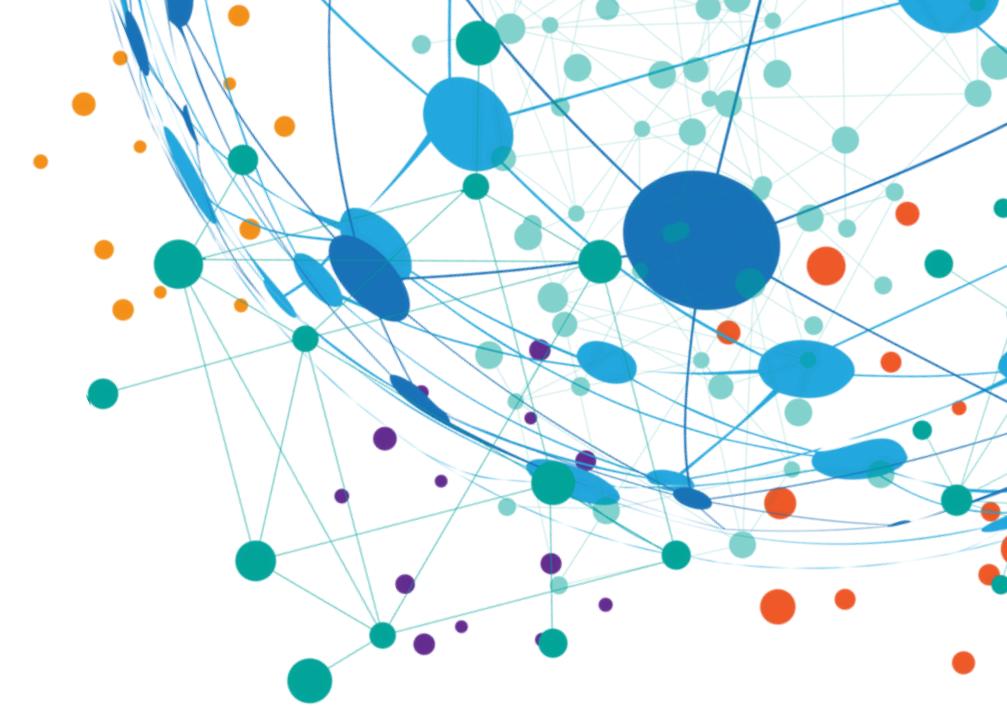




UNIVERSITÀ DEGLI STUDI DI NAPOLI  
**FEDERICO II**

DIPARTIMENTO DI  
**Scienze Sociali**



# Società Iperconnessa e Latenza di Rete: Valutazione della Quality of Experience dell'Utente tramite Predizione dell'RTT in Reti 4G e 5G utilizzando il Machine Learning

## Relatore

Prof.ssa Stefania Zinno

## Correlatore

Prof. Adam Erik Arvidsson

Anno accademico  
2023/2024

## Candidato

Simone Buglione  
Mat. P41000139



# Contesto di ricerca

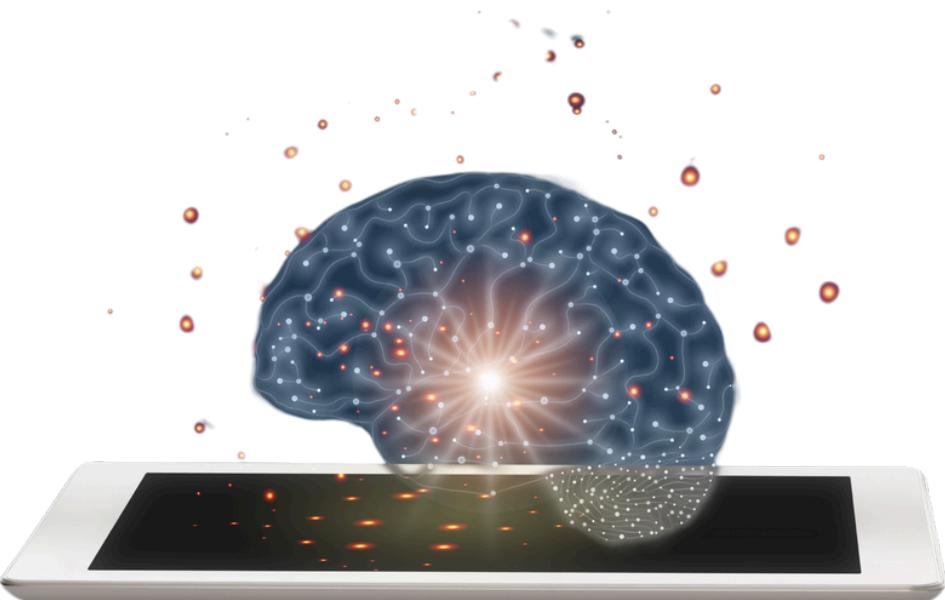
- ▶ Il passaggio dai media tradizionali ai media digitali ha provocato **profondi cambiamenti** nella società, basati sul concetto di **mediatizzazione** e sul "**villaggio globale**" (McLuhan).
- ▶ In passato, si considerava il pubblico come **passivo** e influenzato dai media senza reagire (Klapper), mentre oggi esso **partecipa attivamente** alla ricerca di **gratificazioni** (McQuail e Gurevitch).
- ▶ Nasce il concetto di **iperconnessione** (Quan-Haase e Wellman), che delinea l'attuale **società iperconnessa**, all'interno della quale gli individui sono costantemente raggiungibili e online.
- ▶ Ciò ha accentuato la necessità di **reti sempre più performanti**, requisito cruciale sostenuto dall'industria delle telecomunicazioni allo scopo di garantire elevate prestazioni di rete.





# Obiettivo della ricerca

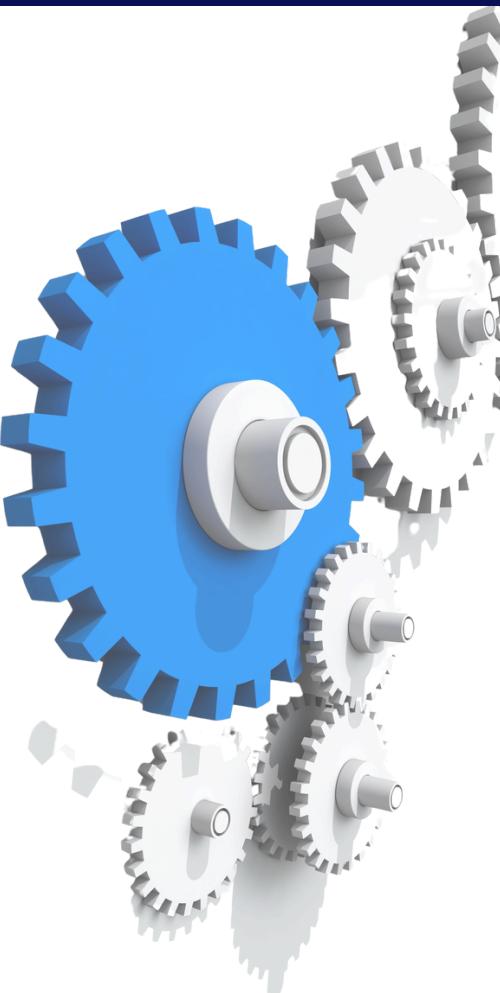
Ottimizzare i **parametri di qualità del segnale** di reti 4G e 5G  
attraverso la previsione della **latenza di rete (RTT)**, mediante  
l'addestramento di modelli predittivi di **Machine Learning (ML)**,  
utili ad migliorare la **qualità del servizio (QoS)** e la **qualità  
dell'esperienza (QoE)** utente online.





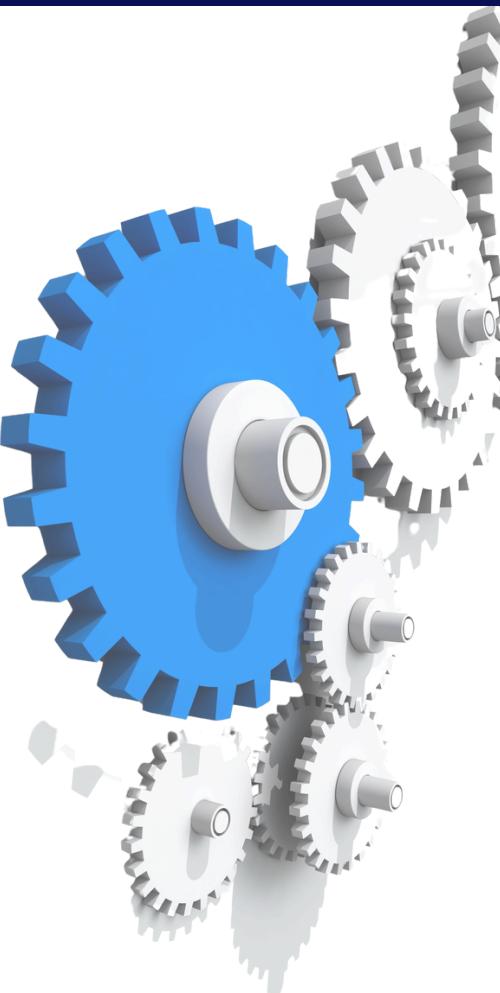
# Metodologia

- ▶ Per la previsione è stata utilizzata la campagna di misurazione condotta su Roma “**Latency test – Online Gaming – Active Measurements**” e basata su **smartphone** in due scenari: **passeggiando all’aperto (OW)** e **guidando un veicolo (OD)**.
  
- ▶ Per l’analisi è stato utilizzato l’editor di codice **Visual Studio Code** con linguaggio di programmazione **Python** e la libreria **Scikit-learn**.





# Metodologia



- ▶ I dati sono stati analizzati utilizzando a scopo di classificazione gli algoritmi **Decision Tree (DT)**, **Random Forest (RF)** e **Gradient Boosting (GB)**.
  
- ▶ I risultati sono stati visualizzati attraverso un **Report di classificazione** e visualizzati graficamente tramite la **curva Receiver Operating Characteristic (ROC)**, la **curva Precision-Recall** e la **curva di apprendimento**.



# Pre-processing

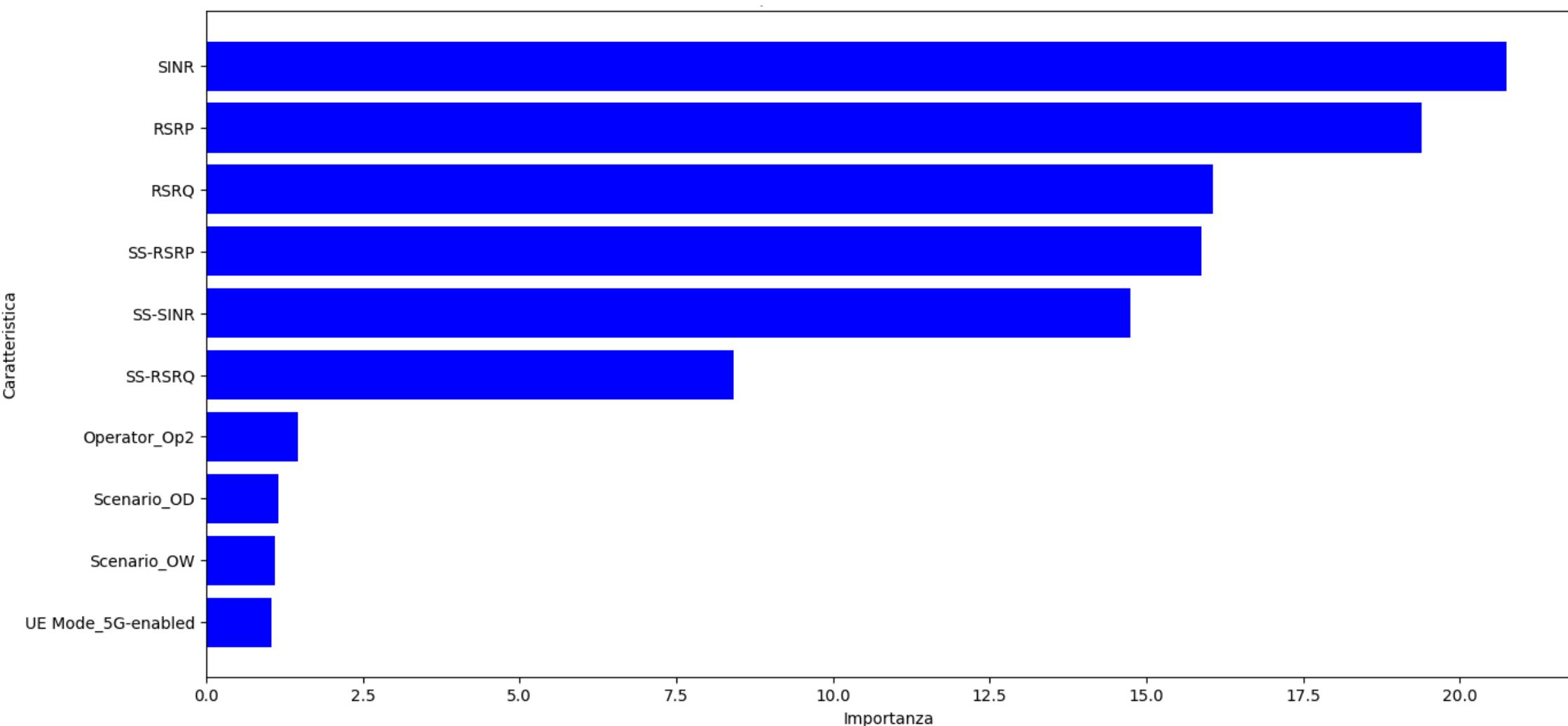
- ▶ E' stata innanzitutto effettuata la gestione e conteggio dei **valori mancanti** (NA) e riempiti attraverso il metodo **Forward Fill (FFILL)**.
- ▶ Le unità di tempo sono state convertite in **millisecondi (Ms)** per il calcolo della latenza.
- ▶ Le variabili categoriali sono state codificate tramite **One-Hot-Encoder** per essere utilizzate nei modelli di Machine Learning.
- ▶ Infine, sono state selezione le **variabili indipendenti** più rilevanti, quali qualità e potenza del segnale ricevuto (RSRQ-RSRP) e rapporto segnale-rumore (SINR) e la **variabile dipendente** del tempo di andata e ritorno (RTT).





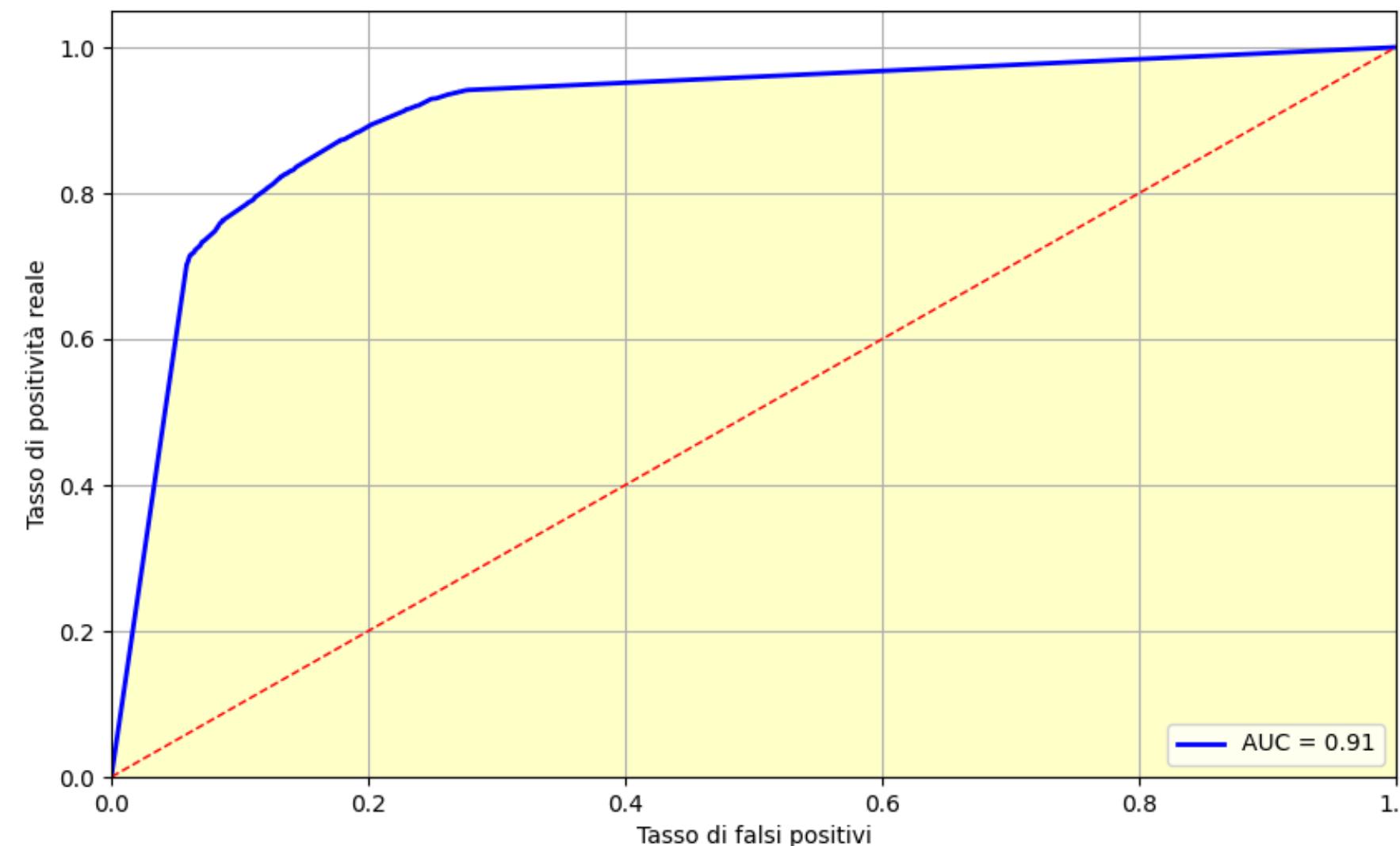
# Decision Tree/Feature importance

Struttura ad albero - Facile interpretazione - Overfitting comune - Algoritmo greedy - Deterministico - Decisioni binarie in ogni nodo



- ▶ Rapporto segnale-rumore e Potenza segnale ricevuto 4G maggiormente influenti sulla predizione

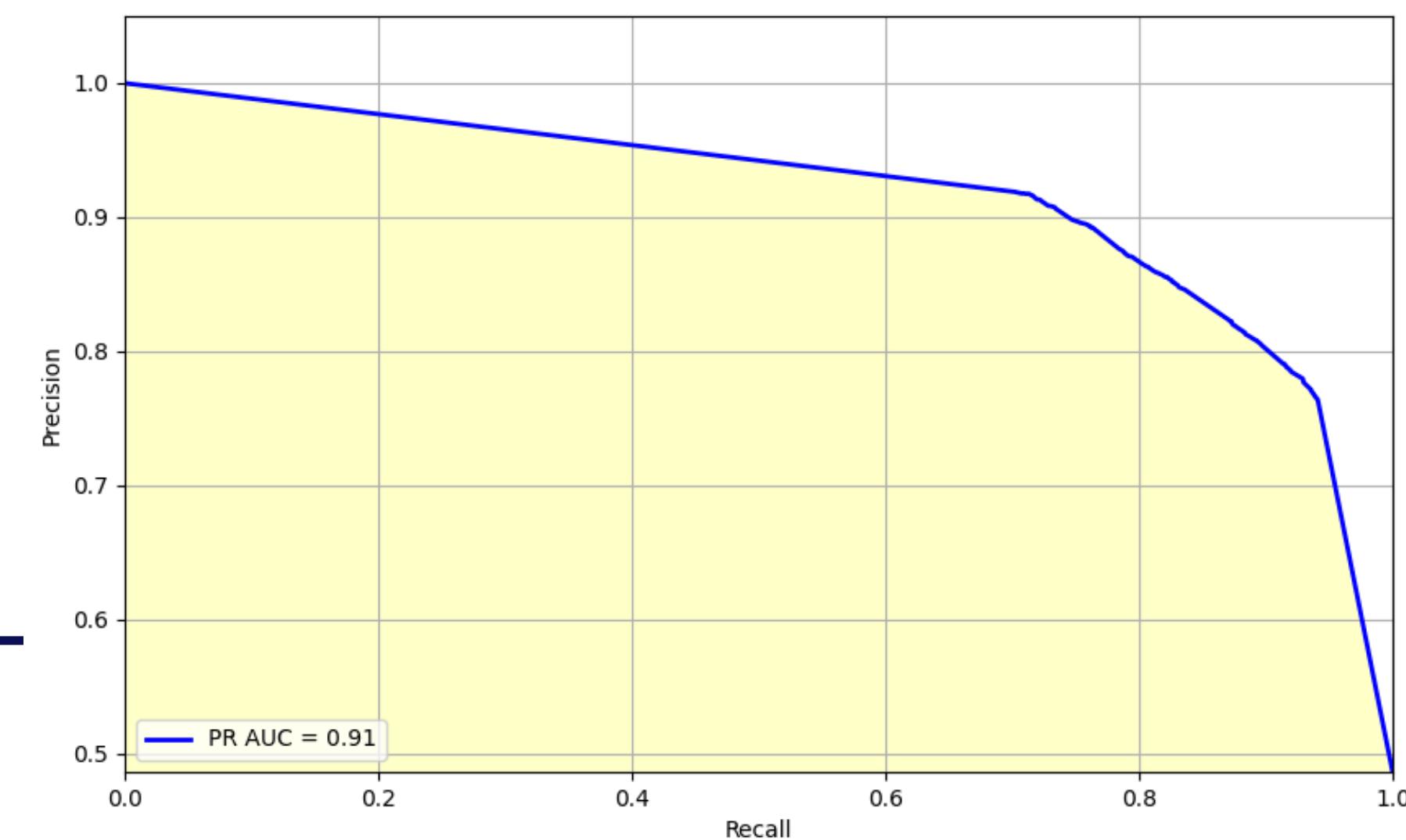
# Curva ROC



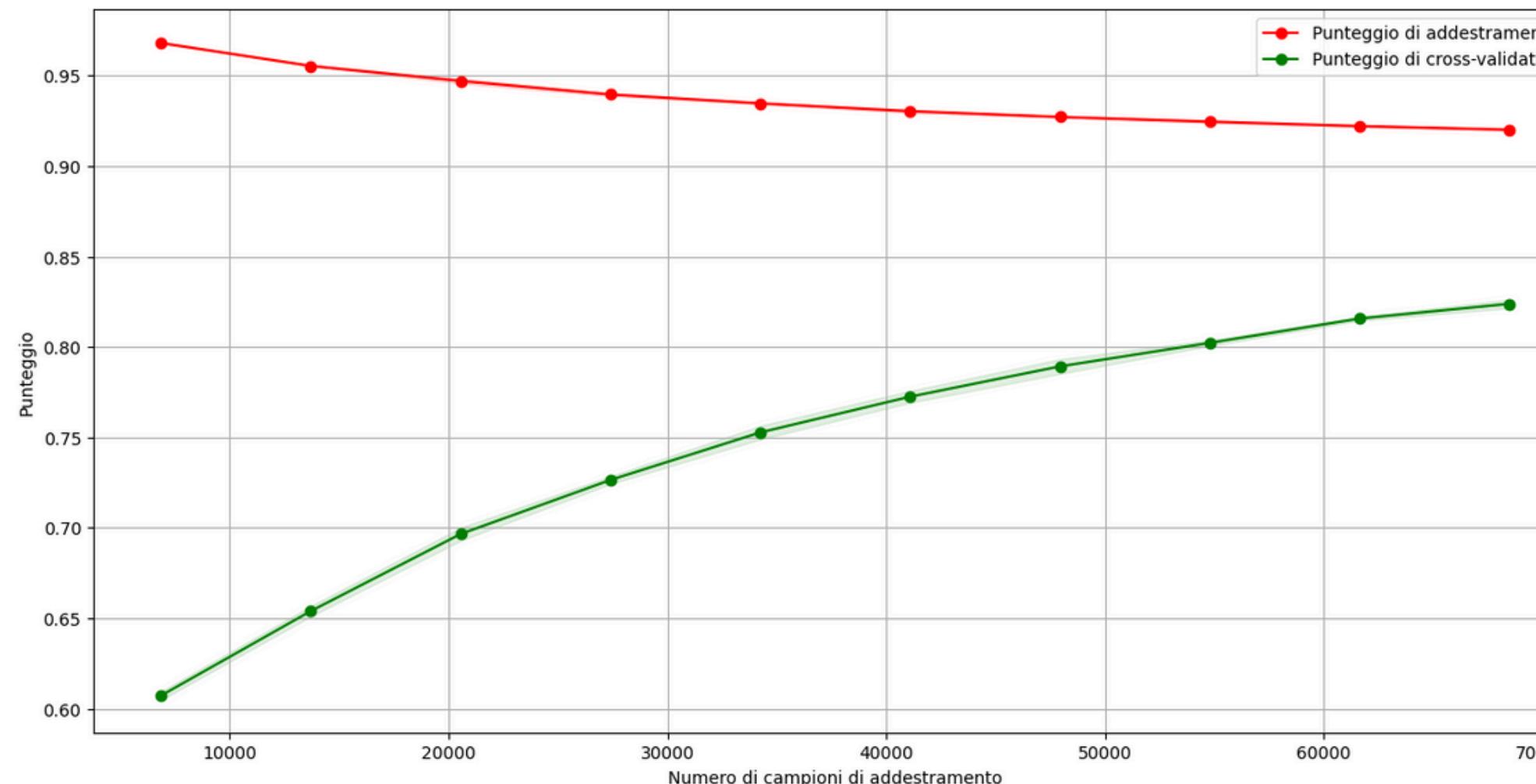
- ▶ Alta precisione
- ▶ Ottime prestazioni anche in situazioni di sbilanciamento

- ▶ Ottima capacità di classificazione
- ▶ Elevata accuratezza

# Curva Precision-Recall



# Curva di apprendimento



Accuratezza sul set di addestramento: 0.916579051891311  
Accuratezza sul set di test: 0.8462688659408439

- 
- ▶ Alto punteggio di addestramento
  - ▶ Punteggio CV in costante aumento
  - ▶ Buona generalizzazione del modello

## Report di classificazione

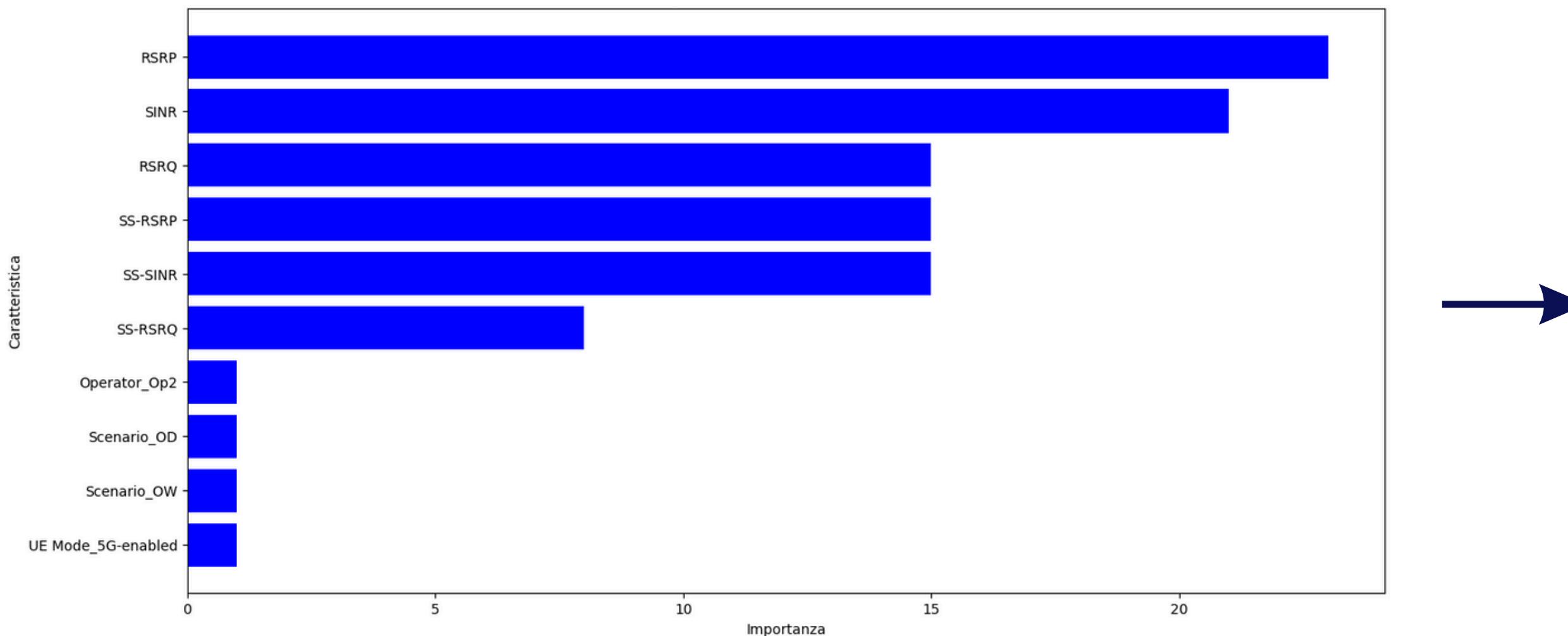
- ◀
- ▶ Alta accuratezza
  - ▶ Equilibrio nelle predizioni tra le classi

	precision	recall	f1-score	support
buono	0.85	0.87	0.86	10993
cattivo	0.86	0.84	0.85	10408
accuracy			0.86	21401
macro avg	0.86	0.86	0.86	21401
weighted avg	0.86	0.86	0.86	21401



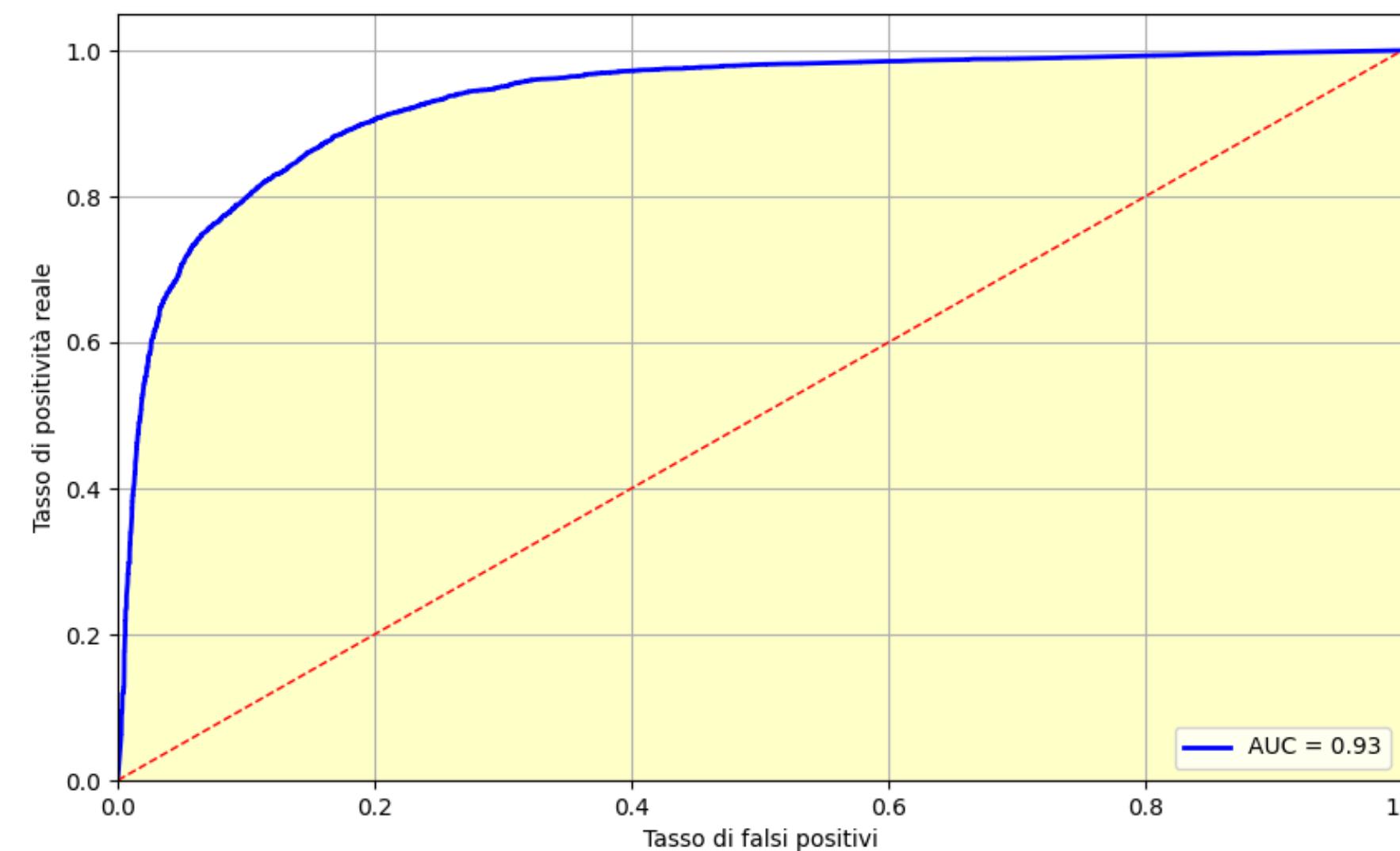
# Random Forest/Feature importance

Ensemble di alberi - Riduzione dell'overfitting - Media delle predizioni -  
Robustezza e stabilità - Alta capacità predittiva



- ▶ Potenza segnale ricevuto e rapporto segnale-rumore 4G maggiormente influenti sulla predizione

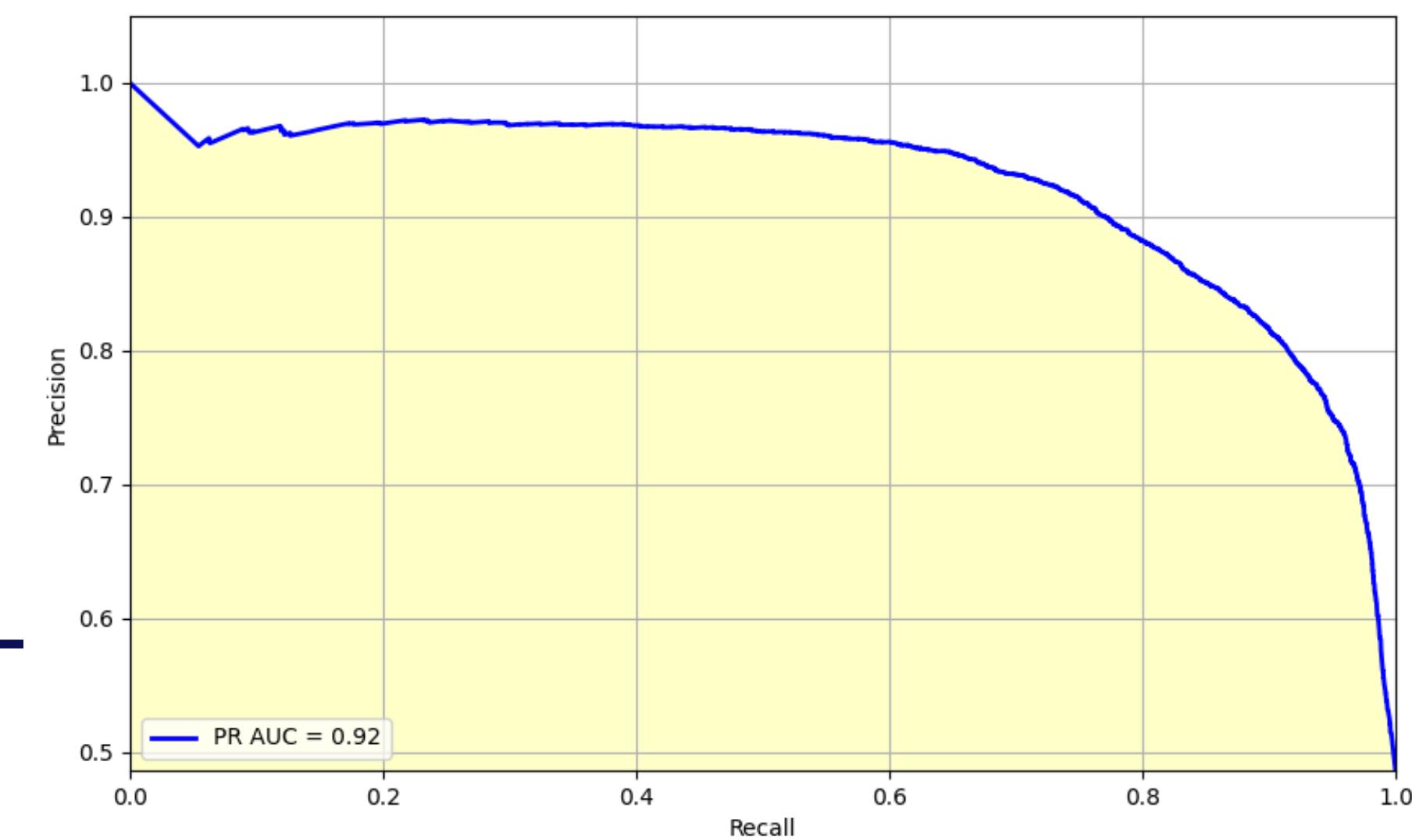
# Curva ROC



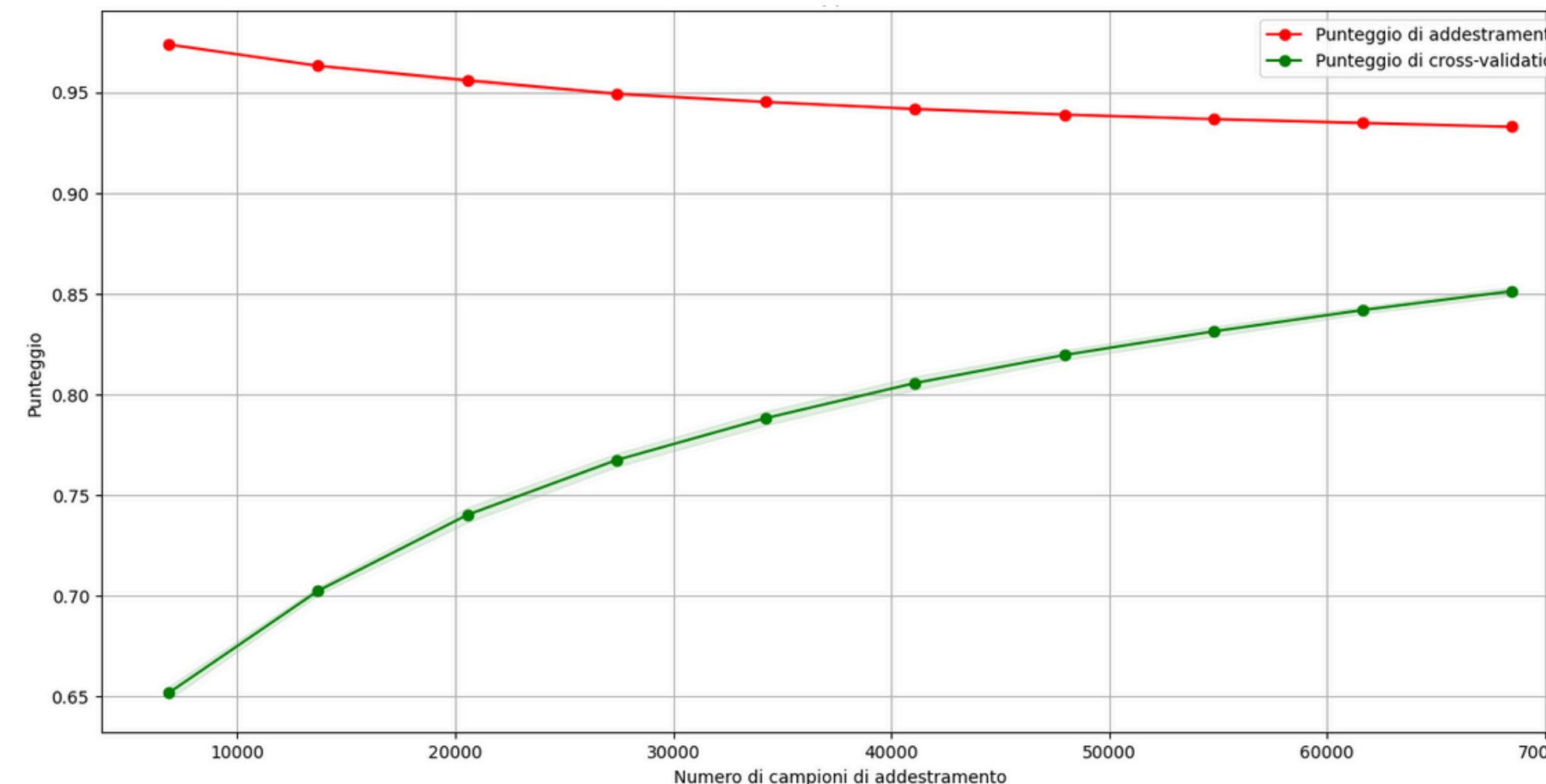
- ▶ Alta precisione
- ▶ Prestazioni robuste anche in situazioni di sbilanciamento

- ▶ Eccellente capacità di classificazione
- ▶ Elevata accuratezza

# Curva Precision-Recall



# Curva di apprendimento



Accuratezza sul set di addestramento: 0.916579051891311  
Accuratezza sul set di test: 0.8558945843652166

- 
- ▶ Elevato punteggio di addestramento
  - ▶ Punteggio CV in costante aumento
  - ▶ Ottima generalizzazione del modello

## Report di classificazione

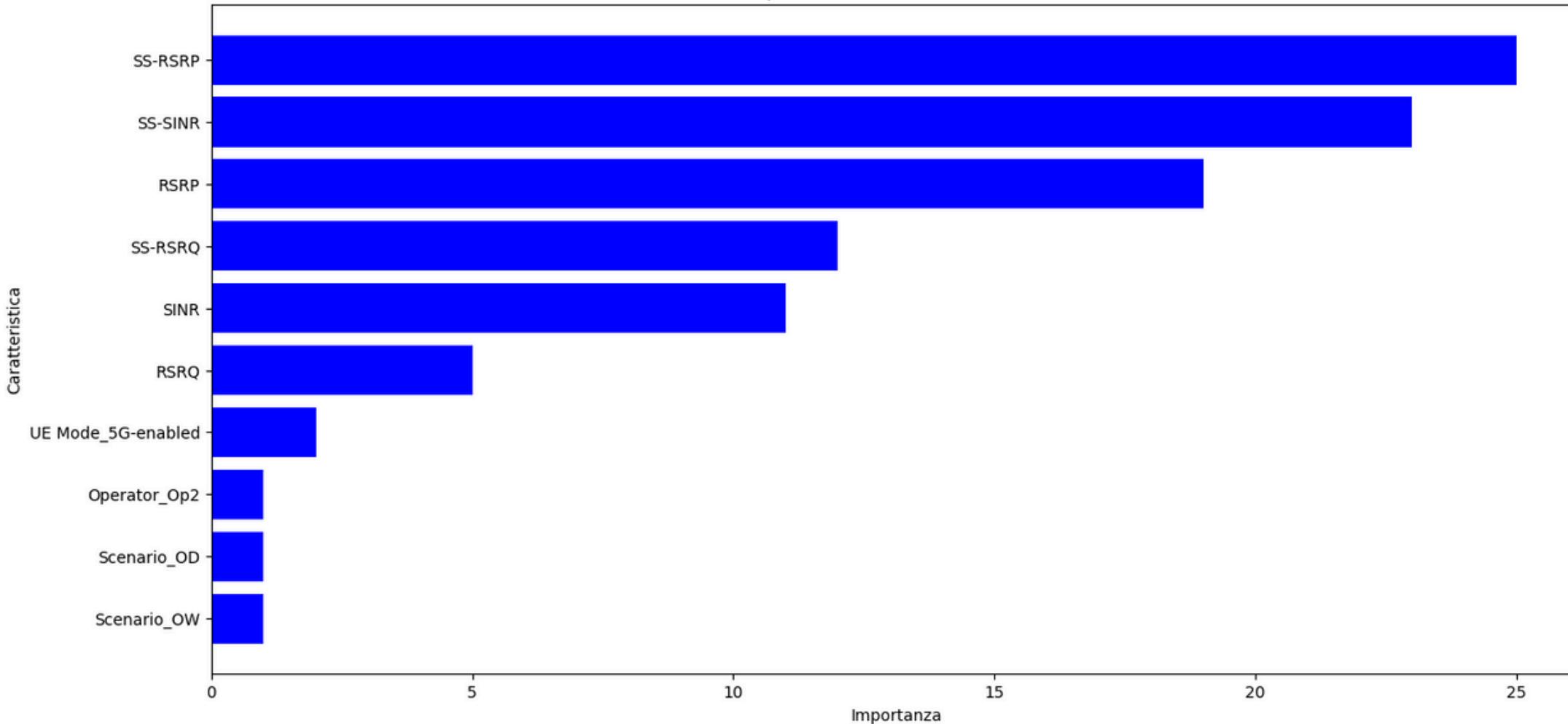
- ◀
- ▶ Elevata accuratezza
  - ▶ Equilibrio nelle predizioni tra le classi

	precision	recall	f1-score	support
buono	0.87	0.87	0.87	10993
cattivo	0.86	0.86	0.86	10408
accuracy			0.87	21401
macro avg	0.87	0.87	0.87	21401
weighted avg	0.87	0.87	0.87	21401



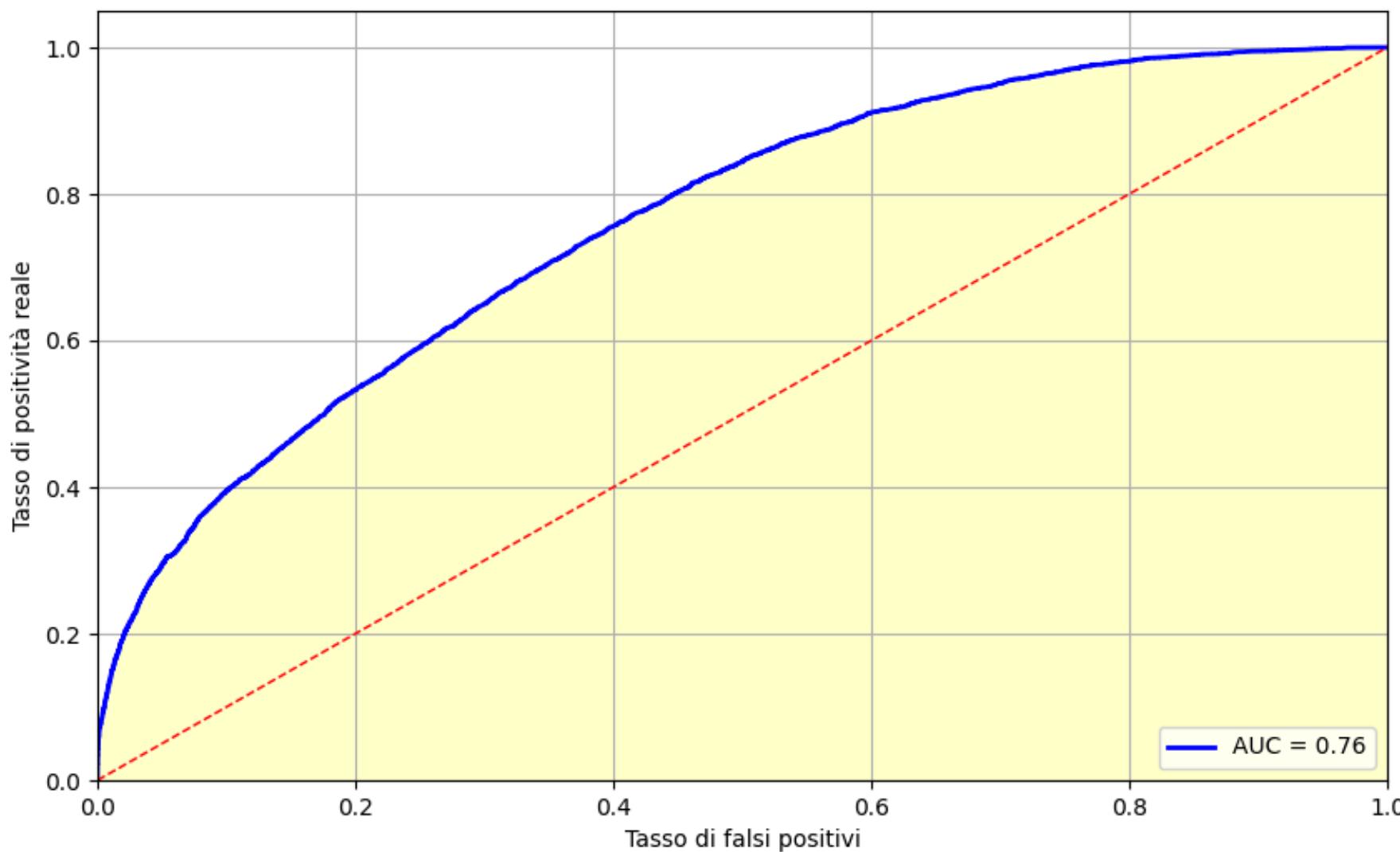
# Gradient Boosting/Feature importance

Ensemble di modelli deboli - Ottimizzazione gradientale - Riduzione dell'errore residuo - Overfitting mitigabile con regularizzazione - Alta capacità predittiva



- ▶ Potenza segnale ricevuto e rapporto segnale-rumore 5G maggiormente influenti sulla predizione

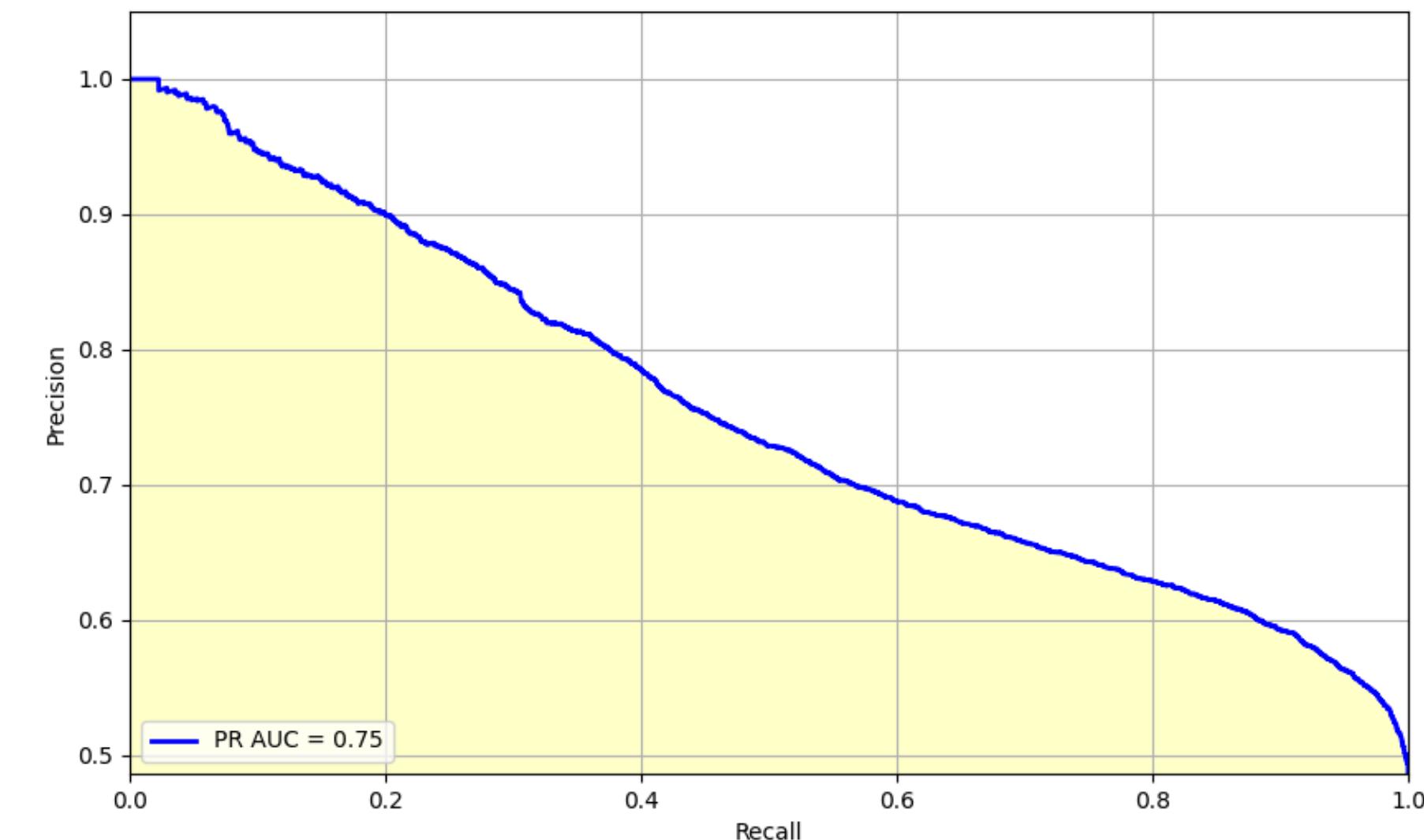
# Curva ROC



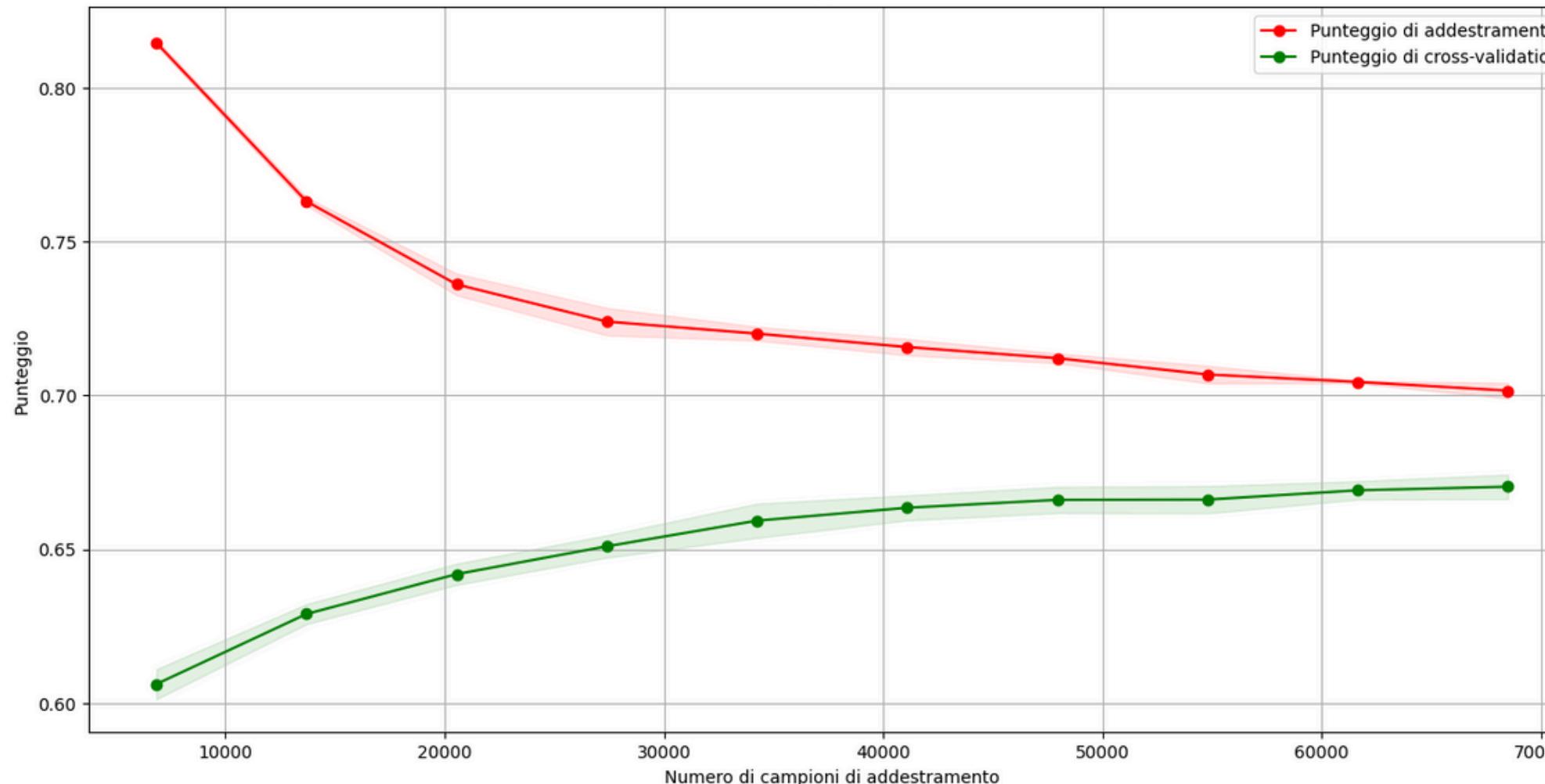
- Discreta precisione
- Discrete prestazioni anche in situazioni di sbilanciamento

- Buona capacità di classificazione
- Buona accuratezza

# Curva Precision-Recall



# Curva di apprendimento



Accuratezza sul set di addestramento: 0.6979626644237285  
Accuratezza sul set di test: 0.6766973505910939

- ▶ Punteggio di addestramento in calo
- ▶ Punteggio CV in aumento
- ▶ Discreta generalizzazione del modello

## Report di classificazione

- ▶ Buona accuratezza
- ▶ Discreto equilibrio nelle predizioni tra le classi ←

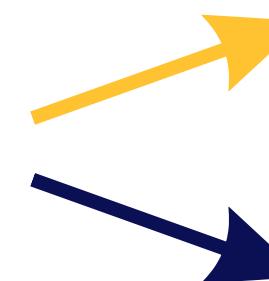
	precision	recall	f1-score	support
buono	0.69	0.68	0.69	10993
cattivo	0.67	0.67	0.67	10408
accuracy			0.68	21401
macro avg	0.68	0.68	0.68	21401
weighted avg	0.68	0.68	0.68	21401



# Conclusioni

- ▶ **Random Forest:** → Miglior modello nel contesto di ricerca
- ▶ **Decision Tree:** → Ottime prestazioni
- ▶ **Gradient Boosting:** → Modello meno performante

Possibile applicazione dei modelli addestrati allo scopo di:



Migliorare la Quality of Service  
in campi sensibili alla latenza

Migliorare la Quality of  
Experience dell'utente in rete

# Limiti

- ▶ Strumentazione e potenza computazionale non ottimale per l'esplorazione di dati su vasta scala
- ▶ Contesto specifico di raccolta dati limitato su Roma

# Prospettive future

- ▶ Esplorare ulteriori contesti geografici
- ▶ Esplorare set di dati contenenti variabili di possibili influenza
- ▶ Testare ulteriori modelli



UNIVERSITÀ DEGLI STUDI DI NAPOLI  
**FEDERICO II**

DIPARTIMENTO DI  
**Scienze Sociali**



# GRAZIE PER L'ATTENZIONE

**Relatore**

Prof.ssa Stefania Zinno

**Correlatore**

Prof. Adam Erik Arvidsson

**Anno accademico**  
2023/2024

**Candidato**

Simone Buglione  
Mat. P41000139