

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 1: “Dati, Informazioni e Conoscenza”



Bing Image Creator

DATA DRIVING

Il data driving è un approccio basato sui dati che le organizzazioni utilizzano per prendere decisioni migliori. Si tratta di raccogliere, analizzare e interpretare i dati per comprendere meglio i propri clienti, le proprie attività e il proprio mercato. Le informazioni ottenute da questi dati possono quindi essere utilizzate per migliorare i prodotti e i servizi, aumentare le vendite e ottimizzare le operazioni.

Modulo 1: “Dati, Informazioni e Conoscenza”

Il data driven è diventato sempre più importante negli ultimi anni, poiché le aziende hanno accesso a più dati che mai. Questo è dovuto in parte alla crescita di Internet e dei dispositivi mobili, che generano una grande quantità di dati sui clienti e sulle loro attività. Inoltre, le aziende stanno investendo sempre di più in tecnologie di analisi dei dati, che permettono loro di estrarre informazioni più utili dai dati raccolti.

Il data driven può essere utilizzato in tutti i settori di attività, ma è particolarmente diffuso nel marketing, nelle vendite e nel servizio clienti. Ad esempio, le aziende possono utilizzare i dati per:

Modulo 1: “Dati, Informazioni e Conoscenza”

- **Identificare** i propri clienti migliori e targettizzare le campagne di marketing di conseguenza.
- **Personalizzare** i prodotti e i servizi in base alle esigenze dei singoli clienti.
- **Prevedere** la domanda dei clienti e pianificare di conseguenza la produzione e le scorte.
- **Risolvere** i problemi dei clienti più rapidamente ed efficacemente.

Alcuni esempi di data driving in azione:

- Un'azienda di e-commerce utilizza i dati sui comportamenti di acquisto dei clienti per personalizzare le raccomandazioni di prodotto.
- Un'azienda di viaggi utilizza i dati sulle preferenze di viaggio dei clienti per creare offerte personalizzate.
- Un'azienda di servizi finanziari utilizza i dati sui comportamenti finanziari dei clienti per prevenire le frodi.
- Un'azienda di produzione utilizza i dati sui sensori delle macchine per prevedere guasti e ridurre i tempi di fermo.

Modulo 1: “Dati, Informazioni e Conoscenza”

Dato

Un dato è un'entità elementare che rappresenta un'osservazione o un'esperienza. È un'unità di misura, un simbolo o un'immagine che, da solo, non ha alcun significato.

Ad esempio, il numero 1 è un dato, così come la parola "cane" o l'immagine di un albero.

Informazione

Un'informazione è un insieme di dati correlati tra loro. Assume significato quando viene inserita in un contesto specifico.

Ad esempio, la frase "Il numero 1 è il primo numero naturale" è un'informazione, perché il numero 1 assume significato solo in relazione al concetto di "numero naturale".

Modulo 1: “Dati, Informazioni e Conoscenza”

Conoscenza

La conoscenza è un insieme di informazioni organizzate e integrate in un sistema coerente. È il risultato di un processo di apprendimento e comprensione.

Ad esempio, la conoscenza del mondo che ci circonda è il risultato dell'elaborazione di un'enorme quantità di informazioni, provenienti da diverse fonti.

Modulo 1: “Dati, Informazioni e Conoscenza”

In altre parole, i dati sono la **materia prima**, le informazioni sono il **prodotto semilavorato** e la conoscenza è il **prodotto finito**.

Ad esempio, i dati relativi alle vendite di un'azienda possono essere utilizzati per creare informazioni sui trend di mercato. Queste informazioni, a loro volta, possono essere utilizzate per prendere decisioni strategiche, che rappresentano un esempio di conoscenza.

Esempi

Ecco alcuni esempi concreti di dati, informazioni e conoscenza:

- **Dati:** il numero di abitanti di un paese, il prezzo di un litro di benzina, la temperatura media di un mese.
- **Informazioni:** il tasso di natalità di un paese, l'inflazione, il clima di una regione.
- **Conoscenza:** la storia di un paese, l'economia di un mercato, le leggi della fisica.

Modulo 1: “Dati, Informazioni e Conoscenza”

DATI STRUTTURATI E NON STRUTTURATI

I dati strutturati e non strutturati sono due tipi di dati che si differenziano per la loro forma e il modo in cui vengono archiviati.

Dati strutturati

I dati strutturati sono dati che sono organizzati in un formato predefinito. Questo significa che i dati sono raggruppati in campi e tabelle, e ogni campo ha un tipo di dato specifico. I dati strutturati sono spesso memorizzati in database relazionali, che sono progettati per archiviare e recuperare dati strutturati in modo efficiente.

Alcuni esempi di dati strutturati includono:

- Dati numerici, come numeri di telefono, indirizzi e codici postali
- Dati di testo, come nomi, cognomi e indirizzi email
- Dati di date e orari, come date di nascita e orari di arrivo

Modulo 1: “Dati, Informazioni e Conoscenza”

Dati non strutturati

I dati non strutturati sono dati che non sono organizzati in un formato predefinito. Questo significa che i dati possono essere di qualsiasi forma o dimensione, e non hanno una struttura definita. I dati non strutturati sono spesso memorizzati in file di testo, immagini, video o audio.

Alcuni esempi di dati non strutturati includono:

- Testo libero, come articoli di notizie, post sui social media e e-mail
- Immagini, come foto e video
- Audio, come file musicali e podcast

Modulo 1: “Dati, Informazioni e Conoscenza”

I tipi di dati si possono classificare nel seguente modo:

Dati numerici

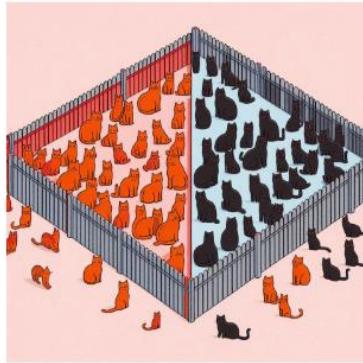
- **Dati discreti:** dati che possono assumere solo un numero finito di valori.
Ad esempio, il numero di figli che ha una persona, il numero di libri in una biblioteca, il numero di punti segnati in una partita di calcio.

Modulo 1: “Dati, Informazioni e Conoscenza”

- **Dati continui:** dati che possono assumere qualsiasi valore all'interno di un intervallo. Ad esempio, l'altezza di una persona, il peso di un oggetto, la temperatura di un ambiente.

Modulo 1: “Dati, Informazioni e Conoscenza”

Dati categorici



Gatti rossi e neri con Image creator

- **Dati nominali:** dati che rappresentano categorie senza un ordine naturale.
Ad esempio, il colore dei capelli/gatti, il paese di nascita, lo stato civile.

Modulo 1: “Dati, Informazioni e Conoscenza”

- **Dati ordinali:** dati che rappresentano categorie con un ordine naturale. Ad esempio, il grado di soddisfazione di un cliente (1 = molto insoddisfatto, 5 = molto soddisfatto), il livello di istruzione (1 = laurea, 2 = diploma, 3 = licenza media), il podio di una gara (1° posto, 2° posto, 3° posto).

Modulo 1: “Dati, Informazioni e Conoscenza”

Ecco alcuni esempi di ciascun tipo di dati:

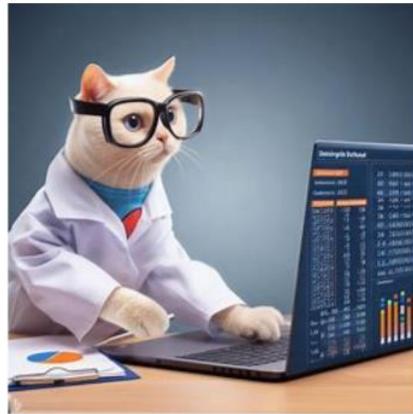
Numerici discreti: Numero di figli, numero di libri in una biblioteca, numero di punti segnati in una partita di calcio

Numerici continui: Altezza di una persona, peso di un oggetto, temperatura di un ambiente

Nominali: Colore dei capelli, paese di nascita, stato civile

Ordinali: Grado di soddisfazione di un cliente, livello di istruzione, podio di una gara

Modulo 1: “Dati, Informazioni e Conoscenza”



Bing Image Creator

Un dataset è un insieme di dati strutturati e organizzati in modo da poter essere utilizzati per scopi di analisi e apprendimento automatico.

I dati in un dataset possono essere di diversi tipi, come numeri, testo, immagini o video. I dati possono essere raccolti da diverse fonti, come sensori, database, o anche manualmente.

I dataset possono essere utilizzati per **una varietà di scopi**, come:

Modulo 1: “Dati, Informazioni e Conoscenza”

- **Analisi statistica:** i dataset possono essere utilizzati per analizzare tendenze, identificare correlazioni e prendere decisioni.
- **Apprendimento automatico:** i dataset possono essere utilizzati per addestrare modelli di apprendimento automatico, che possono poi essere utilizzati per svolgere compiti come la classificazione, la regressione o il riconoscimento di immagini.
- **Ricerca scientifica:** i dataset possono essere utilizzati per condurre ricerche scientifiche, come ad esempio studi epidemiologici o studi sul comportamento umano.

Modulo 1: “Dati, Informazioni e Conoscenza”

I dataset possono essere **classificati in base a diversi criteri**, come:

- **Dimensione**: i dataset possono essere di piccole dimensioni, medie dimensioni o grandi dimensioni.
- **Tipologia di dati**: i dataset possono contenere dati numerici, dati di testo, dati di immagini o dati di video.
- **Fonte dei dati**: i dataset possono essere raccolti da fonti interne, come database aziendali, o da fonti esterne, come sensori o siti web.

Modulo 1: “Dati, Informazioni e Conoscenza”

DATASET TITANIC



Modulo 1: “Dati, Informazioni e Conoscenza”

Il dataset contiene le seguenti colonne:

- **PassengerId**: un identificatore univoco per ogni passeggero.
- **Survived**: un valore booleano che indica se il passeggero è sopravvissuto all'affondamento (1) o no (0).
- **Pclass**: la classe di viaggio del passeggero (1 = prima classe, 2 = seconda classe, 3 = terza classe).
- **Name**: il nome del passeggero.
- **Sex**: il genere del passeggero (M = maschio, F = femmina).
- **Age**: l'età del passeggero.

Modulo 1: “Dati, Informazioni e Conoscenza”

- **SibSp**: il numero di fratelli o sorelle che il passeggero aveva a bordo del Titanic.
- **Parch**: il numero di genitori o figli che il passeggero aveva a bordo del Titanic.
- **Ticket**: il numero del biglietto del passeggero.
- **Fare**: la tariffa di viaggio del passeggero.
- **Cabin**: la cabina del passeggero (se presente).
- **Embarked**: il porto di partenza del passeggero (C = Cherbourg, Q = Queenstown, S = Southampton).

Modulo 1: “Dati, Informazioni e Conoscenza”

Il dataset Titanic è stato utilizzato in molti compiti di analisi e apprendimento automatico, come la classificazione dei passeggeri in base alla probabilità di sopravvivenza o l'identificazione dei fattori che hanno influenzato la sopravvivenza.

Il dataset Titanic è una risorsa preziosa per chiunque voglia imparare a lavorare con i dati e sviluppare modelli di apprendimento automatico. Il dataset è disponibile gratuitamente in diversi formati, tra cui CSV e JSON.

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito

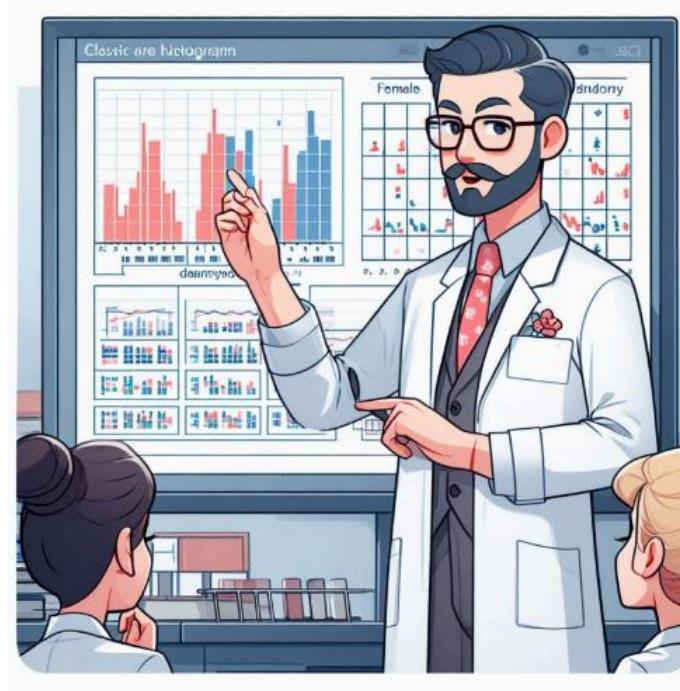


Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 1: “Rappresentazione Grafica dei Dati”



Bing Image Creator

Modulo 1: “Rappresentazione Grafica dei Dati”

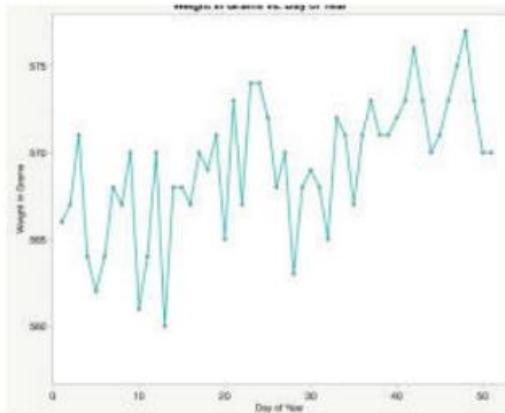
Perché è utile rappresentare i dati con i grafici?

La rappresentazione grafica dei dati è un modo efficace per comunicare informazioni quantitative. I grafici possono aiutare a rendere i dati più comprensibili e facili da interpretare, sia per le persone che non sono esperte di statistica sia per gli esperti.

Ecco alcuni dei motivi per cui è utile rappresentare i dati con i grafici:

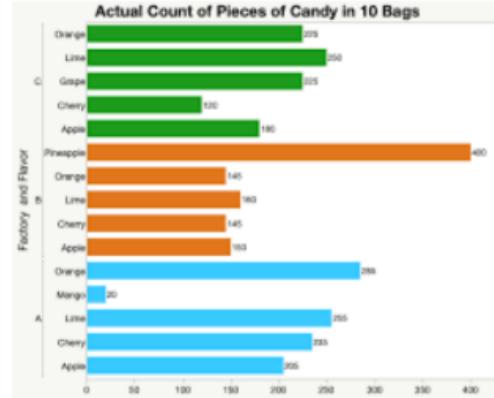
Modulo 1: “Rappresentazione Grafica dei Dati”

- I grafici possono aiutare a visualizzare rapidamente le tendenze e i modelli nei dati. Ad esempio, un grafico a linee può essere utilizzato per visualizzare come cambia un valore nel tempo.



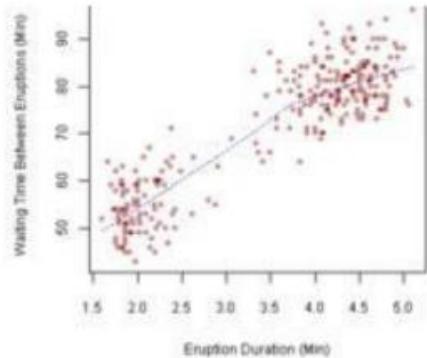
Modulo 1: “Rappresentazione Grafica dei Dati”

- I grafici possono aiutare a confrontare diversi set di dati. Ad esempio, un grafico a barre può essere utilizzato per confrontare i voti di due candidati in un'elezione.



Modulo 1: “Rappresentazione Grafica dei Dati”

- I grafici possono aiutare a evidenziare i valori anomali nei dati. Ad esempio, un grafico a dispersione può essere utilizzato per identificare i punti dati che si discostano dalla tendenza generale.



Modulo 1: “Rappresentazione Grafica dei Dati”

I grafici possono essere utilizzati in una **varietà di contesti**, tra cui:

- **Analisi dei dati** - I grafici vengono utilizzati dagli analisti dei dati per comprendere le informazioni quantitative.
- **Comunicazione dei dati** - I grafici vengono utilizzati per comunicare informazioni quantitative a un pubblico più ampio.
- **Decision making** - I grafici possono essere utilizzati per prendere decisioni informate basate sui dati.

Modulo 1: “Rappresentazione Grafica dei Dati”

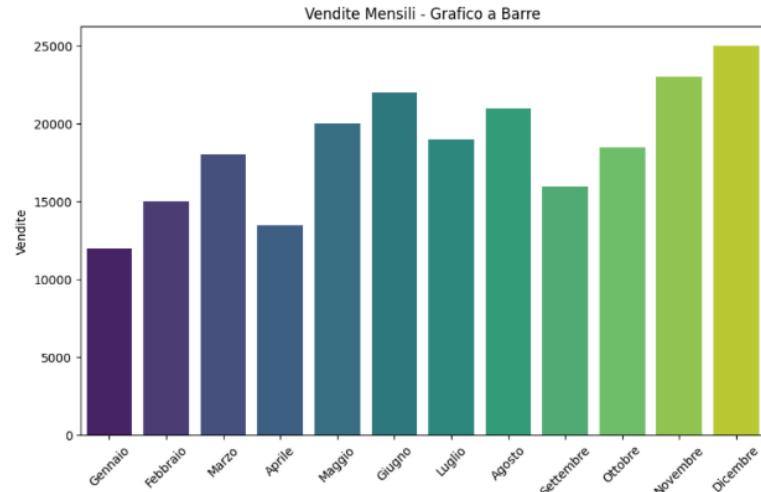
Ci sono diversi tipi di grafici che possono essere utilizzati per rappresentare dati in modo efficace a seconda della natura dei dati stessi e degli obiettivi della rappresentazione.

Di seguito fornisco una panoramica di alcuni dei grafici principali¹ relativamente ai dati:

Modulo 1: “Rappresentazione Grafica dei Dati”

Grafico a barre:

- Descrizione: Utilizzato per rappresentare dati discreti o categorici. Le barre rappresentano le categorie e la lunghezza della barra è proporzionale al valore associato a ciascuna categoria.
- Quando usarlo: Per confrontare valori tra diverse categorie.

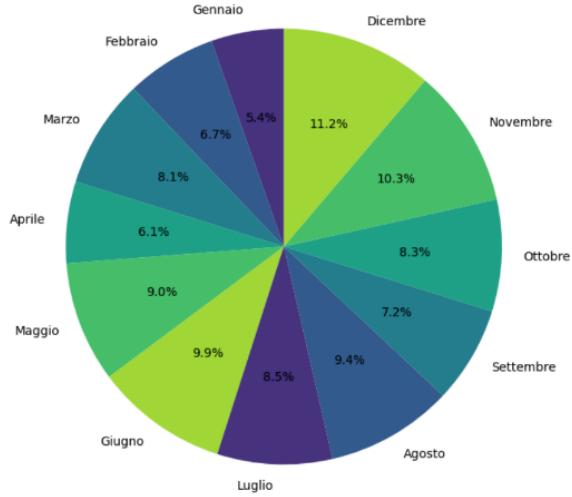


Modulo 1: “Rappresentazione Grafica dei Dati”

Grafico a torta:

- Descrizione: Mostra la proporzione di parti rispetto al tutto. È utile quando si vuole evidenziare la composizione percentuale dei dati.
- Quando usarlo: Per mostrare la distribuzione percentuale di un insieme di categorie.

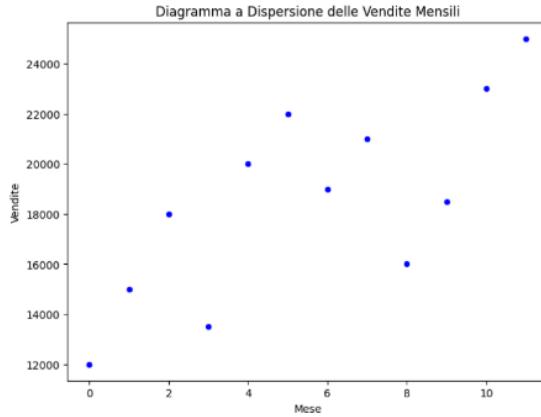
Distribuzione delle Vendite Mensili - Grafico a Torta



Modulo 1: “Rappresentazione Grafica dei Dati”

Diagramma a dispersione (scatter plot):

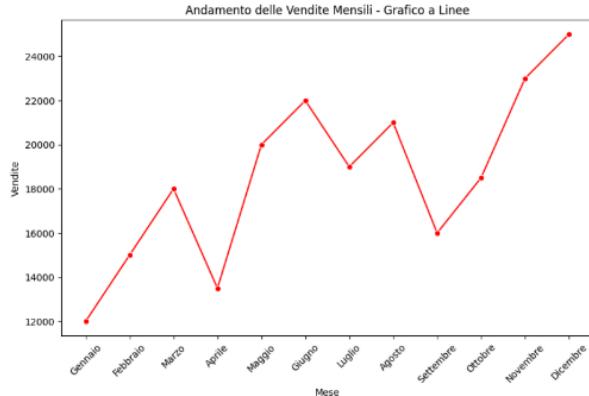
- Descrizione: Rappresenta la relazione tra due variabili. Ogni punto sul grafico rappresenta una coppia di valori.
- Quando usarlo: Per visualizzare la correlazione tra due variabili.



Modulo 1: “Rappresentazione Grafica dei Dati”

Grafico a linee:

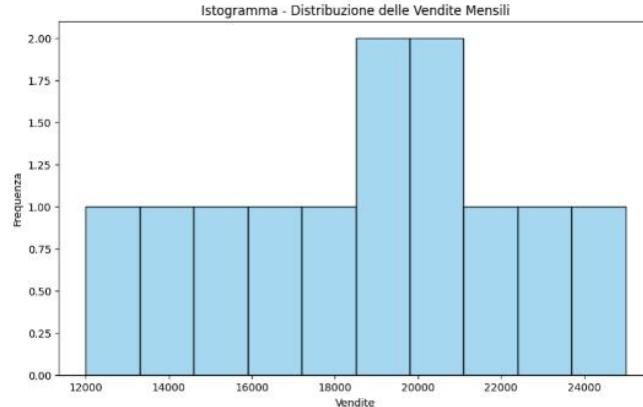
- Descrizione: Mostra la variazione di una variabile rispetto a un'altra attraverso linee che connettono i punti dati.
- Quando usarlo: Per mostrare tendenze temporali o connessioni tra dati ordinati.



Modulo 1: “Rappresentazione Grafica dei Dati”

Istogramma:

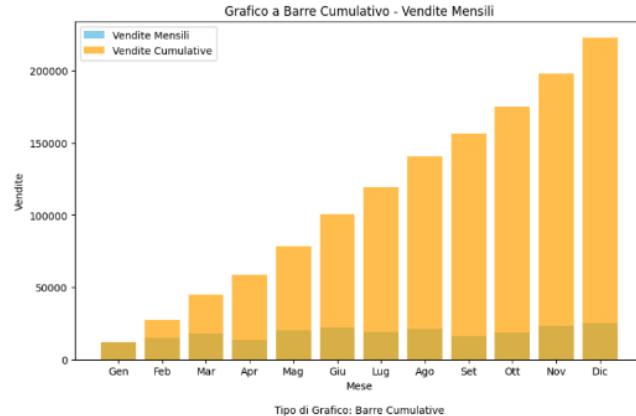
- Descrizione: Rappresenta la distribuzione di una variabile continua attraverso barre verticali. Ogni barra rappresenta un intervallo di valori.
- Quando usarlo: Per visualizzare la distribuzione di dati continui.



Modulo 1: “Rappresentazione Grafica dei Dati”

Diagramma a settori (o a barre cumulativo):

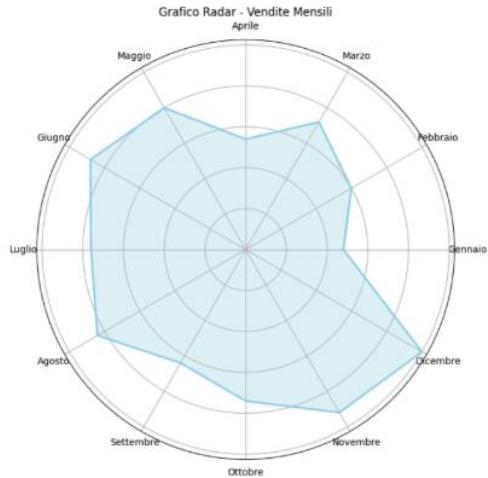
- Descrizione: Mostra come una quantità si suddivide in parti più piccole attraverso cerchi concentrici o barre cumulativa.
- Quando usarlo: Per mostrare la struttura a livelli di una quantità.



Modulo 1: “Rappresentazione Grafica dei Dati”

Grafico a radar:

- Descrizione: Rappresenta dati multivariati in un formato circolare. Utile per mostrare le prestazioni relative in diverse categorie.
- Quando usarlo: Per confrontare le prestazioni in diverse categorie su più variabili.



Modulo 1: “Rappresentazione Grafica dei Dati”



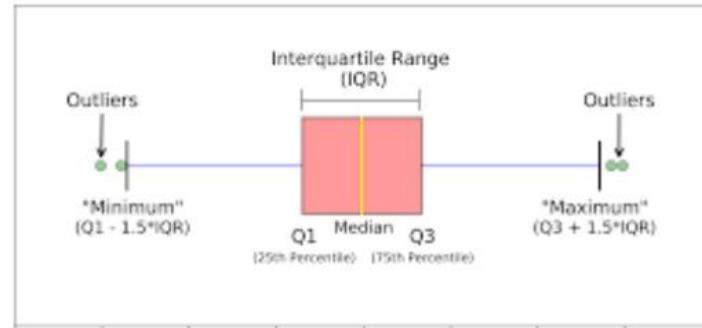
Stravaganza di Bing Image Creator

Modulo 1: “Rappresentazione Grafica dei Dati”

Un **box plot**, noto anche come diagramma a scatola, è un tipo di rappresentazione grafica che mostra la distribuzione di un insieme di dati numerici. Questo tipo di grafico fornisce una visualizzazione sintetica di alcune importanti caratteristiche della distribuzione, come la mediana, il primo e terzo quartile e la presenza di eventuali valori anomali.

Modulo 1: “Rappresentazione Grafica dei Dati”

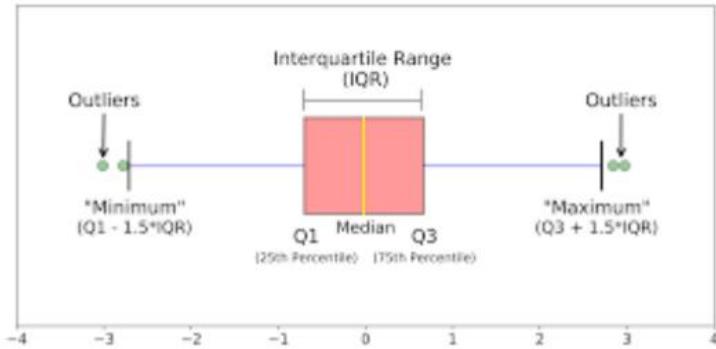
Ecco come interpretare un box plot:



Rettangolo (scatola): Il rettangolo rappresenta il 50% centrale dei dati ed è delimitato da due linee orizzontali. La parte inferiore del rettangolo indica il primo quartile (25% dei dati più bassi), mentre la parte superiore indica il terzo quartile (75% dei dati più bassi).

Modulo 1: “Rappresentazione Grafica dei Dati”

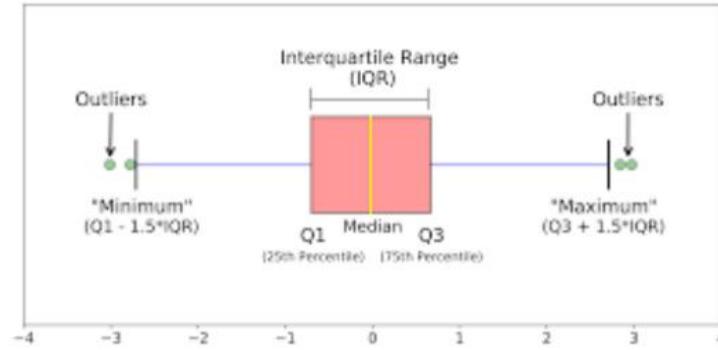
Ecco come interpretare un box plot:



Linea interna (mediana): All'interno del rettangolo c'è una linea orizzontale che rappresenta la mediana, ovvero il valore che divide il set di dati a metà.

Modulo 1: “Rappresentazione Grafica dei Dati”

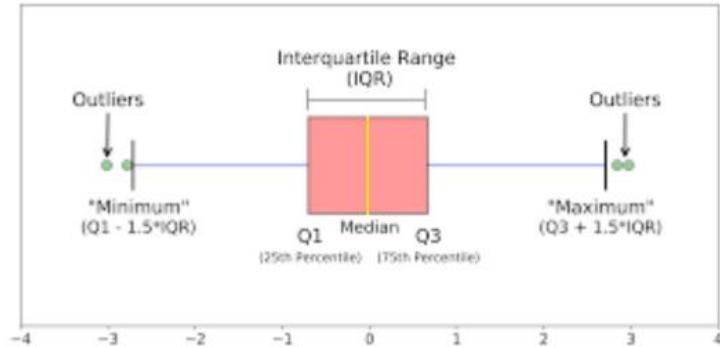
Ecco come interpretare un box plot:



Baffi (whiskers): I baffi si estendono dal rettangolo verso l'alto e verso il basso e rappresentano la variabilità al di fuori del 50% centrale dei dati. I baffi possono essere calcolati in diversi modi; comunemente vengono estesi fino al valore massimo o minimo all'interno di una determinata distanza interquartile.

Modulo 1: “Rappresentazione Grafica dei Dati”

Ecco come interpretare un box plot:

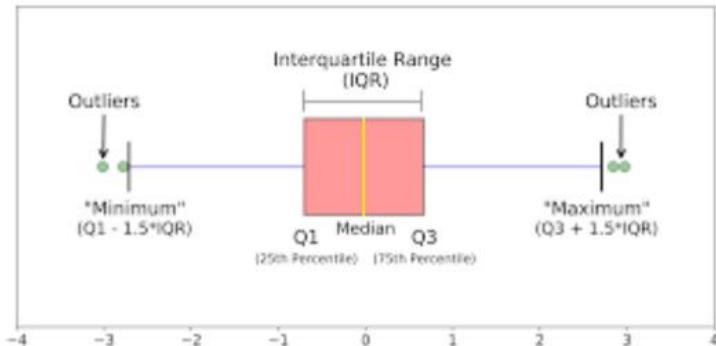


La distanza interquartile (IQR) è calcolata sottraendo il primo quartile (Q1) dal terzo quartile (Q3). La formula per il calcolo dell'IQR è:

$$IQR = Q3 - Q1$$

Modulo 1: “Rappresentazione Grafica dei Dati”

Ecco come interpretare un box plot:



I baffi vengono spesso estesi fino al massimo e al minimo dei dati all'interno di $1.5 \times IQR$ dalla parte superiore e inferiore della scatola. Eventuali dati oltre questi limiti vengono considerati valori anomali e rappresentati separatamente.

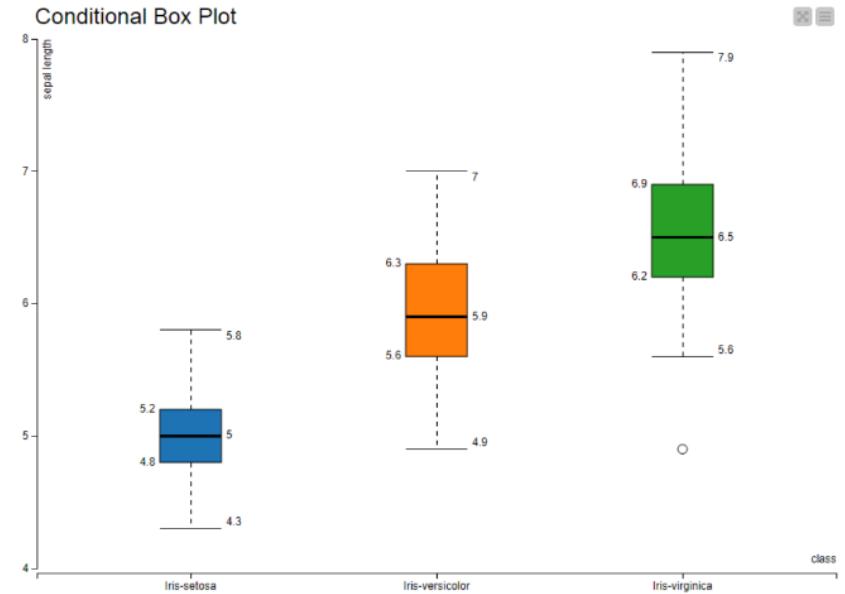
Nel nostro esempio di box plot, possiamo calcolare i valori dei baffi come segue:

$$\text{Limite Inferiore del Baffo} = Q1 - 1.5 \times IQR$$

$$\text{Limite Superiore del Baffo} = Q3 + 1.5 \times IQR$$

Modulo 1: “Rappresentazione Grafica dei Dati”

Il Box Plot del dataset IRIS



FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 1: “Pre-elaborazione dei dati”



Bing Image Creator

FUTURA
LA SCUOLA
PER L'ITALIA DI DOMANI

Finanziato dall'Unione europea
Ministero dell'Istruzione
e del Merito
Italiadomani

Modulo 1: “Pre-elaborazione dei dati”

La fase di pre-elaborazione dei dati è la prima fase del processo di elaborazione dei dati. In questa fase, i dati grezzi vengono puliti e organizzati per la fase successiva,

l'elaborazione. L'obiettivo della pre-elaborazione è garantire che i dati siano accurati, completi e coerenti, in modo che i risultati dell'analisi siano validi.

Modulo 1: “Pre-elaborazione dei dati”

La fase di pre-elaborazione è un processo importante e spesso trascurato. Un'adeguata pre-elaborazione dei dati è essenziale per garantire che l'analisi dei dati sia accurata e affidabile.

Modulo 1: “Pre-elaborazione dei dati”

La fase di pre-elaborazione può essere suddivisa in tre fasi principali:

1. Raccolta dei dati

In questa fase, i dati vengono raccolti da varie fonti, come database, file, sensori e social media. La raccolta dei dati può essere un processo complesso e richiede un'attenta pianificazione.

Esempio pratico: Un'azienda che vuole analizzare le vendite dei suoi prodotti deve raccogliere dati da diversi sistemi, come il sistema ERP, il sistema di e-commerce e il sistema CRM.

Modulo 1: “Pre-elaborazione dei dati”

2. Pulizia dei dati

In questa fase, i dati vengono puliti da errori e anomalie. Gli errori possono essere di tipo diverso, come dati mancanti, dati duplicati, dati non validi e dati incoerenti.

Esempio pratico: Un'azienda che vuole analizzare i dati dei clienti deve pulire i dati da indirizzi email errati, numeri di telefono non validi e date di nascita non coerenti.

Modulo 1: “Pre-elaborazione dei dati”

3. Organizzazione dei dati

In questa fase, i dati vengono organizzati in modo da essere facilmente accessibili e analizzabili. I dati possono essere organizzati in vari modi, come per formato, per fonte, per argomento o per periodo di tempo.

Esempio pratico: Un'azienda che vuole analizzare i dati delle vendite dei suoi prodotti deve organizzare i dati per periodo di tempo, prodotto e canale di vendita.

Modulo 1: “Pre-elaborazione dei dati”

Oltre a queste tre fasi principali, la fase di pre-elaborazione può includere anche altre attività, come:

- Denormalizzazione dei dati
- Normalizzazione dei dati
- Rimozione di duplicati
- Rimozione di outlier
- Aggiunta di valori mancanti

Modulo 1: “Pre-elaborazione dei dati”

I VALORI MANCANTI



I valori mancanti sono dati che sono assenti o sconosciuti. Possono verificarsi per una serie di motivi, come errori di input, mancanza di dati o rifiuto dei partecipanti a fornire informazioni.

Modulo 1: “Pre-elaborazione dei dati”

I valori mancanti possono avere un impatto significativo sull'analisi dei dati. Possono distorcere i risultati, rendendoli meno accurati e affidabili.

Esistono diversi metodi per gestire i valori mancanti, tra cui:

- **Elimina i dati con valori mancanti.** Questo è il metodo più semplice, ma può portare alla perdita di informazioni.
- **Rimpiazza i valori mancanti con valori stimati.** Questo può essere fatto utilizzando una varietà di metodi, come la media dei valori vicini o la regressione lineare.
- **Ignora i valori mancanti.** Questo può essere fatto se i valori mancanti sono pochi e non si ritiene che abbiano un impatto significativo sui risultati.

Modulo 1: “Pre-elaborazione dei dati”

Il metodo migliore per gestire i valori mancanti dipende da una serie di fattori, tra cui il tipo di dati, la quantità di valori mancanti e l'obiettivo dell'analisi.

Ecco alcuni esempi pratici di valori mancanti:

- Un foglio di calcolo con dati demografici potrebbe avere valori mancanti per l'età o il reddito.
- Un database di vendite potrebbe avere valori mancanti per il prezzo o la quantità.
- Un questionario potrebbe avere valori mancanti per le risposte a domande sensibili.

Modulo 1: “Pre-elaborazione dei dati”

È importante notare che i valori mancanti non sono sempre un problema. In alcuni casi, possono essere una parte naturale dei dati. Ad esempio, è normale che i dati medici abbiano valori mancanti per le misurazioni che non sono state effettuate.

Tuttavia, è importante essere consapevoli dell'impatto che i valori mancanti possono avere sull'analisi dei dati. Se i valori mancanti sono significativi, è importante adottare misure per gestirli in modo appropriato.

Modulo 1: “Pre-elaborazione dei dati”

GII OUTLIER



Gli outlier sono dati che si discostano significativamente dal resto del set di dati. Possono essere causati da errori di input, anomalie nel sistema o dati reali che si discostano dalla norma.

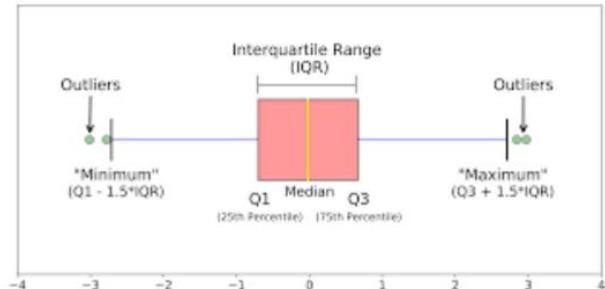
Modulo 1: “Pre-elaborazione dei dati”

Gli outlier possono avere un impatto significativo sull'analisi dei dati. Possono distorcere i risultati, rendendoli meno accurati e affidabili.

Esistono diversi metodi per identificare gli outlier, tra cui:

- **Visualizzazione dei dati.** È possibile identificare gli outlier osservando i dati in un grafico o in un altro formato visivo. Gli outlier appariranno come punti che si discostano dal resto dei dati.

Ecco come interpretare un box plot:



Modulo 1: “Pre-elaborazione dei dati”

- **Utilizzo di tecniche statistiche.** Esistono una varietà di tecniche statistiche che possono essere utilizzate per identificare gli outlier, come l'analisi di boxplot e il test di Grubbs.



Modulo 1: “Pre-elaborazione dei dati”

Esistono diversi metodi per gestire gli outlier, tra cui:

- **Elimina gli outlier.** Questo è il metodo più semplice, ma può portare alla perdita di informazioni.
- **Rimpiazza gli outlier** con valori stimati. Questo può essere fatto utilizzando una varietà di metodi, come la media dei valori vicini o la regressione lineare.
- **Mantieni gli outlier.** Questo può essere fatto se si ritiene che gli outlier siano dati reali che forniscono informazioni utili.

Modulo 1: “Pre-elaborazione dei dati”

Ecco alcuni esempi pratici di outlier:

- In un set di dati di temperature, un outlier potrebbe essere una temperatura di 50 gradi Celsius in un giorno in cui la temperatura media è di 20 gradi Celsius.
- In un set di dati di vendite, un outlier potrebbe essere una vendita di 100.000 euro per un prodotto che normalmente vende per 100 euro.
- In un set di dati di altezza, un outlier potrebbe essere un'altezza di 2 metri per un adulto.

Modulo 1: “Pre-elaborazione dei dati”

ELIMINAZIONE DEI DUPLICATI



La rimozione dei duplicati è un processo che consiste nel rimuovere i dati che si ripetono all'interno di un set di dati. I duplicati possono verificarsi per una serie di motivi, come errori di input, dati incompleti o dati non validi.

Modulo 1: “Pre-elaborazione dei dati”

La rimozione dei duplicati è importante per una serie di motivi, tra cui:

- **Migliorare l'accuratezza e l'affidabilità** dei dati. I duplicati possono distorcere i risultati dell'analisi, rendendoli meno accurati e affidabili.
- **Ridurre le dimensioni del set di dati.** La rimozione dei duplicati può ridurre le dimensioni del set di dati, rendendolo più facile da gestire e analizzare.
- **Migliorare le prestazioni delle query.** La rimozione dei duplicati può migliorare le prestazioni delle query, rendendole più veloci ed efficienti.

Modulo 1: “Pre-elaborazione dei dati”

Esistono diversi metodi per rimuovere i duplicati, tra cui:

- **Ricerca del valore duplicato.** Questo metodo consiste nel confrontare ogni valore del set di dati con tutti gli altri valori. Se un valore viene trovato più di una volta, viene considerato un duplicato.
- **Ricerca del valore chiave.** Questo metodo consiste nel definire una chiave per ogni valore del set di dati. I valori con la stessa chiave vengono considerati duplicati.
- **Ricerca dell'hash.** Questo metodo consiste nell'assegnare un hash a ogni valore del set di dati. I valori con lo stesso hash vengono considerati duplicati.

Modulo 1: “Pre-elaborazione dei dati”

In conclusione, la rimozione dei duplicati è un processo importante che può essere utilizzato in una varietà di contesti per migliorare l'accuratezza, l'affidabilità e la qualità dei dati.



FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 1: “Aggregazione e Correlazione dei Dati”

I dati spesso hanno bisogno di “Essere Ridotti”, mantenendo il contenuto informativo:

- riduzione delle caratteristiche
- riduzione/aggregazione di interi record



Modulo 1: “Aggregazione e Correlazione dei Dati”



Bing Image Creator

Il **binned dei dati** è un processo di aggregazione dei dati in gruppi, in base a un criterio predefinito. Questo processo può essere utilizzato per diversi scopi, tra cui:

Modulo 1: “Aggregazione e Correlazione dei Dati”

- Ridurre la dimensione dei dati: raggruppando i dati in gruppi, è possibile ridurre la dimensione del set di dati, rendendolo più gestibile e veloce da analizzare.
- Migliorare la leggibilità dei dati: raggruppando i dati in gruppi, è possibile rendere i dati più leggibili e comprensibili.
- Scoprire le tendenze nei dati: raggruppando i dati in gruppi, è possibile individuare le tendenze nei dati che potrebbero non essere visibili se i dati fossero analizzati singolarmente.

Modulo 1: “Aggregazione e Correlazione dei Dati”

Esistono diversi modi per eseguire il binned dei dati. Uno dei metodi più comuni è quello di utilizzare intervalli di uguale ampiezza.

Per ogni colonna è possibile definire un certo numero di intervalli, noti come bin. A ciascuno di questi bin viene assegnato un nome univoco (per quella colonna), un intervallo definito e bordi aperti o chiusi. Questi intervalli vengono automaticamente ordinati in ordine decrescente e i bordi degli intervalli sono coerenti. Inoltre, ogni colonna viene sostituita con la colonna binned di tipo stringa, oppure viene aggiunta una nuova colonna binned di tipo stringa.

Modulo 1: “Aggregazione e Correlazione dei Dati”

Ad esempio, se si dispone di un set di dati di temperatura, si potrebbero raggruppare i dati in intervalli di 10 gradi Celsius. In questo modo, tutti i dati compresi tra 0 e 10 gradi Celsius verrebbero raggruppati insieme, tutti i dati compresi tra 10 e 20 gradi Celsius verrebbero raggruppati insieme, e così via.

#	RowID	dati Number (integer)	dati_binned String
9	Row8	17	1-50
10	Row9	45	1-50
11	Row10	66	51-100
12	Row11	77	51-100
13	Row12	66	51-100
14	Row13	44	1-50
15	Row14	99	51-100
16	Row15	25	1-50
17	Row16	25	1-50
18	Row17	26	1-50
19	Row18	27	1-50
20	Row19	34	1-50
21	Row20	35	1-50
22	Row21	36	1-50
23	Row22	45	1-50
24	Row23	46	1-50

Modulo 1: “Aggregazione e Correlazione dei Dati”

Un altro metodo per eseguire il binned dei dati è quello di **utilizzare intervalli di ampiezza variabile**. Ad esempio, se si dispone di un set di dati di vendite, si potrebbero raggruppare i dati in intervalli di ampiezza variabile, in base al valore delle vendite. In questo modo, i dati con valori di vendita più alti verrebbero raggruppati insieme in intervalli più stretti, mentre i dati con valori di vendita più bassi verrebbero raggruppati insieme in intervalli più ampi.

Modulo 1: “Aggregazione e Correlazione dei Dati”

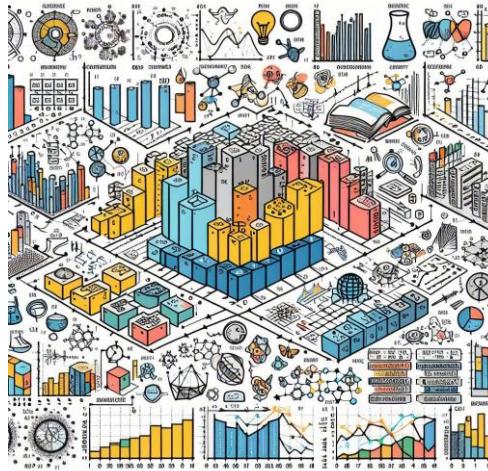
Ecco alcuni esempi di come il binned dei dati può essere utilizzato:

- In un **contesto aziendale**, il binned dei dati può essere utilizzato per analizzare le vendite, i costi e altri dati finanziari. Ad esempio, si potrebbe utilizzare il binned dei dati per individuare le tendenze nelle vendite stagionali o per identificare i prodotti o i servizi che generano il maggior profitto.



Modulo 1: “Aggregazione e Correlazione dei Dati”

- In un contesto scientifico, il binned dei dati può essere utilizzato per analizzare i dati sperimentali. Ad esempio, si potrebbe utilizzare il binned dei dati per individuare le tendenze nei dati biologici o per identificare i fattori che influenzano un particolare fenomeno.



Modulo 1: “Aggregazione e Correlazione dei Dati”

- In un **contesto sociale**, il binned dei dati può essere utilizzato per analizzare i dati demografici o i dati di opinione pubblica. Ad esempio, si potrebbe utilizzare il binned dei dati per individuare le tendenze demografiche o per identificare le opinioni pubbliche su un particolare argomento.



Modulo 1: “Aggregazione e Correlazione dei Dati”



Bing Image Creator

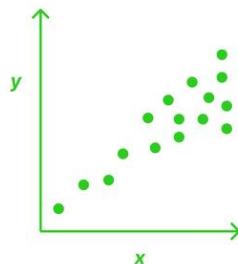
La **correlazione** dei dati è una misura della relazione tra due o più variabili. È una misura statistica che indica quanto i valori di due variabili sono correlati tra loro.

Modulo 1: “Aggregazione e Correlazione dei Dati”

Il **coefficiente di correlazione di Pearson** è il metodo più comune per misurare la correlazione tra due variabili quantitative. Il coefficiente di correlazione di Pearson è un numero compreso tra -1 e 1.

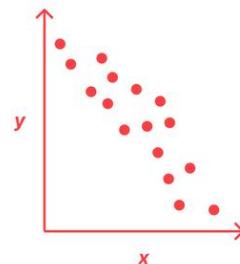
Relazione Positiva

Se aumenta X, aumenta Y



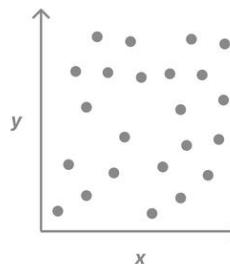
Relazione Negativa

Se aumenta X, diminuisce Y



Relazione Neutra

Se aumenta X,
Y aumenta o diminuisce



Modulo 1: “Aggregazione e Correlazione dei Dati”

- Un valore di $r = 1$ indica una correlazione positiva perfetta. Ciò significa che i valori delle due variabili aumentano o diminuiscono insieme in modo lineare.
- Un valore di $r = -1$ indica una correlazione negativa perfetta. Ciò significa che i valori delle due variabili aumentano o diminuiscono in modo lineare in direzioni opposte.
- Un valore di $r = 0$ indica che non esiste una correlazione tra le due variabili.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$
$$\rho_{AB} = \frac{\text{cov}(A, B)}{\sigma_A \cdot \sigma_B}$$

Modulo 1: “Aggregazione e Correlazione dei Dati”

Ecco alcuni **e esempi di correlazione** tra dati:

- **Correlazione positiva:**
 - L'altezza e il peso di una persona sono correlati positivamente. In generale, le persone più alte sono anche più pesanti.

Modulo 1: “Aggregazione e Correlazione dei Dati”

Ecco alcuni esempi di come la correlazione dei dati può essere utilizzata:

- **In economia:** La correlazione tra il prezzo del petrolio e l'inflazione può essere utilizzata per prevedere l'inflazione futura.
- **In medicina:** La correlazione tra il fumo e il cancro può essere utilizzata per promuovere campagne di sensibilizzazione sulla prevenzione del cancro.

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

The screenshot shows the Datablist website homepage. At the top, there is a navigation bar with the Datablist logo, a search bar, and dropdown menus for "Prodotto", "Casi d'uso", "Prezzo", and "Risorse". To the right are "Login" and "Apri Datablist" buttons. The main content area features a large, bold headline: "Lo strumento migliore per gestire le tue liste di prodotti". Below the headline is a subtext: "Stiamo reinventando la gestione delle liste per aumentare la tua produttività. Dì addio a Excel!". At the bottom of the main section is another "Apri Datablist" button and the text "Nessuna registrazione richiesta".

<https://www.datablist.com/it>

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

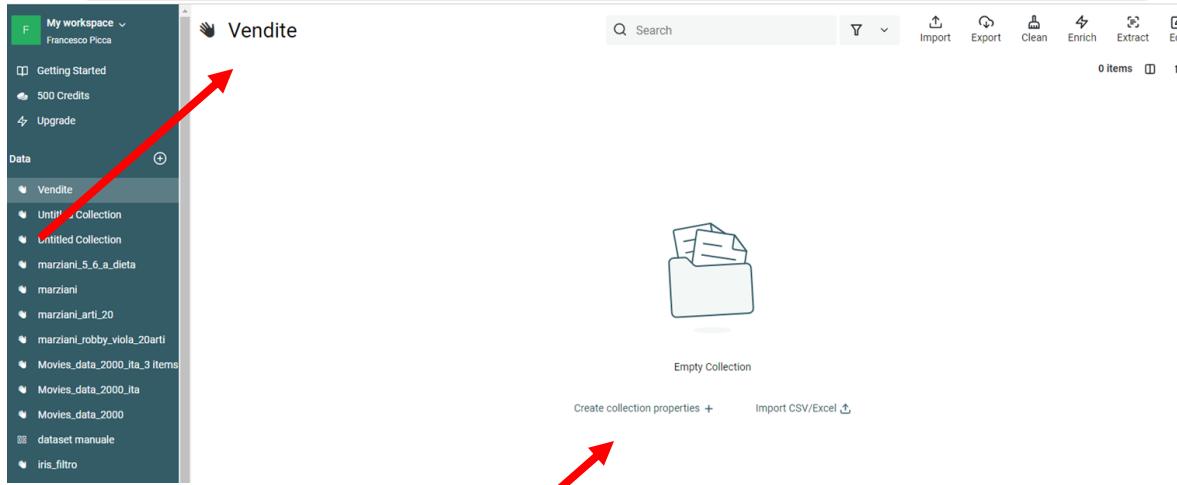
A screenshot of the Datablist web application. On the left is a dark sidebar titled 'My workspace' with a green icon containing a white letter 'F'. It shows the user's name 'Francesco Picca'. Below this are sections for 'Getting Started' (with a document icon), '500 Credits' (with a person icon), and 'Upgrade' (with a gear icon). The 'Data' section lists several collections: 'Untitled Collection' (x2), 'marziani_5_6_a_dieta', 'marziani', 'marziani_arti_20', 'marziani_robbi_viola_20arti', 'Movies_data_2000_Ita_3 items', 'Movies_data_2000_Ita', 'Movies_data_2000', 'dataset manuale', 'iris_filtro', and 'iris'. A red arrow points from the top-left towards the sidebar. The main content area has a light gray background. At the top, it says 'Getting Started' with a hand icon. Below this is a message: 'Thanks for joining Datablist beta program! We are thrilled to have you onboard! ☺ Feel free to play around and if you want to send us your feedback, contact us on contact@datablist.com'. There are two buttons: 'Start with a CSV/Excel file' (with a CSV icon) which says 'Import data from a CSV or Excel file into a collection.', and 'New collection' (with a plus icon) which says 'Start with a empty collection.' Below this is a section titled 'Blazing fast and keyboard friendly' with a list of keyboard shortcuts:

- Cmd + K: Search and jump to a collection
- C: In a collection, press 'C' to create a new item
- N: Press 'N' to create a new collection
- Cmd + ⌘: On a cell, press 'Shift' to open the item detail

**Iniziare un nuovo progetto con una
Nuova collezione o con un file csv/excel
esistente**

Modulo 2: “Analisi dei dati con Datablist”

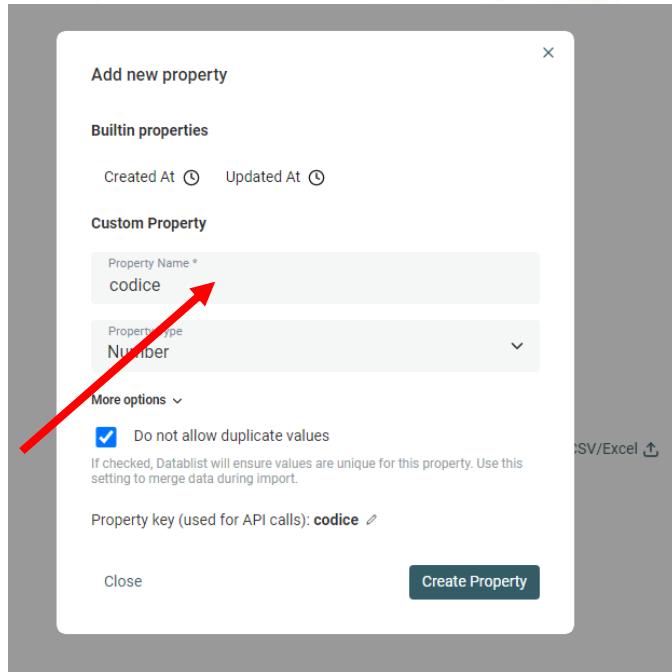
Accesso alla web app e prime elaborazioni



Proprietà della nuova
collezione

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni



Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Vendite

Search Import Export

codice descrizione prezzo +

Empty Collection

Create your first item Import CSV/Excel

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

The screenshot shows the Datablist interface. At the top, there is a toolbar with icons for Import, Export, Clean, Enrich, Extract, Edit, and a dropdown menu. Below the toolbar, a header bar indicates "1 items" and features a "New Item +" button, which is highlighted with a red arrow pointing from a callout box labeled "Aggiungere un nuovo record". The main area displays a table with a single row. The first column contains a search bar with the placeholder "Search" and a dropdown menu. The second column has a "prezzo" field with the value "1" and a "+" button. The third column contains fields for "codice" (value "1"), "descrizione" (value "chiave cat xyz"), and "prezzo" (value "3"). A red callout box labeled "Inserisco i dati" points to the "prezzo" field in the third column.

Aggiungere un nuovo record

Inserisco i dati

Search	+	New #item in Vendite
prezzo	1	<p>codice 1</p> <p>descrizione chiave cat xyz</p> <p>prezzo 3</p>

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Brevi video
tutorial

Remember the basics

Customers

First Name	Last Name	Email
Kenneth	Maxwell	perezmonique@...
Bradley	Diaz	thomas38@hot...
Stephanie	J	mullinssarah@h...
Mary	I	joseph54@hot...
Heather	Johnson	wellssteve@sma...
Jeffrey	Gomez	rebecca33@gm...

Data / Audience

First Name	Last Name	Email
Kenneth	Maxwell	pere...
✓ Diaz	Long	mull...
✓ Mary	Woodard	jose...
✓ Heather	Johnson	wells...
Jeffrey	Gomez	rebe...

Press Enter on a cell to edit its content

Click and drag to select multiple items

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Getting Started

Thanks for joining Datablist beta program! We are thrilled to have you onboard! ☺

Feel free to play around and if you have any questions or feedback, contact us on contact@datablist.com.

Apro un
csv/excel
esistente



Start with a CSV/Excel file

Import data from a CSV or
Excel file into a collection.



New collection

Start with an empty
collection.



Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Import CSV/Excel to Untitled Collection Step 1/3

Apro un csv/excel sistente

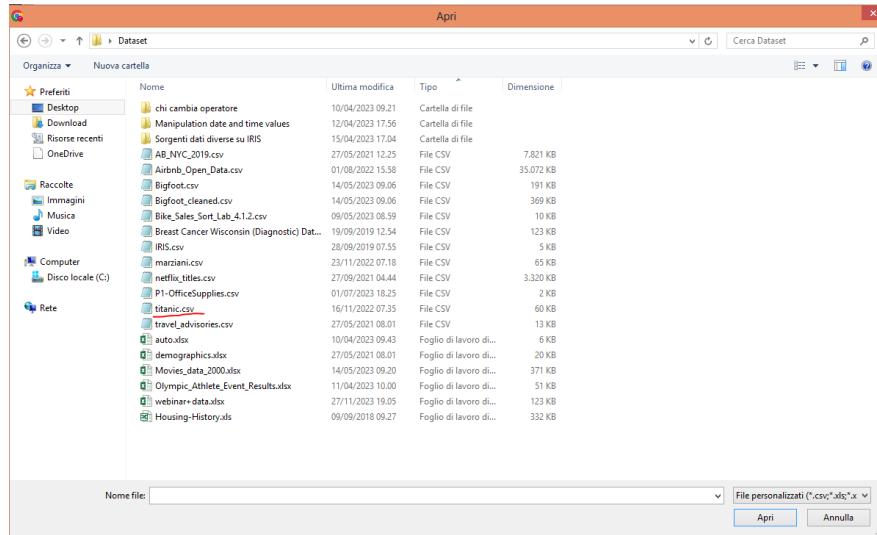
Drop a CSV or Excel file to upload or click here to choose file
The column headers must be in the first row. See "More options" to change this setting.

> [More options](#)

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Apro un
csv/excel
esistente



Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

STEP 1: File importato, leggo le caratteristiche, colonne e righe

Import CSV/Excel to  Untitled Collection

Step 1/3 - titanic.csv

60 KB - titanic.csv

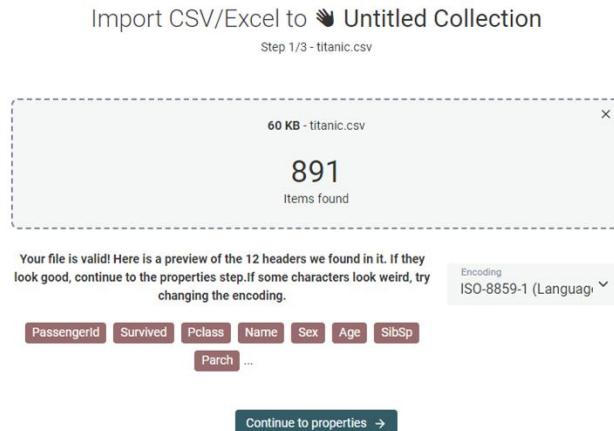
891 Items found

Your file is valid! Here is a preview of the 12 headers we found in it. If they look good, continue to the properties step. If some characters look weird, try changing the encoding.

Encoding: ISO-8859-1 (Language: 

PassengerId Survived Pclass Name Sex Age SibSp
Parch ...

Continue to properties →



Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Customize your properties
Step 2/3 - titanic.csv

Select the file columns you want to import and define your column data types.

Name	Data Type	Import?
Passengerid	Number	
Survived	Checkbox	
Pclass	Number	
Name	Text	
Sex	Text	
Age	Number	
SibSp	Number	
Parch	Number	
Ticket	Text	
Fare	Number	
Cabin	Text	
Embarked	Text	

[Continue →](#)

STEP 2: posso
nascondere
alcune
caratteristiche

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Customize your properties
Step 2/3 - titanic.csv

Select the file columns you want to import and define your column data types.

Name	Data Type	Import?
PassengerId	Number	<input checked="" type="radio"/>
Ticket	Text	<input type="radio"/>
Age	Number	<input type="radio"/>
Sex	Text	<input type="radio"/>
Survived	Number	<input type="radio"/>
Pclass	Text	<input type="radio"/>
Name	Text	<input type="radio"/>
Fare	Number	<input type="radio"/>
Cabin	Text	<input type="radio"/>

Edit property

Property Name *
Sex

Property Type
Text

Do not allow duplicate values
If checked, Datablist will ensure values are unique for this property. Use this setting to merge data during import.

STEP 2: oppure
cambiare il
formato dei dati

Modulo 2: “Analisi dei dati con Datablist”

Accesso alla web app e prime elaborazioni

Review before import

Step 3/3 - titanic.csv

File Data Preview (10/891 items)

PassengerId	Survived	Pclass	Name
1	□	3	Braund, Mr. Owen Harris
2	☒	1	Cumings, Mrs. John Bradley ...
3	☒	3	Heikkinen, Miss. Laina
4	☒	1	Futrelle, Mrs. Jacques Heath...
5	□	3	Allen, Mr. William Henry
6	□	3	Moran, Mr. James
7	□	1	McCarthy, Mr. Timothy J
8	□	3	Palsson, Master. Gosta Leon...
9	☒	3	Johnson, Mrs. Oscar W (Elis...
10	☒	2	Nasser, Mrs. Nicholas (Adele...

Show 10 more preview items

Import 891 items

STEP 3:
concludo la
procedura

titanic

	PassengerId	Survived	Pclass	Name	Sex	A 891 items	TF	Import	Export	Clear	Enrich	Extract	Edit
<input type="checkbox"/>	1	✓	3	Braund, Mr. Owen Harris	male	22	1						
<input type="checkbox"/>	2	✓	1	Cumings, Mrs. John Bradley...	female	38	1						
<input type="checkbox"/>	3	✓	3	Heikkinen, Miss. Laina	female	26	0						
<input type="checkbox"/>	4	✓	1	Futrelle, Mrs. Jacques Heat...	female	35	1						
<input type="checkbox"/>	5		3	Allen, Mr. William Henry	male	35	0						
<input type="checkbox"/>	6		3	Moran, Mr. James	male		0						
<input type="checkbox"/>	7		1	McCarthy, Mr. Timothy J	male	54	0						
<input type="checkbox"/>	8		3	Palsson, Master. Gosta Leo...	male	2	3						
<input type="checkbox"/>	9	✓	3	Johnson, Mrs. Oscar W (Elis...	female	27	0						
<input type="checkbox"/>	10	✓	2	Nasser, Mrs. Nicholas (Adele...	female	14	1						
<input type="checkbox"/>	11	✓	3	Sandstrom, Miss. Margaret...	female	4	1						
<input type="checkbox"/>	12	✓	1	Bonnell, Miss. Elizabeth	female	58	0						
<input type="checkbox"/>	13		3	Saundercock, Mr. William H...	male	20	0						
<input type="checkbox"/>	14		3	Andersson, Mr. Anders Johan	male	39	1						

File Items Parsed: 891
New Items saved in collection: 891

Import another file

Back to collection

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

The screenshot shows the Datablist website homepage. At the top, there is a navigation bar with the Datablist logo, a search icon, and dropdown menus for "Prodotto", "Casi d'uso", "Prezzo", and "Risorse". On the right side of the nav bar are "Login" and a prominent blue button labeled "Apri Datablist". Below the header, a large title reads "Lo strumento migliore per gestire le tue liste di prodotti" (The best tool to manage your product lists). A subtext below the title says "Stiamo reinventando la gestione delle liste per aumentare la tua produttività. Dì addio a Excel!" (We are reinventing list management to increase your productivity. Say goodbye to Excel!). At the bottom of the main section is another "Apri Datablist" button and the text "Nessuna registrazione richiesta" (No registration required).

<https://www.datablist.com/it>

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

Rinominare, Clonare
e Cancellare una
collezione

The screenshot shows the Datablist application interface. On the left, there's a sidebar with a user profile (Francesco Picca) and sections for 'Getting Started', '500 Credits', and 'Upgrade'. Below these are 'Data' and a list of collections: 'titanic', 'Vendite', 'Untitled Collection', 'Untitled Collection', 'marziani_5_6_a_dieta', 'marziani', 'marziani_arti_20', and 'marziani_robbi_viola_20arti'. The 'titanic' collection is selected and expanded, showing its data table. The table has two columns: 'Survived' and 'Pclass'. The data consists of five rows with values: (3, 1), (3, 1), (3, 1), (3, 1), and (2, 1). A context menu is open over the 'titanic' collection name, containing three options: 'Rename', 'Clone collection', and 'Delete'.

Survived	Pclass
3	1
3	1
3	1
3	1
2	1

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

The screenshot shows a data analysis interface with a table of 891 items. The columns are Survived, Pclass, Name, Sex, and Age. A red arrow points from the 'Export' button in the top right to a modal dialog titled 'Export collection'. The dialog has dropdowns for 'Export Format' (set to 'Comma-Separated Values (.csv)') and 'Column Delimiter' (set to 'Comma (default)'). It also includes a 'More options >' link and a large 'Export' button.

Survived	Pclass	Name	Sex	Age
	3	Braund, Mr. Owen Harris	male	22
✓	1	Cumings, Mrs. John Bradley...	female	38
✓	3	Heikkinen, Miss. Laina	female	26
✓	1	Futrelle, Mrs. Jacques Heat...	female	35
			male	35
			male	
			male	54
			male	2
			female	27
			female	14
			female	4
			female	58
	3	Saundercock, Mr. William H...	male	20

Esportare la collezione in formato csv/excel

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

titanic

Passengerid	Survived	Pclass	Name	Sex	Age
1	0	3	Braund, Mr. Owen Harris	male	22
2	1	1	Cumings, Mrs. John Bradley... female	female	38
3	1	3	Heikkinen, Miss. Laina	female	26
4	1	1	Futrelle, Mrs. Jacques Heath... female	female	35

Importare nuovi dati
nella collezione, da
una collezione o
esterni csv/excel

titanic

Passengerid	Survived	Pclass	Name	Sex	Age
1	0	3	Braund, Mr. Owen Harris	male	22
2	1	1	Cumings, Mrs. John Bradley... female	female	38
3	1	3	Heikkinen, Miss. Laina	female	26
4	1	1	Futrelle, Mrs. Jacques Heath... female	female	35

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

Map properties
Step 1/3 - marziani

Map the two collection properties or create new properties by clicking the **+** button.

Imported Property name	Collection Property
specie	Select a property to map Do not import
colore	Select a property to map Do not import
n_arti	Select a property to map Do not import
peso	Select a property to map Do not import
altezza	Select a property to map Do not import
larghezza	Select a property to map Do not import

Continue →

Importare nuovi dati
nella collezione,
mappando i nuovi
dati

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

Name	Sex		
Braund, Mr. Owen Harris	male		
Cumings, Mrs. John Bradley...	female		
Heikkinen, Miss. Laina	female		
Futrelle, Mrs. Jacques Heat...	female		
Allen, Mr. William Henry	male	35	0
Moran, Mr. James	male		0
McCarthy, Mr. Timothy J	male	54	0

Import Export Clean Enrich Extract Edit New Item +

- Duplicates Finder
- Find & Replace
- Clean Spaces/Symbols/HTML
- Normalize Street Names
- Text ⇄ Date, Number, Chec...
- Ask for more

Pulizia dei DATI:
ricerca dei duplicati

Duplicates Finder

Select duplicate check method.

All Properties

Look for items with similar values for all properties. Two items are similar when all of their properties match.

Selected Properties & Multi Collections

Select the properties to be used for similarity check. Two items are similar when they have similar values on all the selected properties.

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

ass	Name	Sex		
	Braund, Mr. Owen Harris	male		
	Cumings, Mrs. John Bradley...	female		
	Heikkinen, Miss. Laina	female		
	Futrelle, Mrs. Jacques Heat...	female		
	Allen, Mr. William Henry	male	35	0
	Moran, Mr. James	male		0
	McCarthy, Mr. Timothy J	male	54	0
	Palsson, Master. Gosta Leo...	male	2	3
	Johnson, Mrs. Oscar W (Elis...	female	27	0

- Duplicates Finder
- Find & Replace
- Clean Spaces/Symbols/HTML
- Normalize Street Names
- Text ⇄ Date, Number, Chec...
- Ask for more

Pulizia dei DATI:
trova e sostituisci

Find & Replace

Add filters or select items to process a subset of items.

Options

Find: female
Replace with: donna

Search in: All Properties

Only properties with texts and numbers will be searched

Match case
If checked, the search will be case sensitive

Match using regular expression
Check our documentation to know more.

Whole words
If checked, the search will match entire words

Ignore accents
If checked, a string with accents will match pattern without accents

214 results were found on 314 items. Click 'Replace all' to replace them with 'donna'.

Results per property

314 Sex

Preview of replacements (not yet committed). Only the first 10 matched items are shown.

tes
donna

Close Replace All

Cancel Find

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

The screenshot shows the Datablist application interface. At the top, there's a toolbar with various icons: Search, Import, Export, Clean (which is highlighted), Enrich, Extract, and Edit. Below the toolbar is a table with columns: Pclass, Name, and Sex. A context menu is open over the 'Clean' button, listing options: Duplicates Finder, Find & Replace, Clean Spaces/Symbols/HTML (which is checked and highlighted with a red box), Normalize Street Names, Text ⇌ Date, Number, Check..., and Ask for more.

A modal window titled "New" is open, showing a list of items with counts: 1, 1, 0, 1, 0, 0. Below the list is a note: "Add filters or select items to process a subset of items." Underneath the list is an "Options" section with a dropdown set to "Clean data on Name". At the bottom of the modal is a "Cleaning Action" dropdown set to "Remove extra spaces".

Pulizia dei DATI:
pulisci spazi e
simboli

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

The screenshot shows a table with two columns: "Nome" (Name) and "Sex". The data includes: "Harris, Mr. Owen" (male), "Bradley, Mrs. John" (female), "Laina, Miss. Laina" (female), "Heat... trelle, Mrs. Jacques" (female), "Henry, Mr. William" (male), and "James, Mr. James" (male). A context menu is open over the table, with the "Clean" button highlighted. The menu options are: Duplicates Finder, Find & Replace, Clean Spaces/Symbols/HTML, Normalize Street Names, Text ⇌ Date, Number, Check..., and Ask for more. The "Text ⇌ Date, Number, Check..." option is underlined with a red line.

Nome	Sex
Harris, Mr. Owen	male
Bradley, Mrs. John	female
Laina, Miss. Laina	female
Heat... trelle, Mrs. Jacques	female
Henry, Mr. William	male
James, Mr. James	male

Pulizia dei DATI:
trasformare testo in
date, numeri ed altro

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

Cabin	Embarked	data Nascita	+	891 items
	S			
C85	C			
	S			
C123	S			
	S			

Aggiungo una nuova proprietà dataNascita di formato Text

iin	Embarked	data Nascita	+	891 items
	S	10/12/1899		
;	C			
	S			

Modulo 2: “Analisi dei dati con Datablist”

Gestire Proprietà

Convert data

Add filters or select items to process a subset of items.

Convert Preview

A new "data Nascita (DateTime)" property will be created to store the converted data. The original property "data Nascita" will stay unchanged.

Preview of changes (not yet committed) of your first 10 items.

data Nascita (DateTime)

data Nascita	data Nascita (DateTime)
12/10/1899, 00:00:00	12/10/1899, 00:00:00
(empty)	

Close Convert All

Converto
dataNascita in
formato data

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Uso della barra di ricerca: passeggeri di nome Elisabeth

titanic

	PassengerId	Survived	Pclass	Name	Sex	A	Show count	Filter	Import	Export	Clean	Enrich	Extract	Edit
	12	✓	1	Bonnell, Miss. Elizabeth	female	58	0							
	54	✓	2	Faunthorpe, Mrs. Lizzie (Eliz...	female	29	1							
	67	✓	2	Nye, Mrs. (Elizabeth Ramell)	female	29	0							
	80	✓	3	Dowdell, Miss. Elizabeth	female	30	0							
	162	✓	2	Watt, Mrs. James (Elizabeth...	female	40	0							
	178		1	Isham, Miss. Ann Elizabeth	female	50	0							
	273	✓	2	Mellinger, Mrs. (Elizabeth A...	female	41	0							
	338	✓	1	Burns, Miss. Elizabeth Marg...	female	41	0							
	342	✓	1	Fortune, Miss. Alice Elizabeth	female	24	3							
	497	✓	1	Eustis, Miss. Elizabeth Mus...	female	54	1							
	514	✓	1	Rothschild, Mrs. Martin (Eliz...	female	54	1							
	610	✓	1	Shutes, Miss. Elizabeth W	female	40	0							
	643		3	Skoog, Miss. Margit Elizabeth	female	2	3							
	671	✓	2	Brown, Mrs. Thomas Willia...	female	40	1							

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Selezione le righe

titanic		Click to show items count		Pclass	Name	Sex	A ↴	Show count
<input checked="" type="checkbox"/>	12	<input checked="" type="checkbox"/>	✓	1	Bonnell, Miss. Elizabeth	female	58	
<input checked="" type="checkbox"/>	54	<input checked="" type="checkbox"/>	✓	2	Faunthorpe, Mrs. Lizzie (Eliz...	female	29	
<input checked="" type="checkbox"/>	67	<input checked="" type="checkbox"/>	✓	2	Nye, Mrs. (Elizabeth Ramell)	female	29	
<input checked="" type="checkbox"/>	80	<input checked="" type="checkbox"/>	✓	3	Dowdell, Miss. Elizabeth	female	30	
<input checked="" type="checkbox"/>	162							
<input checked="" type="checkbox"/>	178							
<input checked="" type="checkbox"/>	273							
<input checked="" type="checkbox"/>	338							
<input checked="" type="checkbox"/>	342	<input checked="" type="checkbox"/>	✓					
<input checked="" type="checkbox"/>	497	<input checked="" type="checkbox"/>	✓					
<input checked="" type="checkbox"/>	514	<input checked="" type="checkbox"/>	✓					
<input checked="" type="checkbox"/>	610	<input checked="" type="checkbox"/>	✓	1	Shutes, Miss. Elizabeth W	female	40	
<input checked="" type="checkbox"/>	643			3	Skoog, Miss. Margit Elizabeth	female	2	
<input checked="" type="checkbox"/>	671	<input checked="" type="checkbox"/>	✓	2	Brown, Mrs. Thomas Willia...	female	40	
<input checked="" type="checkbox"/>	775	<input checked="" type="checkbox"/>	✓	2	Hocking, Mrs. Elizabeth (Eli...	female	54	

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

The screenshot shows a Datablist interface with a search bar at the top containing the name "Elizabeth". Below the search bar, there is a table with columns: Pclass, Name, and Sex. The table lists 15 items selected, with checkboxes next to each row. A modal dialog box titled "Export 15 items" is open in the center. The dialog has two dropdown menus: "Export Format" set to "Comma-Separated Values (.csv)" and "Column Delimiter" set to "Comma (default)". At the bottom of the dialog are "Cancel" and "Export" buttons.

Pclass	Name	Sex
1	Bonnell, Miss. Elizabeth	female
2	Faunthorpe, Mrs. Lizzie (Eliz...	female
2	Nye, Mrs. (Elizabeth Ramell)	female
3	Dowdell, Miss. Elizabeth	female
1	... (repeated)	female
2	... (repeated)	female
3	... (repeated)	female
1	... (repeated)	female
2	... (repeated)	female
3	... (repeated)	female
1	... (repeated)	female
2	... (repeated)	female
3	... (repeated)	female
1	... (repeated)	female
2	... (repeated)	female
3	Skoog, Miss. Margit Elizabeth	female

Posso esportare in
csv/excel la parte
ricercata

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Cancellare le righe cercate

The screenshot shows a Datablist interface with a search bar containing 'Elizabeth'. A red box highlights the trash can icon in the toolbar, which is used to delete selected rows. The table below shows two rows selected for deletion.

	Pclass	Name	Sex
<input checked="" type="checkbox"/>	1	Bonnell, Miss. Elizabeth	female
<input checked="" type="checkbox"/>	2	Faunthorpe, Mrs. Lizzie (Eliz...	female

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

The screenshot shows the Datablist interface. At the top, there is a search bar with the query "Elizabeth" and a dropdown menu. Below the search bar is a table with columns: Pclass, Name, and Sex. Three rows are selected, indicated by green checkmarks in the first column. A red box highlights the collection name "titanic-Elisabeth" in the sidebar under the "Data" section.

Pclass	Name	Sex
1	Bonnell, Miss. Elizabeth	female
2	Faunthorpe, Mrs. Lizzie (Eliz...	female
2	Nye, Mrs. (Elizabeth Ramell)	female

Create a new collection with the found rows

New collection name: titanic-Elisabeth

Create collection

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Search Filter icon Import Export Clean Enrich Extract Edit ...

Pclass	Name	Sex	A 891 items	⋮	New Item +
3	Braund, Mr. Owen Harris	male	22	1	
1	Cumings, Mrs. John Bradley...	female	38	1	

Filtering collection

Where Select a property * Select an operator * Add filter + Add filter group +

Cancel Apply

Uso strumento Filtro:
passeggeri di classe
3

titanic

	PassengerId	Survived	Pclass	Name	Sex
<input type="checkbox"/>	1		3	Braund, Mr. Owen Harris	male
<input type="checkbox"/>	3		3	Heikkinen, Miss. Laina	female
<input type="checkbox"/>	5		3	Allen, Mr. William Henry	male
<input type="checkbox"/>	6		3	Moran, Mr. James	male
<input type="checkbox"/>	8		3	Palsson, Master. Gosta Leo...	male
<input type="checkbox"/>	9		3	Johnson, Mrs. Oscar W (Elli...	female
<input type="checkbox"/>	11		3	Sandstrom, Miss. Margaret...	female
<input type="checkbox"/>	13		3	Saundercock, Mr. William H...	male
<input type="checkbox"/>	14		3	Andersson, Mr. Anders Johan	male
<input type="checkbox"/>	15		3	Vestrom, Miss. Hulda Aman...	female

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

The screenshot shows a table of data with columns for Name, Gender, and Show count. A dropdown menu is open at the top right, showing 'No saved filter' and 'Save filters'. The 'Save filters' option is highlighted with a red box. A modal window titled 'Save Filters' is also visible, containing a 'Filter Name' input field with 'classe 3' typed into it, and 'Cancel' and 'Save' buttons.

Name	Gender	Show count
Braund, Mr. Owen Harr		22
Heikkinen, Miss. Laina	female	26
Allen, Mr. William Henry	male	35

Save Filters

Filter Name *

classe 3

Cancel Save

Uso strumento Filtro:
passeggeri di classe
3, salvo il Filtro

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Filtering collection

Where

Select a property * Pclass

Select an operator * = Enter a value * 3

Operation and

Select a property * Survived

Select an operator * is

Add filter + Add filter group +

Cancel

Apply

Uso strumento Filtro:
passeggeri di classe
3 sopravvissuti

PassengerID	Survived	Pclass	Name	Sex	Age
1	0	3	Heikkinen, Miss. Laina	female	26
2	1	3	Johnson, Mrs. Oscar W (Elle...)	female	27
3	1	3	Sandstrom, Miss. Margaret.	female	4
4	1	3	Masseyman, Miss. Palma	female	15
5	1	3	McGowan, Miss. Anna Amelia	female	38
6	1	3	Aasland, Mrs. Carl Oscar (S...	female	30
7	1	3	O'Dwyer, Miss. Ellen Nelly	female	23
8	1	3	Givens, Miss. Mary Agatha	female	24

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Filtering collection

Where

Pclass = 3

Operation and

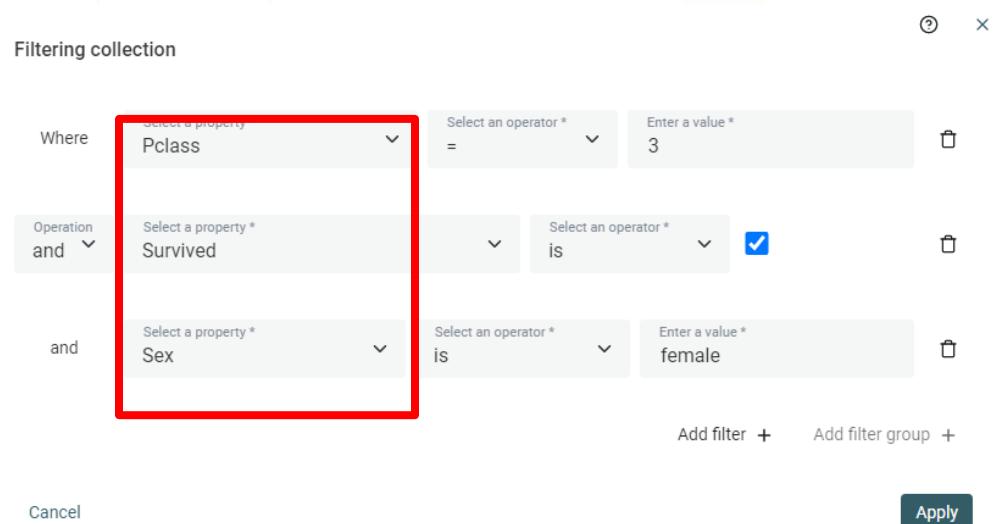
Survived is ✓

and

Sex is female

Add filter + Add filter group +

Cancel



Uso strumento Filtro:
passeggeri Donna di
classe 3
sopravvissuti

titanic

PassengerId	Survived	Pclass	Name	Sex	Age	Class
3	✓	3	Herkimer, Miss. Laina	female	26	0
9	✓	3	Johnsons, Mrs. Oscar W (Eliza)	female	27	0
11	✓	3	Sandstrom, Miss. Margaret.	female	4	1
20	✓	3	Masseimani, Mrs. Fatima	female	0	0
23	✓	3	McGowan, Miss. Anna Anna	female	15	0
26	✓	3	Asplund, Mrs. Carl Oscar (Gösta)	female	38	1
29	✓	3	O'Dwyer, Miss. Ellen "Nellie"	female	0	0
33	✓	3	Glynn, Miss. Mary Agatha	female	0	0
40	✓	3	Nicola-Warred, Miss. Jamilia	female	14	1
43	✓	2	Dorance Miss. Margaret De... (Dorance)	female	19	0
48	✓	3	O'Halloret, Miss. Bridget	female	0	0

Modulo 2: “Analisi dei dati con Datablist”

Ricercare e Filtrare i Dati

Uso strumento Filtro:
anche a partire da un
Filtro posso creare una
nuova collezione o
salvare in formato
csv/excel

The screenshot shows a Datablist interface with a search bar and various filters at the top. A red box highlights the 'Click to show items count' button, which is part of a filter panel. Below the header, there is a table of passenger data with columns for Name, Sex, Show count, and other details. The first few rows of the table are:

			Name	Sex	Show count	⋮
✓	3	✓	Heikkinen, Miss. Laina	female	26	...
✓	9	✓	Johnson, Mrs. Oscar W (Elis...	female	27	...
✓	11	✓	Sandstrom, Miss. Marguerit...	female	4	...
✓	20	✓	Masselmani, Mrs. Fatima	female
✓	23	✓	McGowan, Miss. Anna 'Anni...	female	15	...
✓	26	✓	Asplund, Mrs. Carl Oscar (S...	female	38	...
✓	29	✓	O'Dwyer, Miss. Ellen 'Nellie'	female
✓	33	✓	Glynn, Miss. Mary Agatha	female

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



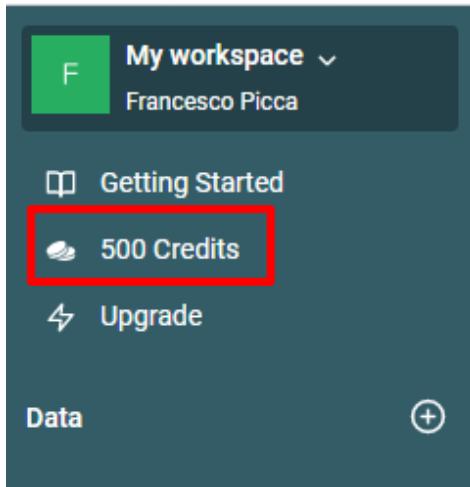
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive



Settings

Subscription

You are currently on the **Free plan**. Upgrade to collaborate with more people and access premium actions and cloud synchronization.

Plan	Cost	Includes
Free	\$0 per user per month	Current Plan Free plan includes: ✓ 1 member ✓ Only Basic Enrichments ✓ One time 500 credits
Standard	\$12 per user per month	Standard plan includes: ✓ Unlimited members ✓ Premium Enrichments ✓ 5000 credits per month ✓ Premium support

Monthly Yearly

Need Help? Contact our [sales team](#) for information about custom plans.

Il SISTEMA dei CREDITI per i servizi extra

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

Come Utilizzare i Crediti per le funzionalità Avanzate

The screenshot shows the Datablist application interface. At the top is a toolbar with the following icons from left to right: a search bar with a magnifying glass icon labeled "Search", a filter icon with a dropdown arrow, an "Import" icon (up arrow), an "Export" icon (down arrow), a "Clean" icon (trash can), an "Enrich" icon (lightning bolt, highlighted with a red box), an "Extract" icon (two overlapping documents), an "Edit" icon (pencil), and a more options menu icon (three dots). Below the toolbar is a table header row with columns for "Name", "Sex", and "Age". It displays "15 items". The table has several rows of data, with the first row partially visible. At the bottom right of the table is a "New Item +" button.

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

Enrichments

Ask for more enrichments Async Enrichments

Filter enrichments

All Companies People Translations Places AI URLs

Inputs Outputs

Premium Email Address Verification Premium 5 credits per deliverability check. Free if the verification stops on syntax or domain errors (disposable/invalid).

Complete email verification service. From syntax validation to checking the mailbox exists and can receive emails.

Email Email Status Reason Suggested Email Free Provider Role Account Domain

Free Email Address Validation Datablist Free

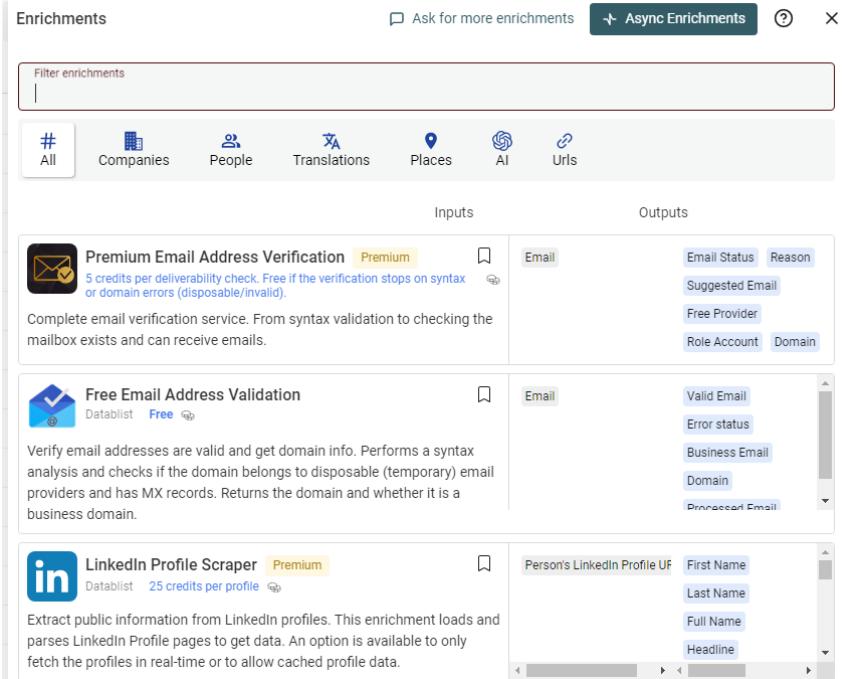
Verify email addresses are valid and get domain info. Performs a syntax analysis and checks if the domain belongs to disposable (temporary) email providers and has MX records. Returns the domain and whether it is a business domain.

Email Valid Email Error status Business Email Domain Processed Email

LinkedIn Profile Scraper Premium Datablist 25 credits per profile

Extract public information from LinkedIn profiles. This enrichment loads and parses LinkedIn Profile pages to get data. An option is available to only fetch the profiles in real-time or to allow cached profile data.

Person's LinkedIn Profile UF First Name Last Name Full Name Headline



Come Utilizzare i Crediti

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

 Free Email Address Validation Datablist Free 	 Email Valid Email Error status Business Email Domain Processed Email
---	---

Verify email addresses are valid and get domain info. Performs a syntax analysis and checks if the domain belongs to disposable (temporary) email providers and has MX records. Returns the domain and whether it is a business domain.

Il SISTEMA dei CREDITI per il SERVIZIO Free Email Address Validation

Verifica che gli indirizzi email siano validi e ottieni informazioni sul dominio. Esegue un'analisi della sintassi e controlla se il dominio appartiene a provider di posta elettronica usa e getta (temporanei)

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

	Nome	Cognome	e-mail	mail	+ Run enrichment #Free Email Address Validation
<input type="checkbox"/>	Maria	Bianchi	mariabianchi@gmail.com	mbianchi@libero.it	<input type="text" value="mrossi#libero.it"/>
<input type="checkbox"/>	Mario	Rossi	mariorossi@gmail.com		

Verify email addresses are valid and get domain info. Performs a syntax analysis and checks if the domain belongs to disposable (temporary) email providers and has MX records. Returns the domain and whether it is a business domain.
Cost: Free

[Go to action documentation](#)

Settings

Check for MX-records in email domain
The MX (Mail eXchange) record specifies the mail server responsible for accepting email messages

Input Property

Enrichments take inputs and return outputs. An input can be a property from your collection, or a custom value with variables.

Email
 Map input with existing property * mail
Keep empty to skip input.

Outputs
For each enrichment output, you can ignore it, map the result with an existing property, or store the result into a new property.

Valid Email
Checked if the syntax is valid and the email is not disposable
Map output with existing property * Valid Email-7

Error status
The reason when the email address is invalid
Map output with existing property * Error status-2

Business Email
Return True if the domain is not from a generic email provider (gmail, yahoo, etc.)
Map output with existing property * Business Email

Domain
Return the domain part of the email. After the @. Example: gmail.com
Map output with existing property * Domain

Processed Email
Checkbox flag to track if an item has been processed.
Map output with existing property * Processed Email

Il SISTEMA dei CREDITI per il SERVIZIO Free Email Address Validation

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

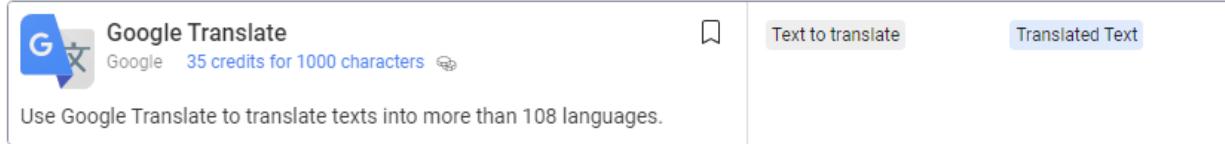
	Nome	Cognome	e-mail	mail	+
<input type="checkbox"/>	Maria	Bianchi	mariabianchi@gmail.com	mbianchi@libero.it	
<input type="checkbox"/>	Mario	Rossi	mariorossi@gmail.com	mrossi@libero.it	

Il SISTEMA dei
CREDITI per il
SERVIZIO Free Email
Address Validation

Clienti						
il	Error status	Business Email	Domain	Processed Email	Valid Email-7	3 items
			picca.com			
		libero.it				
	Invalid format					

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive



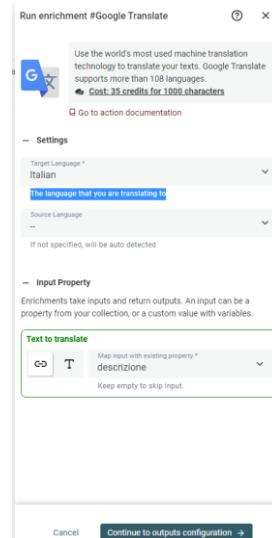
Google Translate
Google 35 credits for 1000 characters

Use Google Translate to translate texts into more than 108 languages.

Utilizza la tecnologia di traduzione automatica più utilizzata al mondo per tradurre i tuoi testi. Google Translate supporta più di 108 lingue.

Costo: 35 crediti per 1000 caratteri

Il SISTEMA dei CREDITI per i servizi extra: Google Translate



Run enrichment #Google Translate

Use the world's most used machine translation technology to translate your texts. Google Translate supports more than 108 languages.
Cost: 35 credits for 1000 characters

Go to action documentation

Settings

Target Language: Italian

The language that you are translating to

Source Language

If not specified, will be auto detected

Input Property

Enrichments take inputs and return outputs. An input can be a property from your collection, or a custom value with variables.

Text to translate

Map input with existing property: descrizione

Keep empty to skip input.

Cancel Continue to output configuration →

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

Il SISTEMA dei
CREDITI per i servizi
extra: Google
Translate

Valid Email-7	Error status-2	descrizione	+
✓			
		the customer is reliable and...	

Si suppone che ci sia un campo descrizione: “the customer is reliable and has changed address recently”, che vogliamo tradurre in Italiano

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

Run enrichment #Google Translate



Use the world's most used machine translation technology to translate your texts. Google Translate supports more than 108 languages.
Cost: 35 credits for 1000 characters

[Go to action documentation](#)

— Settings

Target Language *
Italian

The language that you are translating to

Source Language
English

If not specified, will be auto detected

— Input Property

Enrichments take inputs and return outputs. An input can be a property from your collection, or a custom value with variables.

Text to translate



Map input with existing property *
descrizione

Keep empty to skip input.

— Output

For each enrichment output, you can ignore it, map the result with an existing property, or store the result into a new property.

Translated Text

Translation from Google Traduction



Map output with existing property *
description

— Run Settings

Save options

Store options and mappings for future use.

Test on first 10 items

If checked, a first run on the first 10 items will be done. You will be able to resume on the remaining items afterward.

Existing data rule

Skip items with existing values in output properties

Cancel

Run on first 10 items

descrizione

description

3 items



the customer is reliable and... il cliente è affidabile e ha ca...

Il SISTEMA dei
CREDITI per i servizi
extra: Google
Translate

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

👉 Persone

The screenshot shows a list of names under the heading "Persone". The names are listed in a tree structure:

- Nominativo
 - Martina Stella Bianchi
 - Elisabetta Gentile
 - Mario Andrea Rossi
 - Francesco Picca

At the bottom left is a logo for "Datablist" with the text "0.05 credits per item".

Returns the gender, country, and all name parts (first name, last name, title, etc.) from a person's name.

Il SISTEMA dei CREDITI per i servizi extra: Name Parser

The screenshot shows a "Name" parser interface with the following components:

- First Name
- Middle Name
- Last Name
- Gender
- Title
- Country

At the bottom right is a "Proceed" button.

Questa funzione prende un nome e restituisce il genere, il paese in cui il nome è più comune e analizza il nome nei suoi singoli componenti (nome, secondo nome, cognome, titolo)

Costo: 0,05 crediti per articolo

Modulo 2: “Analisi dei dati con Datablist”

Funzionalità Aggiuntive

Run enrichment #Name Parser

This enrichment takes a name and returns the gender, the country where the name is the most common, and parse the name into its individual components (first name, middle name, last name, title)

 Cost: 0.05 credits per item

— Input Property

Enrichments take inputs and return outputs. An input can be a property from your collection, or a custom value with variables.

Name
Map input with existing property *
Nominativo
Keep empty to skip input.

Persons

	First Name	Middle Name	Last Name	Gender
Nominativo	"First Name" output	"Middle Name" output	"Last Name" output	"Gender" output
Martina Stella Bianchi	Martina	Stella	Bianchi	Female
Elisabetta Gentile	Elisabetta		Gentile	Female
Mario Andrea Rossi	Mario	Andrea	Rossi	Male
Francesco Picca	Francesco		Picca	Male

Run enrichment #Name Parser

Middle Name
Return the Middle Name part from the name
Create new property and store output value

Last Name
Return the Last Name part from the name
Create new property and store output value

Gender
Return Male or Female
Create new property and store output value

Title
Return Title part from the name. For example Dr., Jr., Mr., Mme, etc.
Do not import

Country
Return the country where the name is the most common
Create new property and store output value

Processed
Checkbox flag to track if an item has been processed.
Do not import

Cancel Run on first 10 items

Il SISTEMA dei CREDITI per i servizi extra: Name Parser

Persons

	First Name	Middle Name	Last Name	Gender	Country
Nominativo					4 items
Martina Stella Bianchi	Martina	Stella	Bianchi	Female	Italy
Elisabetta Gentile	Elisabetta		Gentile	Female	Italy
Mario Andrea Rossi	Mario	Andrea	Rossi	Male	Italy
Francesco Picca	Francesco		Picca	Male	Italy

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 3: “Modelli di Apprendimento”

Come Impara una Macchina?



Modulo 3: “Modelli di Apprendimento”

Come Impara una Macchina?

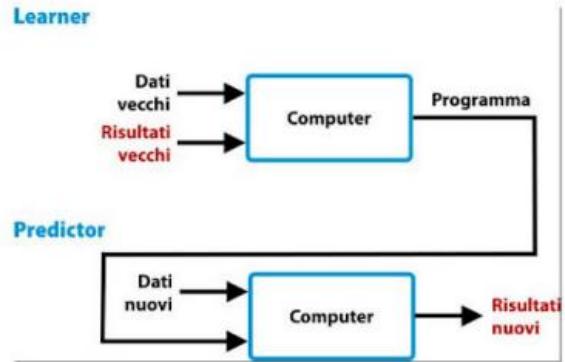
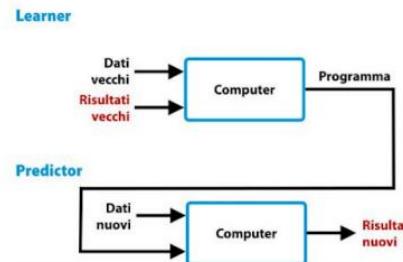


Immagine tratta da Big Data Analytics (Andrea De Mauro - APOGEO)

Modulo 3: “Modelli di Apprendimento”

L'immagine è suddivisa in tre fasi:

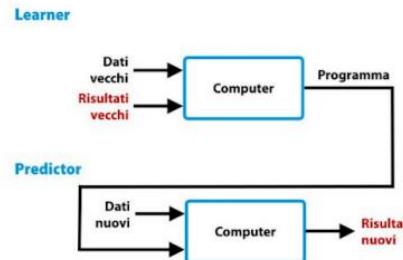
- **Fase di apprendimento:** In questa fase, il computer viene addestrato su un insieme di dati di input. I dati di input sono etichettati, il che significa che ogni dato è associato a un valore di output specifico. Il computer utilizza questi dati etichettati per imparare a fare previsioni sui dati di input.
- **Fase di previsione:** In questa fase, il computer viene utilizzato per fare previsioni su nuovi dati di input. I dati di input nuovi non sono etichettati, quindi il computer deve utilizzare il modello che ha imparato nella fase di apprendimento per fare previsioni.
- **Fase di valutazione:** In questa fase, le previsioni del computer vengono valutate. Le previsioni vengono confrontate con i valori di output reali per determinare l'accuratezza del modello.



Modulo 3: “Modelli di Apprendimento”

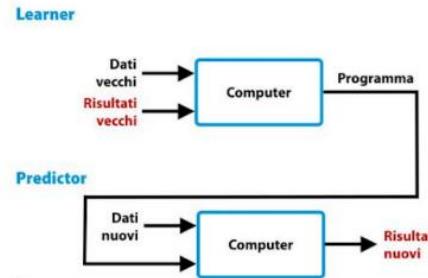
L'immagine mostra i **seguenti elementi**:

- **Dati vecchi:** I dati vecchi sono i dati di input utilizzati per addestrare il computer.
- **Risultati vecchi:** I risultati vecchi sono i valori di output associati ai dati vecchi.



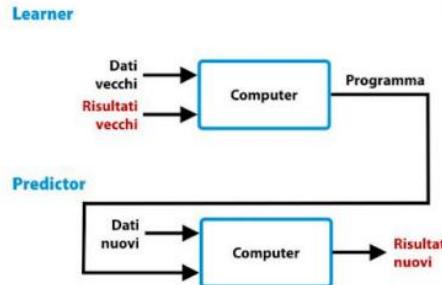
Modulo 3: “Modelli di Apprendimento”

- **Programma:** Il programma è il modello di apprendimento automatico che il computer ha imparato.
- **Computer:** Il computer è l'hardware che viene utilizzato per eseguire il processo di apprendimento automatico.
- **Predictor:** Il predictor è il componente del computer che viene utilizzato per fare previsioni sui dati di input nuovi.



Modulo 3: “Modelli di Apprendimento”

- **Dati nuovi:** I dati nuovi sono i dati di input su cui il computer viene utilizzato per fare previsioni.
- **Risultati nuovi:** I risultati nuovi sono le previsioni fatte dal computer sui dati nuovi.



Modulo 3: “Modelli di Apprendimento”

Definizione generica:

Un modello di apprendimento è un algoritmo o una struttura dati che può apprendere da esperienze passate (dati) per migliorare le sue prestazioni su compiti specifici.



Modulo 3: “Modelli di Apprendimento”

Esempi di modelli di apprendimento:

Regessione Lineare:

- *Descrizione:* Un modello di regressione lineare cerca di trovare la relazione lineare tra una variabile dipendente e una o più variabili indipendenti.
- *Esempio:* Prevedere il prezzo di una casa in base a parametri come la superficie, il numero di camere, ecc.

Modulo 3: “Modelli di Apprendimento”

Reti Neurali:

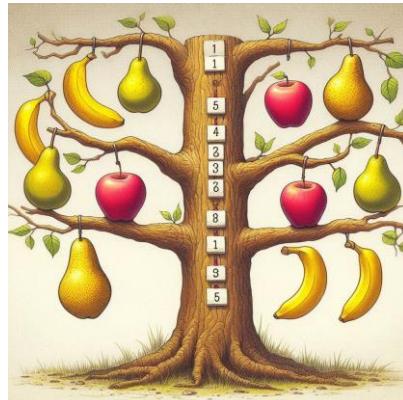
- *Descrizione:* Le reti neurali sono modelli ispirati al funzionamento del cervello umano, composti da strati di neuroni artificiali collegati tra loro.
- *Esempio:* Riconoscimento di immagini, dove la rete apprende a identificare oggetti o pattern nelle immagini.



Modulo 3: “Modelli di Apprendimento”

Alberi Decisionali:

- *Descrizione:* Gli alberi decisionali suddividono iterativamente il dataset in base a condizioni specifiche, creando una struttura ad albero di decisioni.
- *Esempio:* Classificazione di pazienti in base a sintomi per diagnosticare una malattia.

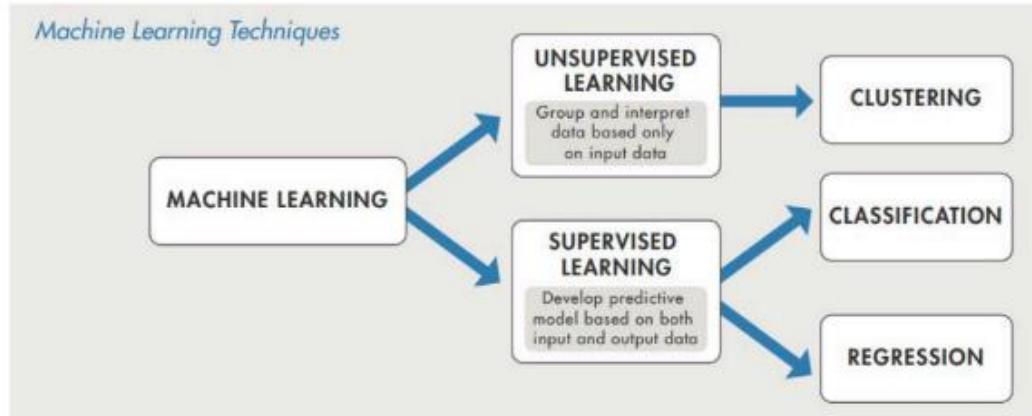


Modulo 3: “Modelli di Apprendimento”

Apprendimento Profondo (Deep Learning):

- *Descrizione:* Una sotto-categoria di apprendimento automatico che utilizza reti neurali profonde con molteplici strati nascosti.
- *Esempio:* Riconoscimento vocale, dove il modello apprende a convertire il parlato in testo.

Modulo 3: “Modelli di Apprendimento”



Differenza-apprendimento-supervisionato-e-non-supervisionato-nel-machine-learning

<https://vitolavecchia.altervista.org/differenza-apprendimento-supervisionato-e-non-supervisionato-nel-machine-learning/>

Modulo 3: “Modelli di Apprendimento”

Apprendimento supervisionato

L'apprendimento supervisionato è un tipo di apprendimento automatico in cui il modello viene addestrato su un insieme di dati che include sia dati di input che dati di output. I dati di input sono le caratteristiche dell'oggetto o del fenomeno che si sta cercando di comprendere, mentre i dati di output sono le classi o le categorie a cui apparterrebbe l'oggetto o il fenomeno.

Modulo 3: “Modelli di Apprendimento”

Le tecniche di apprendimento supervisionato possono essere utilizzate per risolvere una varietà di problemi, tra cui:

- **Classificazione:** l'assegnazione di un'etichetta a un oggetto o a un fenomeno sulla base delle sue caratteristiche. Ad esempio, un modello di apprendimento supervisionato può essere utilizzato per classificare le immagini di animali in base alla loro specie.

Modulo 3: “Modelli di Apprendimento”

- **Regressione:** la previsione del valore di una variabile sulla base dei valori di altre variabili. Ad esempio, un modello di apprendimento supervisionato può essere utilizzato per prevedere il prezzo di un'azione sulla base dei dati storici di prezzo.

Modulo 3: “Modelli di Apprendimento”

Apprendimento non supervisionato

L'apprendimento non supervisionato è un tipo di apprendimento automatico in cui il modello viene addestrato su un insieme di dati che include solo dati di input. Il modello deve imparare a trovare le relazioni tra i dati di input senza l'aiuto di dati di output.

Modulo 3: “Modelli di Apprendimento”

Le tecniche di apprendimento non supervisionato possono essere utilizzate per risolvere una varietà di problemi, tra cui:

- **Raggruppamento:** l'aggregazione di oggetti o fenomeni simili in gruppi. Ad esempio, un modello di apprendimento non supervisionato può essere utilizzato per raggruppare i clienti in base ai loro interessi.

Modulo 3: “Modelli di Apprendimento”

- **Riduzione della dimensionalità:** la riduzione del numero di caratteristiche di un insieme di dati senza perdere informazioni importanti.
- **Rilevamento di anomalie:** l'identificazione di oggetti o fenomeni che si discostano dalle tendenze generali. Ad esempio, un modello di apprendimento non supervisionato può essere utilizzato per identificare transazioni fraudolente.

Modulo 3: “Modelli di Apprendimento”

L'apprendimento per rinforzo è un tipo di apprendimento automatico in cui il modello viene addestrato sperimentando il mondo e ricevendo feedback sulle sue azioni. Il modello deve imparare a prendere le decisioni migliori per raggiungere un obiettivo, senza avere a disposizione dati di input o di output predefiniti.



Modulo 3: “Modelli di Apprendimento”

Le tecniche di apprendimento per rinforzo possono essere utilizzate per risolvere una varietà di problemi, tra cui:

- Controllo di sistemi complessi: l'apprendimento per rinforzo può essere utilizzato per controllare sistemi complessi, come robot o veicoli autonomi.
- Gioco: l'apprendimento per rinforzo può essere utilizzato per giocare a giochi, come scacchi o poker.
- Ottimizzazione: l'apprendimento per rinforzo può essere utilizzato per trovare la soluzione ottimale a un problema, come il percorso più breve tra due punti.

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 3: “Classificazione e Regressione”



Bing Image Creator

La classificazione è una delle attività più comuni nell'apprendimento automatico. Si tratta di un processo che consiste nell'assegnare un'etichetta a un'istanza, sulla base di un insieme di caratteristiche.

Modulo 3: “Classificazione e Regressione”

Nella classificazione, gli output sono divisi in due o più classi e il sistema di apprendimento deve produrre un modello che assegna gli input non ancora visti a una o più di queste. Questo viene affrontato solitamente in maniera supervisionata, in cui il modello viene fornito di un insieme di dati di addestramento, costituito da istanze con le relative etichette.

Modulo 3: “Classificazione e Regressione”

Gli algoritmi di classificazione possono essere classificati in base a diversi criteri, tra cui:

- **Metodo di apprendimento:** gli algoritmi di classificazione possono essere classificati in base al metodo di apprendimento utilizzato. I metodi di apprendimento più comuni sono l'apprendimento supervisionato, l'apprendimento non supervisionato e l'apprendimento semi-supervisionato.
- **Tipo di modello:** gli algoritmi di classificazione possono essere classificati in base al tipo di modello che generano. I modelli di classificazione più comuni sono i modelli lineari, i modelli non lineari e i modelli probabilistici.

Modulo 3: “Classificazione e Regressione”

Alcuni esempi di algoritmi di classificazione sono:

- **Regressione logistica:** è un algoritmo di apprendimento supervisionato che produce un modello lineare per la classificazione binaria.
- **Support Vector Machine:** è un algoritmo di apprendimento supervisionato che produce un modello non lineare per la classificazione binaria e multiclasse.
- **K-nearest neighbors:** è un algoritmo di apprendimento supervisionato che assegna un'istanza alla classe più comune tra le istanze più vicine.
- **Decision Tree:** è un algoritmo di apprendimento supervisionato che crea un albero decisionale per la classificazione.
- **Random Forest:** è un algoritmo di apprendimento supervisionato che crea un insieme di alberi decisionali per la classificazione.

Modulo 3: “Classificazione e Regressione”

La classificazione è una tecnica molto versatile che può essere utilizzata in una vasta gamma di applicazioni, tra cui:

- **Rilevazione di frodi:** gli algoritmi di classificazione possono essere utilizzati per identificare le transazioni fraudolente.
- **Riconoscimento di immagini:** gli algoritmi di classificazione possono essere utilizzati per identificare gli oggetti in un'immagine.
- **Riconoscimento vocale:** gli algoritmi di classificazione possono essere utilizzati per identificare le parole in un discorso.
- **Filtraggio dei contenuti:** gli algoritmi di classificazione possono essere utilizzati per filtrare i contenuti indesiderati.

Modulo 3: “Classificazione e Regressione”

Nella classificazione, la **colonna target** è la colonna che contiene l'etichetta che si desidera predire. Le **feature** sono le colonne che contengono le caratteristiche che vengono utilizzate per predire l'etichetta.

Ad esempio, se si sta tentando di classificare i fiori in base al loro colore, la colonna target potrebbe contenere le etichette "rosso", "blu", "giallo" e "verde". Le feature potrebbero contenere le colonne "lunghezza dei petali", "larghezza dei petali" e "colore del gambo".



Modulo 3: “Classificazione e Regressione”

L'algoritmo di classificazione viene addestrato su un insieme di dati di addestramento, costituito da istanze con le relative etichette. L'algoritmo utilizza le feature per imparare a predire l'etichetta corretta per nuove istanze.



I Fiori Iris Van Gogh (Rivisitazione di Bing Image Creator)

Modulo 3: “Classificazione e Regressione”

La scelta delle feature è un aspetto importante della classificazione. Le feature devono essere rilevanti per l'etichetta che si desidera predire. Se le feature non sono rilevanti, l'algoritmo di classificazione non sarà in grado di imparare a predire l'etichetta con precisione.

Modulo 3: “Classificazione e Regressione”

Ecco alcuni esempi di feature che possono essere utilizzate per classificare i fiori:

- Caratteristiche fisiche: lunghezza dei petali, larghezza dei petali, colore dei petali, colore del gambo, forma dei petali, numero di petali, ecc.
- Caratteristiche chimiche: composizione chimica dei petali, profumo, ecc.
- Caratteristiche biologiche: specie, genere, famiglia, ecc.

Modulo 3: “Classificazione e Regressione”

Il dataset Iris è un dataset multivariato introdotto da Ronald Fisher nel 1936. Consiste in 150 istanze di Iris misurate da Edgar Anderson e classificate secondo tre specie: Iris setosa, Iris virginica e Iris versicolor. Le quattro variabili considerate sono la lunghezza e la larghezza del sepalo e del petalo.

Il dataset Iris è uno dei dataset più utilizzati nell'apprendimento automatico. È un dataset relativamente semplice, ma è molto utile per imparare i principi di base della classificazione.



Wikimedia Commons

Modulo 3: “Classificazione e Regressione”

TRAINING E TESTING SET & K-fold cross validation



Bing Image Creator

Modulo 3: “Classificazione e Regressione”

La **scissione dei dati di addestramento/test** è un processo che consiste nel dividere un insieme di dati in due parti:

un **set di addestramento** e un **set di test**. Il set di addestramento viene utilizzato per addestrare l'algoritmo di classificazione, mentre il set di test viene utilizzato per valutare l'accuratezza dell'algoritmo.

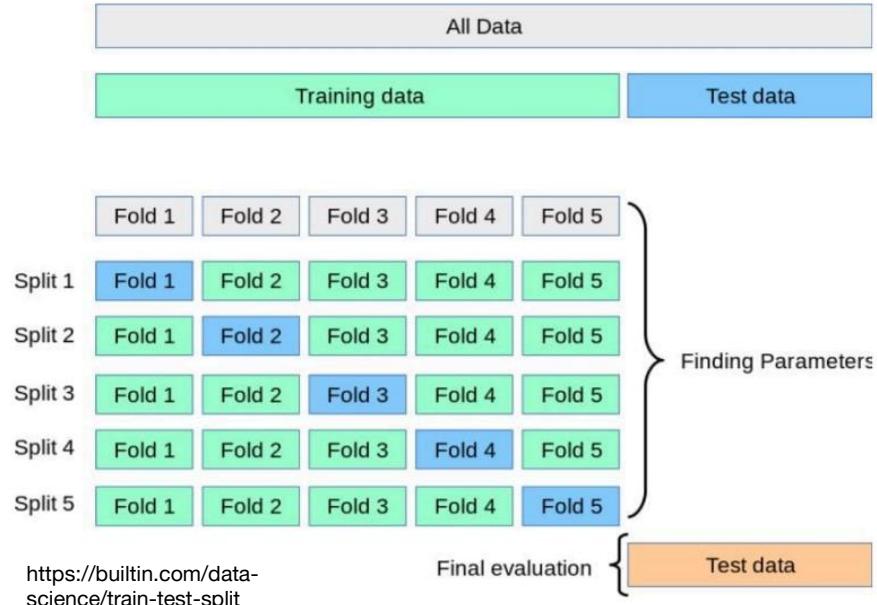
Modulo 3: “Classificazione e Regressione”

La scissione dei dati di addestramento/test è importante per evitare il sovradattamento (overfitting). Il sovradattamento si verifica quando l'algoritmo di classificazione si adatta troppo bene ai dati di addestramento e non è in grado di generalizzare a nuovi dati.

La scissione dei dati di addestramento/test viene solitamente eseguita in modo casuale. Ciò significa che ogni istanza ha la stessa probabilità di essere assegnata al set di addestramento o al set di test.

Modulo 3: “Classificazione e Regressione”

K-fold cross validation.



Modulo 3: “Classificazione e Regressione”

La **K-fold cross validation** è una tecnica di valutazione di un modello di machine learning che consiste nel dividere il dataset in K sottoinsiemi (fold) e poi addestrare il modello su K-1 fold e valutarlo sul fold rimanente. Questo viene ripetuto K volte, utilizzando ogni fold come set di valutazione una volta.

L'accuratezza media del modello sui K fold viene utilizzata come stima dell'accuratezza complessiva del modello. La K-fold cross validation è una tecnica di valutazione molto efficace, poiché consente di valutare il modello su diversi sottoinsiemi del dataset, evitando così l'overfitting.

Modulo 3: “Classificazione e Regressione”

La K-fold cross validation è una tecnica molto versatile che può essere utilizzata con diversi tipi di modelli di machine learning, tra cui la regressione logistica, le support vector machine, i decision tree e i random forest.

Modulo 3: “Classificazione e Regressione”

Regressione Lineare

La regressione lineare è una tecnica statistica utilizzata per studiare la relazione tra una variabile dipendente (o risposta) e una o più variabili indipendenti (o predittori). L'obiettivo è trovare la migliore retta (o piano in caso di più variabili indipendenti) che rappresenti al meglio la relazione tra le variabili.

La forma generale di un modello di regressione lineare semplice è:

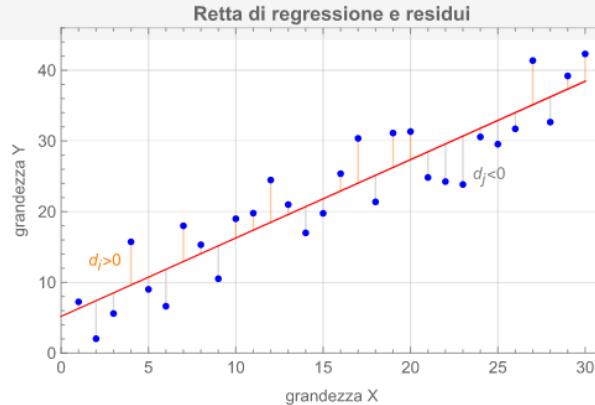
$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Dove:

- Y è la variabile dipendente.
- X è la variabile indipendente.
- β_0 è l'intercetta della retta (il valore di Y quando X è zero).
- β_1 è il coefficiente angolare (la variazione in Y per una variazione unitaria in X).
- ε rappresenta l'errore residuo, ovvero la differenza tra il valore reale osservato e il valore previsto dal modello.

Modulo 3: “Classificazione e Regressione”

L'obiettivo della regressione lineare è trovare i valori ottimali di β_0 e β_1 in modo che il modello si adatti meglio possibile ai dati osservati. Questo è spesso fatto minimizzando la somma dei quadrati degli errori residui (SSE), che è la somma dei quadrati delle differenze tra i valori osservati e quelli previsti dal modello.



Modulo 3: “Classificazione e Regressione”

Regressione Lineare con più variabili indipendenti

La regressione lineare può essere estesa a modelli multipli, in cui sono presenti più variabili indipendenti. In questo caso, il modello prende la forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Dove n è il numero di variabili indipendenti nel modello.

Modulo 3: “Classificazione e Regressione”

Regressione Polinomiale

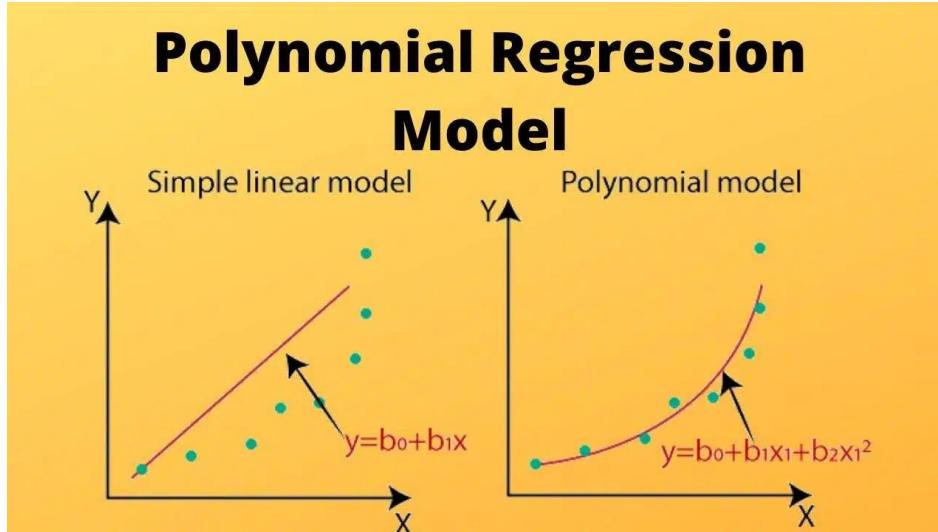
La regressione polinomiale è una tecnica di modellizzazione che estende la regressione lineare introducendo termini polinomiali nelle variabili indipendenti. Il modello di regressione polinomiale di grado n ha la forma:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_n X^n + \varepsilon$$

dove:

- Y è la variabile dipendente,
- X è la variabile indipendente,
- $\beta_0, \beta_1, \dots, \beta_n$ sono i coefficienti del polinomio,
- ε è l'errore residuo.

Modulo 3: “Classificazione e Regressione”



<https://fineproxy.org/it/wiki/polynomial-regression/>

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



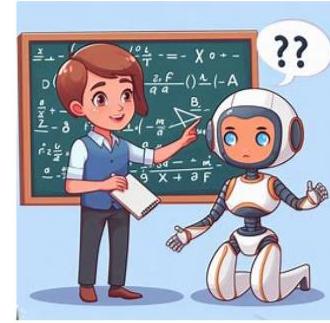
I.T.T. Pacinotti
una scuola per il futuro

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

ESEMPIO 1

Immagina di dover fare un tiro al bersaglio con una freccia. Supponiamo che il centro del bersaglio sia il risultato accurato che vogliamo ottenere con il nostro modello, mentre il punto in cui atterra la freccia rappresenta la previsione del nostro modello.

Bias (Errore sistematico): Se il tuo tiro al bersaglio è costantemente lontano dal centro (il risultato corretto), hai un alto bias. In termini di modelli, un alto bias si verifica quando il modello è troppo semplificato per catturare la complessità dei dati.



Bing Image Creator

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Varianza (Sensibilità alle fluttuazioni): Se, invece, i tuoi colpi sono molto sparsi e alcuni sono vicini al centro mentre altri sono lontani, hai un'elevata varianza. In termini di modelli, l'alta varianza si verifica quando il modello è troppo complesso e si adatta molto ai dati di addestramento, ma non generalizza bene a nuovi dati. Immagina di avere molte frecce di dimensioni diverse: alcune potrebbero avvicinarsi al centro, ma la dispersione è alta.

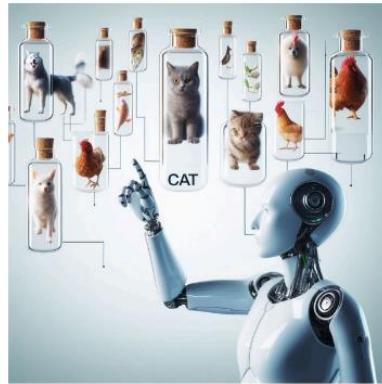


Bing Image Creator

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

ESEMPIO 2

Immagina di dover insegnare a un robot a riconoscere gli animali guardando delle immagini.



Bing Image Creator

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

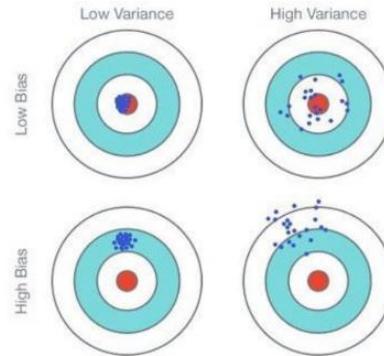
- **Bias (Errore sistematico):** Se il robot semplificasse troppo le cose e decidesse che tutti gli animali sono cani, avrebbe un alto bias.

Indipendentemente dall'immagine di un elefante o di un pesce, il robot insiste nel chiamarli cani.

- **Varianza (Sensibilità alle fluttuazioni):** Al contrario, se il robot si "iper addestra" memorizzando ogni dettaglio di ogni animale nelle immagini di addestramento, potrebbe non essere in grado di riconoscere correttamente nuove immagini di animali, mostrando così una varianza elevata.

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

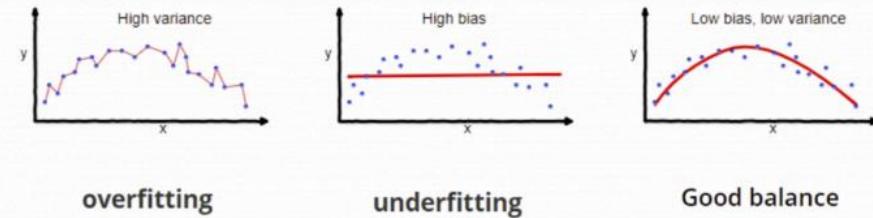
Il **compromesso bias-varianza** qui sarebbe insegnare al robot a riconoscere gli animali in modo abbastanza generale, catturando le caratteristiche comuni senza focalizzarsi su dettagli specifici che potrebbero essere unici solo nei dati di addestramento. In questo modo, il robot avrebbe un buon equilibrio tra la capacità di riconoscere correttamente nuovi animali e la capacità di adattarsi a una varietà di casi.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

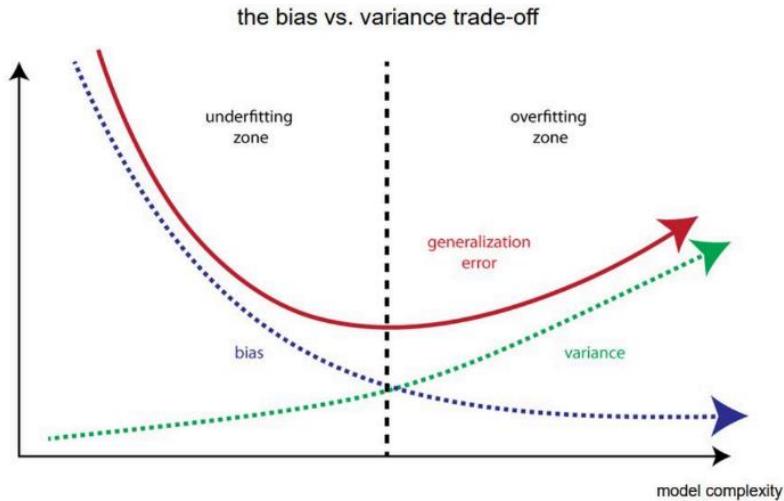
Il **Bias / Variance Tradeoff** è, quindi, un importante concetto da tenere a mente durante la creazione di un modello di machine learning.

Comprenderlo ci aiuterà non solo a creare modelli più precisi ma anche a evitare di imbatterci in *underfitting* e *overfitting*, due avversari temibili.¹



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

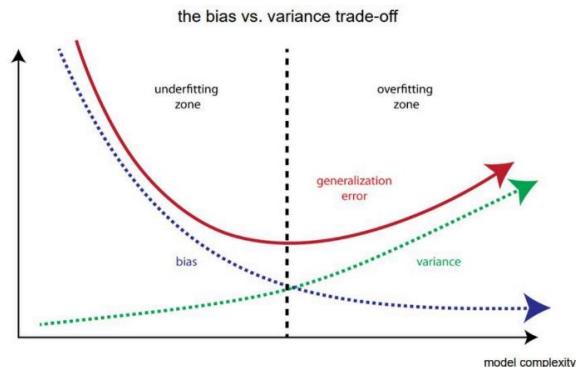
Curva di Bias: Inizia con un errore elevato in quanto il modello è troppo semplice per catturare la complessità nei dati. L'errore di bias diminuisce man mano che la complessità del modello aumenta.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Curva di Varianza: Inizia con un errore basso in quanto il modello si adatta bene ai dati di addestramento. L'errore di varianza aumenta man mano che la complessità del modello cresce, poiché il modello inizia ad adattarsi anche al rumore nei dati.

Il punto in cui queste due curve si incontrano rappresenta il punto ottimale in termini di trade-off bias-varianza. In questo punto, l'errore complessivo è minimo, indicando che il modello è sufficientemente complesso da catturare la struttura nei dati senza adattarsi eccessivamente al rumore. Questo è il punto in cui si cerca di posizionare il modello per ottenere le migliori prestazioni su dati non visti.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Errori nella Regressione



Bing Image Creator

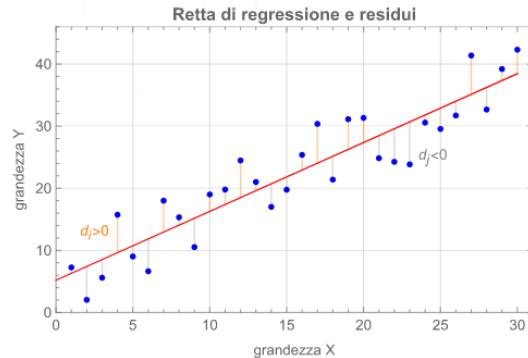
Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

R^2

In termini statistici esso è dato dalla frazione della varianza campionaria di y_i predetta dai regressori x_i .

In formule matematiche l'R-squared può assumere valori compresi fra 0 e 1. Se è pari a 1 allora esiste una perfetta relazione lineare fra il fenomeno analizzato e la sua retta di regressione.

Se pari a 0 non esiste alcuna relazione lineare fra le due variabili mentre i valori compresi fra 0 e 1 forniscono un'indicazione sulla efficacia della retta di regressione di sintetizzare l'oggetto dell'analisi.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

La sua formula è la seguente:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$$

rappresenta la devianza spiegata dal modello
(Explained Sum of Squares)

$$TSS = \sum_{i=1}^n (y_i - \bar{y})^2$$

rappresenta la devianza totale (total Sum of Squares)
e

$$RSS = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

rappresenta la devianza residua (Residual
Sum of Squares)

y_i sono i dati osservati;

\bar{y} è la loro media;

\hat{y}_i sono i dati stimati dal modello ottenuto dalla regressione.

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Errori nella Classificazione (Matrice di Confusione)

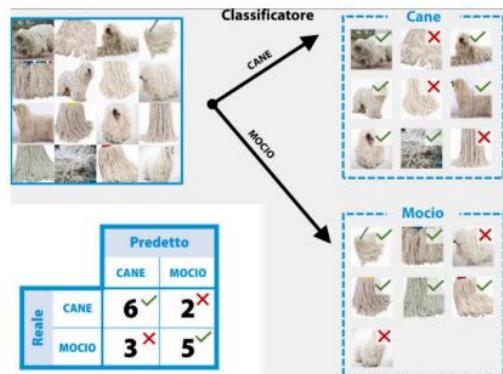
Errore nella Classificazione (Matrice di Confusione)



Bing Image Creator (dove metto questo gattino?!?)

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Errori nella Classificazione (Matrice di Confusione)



Big Data Analytics (A. DE MAURO - APOGEO)

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

La matrice di confusione è una tabella che mostra le prestazioni di un modello di classificazione su un set di dati, organizzando le previsioni in base alle classi reali.

Spesso il Modello commette ERRORI!!!



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Dove:

TP (True Positive): Questi sono i casi in cui il modello ha predetto correttamente la classe positiva. È una previsione corretta.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

TN (True Negative): Questi sono i casi in cui il modello ha predetto correttamente la classe negativa. È una previsione corretta.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

FP (False Positive): Questi sono i casi in cui il modello ha predetto erroneamente la classe positiva quando la classe reale era negativa. È un falso allarme.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

FN (False Negative): Questi sono i casi in cui il modello ha predetto erroneamente la classe negativa quando la classe reale era positiva. È una mancata previsione.



Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Con queste informazioni, è possibile calcolare diverse metriche di valutazione della classificazione, tra cui:

Accuracy (Accuratezza):

Definizione: L'accuratezza è la percentuale totale di istanze correttamente classificate dal modello.

Formula: Accuratezza = $(TP + TN) / (TP + TN + FP + FN)$

		PREDICTED VALUES	
		Positive (1)	Negative (0)
ACTUAL VALUES	Positive (1)	6	1
	Negative (0)	2	11
		TRUE POSITIVE	FALSE NEGATIVE
		FALSE POSITIVE	TRUE NEGATIVE

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Precision (Precisione):

“La precisione è un parametro utile nei casi in cui i falsi positivi costituiscono una preoccupazione maggiore rispetto ai falsi negativi ”

Definizione: La precisione misura la frazione di istanze positive predette dal modello che sono effettivamente positive.

È una misura di correttezza che si ottiene nella previsione vera.

In parole semplici, ci dice quanti pronostici sono effettivamente positivi su tutto il totale positivo previsto.

		PREDICTED VALUES	
		Positive (1)	Negative (0)
ACTUAL VALUES	Positive (1)	6 TRUE POSITIVE	1 FALSE NEGATIVE
	Negative (0)	2 FALSE POSITIVE	11 TRUE NEGATIVE

↑
Precision

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{\text{Predictions Actually Positive}}{\text{Total Predicted positive}} = \frac{6}{6 + 2} = 0.75$$

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

Recall (Sensibilità): "Recall è una metrica utile nei casi in cui il falso negativo prevale sul falso positivo "

Definizione: Il recupero rappresenta la frazione di istanze positive effettive che sono state predette correttamente dal modello.

Formula: Recupero = TP / (TP + FN)

		PREDICTED VALUES	
		Positive (1)	Negative (0)
ACTUAL VALUES	Positive (1)	6 TRUE POSITIVE	1 FALSE NEGATIVE
	Negative (0)	2 FALSE POSITIVE	11 TRUE NEGATIVE

← Recall

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{\text{Predictions Actually Positive}}{\text{Total Actual positive}} = \frac{6}{6 + 1} = 0.85$$

Modulo 3: “Bias e Varianza - Errori (classificazione e regressione)”

F1-Score:

Definizione: L'F1-Score è la media armonica tra precisione e recupero.

$$\text{F1-Score} = 2 * \frac{(\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$$

Il punteggio F1 è un numero compreso tra 0 e 1 ed è la media armonica di precisione e richiamo. Usiamo la media armonica perché non è sensibile a valori estremamente grandi, a differenza delle medie semplici.

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



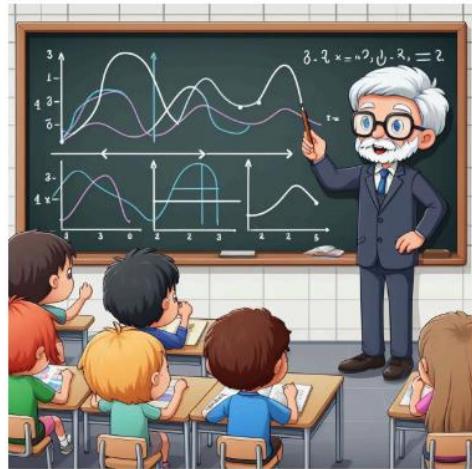
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Regressione Logistica



Bing Image Creator

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

La regressione logistica è una tecnica di analisi dei dati che viene utilizzata per prevedere la probabilità che un evento accada. L'evento in questione può essere rappresentato da una variabile dipendente dicotomica, cioè che può assumere solo due valori, ad esempio "sì" o "no", "positivo" o "negativo", "morts" o "vivo".

La regressione logistica è un metodo potente e versatile che può essere utilizzato in una vasta gamma di applicazioni. Alcuni esempi includono:

- **Rilevamento delle frodi:** la regressione logistica può essere utilizzata per identificare le transazioni fraudolente in base a fattori quali la quantità della transazione, il tipo di carta di credito utilizzata e la posizione della transazione.
- **Marketing:** la regressione logistica può essere utilizzata per prevedere la probabilità che un cliente acquisti un prodotto o un servizio in base a fattori quali il reddito del cliente, i suoi interessi e le sue abitudini di acquisto.
- **Medicina:** la regressione logistica può essere utilizzata per prevedere la probabilità che un paziente sviluppi una malattia in base a fattori quali la sua età, il suo sesso e la sua storia familiare.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Ecco una spiegazione passo per passo della regressione logistica:

1. Definizione del problema:

- La regressione logistica è spesso utilizzata quando si desidera prevedere la probabilità che un'osservazione appartenga a una determinata categoria.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

2. Variabili coinvolte:

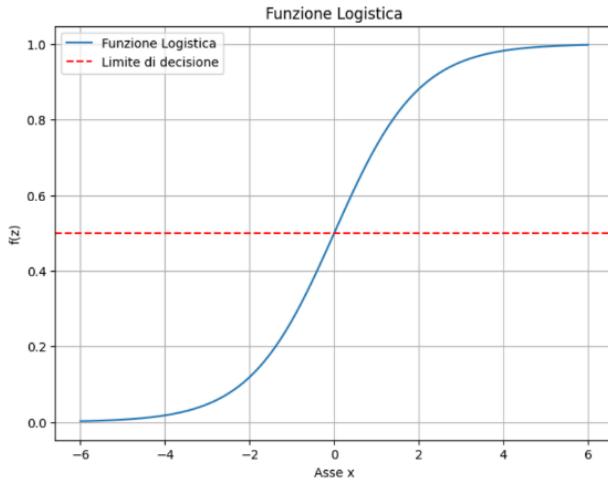
- Variabile dipendente (risposta): è la variabile che si sta cercando di prevedere. È binaria (0 o 1).
- Variabili indipendenti (predittori): sono le variabili che vengono utilizzate per prevedere la variabile dipendente. Possono essere continue o categoriche.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

3. Funzione logistica:

- La regressione logistica utilizza una funzione logistica (o sigmoide) per trasformare una combinazione lineare delle variabili indipendenti in una probabilità compresa tra 0 e 1.
- La formula della funzione logistica è: $P(Y = 1) = \frac{1}{1+e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}$
 - $P(Y = 1)$ è la probabilità che la variabile dipendente sia 1.
 - $\beta_0, \beta_1, \dots, \beta_k$ sono i coefficienti da stimare.
 - X_1, X_2, \dots, X_k sono i valori delle variabili indipendenti.
 - e è la base del logaritmo naturale (euleriano).

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”



4. Stima dei coefficienti:

- I coefficienti $\beta_0, \beta_1, \dots, \beta_k$ vengono stimati utilizzando metodi come la massima verosimiglianza.

5. Decision Boundary:

- In un problema di classificazione binaria, viene definito un valore di soglia (tipicamente 0.5). Se la probabilità calcolata supera questo valore, la previsione è 1; altrimenti, la previsione è 0.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

6. Valutazione del modello:

- Il modello viene valutato utilizzando metriche come l'accuratezza, la precisione, il richiamo e l'area sotto la curva ROC, a seconda dei requisiti specifici del problema.

Come si stimano i parametri?

- I coefficienti $\beta_0, \beta_1, \dots, \beta_k$ vengono stimati utilizzando metodi come la massima verosimiglianza. *Stima dei coefficienti:*

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

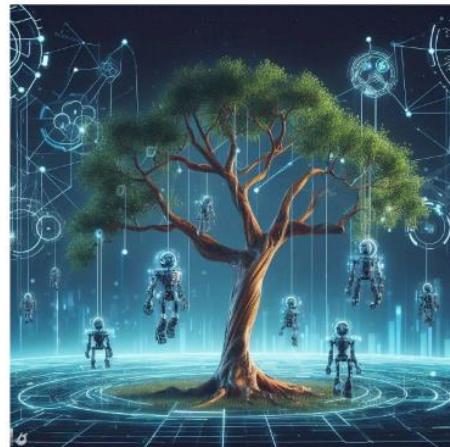
La massima verosimiglianza (Maximum Likelihood, ML) è un metodo statistico utilizzato per stimare i parametri di un modello matematico in modo che la probabilità di osservare i dati osservati sia massimizzata. In altre parole, si cerca di trovare i valori dei parametri del modello che rendono più probabile l'insieme di dati osservato.

[Approfondimento WEB](#)



Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Decision Tree



Bing Image Creator

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Un **albero decisionale** (Decision Tree) è un modello di apprendimento automatico che rappresenta una struttura ad albero in cui ogni nodo interno rappresenta una decisione basata su un attributo, ogni ramo rappresenta l'esito di quella decisione, e ogni foglia rappresenta l'etichetta di classe risultante. Questi alberi possono essere utilizzati sia per problemi di classificazione che di regressione.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Ecco un'esposizione semplificata dei concetti chiave:

Nodi e Rami:

- Nodo Radice: Il primo nodo dell'albero, rappresenta la decisione iniziale o la caratteristica più importante per la divisione.
- Nodi Interni: Gli altri nodi dell'albero, rappresentano decisioni basate su attributi.
- Foglie: Nodi finali dell'albero, rappresentano le etichette di classe o i valori di previsione.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Divisioni:

- Ad ogni nodo interno, l'albero decide quale attributo utilizzare per la divisione del dataset. Questa decisione si basa sulla caratteristica che meglio separa le classi o riduce la varianza nel caso di problemi di regressione.

Criteri di Divisione:

- I criteri comuni per la divisione includono la Gini impurity, l'entropia e l'errore quadratico medio (per problemi di regressione).

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Costruzione:

- L'albero viene costruito ricorsivamente. Inizialmente, si seleziona la caratteristica migliore per dividere il dataset. Questo processo continua fino a quando una condizione di stop è soddisfatta (ad esempio, una profondità massima è raggiunta o il numero minimo di campioni in una foglia è raggiunto).

Overfitting:

- Gli alberi decisionali possono essere suscettibili all'overfitting, cioè possono adattarsi troppo ai dati di addestramento, perdendo la capacità di generalizzare bene su nuovi dati. Strategie come la potatura (pruning) degli alberi possono essere utilizzate per mitigare questo problema.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Vantaggi:

- Interpretabilità: Gli alberi decisionali possono essere facilmente interpretati e visualizzati.
- Adattabilità: Possono gestire dati sia categorici che numerici.
- Non richiedono normalizzazione dei dati.

Svantaggi:

- Sensibilità ai dati di addestramento: Piccole variazioni nei dati di addestramento possono portare a alberi decisionali molto diversi.
- Tendenza all'overfitting.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Criteri di divisione



Bing Image Creator (quale strategia per suddividere?)

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

La suddivisione di un nodo in un albero decisionale avviene in base a criteri specifici che mirano a massimizzare l'informazione ottenuta e migliorare la capacità di classificazione del modello. Ecco i passi principali:

1. Selezione dell'attributo:

- L'algoritmo analizza tutti gli attributi disponibili e sceglie quello che fornisce la migliore discriminazione tra le diverse classi.
- Vengono utilizzati criteri come il guadagno di informazione, l'indice di Gini o la varianza di impurità per valutare l'utilità di ciascun attributo.
- L'attributo con il punteggio più alto viene selezionato per la suddivisione del nodo.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

2. Definizione del punto di split:

- Per l'attributo scelto, si determina il valore che meglio separa le istanze in base alla loro classe.
- Nel caso di attributi numerici, si individua una soglia di divisione.
- Per attributi categorici, si creano rami per ciascuna categoria.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

3. Creazione dei sottoalberi:

- In base al punto di split, il nodo viene suddiviso in due o più sottoalberi.
- I sottoalberi rappresentano i diversi possibili scenari in base al valore dell'attributo selezionato.
- Il processo di suddivisione si ripete ricorsivamente per ciascun sottoalbero, finché non si raggiungono le foglie dell'albero.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Per la scelta della suddivisione dei dati di un NODO, ci sono varie possibilità:

- Misura dell'Errore Quadratico Medio (MSE)
- Misura dell'Entropia
- Indice di Gini



Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Impurità di Gini: se è dato un nodo, si calcola considerando soltanto i campioni presenti in quel nodo. Supponiamo siano N in totale. Ogni campione di questi N ha una propria categoria di appartenenza (quella annotata nel dataset di addestramento). Supponiamo siano esse A, B, C,...e chiamiamo con n_A , n_B , n_C ,... il numero di campioni delle varie categorie, per cui $n_A + n_B + n_C + \dots = N$.

L'impurità di Gini si calcola come:

$$I_G = 1 - (n_A/N)^2 - (n_B/N)^2 - (n_C/N)^2 - \dots$$

Noi utilizziamo solo **alberi decisionali binari**: ogni nodo ha al più due nodi successori. In questo caso il valore massimo possibile di impurità di Gini è 0,5, che si ottiene quando i campioni si dividono esattamente a metà fra le due categorie. Un valore uguale a 0 significa che tutti i campioni appartengono alla stessa categoria.

(Intelligenza Artificiale - Tecnologia Zanichelli)

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

IMPURITA' DI GINI

La Gini impurity è un altro criterio comunemente utilizzato per misurare l'impurità o il disordine in un insieme di dati, soprattutto nei decision tree. Come l'entropia, la Gini impurity valuta quanto un nodo è impuro, ma utilizza una diversa formula matematica.

Per un nodo i con C classi, la Gini impurity è calcolata come:

$$\text{Gini}(i) = 1 - \sum_{c=1}^C (p(c|i))^2$$

La Gini impurity è compresa tra 0 e 1. Un valore di Gini impurity più vicino a 0 indica un nodo più puro, mentre un valore più vicino a 1 indica un nodo più impuro.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Metodi Ensemble e Random Forest



Bing Image Creator

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

I **metodi di ensemble** nell'apprendimento automatico consistono nel combinare le previsioni di più modelli per migliorare le prestazioni complessive e la robustezza. Questi metodi mirano a risolvere le limitazioni dei singoli modelli sfruttando l'intelligenza collettiva di un gruppo di modelli. Ecco alcuni popolari metodi di ensemble:

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Random Forest:

- Idea: Un ensemble di alberi decisionali creati attraverso il bagging. Introduce casualità aggiuntiva selezionando un sottoinsieme casuale di caratteristiche ad ogni suddivisione.
- Esempio: Ogni albero in un Random Forest è un modello di base, e la previsione finale viene fatta mediando (regressione) o votando (classificazione) su tutti gli alberi.



Bing Image Creator

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Ecco i passaggi chiave per comprendere meglio come funziona una Random Forest:

Campionamento Bootstrap:

- Si crea un numero specificato di sottoinsiemi di addestramento (campioni bootstrap) selezionando casualmente con sostituzione dati dal set di addestramento originale. Questo significa che alcuni dati possono apparire più di una volta in un sottoinsieme, mentre altri possono essere omessi.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Creazione degli Alberi Decisionali:

- Per ogni sottoinsieme creato tramite il campionamento bootstrap, si addestra un albero decisionale. Questi alberi sono detti "alberi decisionali deboli" e possono essere addestrati fino a una profondità massima specificata o fino a quando tutte le foglie contengono un numero minimo di campioni.

Modulo 3: “Regressione logistica - Decision Tree - Random Forest”

Voto o Media:

- Per la classificazione, il risultato finale della Random Forest è determinato attraverso un voto maggioritario delle previsioni di tutti gli alberi. Per la regressione, si può prendere la media delle previsioni degli alberi.

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.T.T. Pacinotti
una scuola per il futuro

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Insegna a un computer a giocare

Iniziamo

Per saperne di più

- 1 Raccogli esempi di cose che vuoi essere in grado di riconoscere
- 2 Utilizzare gli esempi per addestrare un computer ad essere in grado di riconoscerli
- 3 Crea un gioco in Scratch che utilizza la capacità del computer di riconoscerli

Machine Learning
For Kids... ma non solo!!

Machine learning for kids

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Cerca Fogli di lavoro Pretrained Stories Book Novità Aiuto Accedi

Machine Learning per bambini

In 6 secondi...

3 ... and use it to make games in App Inventor

45 seconds
Doggie
Monkey
Scissors
Restart Take photo
91% confident you found the Scissors

In 20 minuti...

Machine Learning for Kids

Machine Learning for Kids

machinelearningforkids.co.uk

Che cosa?

Questo strumento introduce il machine learning (o apprendimento automatico) fornendo esperienze pratiche per addestrare i sistemi di apprendimento automatico e costruire cose con loro.

Fornisce un ambiente guidato di facile utilizzo per l'addestramento di modelli di machine learning per la classificazione di testo, numeri o il riconoscimento di immagini.

Ciò si basa sugli sforzi esistenti per introdurre e insegnare la codifica ai bambini, aggiungendo questi modelli a Scratch (una piattaforma di codifica educativa ampiamente utilizzata), consentendo ai bambini di creare progetti e costruire giochi con i modelli di apprendimento automatico che hanno formato. Per saperne di più, guarda:

[Stories](#)

Machine Learning
For Kids... ma non
solo!!

Pagina
introduttiva

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Cerca **Fogli di lavoro** Pretrained Stories Book Novità Aiuto Accedi Language

Progetti di apprendimento automatico

Questi progetti sono guide passo-passo scaricabili, con spiegazioni e immagini che gli studenti devono seguire. Ogni progetto è un'attività a sé stante, scritta per durare una sola lezione, e guiderà i bambini a creare un gioco o un progetto interattivo che dimostrerà l'uso reale dell'intelligenza artificiale e dell'apprendimento automatico. Examples of how to use these worksheets, and a suggested order for using these with a class, can be found in the stories about what students learn. Suggerimenti per nuove schede, suggerimenti per il miglioramento di una qualsiasi delle schede o contributi con nuove schede, sono tutti benvenuti.

Tutti i tipi di progetto ▾ Tutte le difficoltà ▾ Tutti i tipi di marca ▾

Describe the glass Create a game in Scratch that learns when you describe a glass as half-full or half-empty. Teach a computer to predict your answers Difficoltà: Beginner Riconoscendo: numeri  Tag: decision tree learning Download	Pokémon images Make a project in Scratch that predicts the type of Pokémon from how it looks Teach a computer to recognize pictures Difficoltà: Beginner Riconoscendo: immagini  Tag: image classification, supervised learning Download	Aula Intelligente Crea un assistente intelligente in Scratch che ti consente di controllare i dispositivi virtuali. Insegna a un computer a riconoscere il significato dei tuoi comandi Difficoltà: Beginner Riconoscendo: testo  Tag: digital assistants, supervised learning Download	I Spy Make a game in Scratch where you have to guess what the computer recognises. Use a computer that has been trained to recognize photos Difficoltà: Beginner Riconoscendo: immagini  Tag: image recognition, pretrained models Download
Rendermi felice Crea un personaggio in Scratch che sorride se gli dici cose carine e piangi se dici cose cattive. 	Pokémon statistics Make a project in Scratch that predicts the type of Pokémon from their size and fighting stats. 	Snap! Crea un gioco di carte in Scratch che impara a riconoscere le foto della tua 	Quiz Show Use a computer that has been trained to find answers to questions in documents. 

Machine Learning
For Kids... ma non
solo!!

Schede di lavoro
pronte

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

The screenshot shows the top navigation bar with tabs: Circa, Fogli di lavoro, Pre-trained (highlighted with a red box), Storie, Book, Novità, Aiuto, Accedi, and Language. Below the navigation is a section titled "Pre-trained models".

Machine Learning for Kids provides pre-trained models you can use in your projects. Real-world machine learning projects often use models already trained by other people. There are lots of well-trained models that are freely available, and these are useful when you don't have time to collect the amount of training data needed to train your own.

How to use

Iniziamo

Click on the button to go to Scratch.

Go to the version of Scratch 3 available from Machine Learning for Kids.

Pre-trained models are available from the Extensions panel. Click on the blue extensions button in the bottom-left of the Scratch window to find them, then click on the one you want to add to your project.

Speech to text

This model can be used to recognize speech recorded through your microphone.

It gives you a block you can use to record some audio and then give you the text that it recognized, and a block that you can tell it to listen out for a particular word or phrase.

Face detection

This model can be used to recognize your face in the video feed from your webcam.

It gives blocks that will find the x,y coordinates of your eyes, nose and mouth.

The screenshot shows a grid of project thumbnails. Each thumbnail has a title below it:

- Alien Language
- Avoid the monster
- Auto o Tazza
- Carpool Cheats (test)
- Carpool Cheats (train)
- Camaleonte
- Confuso
- Describe the glass
- Explainability
- Face Lock

Machine Learning
For Kids... ma non
solo!!

Modelli pre
addestrati

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

The screenshot shows the Scratch interface with a purple header bar. The header includes the Scratch logo, settings, file, edit, and share buttons, followed by the text "Modelli di progetto" which is highlighted with a red box. To the right of "Modelli di progetto" are "Share" and "Progetto Scratch" buttons. Below the header is a search bar with the placeholder "Cerca" and a "Tutti" button. Underneath the search bar are tabs for "testo", "immagini", "numeri", and "suoni". The main area displays a grid of ten project thumbnails, each with its name below it:

- Alien Language
- Avoid the monster
- Auto o Tazza
- Carpool Cheats (test)
- Carpool Cheats (train)
- Camaleonte
- Confuso
- Describe the glass
- Explainability
- Face Lock

Machine Learning
For Kids... ma non
solo!!

Progetti svolti e
Modelli pre
addestrati

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

L'obiettivo di **Machine Learning for Kids** è aiutare gli studenti a conoscere l'intelligenza artificiale e l'apprendimento automatico. Ciò viene fatto fornendo opportunità di realizzare oggetti utilizzando le tecnologie di apprendimento automatico.

Quando vengono concessi il tempo e la libertà di costruire e sperimentare l'apprendimento automatico in un sandbox familiare come Scratch, gli studenti scoprono molto su come questa tecnologia viene applicata e si comporta.

Queste storie sono esempi: esempi di ciò che gli studenti hanno osservato, cosa hanno fatto e cosa hanno imparato. Si tratta di spunti utili per pianificare come introdurre l'intelligenza artificiale/ML in classe.

Se hai la tua storia, considera di contribuire .

L'apprendimento automatico non ha sostituito la necessità di imparare a programmare

Flusso di lavoro di un progetto di machine learning

La raccolta di più esempi di formazione spesso migliora la precisione

Il crowdsourcing e la gamification possono aiutare a generare dati di formazione

La crescente diversità dei dati di addestramento spesso migliora la precisione

I dati relativi alla formazione e ai test devono essere mantenuti separati

L'apprendimento automatico può imparare a riconoscere diversi tipi di modelli

Machine Learning
For Kids... ma non
solo!!

Lezioni e schede di
lavoro

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

 Circa Fogli di lavoro Preaddestrato Storie Libro Novità **AIuto** [Accedi](#) Lingua

In caso di problemi o domande, chiedere nel gruppo Google ML for Kids

Come cominciare?
Scarica un PDF con una breve guida introduttiva:

 **Conto di classe gestiti**
Per insegnanti scolastici, leader di gruppi di programmazione o altri volontari

 **Conto di classe non gestiti**
Per i dipendenti IBM o gli insegnanti e i leader dei gruppi di programmazione che gestiscono i propri account di classe

Vuoi sapere di più?
Qualcosa non funziona?
L'elenco che segue contiene le domande e i problemi che sento di più spesso. Se vedi qualcosa che assomiglia al tuo problema, clicca sul sommario per maggiori informazioni.
Se sei ancora bloccato o hai riscontrato un problema che non sono incluso nell'elenco, scusa! Fai una domanda nel gruppo Google mlforkids in modo che io possa aiutarti.

Domande sul progetto

Machine Learning
For Kids... ma non
solo!!

Help!!!

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

The screenshot shows the homepage of the MachineLearning ForKids website. At the top, there is a navigation bar with links: Circa, Fogli di lavoro, Preaddestrato, Storie, Libro, Novità, Aiuto, and Accedi. The 'Accedi' button is highlighted with a red box. Below the navigation bar, a large button says "Inizia con il machine learning". Underneath it, there are three main sections: "È la tua prima visita?", "Già registrato?", and "Prova senza registrarti". Each section has a "Prova subito" button. To the right of the first section is a "Iscriviti" button and a link "Perché registrarsi?". To the right of the second section is a "Accedi" button and a link "Hai dimenticato i tuoi dati?".

Machine Learning
For Kids... ma non
solo!!

This screenshot shows the "Perché registrarsi?" (Why register?) page. It explains that registration is not mandatory but offers more features. It compares the "Prova subito" (Try now) and "Crea un account" (Create an account) options through a table.

	Prova subito	Crea un account
Utilizza modelli di machine learning preaddestrati	✓	✓
Utilizza modelli TensorFlow ospitati	✓	✓
Aggiungere modelli di apprendimento automatico per riconoscere...		
testo	✓	✓
numeri	✓	✓
suoni	✓ <small>utilizzare uno da Google</small>	✓ <small>utilizzare uno da Google</small>
immagini	✓ <small>utilizzare uno da Google</small>	✓

Creare un account

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

The screenshot shows the homepage of the MachineLearning ForKids website. At the top, there is a navigation bar with links: Circa, Fogli di lavoro, Preaddestrato, Storie, Libro, Novità, Aiuto, and Accedi. The 'Accedi' button is highlighted with a red box. Below the navigation, a large button says "Inizia con il machine learning". Underneath, there are three main sections: "È la tua prima visita?", "Già registrato?", and "Prova senza registrarti". Each section has a blue "Iscriviti" or "Accedi" button and a "Perché registrarsi?" or "Hai dimenticato i tuoi dati?" link. At the bottom, it says "Accesso fornito da:" followed by the AuthID logo and "Token Based Authentication".

Machine Learning
For Kids... ma non
solo!!

This screenshot shows a simplified version of the website's interface. It features a large "Prova subito" button at the top right and a "Prova senza registrarti" link at the bottom left.

Prova senza
registrazione

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Insegna a un computer a giocare

Vai ai tuoi progetti

- 1 Raccogli esempi di cose che vuoi essere in grado di riconoscere
- 2 Utilizzare gli esempi per addestrare un computer ed essere in grado di riconoscerli
- 3 Crea un gioco in Scratch che utilizza la capacità del computer di riconoscerli

Machine Learning For Kids... ma non solo!!

Prova senza registrazione

I tuoi progetti di machine learning

Fai clic sul pulsante “più” sulla destra per creare il tuo primo progetto.→

+ Aggiungi un nuovo progetto Copia modello

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Inizia con il machine learning

< Indietro

Sei un insegnante o istruttore in un gruppo di coding?

Se desideri dare l'accesso ai tuoi studenti, crea un account di classe qui di seguito.

Esistono due tipi di account di classe, a seconda di chi lo ha impostato.

Crea un account di classe libero

Se preferisci configurare e gestire tu stesso l'account di classe, puoi farlo ora!

Dovrai:

- Creare un account su IBM Cloud
- Creare una chiave API per utilizzare i servizi cloud di Watson, necessari alla creazione dei modelli di machine learning
- Inserisci le chiavi API in questo strumento Machine Learning per bambini

Sono disponibili istruzioni passo-passo (con screenshot) per tutti i passaggi.

Iscriviti

Crea un account classe gestito

Se desideri che Dale configuri l'account della classe, utilizza il pulsante in basso per inviare un'email con le seguenti informazioni:

- Indicativamente, quanti account studenti ti serviranno
- Il nome della scuola o del gruppo di coding per cui crei l'account

This offer is only available to schools and not-for-profit student coding clubs.

Email

Machine Learning
For Kids... ma non
solo!!

Registrazione al
servizio web con 2
possibilità

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Username *
fpicca

Email address *
f.picca@righitarranto.edu.it

6 / 15

Intended use

✓ By checking this box, you accept responsibility for any fees or charges incurred on IBM Cloud.

Please read the notes below. If you're concerned about any of this, or have any questions, please get in touch.

Setting up this tool to allow your students to train machine learning models will require some access codes (called "API keys") from IBM Cloud.

"Do I have to pay for that?"

- No.
- Creating an account on IBM Cloud is free.
- There is a limit to the number of free API keys you can create.
- Watson API keys you can create for free have usage limits. You can choose to pay IBM if you want to increase these limits. It is these charges that are your responsibility.
- An advisory price list is included in the [setup instructions](#) (section 4).

"Will my students need to do this?"

- No.
- This is a one-off setup task that you will need to do as the teacher / group-leader.
- Students in your group will not need to create an account with IBM Cloud.
- Students in your group will not need to visit IBM Cloud.

"How will I know how to do that?"

- Step-by-step instructions with screenshots are available to walk you through this.

"Do all projects require IBM Cloud API keys?"

- No. Machine learning projects to recognize images, numbers or sounds do not require API keys. Projects using pre-trained models do not require API keys.
- Only machine learning projects to recognize text require IBM Cloud API keys.

If you're concerned about any of this, or have any questions, please [get in touch](#).

[CREATE CLASS ACCOUNT](#) CANCEL

Machine Learning
For Kids... ma non
solo!!

Creazione
ACCOUNT

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Create an unmanaged class account

Your account has been created - your password is **ksPF3NTIB8T6**. Check your inbox for a confirmation email. You need to click the link in the email to activate your class account.

Username *
fpicca

Email address *
f.picca@righitaranto.edu.it

6 / 15

Insegna a un computer a giocare

1 Raccogli esempi di cose essere in grado di riconoscere

2 Utilizzare gli esempi per addestrare un computer in grado di riconoscerli

Welcome to Machine Learning for Kids
Your email address has been verified.

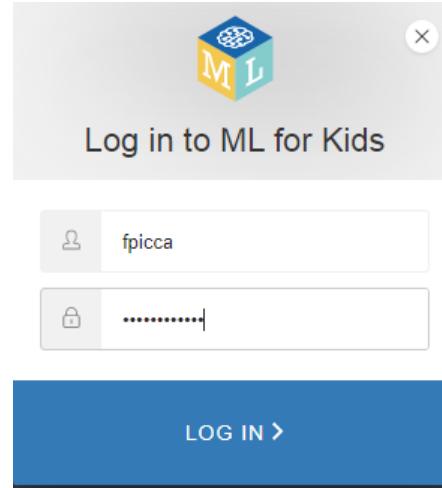
Iniziamo Per saperne di più OK

Machine Learning For Kids... ma non solo!!

Fatto!!

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione



Machine Learning
For Kids... ma non
solo!!

Accesso

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Teach a computer to play a game

[Go to your Admin Page](#)

[Go to your Projects](#)

- 1 Collect examples of things you want to be able to recognise
- 2 Use the examples to train a computer to be able to recognise them
- 3 Make a game in Scratch that uses the computer's ability to recognise them

Machine Learning
For Kids... ma non solo!!

Pagina personale

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Administration

Restrictions
Review limits and restrictions that apply to your class
[Restrictions](#)

API Keys
Incomplete!
Set up IBM Watson credentials to enable your class to create machine learning models
[Watson API Keys](#)

Students
Create, reset and delete logins for students in your class
[Student management](#)

Supervision
Review projects that your students have created
[Supervision](#)

Advanced options

Can you help?
Please share your experiences and ideas for how to make the tool better
[Can you help?](#)

Delete class
This will delete the accounts for you and all of the students in your class
[Delete class](#)

Machine Learning
For Kids... ma non
solo!!

Attivazione delle
API

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Amministrazione: chiavi API IBM Watson

Se non sei sicuro di cosa fare qui, segu la guida passo passo. Se desideri aiuto, contattaci.

L'addestramento di modelli di machine learning per progetti **di testo** utilizza **Watson Assistant**

Non hai credenziali di assistente.
Ciò significa che non sarà possibile addestrare modelli per progetti **di testo**.
Se non sai dove trovarli, scarica la [guida passo passo](#)

+ Aggiungi nuove credenziali

L'addestramento dei modelli di machine learning per i progetti **di immagini** non richiede chiavi API

L'addestramento dei modelli di machine learning per i progetti **numerici** non richiede chiavi API

L'addestramento dei modelli di machine learning per i progetti **audio** non richiede chiavi API

Aggiungi le credenziali di Watson Assistant

Questi verranno utilizzati quando i tuoi studenti formeranno i modelli in progetti di testo. Questo non verrà mostrato agli studenti del tuo gruppo.

Per assistenza nella creazione di nome utente e password o chiave API, scarica la [guida passo passo](#)

IBM Cloud Catalog Docs Support Manage

Watson Services / Assistant / for-machine-learning-for-kids

Resource Group: default Location: US East Add Tags

Get started by launching the tool.

Launch tool Getting started tutorial API reference

Credentials

API Key: `_EsquqMj_gDINWf3tpfvlKuawipZdF9Lz7Kuodz1qPWRJLj`

URL: <https://gateway-wdc.watsonplatform.net/assistant>

Inserisci la chiave API di seguito
Chiave API:

0/44

AGGIUNGERE ANNULLA

Machine Learning
For Kids... ma non
solo!!

Aggiunta API per il
servizio di Testo

Modulo 4: “Machine Learning con MachineLearning ForKids”

Introduzione

Guida a IBM Cloud

Insegnanti/genitori/dirigenti di club di codice con account di classe non gestiti devono fornire “[chiavi API](#)” da IBM Cloud. Si tratta di codici segreti che forniscono l’accesso alla tecnologia di intelligenza artificiale che alimenterà i progetti degli studenti.

I progetti di testo necessitano di una chiave API “Watson Assistant”. (Immagini, numeri e progetti sonori non richiedono chiavi API).

Dovrai creare un account su IBM Cloud per poter creare questi codici. La creazione di un account è gratuita e non sono previsti costi per le chiavi API “Lite”.

[Come creare chiavi API](#)

Scarica una guida passo passo su come creare chiavi API su IBM Cloud

Machine Learning
For Kids... ma non
solo!!

Scarica la GUIDA
per ottenere l’API

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Circa Progetti Fogli di lavoro Preaddestrato Storie Libro Novità Aiuto Esci Lingua

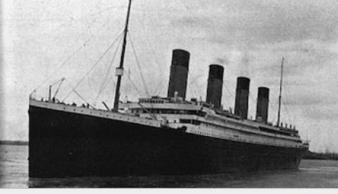
Copia un modello di progetto di machine learning

Puoi creare un nuovo progetto di machine learning basato su uno di questi modelli.

Ognuno include una serie di dati di allenamento che sono stati raccolti per aiutarti a iniziare rapidamente.

Se lo desideri, potrai apportare modifiche ai dati di addestramento oppure puoi procedere direttamente all'addestramento di un modello di machine learning con essi.

Sopravvissuti al Titanic
Prevedere la sopravvivenza sul Titanic
Il 15 aprile 1912, il Titanic colpì un iceberg ed affondò, uccidendo oltre 1500 persone tra passeggeri ed equipaggio. Questo progetto include dati personali su centinaia di passeggeri. Puoi usarlo per addestrare un modello di apprendimento automatico che, se gli fornisci la descrizione di un passeggero, può imparare a prevedere la probabilità che sarebbe sopravvissuto.



Riconoscevo: numeri Importare

Titoli dei giornali del Regno Unito
Riconoscere l'uso del linguaggio nei titoli dei giornali
Diversi giornali usano la lingua in modi diversi. Questo progetto include il titolo principale in prima pagina di quattro giornali britannici. Puoi usarlo per addestrare un modello di machine learning in grado di imparare a prevedere il giornale più probabile da cui probabilmente proviene un nuovo titolo.



Riconoscevo: testo Importare

Cani e gatti
Riconoscere le foto di cani e gatti
Addestrare un modello di machine learning per riconoscere foto di cani e gatti è una buona e semplice introduzione al riconoscimento delle immagini.



Riconoscevo: testo Importare

Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Circa Progetti Fogli di lavoro Preaddestrato Storie Libro Novità Aiuto Esci Lingua

Copia un modello di progetto di machine learning

Puoi creare un nuovo progetto di machine learning basato su uno di questi modelli.

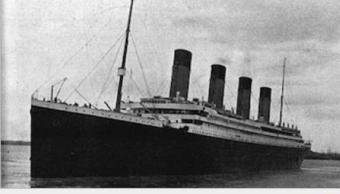
Ognuno include una serie di dati di allenamento che sono stati raccolti per aiutarti a iniziare rapidamente.

Se lo desideri, potrai apportare modifiche ai dati di addestramento oppure puoi procedere direttamente all'addestramento di un modello di machine learning con essi.

Sopravvissuti al Titanic

Prevedere la sopravvivenza sul Titanic

Il 15 aprile 1912, il Titanic colpì un iceberg ed affondò, uccidendo oltre 1500 persone tra passeggeri ed equipaggio. Questo progetto include dati personali su centinaia di passeggeri. Puoi usarlo per addestrare un modello di apprendimento automatico che, se gli fornisci la descrizione di un passeggero, può imparare a prevedere la probabilità che sarebbe sopravvissuto.



Riconosce numeri Importare

Titoli dei giornali del Regno Unito

Riconoscere l'uso del linguaggio nei titoli dei giornali

Diversi giornali usano la lingua in modi diversi. Questo progetto include il titolo principale in prima pagina di quattro giornali britannici. Puoi usarlo per addestrare un modello di machine learning in grado di imparare a prevedere il giornale più probabile da cui probabilmente proviene un nuovo titolo.



"Cani e gatti"

Addestrare un modello di machine learning per riconoscere foto di cani e gatti è una buona e semplice introduzione al riconoscimento delle immagini.

Questo set di dati di addestramento di foto da Wikimedia Commons è un buon set di partenza per creare i propri dati di addestramento di foto di animali

Vuoi utilizzare alcuni dati per i test?

Mantieni lo 0% per i test. Utilizza il 100% per l'allenamento

IMPORTARE

Cani e gatti

Riconoscere le foto di cani e gatti

Addestrare un modello di machine learning per riconoscere foto di cani e gatti è una buona e semplice introduzione al riconoscimento delle immagini.



Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

The screenshot shows the homepage of the MachineLearning ForKids website. At the top, there is a navigation bar with links: Circa, Progetti, Fogli di lavoro, Preaddestrato, Storie, Libro, Novità, Aiuto, and Esci. On the right side of the header, there is a "Lingua" dropdown menu. Below the header, a large banner says "I tuoi progetti di machine learning". Underneath this, there is a project card for "Cani e gatti". The card displays the text "Riconoscendo immagini come gatto o cane". To the right of the card are two buttons: "Aggiungi un nuovo progetto" and "Copia modello". Below the project card is a small trash can icon. The main content area has a red border around the "Addestramento" section.

Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Addestramento

Raccogli esempi di ciò che vuoi che il computer riconosca

Addestramento

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

Impara & Testa

Programma

Utilizza il modello di machine learning che hai addestrato per creare un gioco o un'app, in Scratch, Python o App Inventor

Programma

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

I tuoi progetti di machine learning

Cani e gatti

Riconoscendo immagini come gatto o cane

+ Aggiungi un nuovo progetto Copia modello

Machine Learning
For Kids... ma non
solo!!

Riconoscendo immagini vieni gatto o cane

gatto

cane

14 13

Classificare cani e gatti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

The screenshot shows the homepage of the MachineLearning ForKids website. At the top, there is a navigation bar with links: Circa, Progetti, Fogli di lavoro, Preaddestrato, Storie, Libro, Novità, Aiuto, and Esci. On the right side of the header, there is a "Lingua" dropdown menu. Below the header, a large banner says "I tuoi progetti di machine learning". Underneath this, there is a project card for "Cani e gatti", which is described as "Riconoscendo immagini come gatto o cane". To the right of the project card are two buttons: "Aggiungi un nuovo progetto" and "Copia modello". Below the project card, there is a small trash can icon and a square icon with a minus sign. The main content area has a red border around the text "Machine Learning For Kids... ma non solo!!".

Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Addestramento

Raccogli esempi di ciò che vuoi che il computer riconosca

Addestramento

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

Impara & Testa

Programma

Utilizza il modello di machine learning che hai addestrato per creare un gioco o un'app, in Scratch, Python o App Inventor

Programma

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

< Ritorna ai progetti

Modelli di machine learning

Cos'hai fatto?

Hai addestrato un modello di apprendimento automatico per riconoscere quando le immagini sono gatto o cane.

Hai creato il modello su giovedì 18 gennaio 2024 13:52.

Hai raccolto:

- 14 esempi di gatto ,
- 13 esempi di cane

E adesso che succede?

Prova a testare il modello di apprendimento automatico qui sotto. Inserisci un'immagine di esempio che non hai incluso negli esempi che hai usato per addestrarlo. Ti dirà come lo riconoscerà e quanto è sicuro di sé.

Se il computer sembra aver imparato a riconoscere correttamente le cose, allora si può andare su Scratch e usare ciò che il computer ha imparato per creare un gioco!

Se il computer sta sbagliando troppe cose, potresti voler tornare alla pagina [Addestramento](#) e raccogliere altri esempi.

Una volta fatto questo, clicca sul pulsante qui sotto per addestrare un nuovo modello di apprendimento automatico e vedere quale differenza faranno gli esempi extra!

Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Dati di test:

Hai mantenuto alcuni dati del modello di progetto fuori dal set di training, in modo che possano essere utilizzati per i test. Fare clic sul pulsante in basso per scaricare un file CSV.

Scarica i dati del test

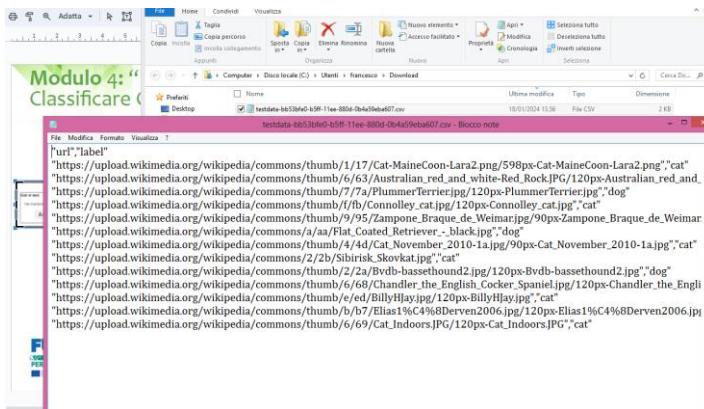
Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Dati di test:

Hai mantenuto alcuni dati del modello di progetto fuori dal set di training, in modo che possano essere utilizzati per i test. Fare clic sul pulsante in basso per scaricare un file CSV.

[Scarica i dati del test](#)



Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Machine Learning
For Kids... ma non
solo!!

Dati di test:

Hai mantenuto alcuni dati del modello di progetto fuori dal set di training, in modo che possano essere utilizzati per i test. Fare clic sul pulsante in basso per scaricare un file CSV.

Scarica i dati del test

Classificare cani e gatti, con
immagine da test

Prova a inserire un'immagine per vedere come viene riconosciuta in base al tuo addestramento.

 Prova con la webcam

 Prova di disegno

 Prova con www

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Machine Learning
For Kids... ma non
solo!!

Prova a inserire un'immagine per vedere come viene riconosciuta in base al tuo addestramento.

 Prova con la webcam

 Prova di disegno

<https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Cat-MaineCoon-Lara2.png/598px-Cat-MaineCoon-Lara2.png>

Prova con www

Classificare cani e gatti

Prova a inserire un'immagine per vedere come viene riconosciuta in base al tuo addestramento.

 Prova con la webcam

 Prova di disegno

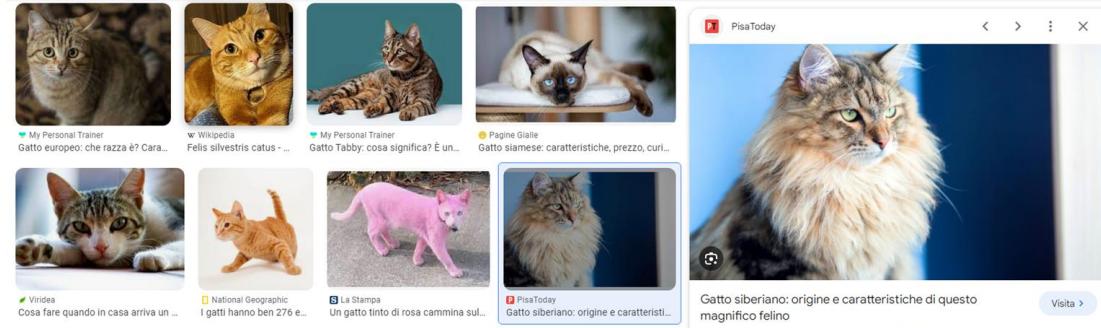
<https://upload.wikimedia.org/wikipedia/commons/thumb/1/17/Cat-MaineCoon-Lara2.png/598px-Cat-MaineCoon-Lara2.png>

Prova con www

Riconosciuto come **gatto**
con il 79% di fiducia

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti



Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti, con
immagine dalla rete

Prova a inserire un'immagine per vedere come viene riconosciuta in base al tuo addestramento.

Prova con **webcam**

Test di **disegno**

https://upload.wikimedia.org/wikipedia/commons/thumb/9/9e/Ginger_european_cat.jpg/220px-Ginger_european_cat.jpg

Prova con **www**

Riconosciuto come **cat**
with 100% confidence

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

The screenshot shows the main interface of the MachineLearning ForKids platform. At the top, there's a navigation bar with links like "Circa", "Progetti", "Fogli di lavoro", "Preaddestrato", "Storie", "Libro", "Novità", "Aiuto", and "Esci". On the right, there's a "Lingua" dropdown. Below the navigation, a header says "I tuoi progetti di machine learning". A red box highlights a project titled "Cani e gatti" which is "Riconoscendo immagini come gatto o cane". To the right of the project are two buttons: "+ Aggiungi un nuovo progetto" and "Copia modello". Below the project list, there are three large cards: "Addestramento", "Impara & Testa", and "Programma".

Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Addestramento

Raccogli esempi di ciò che vuoi che il computer riconosca

[Addestramento](#)

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

[Impara & Testa](#)

Programma

Utilizza il modello di machine learning che hai addestrato per creare un gioco o un'app, in Scratch, Python o App Inventor

[Programma](#)

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Addestramento

Raccolgi esempi di ciò che vuoi che il computer riconosca

Addestramento

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

Impara & Testa

Programma

Utilizza il modello di machine learning che hai addestrato per creare un gioco o un'app, in Scratch, Python o App Inventor

Programma

Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Crea qualcosa con il tuo modello di apprendimento automatico

< Ritorna ai progetti

Scratch 3



Python

Scrivi il codice Python per utilizzare il tuo modello di apprendimento automatico



Python



App Inventor

Crea un'applicazione mobile per il tuo telefono o tablet



App Inventor



Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

Crea qualcosa con il tuo modello di apprendimento automatico

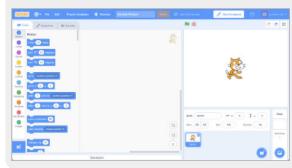
< Ritorna ai progetti

Scratch 3

Usa la nuova versione di Scratch



Scratch 3

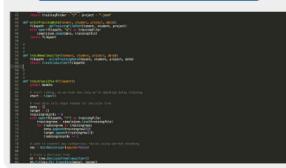


Python

Scrivi il codice Python per utilizzare il tuo modello di apprendimento automatico



Python



App Inventor

Crea un'applicazione mobile per il tuo telefono o tablet



App Inventor

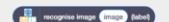


Usare l'apprendimento automatico in Scratch 3

< Ritorna ai progetti

Apri in Scratch 3

Your project will add these blocks to Scratch.



Inserire images nell'input e restituirà l'etichetta con cui il modello di apprendimento automatico lo riconosce.



Cio restituirà la sicurezza del modello di apprendimento automatico che riconosce il tipo di images. (Come un numero compreso tra 0 e 100).

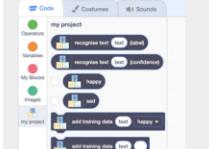


Questi blocchi rappresentano le etichette che hai creato nel tuo progetto, quindi puoi usarne i loro nomi negli script.

Cio significa che puoi fare qualcosa del genere:



Assomiglierà a qualcosa del genere - tranne per il nome del tuo progetto.

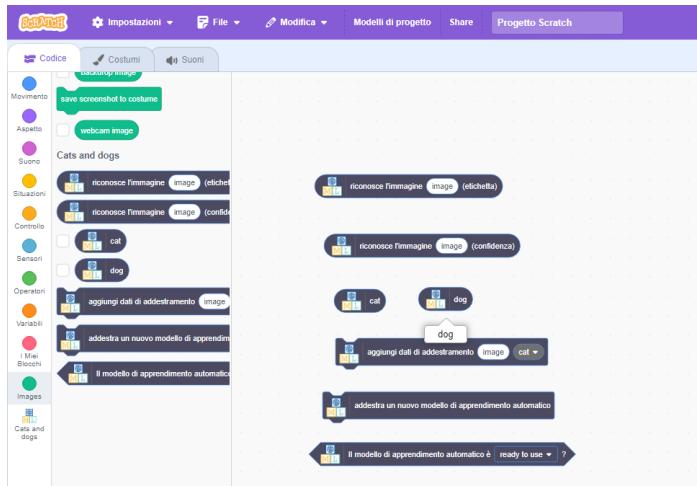


Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti

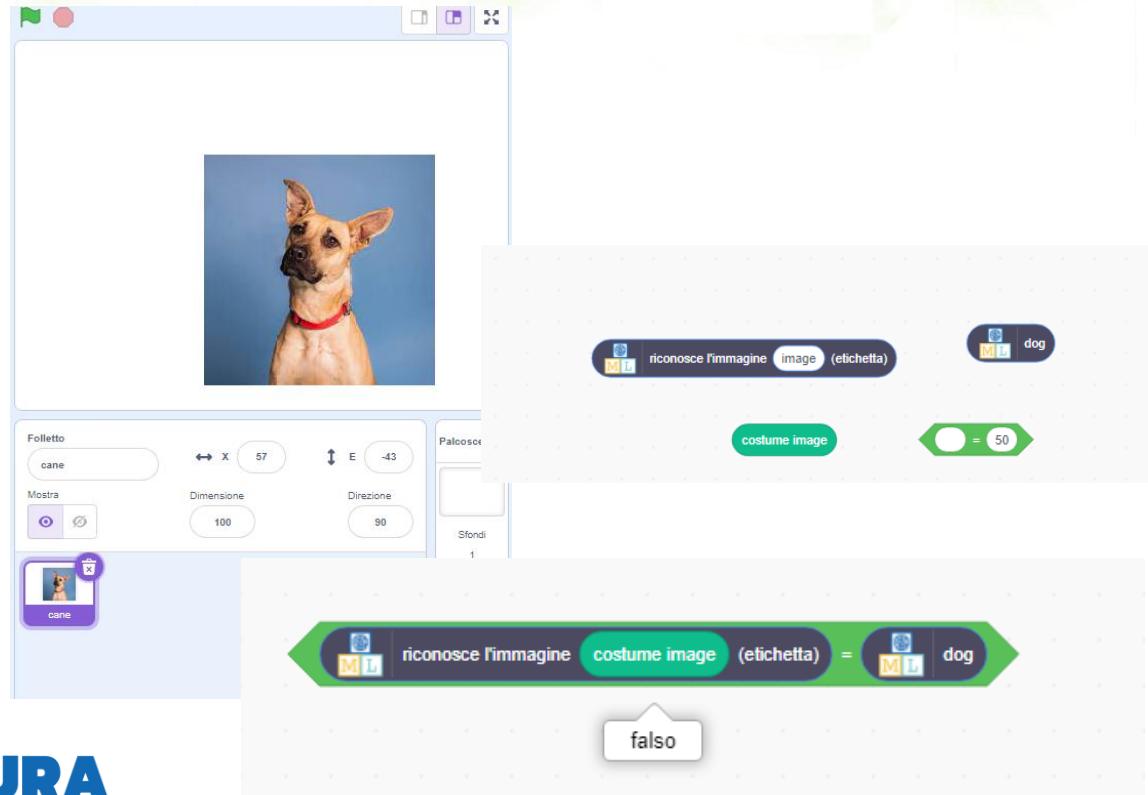


Machine Learning
For Kids... ma non
solo!!

Classificare cani e gatti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Classificare Cani e Gatti



Machine Learning
For Kids... ma non
solo!!

Scelgo uno sprite con
costume cane, il modello
sbaglia!!

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Notizie false

Riconoscere le notizie false

Addestra un modello di machine learning per riconoscere il tipo di scrittura nelle fake news.



Riconoscendo: **testo**

Importare

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

"Notizie false"

Addestra un modello di machine learning per riconoscere il tipo di scrittura nelle fake news.

Le notizie false e fuorvianti sono più comuni poiché le persone si rivolgono ai social media per ottenere le loro notizie. Utilizza questo set di dati per creare un modello di machine learning in grado di provare a distinguere quando le notizie sono vere o false. Attribuzione:

<https://www.uvic.ca/ecs/ece/Isot/datasets/fake-news/> Ringraziamenti: Ahmed H, Traore I, Saad S. "Detecting opinion spams and fake news using text classification", Journal of Security e Privacy, volume 1, numero 1, Wiley, gennaio/febbraio 2018. Ahmed H, Traore I, Saad S. (2017) "Rilevamento di notizie false online utilizzando l'analisi N-Gram e tecniche di apprendimento automatico". In: Traore I., Woungang I., Awad A. (a cura di) Sistemi intelligenti, sicuri e affidabili in ambienti distribuiti e cloud. ISDDC 2017. Lecture Notes in Computer Science, vol 10618. Springer, Cham (pp. 127-138).

✓ Vuoi utilizzare alcuni dati per i test?

Mantieni lo 0% per i test



Utilizzare il 100% per l'allenamento

IMPORTARE

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Circa Progetti Fogli di lavoro Preaddestrato Storie Libro Novità Aiuto Esci Lingua

Riconoscendo **il testo** come **vero o falso**

< Ritorna ai progetti

Aggiungi una nuova etichetta

VERO

- Mercoledì i legislatori ugandesi hanno votato a stragrande ...
- Un ex dirigente della Korea Air Lines che si è infuriato a bor...
- La decisione della Commissione europea di avviare la cosid...
- Sabato il Partito comunista russo ha inaspettatamente abba...

falso

- Donald Trump continua a scavare sempre più in una fossa ...
- È appena venuto alla luce che Stanislav Petrov, ex tenente ...
- Prima che la sua campagna si trasformasse in orde arrabbi...
- Donald Trump sa sicuramente come aggiudicare la beffa al ...

+ Aggiungi un esempio

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
una serie di notizie VERE e
una seconda serie di notizie
FALSE

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Modelli di machine learning

[Ritorna ai progetti](#)

Cos'hai fatto?

Hai raccolto esempi di testo che un computer può utilizzare per riconoscere quando il testo è vero o falso.

Hai raccolto:

- 366 esempi di vero ,
- 364 esempi di falso

E adesso che succede?

Pronto per iniziare l'addestramento?

Fai clic sul pulsante in basso per iniziare ad addestrare un modello di apprendimento automatico utilizzando gli esempi che hai raccolto finora

(Oppure torna alla pagina [Addestramento](#) se vuoi prima raccogliere altri esempi.)

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Dati di test:

Hai mantenuto alcuni dati del modello di progetto fuori dal set di training, in modo che possano essere utilizzati per i test. Fare clic sul pulsante in basso per scaricare un file CSV.

[Scarica i dati del test](#)

Informazioni dal computer di addestramento:

Il modello ha iniziato il suo addestramento alle: Venerdì 19 gennaio 2024 11:49

Stato attuale del modello: Formazione

Ultimo controllo del modello: meno di un minuto fa 

[Annulla l'addestramento](#)

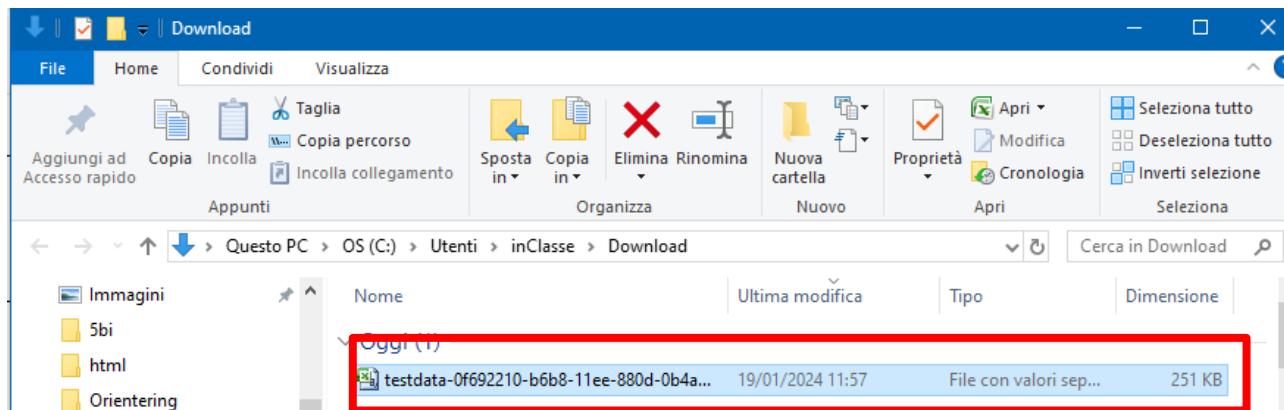
[Descrivi il tuo modello!](#) beta

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
scarichiamo i dati per il
TEST

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News



Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

testdata-0f692210-b6b8-11ee-880d-0b4a59eba607.csv - Blocco note di Windows

File Modifica Formato Visualizza ?

"text","label"

"South Sudanese rebels accused the government army of attacking one of their bases overnight as a new round of peace talks between the warring sides opened in the Ethiopian capital on Monday. Rebel spokesman Lam Paul Gabriel said government troops attacked a rebel base in the town of Lasu in the south of the country late on Sunday. They are in the IO base, he said, referring to the name of the rebel group. Army spokesmen were not immediately available to comment when called by Reuters on Monday afternoon. The talks in Addis Ababa have been convened by the East African bloc IGAD and are aimed at bringing the warring sides back to the negotiating table after a 2015 peace deal collapsed last year during heavy fighting in the capital, Juba. The war began in 2013 between soldiers of President Salva Kiir, an ethnic Dinka, and his former vice president, Riek Machar, a Nuer."

Linea 1, colonna 1 100% Windows (CRLF) UTF-8

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
testiamo una frase FALSA

Prova a inserire del testo per vedere come viene riconosciuto in base al tuo addestramento.

House Majority Leader Kevin McCarthy one of Trump's strongest allies in Congress divulged a hugely ex

Prova

Descrivi il tuo modello! beta

Riconosciuto come falso
con l'86% di confidenza

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Informazioni dal computer di addestramento:

Il modello ha iniziato il suo addestramento alle: Saturday, January 20, 2024 9:45 AM

Stato attuale del modello: Training

Ultimo controllo del modello: less than a minute ago 

[Annulla l'addestramento](#)

[Describe your model! beta](#)

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
descriviamo/spiegabilità del
Modello

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Informazioni dal computer di addestramento:

Il modello ha iniziato il suo addestramento alle: Saturday, January 20, 2024 9:45 AM

Stato attuale del modello: Training

Ultimo controllo del modello: less than a minute ago 

[Annulla l'addestramento](#)

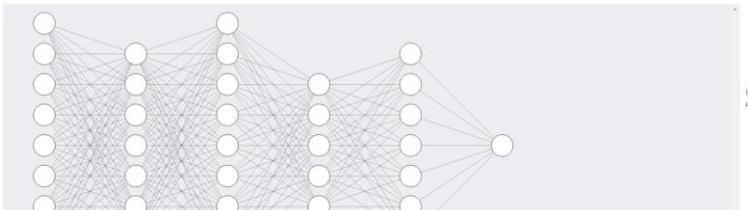
[Describe your model! beta](#)

[Comprendi il tuo modello di machine learning beta](#)

< Torna al modello

Il tipo di modello di machine learning che hai addestrato in questo progetto è chiamato **rete neurale**.
Le prossime pagine spiegheranno come è stato creato il tuo modello.
Questi diagrammi non sono una descrizione esatta del tuo modello. Il tuo modello è più grande e più complicato, ma questi diagrammi sono più facili da spiegare.

[Indietro](#) [Avanti >](#)



per maggiori dettagli prova a cercare:
reti neurali 

Utilizzare questi controlli per ingrandire i diagrammi 

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
descriviamo/spiegabilità del
Modello

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

The screenshot shows the homepage of the MachineLearning ForKids website. At the top, there is a navigation bar with links: Circa, Progetti, Fogli di lavoro, Preaddestrato, Storie, Libro, Novità, Aiuto, and Esci. Below the navigation bar, there is a large central area with three main sections:

- Addestramento**: A section for training the machine learning model. It includes the text "Raccogli esempi di ciò che vuoi che il computer riconosca" and a blue "Addestramento" button.
- Impara & Testa**: A section for learning and testing. It includes the text "Usa gli esempi per addestrare il computer a riconoscere" and a blue "Impara & Testa" button.
- Programma**: A section for programming. It includes the text "Utilizza il modello di machine learning che hai addestrato per realizzare un gioco o un'app, in Scratch o in Python" and a blue "Programma" button.

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
fase 3, passiamo al
programma il modello
addestrato

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News

< Ritorna ai progetti

Scratch 3

Usa la nuova versione di Scratch



Python

Sorvi il codice Python per utilizzare il tuo
modello di apprendimento automatico



EduBlocks

Write Python using a drag and drop coding tool



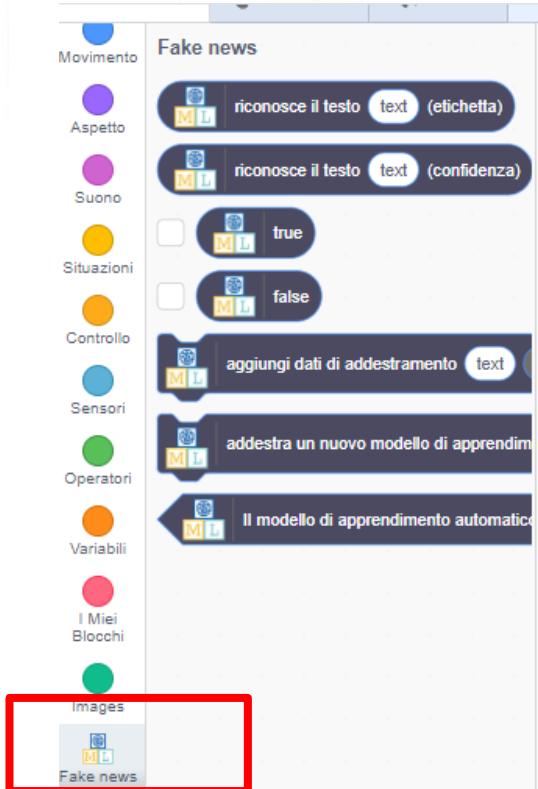
App Inventor

Crea un'applicazione mobile per il tuo telefono
o tablet



Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News



Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
otteniamo i relativi blocchi in
Scratch

Modulo 4: “Machine Learning con MachineLearning ForKids”

Riconoscere Fake News

"Donald Trump's budget is all about punishing the poor while enriching himself. The whole problem with Trump keeping his various business enterprises in his family is that he can use the presidency to only support policies that help him and his family profit financially. And forget about Trump actually doing anything that would harm his own bottom line. Millions of Americans rely on housing aid to make sure they have a roof over their heads and the heads of their children. But that doesn't matter to Trump at all. You see while Trump's budget makes massive cuts to housing and shelter programs that are designed to help the poor and the homeless many of whom are also veterans it protects a federal program that punishes landlords. Acco

Trump himself
of Starrett City
4 percent s", "fa

The screenshot shows a user interface for text recognition. At the top, there is a button labeled "ML" and a status message "riconosce il testo" followed by "text (confidenza)". Below this, a text input field contains the sentence: "'Donald Trump's budget is all about punishing the poor while enriching himself. The whole problem with Trump keeping his various business enterprises in his family is that he can use the presidency to only support policies that help him and his family profit financially. And forget about Trump actually doing anything that would harm his own bottom line. Millions of Americans rely on housing aid to make sure they have a roof over their heads and the heads of their children. But that doesn't matter to Trump at all. You see while Trump's budget makes massive cuts to housing and shelter programs that are designed to help the poor and the homeless many of whom are also veterans it protects a federal program that punishes landlords. Acco'". A large red arrow points from this text to a callout box on the right. At the bottom right of the interface, there is a button labeled "false".

Machine Learning
For Kids... ma non
solo!!

Riconoscere Fake News:
una fake news su Donald
Trump

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

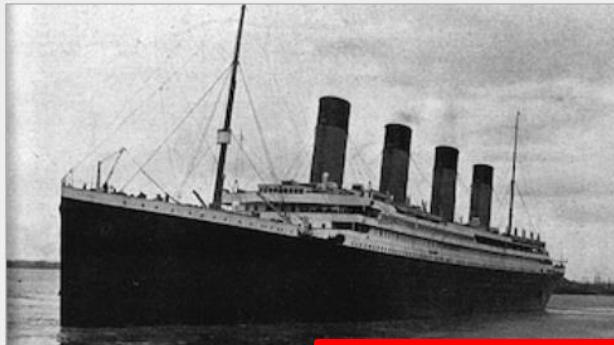
Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Sopravvissuti al Titanic

Prevedere la sopravvivenza sul Titanic

Il 15 aprile 1912, il Titanic colpì un iceberg e affondò, uccidendo oltre 1500 persone tra passeggeri ed equipaggio. Questo progetto include dati personali su centinaia di passeggeri. Puoi usarlo per addestrare un modello di apprendimento automatico che, se gli fornisci la descrizione di un passeggero, può imparare a prevedere la probabilità che sarebbe sopravvissuto.



Riconoscendo: numeri

Importare

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Sopravvissuti al Titanic*

Il 15 aprile 1912, il Titanic colpì un iceberg e affondò, uccidendo oltre 1500 persone tra passeggeri ed equipaggio. Questo progetto include dati personali su centinaia di passeggeri. Puoi usarlo per addestrare un modello di apprendimento automatico che, se gli fornisci la descrizione di un passeggero, può imparare a prevedere la probabilità che sarebbe sopravvissuto.

Il campo "classe biglietto" è un numero che indica il tipo di biglietto per il passeggero: prima classe (1), seconda classe (2) o terza classe (3). Il campo "genere" indica se il passeggero era "maschio" o "femmina". Il campo "età" indica l'età del passeggero. Il "fratello": sp. è il numero di fratelli o sorelle (fratelli) che il passeggero aveva a bordo se erano bambini, oppure il numero di mariti o mogli (coniugi) che il passeggero aveva a bordo se erano adulti. Il "par. cap." campo è il numero di genitori che il passeggero aveva a bordo se erano bambini, oppure il numero di bambini che il passeggero aveva a bordo se erano adulti. Il campo "tariffa biglietto" indica l'importo pagato per il biglietto. Il campo "imbarcato" è il luogo in cui salirono a bordo del Titanic, Cherbourg, Queenstown o Southampton.

✓ Vuoi utilizzare alcuni dati per i test?

Mantieni lo 0% per i test Utilizzare il 100% per l'allenamento

IMPORTARE

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
settare la % di test set ed
importare il template

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

I tuoi progetti di machine learning

Phishing

Riconoscendo **numeri** come **phishing or safe**

+ Aggiungi un
nuovo
progetto

Copia
modello

Condividi



Titoli dei giornali del Regno Unito

Riconoscendo **testo** come **Daily_Mail, Daily_Telegraph or 2 other classes**

Condividi



Sopravvissuti al Titanic

Riconoscendo **numeri** come **survived or did_not_survive**

Condividi



Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
accesso al progetti

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

...
...

Addestramento

Raccogli esempi di ciò che vuoi che il computer riconosca

[Addestramento](#)

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

[Impara & Testa](#)

Programma

Utilizza il modello di machine learning che hai addestrato per realizzare un gioco o un'app, in Scratch o in Python

[Programma](#)

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
addestramento

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Riconoscendo **numeri** sono **sopravvissuto o non sono sopravvissuto**

< Ritorna al progetto

sopravvissuto

classe del biglietto 2 genere femmina età 33 fratello, sp. 1 par. cap. 2 tariffa del biglietto 27.75 imbarcato Southampton	classe del biglietto 2 genere femmina età 32.5 fratello, sp. 0 par. cap. 0 tariffa del biglietto 13 imbarcato Southampton	classe del biglietto 1 genere femmina età 51 fratello, sp. 1 par. cap. 0 tariffa del biglietto 77.9583 imbarcato Southampton
classe del biglietto 3 genere femmina età 36 fratello, sp. 1 par. cap. 0 tariffa del biglietto 17.4 imbarcato Southampton	classe del biglietto 1 genere femmina età 54 fratello, sp. 1 par. cap. 0 tariffa del biglietto 78.2667 imbarcato Cherbourg	classe del biglietto 1 genere femmina età 39 fratello, sp. 1 par. cap. 1 tariffa del biglietto 83.1583 imbarcato Cherbourg
classe del biglietto 2 genere femmina età 29 fratello, sp. 1 par. cap. 0	classe del biglietto 3 genere femmina età 18 fratello, sp. 0 par. cap. 0	classe del biglietto 3 genere femmina età 2 fratello, sp. 0 par. cap. 1

non _è_ sopravvissuto

classe del biglietto 3 genere femmina età 30 fratello, sp. 1 par. cap. 1 tariffa del biglietto 24.15 imbarcato Southampton	classe del biglietto 1 genere maschio età 56 fratello, sp. 0 par. cap. 0 tariffa del biglietto 30.6958 imbarcato Cherbourg	classe del biglietto 3 genere maschio età 22 fratello, sp. 0 par. cap. 0 tariffa del biglietto 7.25 imbarcato Southampton
classe del biglietto 2 genere maschio età 54 fratello, sp. 0 par. cap. 0 tariffa del biglietto 14 imbarcato Southampton	classe del biglietto 3 genere maschio età 44 fratello, sp. 0 par. cap. 0 tariffa del biglietto 8.05 imbarcato Southampton	classe del biglietto 3 genere maschio età 18 fratello, sp. 0 par. cap. 0 tariffa del biglietto 7.775 imbarcato Southampton
classe del biglietto 3 genere maschio età 26 fratello, sp. 1 par. cap. 0	classe del biglietto 3 genere maschio età 8 fratello, sp. 4 par. cap. 1	

+ Aggiungi un esempio

+ Aggiungi una nuova etichetta

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
addestramento

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Addestramento

Raccogli esempi di ciò che vuoi che il computer riconosca

Addestramento

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

Impara & Testa

Programma

Utilizza il modello di machine learning che hai addestrato per realizzare un gioco o un'app, in Scratch o in Python

Programma

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
Impara e testa

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Dati di test:

Hai mantenuto alcuni dati del modello di progetto fuori dal set di training, in modo che possano essere utilizzati per i test. Fare clic sul pulsante in basso per scaricare un file CSV.

[Scarica i dati del test](#)

Informazioni dal computer di addestramento:

Il modello ha iniziato il suo addestramento alle: Sabato 20 gennaio 2024 08:46
Stato attuale del modello: Disponibile

[Cancella questo modello](#)

[Addestra nuovi modelli di machine learning](#)



```
File Modifica Formato Visualizza ?  
testdata-f7c7dc0-b766-11ee-be3e-03ad2519be06.csv - Blocco note  
"ticket class","gender","age","sibl. sp.","par. ch.,"ticket fare","embarked","label"  
3,"male",16,0,0,7.775,"Southampton","did_not_survive"  
2,"female",40,1,1,39,"Southampton","survived"  
2,"female",23,0,0,13.7917,"Cherbourg","survived"  
3,"male",12,1,0,11.2417,"Cherbourg","survived"  
3,"male",26,1,0,7.8542,"Southampton","did_not_survive"  
3,"male",40,0,0,7.8958,"Southampton","did_not_survive"  
3,"male",28,0,0,9.5,"Southampton","did_not_survive"  
1,"female",22,0,1,55,"Southampton","survived"  
2,"male",23,0,0,13,"Southampton","did_not_survive"  
1,"male",45,1,0,83.475,"Southampton","did_not_survive"  
2,"male",29,1,0,21,"Southampton","did_not_survive"  
1,"female",23,3,2,263,"Southampton","survived"  
2,"male",35,0,0,10.5,"Southampton","did_not_survive"  
3,"female",10,0,2,24.15,"Southampton","did_not_survive"  
2,"female",30,0,0,12.35,"Queenstown","survived"
```

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
Scarica il test set

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Cos'hai fatto?

Hai addestrato un modello di apprendimento automatico per riconoscere quando i numeri sono survived o did_not_survive.

Hai creato il modello su Saturday, January 20, 2024 8:46 AM.

Hai raccolto:

- 208 esempi di sopravvissuti ,
- 290 esempi di did_not_survive

E adesso che succede?

Prova a testare il modello di apprendimento automatico qui sotto. Inserisci un esempio di un numero che non hai incluso negli esempi che hai usato per addestrarlo. Ti dirà come lo riconosce e quanto è sicuro di sé.

Se il computer sembra aver imparato a riconoscere correttamente le cose, allora si può andare su Scratch e usare ciò che il computer ha imparato per creare un gioco!

Se il computer sta sbagliando troppe cose, potresti voler tornare alla pagina Addestramento e raccogliere altri esempi.

Una volta fatto questo, clicca sul pulsante qui sotto per addestrare un nuovo modello di apprendimento automatico e vedere quale differenza faranno gli esempi extra!

Prova a inserire alcuni numeri per vedere come viene riconosciuto in base al tuo addestramento.

classe del biglietto	<input type="text"/>
genere	maschio <input type="button" value="▼"/>
età	<input type="text"/>
fratello, sp.	<input type="text"/>
par. cap.	<input type="text"/>
tariffa del biglietto	<input type="text"/>
imbarcato	Cherbourg <input type="button" value="▼"/>

Prova a inserire alcuni numeri per vedere come viene riconosciuto in base al tuo addestramento.

classe del biglietto	2
genere	Donnina <input type="button" value="▼"/>
età	23
fratello, sp.	0
par. cap.	0
tariffa del biglietto	13.7917
imbarcato	Cherbourg <input type="button" value="▼"/>

Riconosciuto come **survived**
con il 100% di fiducia

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
Testa il Modello

```
"ticket class","gender","age","sibl. sp.","par. ch.","ticket fare","embarked","label"  
3,"male",16,0,0,7.775,"Southampton","did_not_survive"  
2,"female",40,1,1,39,"Southampton","survived"  
2,"female",23,0,0,13.7917,"Cherbourg","survived"  
3,"male",12,1,0,11.2417,"Cherbourg","survived"  
3,"male",26,1,0,7.8542,"Southampton","did_not_survive"
```

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Prova a inserire alcuni numeri per vedere come viene riconosciuto in base al tuo addestramento.

classe del biglietto

genere

maschio

età

fratello, sp.

par. cap.

tariffa del biglietto

imbarcato

Cherbourg

Prova

Descrivi il tuo modello!

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
descrizione/spiegabilità del
Modello

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Comprendi il tuo modello di machine learning

< Torna al modello

La tecnica utilizzata per creare il tuo modello di machine learning è chiamata **Decision Tree Classifier**.

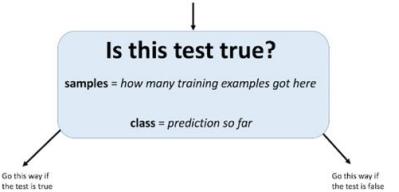
Questo non è l'unico modo per addestrare un modello di machine learning. Lo usiamo perché è molto veloce e facile da addestrare ed è una delle tecniche più facili da capire. Questa pagina mostra l'albero decisionale creato sulla base degli esempi di formazione che hai raccolto.

Quando provi il tuo modello, il computer inizia dalla cima dell'albero e segue un percorso fino a raggiungere il fondo. La **classe** nella parte inferiore dell'albero rappresenta la previsione effettuata dal modello di machine learning.

In ogni casella dell'albero si legge il test descritto nella parte superiore della casella. Se i valori del test superano il test descritto nella casella, segue la freccia a sinistra. Se non supera il test, segue la freccia a destra.

Gli **esempi** mostrati in ogni casella indicano quanti esempi nei dati di addestramento corrispondono a quella parte dell'albero decisionale.

Il **valore** mostrato in ogni casella indica quanti esempi nei dati di addestramento hanno superato il test mostrato in alto (seguendo la freccia a sinistra) e quanti esempi non hanno superato il test in alto (seguendo la freccia a destra).



Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
descrizione/spiegabilità del
Modello

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Comprendi il tuo modello di machine learning

< Torna al modello

La tecnica utilizzata per creare il tuo modello di machine learning è chiamata **Decision Tree Classifier**.

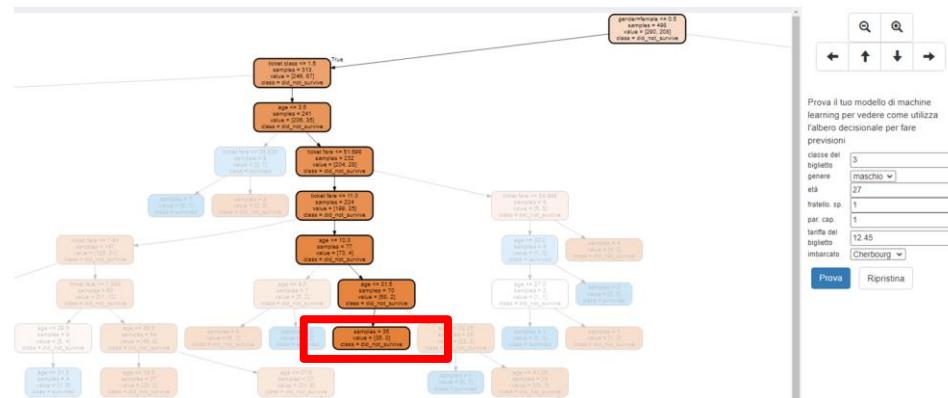
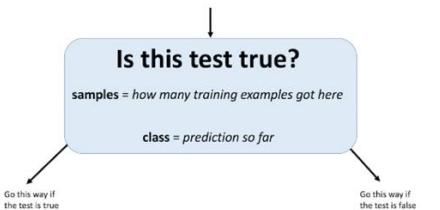
Questo non è l'unico modo per addestrare un modello di machine learning. Lo usiamo perché è molto veloce e facile da addestrare ed è una delle tecniche più facili da capire. Questa pagina mostra l'albero decisionale creato sulla base degli esempi di formazione che hai raccolto.

Quando provi il tuo modello, il computer inizia dalla cima dell'albero e segue un percorso fino a raggiungere il fondo. La **classe** nella parte inferiore dell'albero rappresenta la previsione effettuata dal modello di machine learning.

In ogni casella dell'albero si legge il test descritto nella parte superiore della casella. Se i valori del test superano il test descritto nella casella, segue la freccia a sinistra. Se non supera il test, segue la freccia a destra.

Gli **esempi** mostrati in ogni casella indicano quanti esempi nei dati di addestramento corrispondono a quella parte dell'albero decisionale.

Il **valore** mostrato in ogni casella indica quanti esempi nei dati di addestramento hanno superato il test mostrato in alto (seguendo la freccia a sinistra) e quanti esempi non hanno superato il test in alto (seguendo la freccia a destra).



Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
descrizione/spiegabilità del
Modello

FUTURA
LA SCUOLA
PER L'ITALIA DI DOMANI

Finanziato dall'Unione europea
e dal Mezzogiorno
Italiadomani

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic

Addestramento

Raccogli esempi di ciò che vuoi che il computer riconosca

Addestramento

Impara & Testa

Usa gli esempi per addestrare il computer a riconoscere

Impara & Testa

Programma

Utilizza il modello di machine learning che hai addestrato per realizzare un gioco o un'app, in Scratch o in Python

Programma

Machine Learning
For Kids... ma non
solo!!

Sopravvissuti del Titanic:
Programma

Crea qualcosa con il tuo modello di apprendimento automatico

< Ritorna ai progetti

Scratch 3

Usa la nuova versione di Scratch



Scratch 3

Python

Scrivi il codice Python per utilizzare il tuo modello di apprendimento automatico



Python

EduBlocks

Write Python using a drag and drop coding tool



EduBlocks

App Inventor

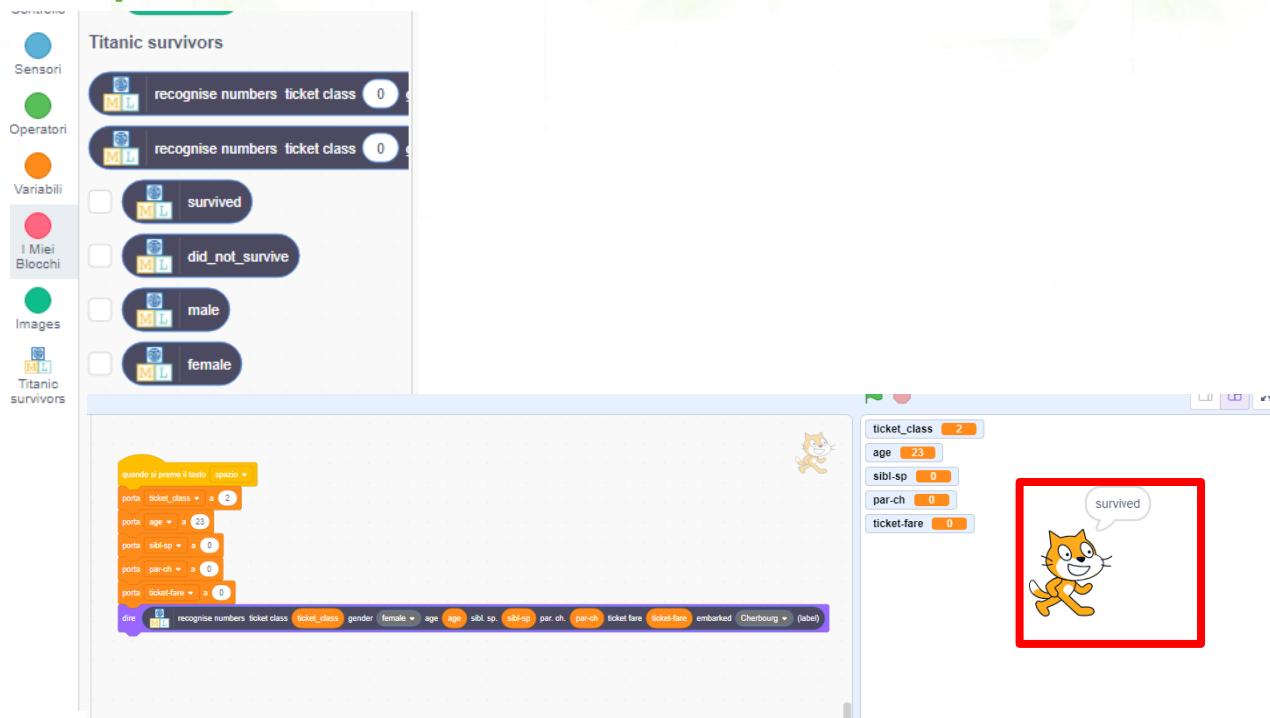
Crea un'applicazione mobile per il tuo telefono o tablet



App Inventor

Modulo 4: “Machine Learning con MachineLearning ForKids”

Sopravvissuti del Titanic



Machine Learning For Kids... ma non solo!!

Sopravvissuti del Titanic: Programma

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



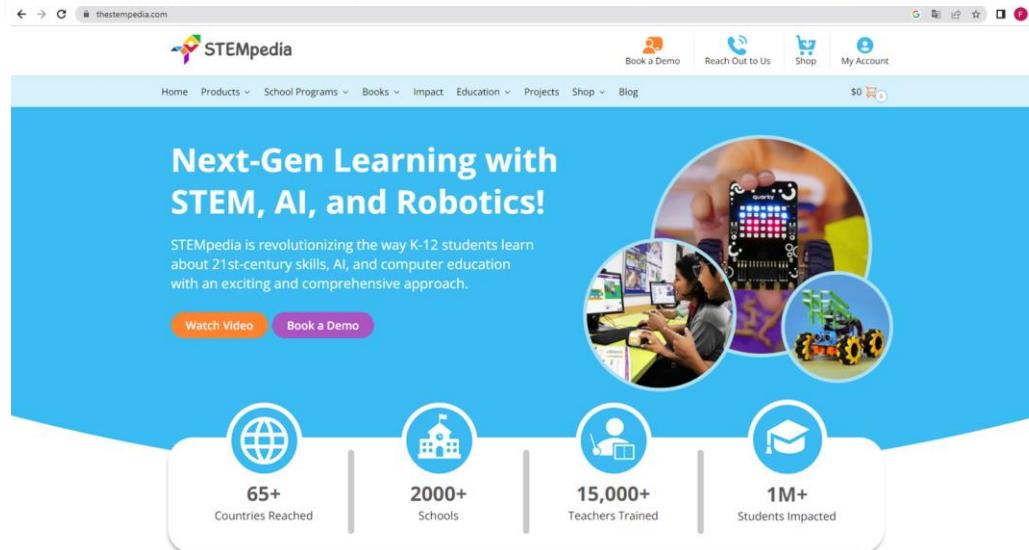
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



Machine Learning
PictoBlox

Accesso al servizio web e
portale

<https://thestempedia.com/>

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

PictoBlox Software



Discover the ultimate **creator platform** for students and teachers with PictoBlox! Easily create projects in **AI, ML, AR, VR, IoT, and hardware with Arduino, LEGO, and Quarky** while enjoying the ease, fun, and engagement that PictoBlox's amazing features offer. From beginner coders to experienced developers, discover why PictoBlox is the best choice for **all your coding needs!**

Watch Video

Explore PictoBlox



Block Coding

Learn to code with colorful puzzles like drag and drop block based coding for Kids.



Python Coding

Create robust programs with simplified interactive codes in Python programming for teens.



Machine Learning

Analyze complex data, train and import models to coding environment with ease.



3D, AR, VR Environment

Create 3D Environments and explore new dimensions of creativity with AR & VR.

Machine Learning
PictoBlox

Accesso al servizio web e
portale

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

PictoBlox: Creator Tool for Coding, AI, AR-VR and Robotics

Revolutionary Block-based and Python programming platform for kids and teens to learn coding through creative and engaging projects.



Windows Apple Linux Chrome OS Android iOS

[Try Now](#) [Download](#)

95+ Countries Reached 4.5+ Rating on Play Store 15,000+ Teachers Trained 1M+ Students Impacted

Machine Learning
PictoBlox

Accesso al servizio web e
portale: Try Now

<https://pictoblox.ai/>

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

What would you like to do?



Block Coding

Start coding by just dragging puzzle-shaped blocks.



Python Coding

Make programs with text (syntax) based coding in Python.

Set as a default coding environment



Machine Learning Environment

Train ML models for image, object, sound, body/hand pose, NLP, and numbers.



AR, VR and 3D Environment

Create interactive experience with Augmented Reality and 3D objects.

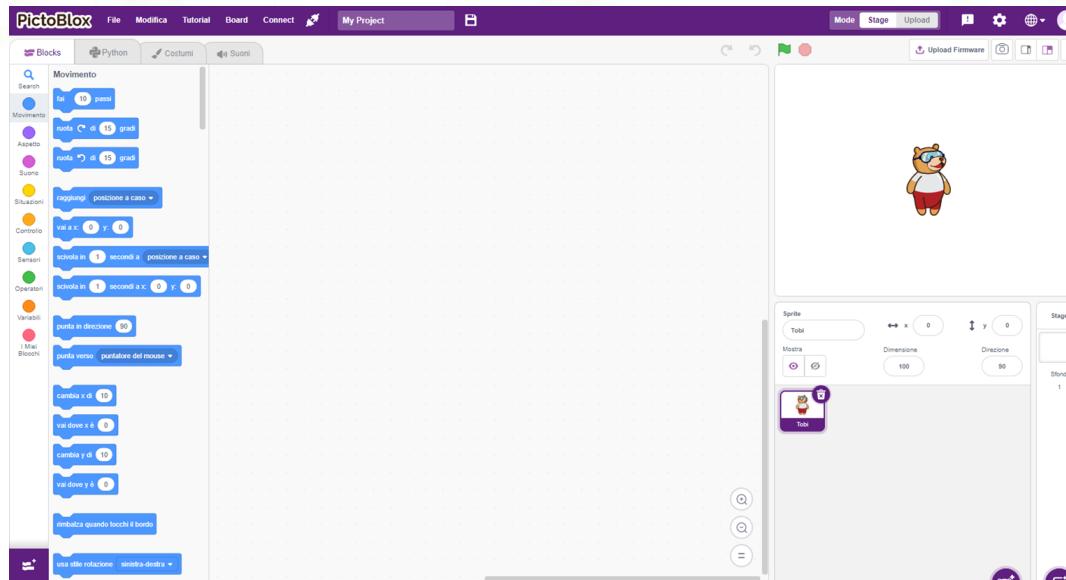
Machine Learning
PictoBlox

Accesso al servizio web e
portale: Block Coding

<https://pictoblox.ai/>

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



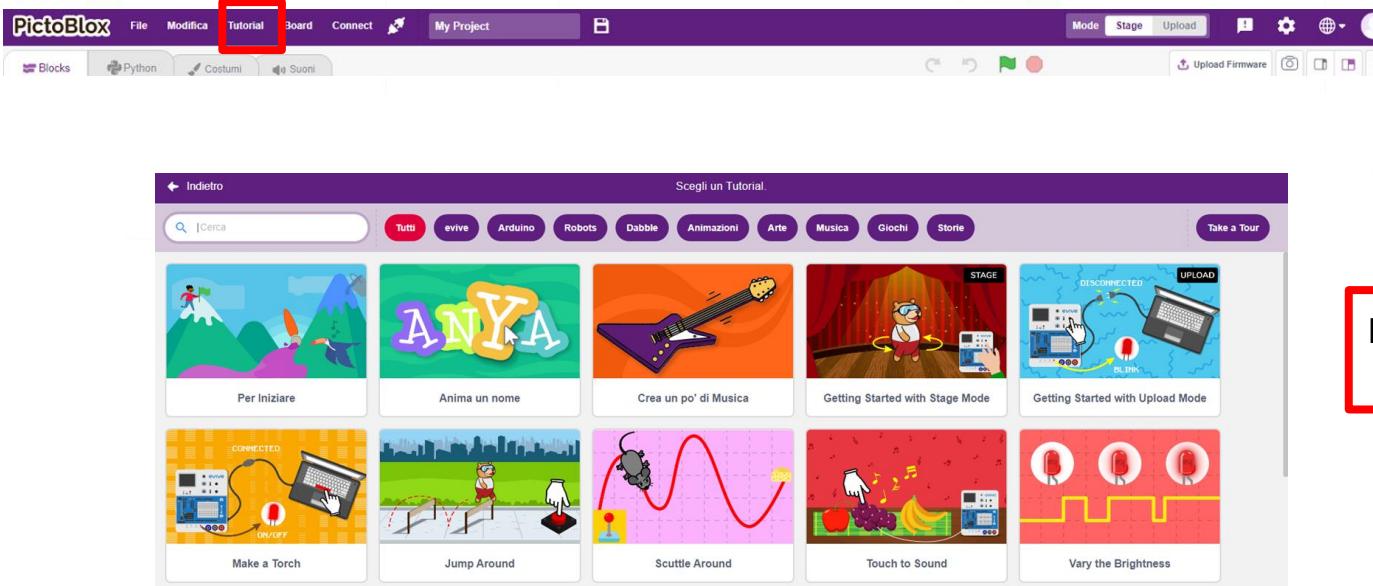
Machine Learning
PictoBlox

Accesso al servizio web e
portale: Block Coding

<https://pictoblox.ai/>

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



The screenshot shows the PictoBlox web interface. At the top, there is a purple navigation bar with tabs for "PictoBlox", "File", "Modifica", "Tutorial" (which is highlighted with a red box), "Board", "Connect", "My Project", "Mode", "Stage", "Upload", and various icons for settings and help. Below the navigation bar, there are tabs for "Blocks", "Python", "Costumi", and "Suoni". The main area is titled "Scegli un Tutorial" and contains a grid of 10 cards, each with an illustration and a title:

- Per Iniziare
- Anima un nome
- Crea un po' di Musica
- Getting Started with Stage Mode
- Getting Started with Upload Mode
- Make a Torch
- Jump Around
- Scuttle Around
- Touch to Sound
- Vary the Brightness

Machine Learning
PictoBlox

Per iniziare: Tutorial

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

The screenshot shows the PictoBlox software interface. At the top, there is a menu bar with "PictoBlox", "File", "Modifica", "Tutorial", "Board" (which is highlighted with a red box), "Connect", "My Project", and a save icon. Below the menu is a toolbar with icons for "Blocks", "Python", "Costumi", and "Suoni". The main area is divided into two sections: "Select Board" on the left and a "Project Workspace" on the right.

Select Board: This section displays various board options with their respective icons. The "Arduino Uno" icon is highlighted with a red box. Other boards shown include evive, Quarky, Wizbot (Alpha), Arduino Mega, Arduino Nano, ESP32, T-Watch, micro:bit, TECbits, LEGO EV3, and LEGO BOOST.

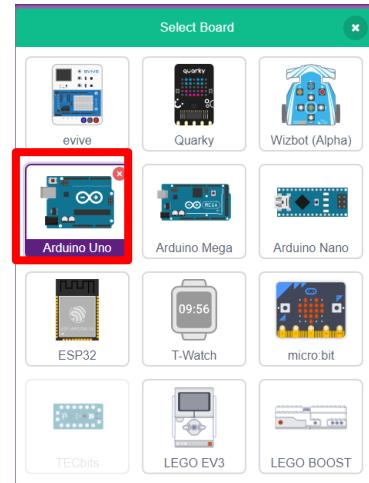
Project Workspace: This section shows a project titled "Arduino Uno" with the following blocks:

- when Arduino Uno starts up
- read status of digital pin 2
- read analog pin A0
- set digital pin 13 output as HIGH
- set PWM pin 3 output as 255
- play tone on 3 of note C2 &
- get timer value
- reset timer
- cast 15.5 to integer
- map 50 from 0 ~ 255 to

Actuators: This section contains blocks for connecting motors, running motors, freeing motors, and setting servos.

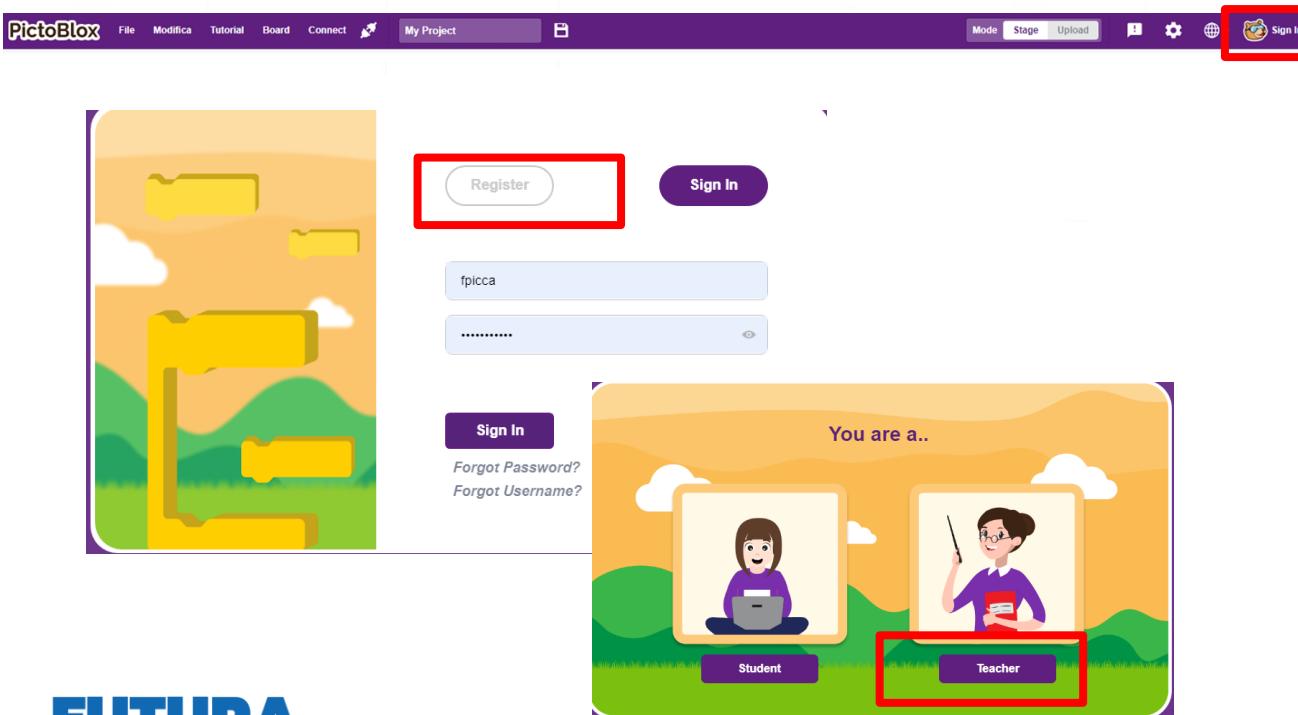
Machine Learning
PictoBlox

Per iniziare: Board (collega schede hardware) e Connect



Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

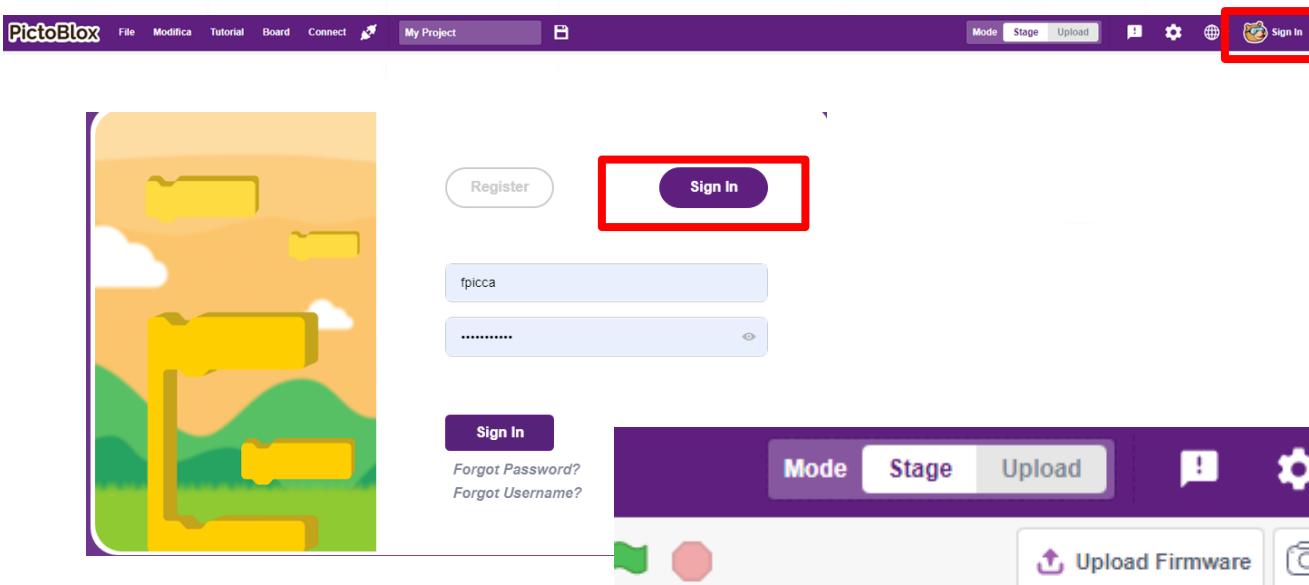


Machine Learning
PictoBlox

Per iniziare: Registrazione
ed accesso

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



Machine Learning
PictoBlox

Per iniziare: Registrazione
ed accesso

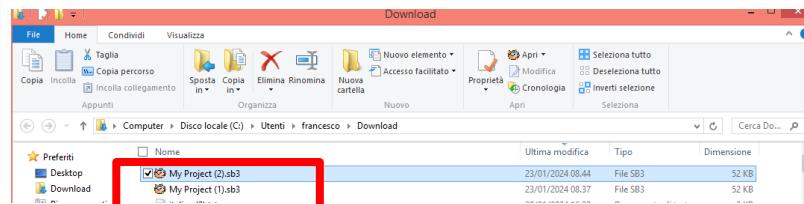
Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

The screenshot shows the PictoBlox web application interface. At the top, there's a purple header bar with the PictoBlox logo, navigation links like 'File', 'Modifica', 'Tutorial', 'Board', 'Connect', and a search icon. Below the header is a toolbar with 'Mode', 'Stage', 'Upload', settings, and a 'Sign In' button. A red box highlights the 'My Project' button in the top right corner of the header. The main area has a purple sidebar with categories like 'Blocks', 'Search', 'Movimento', 'Aspetto', 'Suono', and 'Cittadina'. The main workspace shows some code blocks. A red box highlights the 'File' menu, which contains options: 'Nuovo', 'Carica dal computer', 'Salva sul tuo computer' (which is bolded), 'File sharing', 'Open file from link', 'Open file from QR code', and 'Examples'.

Machine Learning
PictoBlox

Per iniziare: Salva in locale
un progetto



Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

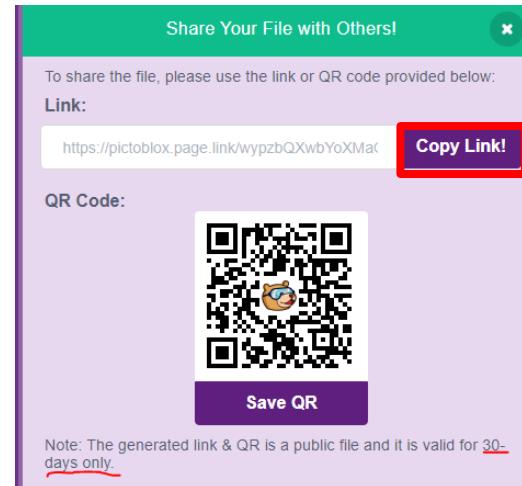
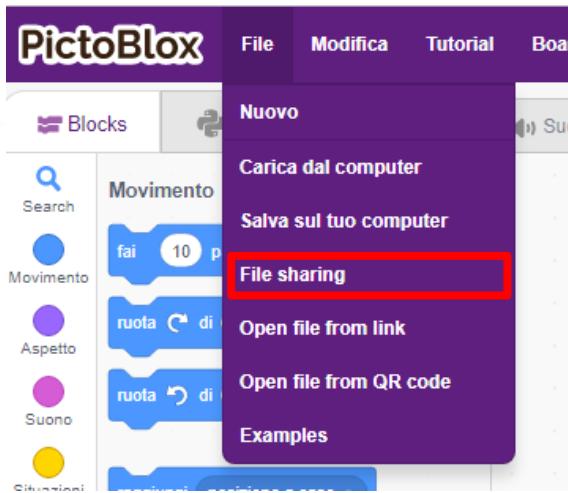
The screenshot shows the PictoBlox web application interface. At the top, there's a purple header bar with the PictoBlox logo, navigation links like 'File', 'Modifica', 'Tutorial', 'Board', 'Connect', and a 'Sign In' button. Below the header is a search bar with 'My Project' and a file icon, which is highlighted with a red box. The main area has a purple sidebar on the left with categories: 'Blocks', 'Search', 'Movimento' (with icons for 'fai', 'ruota', 'Aspetto', 'Suono', and 'Cittadino'), and a 'Movimento' section with blue blocks. The central workspace shows a 'Nuovo' (New) project with a dropdown menu open. The 'Carica dal computer' (Upload from computer) option is highlighted with a red box. Other options in the menu include 'Salva sul tuo computer' (Save to your computer), 'File sharing', 'Open file from link', 'Open file from QR code', and 'Examples'. To the right of the workspace, there's a preview area showing various project icons.

Machine Learning
PictoBlox

Per iniziare: Upload locale di
un progetto

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



Machine Learning
PictoBlox

Per iniziare: Genera il LINK
del Progetto

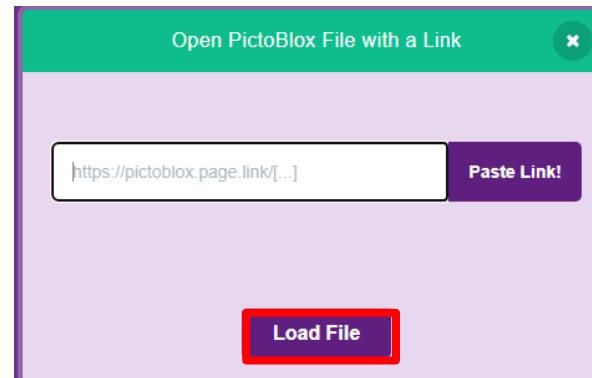
Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

The screenshot shows the PictoBlox web application interface. At the top, there's a purple header bar with the PictoBlox logo, navigation links like 'File', 'Modifica', 'Tutorial', 'Board', 'Connect', and a 'My Project' button. Below the header is a toolbar with icons for 'Mode', 'Stage', 'Upload', settings, and a user profile. The main area has a dark background with a sidebar on the left containing categories like 'Blocks', 'Search', 'Movimento', 'Aspetto', 'Suono', and 'Cittadella'. A central workspace shows some code blocks. A red box highlights the 'File' menu, which is open and displays options: 'Nuovo', 'Carica dal computer', 'Salva sul tuo computer', 'File sharing', 'Open file from link' (which is also highlighted with a red box), 'Open file from QR code', and 'Examples'.

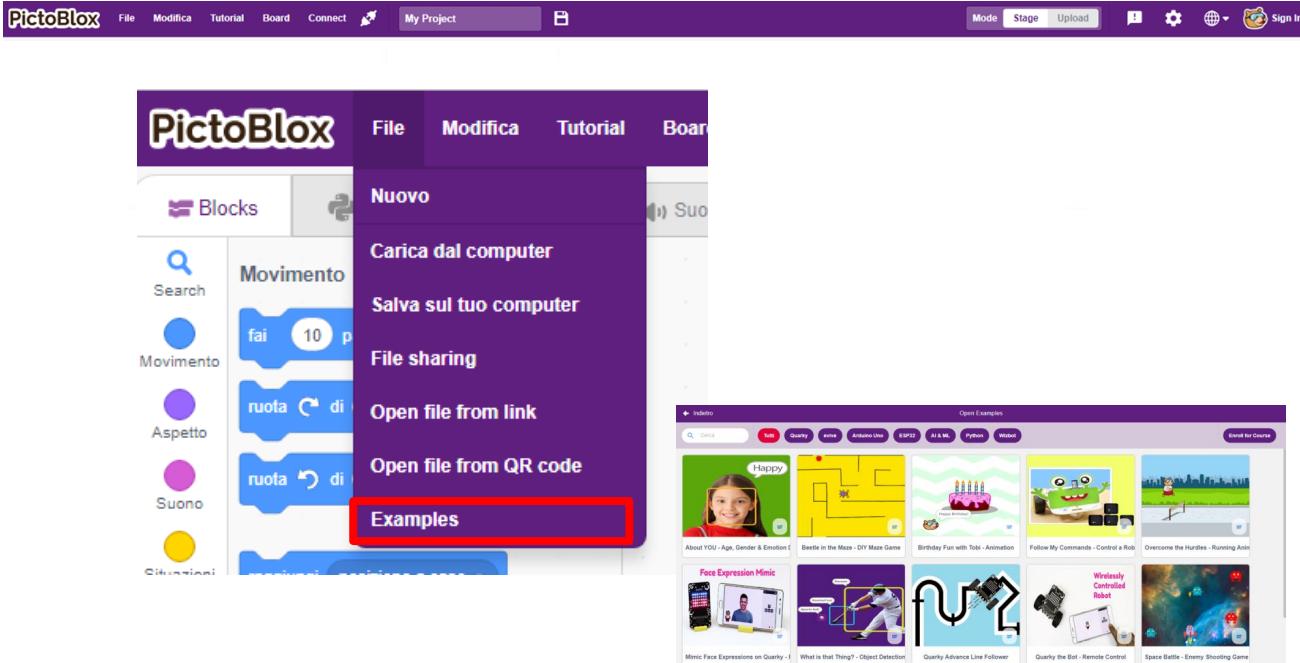
Machine Learning
PictoBlox

Per iniziare: Accedi al LINK
di un Progetto



Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



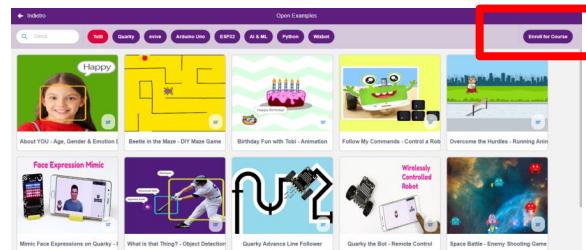
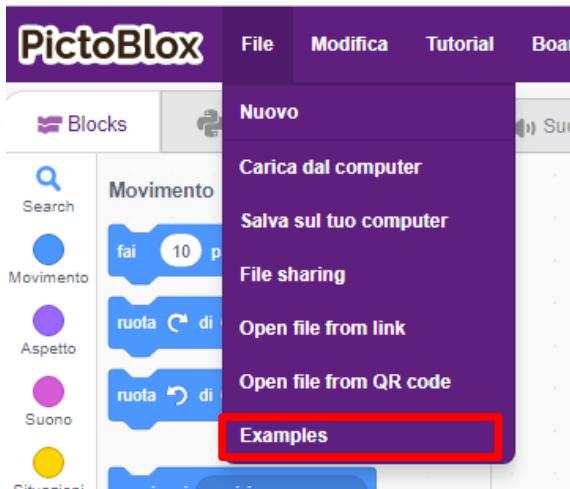
The screenshot shows the PictoBlox web application interface. At the top, there is a purple header bar with the PictoBlox logo, navigation links (File, Modifica, Tutorial, Board, Connect), and a sign-in button. Below the header is a toolbar with icons for Mode, Stage, Upload, and settings. The main area has a sidebar on the left containing categories like Blocks, Movimento, Aspetto, Suono, and Comunicazione, each with corresponding icons and a list of blocks. The central workspace shows a project with some blocks. A dropdown menu is open under the 'File' tab, listing options: Nuovo, Carica dal computer, Salva sul tuo computer, File sharing, Open file from link, Open file from QR code, and Examples. The 'Examples' option is highlighted with a red box. To the right of the workspace, there is a preview window showing various project examples.

Machine Learning
PictoBlox

Per iniziare: Accedi agli
ESEMPI

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale



Machine Learning
PictoBlox

Per iniziare: Accedi alla
FORMAZIONE

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

The screenshot shows the PictoBlox web interface. At the top, there's a navigation bar with links for File, Modifica, Tutorial, Board, Connect, My Project, Mode, Stage, Upload, Settings, and Sign In. Below the navigation is a search bar and a "Cerca" button. A "Open Examples" section displays several thumbnail images of projects: "About YOU - Age, Gender & Emotion Detection", "Beetle in the Maze - DIY Maze Game", "Birthday Fun with Tobi - Animation", "Follow My Commands - Control a Robo", "Overcome the Hurdles - Running Animation", "Face Expression Mimic", "What is that Thing? - Object Detection", "Quarky Advance Line Follower", "Wirelessly Controlled Robot", and "Quarky the Bot - Remote Control". A red box highlights the "Enroll for Course" button in the top right corner of this section.

Machine Learning
PictoBlox

Per iniziare: Accedi alla
FORMAZIONE (stesso
account ACCESSO
PictoBlox)

The screenshot shows the STEMpedia Educator portal. It features a sidebar with icons for Home, Courses, Projects, Gallery, Shop, Logout, and a user profile for "Francesco Picca" (Joined Aug 2023 • Active now). The main area has tabs for "My Courses", "My Certificates", "My Badges", and "Profile". Under "My Courses", there's a card for "121 lessons Empowering Educators: An International AI, Python, C++ course designed to kick start your educator bootcamp... 100% Complete".

FUTURA
LA SCUOLA
PER L'ITALIA DI DOMANI

Progetto dell'Unione europea
Italiadomani

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

The screenshot shows the STEMpedia Education website. At the top, there's a navigation bar with icons for dashboard, courses, projects, gallery, shop, and logout. A search bar and a user profile for 'Francesco Picca' are also present. Below the bar, a sidebar on the left contains icons for dashboard, courses, certificates, badges, and profile. The main area displays 'My Courses', 'My Certificates', 'My Badges', and 'Profile'. A red box highlights the 'Open Resources' button in the top navigation bar.

The screenshot shows the PictoBlox Documentation page. The header includes a sidebar with icons for dashboard, courses, certificates, badges, and profile, followed by the 'STEMpedia' logo and a search bar. The main content area has a purple header with the title 'PictoBlox - Documentation'. Below it, there are several cards for 'Getting Started with Machine Learning Environment'. Each card includes a preview image of the software interface, a title, a brief description, and a 'Explore More Tutorials' button at the bottom. The cards are for 'IMAGE CLASSIFIER (ML) in Block Coding', 'IMAGE CLASSIFIER (ML) in Python', 'OBJECT DETECTION (ML) in Block Coding', and 'OBJECT DETECTION (ML) in Python Coding'.

Machine Learning
PictoBlox

Per iniziare: Accedi a TUTTE
le Open Resources di PICTO
BLOX

Modulo 5: “Machine Learning con PictoBlox”

Accesso al servizio web e Portale

PictoBlox: Creator Tool for Coding, AI, AR-VR and Robotics

Revolutionary Block-based and Python programming platform for kids and teens to learn coding through creative and engaging projects.

Try Now Download

95+ Countries Reached 4.5+ Rating on Play Store 15,000+ Teachers Trained 1M+ Students Impacted

A GESS EDUCATION AWARDS 2022 badge is displayed above the software interface.

Machine Learning
PictoBlox

Accesso al servizio web e portale: Download
(necessario per utilizzare la
Machine Learning Environment)

<https://thestempedia.com/product/pictoblox/download-pictoblox/>

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



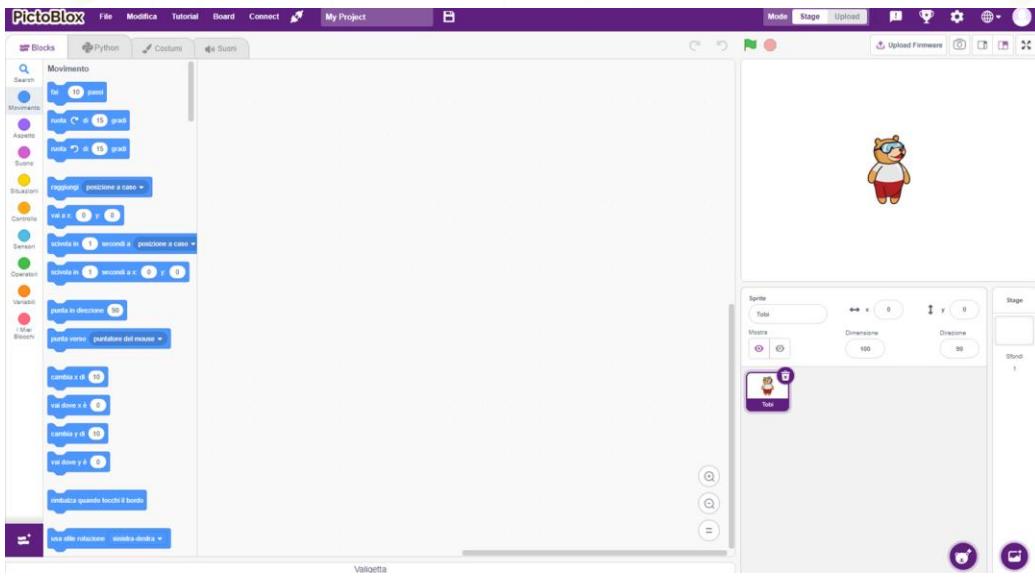
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING

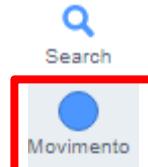


Machine Learning
PictoBlox

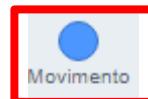
Riallineamento di CODING

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



Search



Movimento



Aspetto



Suono



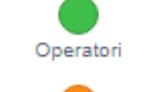
Situazioni



Controllo



Sensori



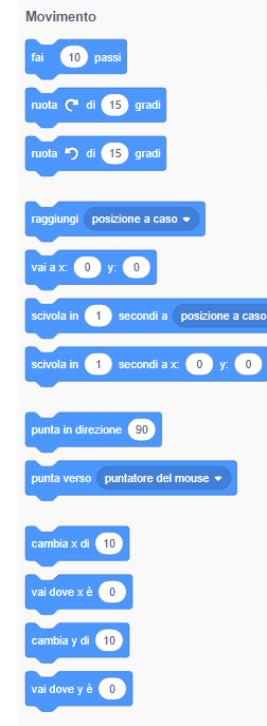
Operatori



Variabili



I Miei
Blochi

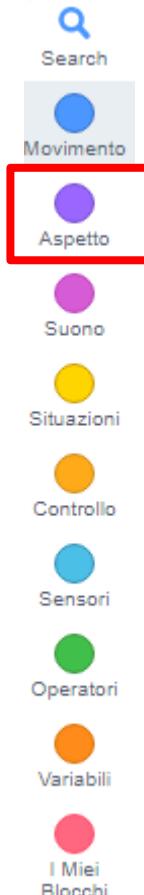


Machine Learning
PictoBlox

Riallineamento di CODING:
MOVIMENTO DEGLI
SPRITE

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



```
dire Ciao! per 2 secondi
dire Ciao!
pensa Uhm... per 2 secondi
pensa Uhm...
passa al costume Tobi Jump
passa al costume seguente
passa allo sfondo backdrop1
passa allo sfondo seguente
cambia dimensione di 10
porta dimensione a 100 %
cambia effetto colore di 25
porta effetto colore a 0
rimuovi effetti grafici
mostra
nasconde
```

Machine Learning
PictoBlox

Riallineamento di CODING:
ASPETTO DEGLI SPRITE

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



Suono

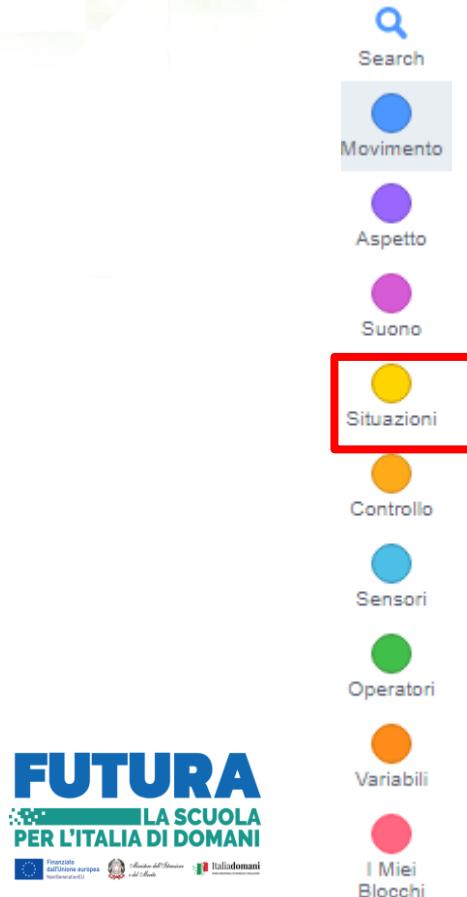
- riproduci suono Grunt ▾ e attendi la fine
- avvia riproduzione suono Grunt ▾
- ferma tutti i suoni
- cambia effetto frequenza ▾ di 10
- porta effetto frequenza ▾ a 100
- rimuovi effetti audio
- cambia volume di -10
- porta volume a 100 %
- volume

Machine Learning
PictoBlox

Riallineamento di CODING:
SUONO

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



Situazioni

- quando si clicca su
- quando si preme il tasto spazio ▾
- quando si clicca questo sprite
- quando lo sfondo passa a backdrop1 ▾
- quando volume microfono ▾ > 10
- quando ricevo messaggio1 ▾
- invia a tutti messaggio1 ▾
- invia a tutti messaggio1 ▾ e attendi

Machine Learning
PictoBlox

Riallineamento di CODING:
SITUAZIONI

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



Machine Learning
PictoBlox

Riallineamento di CODING:
CONTROLLO

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING

The screenshot shows the PictoBlox software interface. On the left, there is a sidebar with various categories: Search (blue magnifying glass), Movimento (blue circle), Aspetto (purple circle), Suono (pink circle), Situazioni (yellow circle), Controllo (orange circle), Sensori (blue circle, highlighted with a red box), Operatori (green circle), Variabili (orange circle), and I Miei Blocchi (red circle). The main area is titled "Sensori" and contains the following blocks:

- sta toccando puntafore del mouse
- sta toccando il colore
- il colore sta toccando il colore
- distanza da puntafore del mouse
- chiedi Come ti chiami? e attendi
- risposta
- tasto spazio premuto
- pulsante del mouse premuto
- x del mouse
- y del mouse
- rendi lo sprite trascinabile
- volume microfono
- cronometro
- azzerà cronometro
- numero dello sfondo di Stage

Machine Learning
PictoBlox

Riallineamento di CODING:
SENSORI

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING

The screenshot shows the PictoBlox software interface. On the left, there is a sidebar with various categories: Search (blue magnifying glass), Movimento (blue circle), Aspetto (purple circle), Suono (pink circle), Situazioni (yellow circle), Controllo (orange circle), Sensors (light blue circle), and Operatori (green circle). The 'Operatori' category is highlighted with a red border. Below the sidebar, there are several code blocks categorized under 'Operatori'. These include:

- Operatori: +, -, *, /
- numero a caso tra 1 e 10
- > 50, < 50, = 50
- e, o, non
- unione di mela e banana
- lettera 1 di mela
- lunghezza di mela
- mela contiene a

Machine Learning
PictoBlox

Riallineamento di CODING:
OPERATORI

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



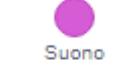
Search



Movimento



Aspetto



Suono



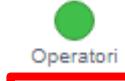
Situazioni



Controllo



Sensori



Operatori



Variabili



I Miei
Blocci

Variabili

Crea una Variabile

my variable

porta my variable a 0

cambia my variable di 1

mostra la variabile my variable

nasconde la variabile my variable

Crea una Lista

Machine Learning
PictoBlox

Riallineamento di CODING:
VARIABILI E LISTE

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING

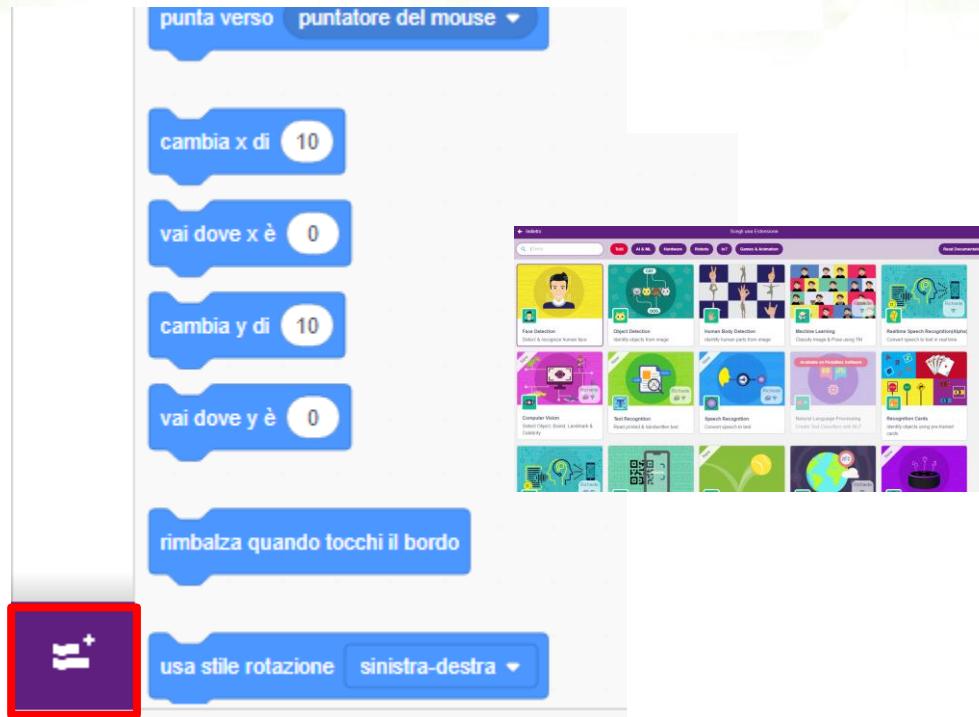
The screenshot shows the PictoBlox software interface. On the left, there is a sidebar with icons for Search, Movimento, Aspetto, Suono, Situazioni, Controllo, Sensori, Operatori, and Variabili. Below this sidebar is a button labeled "I Miei Blocchi" with a red border. In the center, there is a window titled "I Miei Blocchi" with a "Crea un Blocco" button. Below it is a larger window titled "Crea un Blocco" containing a block labeled "sommaNumeri [a b]". At the bottom of this window are four options: "Add an input number", "Add an input text", "Aggiungi un argomento booleano", and "Aggiungi una scritta". There is also a checkbox for "Esegui senza aggiornare lo schermo" and two buttons: "Annulla" and "OK".

Machine Learning
PictoBlox

Riallineamento di CODING: I
MIEI BLOCCHI PER
CREARE FUNZIONI
PROPRIE

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING

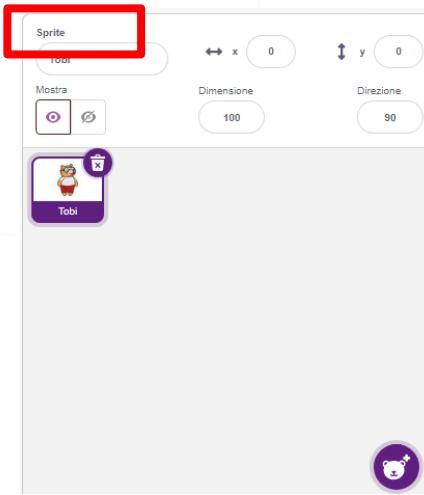


Machine Learning
PictoBlox

Riallineamento di CODING:
estensioni del linguaggio

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING

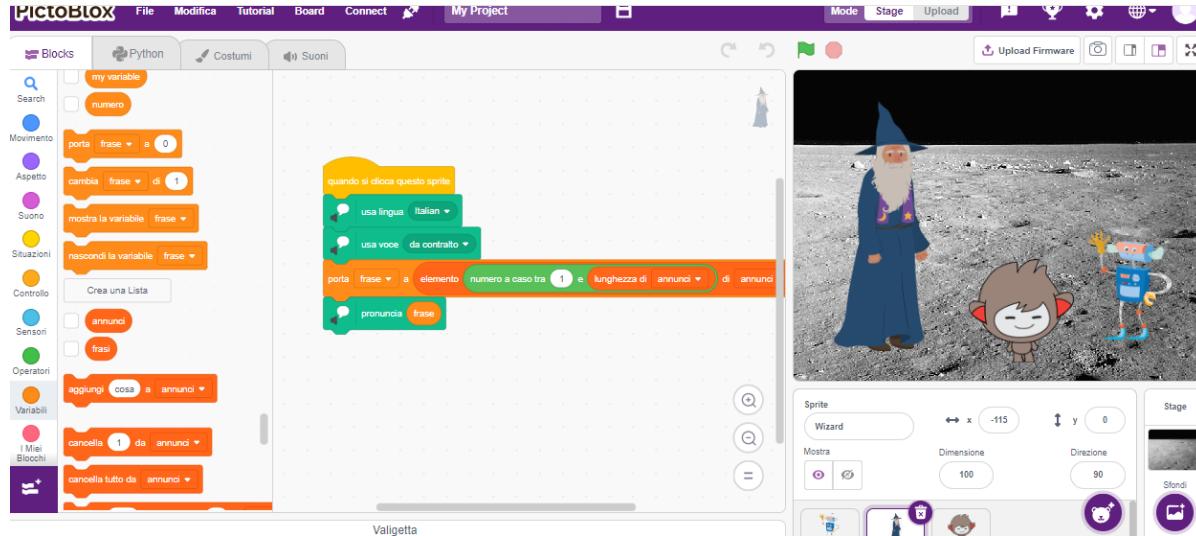


Machine Learning
PictoBlox

Riallineamento di CODING:
GLI SPRITE E LO STAGE

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



Machine Learning
PictoBlox

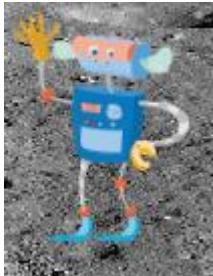
Riallineamento di CODING:
UN ESEMPIO >>

- uso di variabili e liste
- estensione da testo a voce
- traduci
- numeri casuali

<https://pictoblox.page.link/fbHXm3V9mh2pcbLo9>

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



```
when green flag is clicked
  [background v] set [spazio v]
  [say v for (1) steps]
  [change background v to Moon]
  [say v for (1) steps]
  [add (random (1 [1 v] to [10 v])) of [list v] to [frasi v]]
  [repeat (10) [
    [say v for (1) steps]
    [change background v to Moon]
    [say v for (1) steps]
    [add (random (1 [1 v] to [10 v])) of [list v] to [frasi v]]
  ]]
when space key is pressed
  [background v] set [spazio v]
  [say v for (1) steps]
  [change background v to Moon]
  [say v for (1) steps]
  [ask [name v] and wait]
  [add (random (1 [1 v] to [10 v])) of [list v] to [frasi v]]
  [repeat (10) [
    [say v for (1) steps]
    [change background v to Moon]
    [say v for (1) steps]
    [add (random (1 [1 v] to [10 v])) of [list v] to [frasi v]]
  ]]
```

Machine Learning
PictoBlox

Riallineamento di CODING:
UN ESEMPIO >> frasi
casuali e da testo a voce

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



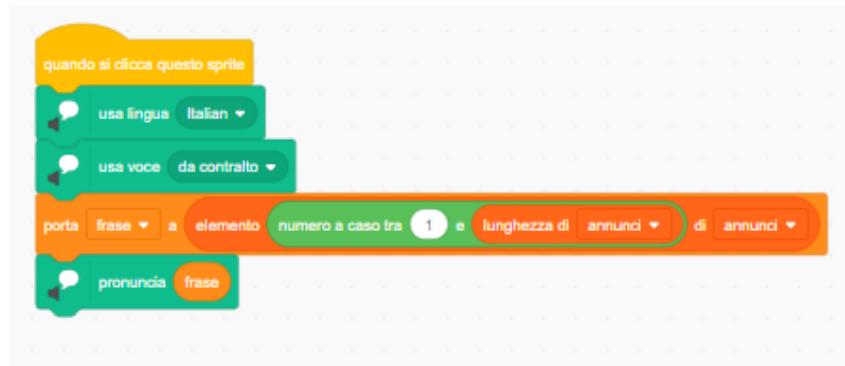
```
when green flag clicked
  say [Italian]
  say [Giochiamo alla tombola]
  set [numero v] to [1]
  say [Numero generato]
  say [numero]
  say [English]
  say [numero]
  say [2 seconds]
  say [traduci union di Numero generato e numero in Inglese]
```

Machine Learning
PictoBlox

Riallineamento di CODING:
UN ESEMPIO >> numeri
casuali

Modulo 5: “Machine Learning con PictoBlox”

Riallineamento di CODING



Machine Learning
PictoBlox

Riallineamento di CODING:
UN ESEMPIO

FUTURA

LA SCUOLA
PER L'ITALIA DI DOMANI



Finanziato
dall'Unione europea
NextGenerationEU

Analisi dei dati e Modelli di Machine Learning attraverso soluzioni block oriented

Esperto Prof. Francesco Picca



Ministero dell'Istruzione
e del Merito



Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA



I.I.T. Pacinotti
una scuola per il futuro

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

The screenshot shows a web browser window with the URL <https://ai.thestempedia.com/docs/pictoblox/machine-learning-environment/>. The page title is "Machine Learning Environment - PictoBloX". The main content area features a diagram titled "MACHINE LEARNING IN PICTOBLOX" showing various machine learning models represented as hexagons: Object Detection (Cat, Dog), Pose Classifier (Dancer), Text Classifier (Document), Audio Classifier (Headphone), Image Classifier (Image), Hand-Pose Classifier (Hand), and Number Classifier & Regression (Graph). Below the diagram, a text block explains that PictoBloX provides a comprehensive Machine-Learning Environment for kids and beginners, offering 7 models including Image, Object Detection, Pose, Hand-Pose, Audio, and Number Classification, as well as Regression and Text Classification. Training is done in three steps: Collect data, Train the model, and Test the model. It includes a step-by-step guide and video tutorials.

Machine Learning
PictoBloX

Machine-Learning-Environment
(solo versione PictoBloX
Desktop)

<https://ai.thestempedia.com/docs/pictoblox/machine-learning-environment/>

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

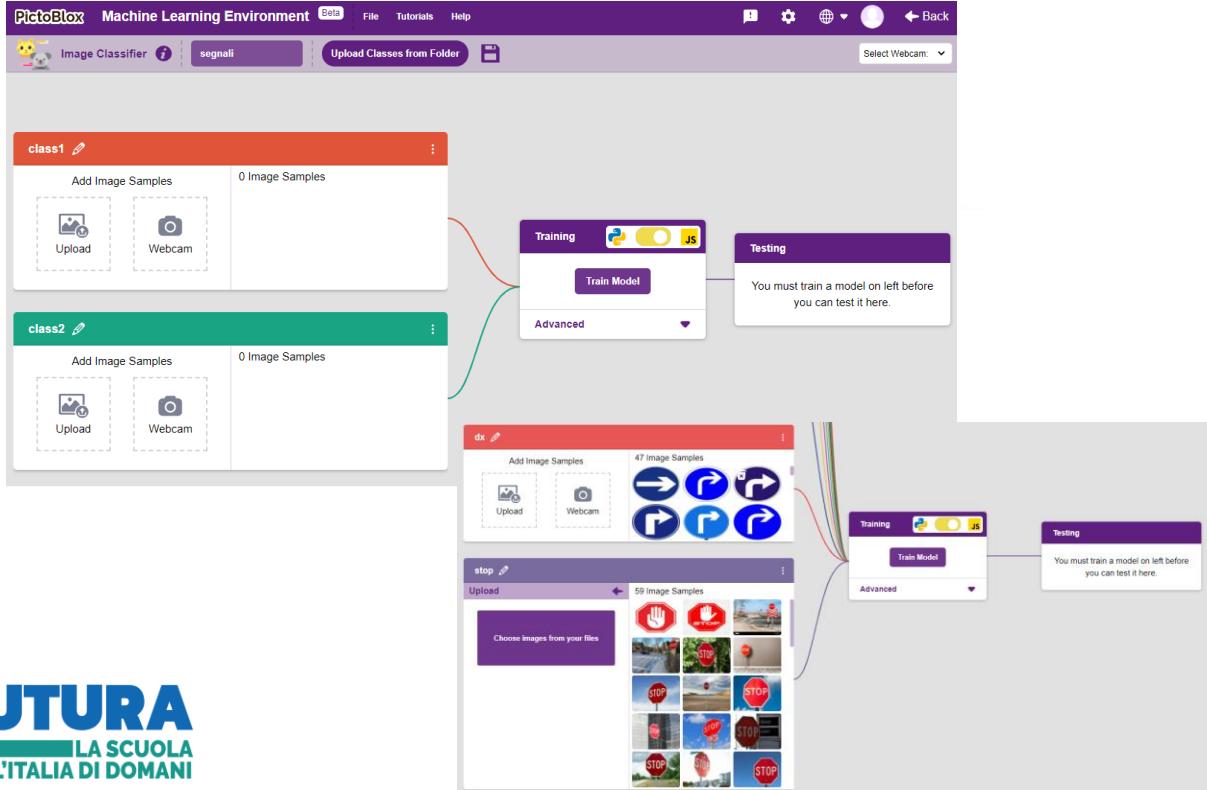
The screenshot shows the PictoBlox Machine Learning Environment interface. At the top, there's a purple header bar with the PictoBlox logo, "Machine Learning Environment", and navigation links like "File", "Tutorials", and "Help". Below the header is a toolbar with "My Projects", a search bar, and buttons for "Create New Project" and "Open Project". The main area has a title "Project Details" and columns for "Type", "No. of Classes", "Last Updated", and "Status". On the left, there's a circular icon with a cartoon character and some icons. In the center, a modal window titled "Create New Project" is open, prompting for "Enter Project Details" (Project Name and Description) and "Selected Project Type". A red box highlights the "Image Classifier" option, which features three small images of animals (cat, dog, bear). Other project types shown include "Object Detection", "Pose Classifier", "Hand Pose Classifier", "Audio Classifier", "Numbers(C/R)", and "Text Classifier". At the bottom right of the modal is a "Create Project" button.

Machine Learning
PictoBlox

Classificazione dei Segnali
stradali

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

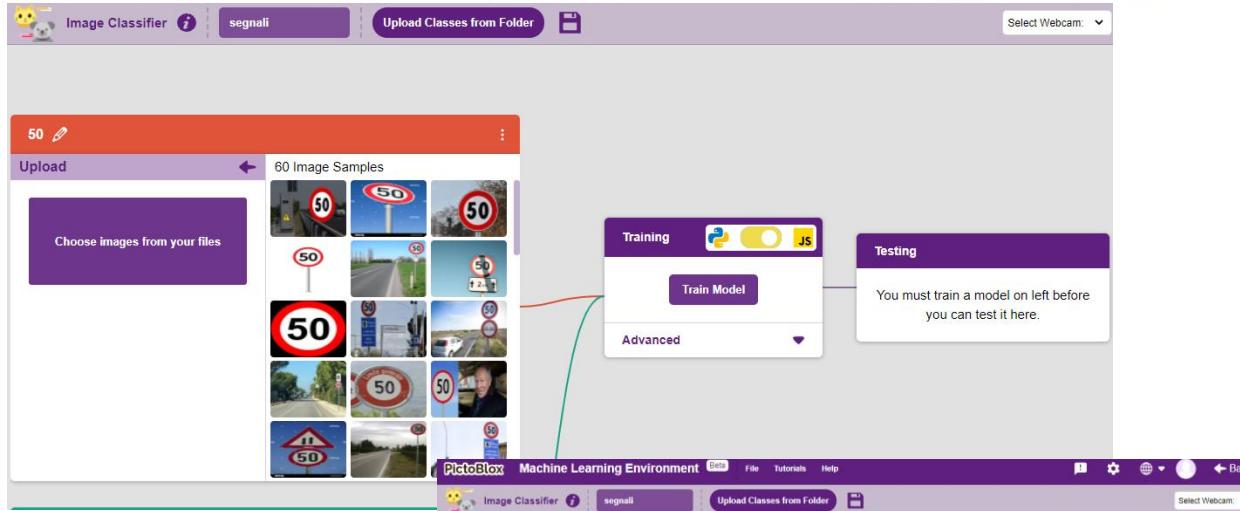


Machine Learning
PictoBlox

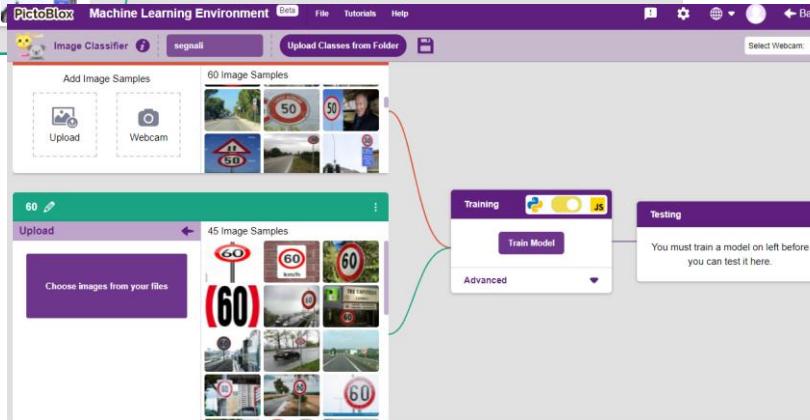
Classificazione dei Segnali
stradali >> definizione delle
classi di segnali immagine

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini



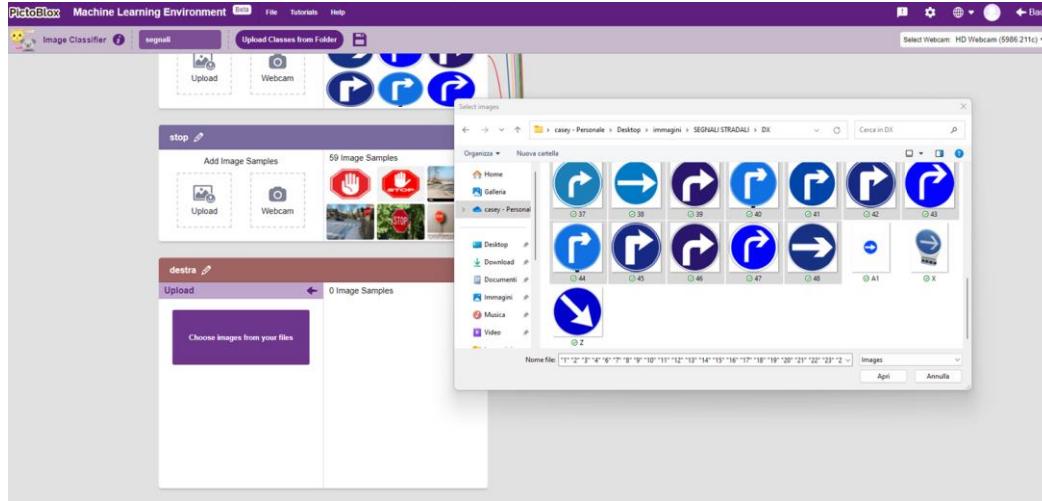
Machine Learning
PictoBlox



Classificazione dei Segnali
stradali >> definizione delle
classi di segnali immagine

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini



Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> 3 immagini non sono
utilizzate per il training e
saranno utilizzate per il testing

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

The screenshot shows the PictoBlox Machine Learning Environment interface. It features two parallel training sessions for "Image Classifier" models.

Top Left Session: This session is for "segnali". It has "49 Image Samples" and displays a graph titled "Training" showing Accuracy Vs Epochs. The accuracy starts at ~0.15 and rises to ~0.55 over 10 epochs. A "Testing" panel indicates that training must be completed before testing. Buttons for "Upload" and "Webcam" are available.

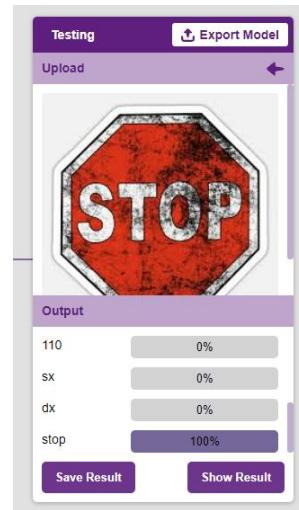
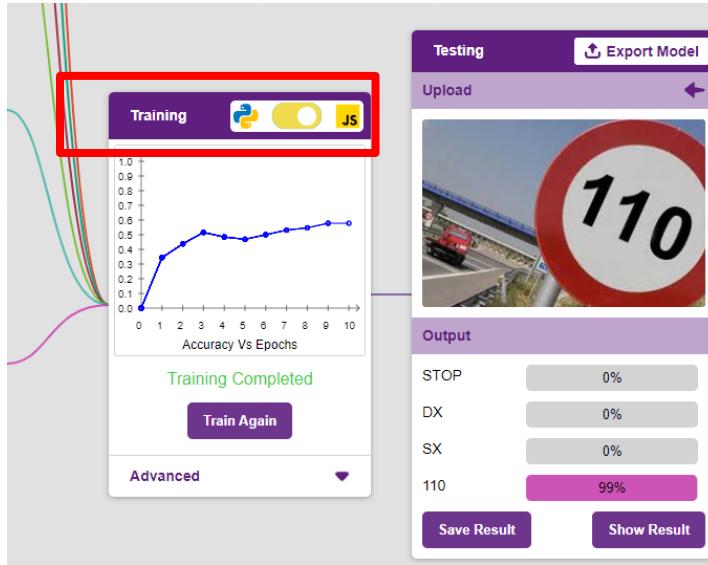
Bottom Right Session: This session is for "110". It has "50 Image Samples" and displays a similar "Training" graph. The accuracy starts at ~0.1 and rises to ~0.6 over 10 epochs. A "Testing" panel shows "Training Completed" and includes an "Export Model" button. Buttons for "Upload" and "Webcam" are also present.

Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> addestramento

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

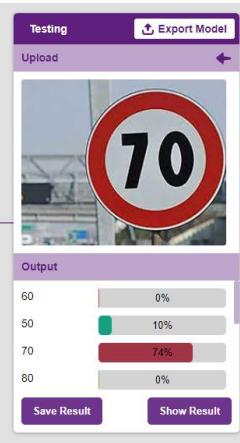
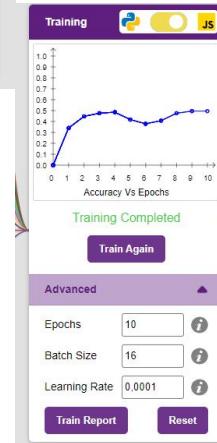


Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> testing

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

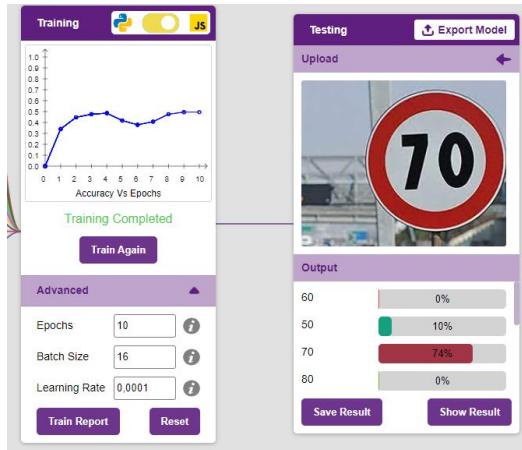


Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> testing

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini



Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> modifichiamo i
parametri di addestramento e
aumentiamo le EPOCHE

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

Ci sono tre iperparametri con cui puoi giocare qui:

- 1 Epoch** : il numero totale di volte in cui i dati verranno inseriti nel modello di training. Pertanto, in 10 epoch, il set di dati verrà alimentato attraverso il modello di training **10 volte** . **L'aumento del numero di epoch può spesso portare a prestazioni migliori.**
- 2 Dimensione batch** : la dimensione del set di campioni che verrà utilizzato in un unico passaggio. Ad esempio, se nel set di dati sono presenti 160 campioni di dati e la dimensione del batch è pari a 16, ogni epoca verrà completata in **160/16=10 passaggi** . **Raramente avrai bisogno di modificare questo iperparametro.**
- 3 Tasso di apprendimento** : determina la velocità con cui il modello aggiorna i pesi dopo aver ripetuto un passaggio. **Anche piccoli cambiamenti in questo parametro possono avere un enorme impatto sulle prestazioni del modello** . L'intervallo abituale è compreso tra **0,001 e 0,0001**.

Machine Learning
PictoBlox

IPERPARAMETRI

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

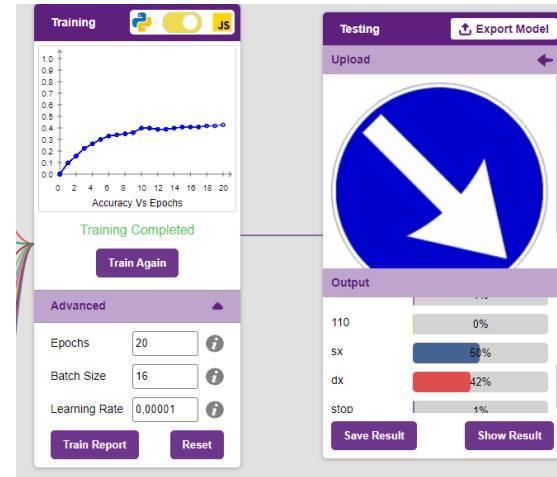
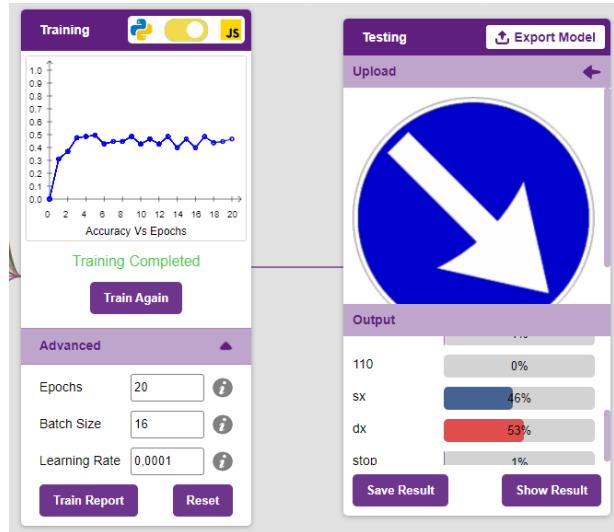


Machine Learning
PictoBlox

IPERPARAMETRI: da 10 a 20
EPOCHE di addestramento,
migliora l'ACCURATEZZA

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

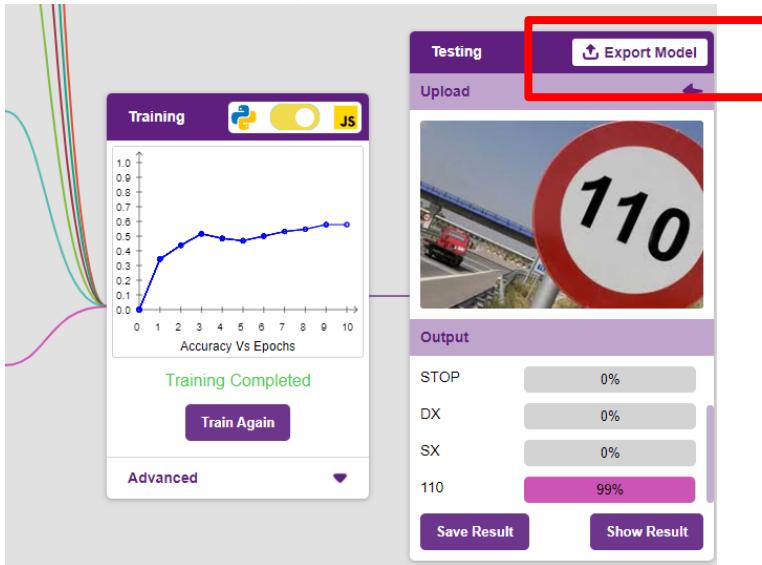


Machine Learning
PictoBlox

IPERPARAMETRI: da 10 a 20
EPOCHE di addestramento e
LEARNING RATE da 0.0001 a
0.00001, peggiora
l'ACCURATEZZA

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

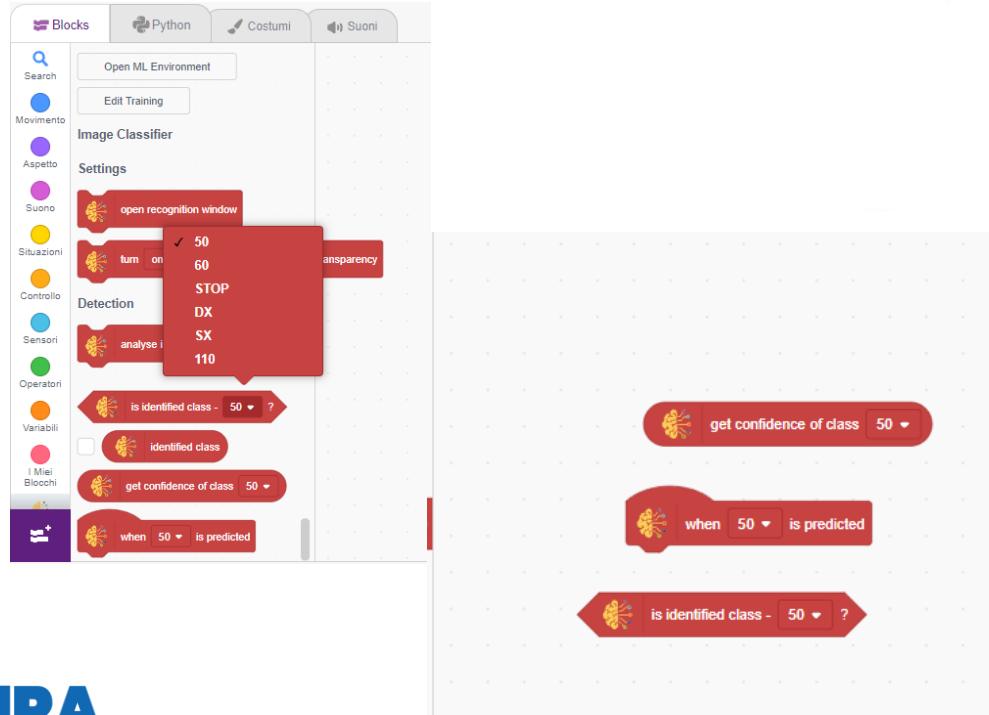


Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> esporta il MODELLO

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

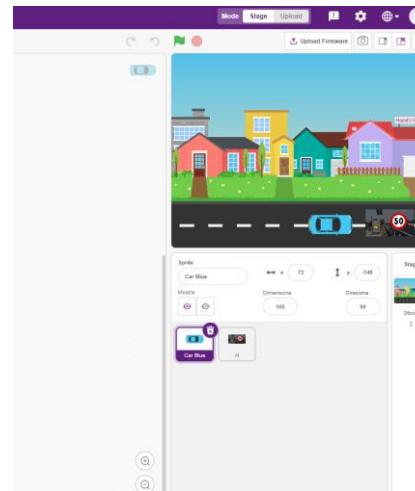
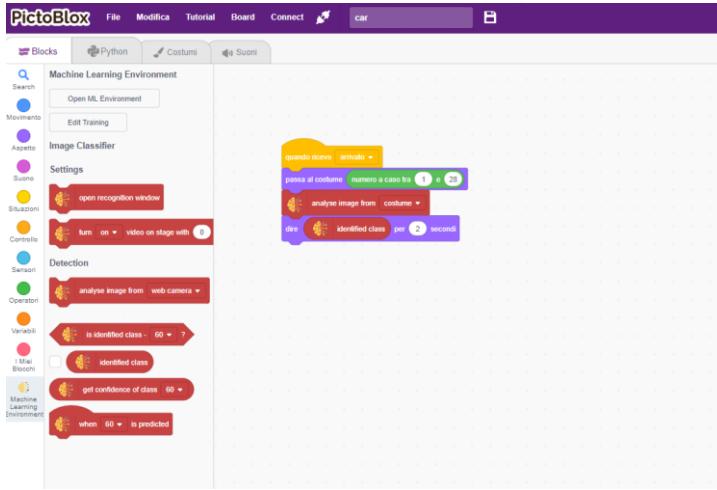


Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> esporta il MODELLO

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini

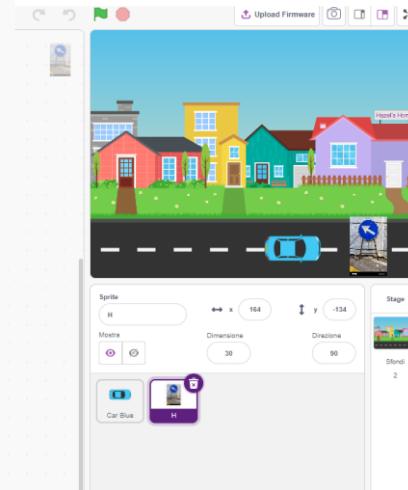
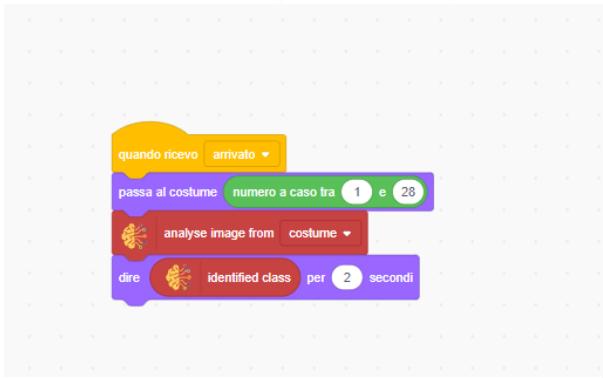


Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> classifico con la APP
attraverso una semplice
ANIMAZIONE

Modulo 5: “Machine Learning con PictoBlox”

Classificazione di Immagini



Machine Learning
PictoBlox

Classificazione dei Segnali
stradali >> classifico con la APP
attraverso una semplice
ANIMAZIONE