IA, uno sguardo dietro le quinte

IPPOG Masterclass

INFN-LNF

Leonardo Plini
National PhD student in Artificial Intelligence
INFN-LNF and Sapienza University of Rome
leonardo.plini@lnf.infn.it

INFN
Istituto Nazionale di Fisica Nucleare
Laboratori Nazionali di Frascati

PINlab
PERCEPTION AND INTELLIGENCE

SAPIENZA
UNIVERSITÀ DI ROMA

Chi sono

- Diploma Liceo Classico
- Laurea Triennale in Ingegneria Matematica, Politecnico di Milano
- Laurea Magistrale in Data Science, Sapienza Università di Roma
- Dottorando in Intelligenza Artificiale, Sapienza Università di Roma e INFN - LNF

Intelligenza Artificiale

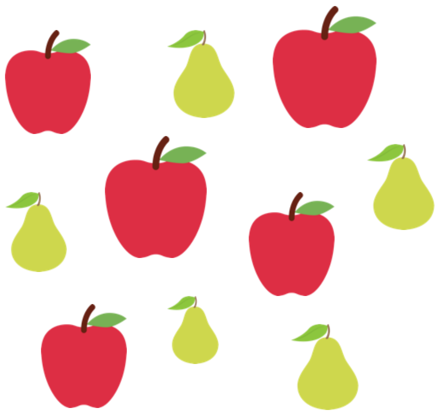
Intelligenza Artificiale

SCAN ME

SCAN ME

Cosa definisce una AI

“Un programma informatico si dice che apprenda dall'**esperienza E** rispetto a una classe di **task T** e una misura di **performance P**, se la sua performance nei compiti in T, misurata da P, migliora con l'esperienza E.”



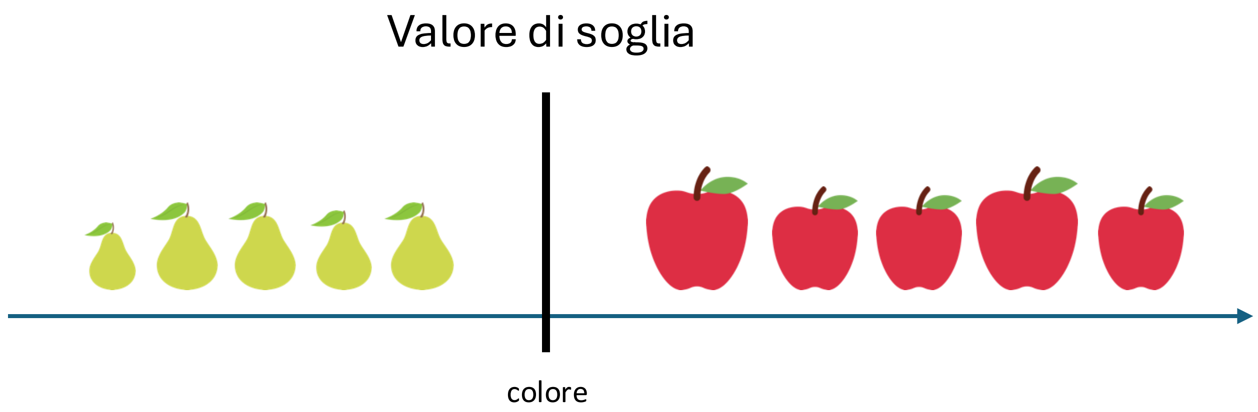
Esperienza: insieme di tutti i dati a disposizione

Task: classificare la frutta

Performance: quanto bene ho separato la frutta?

Cosa definisce una AI

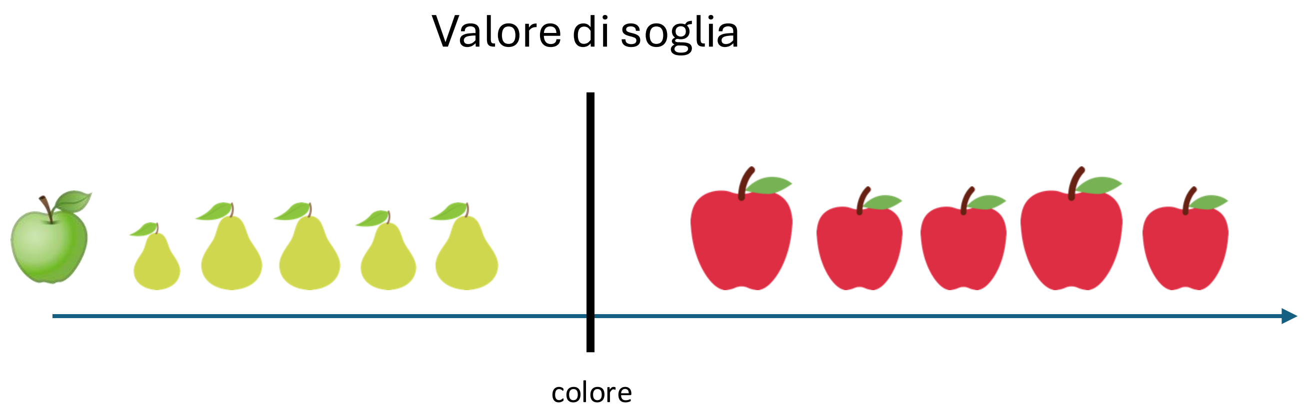
“Un programma informatico si dice che apprenda dall'**esperienza E** rispetto a una classe di **task T** e una misura di **performance P**, se la sua performance nei compiti in T, misurata da P, migliora con l'esperienza E.”



Il mondo reale è complesso

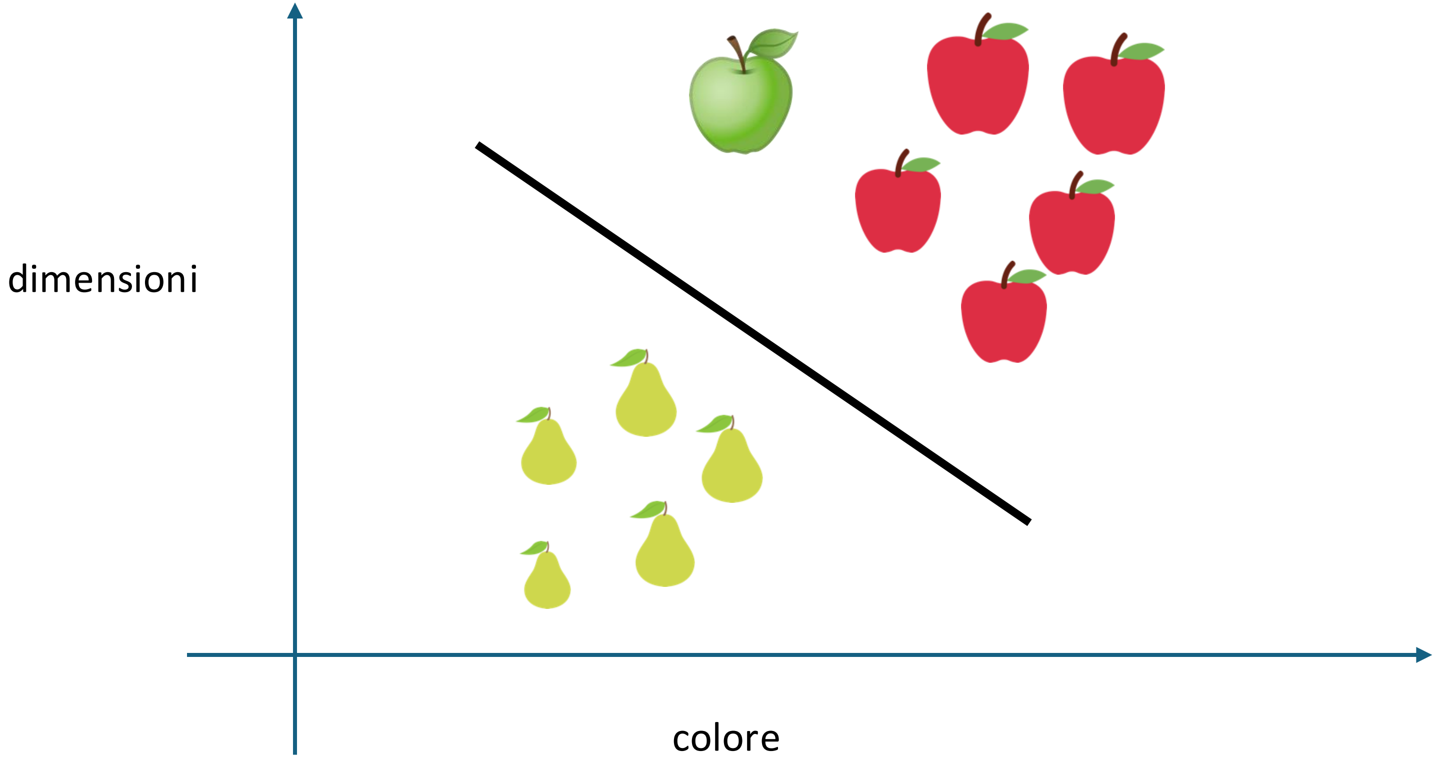
Cosa definisce una AI

"Un programma informatico si dice che apprenda dall'**esperienza E** rispetto a una classe di **task T** e una misura di **performance P**, se la sua performance nei compiti in T, misurata da P, migliora con l'esperienza E."



Il mondo reale è complesso

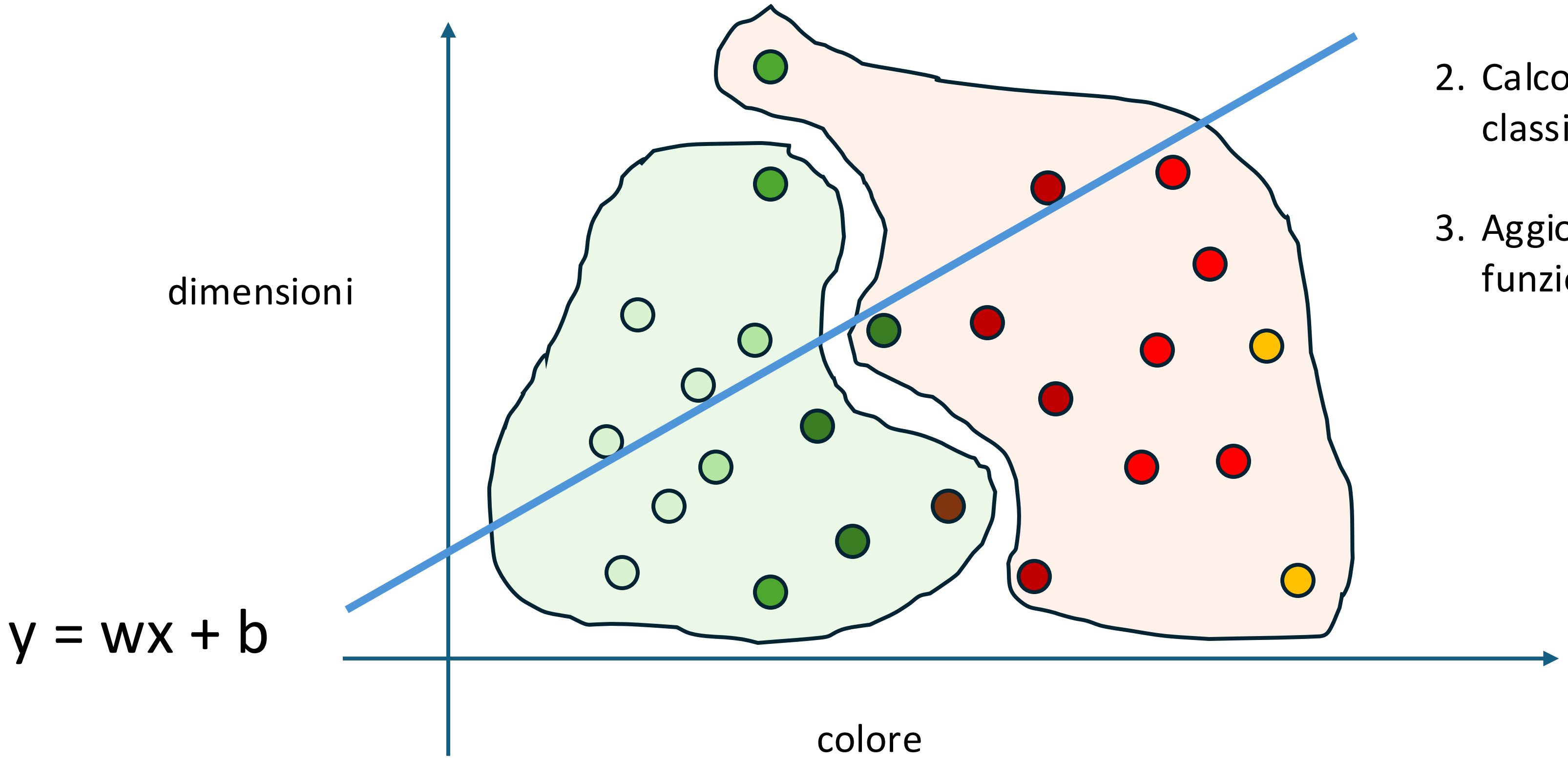
Cosa definisce una AI



Performance: 100%

Cosa «impara» il modello di AI?

- Il mondo reale può avere elementi **non facilmente classificabili**
- Soluzione semplice: Cercare quale sia la retta che **Massimizza** le performance di classificazione

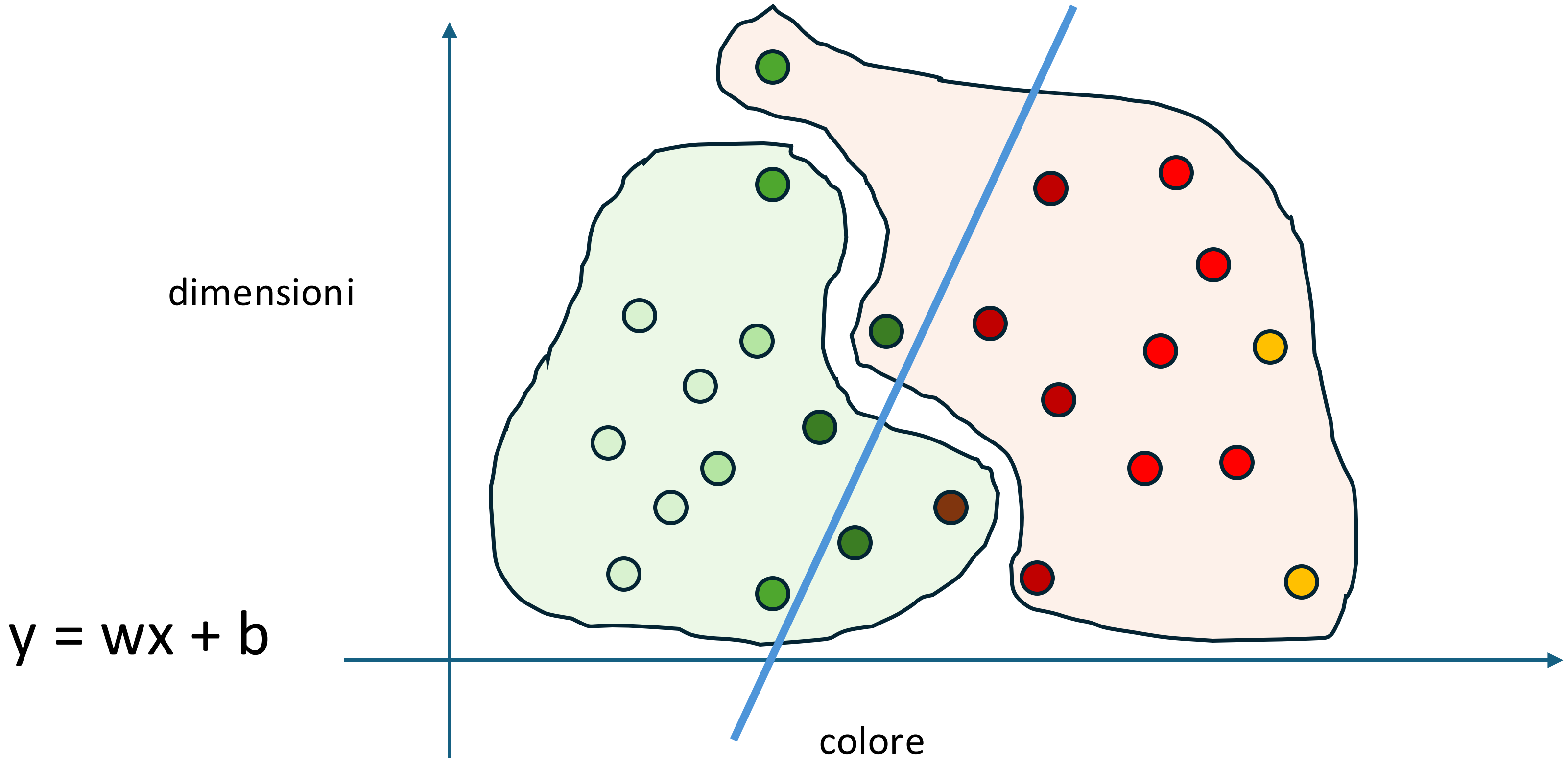


1. Divisione casuale dei data points
2. Calcolo dell'errore (quanti elementi ho classificato male)
3. Aggiornamento dei parametri della funzione

$\boxed{w}$ $\boxed{b}$

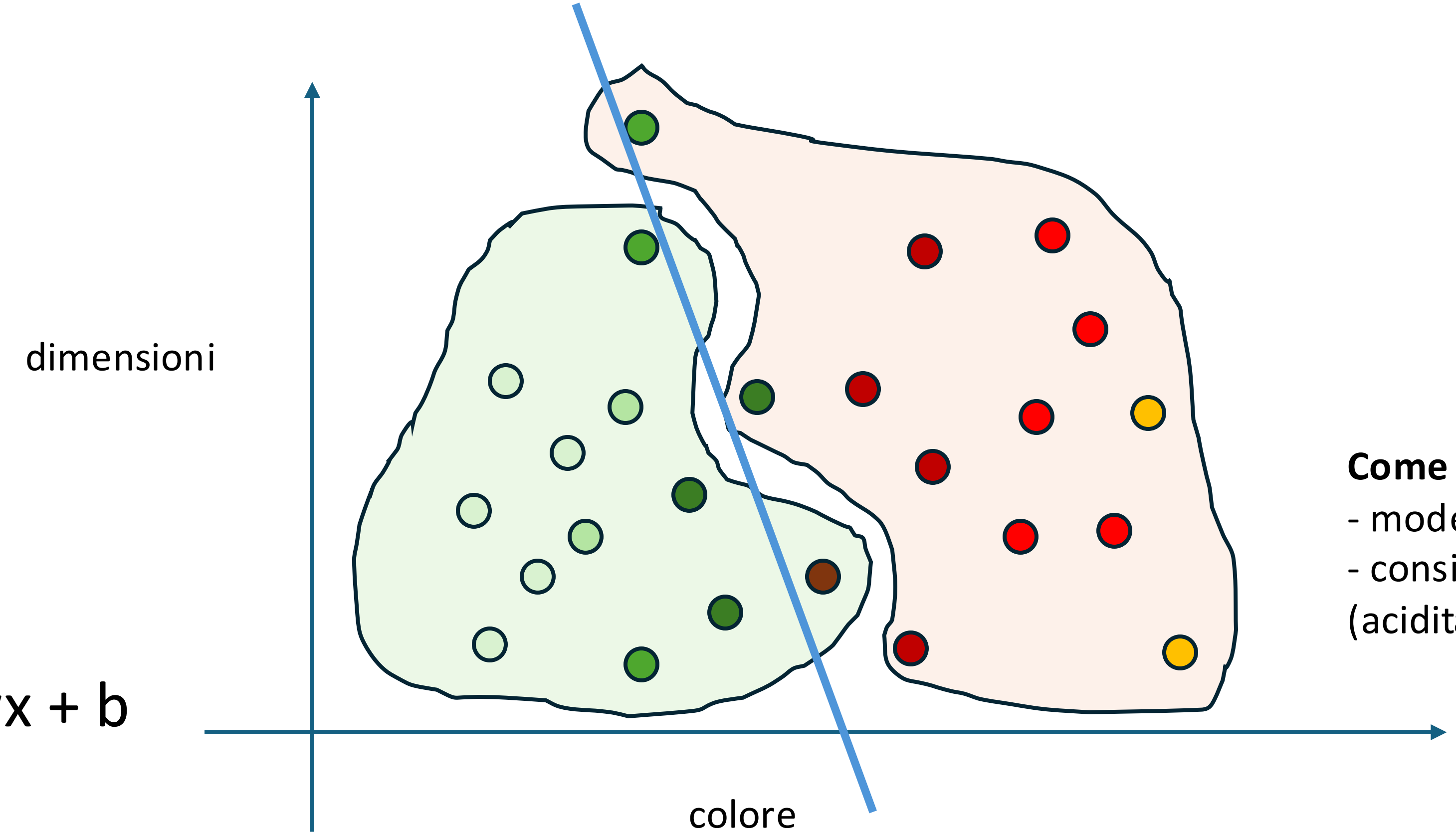
Cosa «impara» il modello di AI?

- Il mondo reale può avere elementi **non facilmente classificabili**
- Soluzione semplice: Cercare quale sia la retta che **Massimizza** le performance di classificazione



Cosa «impara» il modello di AI?

- Il mondo reale può avere elementi **non facilmente classificabili**
- Soluzione semplice: Cercare quale sia la retta che **Massimizza** le performance di classificazione

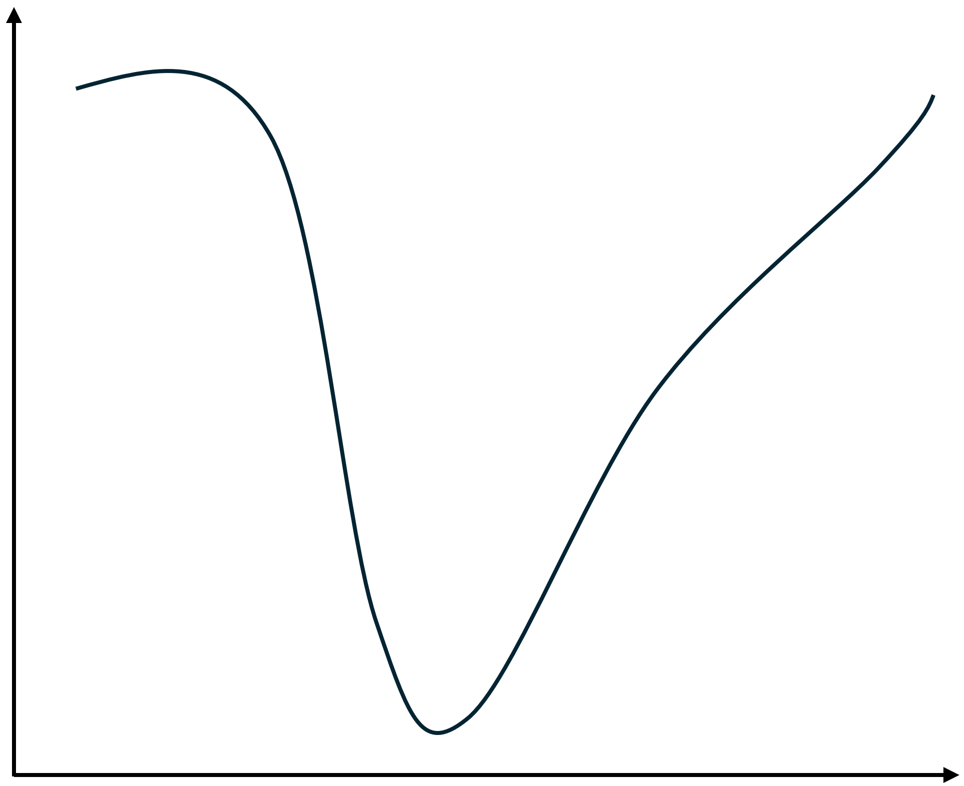


$y = wx + b$

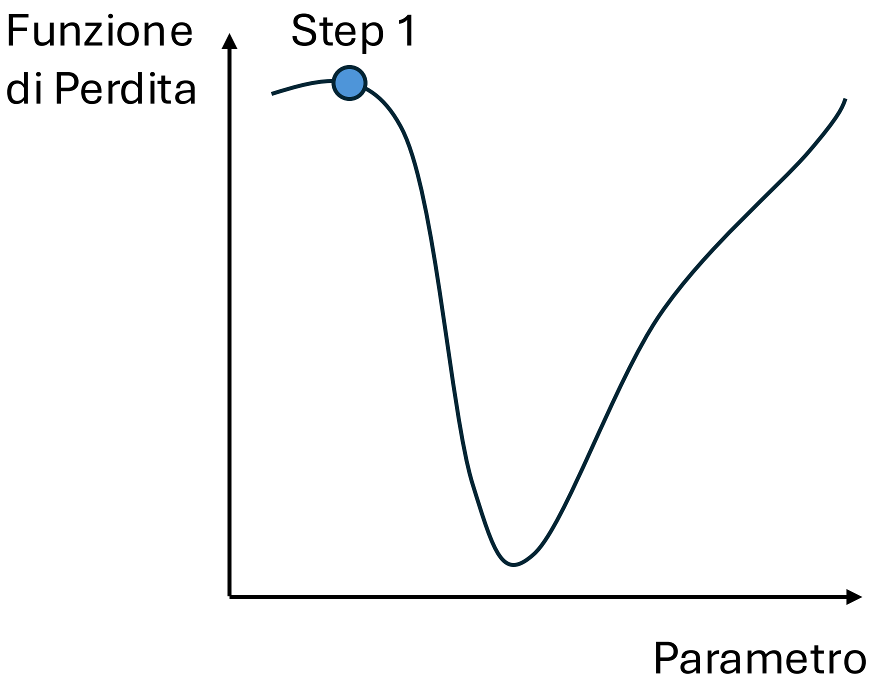
Come migliorare?
- modelli più complessi (non lineari)
- considerare più caratteristiche (acidità, mese di raccolta, ..)

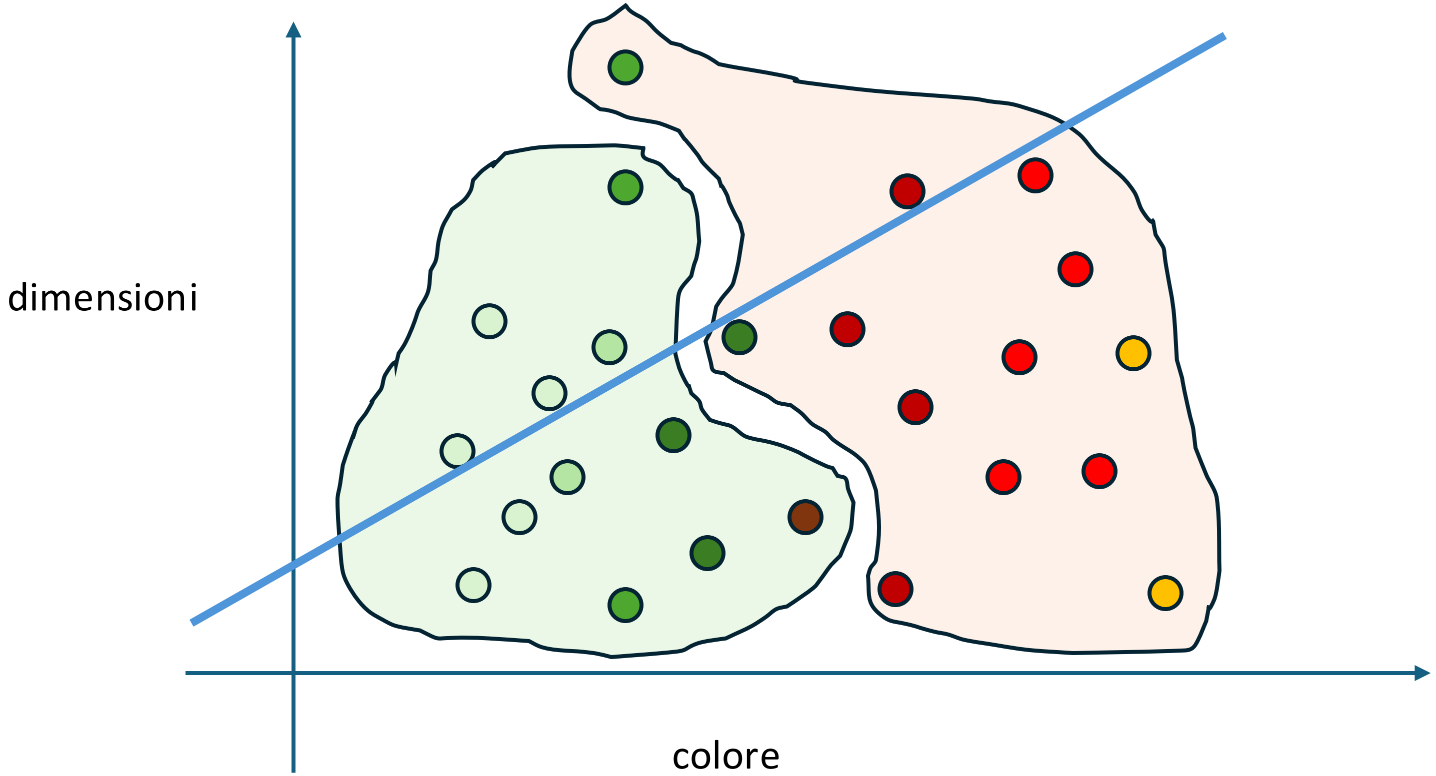
Quanto siamo bravi nel nostro Task T?

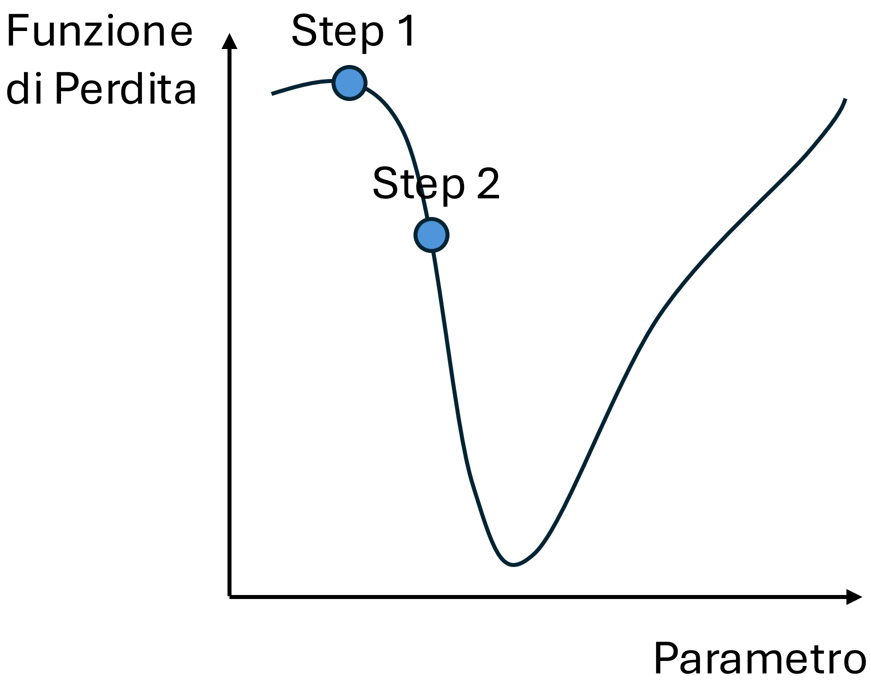
Funzione di
Perdita: L

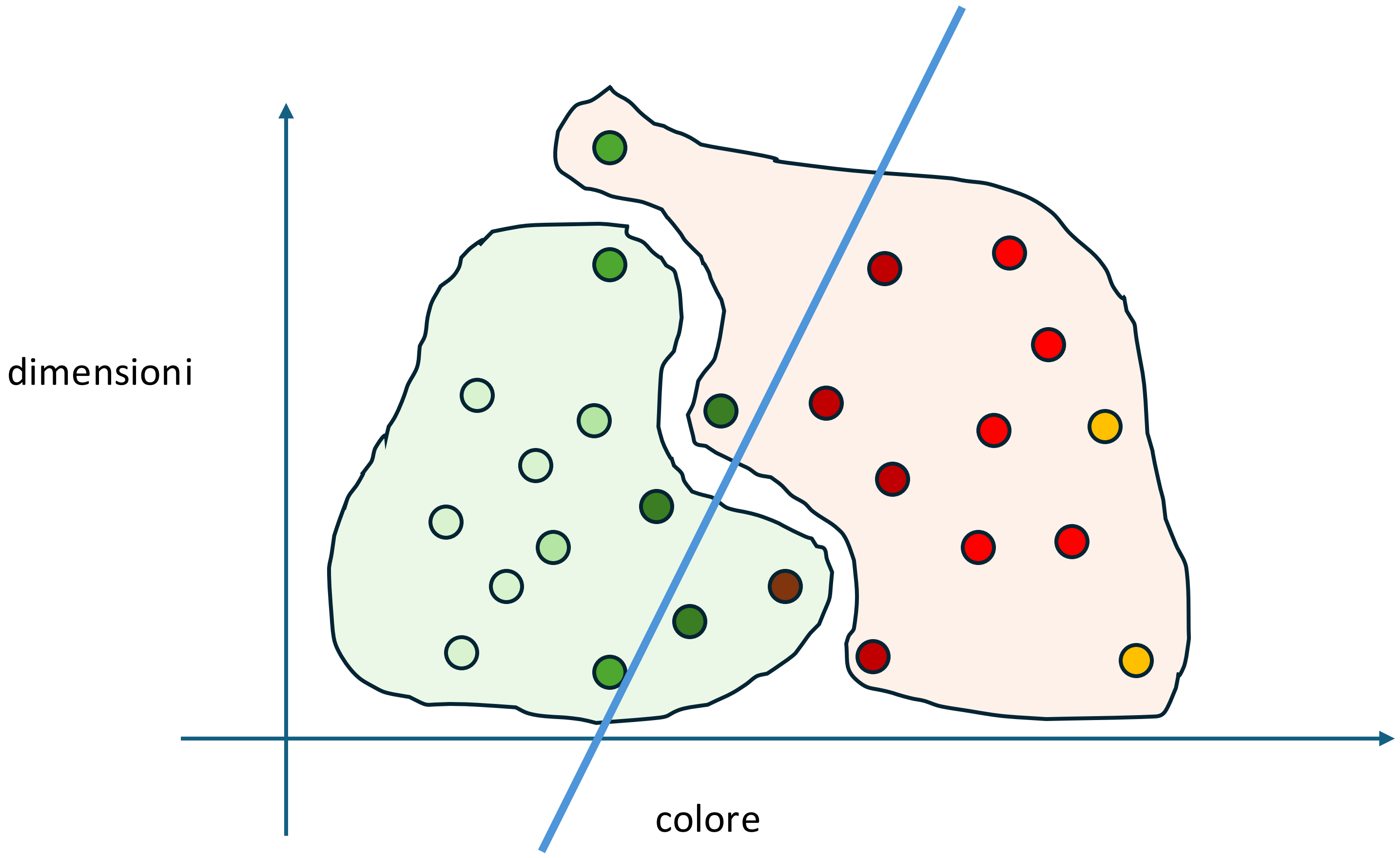


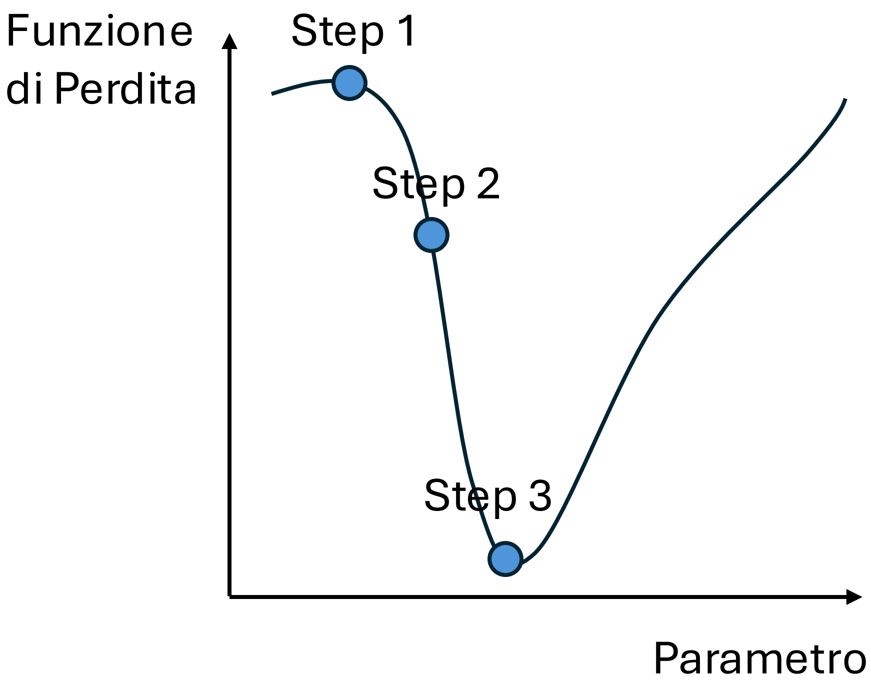
Parametro: w

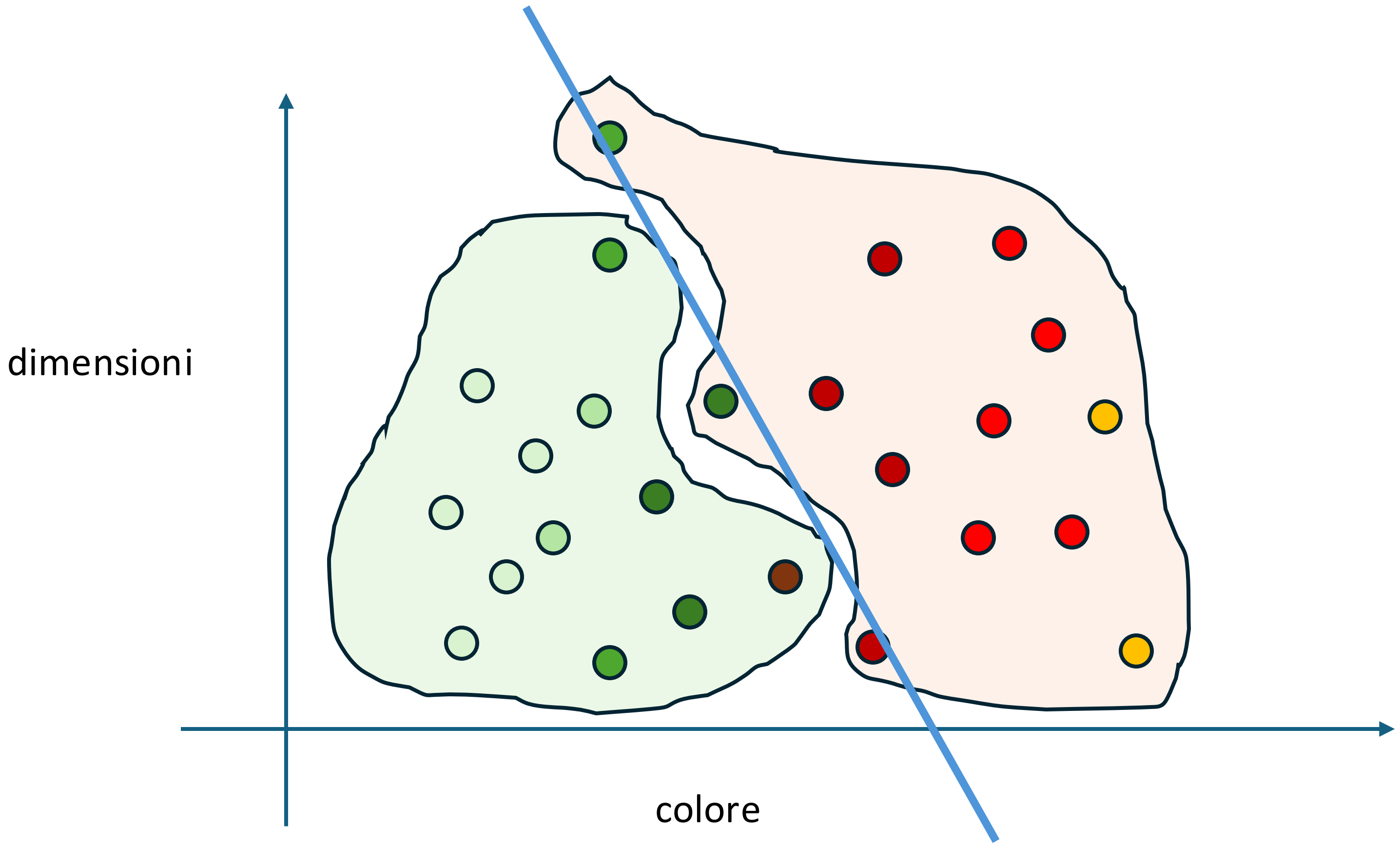

Funzione
di Perdita
Step 1
Parametro


dimensioni
colore


Funzione
di Perdita
Step 1
Step 2
Parametro


dimensioni
colore

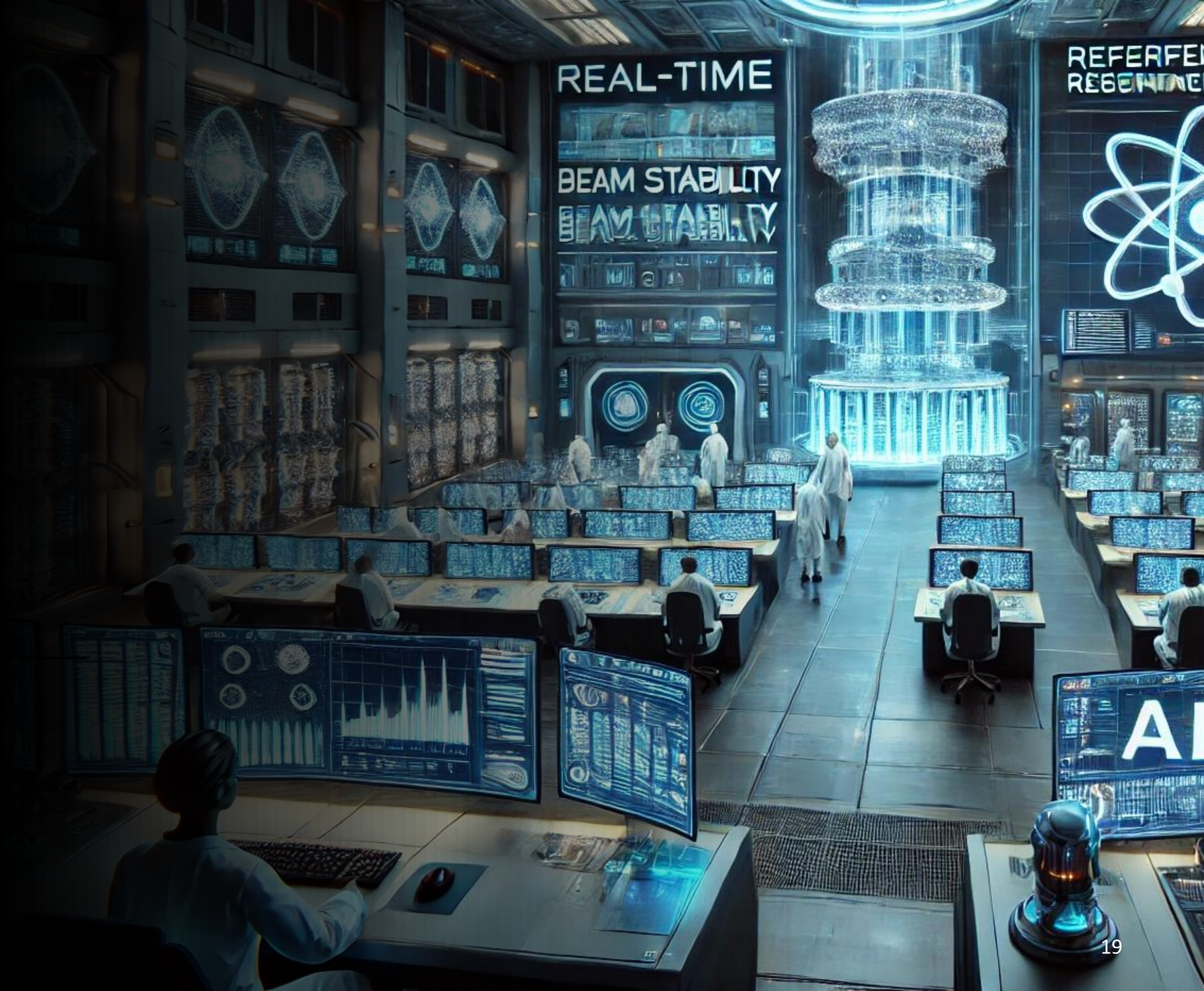

Funzione
di Perdita
Step 1
Step 2
Step 3
Parametro


dimensioni
colore

Come possiamo usare l’IA nella fisica?

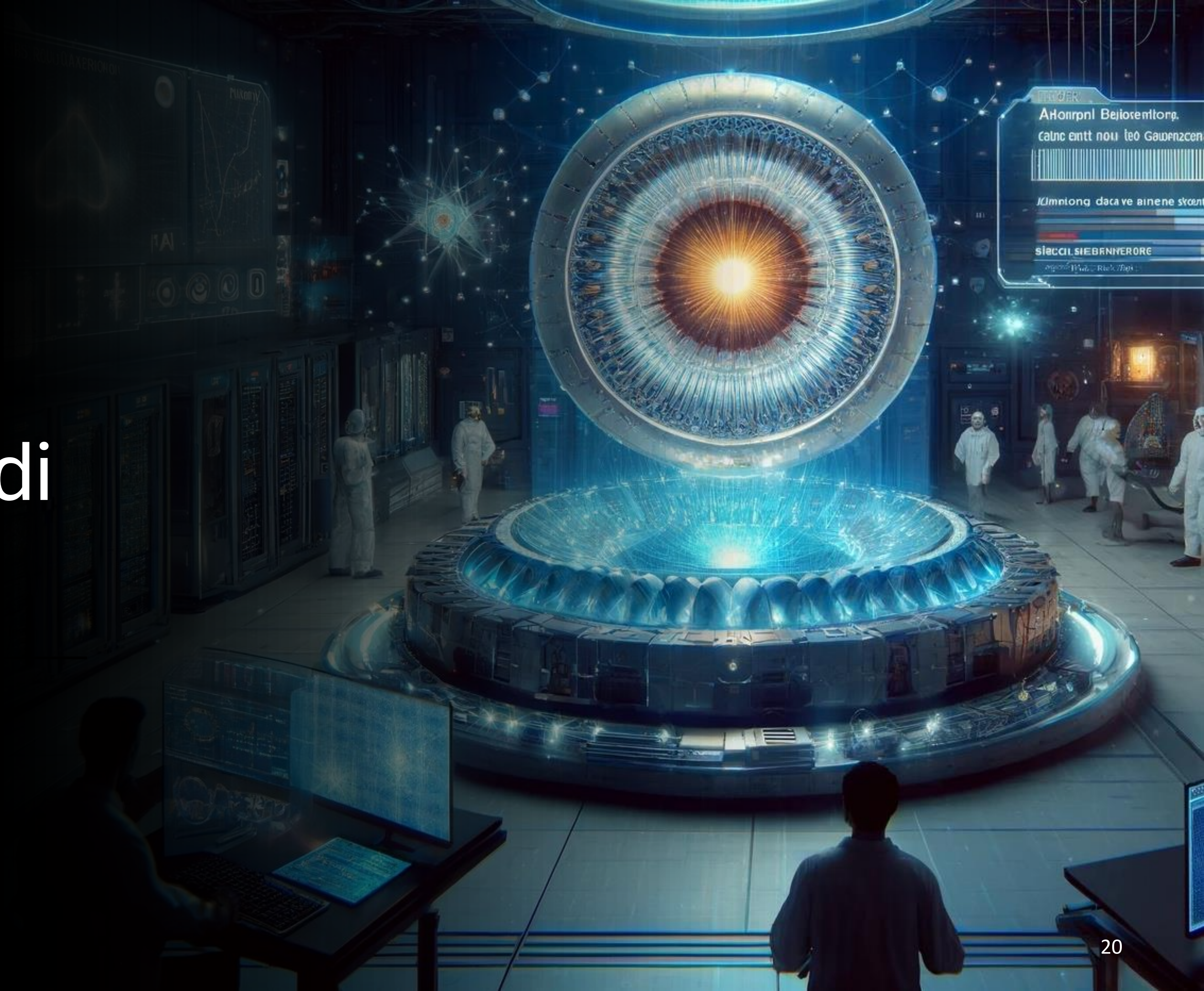
Immagine di una Sala di Controllo

Sistemi di Controllo



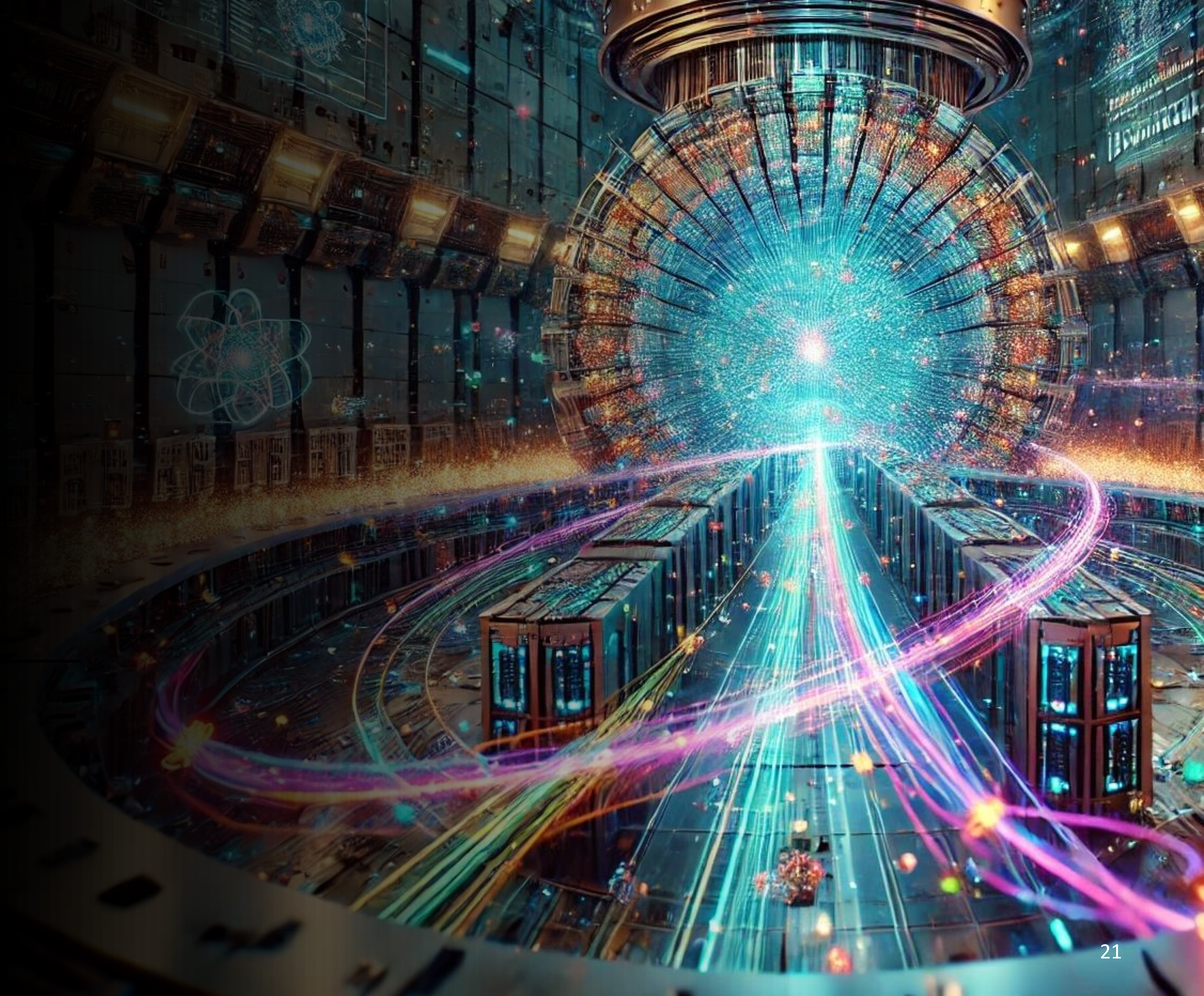
* Immagine generata con AI

Rilevamento di Eventi Rari



* Immagine generata con AI

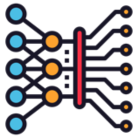
Tracciamento delle Particelle



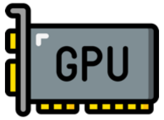
* Immagine generata con AI

Motivazioni

Nonostante la varietà di approcci e modelli teorici testati negli esperimenti fisici, ciò che li accomuna è l'enorme quantità di dati complessi che producono

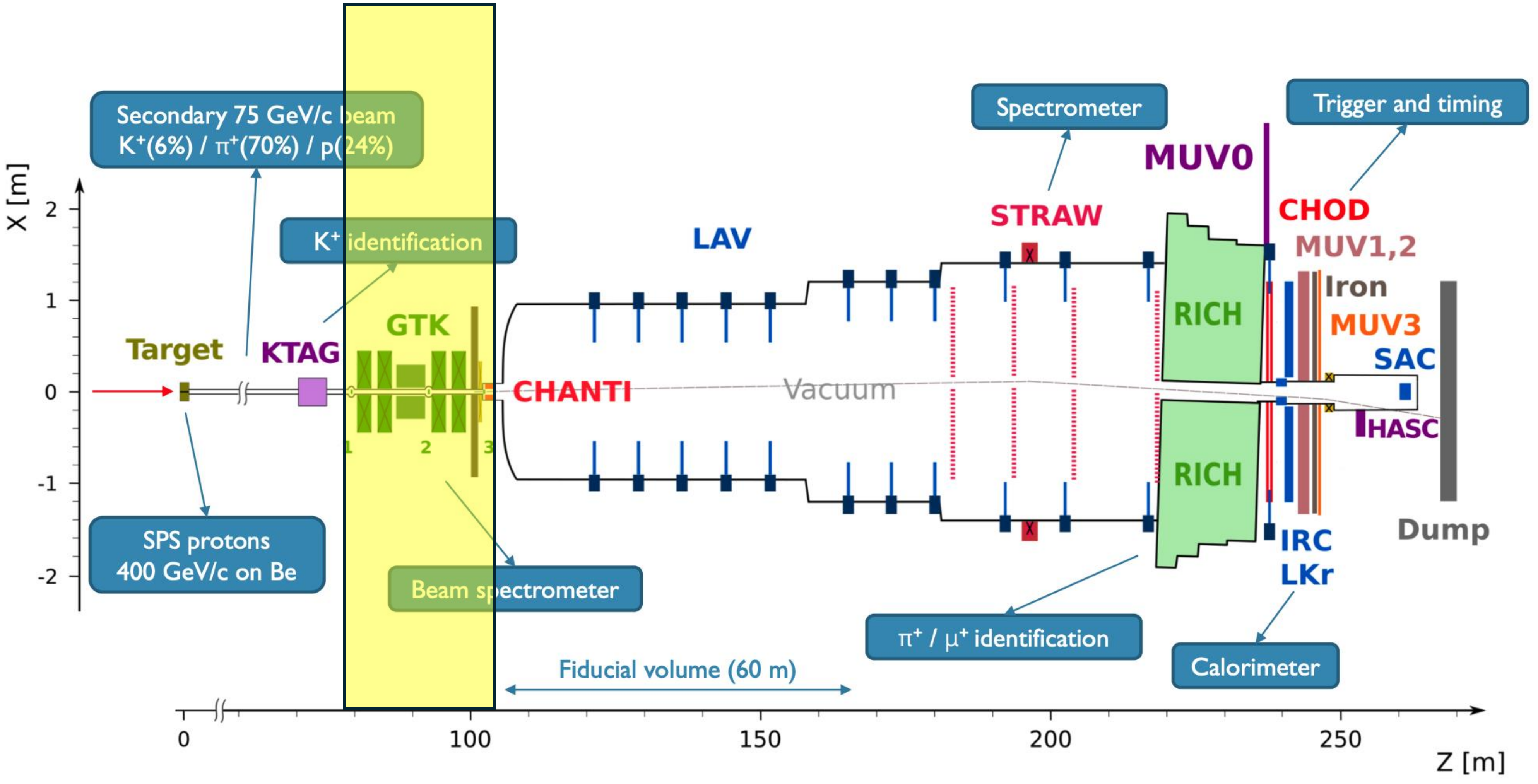


Questa sfida legata ai dati richiede potenti metodi computazionali, come il Machine Learning e il Deep Learning

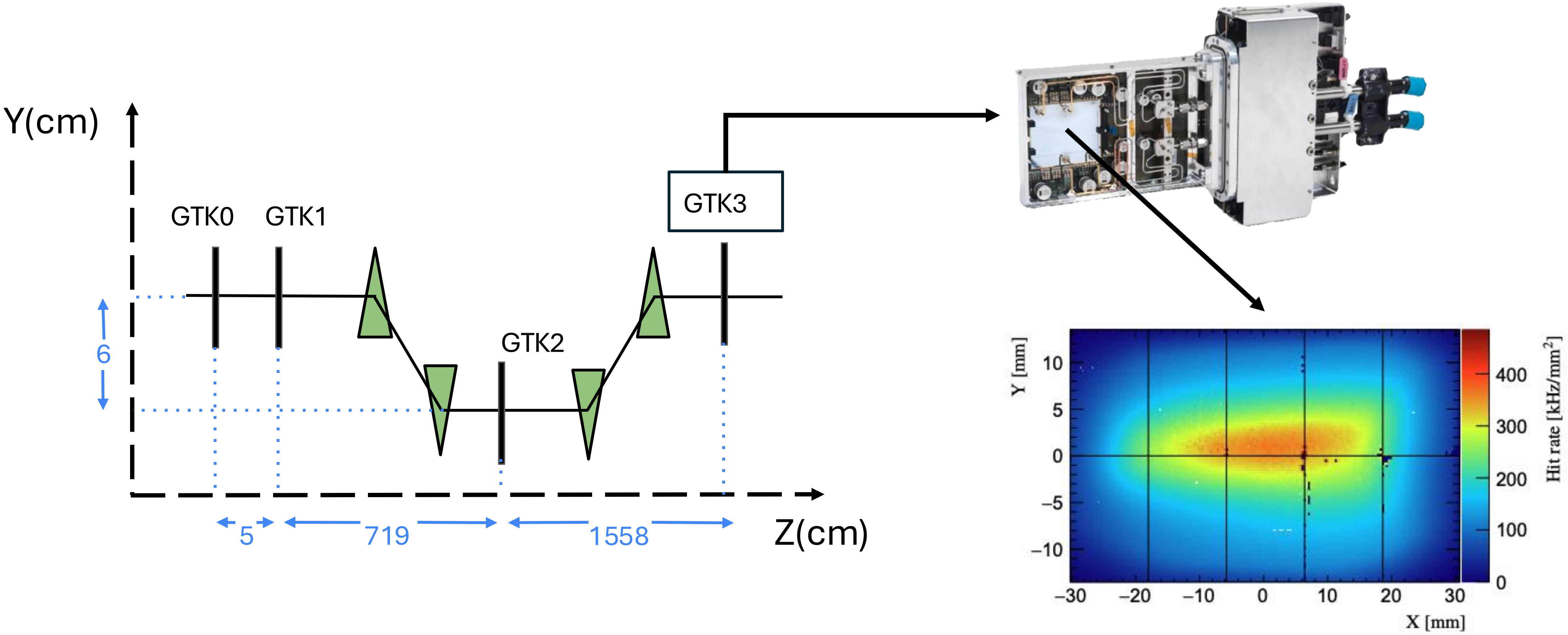


La parallelizzazione tramite GPU (Graphical Processing Unit) e la loro capacità di memoria consentono di ridurre drasticamente i tempi di calcolo.

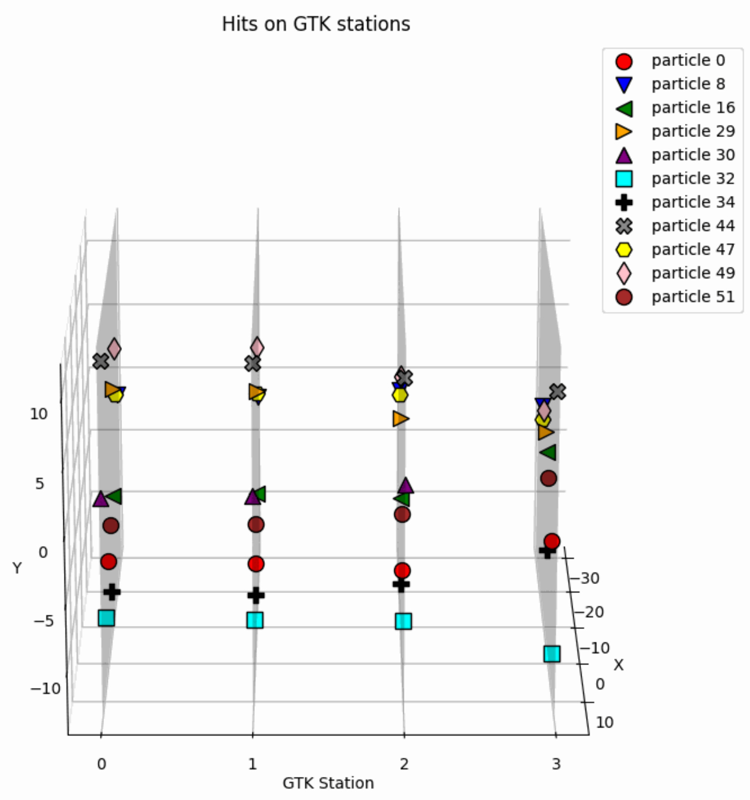
Tracciamento in NA62



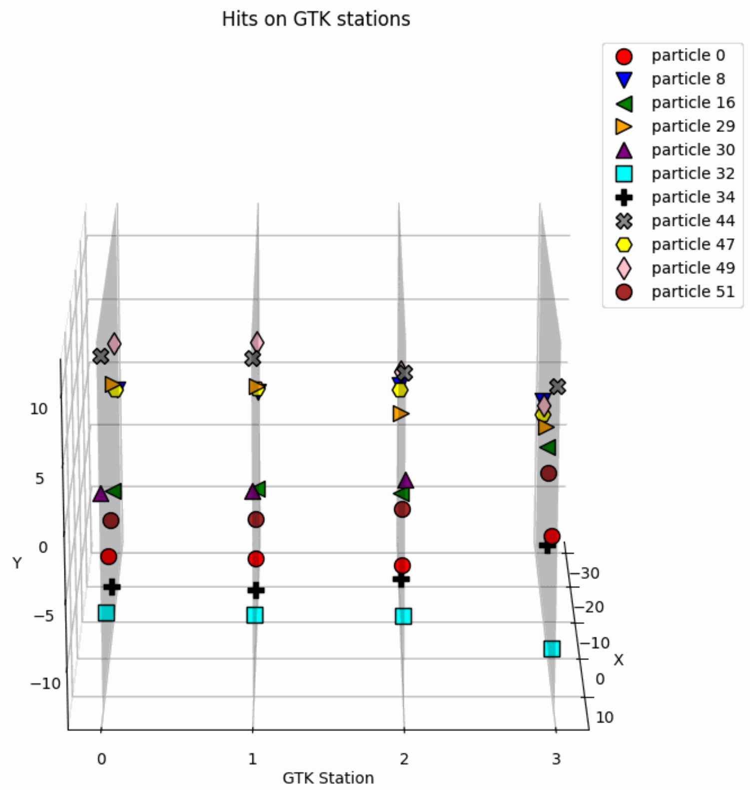
Giga Tracker

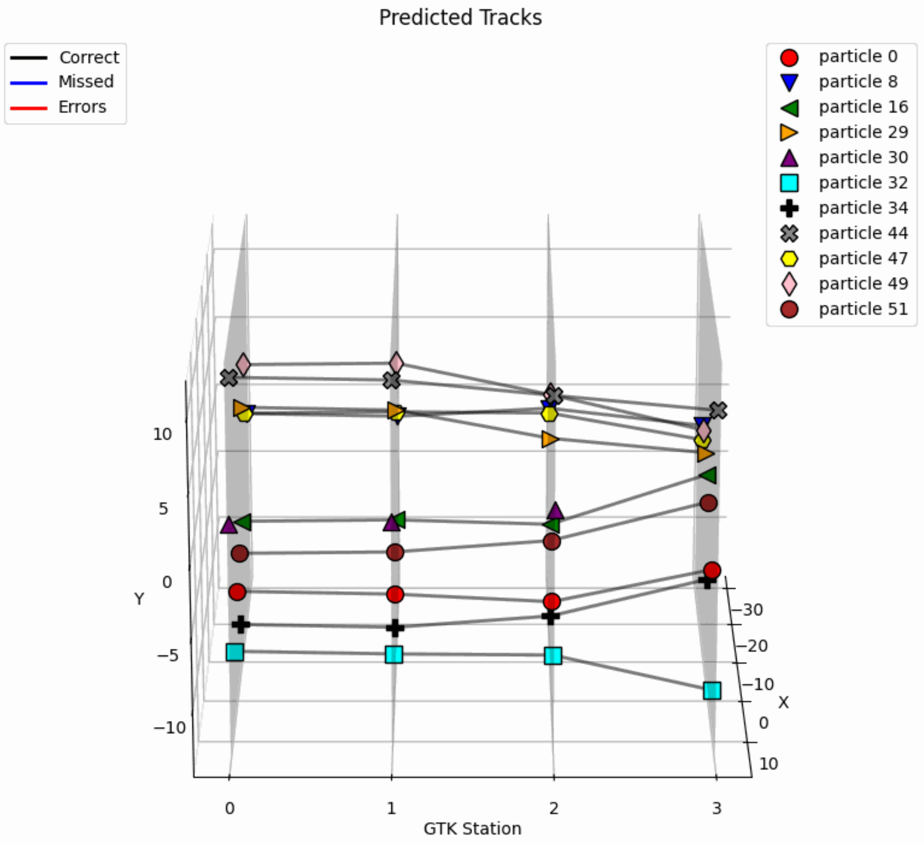


Particle Tracking

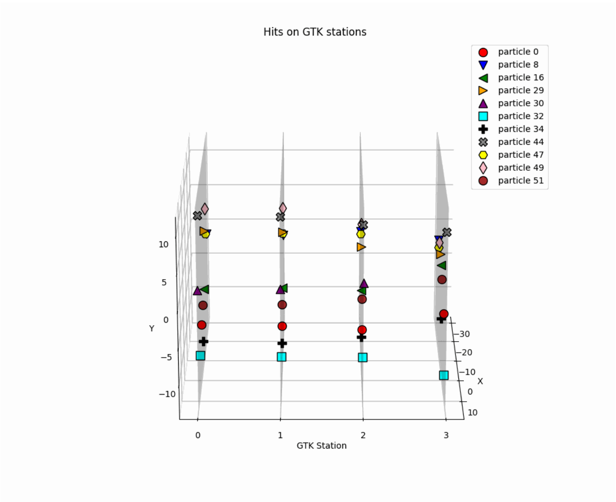


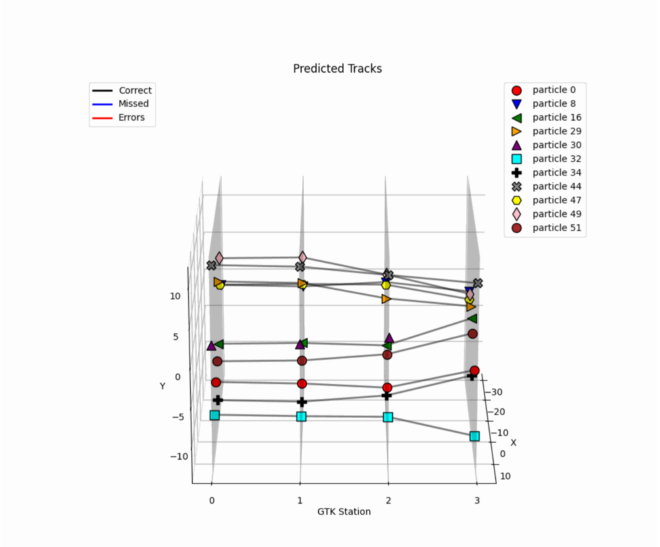
Particle Tracking



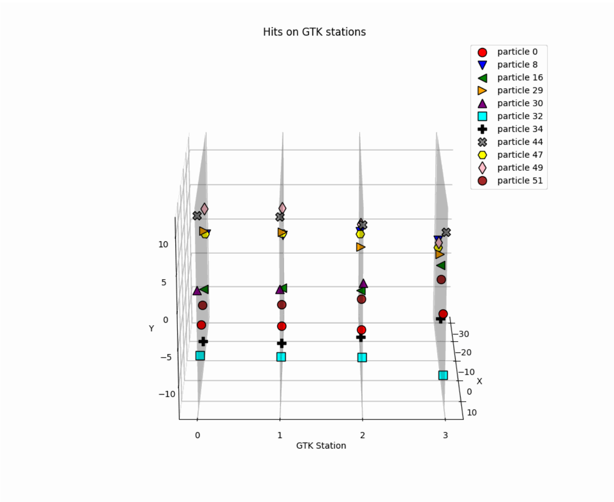


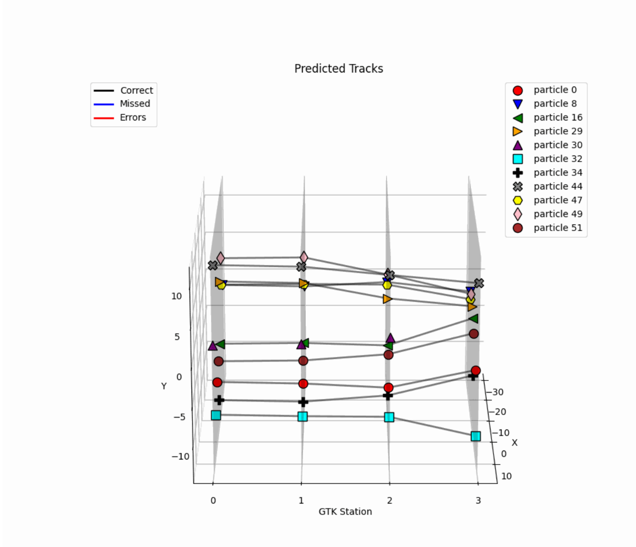
Particle Tracking



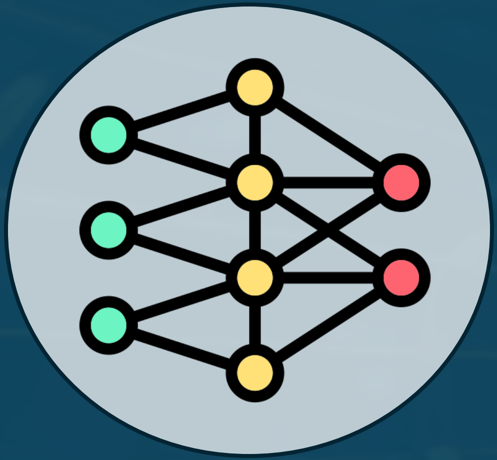


Particle Tracking



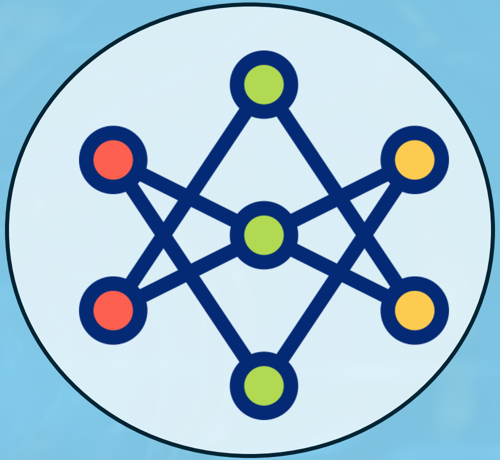


Proposal



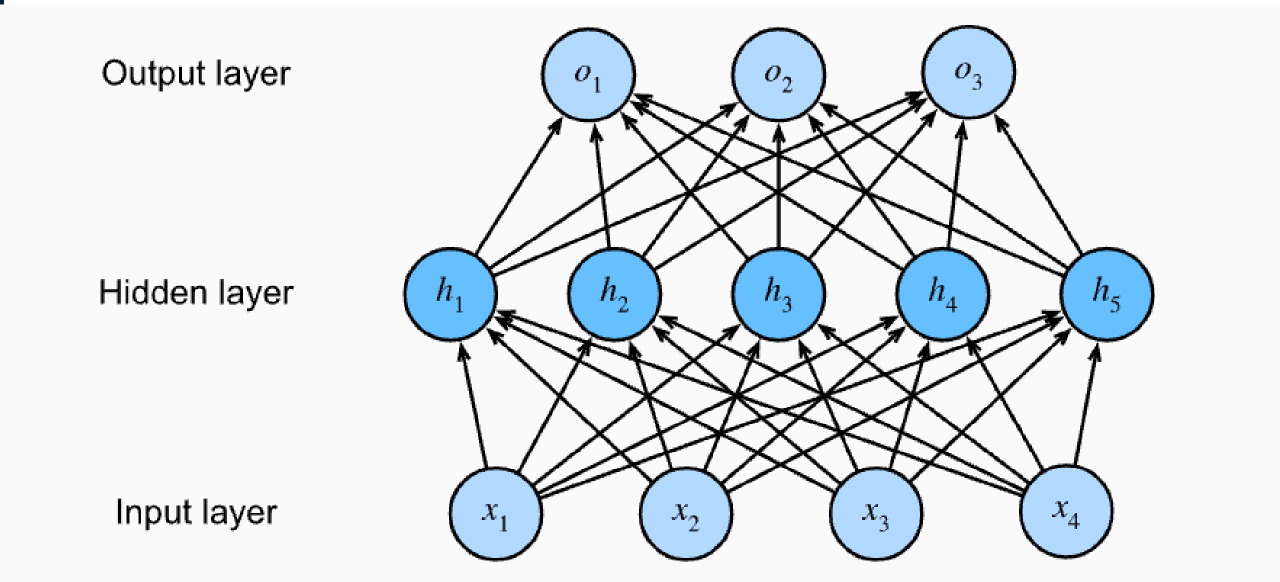
Multi-Layer Perceptron

Transformer



Graph Neural Network

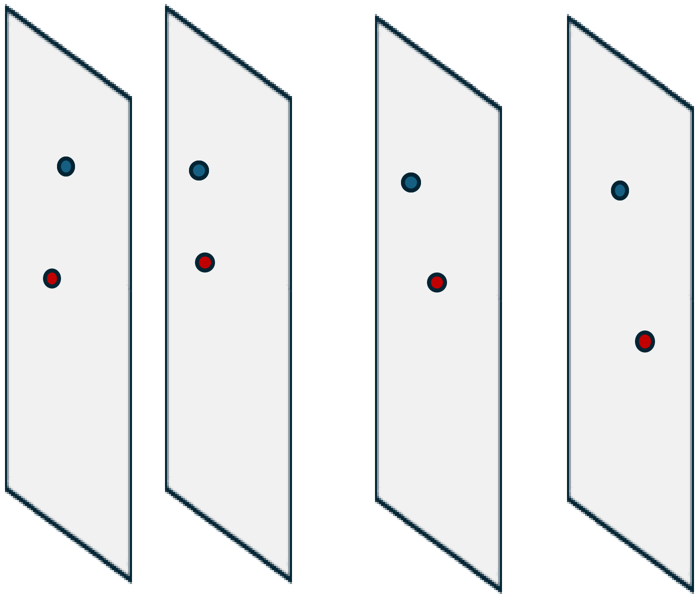
Multi-Layer Perceptron



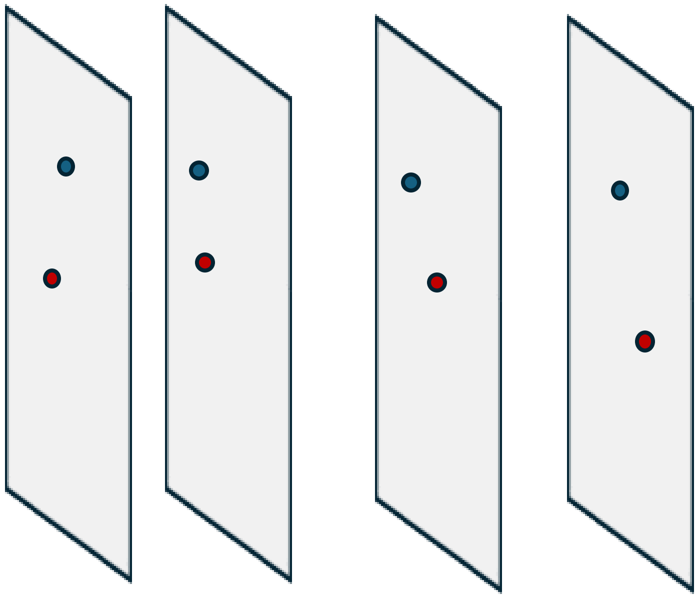
Il Multi-Layer Perceptron è un'architettura composta da strati componibili e differenziabili che ottimizza i suoi pesi (W, b) mediante Back-Propagation per minimizzare una funzione di perdita L

* Image from Dive into Deep Learning, chapter 5

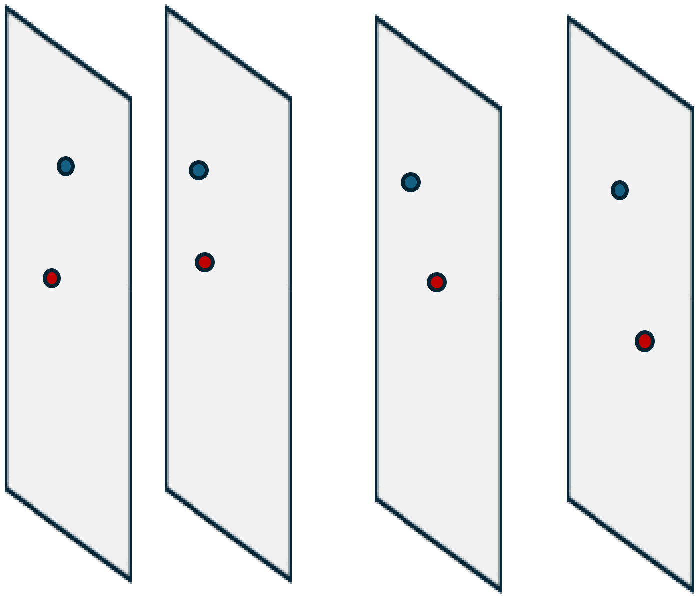
Multi-Layer Perceptron



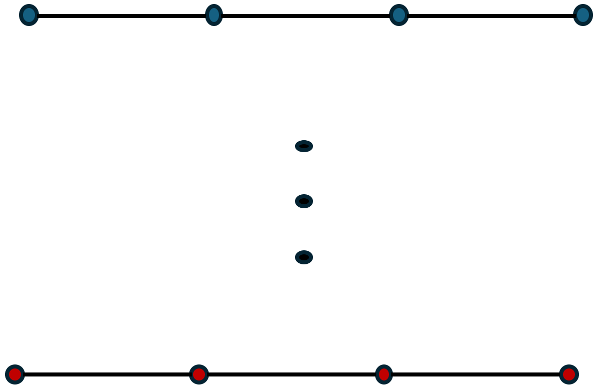
Multi-Layer Perceptron



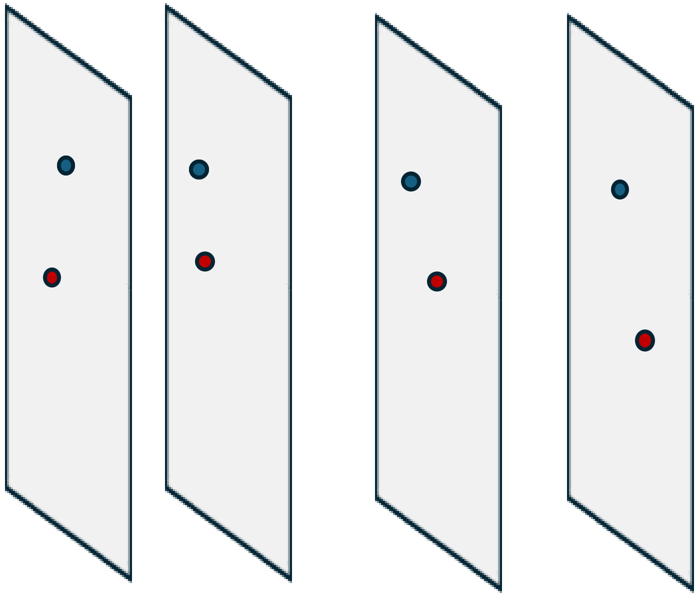
Multi-Layer Perceptron



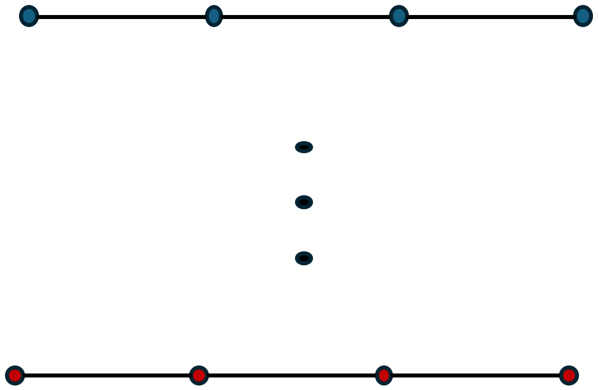
Tracce Ammissibili



Multi-Layer Perceptron

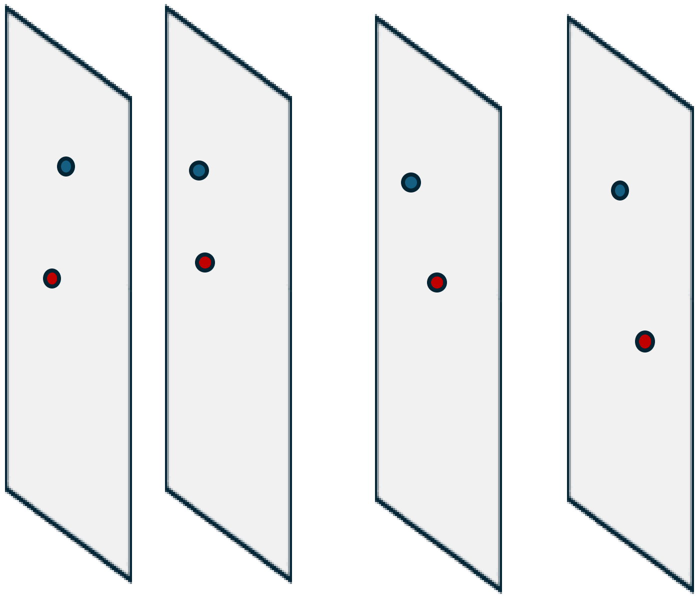


Tracce Ammissibili

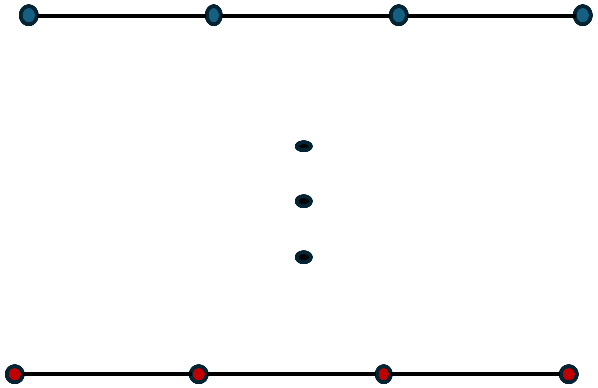


Tracce Non Ammissibili

Multi-Layer Perceptron

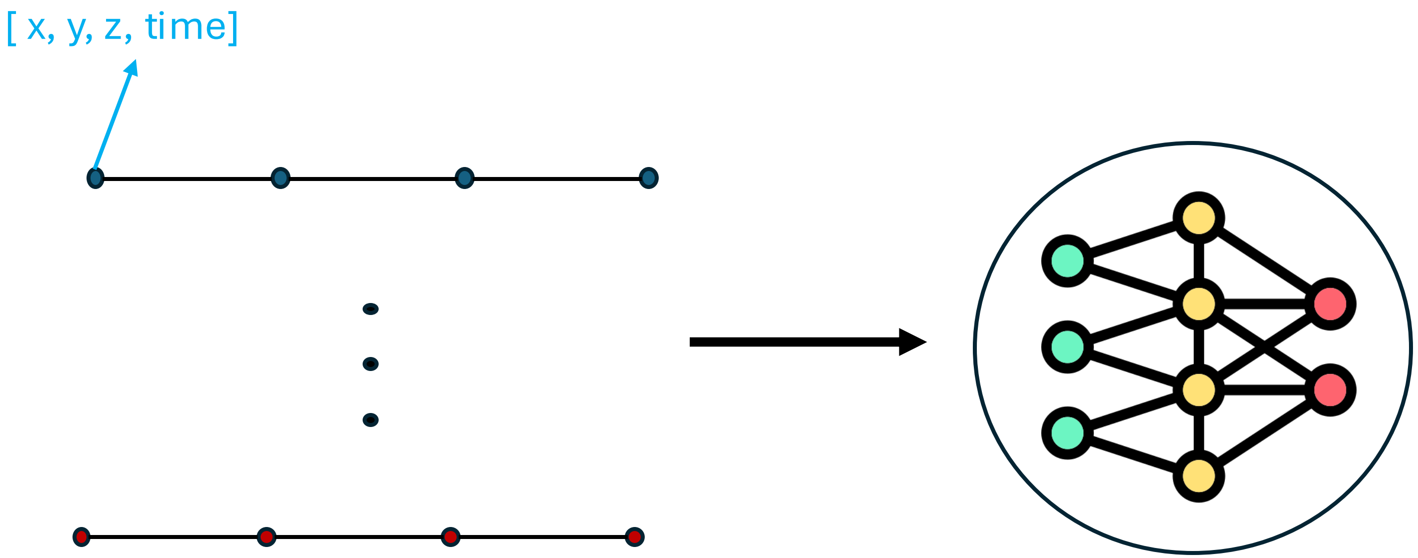


Tracce Ammissibili

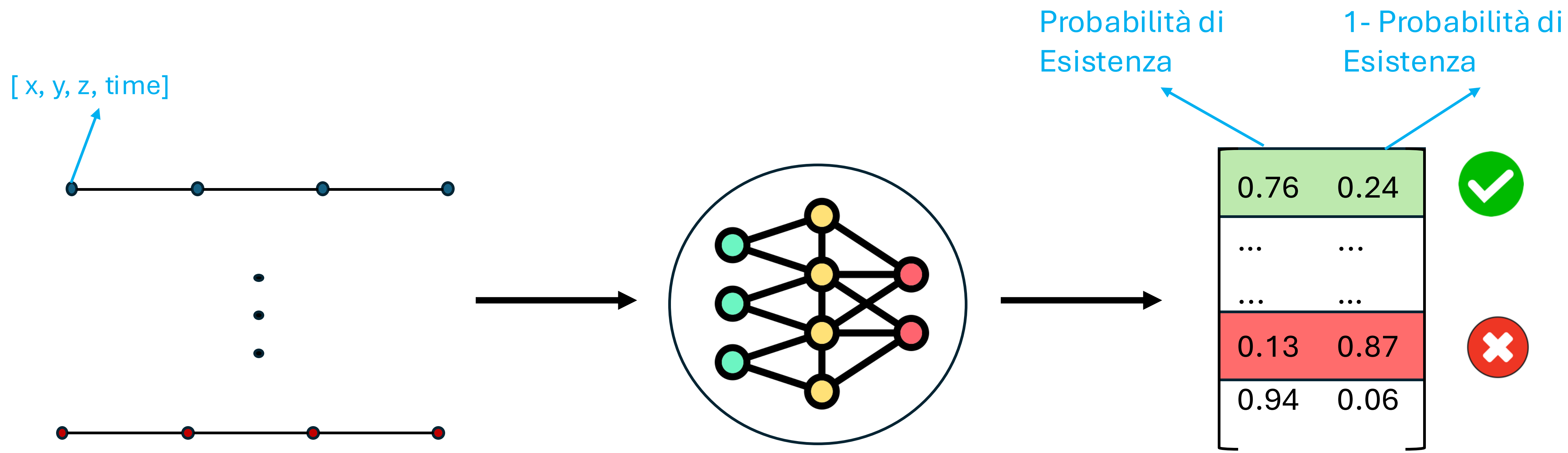


Tracce Non Ammissibili

Multi-Layer Perceptron



Multi-Layer Perceptron



Multi-Layer Perceptron: results

Efficienza $= \dfrac{\#\,Tracce\ Predette\ Correttamente}{\#\,Totale\ Tracce\ Vere}$

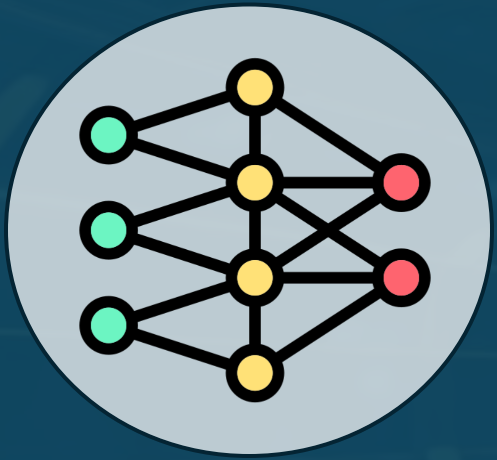
Purezza $= \dfrac{\#\,Tracce\ Predette\ Correttamente}{\#\,Totale\ Tracce\ Predette}$

Tracce False $= \dfrac{\#\,Tracce\ Predette\ con\ Errore}{\#\,Totale\ Tracce\ Vere}$

Efficienza	Purezza	Tracce False
70.06 %	92.3 %	39.45 %

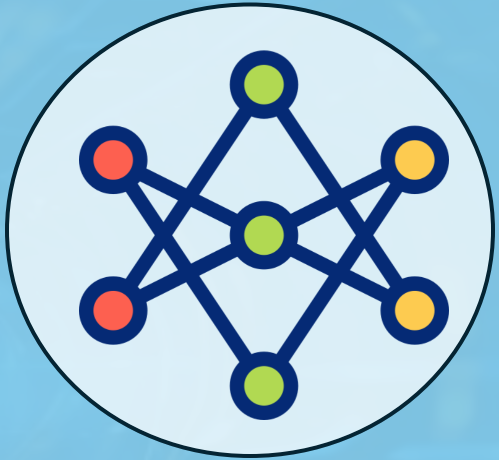
Multi-Layer Perceptron:

- Conoscenza Locale
- Le performance non sono fantastiche
- Ci serve molta memoria per processare tutte le combinazioni
- Diventiamo Lenti
- **Non c’è bilanciamento fra il numero di tracce vere e false**



Multi-Layer Perceptron

Transformer



Graph Neural Network

Transformer

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Presentati nel 2017 e ora ampiamente adottati grazie alle loro incredibili prestazioni e alla parallelizzazione.

Hanno rappresentato una rivoluzione nell'Elaborazione del Linguaggio Naturale, nella Computer Vision e nell'Apprendimento Multimodale.

Transformer

Transformer

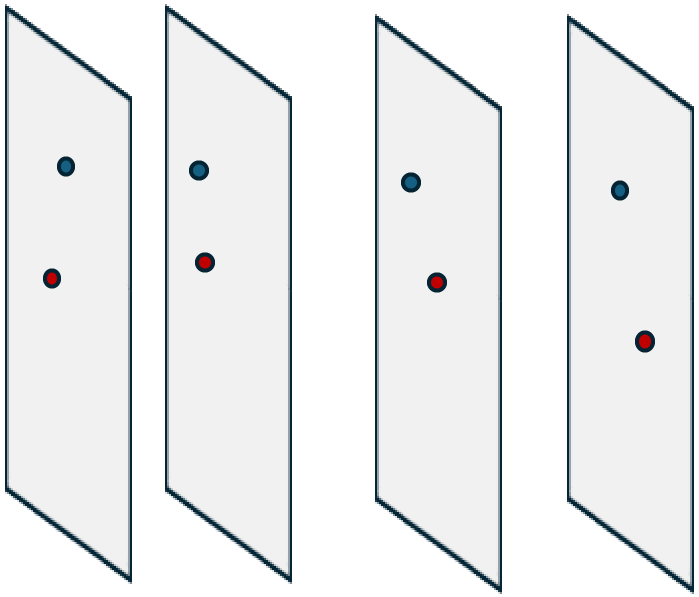
Conoscenza Locale

Molte Combinazioni

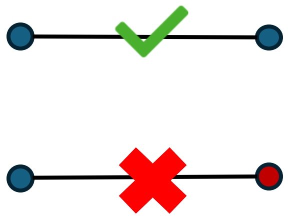
Conoscenza Globale

Cosa succede se invece di considerare la tracce intera consideriamo le connessioni fra hit

Transformer

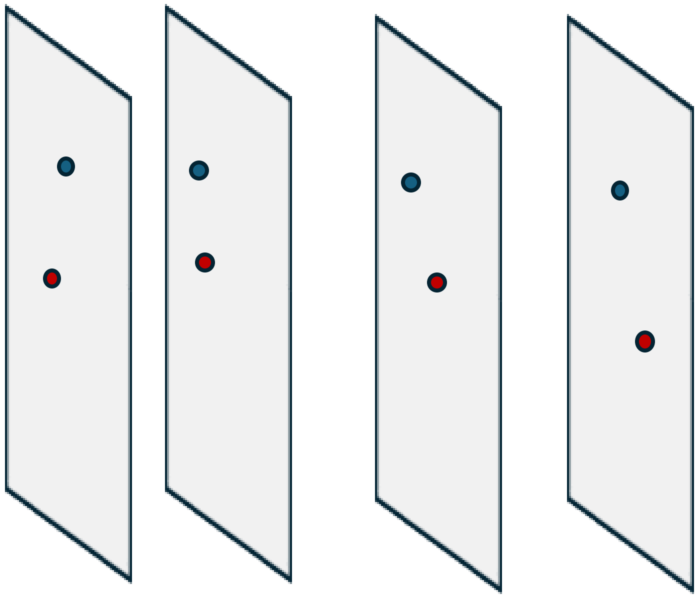


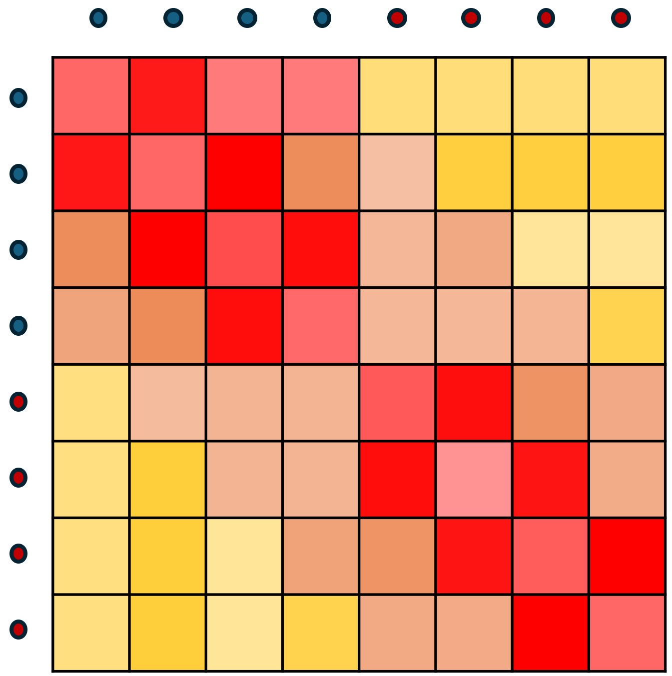
Classificazione binaria: esiste o non esiste



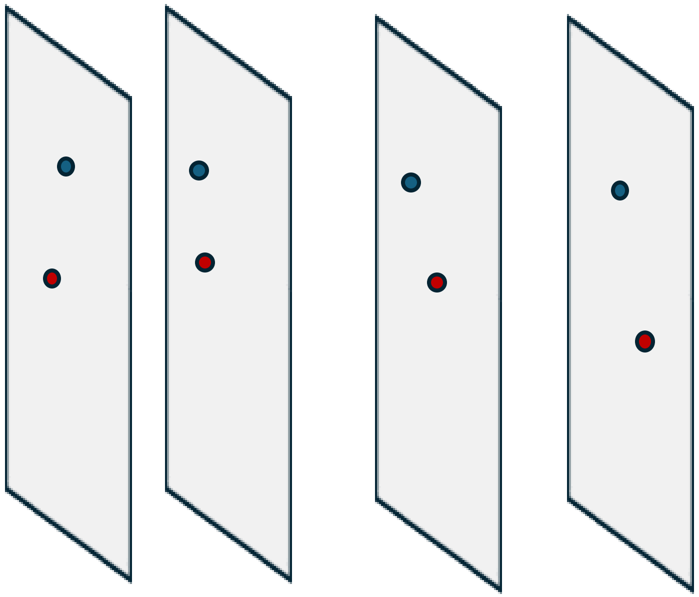
Riduciamo di molto il numero di candidati da considerare

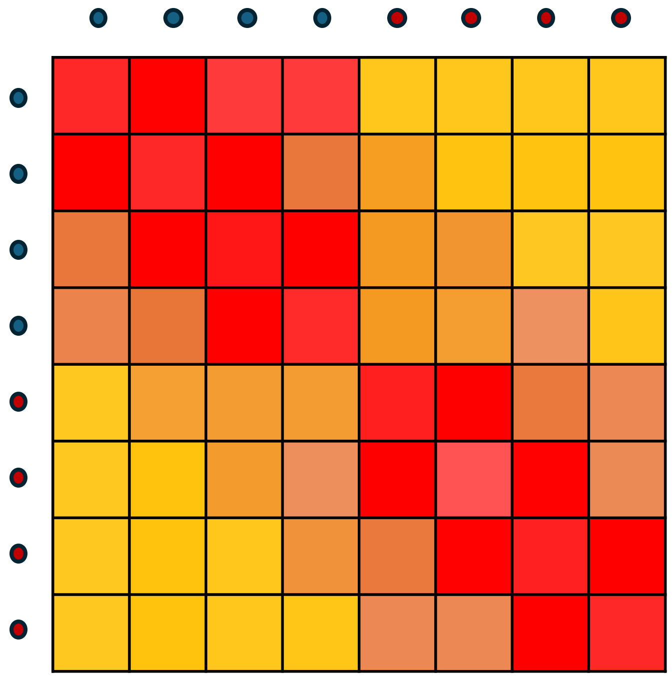
Transformer



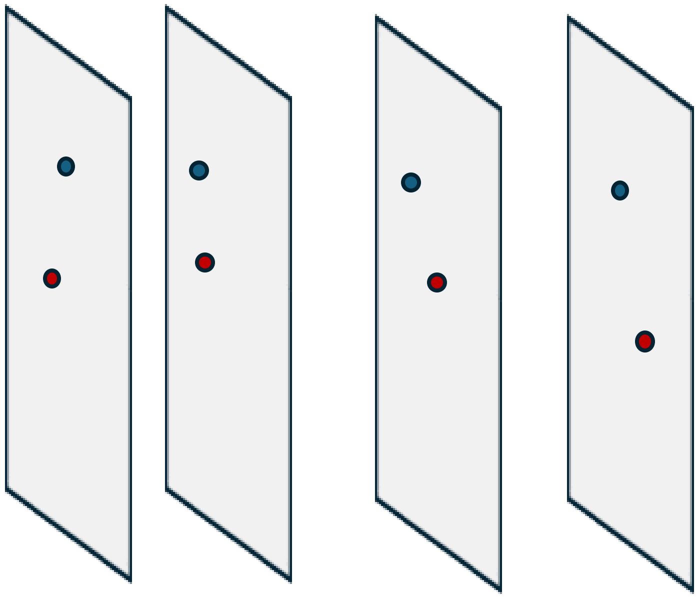


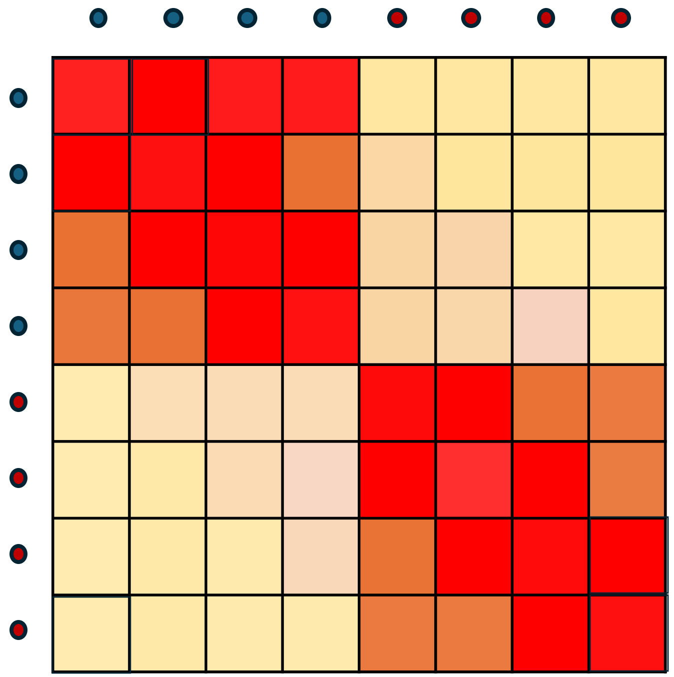
Transformer



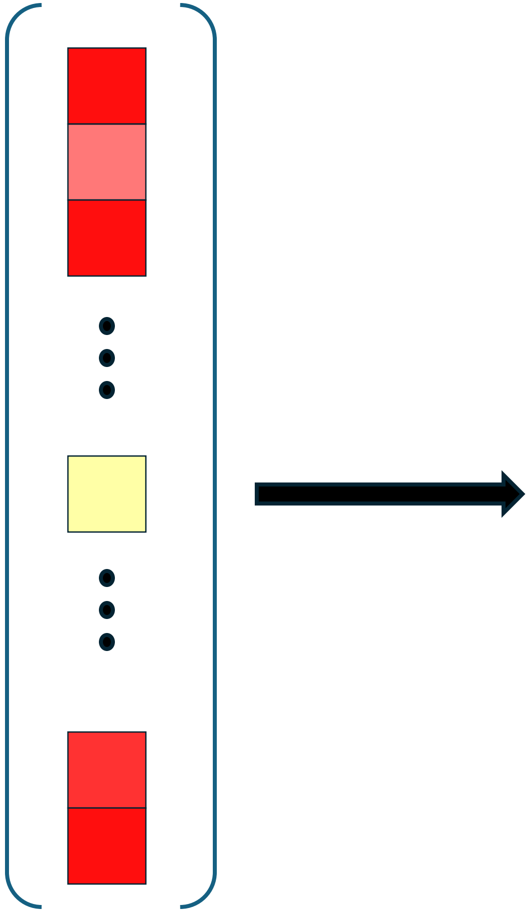


Transformer



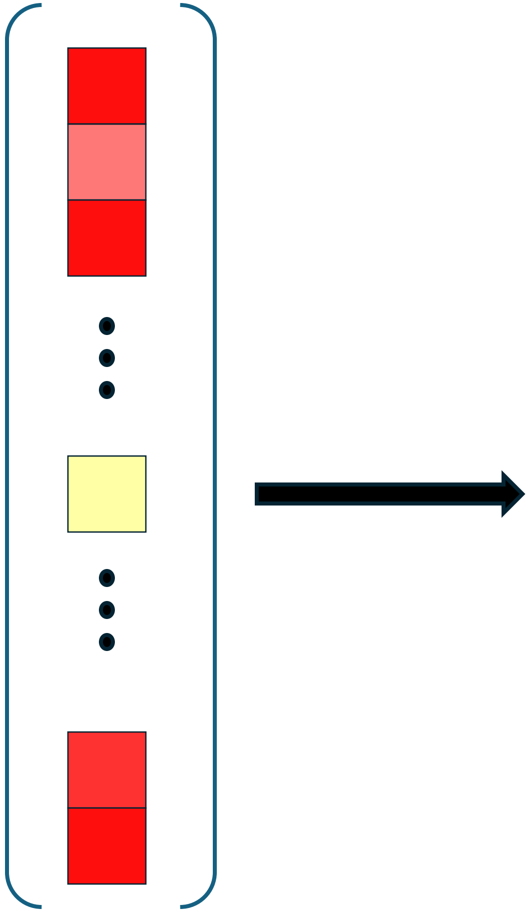


Transformer



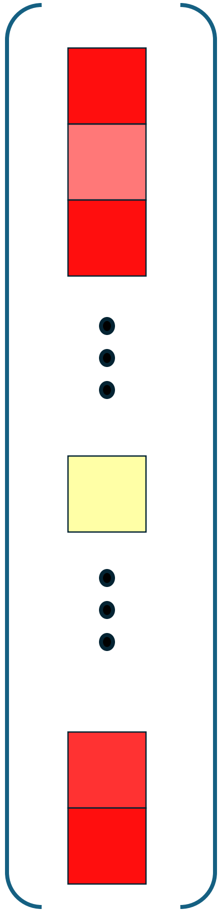
0.99	0.01
...	...
...	...
...	...
0.12	0.88
...	...
...	...
0.99	0.01

Transformer

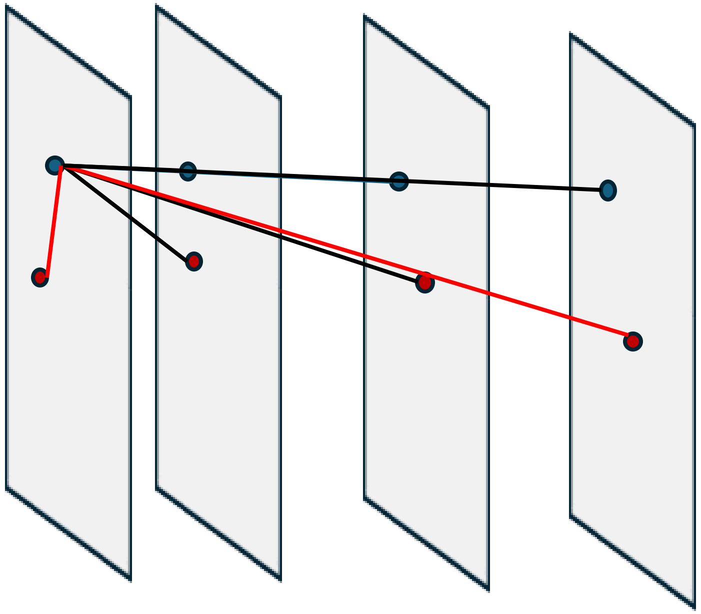


0.99	0.01
...	...
...	...
...	...
0.12	0.88
...	...
...	...
0.99	0.01

Transformer



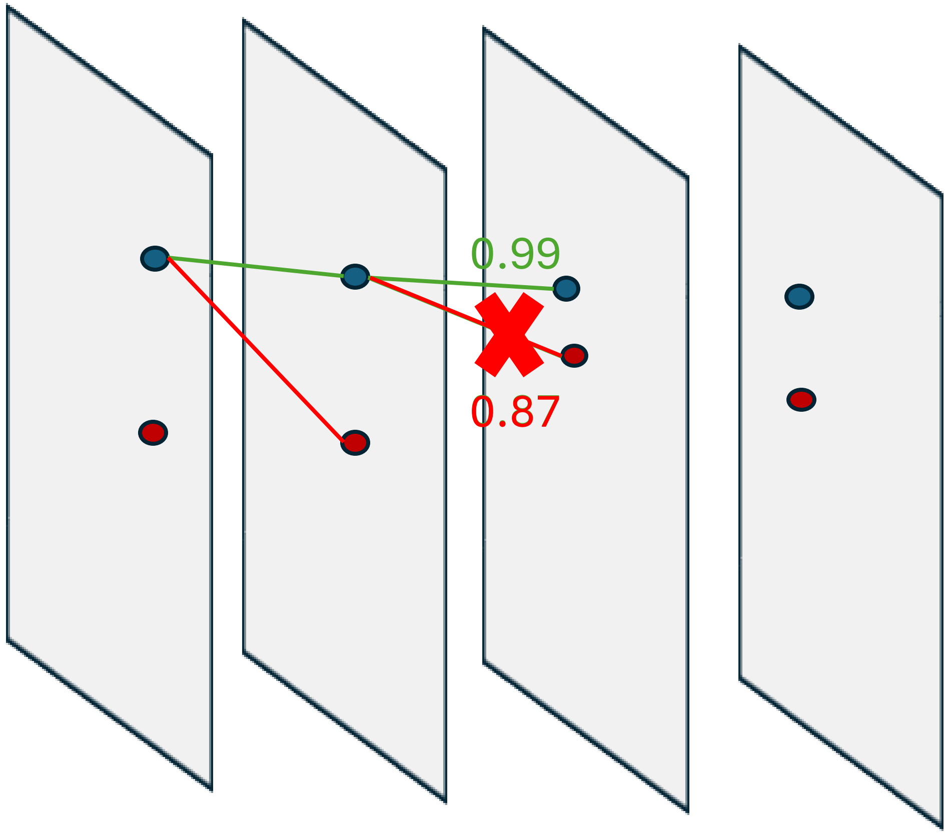
0.99	0.01
…	…
…	…
…	…
0.12	0.88
…	…
…	…
0.99	0.01



Transformer

Selezione degli archi: quale prendiamo per costruire una sola tracce per hit?

0.99	0.01
…	…
…	…
…	…
0.12	0.88
…	…
…	…
0.99	0.01



Transformer

Efficienza = $$\frac{\#\,Tracce\ Predette\ Correttamente}{\#\,Totale\ Tracce\ Vere}$$

Purezza = $$\frac{\#\,Tracce\ Predette\ Correttamente}{\#\,Totale\ Tracce\ Predette}$$

Tracce False = $$\frac{\#\,Tracce\ Predette\ con\ Errore}{\#\,Totale\ Tracce\ Vere}$$

Multi-Layer Perceptron

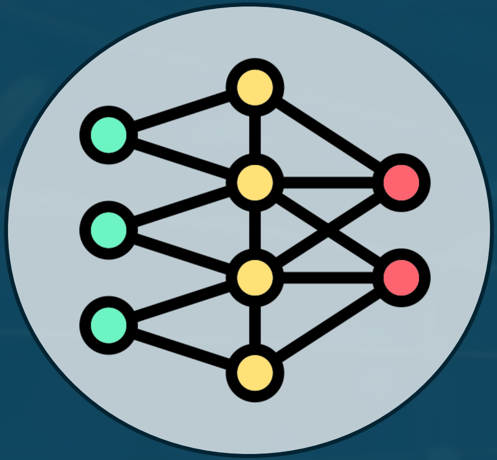
Efficienza	Purezza	Tracce False
70.06 %	92.3 %	39.45 %

Transformer

Efficienza	Purezza	Tracce False
95.95 %	98.62 %	1.22 %

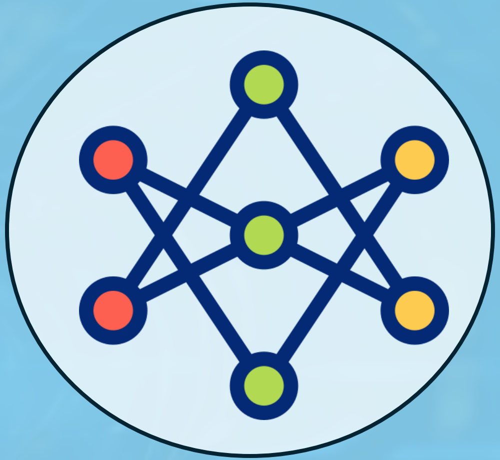
Transformer

- Risolto il problema di avere tanti esempi di tracce false
- Conoscenza globale
- Molto efficienti con i calcoli
- Dobbiamo considerare tutte le connessioni fra le hit



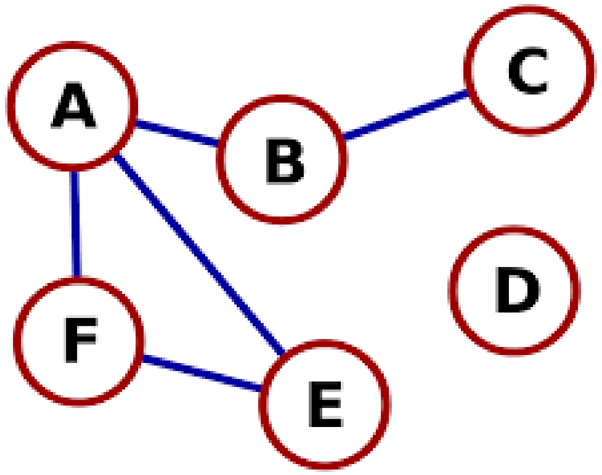
Multi-Layer Perceptron

Transformer



Graph Neural Network

Graph Neural Network



Nodi: A, B, ..., F

Connessioni

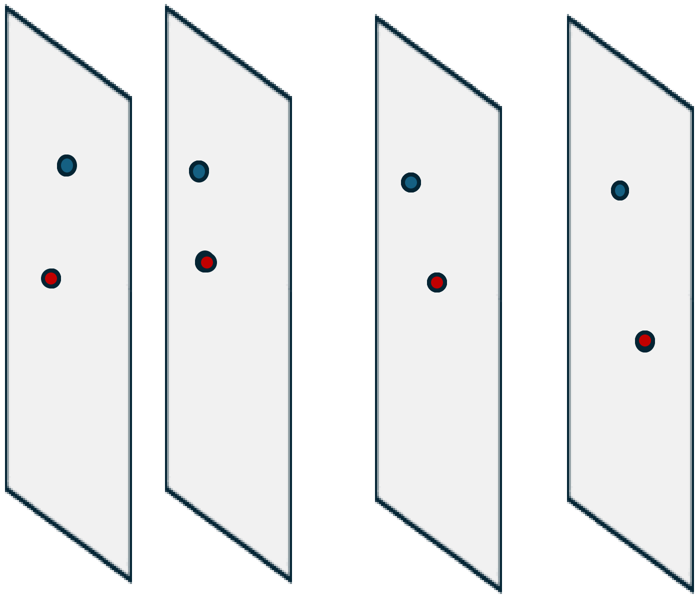
Le **Graph Neural Networks** sono un tipo di rete neurale progettata per lavorare con dati strutturati sotto forma di grafo.
Propagano e aggregano informazioni tra i nodi e i lati di un grafo, consentendo di apprendere rappresentazioni che catturano le relazioni e la struttura dei dati.
Vengono utilizzate per compiti come la classificazione dei nodi, la previsione dei collegamenti e la classificazione dei grafi.

Graph Neural Network

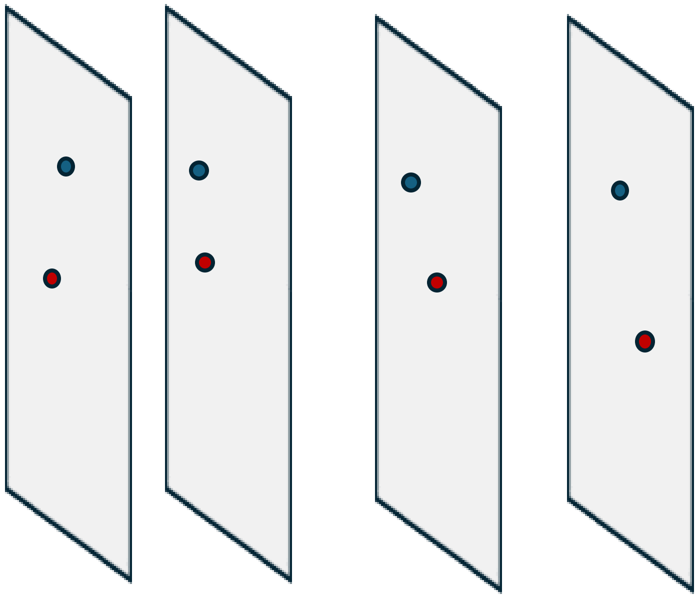
Dobbiamo considerare tutte le connessioni fra le hit

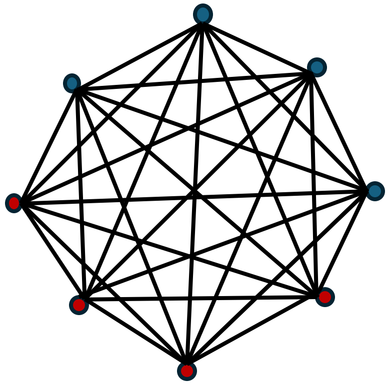
Con i grafi decidiamo in maniera molto naturale le connessioni

Graph Neural Network

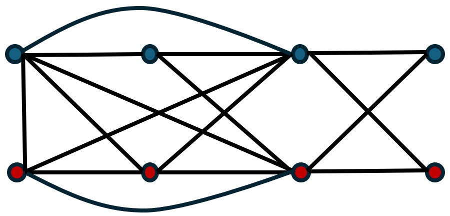


Graph Neural Network



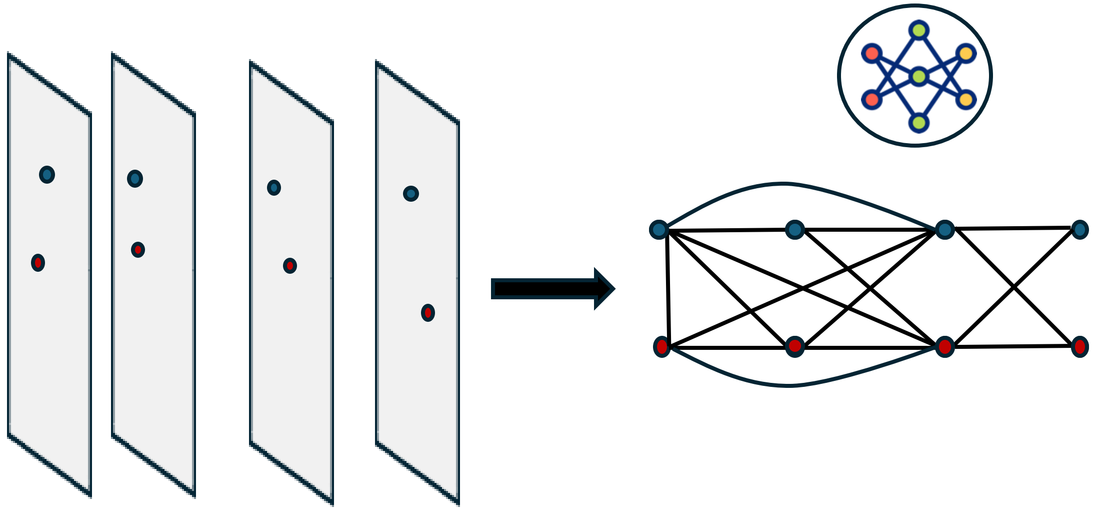


Complete Graph

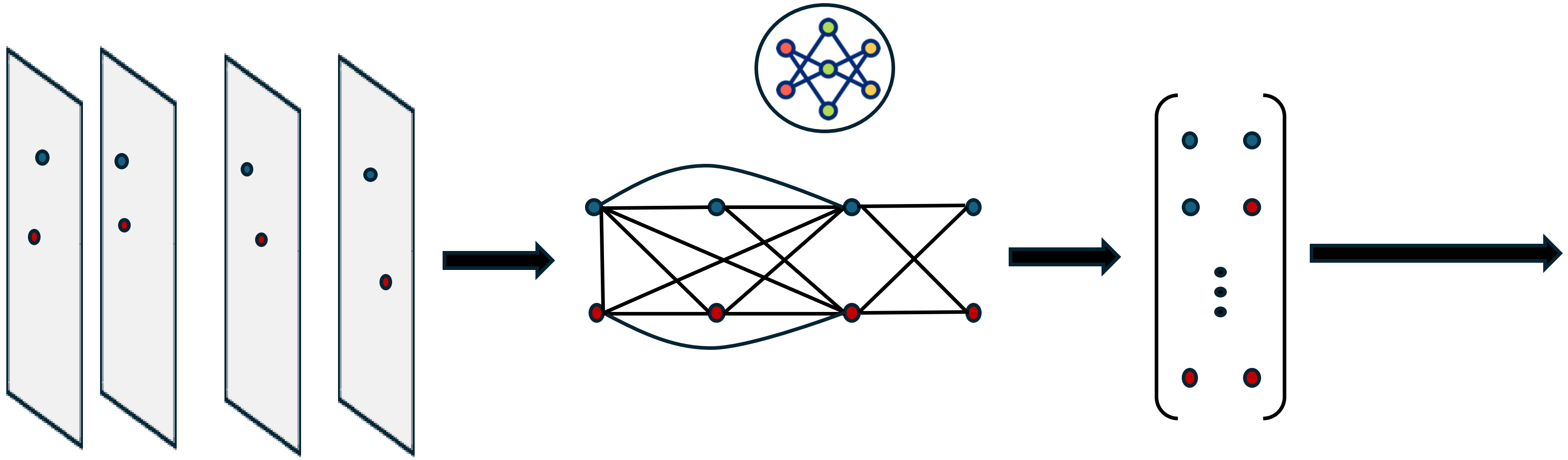


Sparse Graph

Graph Neural Network

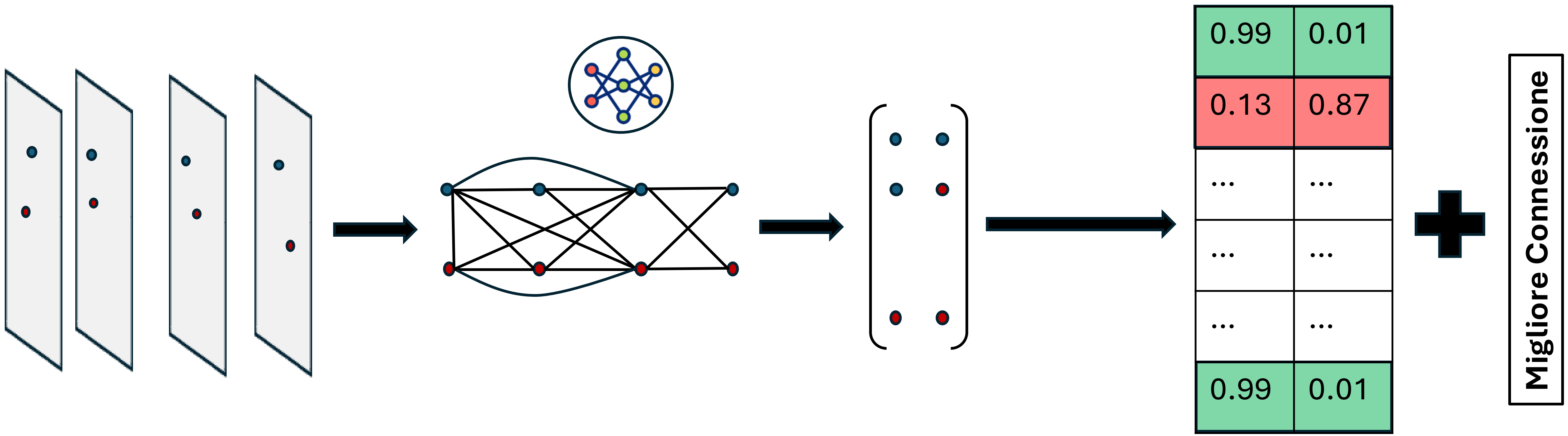


Graph Neural Network



0.99	0.01
0.13	0.87
...	...
...	...
...	...
0.99	0.01

Graph Neural Network



Graph Neural Network

Efficienza $= \dfrac{\#\,Tracce\ Predette\ Correttamente}{\#\,Totale\ Tracce\ Vere}$

Purezza $= \dfrac{\#\,Tracce\ Predette\ Correttamente}{\#\,Totale\ Tracce\ Predette}$

Tracce False $= \dfrac{\#\,Tracce\ Predette\ con\ Errore}{\#\,Totale\ Tracce\ Vere}$

Multi-Layer Perceptron

Efficienza	Purezza	Tracce False
70.06 %	92.3 %	39.45 %

Transformer

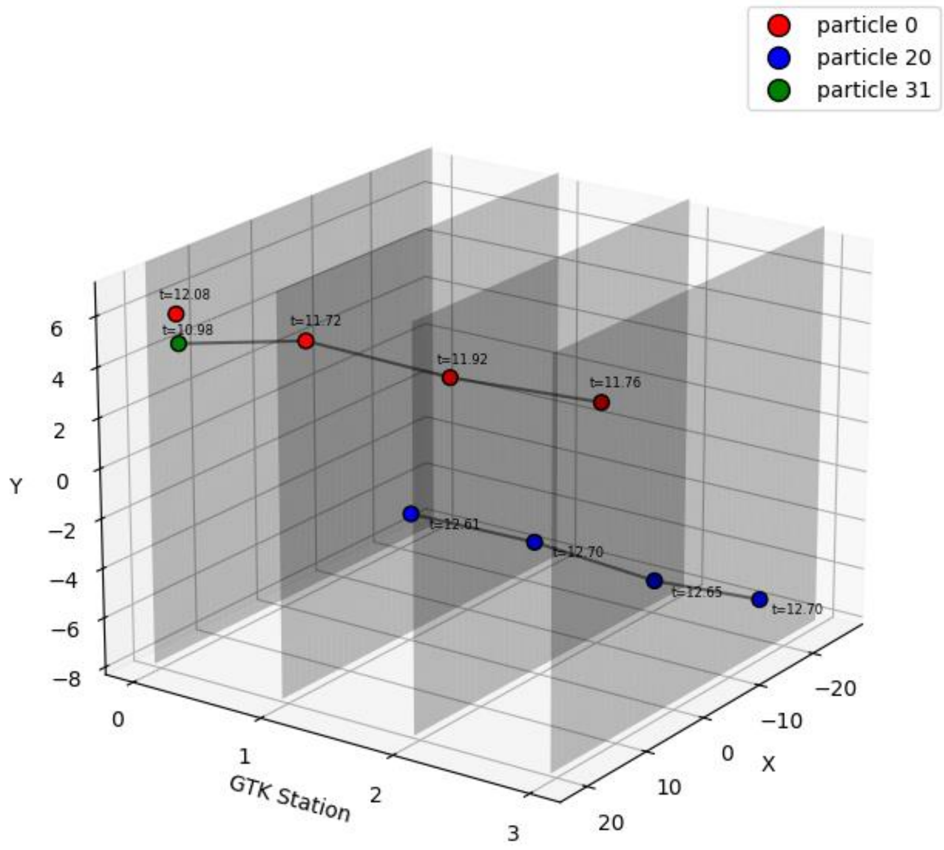
Efficienza	Purezza	Tracce False
95.95 %	98.62 %	1.22 %

Graph Neural Network*

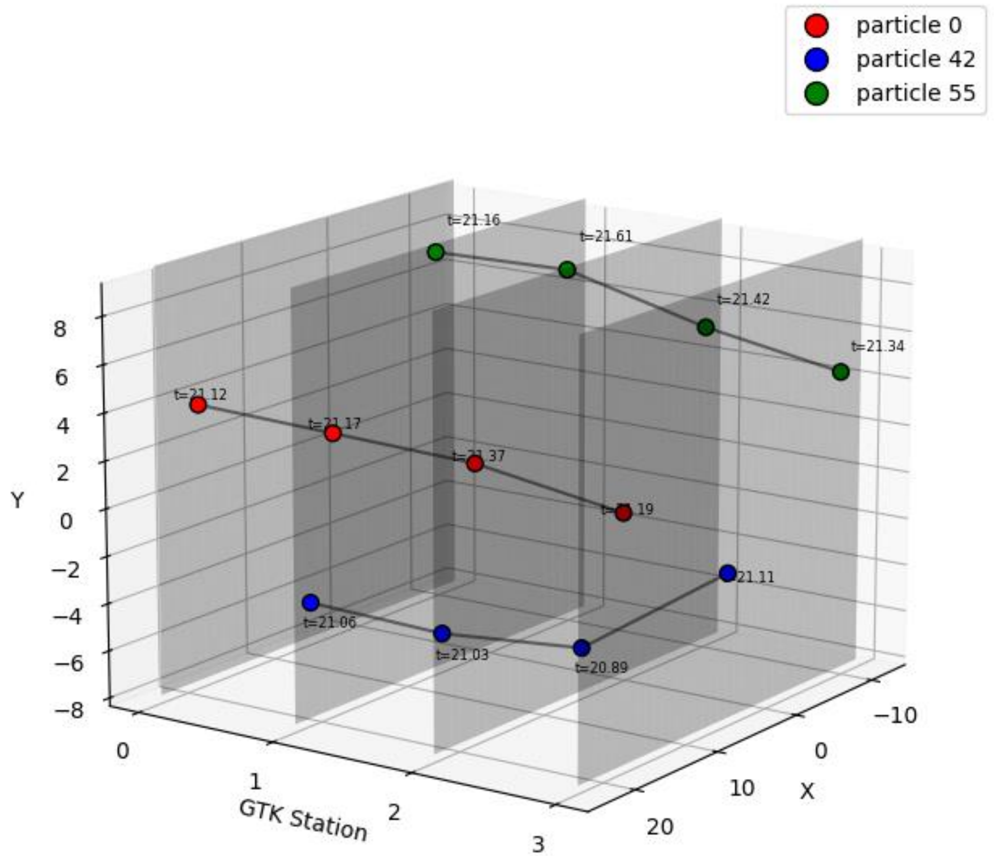
Efficienza	Purezza	Tracce False
94.78 %	99.78 %	0.21 %

Graph Neural Network

Un errore



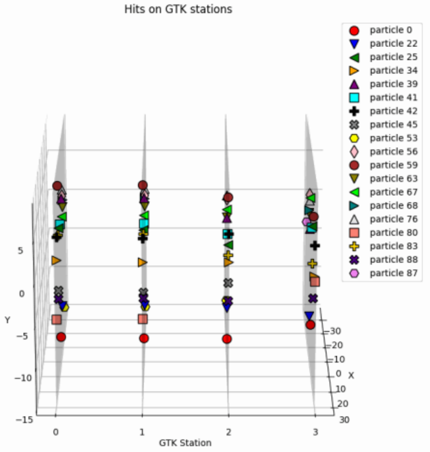
Predizione buona



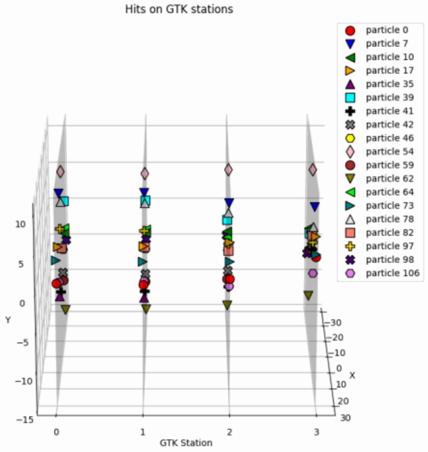
Graph Neural Network

- Risolto il problema del bilanciamento
- Conoscenza globale
- Efficienti nei calcoli
- Decisione delle connessioni flessibile

Graph Neural Network



● ● ●



Graph Neural Network