



FIȘA DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea “Alexandru Ioan Cuza” din Iași
1.2 Facultatea	Facultatea de Informatică
1.3 Departamentul	Departamentul de Informatică
1.4 Domeniul de studii	Informatică
1.5 Ciclu de studii	Licență
1.6 Programul de studii / Calificarea	Informatică/Licențiat în Informatică

2. Date despre disciplină

2.1 Denumirea disciplinei	Invatare automata						
2.2 Titularul activităților de curs	CONF. DR. LIVIU CIORTUZ						
2.3 Titularul activităților de seminar	Conf. dr. L. Ciortuz, Conf. dr. A. Ignat						
2.4 An de studiu	III	2.5 Semestru	1	2.6 Tip de evaluare	E	2.7 Regimul disciplinei*	OB

* OB – Obligatoriu / OP – Opțional

3. Timpul total estimat (ore pe semestru și activități didactice)

3.1 Număr de ore pe săptămână	4	din care: 3.2 curs	2	3.3 seminar/laborator	2
3.4 Total ore din planul de învățământ	56	din care: 3.5 curs	28	3.6 seminar/laborator	28
Distribuția fondului de timp					ore
Studiu după manual, suport de curs, bibliografie și altele					28
Documentare suplimentară în bibliotecă, pe platformele electronice de specialitate și pe teren					14
Pregătire seminarii/laboratoare, teme, referate, portofolii și eseuri					14
Tutoriat					-
Examinări					4
Alte activități					-
3.7 Total ore studiu individual					56
3.8 Total ore pe semestru					116
3.9 Număr de credite					6

4. Precondiții (dacă este cazul)

4.1 De curriculum	Cunoștințe de bază din matematica de liceu (studiul funcțiilor, geometrie elementară), precum și analiză matematică, probabilități și algoritmică din facultate.
4.2 De competențe	

5. Condiții (dacă este cazul)

5.1 De desfășurare a cursului	Prezența la curs este recomandată călduros.
5.2 De desfășurare a seminarului/ laboratorului	Prezența la seminar este obligatorie.

**6. Competențe specifice acumulate**

Competențe specifice acumulate	C1. Capacitatea de a înțelege, a implementa și utiliza algoritmi de bază din domeniul învățării automate.
Competențe transversale	CT1. Abilitatea de utilizare a instrumentelor matematice (analiză matematică, probabilități și statistică) pentru proiectarea și analiza algoritmilor de bază din învățarea automată. CT2. Capacitatea de a lucra în paradigma „data-driven programming”: explorarea datelor, crearea automată de modele, evaluarea modelelor.

7. Obiectivele disciplinei (din grila competențelor specifice acumulate)

Obiectiv general 7.1	Însușirea conceptelor de bază din domeniul învățării automate; înțelegerea și folosirea unor algoritmi de bază din acest domeniu.
Obiective specifice 7.2	O1. Înțelegerea algoritmilor de bază pentru crearea arborilor de decizie (ID3 și AdaBoost), clasificare bayesiană (Bayes Naiv și Bayes Comun), învățarea bazată pe memorare (k-NN) și clusterizare (clust. ierarhică, K-means și regresia logistică). Aplicarea acestor algoritmi pe seturi relativ mici de date.

8. Conținut

8.1		Metode de predare	Observații (nr. de ore)
1.	Noțiuni introductive de Învățare Automată. Tipuri de algoritmi de învățare automată: clasificare, clusterizare, regresie, ranking, selecția trăsăturilor, sisteme de recomandare.	Expunere	2
2.	Recapitularea noțiunilor fundamentale de probabilități și statistică: evenimente aleatoare, funcție de probabilitate, independența evenimentelor aleatoare; variabile aleatoare (discrete și continue), independența variabilelor aleatoare. Medii și varianțe. Distribuții probabile uzuale. Proprietăți de bază.	Expunere	2
3.	Metode de estimare a parametrilor unor distribuții probabile: MLE și MAP. Aplicare la distribuția Bernoulli și distribuția gaussiană (cazul uni-variabil).	Expunere	2
4.	Elemente de bază din teoria informației: entropie, entropie condițională,	Expunere	2



	câștig de informație, entropie relativă, cross-entropie. Proprietăți de bază. Arbori de decizie. Algoritmul ID3.		
5.	Variante / extensii ale algoritmului ID3. Evaluarea algoritmilor de clasificare: acuratețe, eroarea la antrenare, eroarea la [cross-]validare, precizie, recall etc. Overfitting. Strategii de pruning pentru arbori de decizie.	Expunere	2
6.	Tipuri de ipoteze în învățarea automată: ipoteze ML și ipoteze MAP. Clasificare bayesiană. Algoritmii Bayes Naiv și Bayes Comun/Optimal.	Expunere	2
7.	Calculul ratei medii a erorii pentru clasificatori de tip bayesian. Natura separatorului decizional în clasificarea bayesiană. (Legătura dintre clasificarea bayesiană și regresia logistică.) Clasificare bayesiană gaussiană.	Expunere	2
8.	Recapitulare.	Examen parțial	2
9.	Învățare bazată pe memorare. Algoritmul k-NN. Relația dintre rata erorii asimptotice a algoritmului 1-NN și rata erorii algoritmului Bayes Comun.	Expunere	2
10.	Metode ansambliste bazate pe arbori de decizie: Algoritmul AdaBoost.	Expunere	2
11.	Învățare supervizată (clasificare) vs. învățare nesupervizată (clusterizare). Tipuri de algoritmi de clusterizare. Clusterizare ierarhică aglomerativă și Clusterizare ierarhică divizivă. Tipuri de funcții de similaritate (single-link, complete-link, average-linkage).	Expunere	2
12.	Clusterizare partițională (i.e., clusterizare neierarhică cu asignare „hard” a instanțelor la clustere): algoritmul K-means.	Expunere	2
13.	Monotonia criteriului de „distorsiune” (J) la aplicarea algoritmului K-means.	Expunere	2
14.	Regresia Logistică.	Expunere	2
15.	Recapitulare.	Examen final	2

Bibliografie**Referințe principale:**

- Tom Mitchell. Machine Learning. McGraw Hill. 1997.
- Liviu Ciortuz, Alina Munteanu, Elena Bădăraș. Exerciții de învățare automată. Ediția a III-a revizuită (online), Iași, varianta 2023f (www.info.uaic.ro/~ciortuz/ML.ex-book/editia-2023f/ex-book.20sept2023)
- Christopher Manning, Heinrich Schütze, Foundations of Statistical Natural Language Processing, MIT Press, 2000.

Referințe suplimentare:

Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning. Springer, 2009.

8.2	Seminar	Metode de predare	
1.	Sisteme/programe de învățare automată: câteva studii de caz.	Demo-uri	2
2.	Exerciții recapitulative de probabilități și statistică.	Lucru la tabla, lucru individual, metode interactive, programe	2



		demonstrative, teste.	
3.	Estimarea parametrilor distribuțiilor probabile în sens MLE: exerciții.	Idem	2
4.	Aplicarea algoritmului ID3 folosind attribute numerice continue: exerciții. Aplicarea diverselor strategii posibile de pruning pe arbori de decizie: exerciții.	Idem	2
5.	Exerciții pentru calcularea de entropii, câștiguri de informație, demonstrarea unor proprietăți. Exerciții cu arbori de decizii. Aplicarea algoritmului ID3. Calculul erorilor la antrenare și validare încrucișată cu “metoda leave-one-out”.	Idem	2
6.	Aplicarea algoritmilor Bayes Naiv și Bayes Comun: exerciții.	Idem	2
7.	Exerciții de calcul pentru rata medie a erorilor în clasificarea bayesiană. Legătura dintre algoritmul Bayes Naiv și Regresia Logistică. Clasificare bayesiană gaussiană: exerciții.	Idem	2
8.	Examen parțial 1.	-	-
9.	Aplicarea algoritmului k-NN: exerciții.	Idem	2
10.	Aplicarea algoritmului AdaBoost: exerciții.	Idem	2
11.	Aplicarea algoritmilor de clusterizare ierarhică aglomerativă și divizivă: exerciții.	Idem	2
12.	Aplicarea algoritmului K-means: exerciții.	Idem	2
13.	Proprietăți ale algoritmului K-means: monotonie, convergență etc.	Idem	2
14.	Exerciții de aplicare a regresiei logistice.	Idem	2
15-16	Examen parțial 2.	-	-

Bibliografie:

Liviu Ciortuz, Alina Munteanu, Elena Bădăraș. Exerciții de învățare automată. Ediția a III-a revizuită (online), Iași, varianta 2023f (www.info.uaic.ro/~ciortuz/ML.ex-book/editia-2023f/ex-book.20sept2023.pdf).

9. Coroborarea conținutului disciplinei cu așteptările reprezentanților comunității, asociațiilor profesionale și angajatorilor reprezentativi din domeniul aferent programului

Conținutul disciplinei este coroborat cu tematica concursurilor pentru ocuparea posturilor în companiile IT.



Competențe minimale: Studentul va ști la finalul acestui semestru să aplice o serie de algoritmi de bază din învățarea automată: ID3, AdaBoost, Bayes Naiv, k-NN, algoritmul de clusterizare ierarhică bottom-up, k-Means, regresia logistică.

10. Evaluare

Tip activitate	10.1 Criterii de evaluare	10.2 Metode de evaluare	10.3 Pondere în nota finală (%)
10.4 Curs	- aplicarea corectă a algoritmilor de învățare automată predați, în funcție de diverse cerințe specifice; - calitatea formulării răspunsurilor; - estimarea parametrilor unor distribuții probabiliste; - demonstrarea unor proprietăți teoretice legate de chestiunile de mai sus.	2 examene scrise (sapt. 8 și 15-17).	40% (20 puncte).
10.5 Seminar	- estimation of the parameters of some probabilistic distribuții; - proving certain theoretical properties.; - aplicarea concretă a algoritmilor ID3, AdaBoost, Bayes Naiv/Comun, k-NN, clusterizare ierarhică, K-means și regresia logistică pe dataset-uri didactice; - estimarea parametrilor unor distribuții probabiliste; - demonstrarea unor proprietăți teoretice legate de chestiunile de mai sus; - elaborarea corectă și eficientă a unor implementări cu caracter demonstrativ.	2 teste și o prezentare / tema (plus participare activă) la seminar.	40% (2 x 6 + 8 = 20 puncte).
10.6 Standarde minime de performanță			
Pentru promovare trebuie îndeplinite simultan următoarele criterii: • Minim 1.5 (din 6) puncte la fiecare din cele două teste de la seminar (T1 și T2); • Minim 2 (din 8) puncte la tema / prezentarea (plus participarea activă) de la seminar (S); • Minim 2.5 (din 10) puncte la fiecare cele două examene parțiale (P1 și P2); • Nota (vezi formula de mai jos): minim 4,5 (după rotunjire: 5). • Absențele la seminar se penalizează cu câte 0,2 puncte începând de la a doua absență încolo. (Depunctarea se aplică la final de semestru.) Nota: $1/5 \times (10 + \text{punctajele de la seminar} + \text{punctajele de la cele două examene parțiale} - \text{depunctarea pentru absente}).$			

11. Activități online: ne vom conforma deciziilor conducerii Facultății de Informatică, UAIC-Iași.

Observație: Informații mai detaliate despre desfășurarea acestui curs se găsesc în fișierele <http://profs.info.uaic.ro/~ciortuz/SLIDES/ml0.pdf> (vedeți secțiunea Administrativă) și <http://profs.info.uaic.ro/~ciortuz/what-you-should-know.pdf>.



Data completării
25 septembrie 2023

Titular de curs
Conf. dr. Liviu Ciortuz

Titular de seminar
Conf. dr. Liviu Ciortuz

Data avizării în departament

Director de departament