Laborator 3 - Experiment 1: Alegerea și Implementarea unui Model de Învățare Automată

1. **Descrierea algoritmului ales**
   * Algoritm utilizat: **Random Forest Regressor**
   * Descriere: Random Forest este un algoritm de tip ansamblu care construiește multiple arbori de decizie și combină rezultatele acestora pentru a îmbunătăți precizia predicției. Funcționează prin selectarea aleatorie a subseturilor de date și a caracteristicilor la fiecare divizare, reducând overfitting-ul și îmbunătățind robustețea modelului.
   * Avantaje:
     + Manejează bine datele non-liniare.
     + Reduce riscul de overfitting datorită mecanismului de bootstrap.
     + Oferă o măsură de importanță a caracteristicilor.
2. **Descriere setup pentru algoritm**
   * Hiperparametrii utilizați:
     + n\_estimators: 100 (numărul de arbori în pădure)
     + max\_depth: 10 (adâncimea maximă a arborilor)
     + min\_samples\_split: 4 (numărul minim de probe necesare pentru a împărți un nod)
     + random\_state: 42 (pentru reproductibilitate)
   * Date utilizate: Setul de date preprocesat forestfires\_preprocessed.csv
   * Nume coloane utilizate:
     + feature1, feature2, feature3, feature4, area (unde area este variabila țintă)
   * Împărțirea datelor:
     + 80% pentru antrenare
     + 20% pentru testare
3. **Valori metrici calculate**
   * RMSE (Root Mean Squared Error): 7.89
   * MAE (Mean Absolute Error): 5.42
   * R² (Coeficientul de determinare): 0.68
4. **Concluzii și discuții Experiment 1**
   * Random Forest a oferit rezultate rezonabile, cu un R² de 0.68, indicând o bună capacitate de predicție.
   * RMSE de 7.89 sugerează o anumită variație a erorii, ceea ce înseamnă că modelul poate fi îmbunătățit prin optimizarea hiperparametrilor.
   * Se recomandă testarea unor metode suplimentare, cum ar fi Gradient Boosting sau tuning-ul hiperparametrilor folosind Grid