

Lab. Applicazioni di Intelligenza Artificiale



Ing. Marco Cantone - PhD

Università degli Studi di Cassino e del Lazio Meridionale

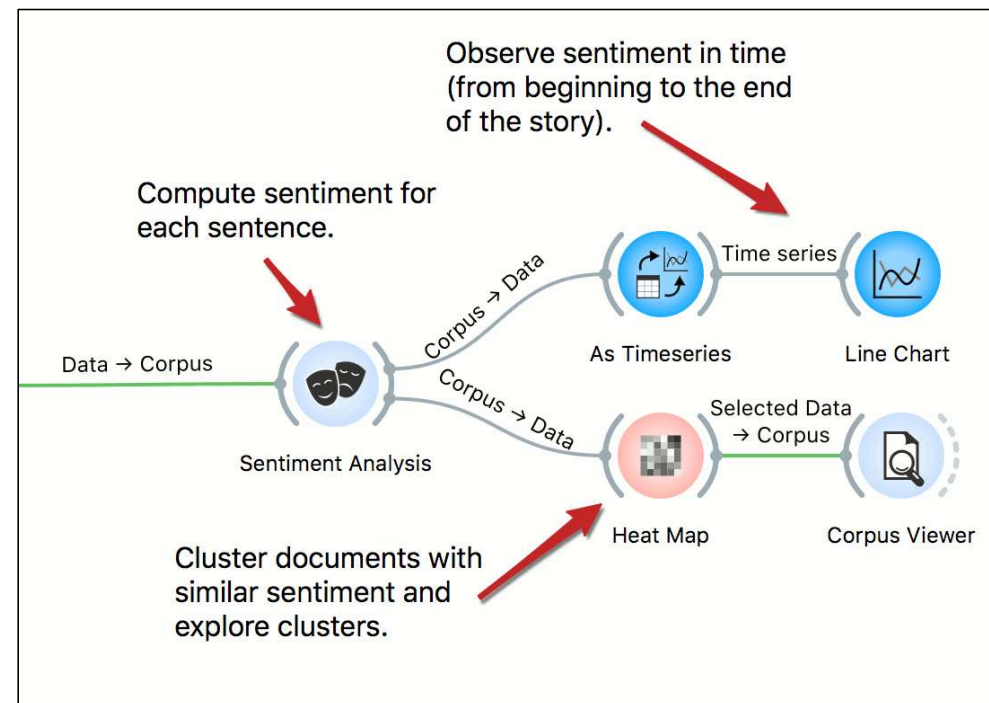


24 febbraio 2026

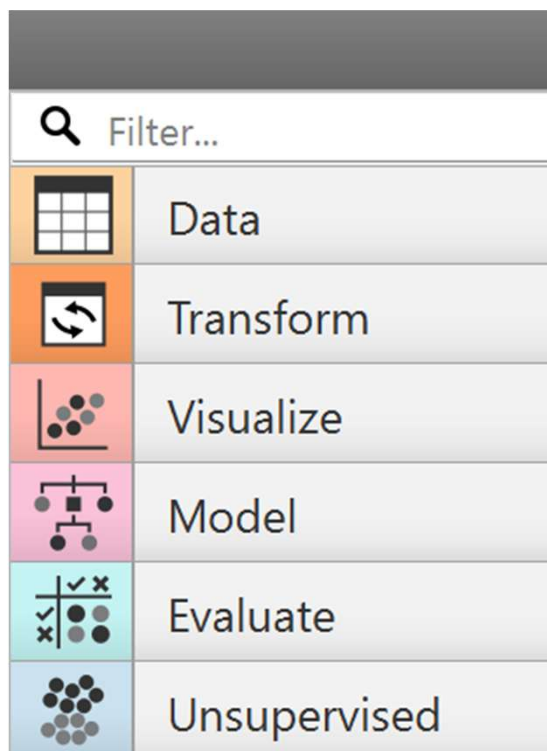
01 — Orange Data Mining

Orange Data Mining

- **Orange Data Mining** è un **software gratuito e visuale** per l'analisi dei dati e il **machine learning**, pensato anche per utenti non esperti.
- L'analisi avviene costruendo **workflow**, cioè sequenze di operazioni, collegando tra loro dei **widget**.
 - **Widget**: blocchi grafici dedicati per eseguire singole funzioni (caricare dati, visualizzarli, applicare modelli, valutare i risultati).
 - **Workflow**: sequenza ordinata di operazioni.



Widget



Data – widget per **caricare, visualizzare e preparare** i dati da analizzare.

Transform – widget per **modificare o filtrare** i dati (es. selezionare colonne, campionare, imputare).

Visualize – widget per **creare grafici e visualizzazioni** che aiutano a capire i dati.

Model – widget per **creare modelli predittivi** tramite algoritmi di machine learning supervisionato.

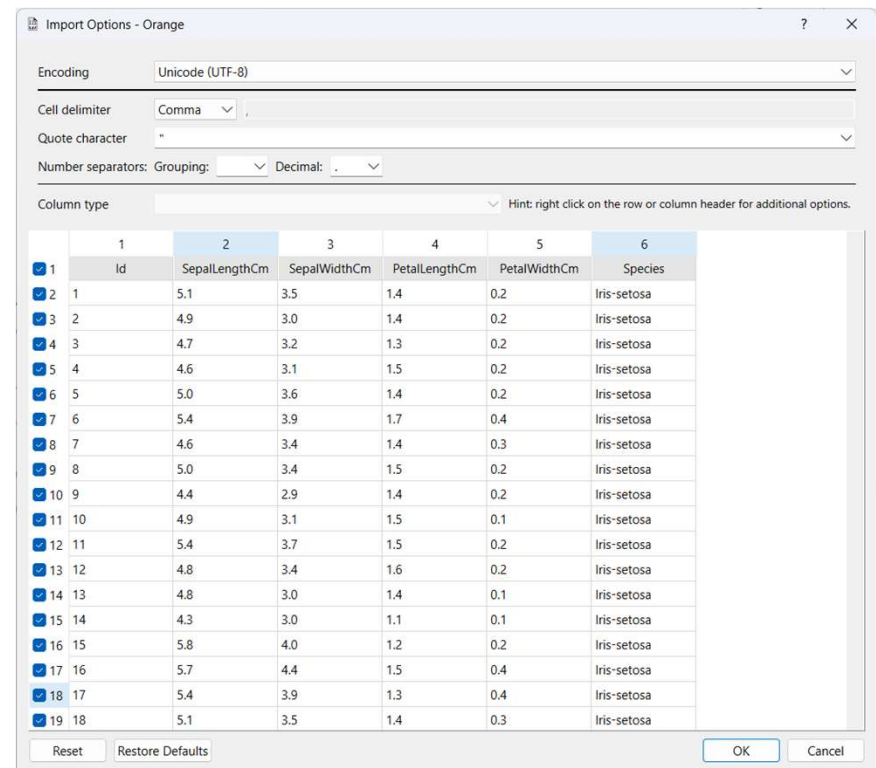
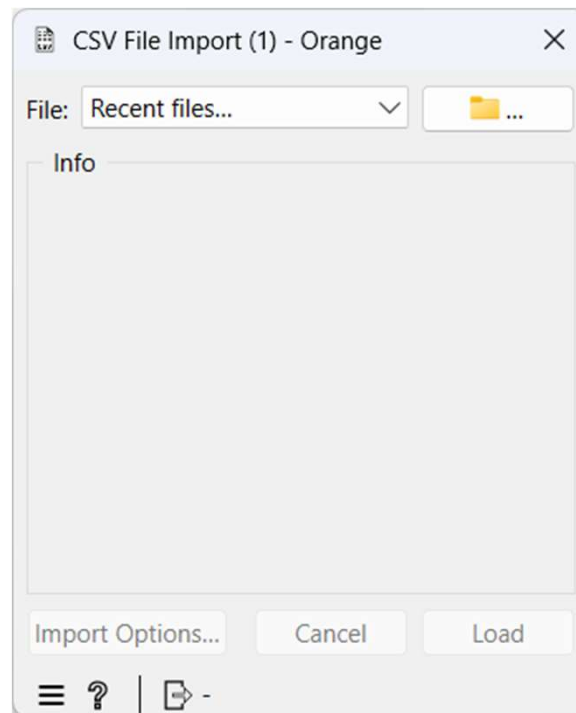
Evaluate – widget per **valutare e confrontare** le prestazioni dei modelli costruiti.

Unsupervised – widget per **scoprire pattern e gruppi nei dati** senza una variabile target.

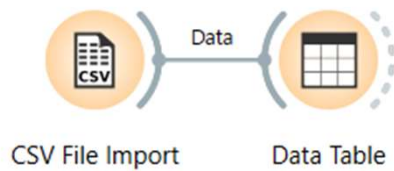
Caricare i dati



CSV File Import



Visualizzare il dataset



- Il widget «table data» permette la visualizzazione in formato tabellare dei dati
- Il dataset è formato da **campioni**, ognuno dei quali possiede diverse caratteristiche (**features**)

Data Table - Orange

Info
150 instances (no missing data)
6 features
No target variable.
No meta attributes.

Variables
☒ Show variable labels (if present)
☐ Visualize numeric values
☒ Color by instance classes

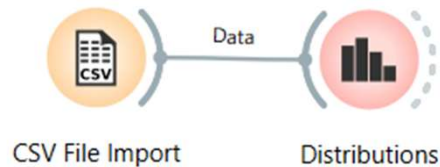
Selection
Clear Selection
☒ Select full rows

Restore Original Order
☒ Send Automatically

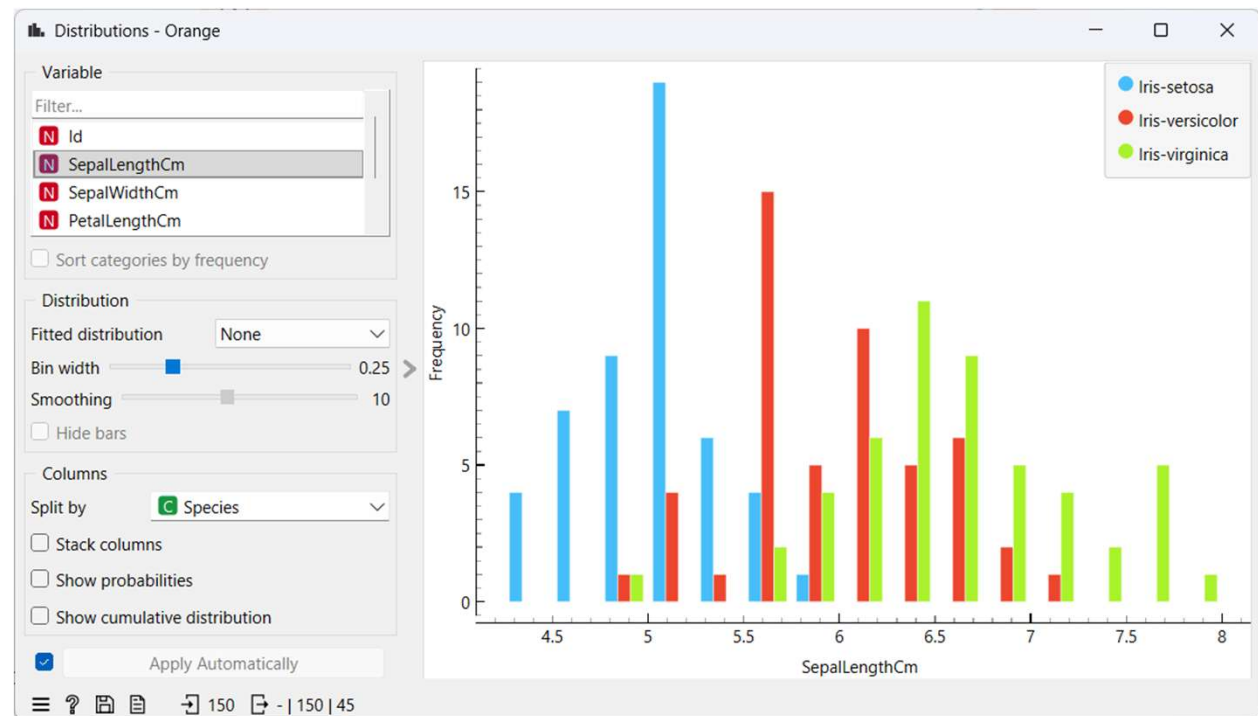
	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	
1	1	5.1	3.5	1.4	0.2	Ir
2	2	4.9	3	1.4	0.2	Ir
3	3	4.7	3.2	1.3	0.2	Ir
4	4	4.6	3.1	1.5	0.2	Ir
5	5	5	3.6	1.4	0.2	Ir
6	6	5.4	3.9	1.7	0.4	Ir
7	7	4.6	3.4	1.4	0.3	Ir
8	8	5	3.4	1.5	0.2	Ir
9	9	4.4	2.9	1.4	0.2	Ir
10	10	4.9	3.1	1.5	0.1	Ir
11	11	5.4	3.7	1.5	0.2	Ir
12	12	4.8	3.4	1.6	0.2	Ir
13	13	4.8	3	1.4	0.1	Ir
14	14	4.3	3	1.1	0.1	Ir
15	15	5.8	4	1.2	0.2	Ir
16	16	5.7	4.4	1.5	0.4	Ir

150 | 150 | 150

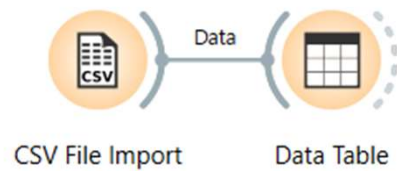
Visualizzare il dataset



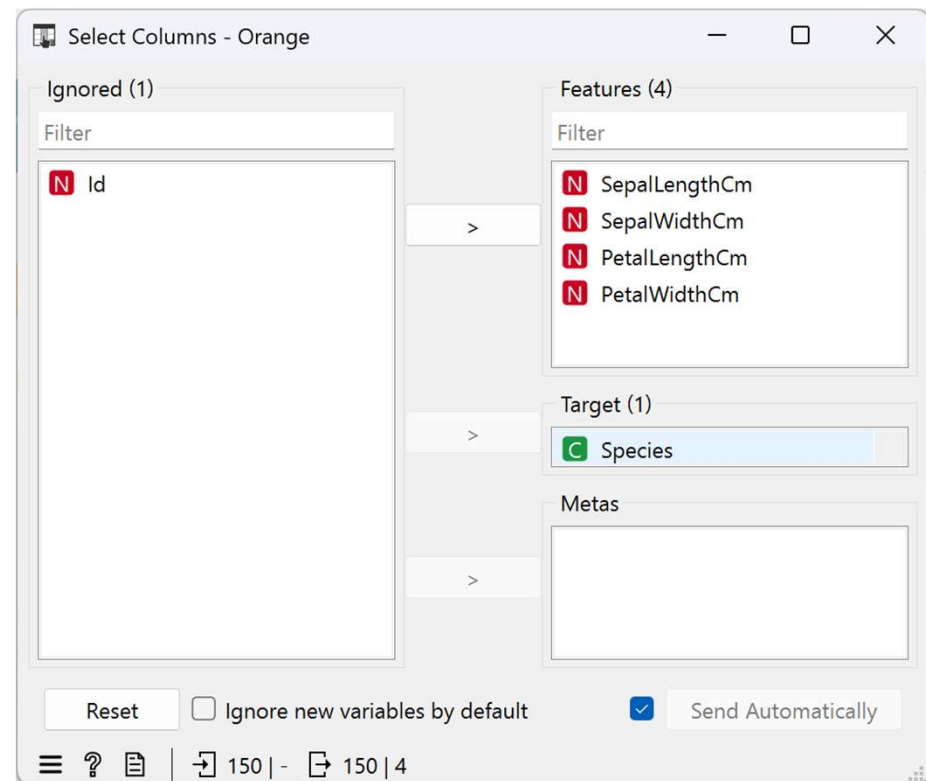
- Il widget «distributions» può essere utilizzato per visualizzare la distribuzione di una feature
- Esiste la possibilità di evidenziare la classe dei campioni



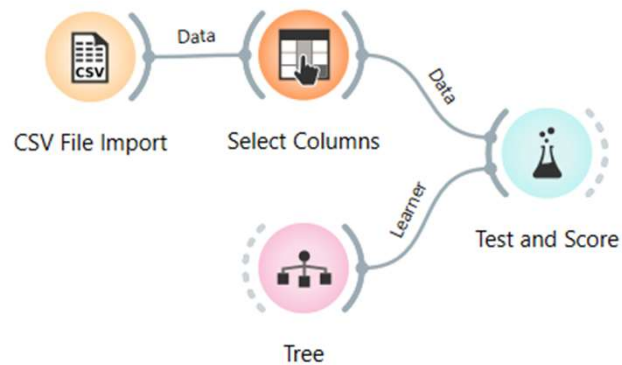
Preparare i dati per l'addestramento



- «select columns» permette di specificare qual è l'etichetta (da utilizzare per l'apprendimento supervisionato) ed escludere alcune colonne



Addestrare un modello



- Il widget «Test and Score» viene utilizzato per addestrare un modello tramite la k-fold cross-validation. Richiede in ingresso il modello da addestrare e i dati

Test and Score - Orange

☒ Cross validation

Number of folds: 5

☒ Stratified

☐ Cross validation by feature

☐ Random sampling

Repeat train/test: 10

Training set size: 66 %

☒ Stratified

☐ Leave one out

☐ Test on train data

☐ Test on test data

Evaluation results for target (None, show average over classes)

Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.974	0.947	0.947	0.947	0.947	0.920

Compare models by: Area under curve

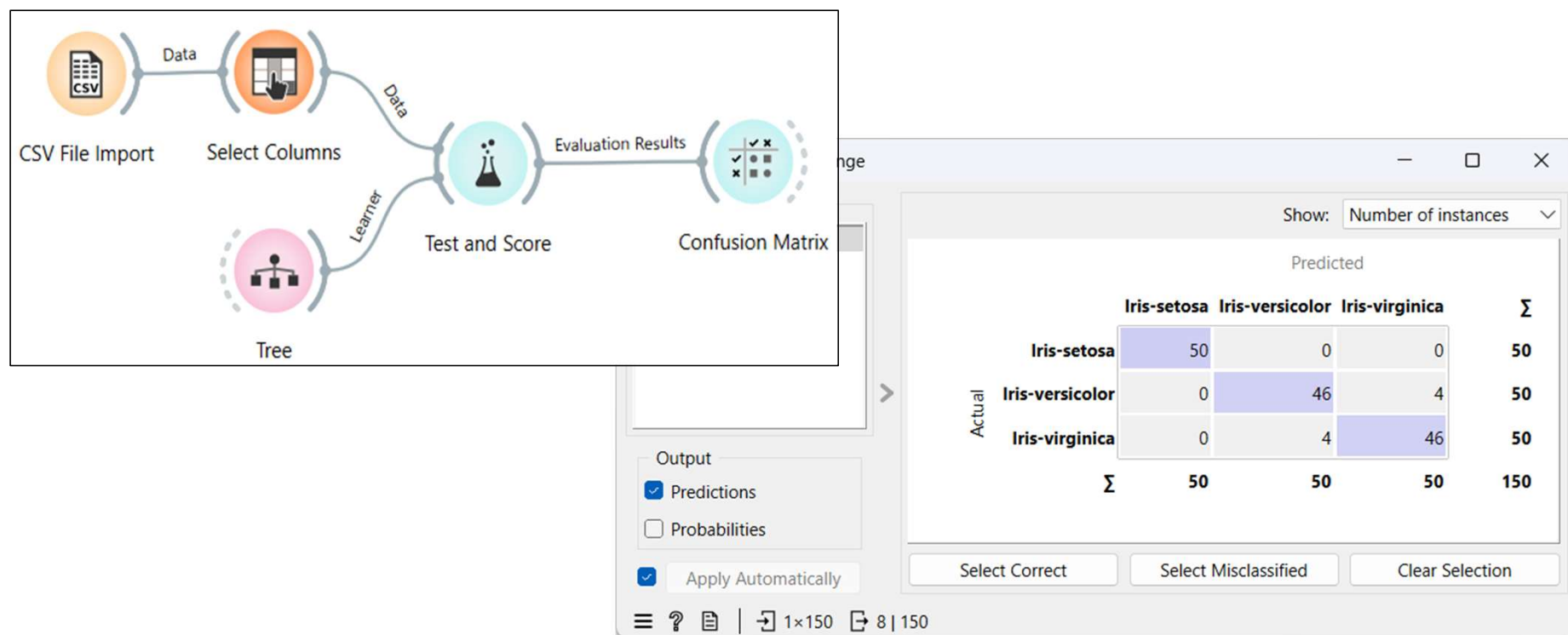
☐ Negligible diff.: 0.1

	Tree
Tree	

Table shows probabilities that the score for the model in the row is higher than that of the model in the column. Small numbers show the probability that the difference is negligible.

150 | 150 | 1x150

Visualizzare la matrice di confusione



Esercitazione 1

- Caricare il dataset Iris
 - Quanti sono i campioni? Quante features?
- Visualizzare la distribuzione di ogni feature
 - Quali sono le features migliori per classificare ogni campione nella specie corretta?
- Addestrate un SVM ed un decision tree e confrontarne le performance
 - Quale dei 2 modelli commette più errori?
- Visualizzare il decision tree addestrato (widget Tree Viewer)

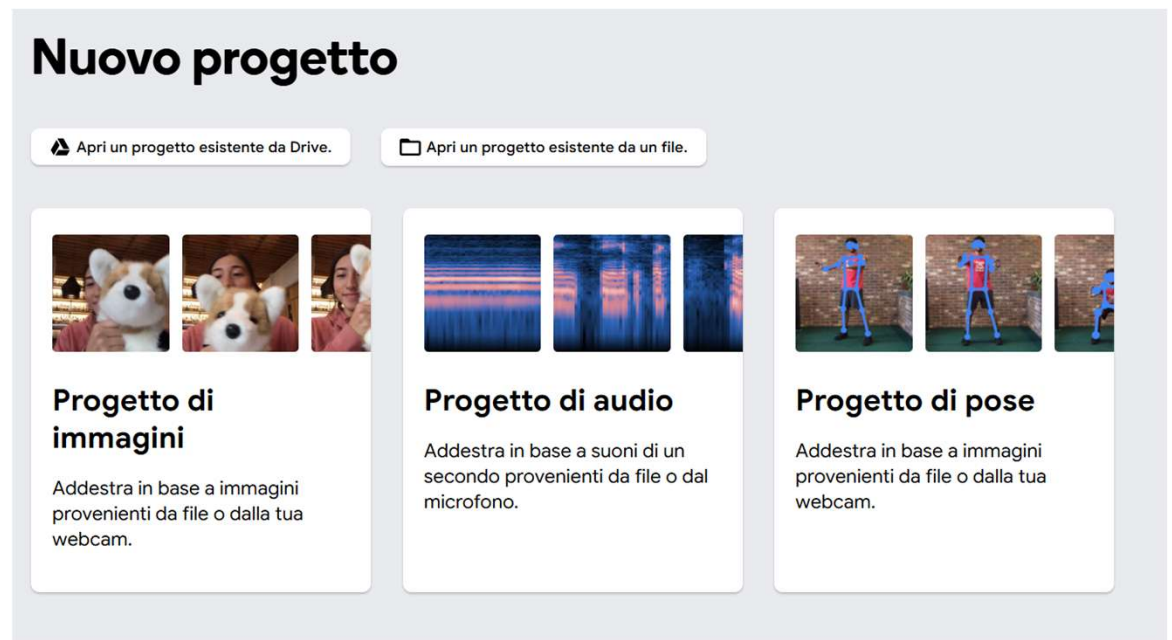
Esercitazione 2

- Caricare il dataset «Advertising» (widget *Datasets*)
- Visualizzare 2 scatterplot: TV-newspaper e TV-radio colorando ogni campione in base al valore delle vendite. Cercare di capire qual è la feature più informativa per prevedere il valore delle vendite
- Controllare le features più informative utilizzando il widget «Rank»
- Addestrare un modello di logistic regression per prevedere le vendite in base ai budget dei 3 canali pubblicitari.
 - Addestrare un altro modello che utilizza solo il valore del budget TV
 - Addestrare un altro modello che utilizza feature reduction utilizzando un PCA con 2 features in uscita
- Visualizzare i campione erroneamente classificati dal modello che viene addestrato sulle features in uscita dalla PCA

02 — Teachable Machine

Teachable Machine

- Teachable Machine è uno strumento gratuito di Google che permette di **addestrare modelli di deep learning** senza programmare.
- Consente di creare facilmente progetti di classificazione di **immagini, audio e pose**, usando esempi forniti dall'utente.



Caricare i dati e addestrare il modello

Acquisire campione usando la webcam

Acquisire campione usando la webcam

Differenti classi

Addestramento

Addestra modello

Avanzate

Periodi: 50

Dimensione del batch: 16

Tasso di apprendimento: 0,001

Ripristina impostazioni predefinite

Funzioni avanzate

Periodi: numero di volte che le immagini del dataset vengono fornite al modello per l'addestramento

Dimensione del batch: Numero di immagini fornite contemporaneamente in un'iterazione di addestramento

Tasso di apprendimento: influenza la velocità di aggiornamento dei parametri della rete

Esercitazione 3

- Creare un progetto di immagini standard con Teachable Machine (immagini a colori 224x224)
- Acquisire un dataset da webcam contenente immagini di una mano che esegue le pose della morra cinese (3 classi: sasso-carta-forbici)
- Addestrare un modello
- Testare il modello in inferenza utilizzando la webcam

03 — Chatbot

Esercitazione 4 (1/4)

- Un'azienda e-commerce riceve centinaia di email al giorno. Vuole automatizzare l'estrazione delle informazioni per integrarle nel proprio sistema gestionale. Il sistema gestionale accetta solo JSON perfettamente formattati secondo il seguente schema:

```
{  
  "customer_name": "",  
  "order_id": "",  
  "category": "",  
  "priority": "",  
  "sentiment": "",  
  "summary": ""  
}
```

category ∈ { "refund", "complaint", "information", "technical_issue" }
priority ∈ { "low", "medium", "high", "critical" }
sentiment ∈ { "positive", "neutral", "negative" }
summary = max 20 parole

- Utilizzare diverse strategie di prompting per analizzare automaticamente le mail utilizzando un LLM

Esercitazione 4 (2/4)

1. Scrivi un prompt senza esempi (zero-shot)
2. Migliora il prompt rendendolo più vincolante (inpedisci testo fuori dal JSON, specifica chiaramente i valori ammessi)
3. Progettare un prompt con esempi.
 - Gli esempi devono: rispettare perfettamente il formato di output richiesto, rappresentare il più possibile le casistiche.
4. Integrare nel prompt regole decisionali implicite, ad esempio
 - Presenza di parole come “avvocato”, “denuncia”, “azione legale” implica priority=“critical”
 - Non scrivere esplicitamente le regole nel prompt: devono emergere solo dagli esempi.
- Per ogni diversa strategia testare il modello su differenti email presenti nel dataset e valutare errori di formato o di classificazione.

Esercitazione 4 (3/4)

Dataset da utilizzare per valutare i risultati ottenuti

1. Buongiorno, sono Marco Bianchi. Ordine #45821. Non ho ancora ricevuto il pacco nonostante il pagamento sia stato effettuato 10 giorni fa.
2. Salve, Anna Rossi qui. Ordine #98765. Il prodotto è arrivato rotto, chiedo rimborso immediato.
3. Sono Luca Verdi, ordine #12345. Potreste indicarmi quando è prevista la spedizione?
4. Buonasera, sono Chiara Neri. Ordine #88321. Il dispositivo non si accende, credo sia difettoso.
5. Ordine #11902. Non funziona nulla, servizio pessimo.
6. Sono Davide Conti, ordine #11902. Vorrei sapere solo lo stato della consegna.
7. Salve, sono Laura Galli. Ordine #77654. Manca un articolo nel pacco ricevuto.
8. Ordine #44567. Se non ricevo rimborso entro 48 ore mi rivolgerò al mio avvocato.
9. Ciao, ordine #22331 arrivato velocemente. Tutto perfetto, grazie!
10. Buongiorno, Paolo Ferri. Ordine #30011. Ho ricevuto un modello sbagliato rispetto a quello acquistato.
11. Sono Elisa Fontana, ordine #44109. Il prodotto è difettoso e non ho ricevuto assistenza. Sto valutando un'azione legale.
12. Ordine #55320. È la terza volta che vi scrivo senza risposta. Procederò con denuncia se non risolvete immediatamente.

Esercitazione 4 (4/4)

Possibili esempi da utilizzare per il few-shot prompting

Altri esempi potranno essere creati per migliorare le prestazioni del sistema. Utilizzare le mail presenti nel dataset per creare nuovi esempi ma ricordarsi di non utilizzare le stesse mail per valutare le prestazioni del sistema

Email: Sono Luca Verdi, ordine #12345.
Vorrei sapere quando verrà spedito il mio ordine.

Output:

```
{
  "customer_name": "Luca Verdi",
  "order_id": "12345",
  "category": "information",
  "priority": "low",
  "sentiment": "neutral",
  "summary": "Richiesta informazioni su stato spedizione."
}
```

Email: Buongiorno, sono Anna Rossi. Ordine #98765. Il prodotto è arrivato rotto e vorrei il rimborso al più presto.

Output:

```
{
  "customer_name": "Anna Rossi",
  "order_id": "98765",
  "category": "refund",
  "priority": "high",
  "sentiment": "negative",
  "summary": "Prodotto danneggiato, richiesta rimborso."
}
```

Email: Sono Marco Bianchi, ordine #45821.
Se non ricevo il rimborso entro 24 ore mi rivolgerò al mio avvocato.

Output:

```
{
  "customer_name": "Marco Bianchi",
  "order_id": "45821",
  "category": "refund",
  "priority": "critical",
  "sentiment": "negative",
  "summary": "Minaccia azione legale per rimborso."
}
```

Esercitazione 5

- Realizzare un documento Word che presenti lo stato dell'arte approfondito sulla previsione di potenza (Power Forecasting) tramite Machine Learning e Deep Learning.
- Utilizzare la funzione «Deep Research» di Gemini:
 - Identificare articoli scientifici rilevanti
 - Analizzare contributi recenti (ultimi 5–7 anni)
 - Confrontare modelli ML e Deep Learning
 - Individuare trend emergenti e criticità aperte
- Selezionare la funzione «Deep Research» e scrivere un prompt coerente con la traccia. Confermare, ed eventualmente modificare, il piano di ricerca proposto dal modello rispettando i vincoli della traccia
- Generare un'infografica e un quiz basato sui risultati ottenuti (funzione «crea» su Gemini)
- Esportare il risultato su Google Docs e salvare in locale come .docx

Esercitazione 6

- Scrivere un prompt dettagliato per un modello di generazione di immagini che produca un'immagine il più possibile simile a quella fornita.
- Non potete utilizzare image-to-image.
- Dovete lavorare solo tramite prompt testuale.
- Si può lavorare per iterazioni successive.



04 — NotebookLM

NotebookLM

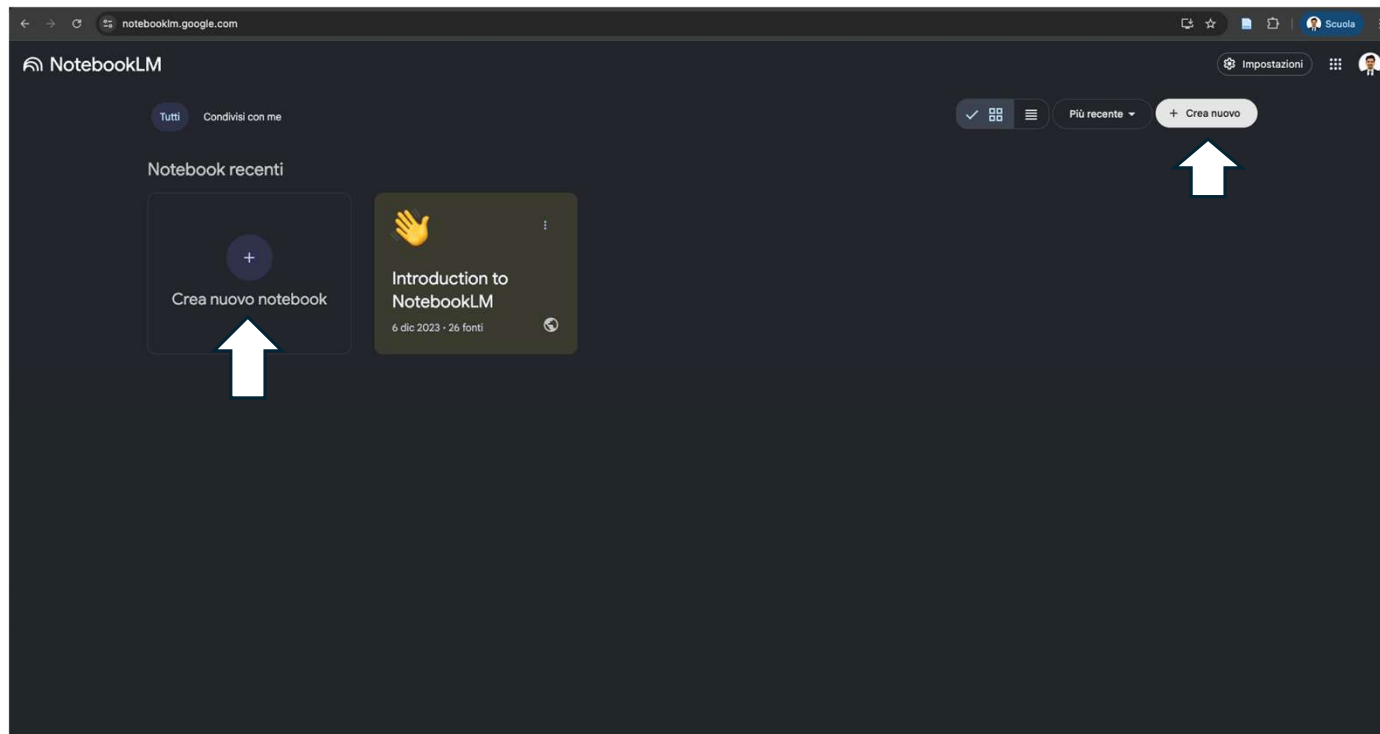
- **NotebookLM** è un'applicazione sviluppata da **Google DeepMind** per creare notebook intelligenti alimentati da modelli linguistici (LLM).
 -  Organizza i documenti e appunti.
 -  Comprende e riassume i contenuti
 -  Genera testi o spiegazioni basate sulle fonti



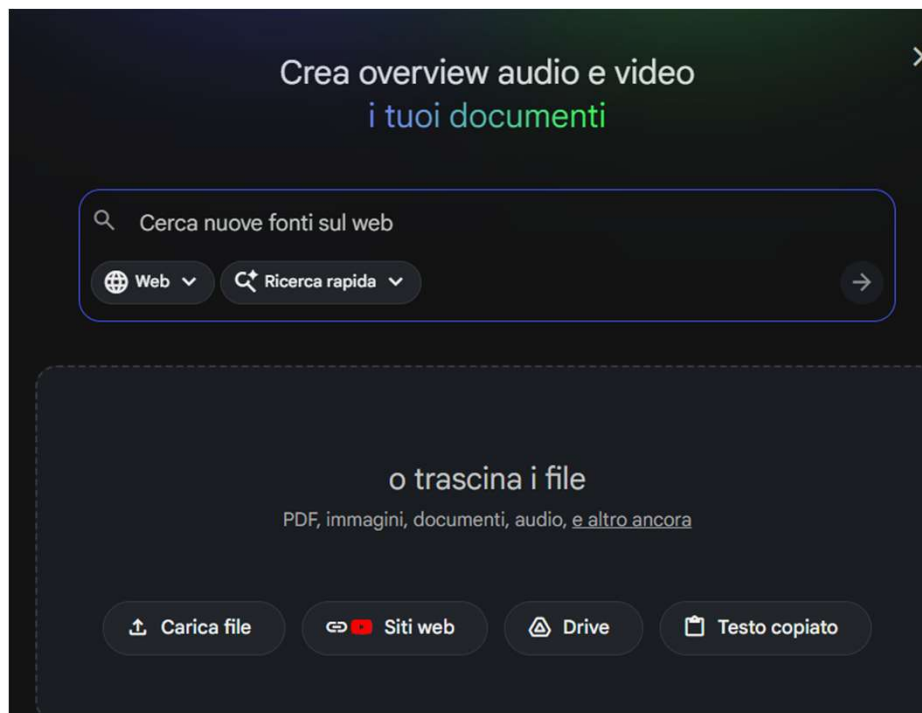
Un **assistente personale** che studia i materiali forniti e aiuta ad elaborarli in modo chiaro e coerente



L'interfaccia di NotebookLM

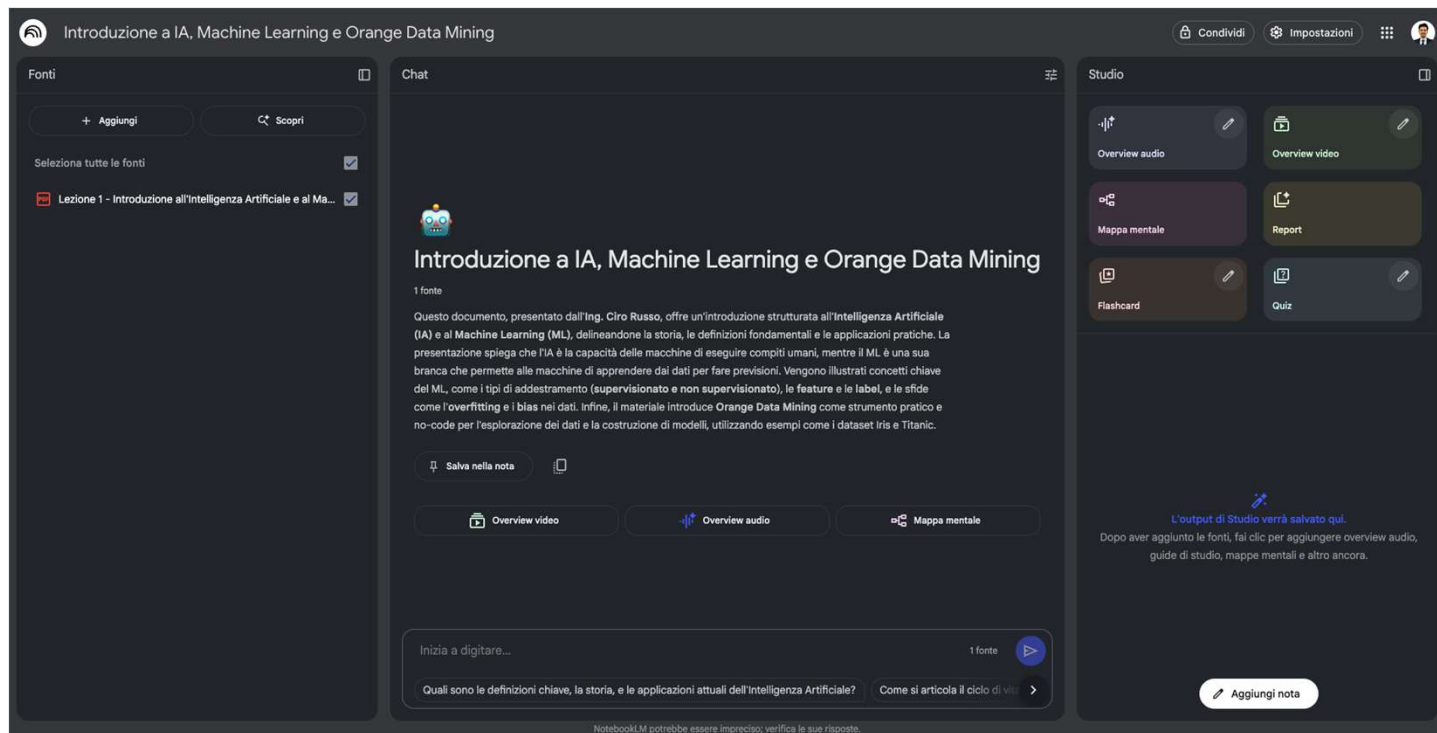


Aggiungere le fonti



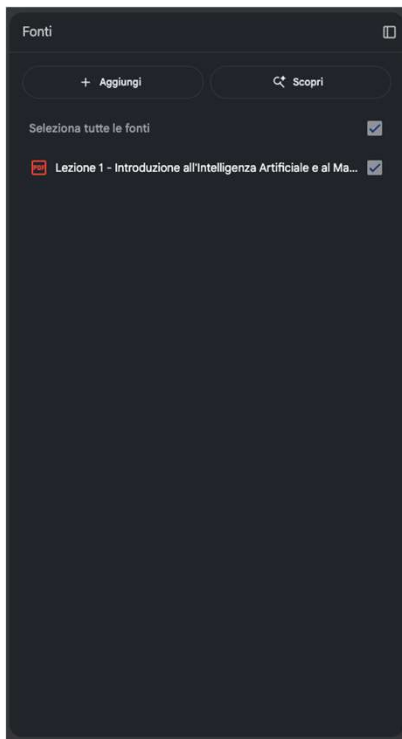
- Possono essere caricati file presenti sul proprio computer o tramite link di Google Drive.
- Posso essere caricati file di tipo testuale, immagini o audio. Sono ammessi diversi formati, tra cui: pdf, txt, docx, csv, bmp, gif, png.
- Possibilità di trovare nuove fonti sul web tramite ricerca normale o deep research.

Il Notebook





Il Notebook - Fonti



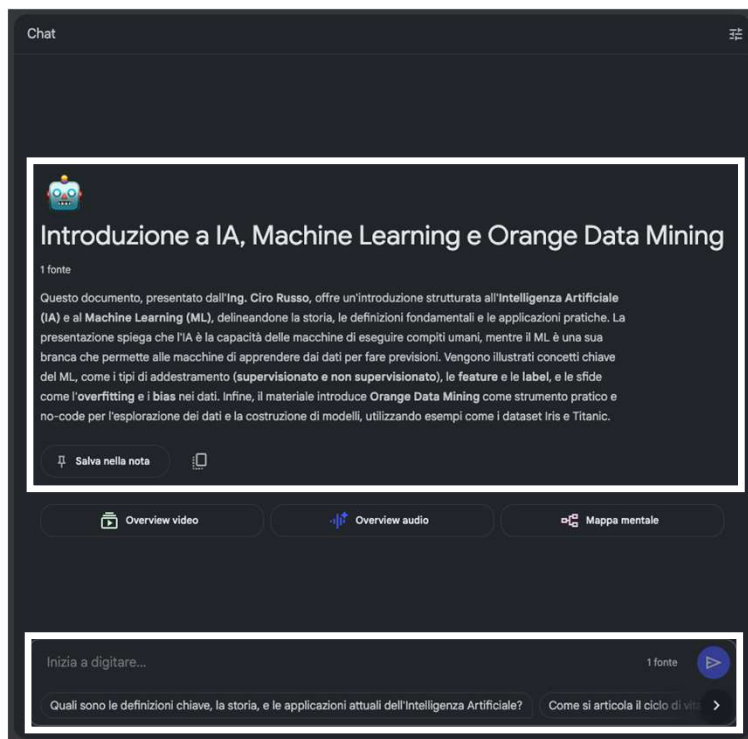
- **Fonti**

- Contiene i documenti caricati
 - si possono aggiungere, selezionare o rimuovere le fonti del notebook
 - ogni fonte può essere un PDF, un testo, un link o un file audio



È la "**memoria**" dell'assistente

Il Notebook - Chat



- **Chat**
 - È il **cuore interattivo**: si possono fare domande, chiedere riassunti o spiegazioni.
 - le risposte si basano **solo** sulle fonti caricate
 - troviamo già un **primo riassunto automatico** del documento

Il Notebook - Studio



- **Studio** 🎧
 - Strumenti di **approfondimento automatico**
 - Overview audio 🎧
 - sintesi vocale automatica dei contenuti caricati
 - Overview video 🎬
 - breve video riassuntivo che presenta i contenuti
 - Mappa mentale 🗺️
 - mappa dei concetti principali del documento
 - Report 📁
 - riassunto testuale strutturato dei contenuti
 - Flashcard 📌
 - schede domanda-risposta per l'apprendimento
 - Quiz 📝
 - domande a risposta multipla basate
 - Presentazione
 - domande a risposta multipla basate

Esercitazione 7

- Creare un notebook su Notebook LM, caricando come unica fonte il file Word sullo stato dell'arte del power forecasting basato su ML e DL.
- Creare una presentazione adatta ad una discussione dei risultati della ricerca bibliografica al proprio gruppo di ricerca.
- Sperimentare utilizzando diversi prompt e la funzione “rivedi” per effettuare modifiche puntuali sulle singole slide

Esercitazione 8

- Utilizzare NotebookLM per generare automaticamente la bozza della delibera di avvio del progetto "DigitalDocs", sintetizzando le informazioni contenute nei documenti dell'Ente.
- Materiale a disposizione (da caricare):
 - <https://drive.google.com/drive/folders/1B8FqEjcgWJut-MZvwnemb5nit7UNClo?usp=sharing>
 1. *Fonte-1-Regolamento.pdf* (direttive sull'innovazione digitale)
 2. *Fonte-2-Verbale.pdf* (decisioni emerse nella riunione direzionale)
 3. *Fonte-3-NotaTecnica.pdf* (dettagli del progetto e richieste formali)
- Scrivere un prompt che rispetti questi 3 vincoli:
 1. **Esclusività:** Il testo generato deve basarsi *solo* sulle fonti caricate, senza inventare nulla.
 2. **Struttura:** Il documento deve essere diviso esattamente in tre sezioni:
 1. **Premesso che:** sintesi dei riferimenti normativi
 2. **Considerato che:** sintesi di quanto emerso dal verbale e dalla nota tecnica
 3. **Delibera:** le decisioni finali
 3. **Formattazione:** I punti della sezione "Delibera" devono essere un **elenco numerato** e iniziare tutti con un **verbo all'infinito**.