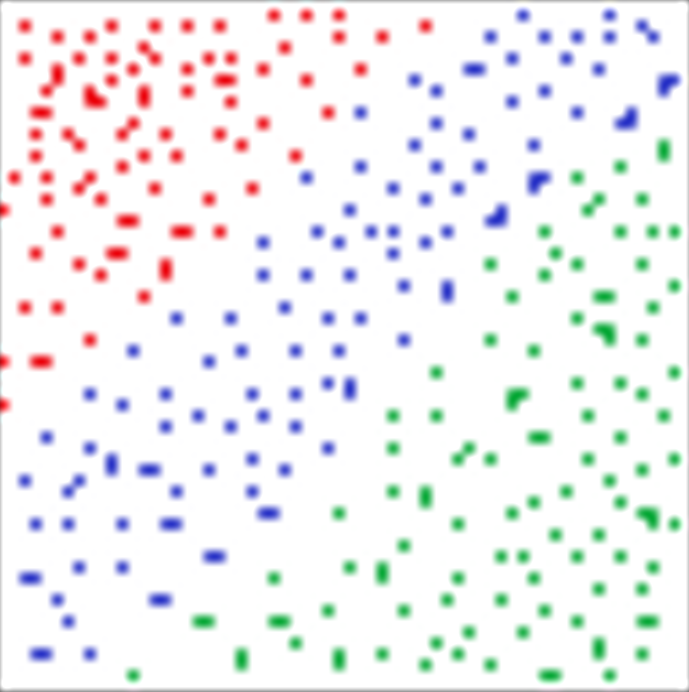
Proiect Machine Learning

Pentru acest proiect am creat 2 imagini de tipuri diferite; una dintre acestea este de tipul linear separabil și cealaltă neliniar separabil, ambele având dimensiunile de 64x64 pixeli. Acestea sunt prezentate mai jos:



**Figura 1. Neliniar separabilă (data2)**



**Figura 2. Liniar separabilă (data1)**

Pentru aceste imagini am ales următorii 3 clasificatori:

* Decision Tree Classifier
* SVC
* KNeighbor’s Classifier

Pentru clasificatorul la alegere, am ales K-Neighbor’s Classifier deoarece este destul de simplu de folosit și nu necesită date de antrenament.

În cele ce urmează, voi explica modul de lucru:

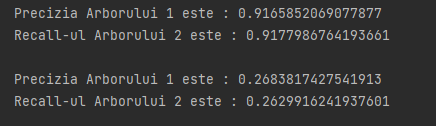
* Pentru a crea seturile de date, am creat fișierele cu numele respectiv, apoi le-am apelat cu funcția read\_data, unde am specificat locația acestora.
* Pentru a putea face un split la date pentru a le folosi pentru train si test la clasificatorii menționați mai sus am apelat la funcția train\_test\_split din biblioteca sklearn.model\_selection
* La subpunctul 7, am definit precizia și recall-ul, după cum urmează:

Precizia reprezintă un indicator al performanței modelului antrenat, în învățarea automată, indicând nivelul calitativ al prezicerilor pozitive. Se măsoară prin împărțirea rezultatelor corecte, la numărul tuturor rezultatelor returnate.

Recall-ul este capacitatea modelului antrenat de a detecta cazuri pozitive. Cu cât recall-ul este mai mare, cu atât numărul de scenarii pozitive este mai mare. Se calculează ca raport intre numărul de scenarii clasificate corect ca fiind pozitive și numărul de cazuri pozitive.

Evaluarea acuratetii, preciziei si recall-ului au fost calculate și prezentate mai jos pentru modelele făcute în proiect.

* **Decision Tree**



* Cross-validation

Figura 2

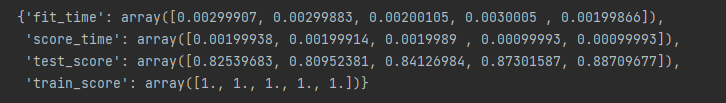
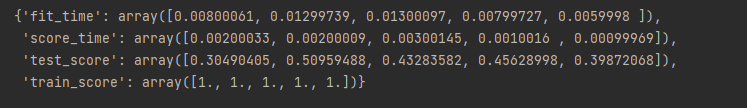
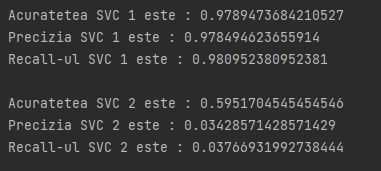


Figura 1

* SVM



Concluzii intre acuratețile obținute

* **Decision Tree Classifier**

În cadrul acestui clasificator, pentru imaginile neliniar separabile, acuratețea scade cu aproximativ 25%.

* **SVM**

După cum se poate observa, aceasta metodă nu atinge pragul maxim de 100%, dar este mai apropiat de acesta, față de Decision Tree Classifier.

Ca fel și in cazul Decision Tree, pentru imaginile neliniar separabile, acuratețea suferă o scădere, însă, in cazul de fata, aceasta este una considerabile, de aproximativ 38%.

**Comparație intre imagini**

Se poate observa ca Decision Tree si KNN sau descurcat excelent in cazul imaginii neliniar separabile, dar nu si SVM. Fiind cele mai apropiate de original.

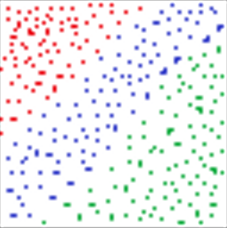
In cazul imagini liniar separabile, SVM s-a descurcat excelent deoarece a fost cea mai apropiată de original.



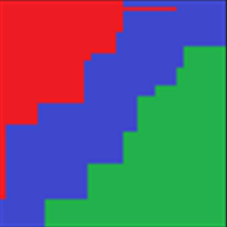
**SVM**



**KNN**



**Original**



**Decision Tree**



**SVM**



**KNN**



**Decision Tree**



**Original**