Notițe Seminar 1

Intro: Mai întâi, ce înseamnă Machine Learning/Învățare automată? Sau , mai bine zis, ce vom face mai exact în acest semestru?

Tipuri de învățare automată

1. **supervizată**
   - **clasificare**

Exemplu: date despre apartamente. Vrem să prezicem dacă o casă este locuibilă.

număr m^2	număr camere	este locuibilă?
10	1	da
100	2	nu
...	...	...

Pentru un nou apartament, care este prețul?

număr m^2	număr camere	este locuibilă?
500	10	???

   - **regresie**

Exemplu: date despre apartamente. Vrem să prezicem prețul unei case.

număr m^2	număr camere	preț
10	1	100.522
100	2	200
...	...	...

Pentru un nou apartament, care este prețul?

număr m^2	număr camere	preț
500	10	???

Ce diferă între clasificare și regresie?

2. **nesupervizată**
   - **clusterizare**

Exemplu: date despre apartamente. Vrem să împărțim casele în grupuri: de exemplu, case locuibile vs case nelocuibile.

număr m^2	număr camere
10	1
100	2
...	...

De ce se cheamă supervizată/nesupervizată?

3. **cu întărire/*reinforcement***: doar la cursul de *Rețele neuronale*

Vom începe cu o recapitulare a noțiunilor de **PS (Probabilități și statistică)** – primele 2-3 seminarii.

De ce? Pentru că unii (nu toți) algoritmi de ML folosesc PS.

Experiment aleator = acțiune/procedură în urma căreia obținem un rezultat din mai multe rezultate posibile

Exemple: aruncarea unui ban, aruncarea 2 monede etc.

Cum formalizăm un experiment aleator cu ajutorul probabilităților?

(___, ___, ___) – **spațiu de probabilitate discret** (nu _________)

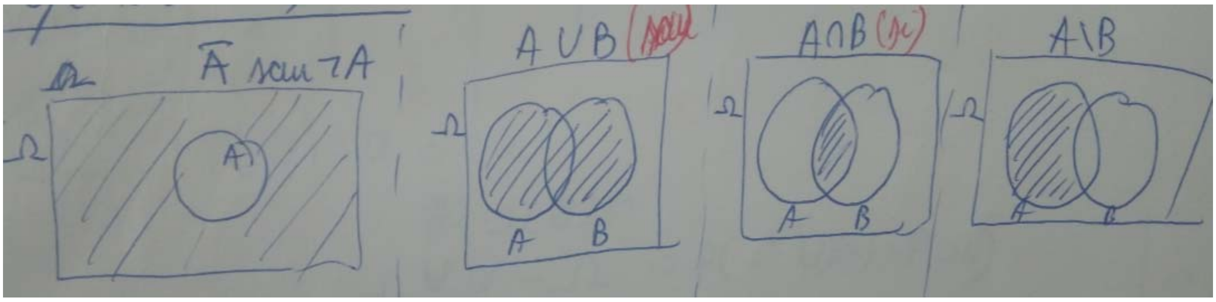
- ___ - **spațiu de eșantionare** (*sample space*) = o mulțime ce conține ca elemente rezultatele posibile ale experimentului aleator
  Intuție: discret vs continuu?
    Exemplul 1: {1,3,9} sau [1,9]
    Exemplul 2: mulțimea numerelor naturale sau mulțimea numerelor reale
- ___ - **spațiu de evenimente**: _____

  ______ A - eveniment aleator
    - De obicei, notat cu litere mari de la începutul alfabetului
    - Se produce/se realizează dacă, în urma experimentului aleator, rezultatul aparține lui A
      Exemplu:
- ___ - **măsură/funcție de probabilitate**
    - Proprietatea de **aditivitate numărabilă**:

    Vizual:

    Exemplu:

Observații: Operații cu mulțimi



Exercițiul 1:

$\Omega = \{a,b,c\}$

$F = 2^{\Omega} =$ ___________________________

Fie următoarele funcții: $f_1, f_2, f_3, f_4 : \mathcal{F} \to \mathbb{R}$ definite parțial. Există deja inconsistențe/greșeli în (parțial) definirea funcțiilor ca funcții de probabilitate? Dacă da, care sunt inconsistențele? Dacă nu, completați celelalte căsuțe asociate funcției folosind cele 2 axiome din definiția funcției de probabilitate?

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
$\emptyset$		1/2		
{a}	1/3			1.5
{b}	1/3			
{c}	1/3		2/3	
{a,b}			1/3	
{a,c}			1/3	
{b,c}				
{a,b,c}	1			

Observație: P($\emptyset$) = 0. De ce?

Exercițiul 2:

$\Omega = \{a_1, \ldots, a_n\}$
$P(\{a_1\}) = \cdots = P(\{a_n\})$ - rezultate echiprobabile
$A = \{a_{i_1}, \ldots, a_{i_k}\}$
Calculați $P(A)$.

Observație importantă: o puteți considera prima presupunere pe care o puteți face când rezolvați o problemă unde trebuie să calculați o probabilitate, dar nu există informații suplimentare.

Exercițiul 3:

a) Care este probabilitatea ca la aruncarea unui zar să iasă un număr par?

b) Care este probabilitatea ca la aruncarea unui zar să iasă un număr par, știind că

P(iese un număr impar) = 1/6?

Exercițiul 4: Într-o grupă de copii de la o creșă, 30% dintre ei au ochi căprui, 50% au ochi albaștri, iar restul de 20% au ochi de alte culori. Care este probabilitatea ca un copil ales în mod aleatoriu din această grupă să aibă ochi albaștri?

Observație: „rezultatele experimentului sunt **echiprobabile**” = „în mod aleatoriu **uniform**”

Remember (nu trebuie să știți să le demonstrați):

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$
$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

Proprietăți (trebuie să știți să le **demonstrați** folosind proprietatea de aditivitate numărabilă; în exerciții, dacă nu se menționează că trebuie demonstrate, le puteți folosi fără a le demonstra)

1. $P(\emptyset) = 0$
2. $P(\overline{A}) = 1 - P(A)$
3. $A \subseteq B \Rightarrow P(A) \leq P(B)$
4. $P(A \setminus B) = P(A) - P(A \cap B)$
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Idei de demonstrație	
	1.
	2.
	3.
	4.
	5

Exercițiu:

61. (Calcul de probabilităţi elementare)

CMU, 2004 fall, T. Mitchell, Z. Bar-Joseph, HW1, pr. 1.3

Doi soldaţi A şi B trag la ţintă. Probabilitatea ca soldatul A să greşească ţinta este de $1/5$. Probabilitatea ca soldatul B să greşească ţinta este de $1/2$. Probabilitatea ca ambii soldaţi să greşească simultan ţinta este de $1/10$.

a. Care este probabilitatea ca cel puţin unul din soldaţi să greşească ţinta?

b. Care este probabilitatea ca exact unul din cei doi soldaţi să greşească ţinta?

Probabilități condiționate

Notație, Intuiție:

P(A|B) = probabilitatea să se realizeze A, știind că s-a realizat B

= probabilitatea să se realizeze A, dacă că s-a realizat B

= probabilitatea să se realizeze A, când spațiul de eșantionare devine B

Este adevărat că P(A) = P(A| Ω)?

Definiție: $$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

De ce?

Observații:

- P(B) $\neq$ 0
- $P(A|B) \neq P(A \cap B)$ în general

Alte 3 formule utile:

Regula de multiplicare: $P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$

Generalizare: **regula lanțului**: $$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1, A_2) \ldots P(A_n \mid A_1, A_2, \ldots, A_{n-1})$$

Formula lui Bayes: $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

P(A|B) – probabilitate **a posteriori**

P(A) – probabilitate **a priori**

Formula probabilității totale: $P(A) = P(A \mid B)P(B) + P(A \mid \neg B)P(\neg B)$

Generalizare: $\textbf{if } A \subseteq \cup B_i \textbf{ and } \forall i \neq j\ B_i \cap B_j = \emptyset\textbf{, then}$
$P(A) = \sum_i P(A \mid B_i)P(B_i)$

Idei de demonstrație

Exercițiu: Avem două urne. Prima urnă conține 11 bile albe și 4 bile roșii. Cea de-a doua urnă conține 8 bile albe și 5 bile roșii. Se alege în mod aleatoriu cu probabilitate uniformă una din cele două urne. Apoi se extrage o bilă din urna aleasă.

a) Care este probabilitatea ca bila extrasă să fie albă?

b) Dacă bila extrasă este albă, are este probabilitatea ca ea să provină din prima urnă?

Evenimente independente

Intuiție: când producerea unui eveniment nu modifică probabilitatea de realizare a celuilalt.

Definiția 1 (slabă, dar intuitivă): A, B – independente, P(B) $\neq$ 0: P(A|B) = P(A)

Definiția 2 (tare; la aceasta ne vom referi de acum încolo; include și cazul în care P(B) = 0):

$$P(A \cap B) = P(A)P(B)$$

Observație importantă: Puteți considera independența evenimentelor a doua presupunere pe care o puteți face când rezolvați o problemă unde trebuie să calculați o probabilitate, dar nu există informații suplimentare. (un algoritm de ML face acest lucru...)

Exercițiul 1: Probabilitatea ca studentul X să obțină 6 puncte la un test este de 0.5. Probabilitatea ca studentul Y să obțină 6 puncte la un test este de 0.2. Care este probabilitatea ca și X, și Y să obțină 6 puncte?

Exercițiul 2: Probabilitatea ca studentul X să obțină 6 puncte la un test este de 0.5. Probabilitatea ca studentul Y să obțină 6 puncte la un test este de 0.2. Probabilitatea ca și X, și Y să obțină 6 puncte este de 0.11. Evenimentele „studentul X obține 6 puncte”, „studentul Y obține 6 puncte” sunt independente?

Observație: În general, dacă avem o definiție/formulă, există și **varianta ei condițională** (P(C) $\neq$ 0):

1. A, B – evenimente independente condițional față de C:

   $$P(A \cap B|C) = P(A|C)P(B|C)$$

2. $P(\emptyset) = 0 \to P(\emptyset|C) = 0$
3. $P(\overline{A}) = 1 - P(A) \to$ ....
4. $A \subseteq B \Rightarrow P(A) \leq P(B) \to$ ....
5. $P(A \cap B) = P(A|B)P(B) \to$ ....

Observații:

1. $P(\overline{A}|B) = 1 - P(A|B)$.
   În general, $P(A|\overline{B})$ nu se poate scrie în funcție de $P(A|B)$.
2. $P(A) = 0 \not\Rightarrow A = \emptyset$
3. $P(A) = 1 \not\Rightarrow A = \Omega$

Schemă de final

- Experiment aleator
- Spațiu de probabilitate
  - Ω: spațiu de eșantionare
  - F: spațiu de evenimente
  - P: funcție de probabilitate
    - P(Ω) = 1
    - Aditivitatea numărabilă
- Probabilități condiționate
  - Definiție
  - Formula de multiplicare + generalizare
  - Formula lui Bayes
  - Formula probabilității totale + generalizare
- Evenimente independente
- Alte formule cu probabilități
- Formule cu probabilități în varianta condițională
- Cele 2 presupuneri
  - Rezultate echiprobabile
  - Evenimente independente