrare un algoritmo a capire se un cliente smetterà di acquistare o resterà fedele in base al comportamento di acquisto. Il sistema di intelligenza artificiale si addestra con un insieme di dati in cui l'input è il comportamento del cliente (acquisti ricorrenti, numero di prodotti acquistati, soddisfazione ecc.).

Ogni riga di input dell'addestramento è basata su un reale comportamento che viene etichettato manualmente (smette di acquistare/resta fedele).

- **ML non supervisionato:** Questa tipologia di algoritmi usa un approccio più indipendente, in cui un computer impara a identificare processi e schemi complessi senza la guida attenta e costante di una persona. Il ML non supervisionato implica una fase di addestramento basata su dati privi di etichette e per i quali non è stato definito un output specifico.

Ad esempio, possiamo dare al sistema in input un dataset di addestramento con le caratteristiche di 5 specie di piante ma senza classificare la specie. In questo caso, l'algoritmo genera automaticamente un modello di classificazione in base alle caratteristiche in comune che ha rilevato e raggruppa le piante sfruttando esclusivamente le informazioni che è riuscito a ricavare dai dati, senza che sia l'essere umano a raggrupparli a priori.

## 3,2 Classificazione, Regressione e Clustering

In apprendimento supervisionato le tipologie di problemi che è possibile risolvere si dividono in due grandi branche: classificazione e regressione.

La classificazione cerca di risolvere problemi il cui risultato è finito e limitato ad un gruppo di risposte. Esempi di problemi di classificazione sono:

- Sulla base di un insieme di informazioni storiche bancarie il cliente si dimostrerà insolvente -> (si/no)
- Sulla base di una serie di proprietà e caratteristiche estetiche e di forma questa pianta è una pianta carnivora → (si/no)
- In relazione ad un insieme di informazioni sul clima, habitat circostante, flora e fauna, in quale continente vive una determinata persona. persona → (Europa, Asia, Africa, America, Australia)

Un metodo di classificazione è il “Decision Tree” ovvero quello relativo a “gli alberi decisionali”: molto comuni nell’ambito del Machine Learning supervisionato perché veloci e semplici da interpretare. In questo algoritmo i dati che vengono forniti come input vengono continuamente divisi in gruppi in base a certi criteri. Si basano su:

- **Nodi:** ovvero i luoghi in cui avviene la divisione.
- **Foglie:** sono i risultati intermedi o finali, ovvero i luoghi dove finiscono i dati una volta separati.

L’obiettivo è quello di dividere l’insieme iniziale di dati in gruppi, i quali dovrebbero essere il più omogenei internamente (caratteristiche simili tra i compo-

# LEZ. 3: MACHINE LEARNING

nenti del singolo gruppo) ed eterogenei tra loro (caratteristiche dissimili tra i vari gruppi). Per capire meglio questo procedimento è utile aiutarsi con un esempio. Supponiamo di dover decidere in base al meteo e alla temperatura esterna se giocare o non giocare a tennis. Il problema in questione, naturalmente, ha solo 2 possibili risultati ma è necessario valutare le combinazioni delle variabili considerate per capire se giocare o meno a tennis.



I nodi sono i luoghi dove avvengono le divisioni, quindi si considera la situazione in cui piove oppure no, mentre le foglie sono i risultati dove possiamo trovare i dati a seguito di una divisione come vediamo nell'ultimo livello.

Considerando la situazione a sinistra possiamo notare che in caso di tempo soleggiato, con la temperatura esterna inferiore a 30 gradi si può andare a giocare a tennis. Qualora la temperatura fosse sopra i 30 gradi, non è consigliato giocare a tennis. Quello rappresentato è chiaramente un esempio molto semplice di un albero decisionale mentre i procedimenti che si basano su Machine Learning considerano centinaia di variabili e prevedono risultati derivati da migliaia di divisioni dell'albero e diversi livelli di profondità. Diventa quindi comprensibile come prendere decisioni accurate a problemi ben più complessi sia impossibile, se si utilizza solo il ragionamento umano per questo è utile fare affidamento su tecnologie avanzate.

Gli algoritmi di **regressione**, rispondono a domande che non hanno un valore finito come risposta. Esempi di problemi di regressione sono:

- In relazione all'inflazione e all'andamento dei prezzi di mercato del grano, quanto costerà la pasta al kg il mese prossimo -> (1€,2€,3€, ecc.)
- Qual è la velocità massima di una macchina con determinate caratteristiche -> (220 km/h,220.5 km/h,220.566 km/h, ecc.)

In tal senso, anche la previsione della vendita è un problema di regressione ed è possibile formulare il problema nella forma seguente:



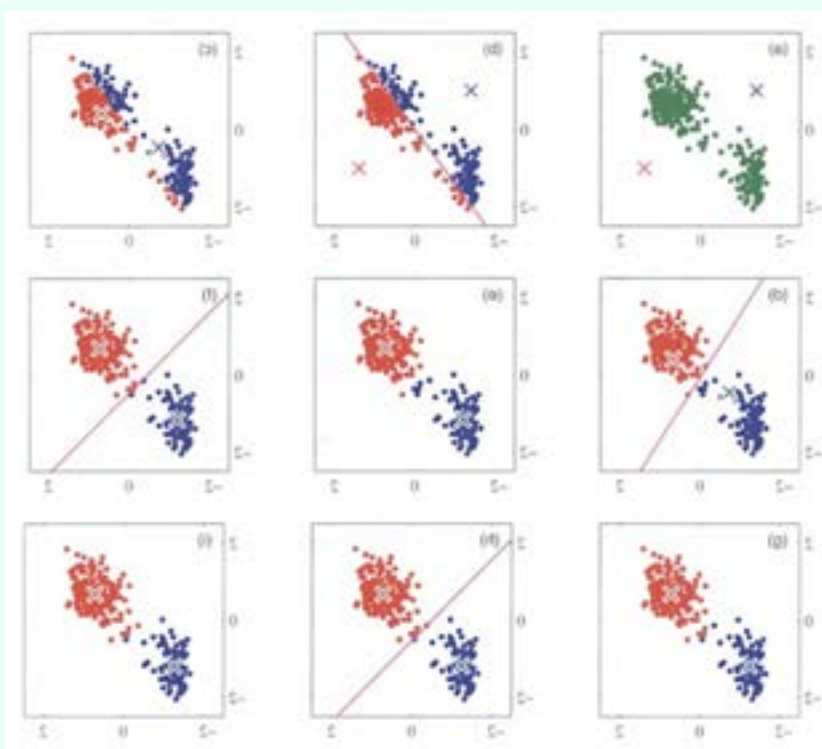
# LEZ. 3: MACHINE LEARNING

- In relazione ai prezzi, alle vendite storiche, ai prezzi dei competitor, ad eventi esogeni e a relazioni con altri prodotti, qual è la previsione di vendita per domani di un determinato prodotto? -> (1 pezzo, 2 pezzi, 3 pezzi, ecc.).

Alcune delle più popolari tecniche non supervisionate utilizzate all'interno del retail sono sicuramente gli algoritmi di **clustering** che identificano autonomamente le somiglianze tra le osservazioni e le divide sulla base di caratteristiche comuni. Questi raggruppamenti sono denominati "cluster".

Ad esempio, si prenda in considerazione i clienti online di uno specifico e-commerce di articoli sportivi. Ognuno di essi presenta delle caratteristiche personali e dei comportamenti di acquisto come ad esempio, sesso, età, frequenza di acquisto, ecc. Utilizzando un algoritmo di clustering è possibile raggruppare i clienti al fine di rivolgere loro offerte speciali in relazione al comportamento di acquisto. A priori non è noto il comportamento di acquisto e non è possibile sapere se sono dei clienti che praticano sport a livello agonistico o dilettantistico, se sono dei clienti occasionali o affezionati. Sarà compito dell'algoritmo stesso scoprire se tra i comportamenti di acquisto dei clienti vi sono delle caratteristiche che consentano di definire dei gruppi di consumatori.

Il k-means è uno degli algoritmi di clustering più diffuso e performante. L'obiettivo di questo algoritmo è quello di raggruppare le osservazioni simili, scoprendo pattern nascosti e dividendo il dataset in un numero  $k$  di clusters. Un cluster è quindi un insieme di osservazioni che condividono caratteristiche simili. Definiti il numero di  $K$  cluster desiderati, sarà lo stesso algoritmo a raggruppare le osservazioni su cui viene addestrato.



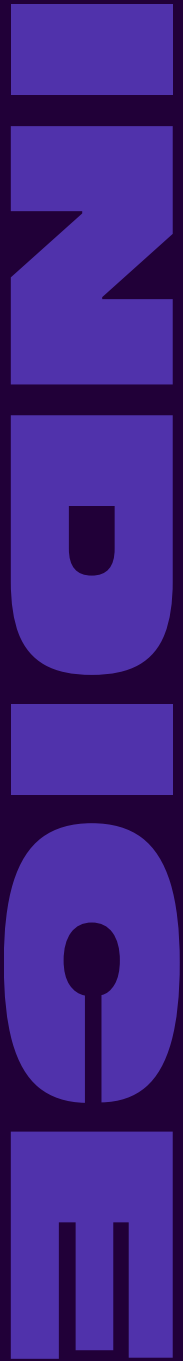
Nell'immagine è possibile notare come durante l'apprendimento l'algoritmo cerca di dividere progressivamente, in maniera sempre più netta un insieme di osservazioni. Alla nona iterazione (in basso a destra) l'algoritmo ha delineato 2 gruppi differenti di osservazioni ben distinguibili.

Una straordinaria caratteristica del ML è la sua capacità predittiva. In passato, le decisioni aziendali venivano spesso prese sulla base di risultati storici mentre oggi, come abbiamo visto grazie a questa tecnologia, si possono utilizzare analisi molto più avanzate che tengono in considerazione molti più fattori, come per esempio le variabili esterne.

Come analizzato nella seconda lezione in cui si esponeva la metodologia CRISP-DM, il problema era prevedere le vendite in caso di promozione (pag. 8) e quello che si ricercava era quindi un'analisi predittiva. Tuttavia, se dovessimo risolvere un problema legato alla decisione di acquisto o rifornimento prodotti, quest'ultima non sarebbe sufficiente. Infatti, al fine di raggiungere questo obiettivo è necessario introdurre il concetto di analisi prescrittiva.

## Lezione 4

# **DIFFERENZA TRA ANALISI PREDITTIVA E PRESCRITTIVA**



## 4.1 Differenza tra analisi predittiva e prescrittiva

# LEZ. 4: DIFFERENZA TRA ANALISI PREDITTIVA E PRESCRITTIVA

## 4.1 Differenza tra analisi predittiva e prescrittiva

Un modello predittivo analizza dati attuali e storici per fare previsioni, come la probabilità che un evento si ripeta in una certa situazione. Ad esempio, è possibile prevedere basandosi su dati storici e variabili esterne, quanti chili di banane verranno vendute la prossima settimana in un determinato punto vendita a seconda delle condizioni metereologiche che ci saranno. Ma questo è sufficiente per massimizzare il profitto dei retailer? No, perché vi sono altre serie di variabili che devono essere considerate per comprendere la corretta quantità di prodotto da riordinare come le giacenze e i vincoli a fornitore.

Per questo motivo, i retailer dovrebbero piuttosto adottare un'analisi prescrittiva, attraverso la quale, partendo da risultati di previsione, la macchina è in grado di fornire suggerimenti dettagliati per prendere decisioni che massimizzino le performance economiche dell'azienda. Un esempio lato produzione: **analisi prescrittiva vs analisi predittiva**

Tramite il Machine Learning, come abbiamo già visto, è possibile attingere ad una vasta gamma di dati interni ed esterni, rivelare pattern e correlazioni precedentemente sconosciuti che si trovano a livello di singolo negozio, prodotto o scaffale. In seguito, è possibile ottimizzare i risultati ottenuti al fine di fornire raccomandazioni settimanali o anche giornaliere per adattare i prezzi, le promozioni o l'assortimento di ogni punto vendita con l'obiettivo di massimizzare i profitti. Secondo dati di McKinsey, l'analisi prescrittiva può aumentare le vendite del 2-5%.

Ma cosa manca al ML per raggiungere questo risultato? È necessario tenere in considerazione molti altri vincoli come, nel caso di un retailer o di un grossista, il lead time, la quantità minima da riordinare, scorte di copertura, vincoli di pedana, gli sconti a valore, etc.

Per tenere in considerazione tutte queste variabili contemporaneamente è necessario utilizzare algoritmi matematici di ottimizzazione lineare, che incrociano previsioni di vendita e vincoli sopra elencati scegliendo così l'opzione che massimizzi la differenza tra ricavi e costi, ovvero il margine.

Questo è l'obiettivo dell'analisi prescrittiva, analizzare svariate situazioni e suggerire quella economicamente più profittevole.

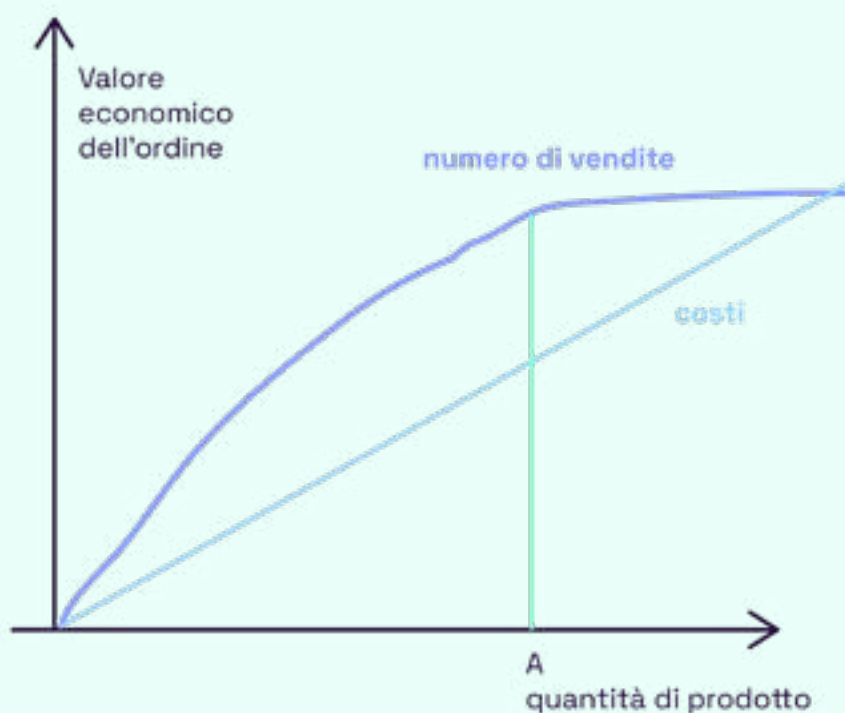
Il risultato principale dell'analisi prescrittiva nel settore del retail e della produzione è sicuramente quello di prendere delle decisioni economicamente più corrette, tuttavia le conseguenze dell'adozione di tale approccio sono molte. Ad esempio, partendo dal retail, è possibile capire se mantenere sullo scaffale

## LEZ. 4: DIFFERENZA TRA ANALISI PREDITTIVA E PRESCRITTIVA

prodotti basso vendenti perché motivo di attrazione per i migliori acquirenti, o se rimuoverli per introdurre nuovi articoli che potrebbe aumentare la marginalità di specifici settori del punto vendita. Questo è qualcosa che l'analisi predittiva da sola non potrebbe fare, perché limitata a prevedere solo le future vendite dei prodotti.

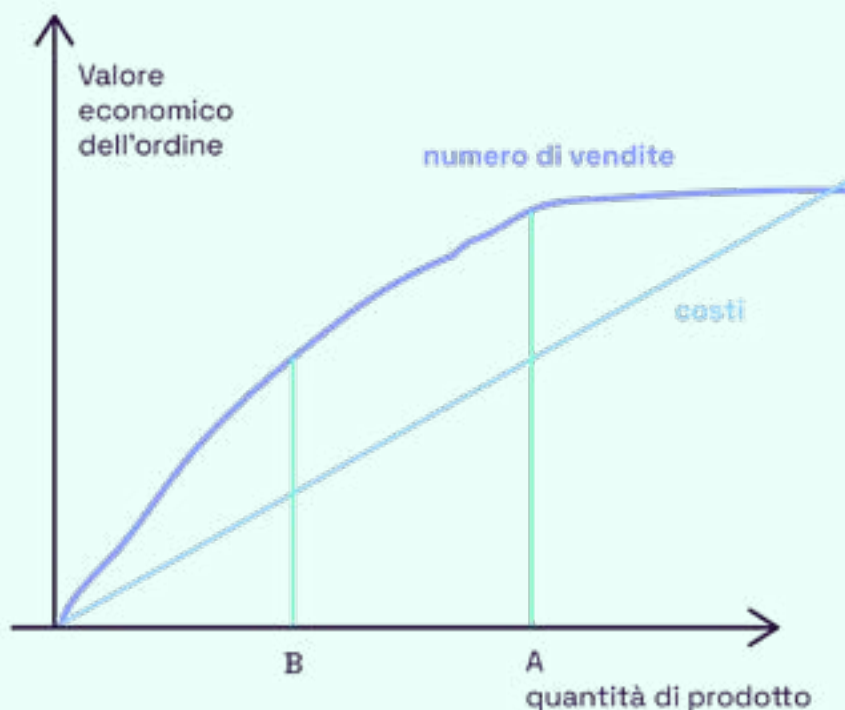
Un altro esempio di applicazione di tale approccio è il riconoscimento di diversi pattern, per raccomandare la tempistica delle promozioni e dei cambiamenti di prezzo dei prodotti. Queste correlazioni, tuttavia, non sono semplici da riconoscere e le variabili da tenere in considerazione sono sempre molto complesse e varie. Per esempio, in alcune città le vendite di avocado e patatine potrebbero aumentare del 5% il venerdì prima di un tipico evento sportivo ma lo strumento prescrittivo potrebbe raccomandare di stoccare il 20% in più di questi due prodotti nei negozi dove è più probabile che le squadre locali raggiungano i tempi supplementari.

Questa analisi può quindi raccomandare specifici cambiamenti di assortimento e stimare il valore monetario di tali cambiamenti. In ambito approvvigionamento, si può capire quindi quanto è importante l'analisi prescrittiva per le aziende del settore alimentare poiché il suggerimento che fornisce è la quantità da riordinare che cerca di massimizzare il profitto trovando un equilibrio tra il livello massimo di vendite raggiungibili e il livello minimo di costi sostenibili. Per spiegare meglio questo concetto può essere utile guardare questo grafico:



## LEZ. 4: DIFFERENZA TRA ANALISI PREDITTIVA E PRESCRITTIVA

Si potrebbe essere portati a pensare che la quantità ottimale di prodotti da ordinare/produire sia in corrispondenza del punto A poiché corrisponde a un massimo di vendite. In realtà però un modello basato su analisi prescrittiva potrebbe suggerire che la quantità ideale sia in corrispondenza del punto B, perché corrisponde a quantità minori di prodotto ma anche a costi proporzionalmente minori. Questo sta a significare che ciò che si ricerca non è la quantità massima di vendita ma la quantità di prodotto che massimizzi la differenza tra ricavi e costi.



Questo accade perché le vendite non sono l'unico fattore determinante ma ci sono altri fattori da tenere in considerazione come l'aumento della quantità di prodotto, che non è direttamente proporzionale all'aumento delle vendite ma può portare ad un maggiore rischio di overstock. Dunque il beneficio a cui aspirare per le aziende non è tanto un miglioramento dell'accuratezza di previsione, piuttosto si deve ambire a prendere decisioni più economicamente performanti. Alcuni di questi indicatori di performance obiettivo ottenibili sono:

- Diminuzione degli esuberanti di magazzino fino al 46%
- Incremento dell'efficienza di vendita del 18%
- Diminuzione dei costi di immobilizzazione di magazzino -8%

**Perché ricordati, come diciamo sempre, il futuro è di chi è in grado di anticiparlo!**

# Il corso finisce qui!

Clicca sul link per eseguire il quiz, testa la tua conoscenza e ricevi l'attestato. Mandaci il tuo feedback per permetterci di migliorare il contenuto e suggerisci i prossimi argomenti da trattare!

## QUIZ

Completa le domande sulle 4 lezioni del corso  
e scopri quanto hai imparato!

**CLICCA QUI**

Alla fine del test riceverai il tuo attestato  
di partecipazione al corso.