Logistic regression – 2 clase

* Particularitățile sarcinilor de clasificare
* Funcția sigmoid
* Relația functie sigmoid intre valoarea prezisa si coeficienții caracteristicilor
* Noțiune de cotă (odds)
* Logaritmul cotei ca relație dintre valoarea prezisa si coeficienții caracteristicilor
* Semnul coeficientului vs variația caracteristicii
* Compararea coeficienților vs importanța caracteristicilor
* Relația dintre logaritmul cotei si probabilitate
* Algoritmul de analiza prin explorarea datelor (EDA) în Scikit-Learn:
* vizualizarea descrierii statistice a datelor din df (describe())
* vizualizarea numărului de valori in fiecare clasa (value\_counts())
* vizualizarea grafica a numărului de valori in fiecare clasa (countplot())
* vizualizarea distribuției valorilor clasei in functie de categorii (boxplot())
* vizualizarea dependenței dintre caracteristici in functie de clase (scatterplot())
* vizualizarea distribuției caracteristicilor in functie de clase (pairplot())
* vizualizarea corelației dintre coloane (heatmap())
* Algoritmul de elaborare a modelului logistic regression în Scikit-Learn:
* crearea setului X si y
* crearea seturilor de train si de test
* scalarea datelor
* importul si crearea modelului logistic regression
* ajustarea modelului pe datele de training
* vizualizarea coeficienților
* realizarea predicției
* realizarea predicției cu afișarea probabilităților
* Rezultate pozitive false și rezultate negative false
* Confusion matrix
* Determinarea acurateței
* Paradoxul acurateței
* Determinarea recall (sensibilității)
* Determinarea precison (preciziei)
* Determinarea f1-score
* Curba ROC si nivelul de prag
* Noțiunea de AUC - aria under curve
* Algoritmul de evaluarea a modelului logistic regression în Scikit-Learn:
* determinarea acurateti modelului (accuracy\_score())
* determinarea confusion matrix a modelului (confusion\_matrix())
* vizualizarea grafica a confusion matrix a modelului (plot\_confusion\_matrix())
* elaborarea classificasion report a modelului (classification\_report())
* determinarea precision a modelului (precision\_score())
* determinarea recall a modelului (recall\_score())
* determinarea f1-score a modelului (f1\_score())
* vizualizarea ROC (plot\_roc\_curve ())
* vizualizarea curbei precision-recall (plot\_precision\_recall\_curve())

Logistic regression – clase multiple

* Algoritmul de elaborarea a modelului logistic regression multi-class în Scikit-Learn
* vizualizarea informatie despre date (info())
* vizualizarea descrierii statistice a datelor (describe())
* vizualizarea numarului de clase din coloana label (value\_counts())
* vizualiazarea grafica a numarul de valori in fiecare clasa a coloanei eticheta (countplot)
* vizualizarea grupării claselor in funcție de 2 caracteristici (scatterplot)
* vizualizarea corelației dintre coloane (heatmap)
* crearea setului X si y
* crearea seturilor de train si de test
* Scalarea datelor
* Importul si crearea modelului logistic regression de baza
* crearea unor liste cu valorile hiperparametrilor ce vor fi variati
* crearea unui dicționar cu hiperparametrii ce vor fi modificati
* crearea modelului grid
* ajustarea modelului pe datele de training
* verificare celor mai buni hiperparametri
* realizarea predicitei pe datele de test
* determinarea acurateti modelului (accuracy\_score())
* determinarea confusion matrix a modelului (confusion\_matrix())
* vizualizarea grafica a confusion matrix a modelului (plot\_confusion\_matrix())
* elaborarea classificasion report a modelului (classification\_report())