



TEXT CLASSIFICATION PER INDIVIDUARE GLI ARTISTI A PARTIRE DAL TESTO DELLE LORO CANZONI

A cura di:

Inzadi Greta – 813649

Portaluppi Alessandro – 816090

Testa Luca – 816000

STRUTTURA DEL PROGETTO



PRE – PROCESSING
DEL TESTO



RAPPRESENTAZIONE
E CLASSIFICAZIONE

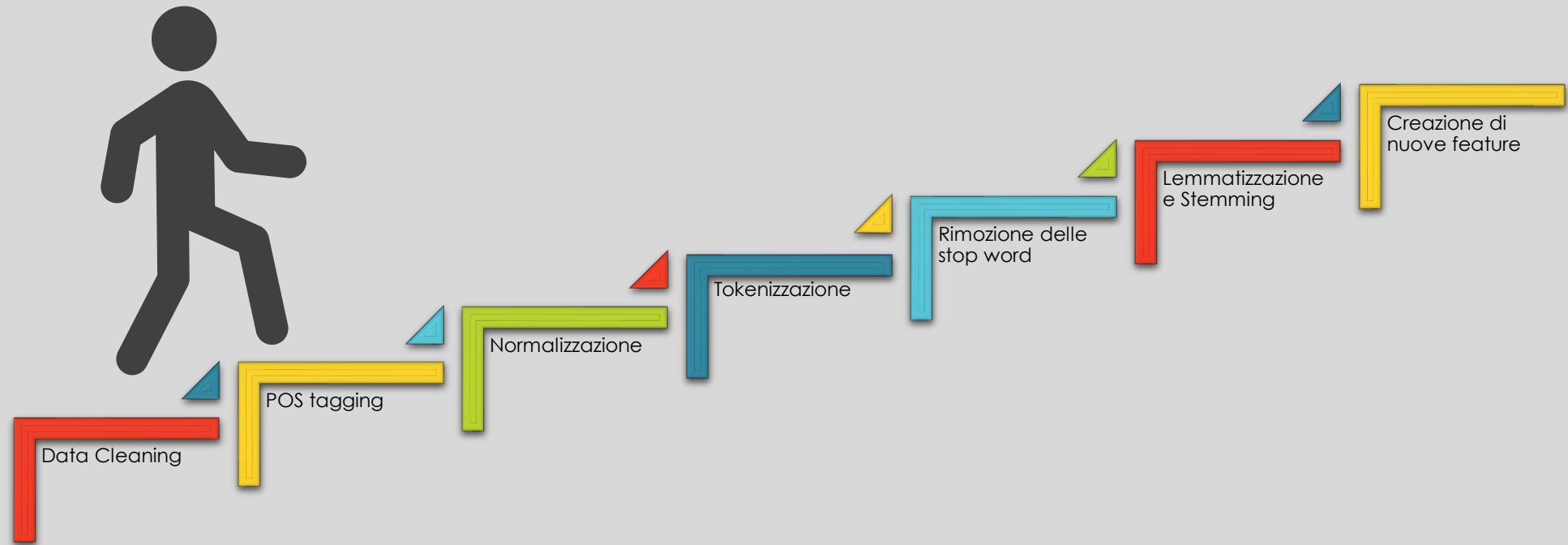


VALUTAZIONE



VISUALIZZAZIONE

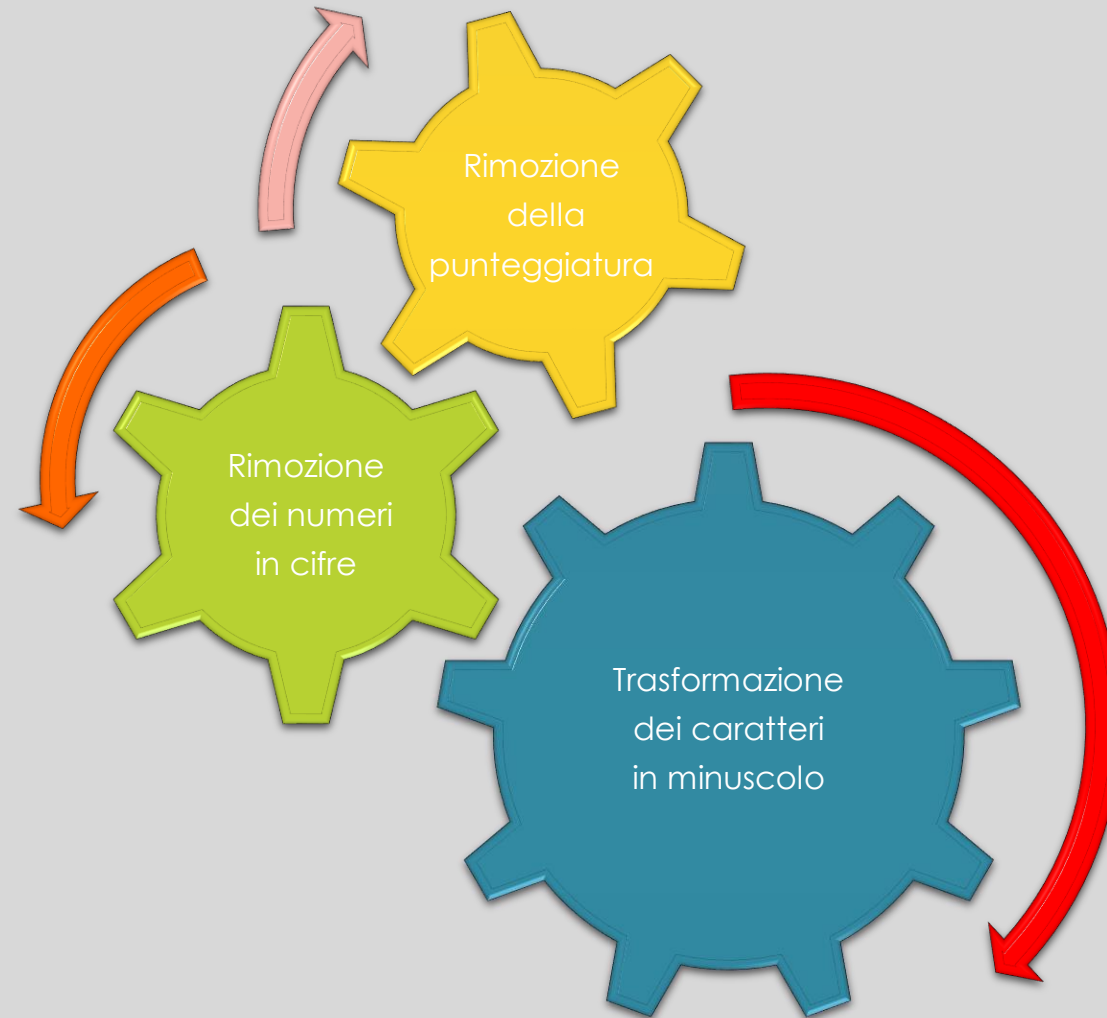
PRE – PROCESSING DEL TESTO ED ESTRAZIONE DELLE FEATURE



DATA CLEANING E POS TAGGING



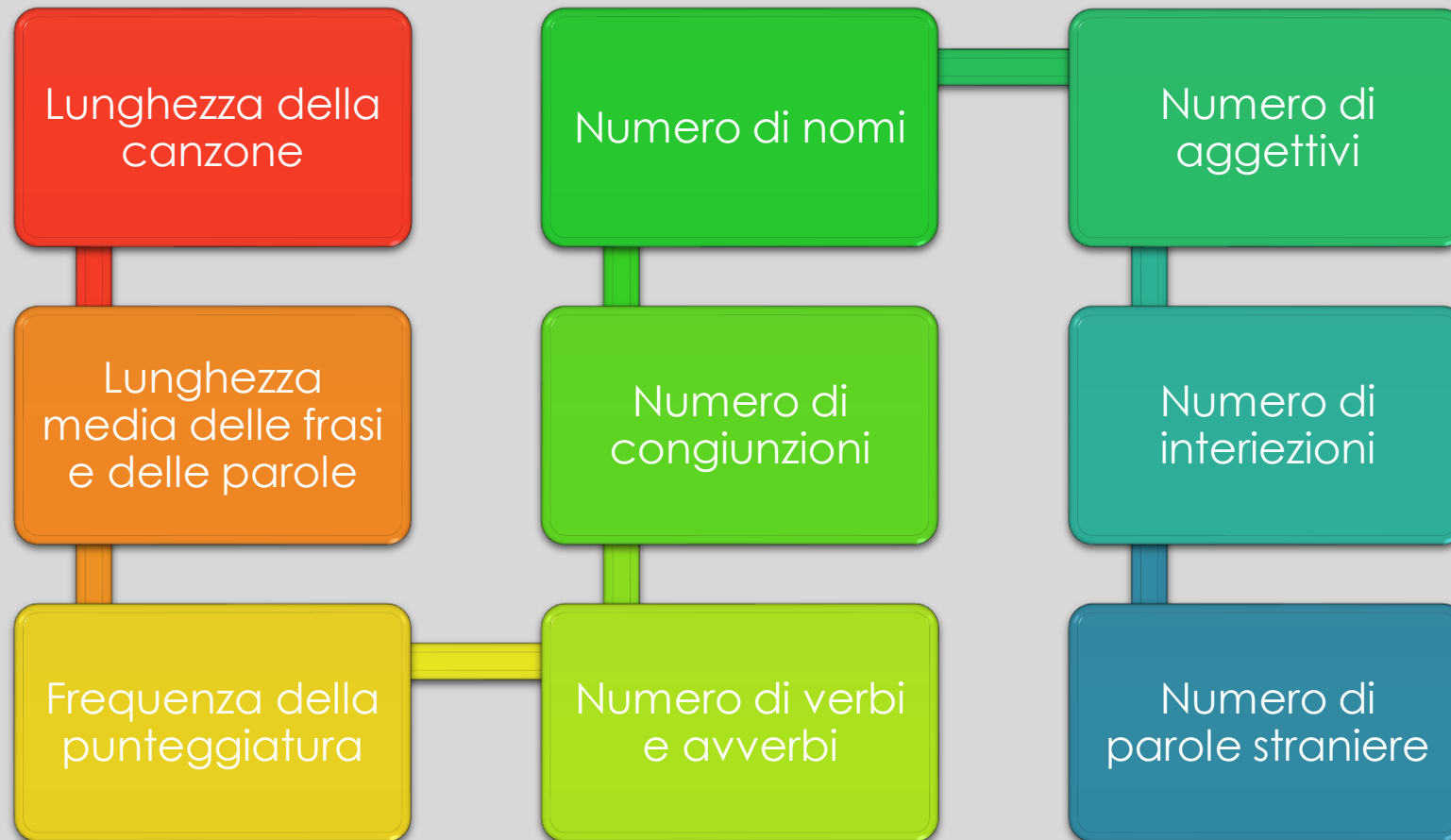
NORMALIZZAZIONE



TOKENIZZAZIONE, STOP WORD, LEMMATIZZAZIONE E STEMMING



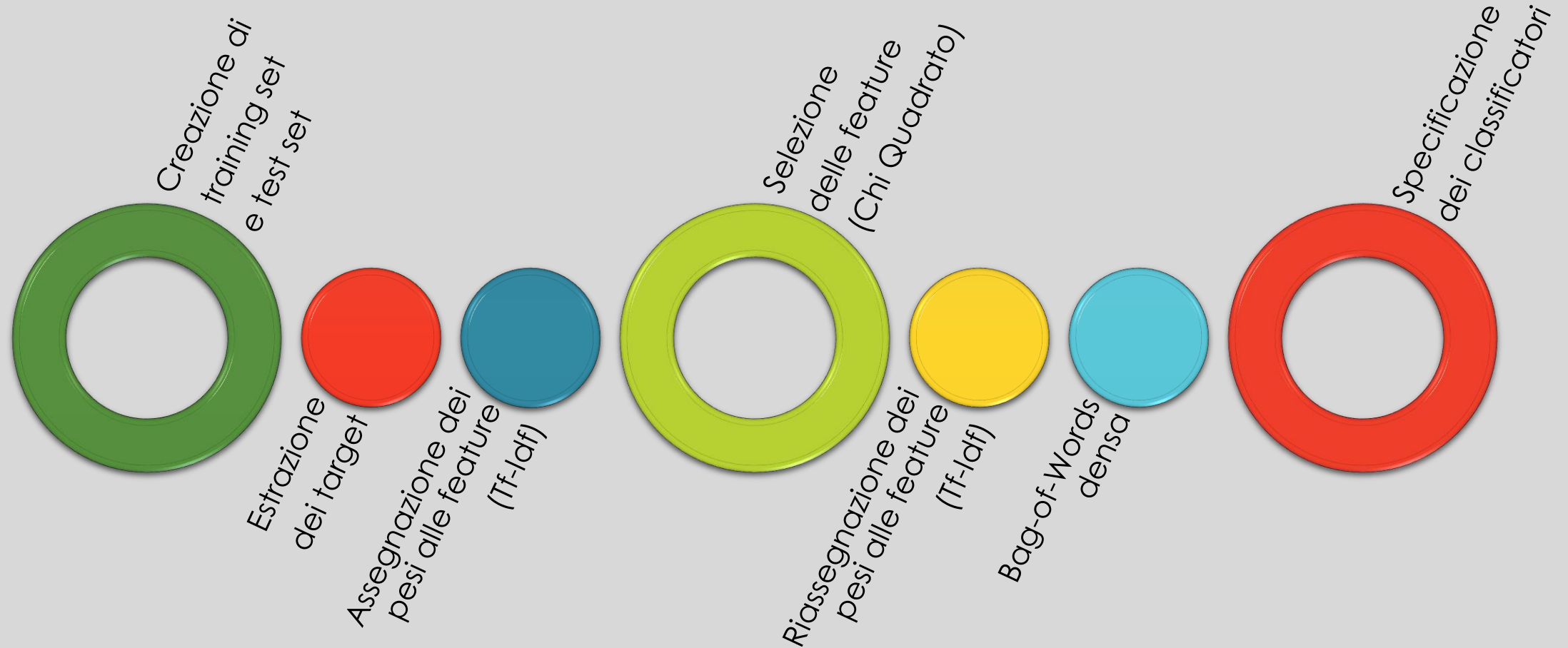
CREAZIONE DI NUOVE FEATURE



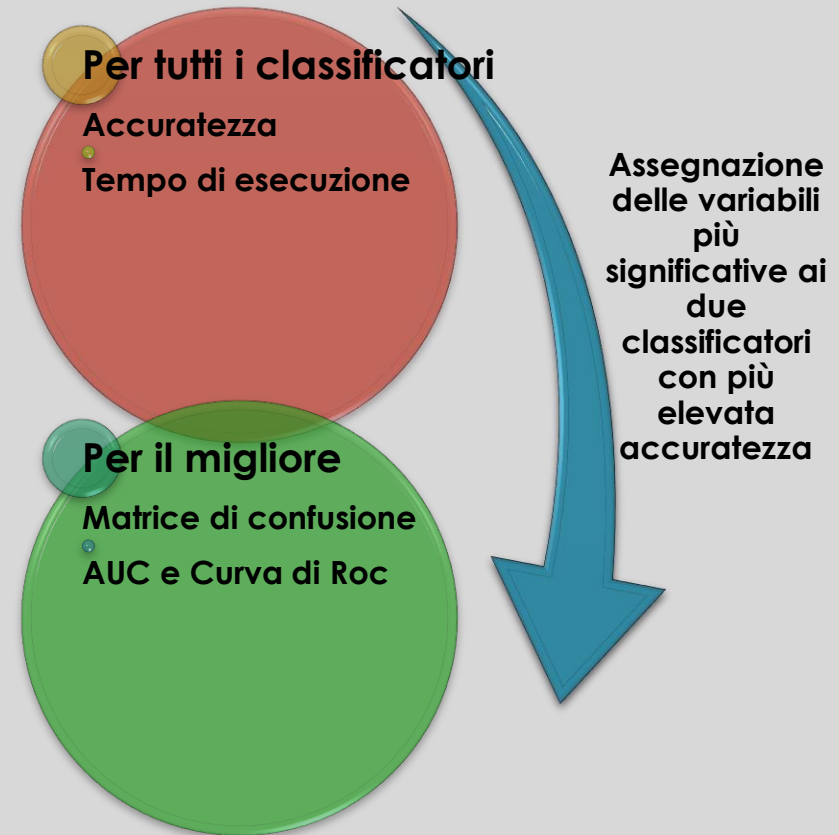
RAPPRESENTAZIONE E CLASSIFICAZIONE DEL TESTO



PREPARAZIONE



CLASSIFICAZIONE E METRICHE



VALUTAZIONE

Accuratezza



Decision Tree

Random Forest

Ada Boost Classifier

Gradient Boosting Classifier

XGB Classifier

LGBM Classifier

KNN

Performance



VISUALIZZAZIONE

Genere Pop



Genere Hip-Pop



Genere Rock

CONCLUSIONI E SVILUPPI FUTURI



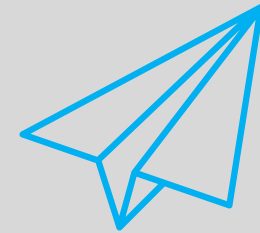
Approfondimento e applicazione della stilometria



Utilizzo di reti neurali e altri classificatori (e metodi di word embedding)



Incremento dei testi per artista nel training set per migliorare l'accuratezza



Applicazione Web



GRAZIE PER L'ATTENZIONE