

# ***HOMEWORK 2***

- Caricamento del dataset
- Diminuisco il dataset fino al 20%
- Appiattisco le immagini
- Creo i modelli: k-NN, regressione logistica, SVM, decision tree
- Trovo il modello migliore
- Visualizzo le matrici di confusione

# CIFAR-10

CARATTERISTICA	DESCRIZIONE
Numero totale immagini	60.000
Training set	50.000 immagini
Test set	10.000 immagini
Dimensioni immagini	32x32 pixel
Tipo di dato	Array di interi(valori tra 0 e 255)
Formato immagine	32x32x3→3072 features(dopo flattening)
Tipo di task	Classificazione multiclasse(10 classi)
Distribuzione classi	uniforme (6.000 immagini per classe)

Le 10 classi sono:aeroplano,automobile,uccello,gatto,cervo,cane,rana,cavallo,nave,camion.

# Risultati

## k-NN

Accuracy su validation set: 0.408

Accuracy su test set: 0.2795

Confusion Matrix - K-NN

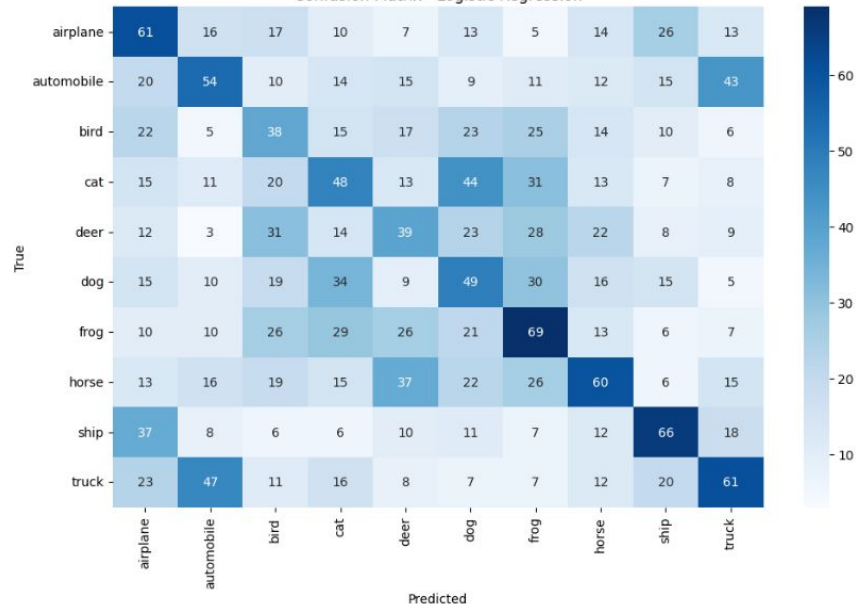


## Regressione Logistica

Accuracy on validation set: 0.268

Accuracy on test set: 0.2725

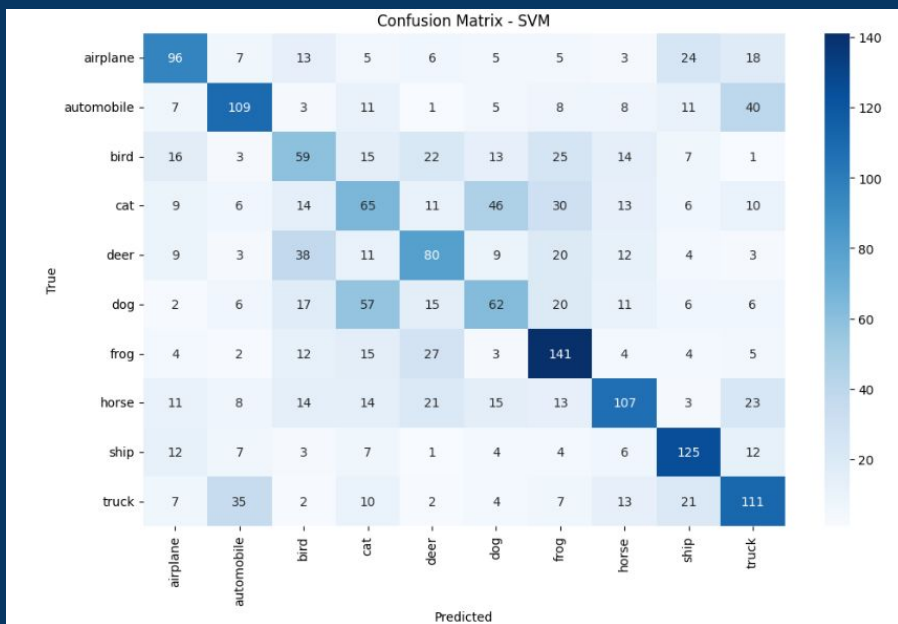
Confusion Matrix - Logistic Regression



# SVM

Accuracy on validation set: 0.7185

Accuracy on test set: 0.4775



# Decision Tree

Accuracy on test set: 0.2535

Accuracy on validation set: 0.235

