Student: Neferu Ana-Maria

Grupa: 232

# Documentatie Proiect IA

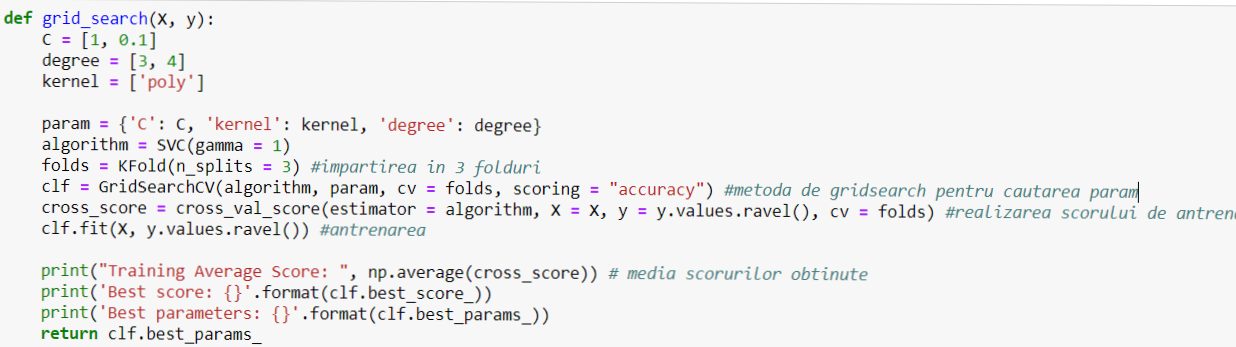
Kernel (default:rbf, linear, sigmoid,precomputed, poly) setat cu poly.

Algoritmul este un SVM, am aplicat clasificatorul SVC care are setat doar parametrul gamma cu 1.

In functia denumita grid\_search:

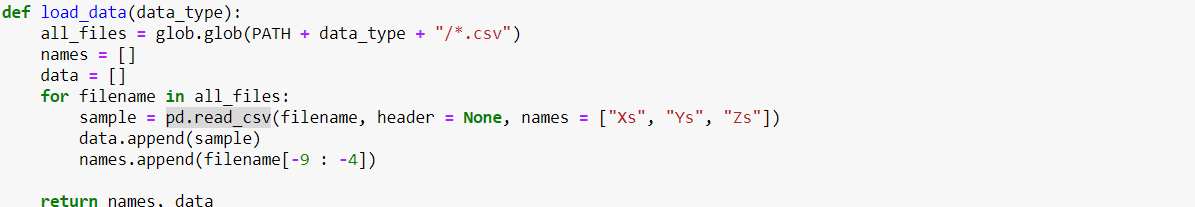
Am folosit KFold() pentru impartirea in 3 fold-uri ( 3 fiind si nr default). Am aplicat metoda GridSearchCV() din sklearn.model\_selection care primeste ca parametrii SVC-ul, C-ul, kernelul si degree, foldurile si are setat parametru scoring cu accuracy pentru a cauta parametrii.

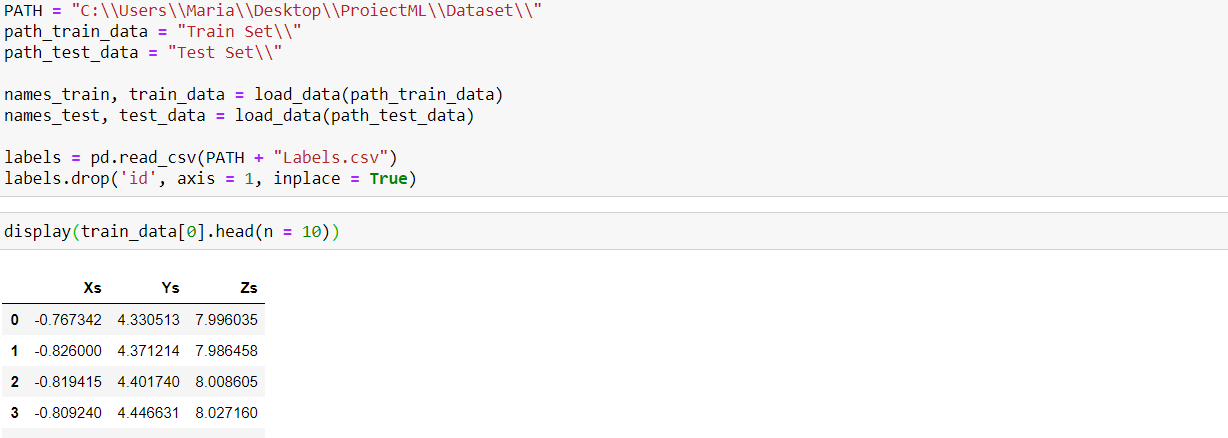
Metoda Cross\_val\_score() din sklearn.model\_selection unde am setat parametrul estimator pentru a face fit datelor cu algoritmul folosit (SVC), datele care trebuiesc fit cu x, y ce reprezinta variabila tinta pentru a incerca sa prezica: cu y.values.ravel() si parametrul cv care stabileste strategia de splitting cu rezultatul dupa aplicarea KFold-ului, toate acestea cu scopul de a realiza scorul de atrenare cu cross validarea. Am facut antrenarea, dupa care am afisat si media scorurilor obtinute cu np.average(), scorul cel mai bun cu clf.best\_score() si cei mai buni parametrii cu clf.best params\_



In functia load\_data:

Am citit si am incarcat datele de train si de test cu ajutorul metodei load\_data() care a primit ca parametru calea completa a datelor primite si cu pd.read\_csv din panda am citit datele din labels punandu-le pe coloanele denumite cu Xs, Ys si Zs dupa care am si afisat primele 10 linii cu metoda display pentru a verifica daca au fost citite corect datele de antrenare.





In functia get\_statistics se afla parametrii extrasi care vor servi ca si caracteristici pentru problema de clasificare:

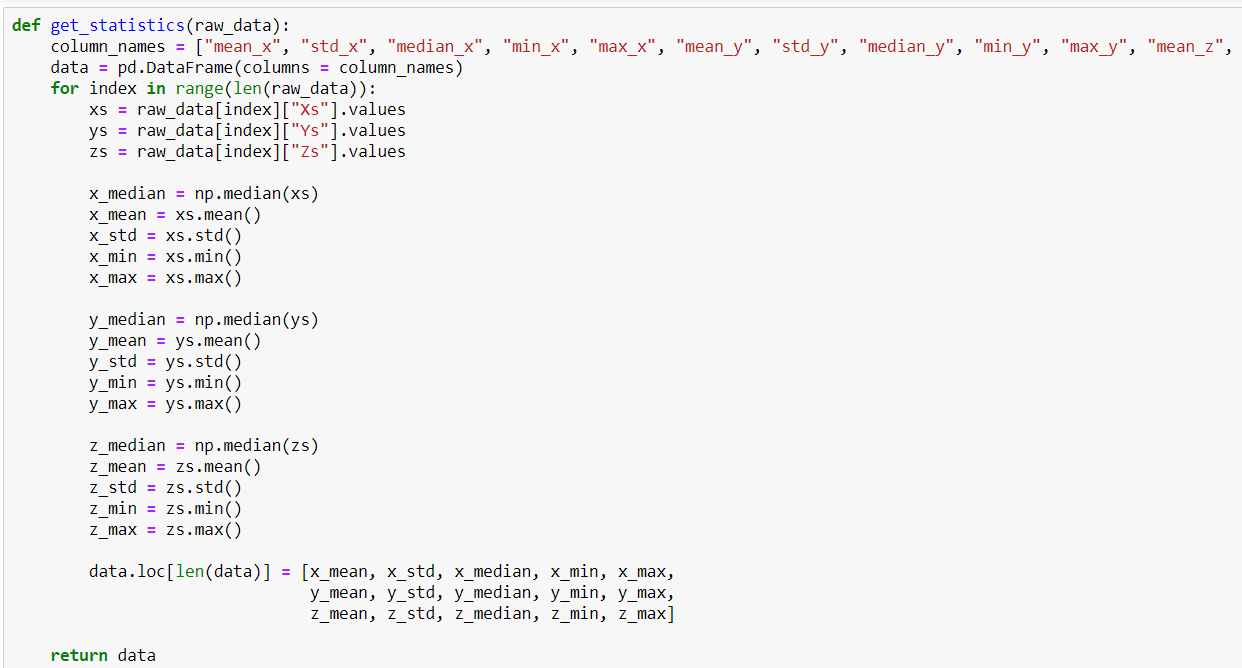
mean(): valorile medii ale datelor accelerometrului triaxial

median(): valorile medianei

std(): valorile deviatiei standard

min(): valorile minime

max(): valorile maxime



train\_test\_split() pentru impartirea listei în trepte aleatoare și subseturi de test: 80% date de antrenare, 20% date de test

Am aplicat clasificatorul pe datele de antrenare cu metoda fit() si am facut prezicerea cu metoda predict() care a primit datele de test.

Am incercat de asemenea si NaïveBays , fara success in comparatie cu scorul mediu obtinut pe datele de train de 89% si 91% pe datele de testare .