

# Consum Caloric

Implementați un model de AI/ML pentru a prezice **consumul caloric** al unei activități. Se utilizează setul de antrenare `train_data.csv` și setul de test `test_data.csv`, pentru care trebuie să realizați predicția.

## Setul de date

Setul de date conține următoarele coloane:

- **User\_ID**: ID-ul utilizatorului care a realizat activitatea
- **Gender**: genul utilizatorului
- **Age**: vârsta utilizatorului
- **Height**: înălțimea în cm
- **Weight**: greutatea în kg
- **Duration**: durata activității
- **Heart\_Rate**: ritmul cardiac mediu
- **Body\_Temp**: temperatura medie
- **Calories**: numărul de calorii consumate (prezent doar în setul de antrenare)

**Notă** Atributul-țintă este **Calories**. Metrica de evaluare: MAE.

## Cerințe

### Partea 1 (20p)

Bazat pe setul de date de antrenare calculați:

- **Samples**: numărul de linii din setul de date de antrenare
- **No.Males**: numărul de exemple pentru bărbați
- **AverageDuration**: durata medie a activităților
- **SeniorUsers**: numărul de utilizatori cu vârsta  $\geq 75$  ani

Fiecare subtask valorează 5 puncte.

## Partea 2 (80p)

- **Subtask 5 (60p)**: Dezvoltați un model AI/ML și preziceți **Calories** pentru fiecare exemplu din test cu coloana **Subtask** = 5.
- **Subtask 6 (20p)**: Estimarea consumului caloric pentru o echipă de handbal masculin, pentru fiecare exemplu din test cu coloana **Subtask** = 6.

## Format de ieșire

Fișier `output.csv` cu coloanele:

- **subtaskID** – numărul subtaskului (1, 2, 3, 4, 5 sau 6)
- **datapointID** – corespunzător **User\_ID** din `test_data.csv`
- **answer** – răspunsul pentru datapoint-ul respectiv și subtask

**Notă:** Pentru subtask-urile 1-4 (un singur răspuns pentru întreg setul) utilizați o singură linie cu **datapointID** = 1. Trimiteți un singur fișier CSV pentru toate subtask-urile. Consultați `sample_output.csv` pentru exemplu.