

UNIVERSITATEA TITU MAIORESCU

FACULTATEA DE INFORMATICĂ

**MASTER ȘTIINȚA DATELOR ȘI INTELIGENȚĂ ARTIFICIALĂ
FUNDAMENTELE ȘTIINȚEI DATELOR**

**CURSUL 6
EVALUAREA ȘI VALIDAREA MODELELOR**

**Autor:
Prof.univ.dr.ing.Titi PARASCHIV**

**București
2024**

IMPORTANT A EVALUĂRII ȘI VALIDĂRII

Evaluarea și validarea modelelor sunt esențiale pentru a asigura acuratețea, precizia și fiabilitatea modelelor în aplicarea lor pe date noi.

Rol: Previn supraantrenarea și asigură că modelele generalizează bine la date neutilizate anterior.

METODE ȘI TEHNICI DE EVALUARE

- Metrice Supervizate:
Acuratețea, precizia, recall, F1-score, ROC curve.

- Validare încrucișată (cross-validation): Împărțirea setului de date în multiple subseturi pentru antrenare și testare.

- Evaluare cu set de testare independent: Modelul este antrenat pe un set și testat pe un set separat.

METRICI DE EVALUARE SUPERVIZATE

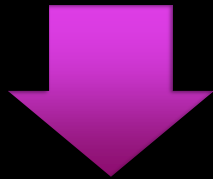
- Acuratețe (Accuracy): Proporția corectă a clasificărilor făcute de model.

- Precizie (Precision): Proporția de exemple pozitive corect identificate.

- Recall: Proporția de exemple pozitive identificate corect.

- F1-Score: Media armonică între precizie și recall.


Validare încrucișată



(cross-
validation)

- • Definiție: Împărțirea setului de date în k subseturi pentru antrenament și testare repetate.
- • Scop: Estimarea performanței modelului pe date noi și prevenirea supraantrenării.

EVALUAREA CU SET DE TESTARE INDEPENDENT

- ▶ • Definiție: Separarea unui set de date pentru testare independentă.
 - ▶ • Scop: Estimarea capacității modelului de a generaliza pe date necunoscute.
- 
- A series of several parallel white diagonal lines in the bottom right corner of the slide, slanting upwards from left to right.

VALIDARE A CU SET DE VALIDARE

- Definiție: Folosirea unui set separat pentru ajustarea parametrilor modelului.

- Scop: Optimizarea modelului înainte de testarea pe setul final.

METRICI DE EVALUARE NESUPERVIZATE

- Exemple: Silhouette score, Coerența, Rand Index, Dunn Index, Inertie.



- Utilizare: Evaluarea clustering-ului și a calității grupărilor.

COMPARA REA CU MODELELE DE REFERINȚĂ

- Definiție: Compararea performanței modelului cu modele de referință sau standarde acceptate.

- Scop: Evaluarea calității și performanței relative a modelului.

VALIDARE PRIN CONFIGURA REA HIPERPARAM ETRILOR

- Definiție: Ajustarea parametrilor modelului pentru a maximiza performanța.

- Tehnici: Grid search, Random search, evaluarea pe setul de validare.

EVALUAR EA ÎN TIMP REAL

- Definiție: Testarea și evaluarea modelelor pe date generate sau colectate în timp real.

- Scop: Adaptarea modelului la schimbările din date și context.

INTERPRETAREA REZULTATELOR

- Definiere: Analiza performanței modelului și implicațiile pentru aplicația specifică.

- Tehnici: Analiza curbei de învățare, matricea de confuzie, analiza importanței caracteristicilor.

AJUSTAREA MODELULUI

- Definiere: Optimizarea performanței prin modificarea parametrilor sau caracteristicilor.

- Tehnici: Regularizare, ajustarea hiperparametrilor, ajustarea arhitecturii.

APLICAȚII PRACTICE

- Domenii: Cercetare medicală, marketing, finanțe, tehnologie, cercetare științifică.

- Exemplu: Evaluarea performanței unui model de machine learning pentru prezicerea cotelor de piață.

BIAS ȘI ERORI DE SISTEM

- Bias: Tendința modelului de a produce rezultate distorsionate.

- Erori de Sistem: Erori tehnice sau de modelare care afectează precizia și fiabilitatea.

TEHNICI DE PREVENIRE A BIASURILOR ȘI ERORILOR

- Date reprezentative:
Asigurarea că setul de date
reflectă distribuția reală.

- Validare: Utilizarea seturilor
de date independente pentru
identificarea biasurilor.

- Modele multiple:
Combinarea rezultatelor mai
multor modele pentru a
minimiza biasul.