Divizarea train - test

* Necesitatea divizării datelor în setul de training și setul de test
* Importul și rolul funcției train\_test\_split
* Algoritmul clasic de elaborarea a unui model
* Prelucrarea și ajustarea datelor în setul X și y
* Divizarea seturilor X și y în seturi train și seturi test
* Crearea și ajustarea unui obiecte de scalarea pe X\_train
* Scalarea datelor X\_train
* Scalarea datelor X\_test
* Crearea modelului
* Ajustarea modelului pe datele X\_train, y\_train
* Realizarea predicție (y\_pred) pe datele X\_test
* Evaluarea modelului comparând y\_pred cu y\_test
* Ajustarea hiperparametrilor modelului si reperarea ultimilor pasi

Divizarea train - valid - test

* Necesitatea setului de validare
* Algoritmul de elaborarea a unui model cu date de validare
* Prelucrarea și ajustarea datelor în setul X și y
* Divizarea seturilor X și y în seturi train, seturi valid și seturi test
* Crearea și ajustarea unui obiecte de scalarea pe X\_train
* Scalarea datelor X\_train
* Scalarea datelor X\_test
* Scalarea datelor X\_valid
* Crearea modelului
* Ajustarea modelului pe datele X\_train, y\_train
* Realizarea predicție (y\_pred) pe datele X\_valid
* Evaluarea modelului comparând y\_pred cu y\_valid
* Ajustarea hiperparametrilor modelului si reperarea ultimilor pasi
* Realizarea predicție finale (y\_pred\_final) pe datele X\_test
* Evaluarea finală a modelului comparând y\_pred\_final cu y\_test

Cross validation – functia cross\_val\_score

* Esența procedurii de cross validation
* Algoritmul de elaborarea a unui model utilizând cross validation cu funcția cross\_val\_score
* Prelucrarea și ajustarea datelor în setul X și y
* Divizarea seturilor X și y în seturi train și seturi test
* Crearea și ajustarea unui obiecte de scalarea pe X\_train
* Scalarea datelor X\_train
* Scalarea datelor X\_test
* Crearea modelului
* Realizarea procedurii cross validation si determinarea erorii medii
* Ajustarea hiperparametrilor modelului si reperarea ultimilor pasi
* Ajustarea modelului pe datele X\_train, y\_train
* Realizarea predicție (y\_pred) pe datele X\_test
* Evaluarea finala a modelului comparând y\_pred cu y\_test

Cross validation – functia cross\_validate

* Avantajele funcției cros\_validate
* Algoritmul de elaborarea a unui model utilizând cross validation cu funcția cross\_val\_score
* Prelucrarea și ajustarea datelor în setul X și y
* Divizarea seturilor X și y în seturi train și seturi test
* Crearea și ajustarea unui obiecte de scalarea pe X\_train
* Scalarea datelor X\_train
* Scalarea datelor X\_test
* Crearea modelului
* Realizarea procedurii cross validation si afisarea diferitor parametri ai acesteia
* Determinarea valorilor medii ale parametrilor cross validation
* Ajustarea hiperparametrilor modelului si reperarea ultimilor pasi
* Ajustarea modelului pe datele X\_train, y\_train
* Realizarea predicție (y\_pred) pe datele X\_test
* Evaluarea finala a modelului comparând y\_pred cu y\_test

Grid Search

* Esența procedurii Grid Search
* Algoritmul de ajustarea a hiperparametrilor unui model:
* Prelucrarea și ajustarea datelor în setul X și y
* Divizarea seturilor X și y în seturi train și seturi test
* Crearea și ajustarea unui obiecte de scalarea pe X\_train
* Scalarea datelor X\_train
* Crearea modelului de baza cu parametri impliciți
* Crearea unui dicționar cu parametrii modelului si valorile posibile ale acestora
* Crearea modelului cu Grid Search
* Ajustarea modelului pe datele X\_train, y\_train
* Vizualizarea celor mai buni parametri si erori
* Realizarea predicție (y\_pred) pe datele X\_test
* Evaluarea finala a modelului comparând y\_pred cu y\_test