

Classificazione immagini con CIFAR10

Model selection e confronto tra modelli.

Nome: Sciacca Erika · Data: 02/05/2025

Descrizione del Dataset CIFAR10

Dimensioni

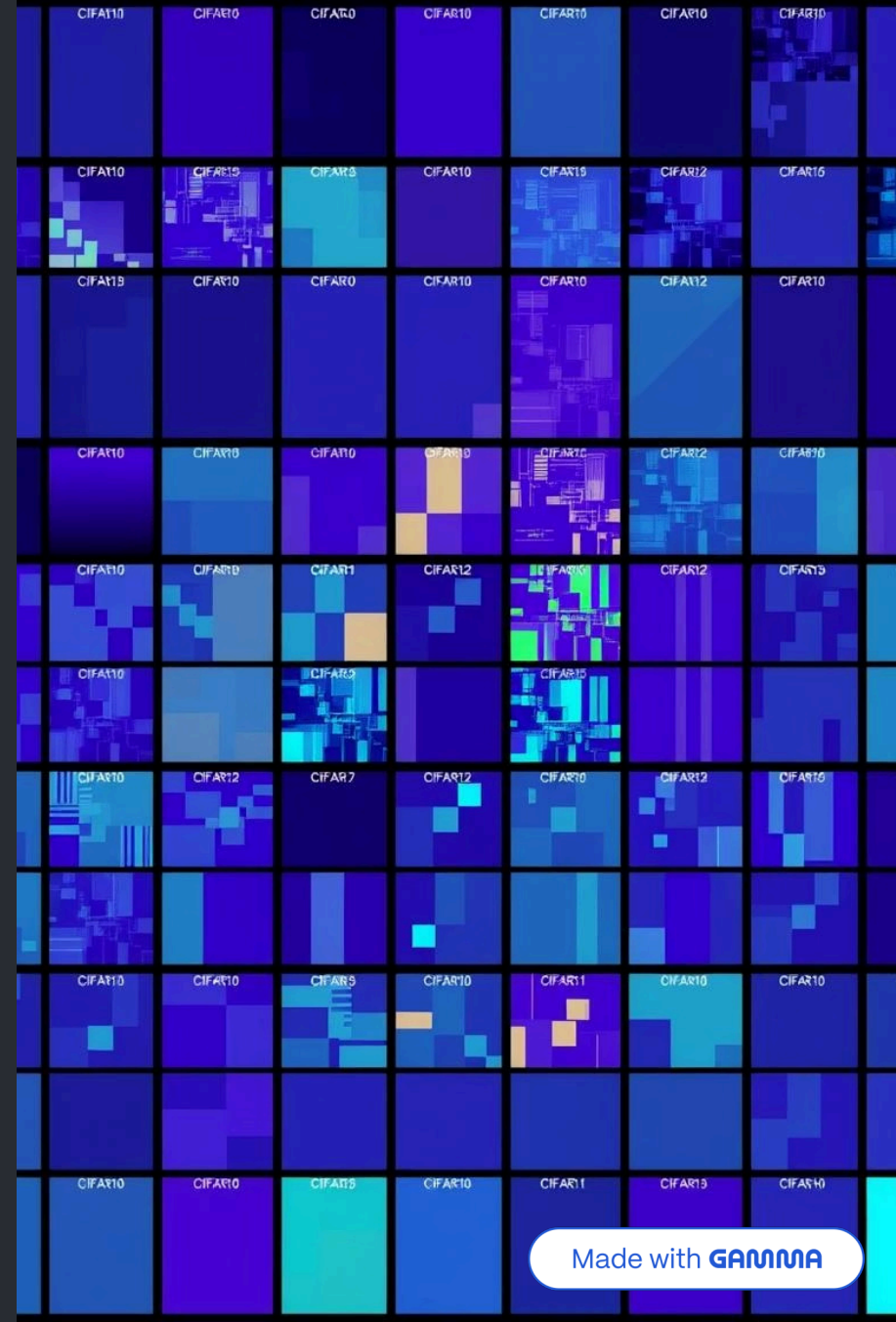
60.000 immagini RGB 32x32 pixel, 10 classi.

Suddivisione

50.000 per training, 10.000 per test.

Obiettivo

Classificare correttamente le immagini per classe.



Preprocessing e Sottocampionamento

☐ Sotto-campionamento

Riduce il dataset per modelli computazionali pesanti.

In questo caso ho scelto di considerare 10.000 immagini su 60.000 per cercare di effettuare una model selection accurata diminuendo i tempi di esecuzione.

☐ Vantaggi

Accelera il training e riduce l'overfitting.

☐ PCA

Riduce la dimensionalità mantenendo l'informazione principale.

Considero $n_components = 200$ che mantiene circa il 95% della varianza.



Strategia di Model Selection

Metodo 1

- Train e validation usati per la selezione modello.
- Riaddestro il modello sul training completo (train + validation) e testo sul set di test per valutare le performance.

Metodo 2

- Unione train + validation per rifare la selezione del modello
- Testo il modello sul set di test

Obiettivo

Garantire scelta robusta:

Il primo metodo mi permette di capire **quale modello scegliere** con rigore statistico.

Il secondo metodo mi permette di capire **l'effetto del maggior volume di dati** e la **robustezza** della selezione iniziale, **ma non per rifare la scelta del modello.**



SVM alecming



K-NN.

Modelli Utilizzati

Regressione Logistica

SVM (Support Vector Machine)

K-NN (K Nearest Neighbors)

Decision Tree

Tutti ottimizzati tramite GridSearchCV per iperparametri(tranne Regressione Logistica)

Perché GridSearchCV

Ricerca esaustiva

Scansione completa su griglia predefinita parametri.

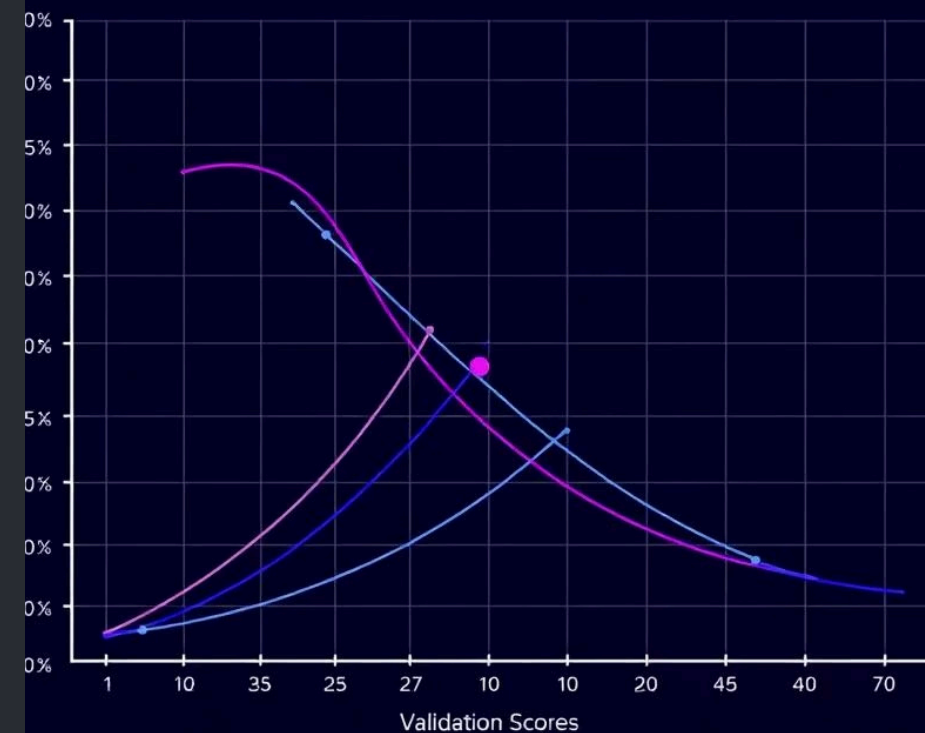
Validazione incrociata

Valuta performance media per evitare overfitting.

Applicazioni

- SVM: kernel, C
- K-NN: numero di vicini k, metrica
- Decision Tree: profondità, min samples split, min samples leaf

Hyperparameter Tuning with Validation Score:



Metriche di Valutazione

Accuracy

Percentuale di classificazioni corrette totali.

Confusion Matrix

Analisi dettagliata di errori su classi sbilanciate.

Classification Report

Precision, recall e F1-score per ogni classe.



Risultati Comparativi Primo Metodo (Validation Set)

Modello	Accuracy	F1-score	Precision	Recall
Logistic Regression	29%	0.2907	0.2923	0.2907
SVM	45%	0.4536	0.4562	0.4533
K-NN	32%	0.3035	0.4108	0.3247
Decision Tree	24%	0.2133	0.2266	0.2413



Risultati Comparativi Secondo Metodo(Test Set)

Modello	Accuracy	F1-score	Precision	Recall
Logistic Regression	27%	0.2684	0.2702	0.2680
SVM	47%	0.4767	0.4815	0.4747
K-NN	30%	0.2873	0.3787	0.2993
Decision Tree	27%	0.2642	0.2718	0.2680

Classifica Modelli

(per entrambi i metodi)

1

SVM

lperparametri ottimali: {'C': 10, 'kernel': 'rbf'}

2

K-NN

lperparametri ottimali: {'metric': 'manhattan', 'n_neighbors': 7}

3

Logistic Regression

lperparametri : {'max_iter':5000, 'solver': 'liblinear', 'C': 1.0}

4

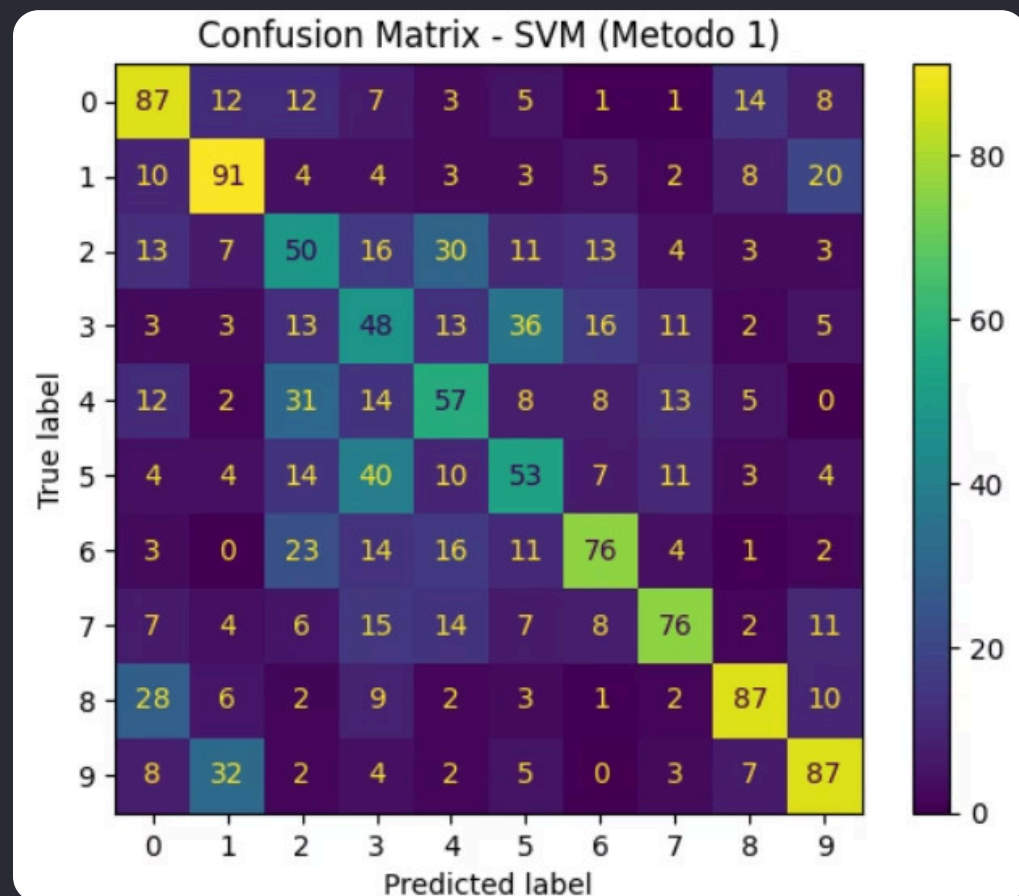
Decision Tree

lperparametri ottimali: {'criterion': 'entropy', 'max_depth': 7, 'min_samples_leaf': 2, 'min_samples_split': 2}

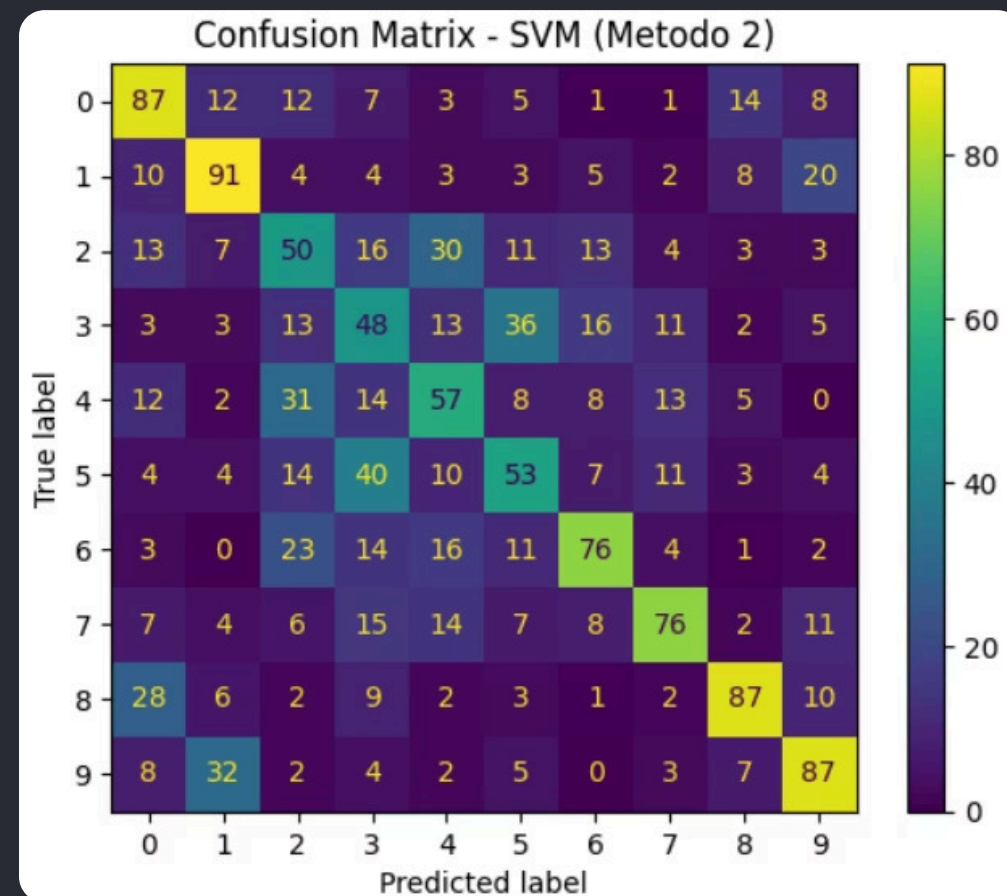
Confusion Matrices del Miglior Modello

In entrambi i metodi è stato scelto il modello SVM con la stessa configurazione degli iperparametri

Primo Metodo



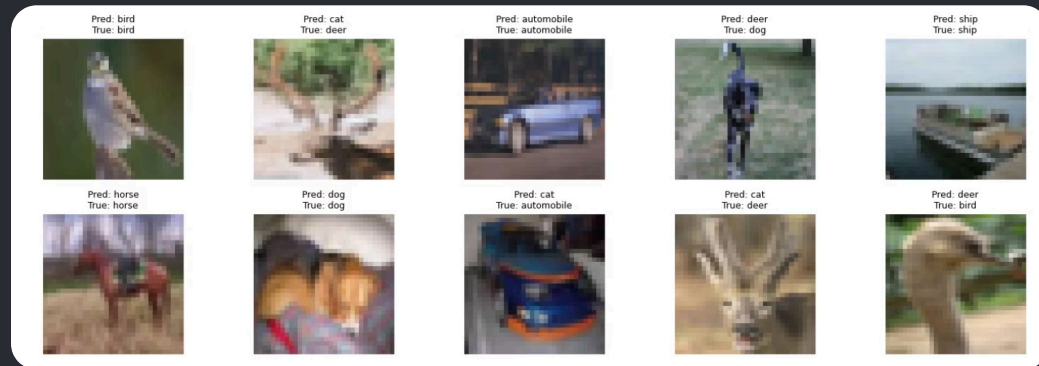
Secondo Metodo



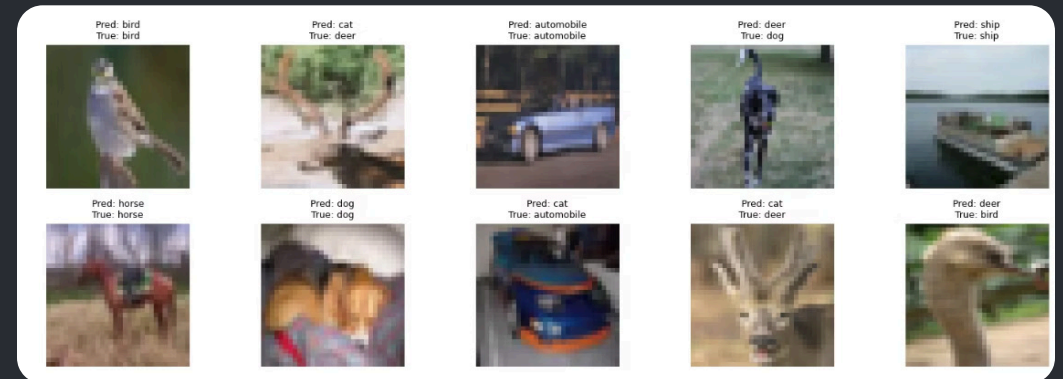
Predizioni delle immagini del Miglior Modello

In entrambi i metodi è stato scelto il modello SVM con la stessa configurazione degli iperparametri

Primo Metodo



Secondo Metodo



Conclusioni

Efficienza e prestazioni

Sottocampionamento e PCA
essenziali per modelli complessi.

Ottimizzazione

GridSearchCV ha identificato
iperparametri ottimali.

Robustezza

Abbiamo osservato che il modello selezionato nella prima fase (SVM) mantiene buone prestazioni anche dopo il riaddestramento con più dati. Il confronto con altri modelli riaddestrati mostra che SVM continua a offrire prestazioni competitive, confermando la validità della scelta iniziale.

E' bene precisare che:

Anche se un altro modello (es. k-NN) avesse ottenuto una leggera performance superiore nel secondo metodo, non potevamo comunque considerarlo il migliore in assoluto poiché tale valutazione è avvenuta sul test set. Dunque il secondo metodo viene utilizzato solo per valutare l'effetto del maggior volume di dati e la robustezza della selezione iniziale

