**Studiu de caz**

**Documentatie**  
  
<https://glenn-viroux.medium.com/creating-a-music-genre-classifier-using-a-convolutional-neural-network-548d06658cee>

<https://cs229.stanford.edu/proj2018/report/21.pdf> = (1)

<https://www.analyticsvidhya.com/blog/2022/03/music-genre-classification-project-using-machine-learning-techniques/> = (2)

<https://www.projectpro.io/article/music-genre-classification-project-python-code/566#mcetoc_1fsm03988d> = (3)

<https://medium.com/@shogulomkurganov73/discover-the-magic-of-music-genre-classification-with-machine-learning-cba2b24febd6> = (4)

<https://scholarworks.calstate.edu/downloads/73666b68n#:~:text=The%20GTZAN%20Dataset%20%2D%20Music%20Genre,pop%2C%20reggae%2C%20and%20rock>. = (5)

<https://indjst.org/articles/comparative-study-of-mfcc-and-mel-spectrogram-for-raga-classification-using-cnn> = (6)

**Cele mai folosite metode pentru clasificare gen muzical:**

* KNN
* Se caută între date cele mai apropiate de cel de interes
* Se determină valoarea K prin trial-and-error, ca să obținem cea mai bună performanță
* Rezultate din (1): acuratețe 54%, din (2): acuratețe 70%, din (4): acuratețe 85%
* Support Vector Machine
* Învățare automată supervizată
* Clasificare lineară, datele trebuie analizate pe bază de o formulă specificată utilizând un kernel anume, ales pentru cazul de față
* Rezultate din (1): acuratețe 60%, din (5): acuratețe 89%, timp de execuție 2,45s
* FF Neural Network
  + - Fully connected neural network, funcție de activare cea mai folosită: ReLu
    - Numărul de straturi se alege a fi nu mare, dar nici prea mic (ex 6)
    - Trebuie folosite metode the regularizare (dropout, L2 etc.) pentru rezultate mai bune
    - Rezultate din (1): acuratețe 54%
    - Exemplu specific: MLPClassifier, rezultate din (4): acuratețe 89%
* Random Forest
  + - Ansamblu de algoritmi
    - Folosește copaci de decizie în faza de antrenare
    - Rezultate din (4): acuratețe 91%, din (5): acuratețe 88%, timp de execuție 8,52s
* CNN
  + - Cel mai des folosit (în ultimii 5-7 ani)
    - Fiecare strat din rețea are substrat de pooling, regularizare, iar în a doua fază sunt câteva straturi de fully connected, urmând straturi de finalizare (softmax output, cross entropy)
    - Rezultate din (1): acuratețe 82%, din (3): acuratețe 88%, din (5): acuratețe 87%, timp de execuție 29s

**Pregătirea datelor**

Dataset: GTZAN

<https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification>

* 10 genuri muzicale, câte 100 de mostre pentru fiecare

Feature extraction

* Spectrograma MEL:
  + - Reprezentarea frecvențelor în domeniul timp
    - În general este folosit un color-map pentru reprezentarea grafică
    - Axa Y este în scala Mel
* Coeficienții MFCC:
  + - Coeficienți care arată caracteristici spectrale
    - Sunt preluate din spectrograma Mel, dar sunt procesate după

Pentru clasificare genuri: conform (1) și (6) => Spectrograma Mel este mai eficientă