

Corso di Machine Learning 2025/2026

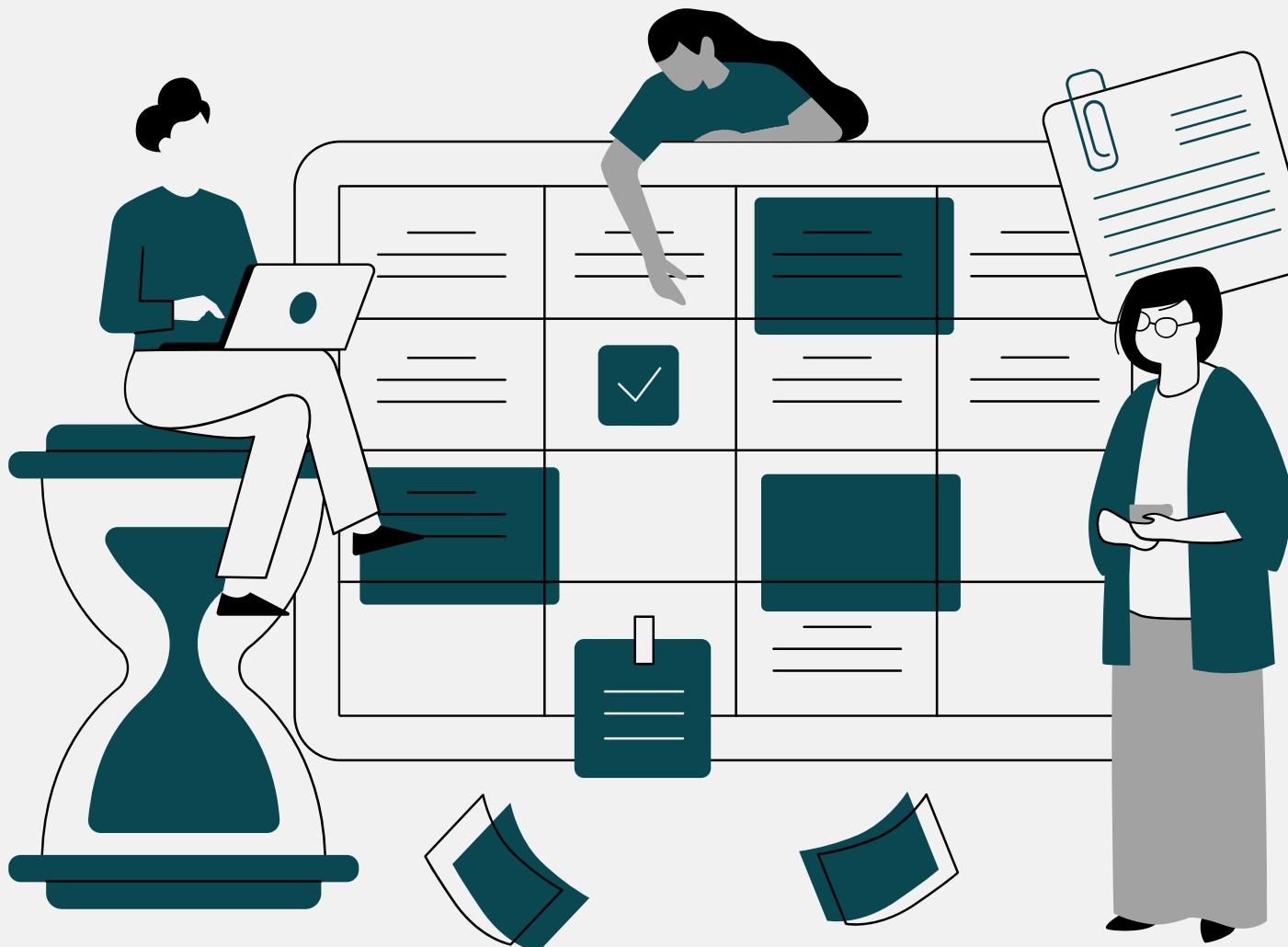
# Unisa Mortgage-AI

Autore:  
Antonio Clavino

Docenti:  
Polese Giuseppe  
Caruccio Loredana



# Il Contesto: L'AI nel settore Fintech



- **L'Evoluzione del Credito:**  
Il settore bancario sta passando da valutazioni manuali e rigide a sistemi data-driven. La gestione tradizionale delle pratiche di mutuo è spesso lenta, costosa e soggetta all'errore umano.
- **La necessità di Automazione:**  
Con l'aumento delle richieste digitali, le banche necessitano di algoritmi capaci di processare migliaia di parametri in tempo reale per stimare la probabilità di insolvenza.
- **Obiettivo del Progetto:**  
Creare un sistema che non si limiti a calcolare un punteggio, ma simuli il ragionamento di un analista finanziario esperto per approvare o respingere una richiesta.

# Il Problema e la Soluzione Proposta



## Il Problema: L'asimmetria informativa

Le banche spesso non riescono a distinguere un buon pagatore da uno cattivo solo basandosi sul reddito. Esistono variabili nascoste che pesano sulla decisione.

## La Sfida della "Black Box":

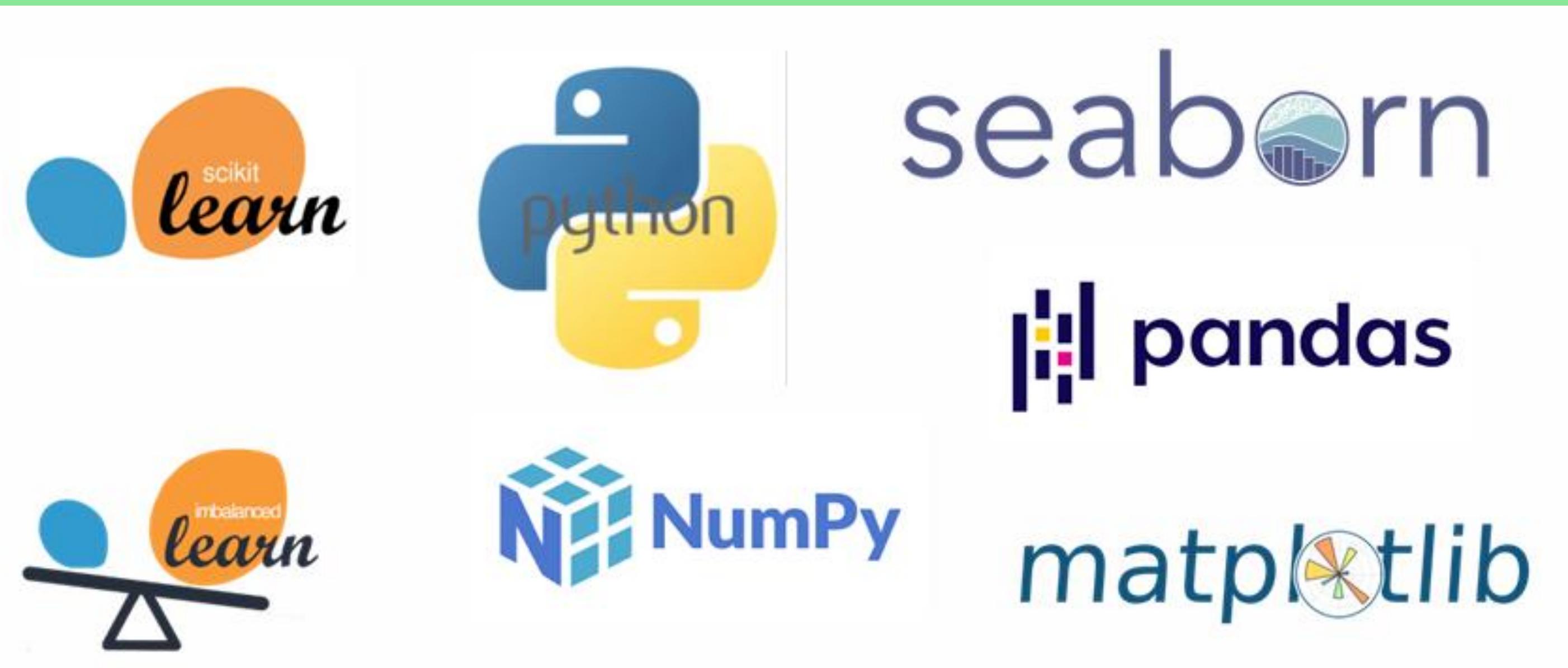
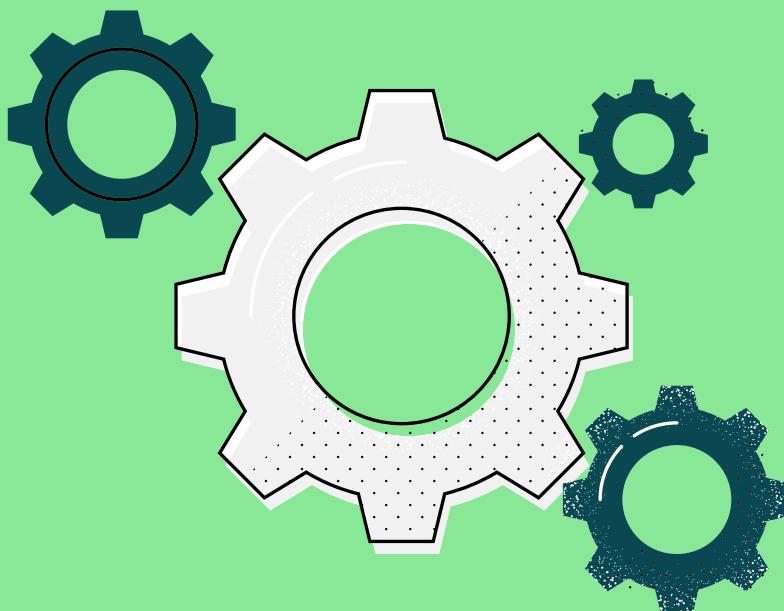
Gli algoritmi avanzati spesso sono «Black Box» incomprensibili. Tuttavia, per normative GDPR e fiducia, è fondamentale che il sistema spieghi il perché di un rifiuto

## La Soluzione Implementata:

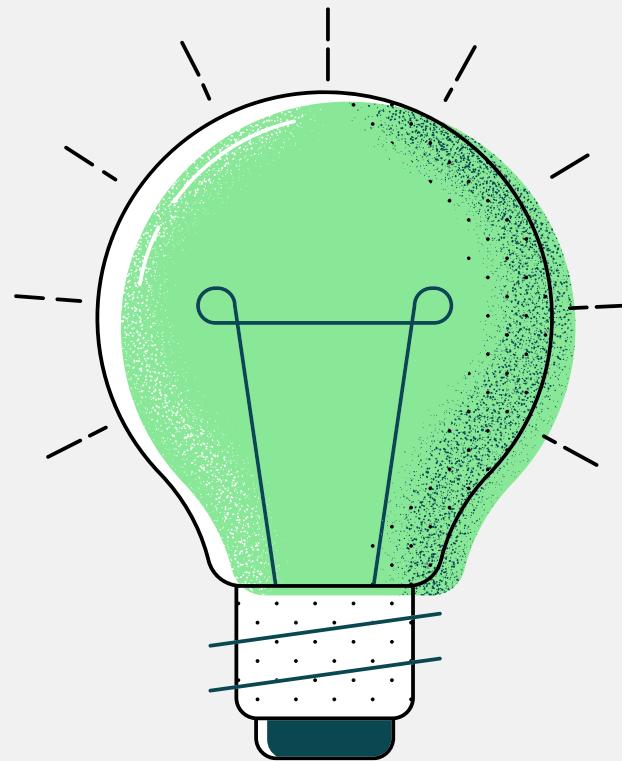
Sviluppo di una Web App interattiva che utilizza algoritmi di classificazione per:

- Prevedere l'esito: Approvazione o Rifiuto del mutuo.
- Matching Bancario: Suggerire l'istituto di credito più affine al profilo del cliente.
- Spiegare la decisione: Mostrare quali fattori hanno determinato l'esito.

# Le Tecnologie Utilizzate



# Dataset e Feature Selection



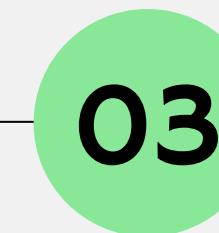
Origine dei Dati: Il nostro Dataset contenente 4.269 record reali di richieste di mutuo.

Variabili di Input: Mostra una tabella o un elenco diviso in categorie:

- Economiche: Reddito Annuo, Importo Mutuo, Rata, Patrimonio.
- Anagrafiche: Istruzione, Tipo Lavoro, Familiari a carico.
- Rischio: Score CRIF, Segnalazioni.

Target:

- Variabile Target 1: mutuo\_approvato.
- Variabile Target 2: banca\_consigliata.



**Raw Data**

**Structured  
Data**

**Data  
Preprocessing**

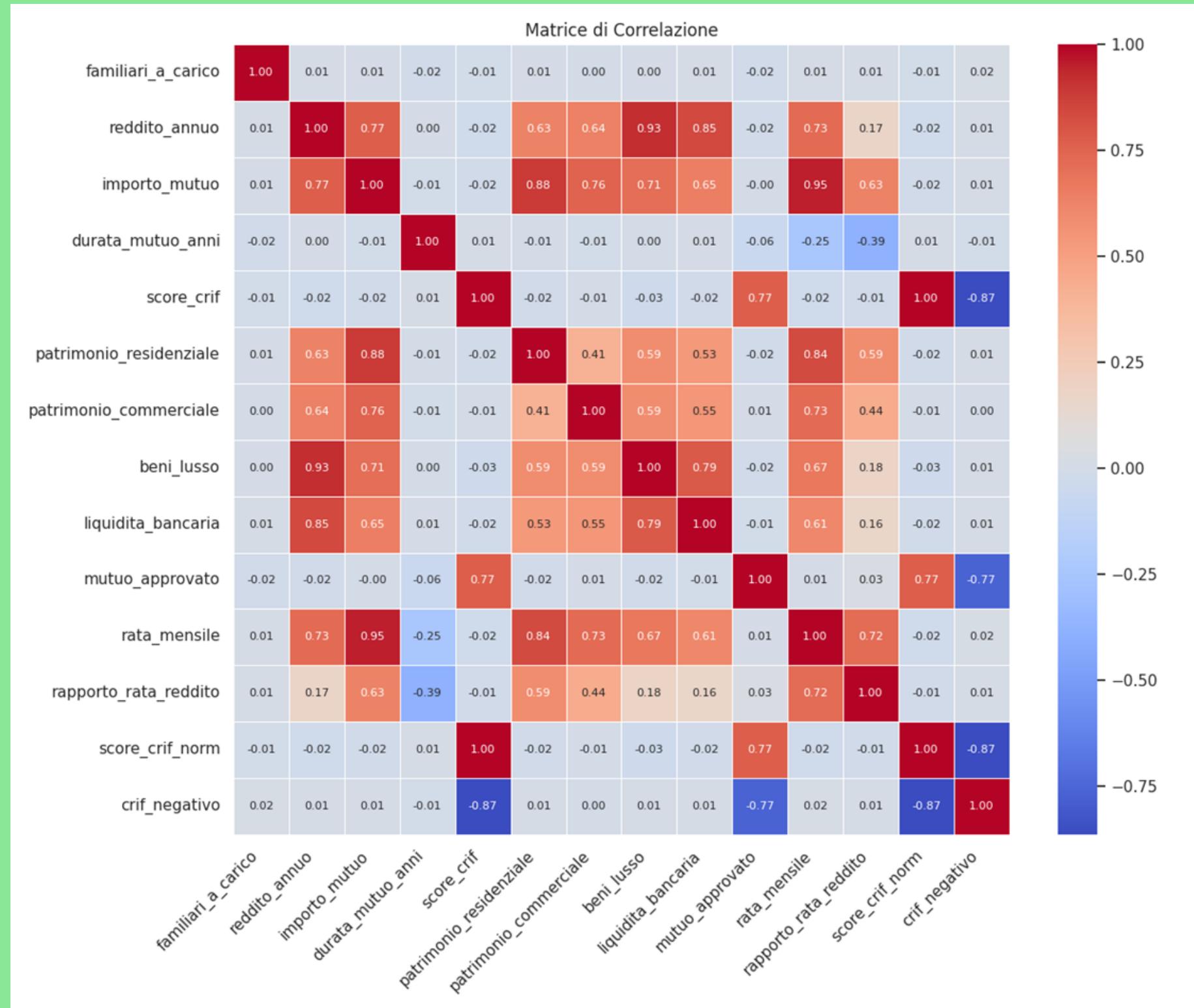
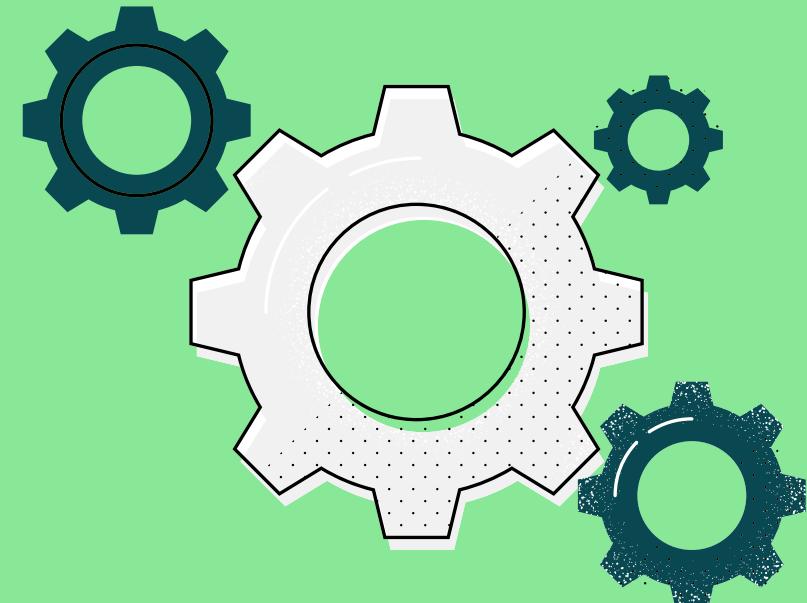
**Exploration  
Data Analysis**

**Insight,  
Reports, Visual  
Graphs**

- Pulizia: Rimozione colonne non predittive per evitare Overfitting
- Encoding: Applicazione del One-Hot Encoding per le variabili categoriche
- Feature Scaling: Utilizzo di MinMaxScaler per normalizzare i dati numerici in un range [0, 1]. Per evitare che variabili con numeri grandi dominino su variabili piccole durante l'addestramento.

# Data Preprocessing & Cleaning

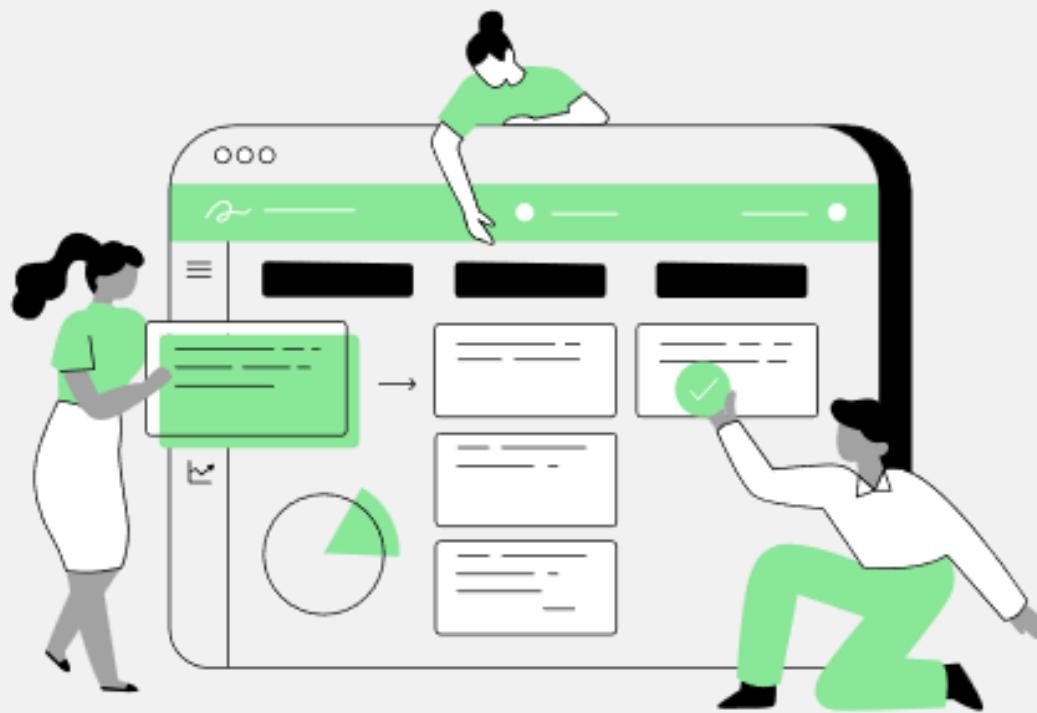
# Analisi Esplorativa



## Insight Chiave:

- Forte correlazione positiva tra Reddito e Importo Mutuo.
- Forte correlazione negativa tra Score CRIF e Rifiuto.
- Identificazione di feature ridondanti ed eliminazione del rumore.

# Scelta del modello



**Configurazione:**  
`n_estimators = 100`  
`max_depth = 7`

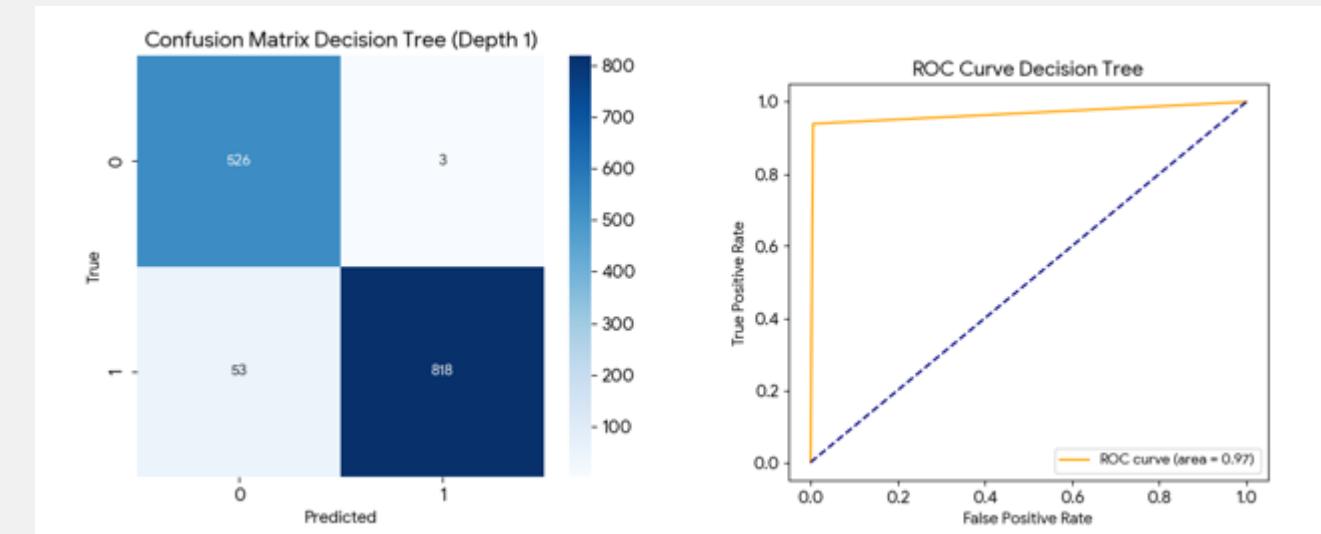
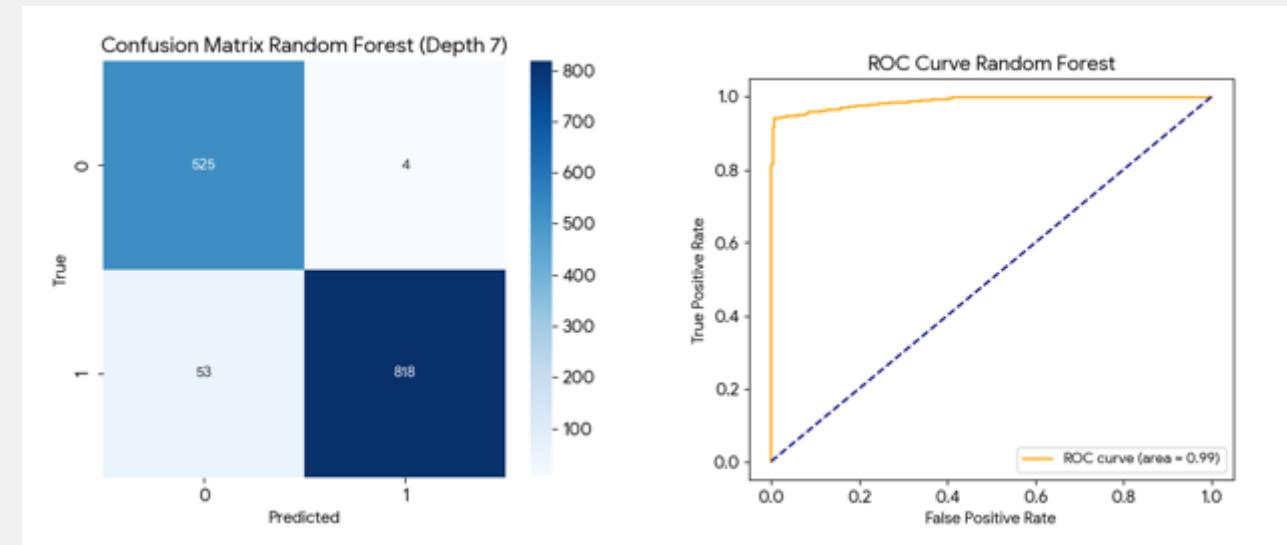
**Algoritmo Scelto:**

**Random Forest Classifier**

**Si è scelto un approccio Ensemble Learning rispetto a un singolo Decision Tree.**

**Perché Random Forest?**

- **Robustezza:** Riduce drasticamente il rischio di Overfitting mediando il risultato di 100 alberi diversi.
- **Gestione Dati Tabulari:** È lo stato dell'arte per dataset strutturati come quelli bancari.
- **Feature Importance:** Permette di estrarre facilmente quali variabili hanno pesato sulla decisione.



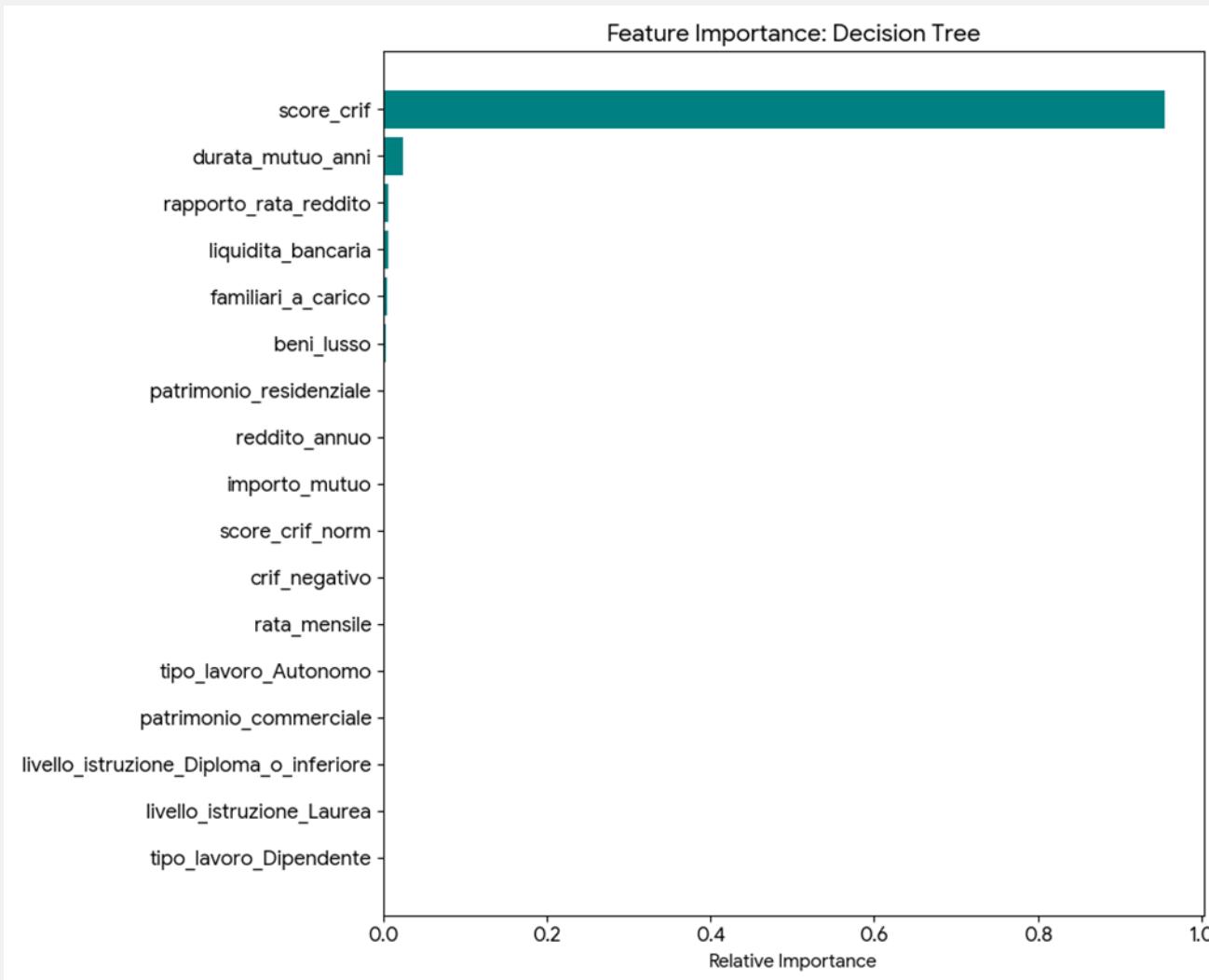
# Analisi delle Performance

Metriche di Valutazione:

- Accuracy: 96%.
- Precision & Recall: Fondamentali per minimizzare i Falsi Positivi e i Falsi Negativi

- Matrice di Confusione: Mostra come il modello sbaglia pochissimo sulle classi critiche.
- Curva ROC: Valore AUC vicino a 0.99, indicando un'eccellente capacità di separazione tra le classi "Approvato" e "Respinto".

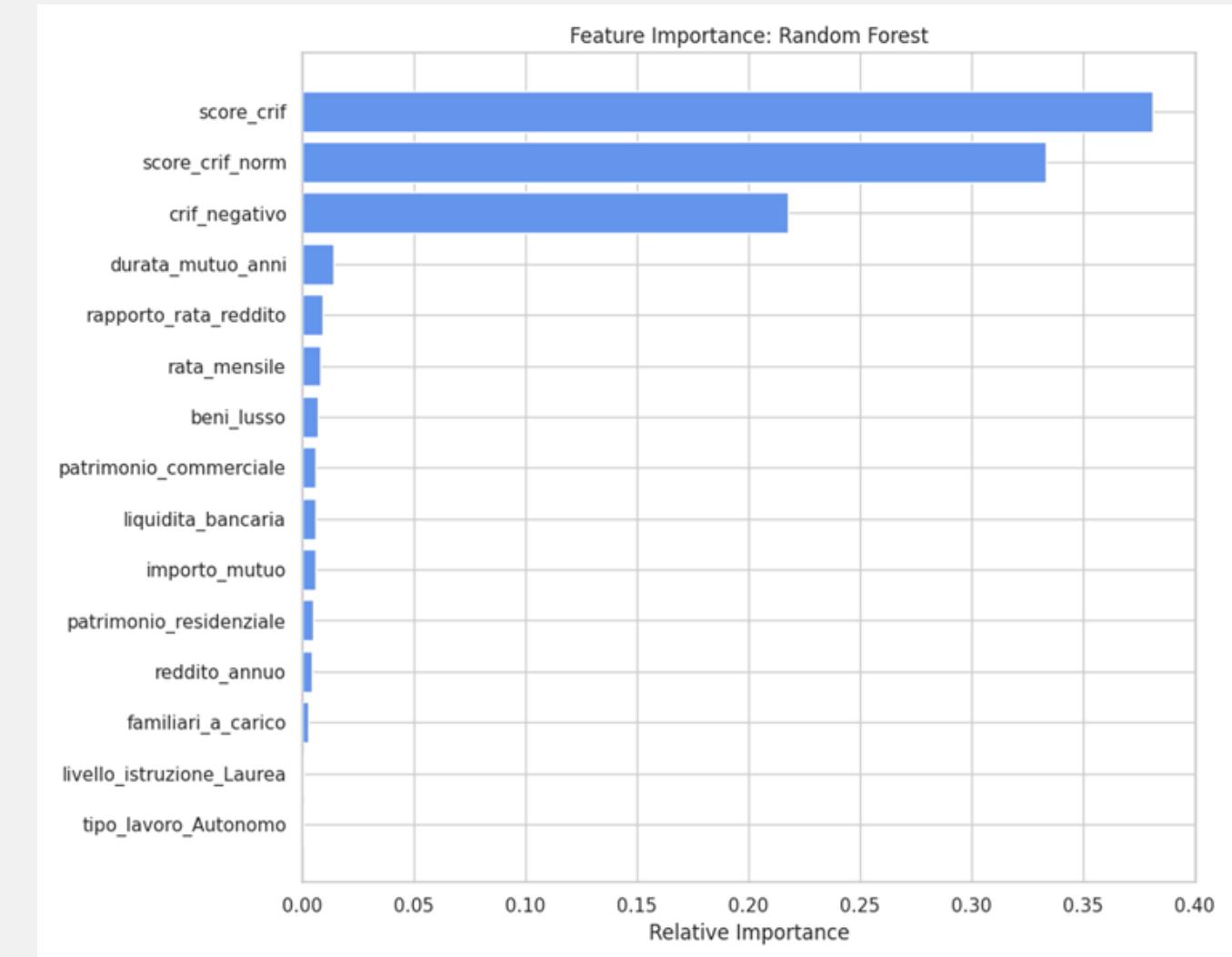
# Explainability



## Feature Importance:

I grafici mostrano che le variabili più influenti sono:

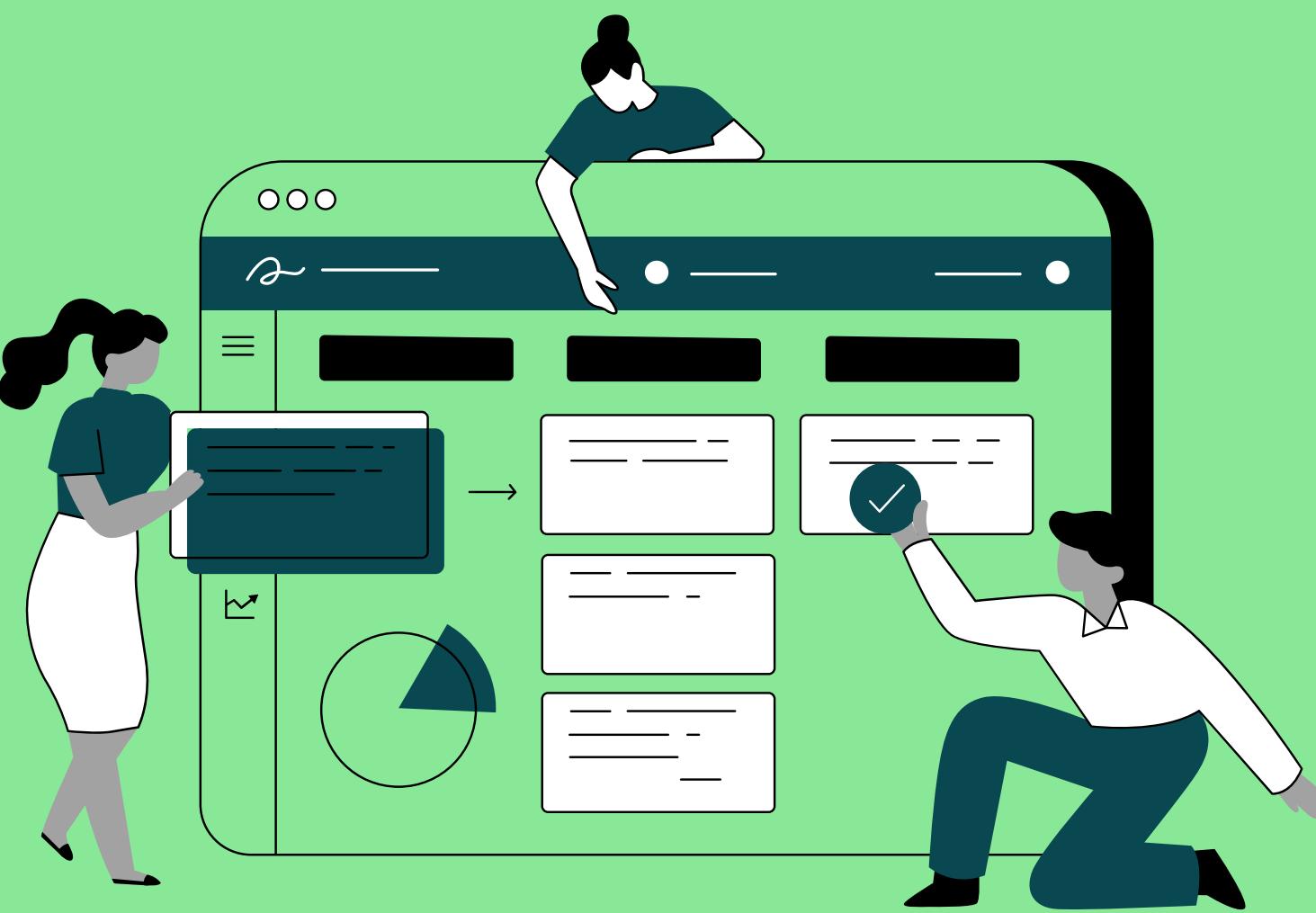
- **Score CRIF.**
- **Reddito Annuo.**
- **Rapporto Rata/Reddito.**



Interpretazione: Il modello ha imparato regole finanziarie corrette e non bias casuali.

Confronto Modelli: Abbiamo implementato un secondo modello specifico per il Matching Bancario, che apprende le politiche di credito delle diverse banche basandosi sui dati storici di erogazione.

# L'Applicazione



**Tecnologia:**  
Sviluppata in Python con framework Streamlit.

**Architettura Model-as-a-Service:**  
Il modello addestrato viene caricato in memoria e interrogato in tempo reale.

The screenshot shows a detailed view of the Credit Scoring & Matching System. On the left, there's a sidebar with sections for 'Parametri Pratica' (Annual Gross Income: 70000, Requested Amount: 600000, Amortization: 20 years), 'Profilo di Rischio' (Credit Score CRIF: 601), and 'Situazione Patrimoniale' (Comprehensive Assets: 60000). The main area is titled 'Credit Scoring & Matching System' and shows a 'Rapporto Rata/Reddito' of 45.0% and a 'Rating Cliente' of 601/900, with a 'Rata Mensile Stimata' of € 2625.00. A button labeled 'CALCOLA MERITO CREDITIZIO' is visible. Below, under 'Risultato Valutazione', it says 'ESITO: POSITIVO' with a note about compatible parameters and an 81.3% approval probability. In the 'Matching Istituti Bancari' section, it suggests 'Banca Premium' based on historical profile.

## Flusso Operativo:

- Input Utente:** Inserimento dati economici/anagrafici via interfaccia web.
- Real-time Preprocessing:** Il sistema applica lo stesso Scaling e Encoding usato nel training.
- Inferenza:** I due modelli girano in sequenza.
- Output:** Dashboard decisionale con semaforo e suggerimenti.

# Simulazione Scenari

## Scenario A

## Scenario B

### Credit Scoring & Matching System

Dati Economici

Reddito Lordo Annuo (€) 45000 - +

Importo Richiesto (€) 500000 - +

Ammortamento (Anni) 30

Rapporto Rata/Reddito 38.9%

Rating Cliente 650/900

Rata Mensile Stimata € 1458.33

**CALCOLA MERITO CREDITIZIO**

**Risultato Valutazione**

**ESITO: POSITIVO** ✓

La pratica presenta parametri compatibili con l'erogazione.

Probabilità di Approvazione: 85.5%

**Matching Istituti Bancari**

In base al profilo storico, l'istituto con maggiore affinità è: [Intesa Sanpaolo](#)

Dati Economici

Reddito Lordo Annuo (€) 50000 - +

Importo Richiesto (€) 600000 - +

Ammortamento (Anni) 30

Rapporto Rata/Reddito 42.0%

Rating Cliente 458/900

Rata Mensile Stimata € 1750.00

**CALCOLA MERITO CREDITIZIO**

**Risultato Valutazione**

**ESITO: NEGATIVO** ⚠

La pratica presenta indicatori di rischio elevati.

Probabilità di Approvazione: 45.8%

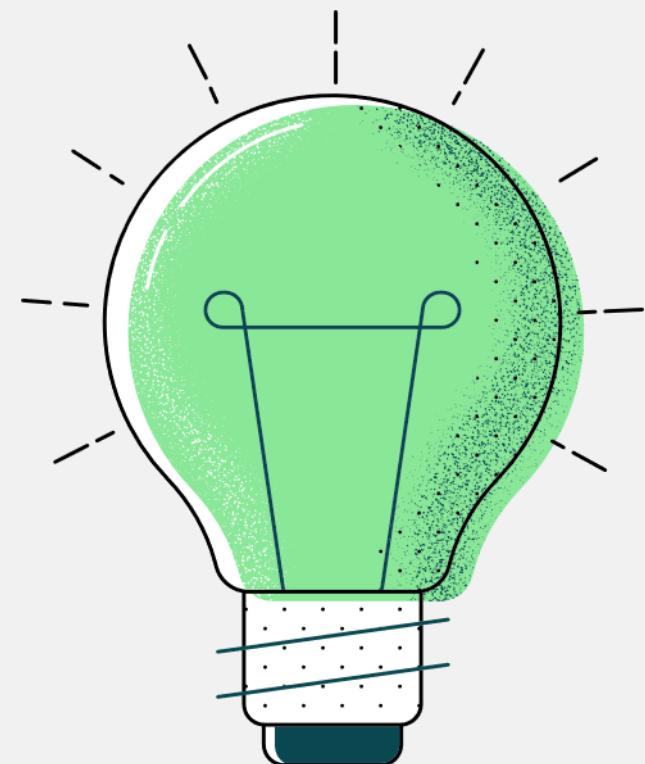
**Note di Analisi**

Il rapporto Rata/Reddito eccede i limiti prudenziali (>35%).

Lo storico creditizio (CRIF) risulta sotto la soglia di sicurezza.



# Conclusioni e Sviluppi Futuri



## Conclusioni:

**Il sistema ha raggiunto un'accuratezza del 96%, dimostrando che i dati storici contengono pattern chiari per l'automazione del credito. L'approccio Double-Model ha permesso di risolvere sia il problema del rischio che quello commerciale.**

## Limiti Attuali:

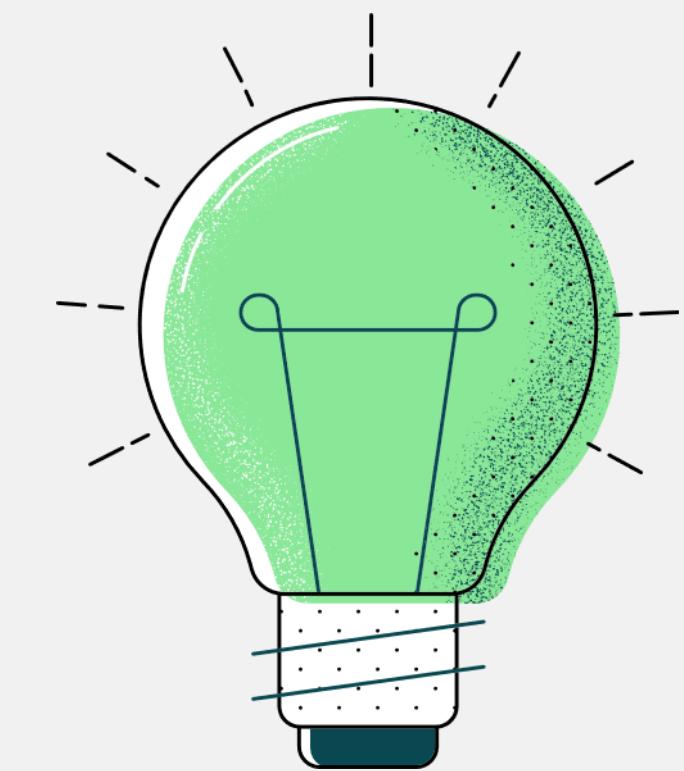
**Il dataset è statico e non si aggiorna con l'andamento dei tassi di interesse attuali.**

## Sviluppi Futuri:

**API Banking: Integrazione con API reali per scaricare i movimenti bancari veri del cliente invece di chiedere il reddito a voce.**

**Deep Learning: Testare Reti Neurali per gestire dati non strutturati.**

**MLOps: Implementare un sistema di ri-addestramento automatico ogni mese con i nuovi mutui erogati.**



# Grazie per l'attenzione!

Corso di Machine Learning 2025/2026

**Antonio Clavino**

Mat. 0512119692

**Indirizzo e-mail**

a.clavino@studenti.unisa.it

**Repository Github**

<https://github.com/An7onyOnly/Mortgage-AI>

