Identificarea setului de date:

* Setul de date: <https://www.kaggle.com/datasets/sriramr/fruits-fresh-and-rotten-for-classification/download?datasetVersionNumber=1>
* Este un set de date public de pe Kaggle

Colectarea setului de date:

* Descarcare programatica facuta in notebook. Mai multe detalii in markdown

Intelegerea setului de date + Analiza exploratorie:

* Numar total de imagini:
* CATEGORII MARI: 10901
  + Numar imagini pe categorii de fructe: mere 4035, banane: 3805, portocale: 3061
  + Numar de imagini pe categorii de fresh/rotten: fresh: 4740, stricate: 6161
* CATEGORII GRANULARE: (fresh, stricate)
  + Mere: 1693, 2342
  + Banane: 1581, 2224
  + Portocale: 1466, 1595
* HEAT MAP:
* O imagine care conține text, captură de ecran, număr, Font

  Descriere generată automat
* Exemplu imagine:
* O imagine care conține fruct, Alimente naturale, alimente, legume și fructe

  Descriere generată automat Categorie: Mar fresh
* Dimensiune inainte de redimensionare ~ 450x450 px (dimensiuni variabile)
* Obs:
  + Imaginiile conțin rotiri
  + Condiții de lumină aceleași la fiecare imagine
  + Unele fotografii sunt făcute de sus, unele din lateral

Preprocesarea setului de date:

* Imaginiile sunt mari, și de dimensiuni variabile, deci le vom redimensiona la 32x32
* Pentru o mai buna predictie, chiar daca conditiile de lumina sunt cam la fel, imaginiile tot vor fi normalizate
* Exemplu imagine dupa redimensionare si normalizare
* O imagine care conține captură de ecran, Color, Software multimedia, sferă

  Descriere generată automat
* Featureuriile au fost encodate din clase (mere,banane,portocale) si (fresh,rotten) in 2 featurears.
* Encode-ul este unul de tip one-hot
* [0,1,2] -> tip de fruct
* [0,1] -> fresh/rotten
* Asa ca o imagine de tipul mar/fresh va avea
* Pereche([1,0,0], [1,0])
* Pentru viitorul model dezvoltat vor fi nevoie de fie 2 modele, fie de unul care are shared knowledge, deoarece theoretic avem 2 features pe care trebuie sa le prezicem (tipul de fresh, si daca acesta e fresh/stricat)

Link catre codul sursa: <https://github.com/Catalin-web/MIASC>