

# HOMEWORK 3

Nicotra Giulio 1000063414



## CLASSIFICAZIONE CON IMMAGINI RUMOROSE

In questo homework si è affrontato il problema della classificazione delle immagini CIFAR-10 alterate da quadrati colorati. L'obiettivo principale è stato quello di costruire un sistema in grado di classificare correttamente le immagini, nonostante la presenza del **rumore**, evitando fenomeni di **overfitting**.

---



## DESCRIZIONE DEL DATASET

Il dataset utilizzato è **CIFAR-10**, composto da immagini a colori di dimensione  $32 \times 32$  pixel, distribuite in 10 classi (come aerei, automobili, animali, ecc.). Le immagini di addestramento erano corrotte con **quadrati** dai colori vivaci (rosso, verde, blu, giallo e blu scuro), inseriti in posizioni casuali. Il test set invece era stato preservato integro, per valutare la capacità del modello di generalizzare su dati non disturbati.

---



## METODOLOGIA: FASE DI PREPROCESSING

Ogni immagine è stata convertita in formato **uint8** e sottoposta ad analisi dei valori **RGB**. Utilizzando soglie sui canali colore, si è creata una **maschera binaria** per identificare le aree contenenti quadrati colorati. Successivamente, si è applicata una tecnica di **inpainting** basata su **cv2.inpaint()**, che ha ricostruito le regioni corrotte in modo coerente con il contesto circostante, restituendo immagini visivamente pulite.

---



## METODOLOGIA: FASE DI CLASSIFICAZIONE

Le immagini sono state normalizzate e appiattite in vettori. Per estrarre rappresentazioni significative, si è utilizzato un **MLPClassifier** di scikit-learn: le probabilità in uscita dalla rete sono state interpretate come **feature non lineari**. Queste feature sono state usate come input di un **XGBoost**, scelto per la sua capacità di migliorare la generalizzazione e rafforzare la classificazione finale.

---



## RISULTATI

Il modello finale, valutato sul test set pulito, ha ottenuto un'**accuratezza** del **48%**. La **matrice di confusione** evidenzia una distribuzione equilibrata delle predizioni tra le classi, senza bias marcati. L'approccio **MLP + XGBoost** si è dimostrato più efficace del solo MLP in termini di accuratezza e capacità di distinguere le diverse categorie.

---

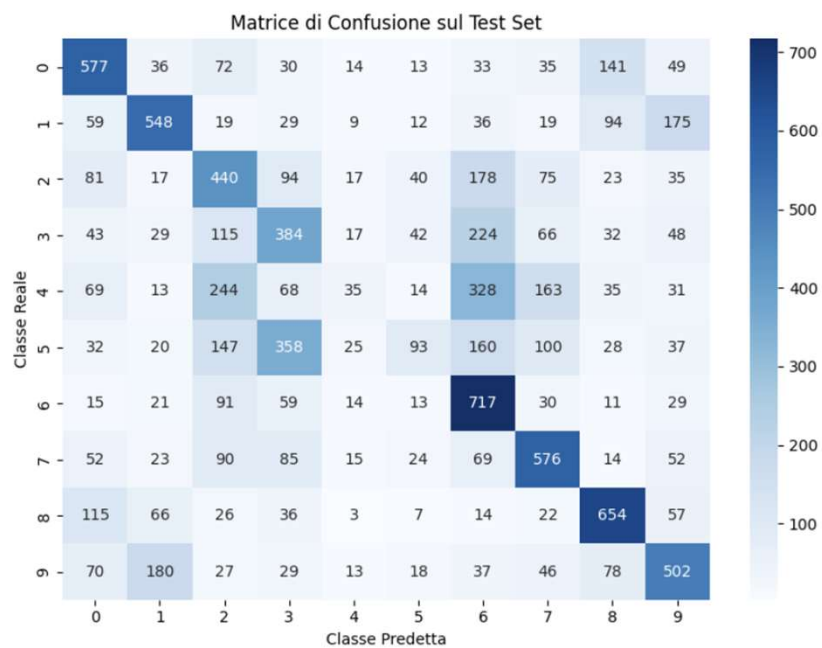


Immagine originale



Immagine ripulita



Immagine originale



Immagine ripulita



Immagine originale



Immagine ripulita



Immagine originale



Immagine ripulita



Immagine originale



Immagine ripulita

