Inteligență Artificială

Bogdan Alexe
bogdan.alexe@fmi.unibuc.ro

Secția Tehnologia Informației, anul III, 2022-2023
Cursul 3

Schimbare de orar

De săptămâna viitoare facem permutarea între ASC și IA

Cursul de Inteligență Artificială va fi joi, în intervalul 10-12, cursul de Arhitectura Sistemelor de Calcul va fi joi, în intervalul 12-14.

Recapitulare – cursul trecut

1. Învățarea Automată: definiție și terminologie

2. Paradigme de învățare

3. Metrici de măsurare a performanței pentru clasificare

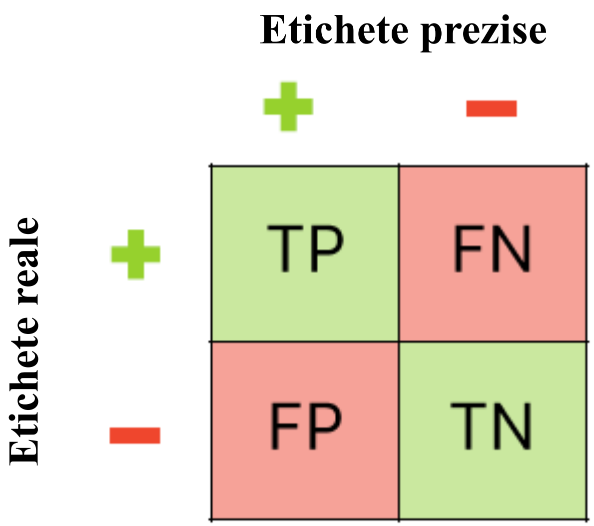
Terminologie în învățarea automată

- **etichetă** (label) ⟶ ce încercăm să prezicem
- **caracteristică** (feature) ⟶ proprietate măsurabilă a unui punct din exemplele din date
- **model** (ipoteză) ⟶ relația dintre caracteristici și etichete
- **antrenare** ⟶ stabilirea relației pe baza unei mulțimi de puncte
- **inferență** ⟶ realizarea de predicții pe date necunoscute
- **algoritm** ⟶ definește cum se face învățarea, poate fi restricționat la o anumită mulțime de ipoteze, fie prin alegerea unui anumit spațiu, fie pe baza hiperparametrilor

Paradigme de învățare

- **Învățare Supervizată** ⟶ prezicem o etichetă
  - etichetă continuă ⟶**Regresie**
  - etichetă discretă ⟶ **Clasificare**
- **Învățarea Nesupervizată**⟶descoperirea de structuri în date
  - gruparea itemilor similari ⟶ **Clusterizare**
  - pattern-uri frecvente ⟶**Căutarea regulilor de asociere**
- **Învățare pe bază de recompense**⟶fără etichete, ci numai recompense sau penalități pentru realizarea unor acțiuni
- **Învățare semi-supervizată**⟶ câteva date etichetate, multe neetichetate
- **Învățarea prin transfer**⟶folosirea unui model antrenat pe o problemă pentru învățarea mai rapidă pentru altă problemă

Acuratețe, Precizie și Recall



Matrice de confuzie

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F_1 = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = \frac{2(\text{Precision} * \text{Recall})}{\text{Precision} + \text{Recall}}$$

TP – true positive
FN – false negative
FP – false positive
TN – true negative

Cuvinte cheie

Învățare automată

Etichetă | Caracteristică | Vector de caracteristici

Antrenare | Inferență | Model | Ipoteză | Algoritm

Supervizare | Clasificare | Regresie | Nesupervizare

Semi-supervizare

Acuratețe | Precizie | Recall | Matrice de confuzie

Cuprinsul cursului de azi

1. Modelul celor mai apropiați k vecini (k-nearest neighbors) – continuare

2. Normalizarea caracteristicilor

3. Clasificatorul naïve Bayes

4. Evaluarea performanței unui model

Modelul celor mai apropiați k-vecini (k-nearest neighbors)

Modelul celor mai apropiați k-vecini (kNN)

- Modelul kNN prezice eticheta unui exemplu test ca fiind eticheta predominantă ale celor mai apropiate *k* exemple de antrenare (cei mai apropiați *k* vecini) din spațiul caracteristicilor (*feature space*).

- Învățare supervizată (avem etichete).

- Modelul kNN nu învață explicit un model:
  - în schimb, memorează datele de antrenare și le folosește în realizarea de predicții

- k = numărul de vecini este hiperparametru

Laborator

- kNN aplicat pe o parte din setul de date MNIST (cifre scrise de mână). Setul original conține 50.000 de exemple de antrenare, 10.000 de exemple de testare.

Vizualizare date de antrenare

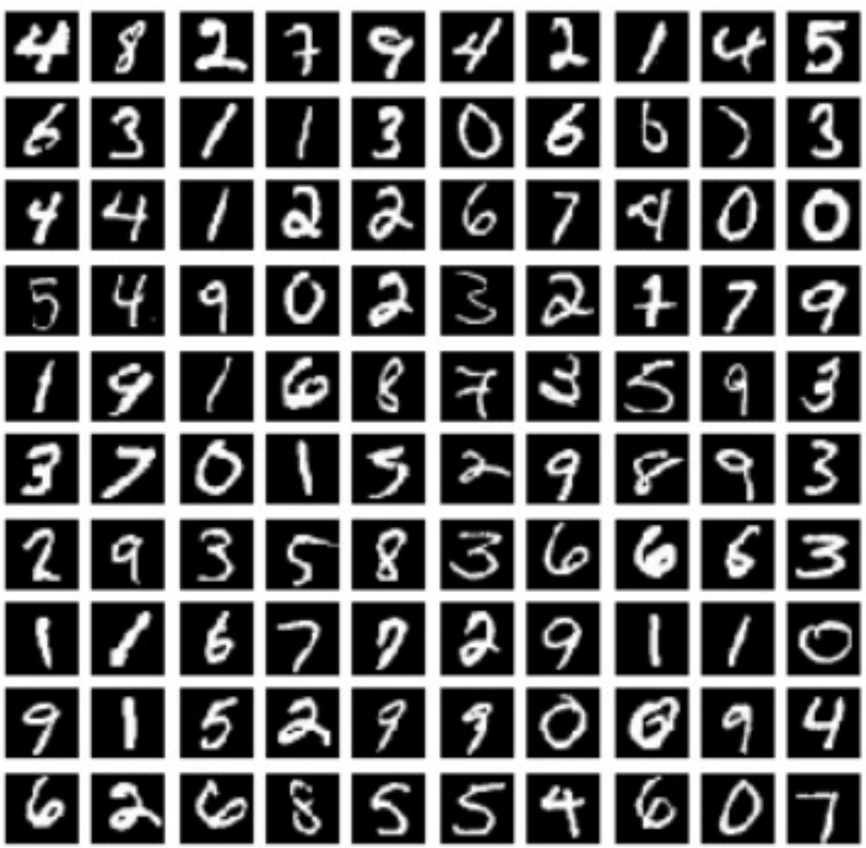
- kNN aplicat pe o parte din setul de date MNIST (cifre scrise de mână). Setul original conține 50.000 de exemple de antrenare, 10.000 de exemple de testare.

```python
#plot the first 100 training images with their labels in a 10 x 10 subp
nbImages = 10
plt.figure(figsize=(5,5))
for i in range(nbImages**2):
    plt.subplot(nbImages,nbImages,i+1)
    plt.axis('off')
    plt.imshow(np.reshape(train_images[i,:],(28,28)),cmap = "gray")
plt.show()
labels_nbImages = train_labels[:nbImages**2]
print(np.reshape(labels_nbImages,(nbImages,nbImages)))
```

Vizualizare date de antrenare

- kNN aplicat pe o parte din setul de date MNIST (cifre scrise de mână). Setul original conține 50.000 de exemple de antrenare, 10.000 de exemple de testare.

Primele 100 exemple de antrenare



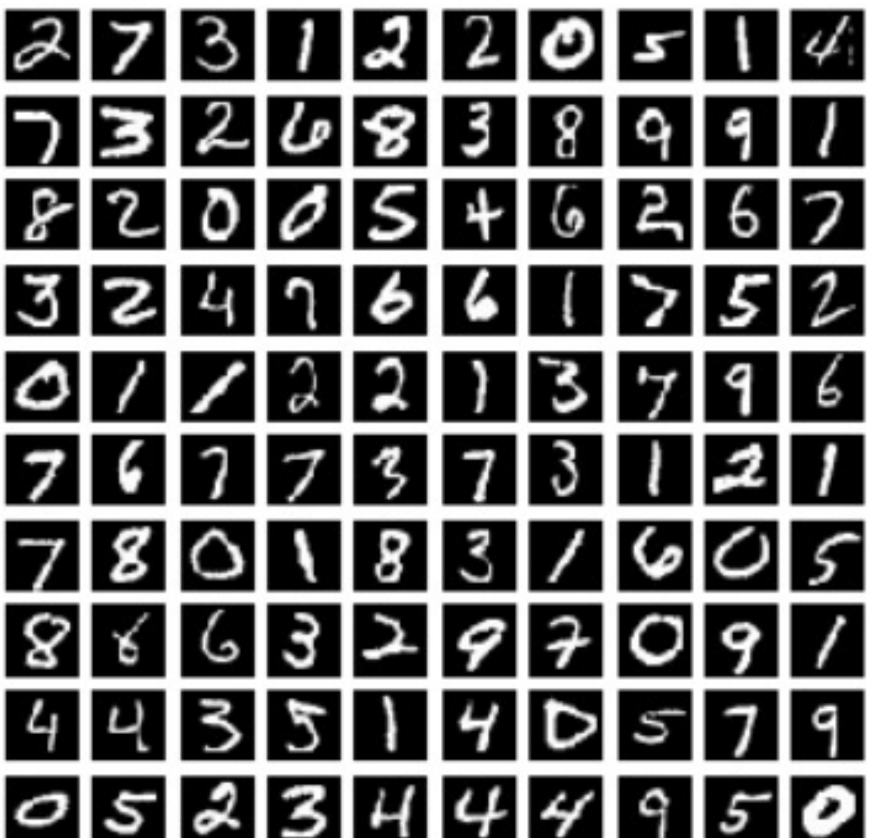
Etichetele corespunzătoare primelor 100 exemple de antrenare

```
[[4 8 2 7 9 4 2 1 4 5]
 [6 3 1 1 3 0 6 6 7 3]
 [4 4 1 2 2 6 7 4 0 0]
 [5 4 9 0 2 3 2 7 7 9]
 [1 9 1 6 8 7 3 5 9 3]
 [3 7 0 1 3 2 9 8 9 3]
 [2 9 3 5 8 3 6 6 6 3]
 [1 1 6 7 7 2 9 1 1 0]
 [9 1 5 2 9 9 0 0 9 4]
 [6 2 6 8 5 5 4 6 0 7]]
```

Vizualizare date de testare

- kNN aplicat pe o parte din setul de date MNIST (cifre scrise de mână). Setul original conține 50.000 de exemple de antrenare, 10.000 de exemple de testare.

Primele 100 exemple de testare



Etichetele corespunzătoare primelor 100 exemple de testare

```
[[2 7 3 1 2 2 0 5 1 4]
 [7 3 2 6 8 3 8 9 9 1]
 [8 2 0 0 5 4 6 2 6 7]
 [3 2 4 7 6 6 1 7 5 2]
 [0 1 1 2 2 1 3 7 9 6]
 [7 6 7 7 3 7 3 1 2 1]
 [7 8 0 1 8 3 1 6 0 5]
 [8 6 6 3 2 9 7 0 9 1]
 [4 4 3 5 1 4 0 5 7 9]
 [0 5 2 3 4 4 4 9 5 0]]
```

Vizualizare 7-NN

- Aplicăm 7-NN pentru primul exemplu de testare

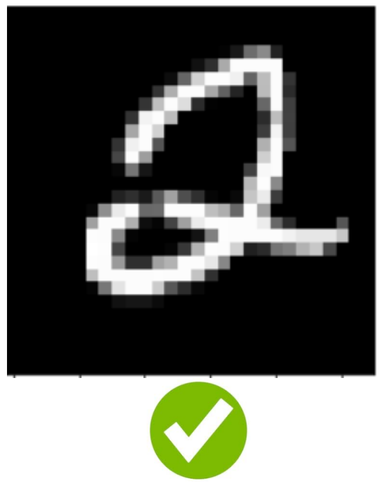
Exemplu testare

Vizualizare 7-NN

- Aplicăm 7-NN pentru primul exemplu de testare

Exemplu testare

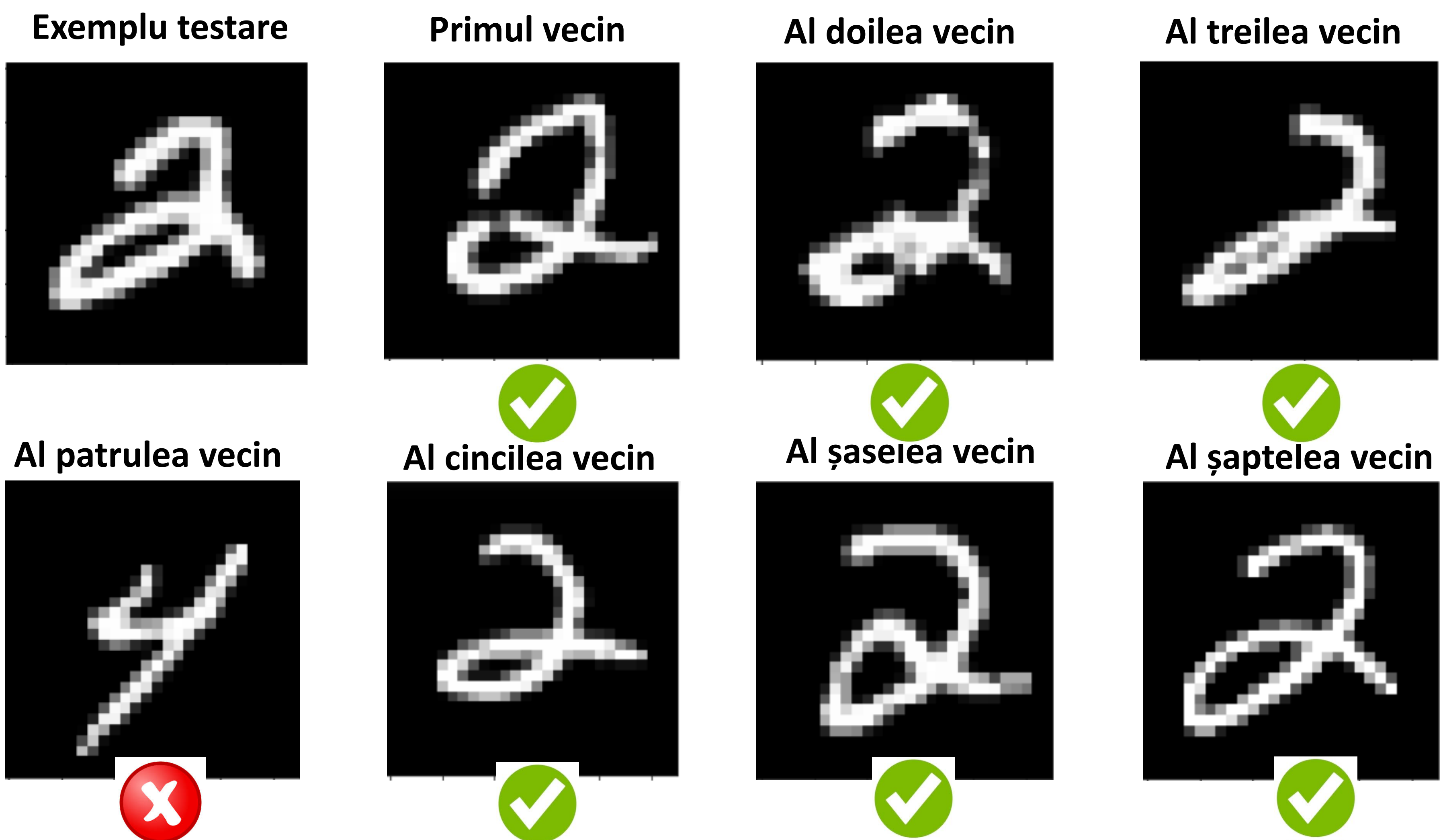
Primul vecin →



Cel mai apropiat exemplu din setul de antrenare considerând distanța Euclidiană calculată pe cele 784 de componente

Vizualizare 7-NN

- Aplicăm 7-NN pentru primul exemplu de testare



Laborator 3 – Matrice confuzie

		Eticheta prezisa									
		0	1	2	3	4	5	6	7	8	9
Eticheta reala	0	51	0	0	0	0	1	1	0	0	0
	1	0	52	0	0	0	0	0	0	0	0
	2	1	6	47	1	0	0	1	2	0	0
	3	0	0	0	51	0	1	0	0	0	1
	4	0	0	0	0	44	0	0	0	0	2
	5	2	1	1	6	0	40	1	0	0	1
	6	0	0	0	0	0	1	47	0	0	0
	7	1	2	0	0	1	0	0	46	0	0
	8	1	0	2	2	1	1	1	1	36	1
	9	0	0	1	1	3	1	0	1	0	35

-nu e simetrică

-numărul de exemple clasificate corect – pe diagonala principală

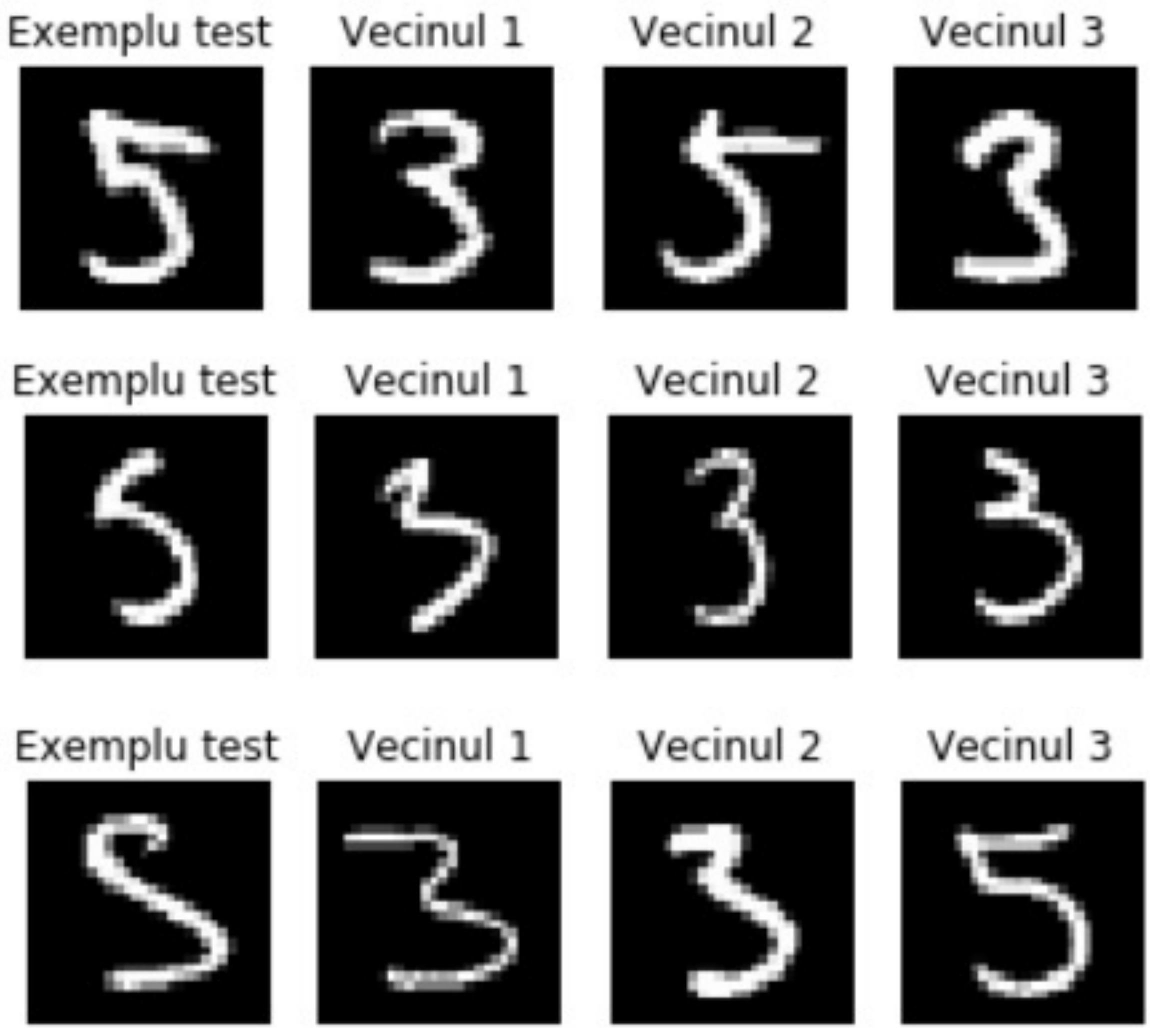
-numărul de exemple clasificate greșit – restul matricei

Laborator 3 – Matrice confuzie

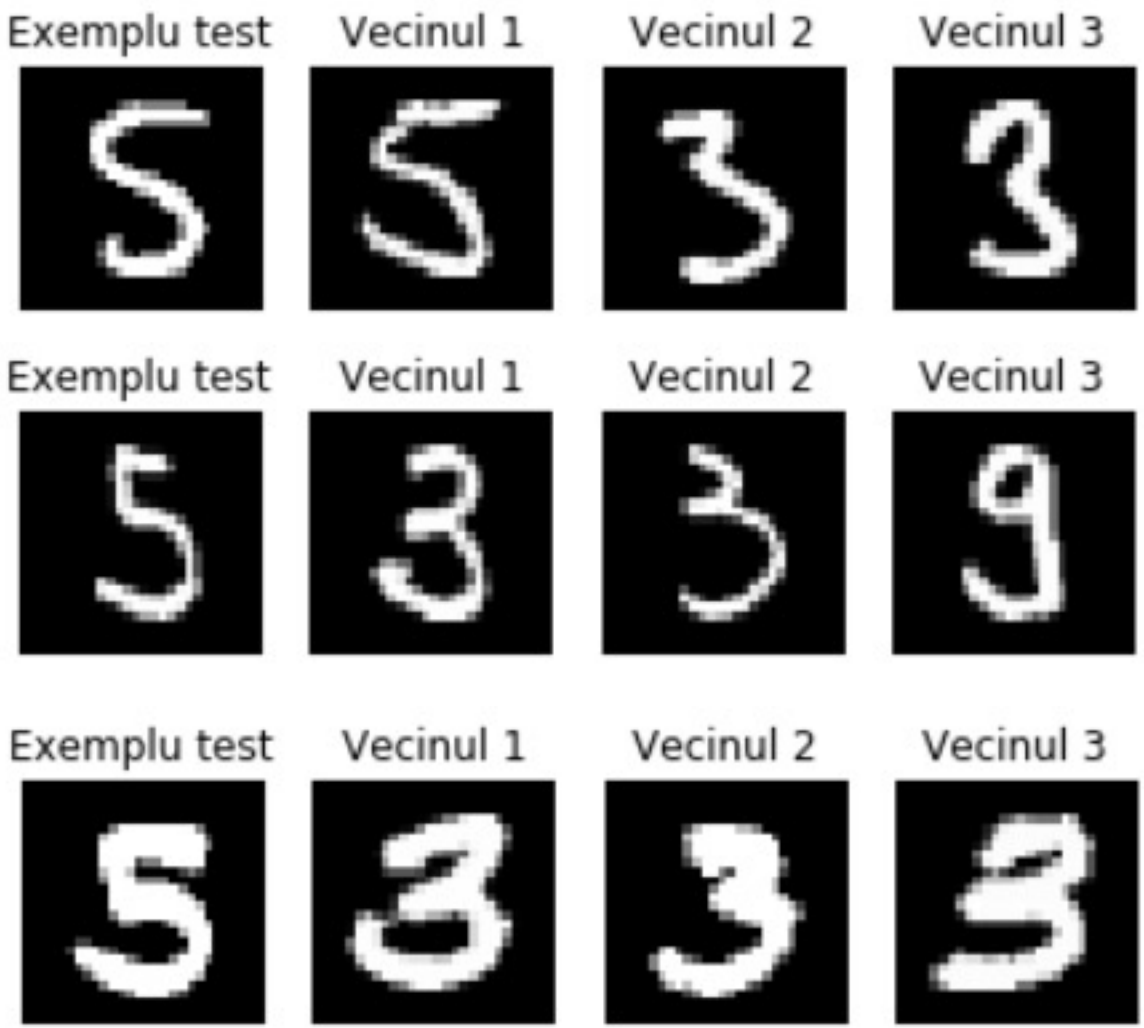
	0	1	2	3	4	5	6	7	8	9
0	51	0	0	0	0	1	1	0	0	0
1	0	52	0	0	0	0	0	0	0	0
2	1	6	47	1	0	0	1	2	0	0
3	0	0	0	51	0	1	0	0	0	1
4	0	0	0	0	44	0	0	0	0	2
5	2	1	1	6	0	40	1	0	0	1
6	0	0	0	0	0	1	47	0	0	0
7	1	2	0	0	1	0	0	46	0	0
8	1	0	2	2	1	1	1	1	36	1
9	0	0	1	1	3	1	0	1	0	35

Numărul de exemple de testare cu cifra 5 care au fost clasificate drept cifra 3

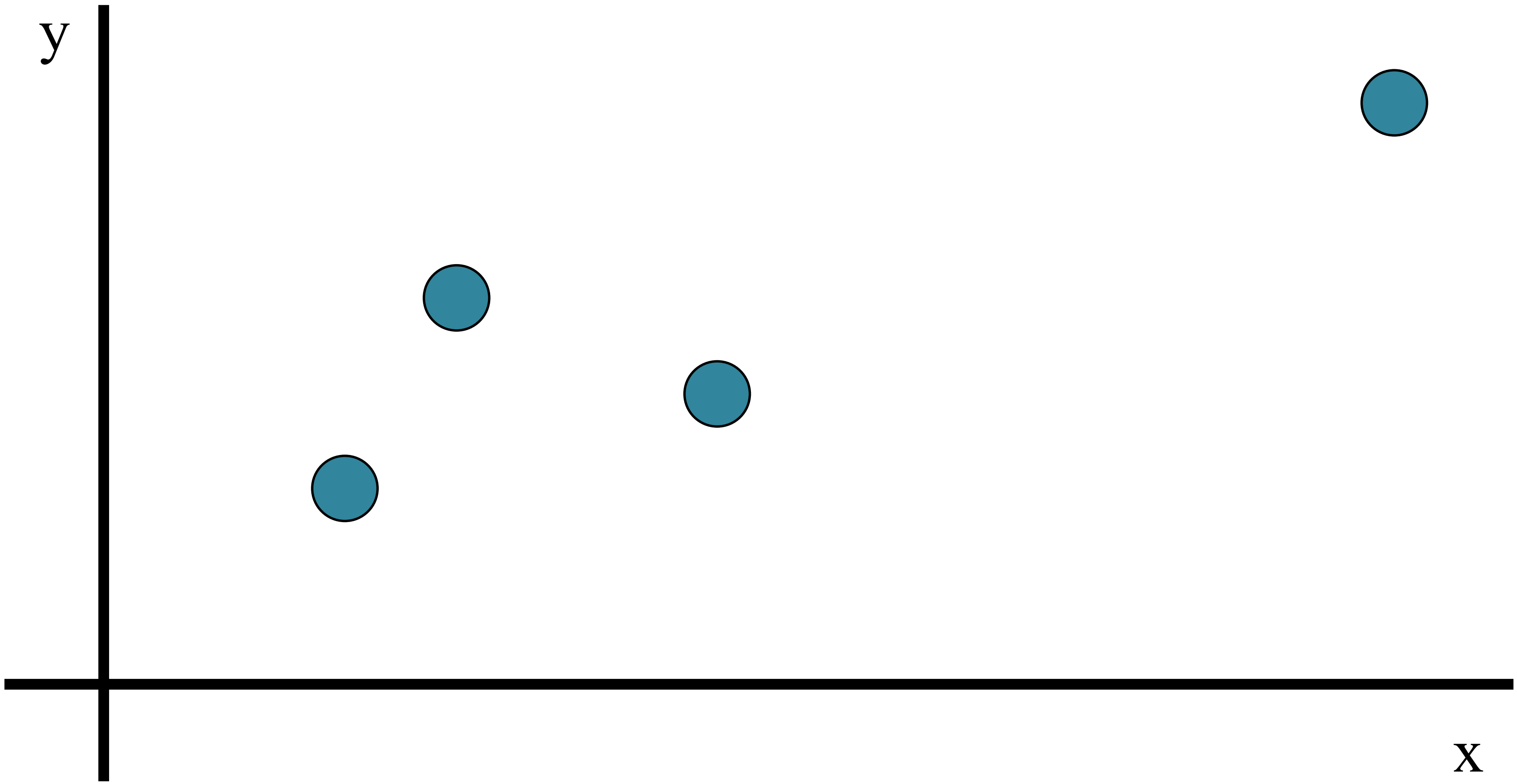
Laborator 3 – Exemple de testare cu cifra 5 clasificate drept cifra 3



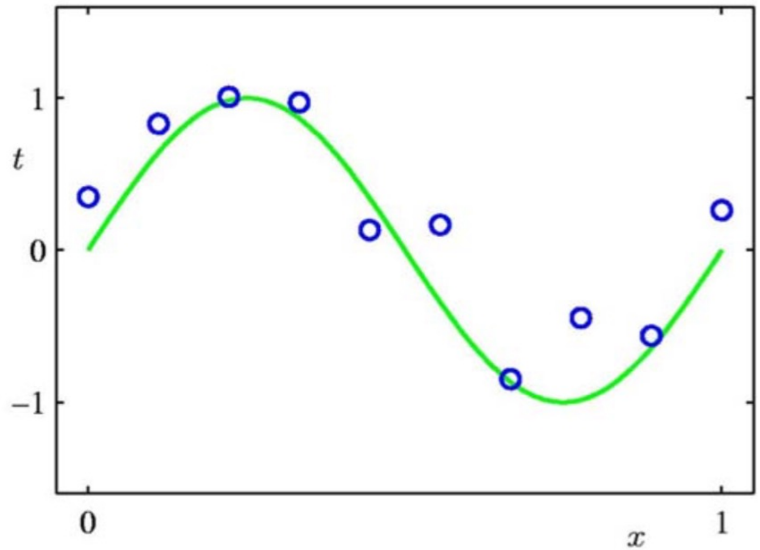
Laborator 3 – Exemple de testare cu cifra 5 clasificate drept cifra 3



Modelul celor mai apropiați k-vecini pentru probleme de regresie



Regresie din exemple etichetate



- Presupunem că avem un set de N exemple de antrenare:

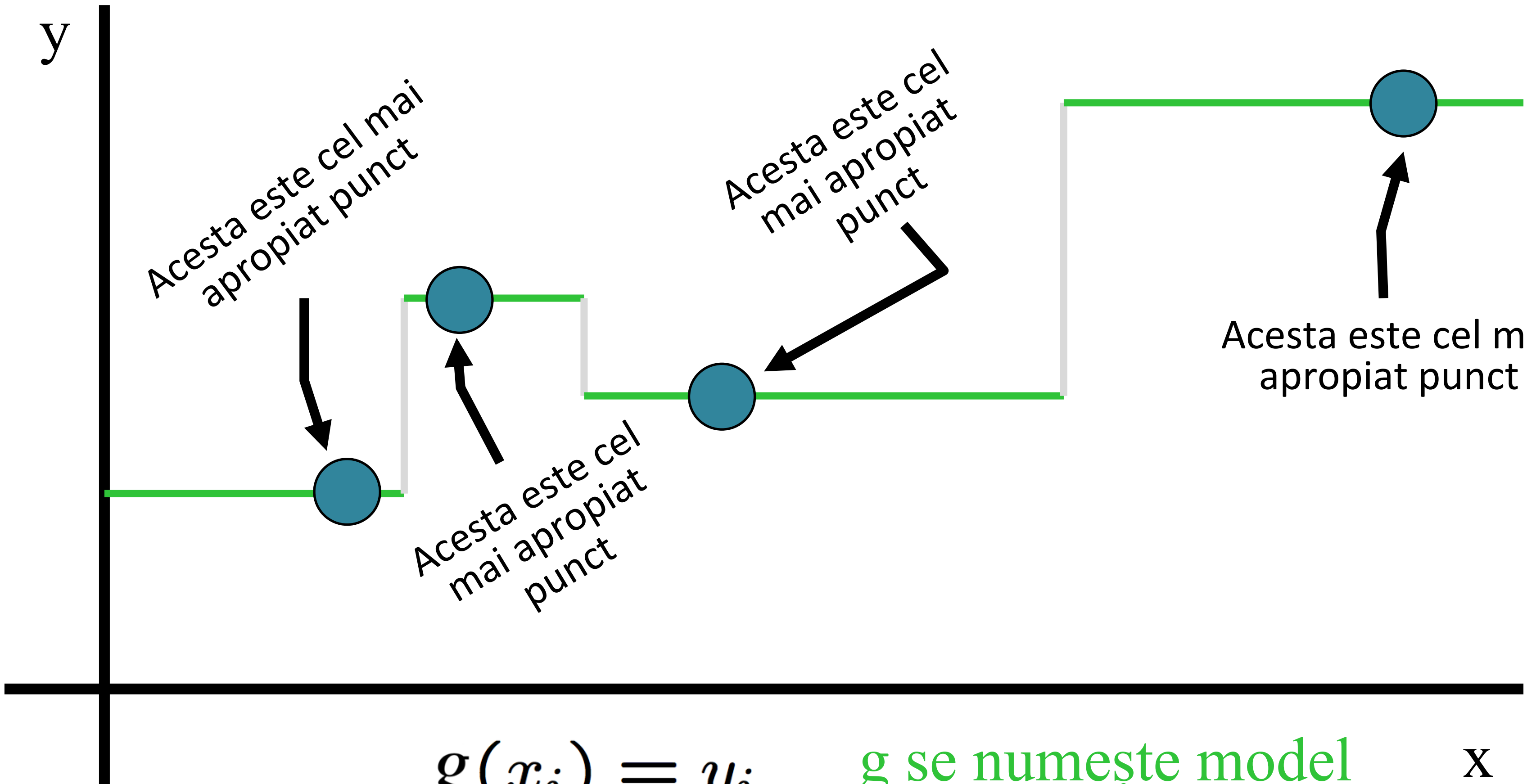
$(x_1, \ldots, x_N)$ și $(y_1, \ldots, y_N), x_i, y_i \in \mathbb{R}$

- Problema regresiei constă în estimarea funcției g(x) a.î.:

$$g(x_i) = y_i$$ g se numește model

Modelul celor mai apropiați k-vecini pentru probleme de regresie

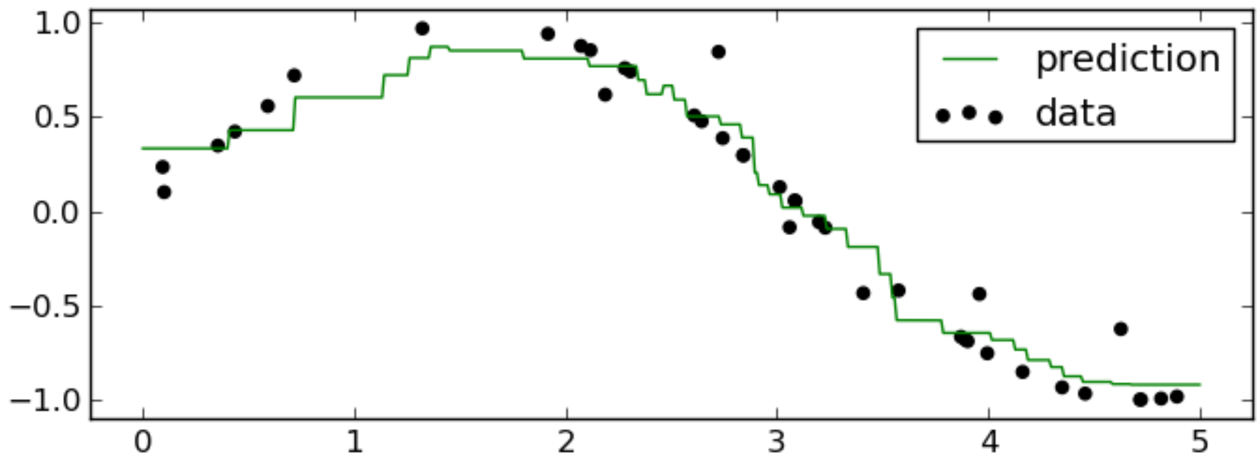
K = 1



$$g(x_i) = y_i$$

g se numește model

Modelul celor mai apropiați k-vecini pentru probleme de regresie



Algoritmul de regresie bazat pe cei mai apropiați k-vecini:

1) Pentru fiecare exemplu de test x, găsim cei mai apropiați k vecini și etichetele lor

2) Predicția este media etichetelor celor k vecini

$$f(\mathbf{x}) = \frac{1}{K} \sum_{i=1}^{K} y_i$$

Modelul celor mai apropiați k-vecini (kNN)

Avantaje:

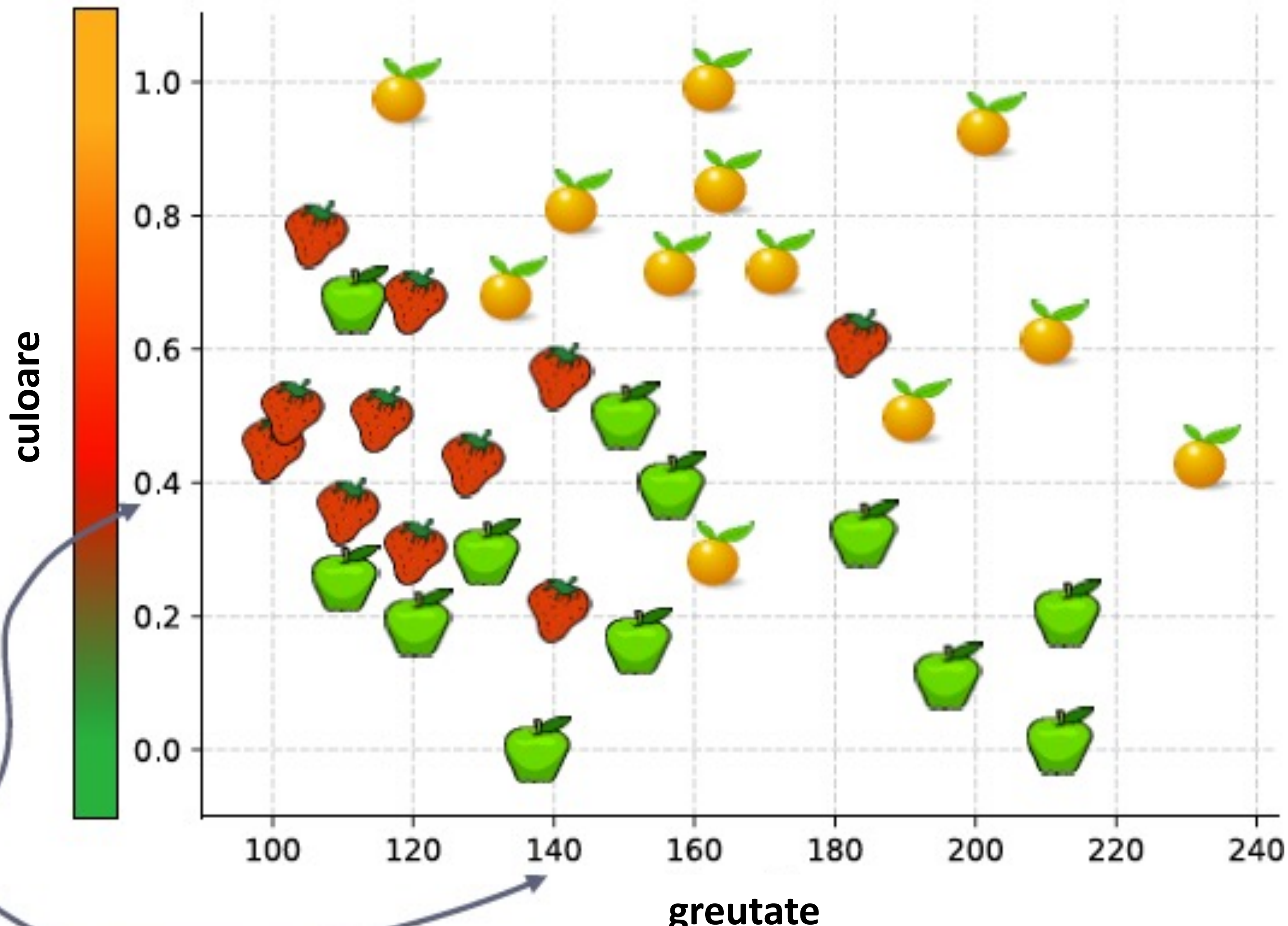
- intuitiv, simplu de înțeles, ușor de implementat
- nu pornește de la anumite presupuneri, poate fi aplicat oricând
- nu are parametri
- poate fi folosit în clasificare binară sau cu mai multe clase dar și în regresie

Dezavantaje:

- timpul de rulare este proporțional cu datele stocate în memorie
- cum alegem hiperparametrul K (= numărul de vecini) ?
- nu învață nimic (memorează datele), spațiu de stocare mare
- ce distanță alegem? caracteristicile trebuie normalizate (să aibă aceeași scală)
- blestemul dimensionalității (ce se întâmplă în multe dimensiuni)
- senzitiv la date debalansate, punctele outlier, caracteristici lipsă

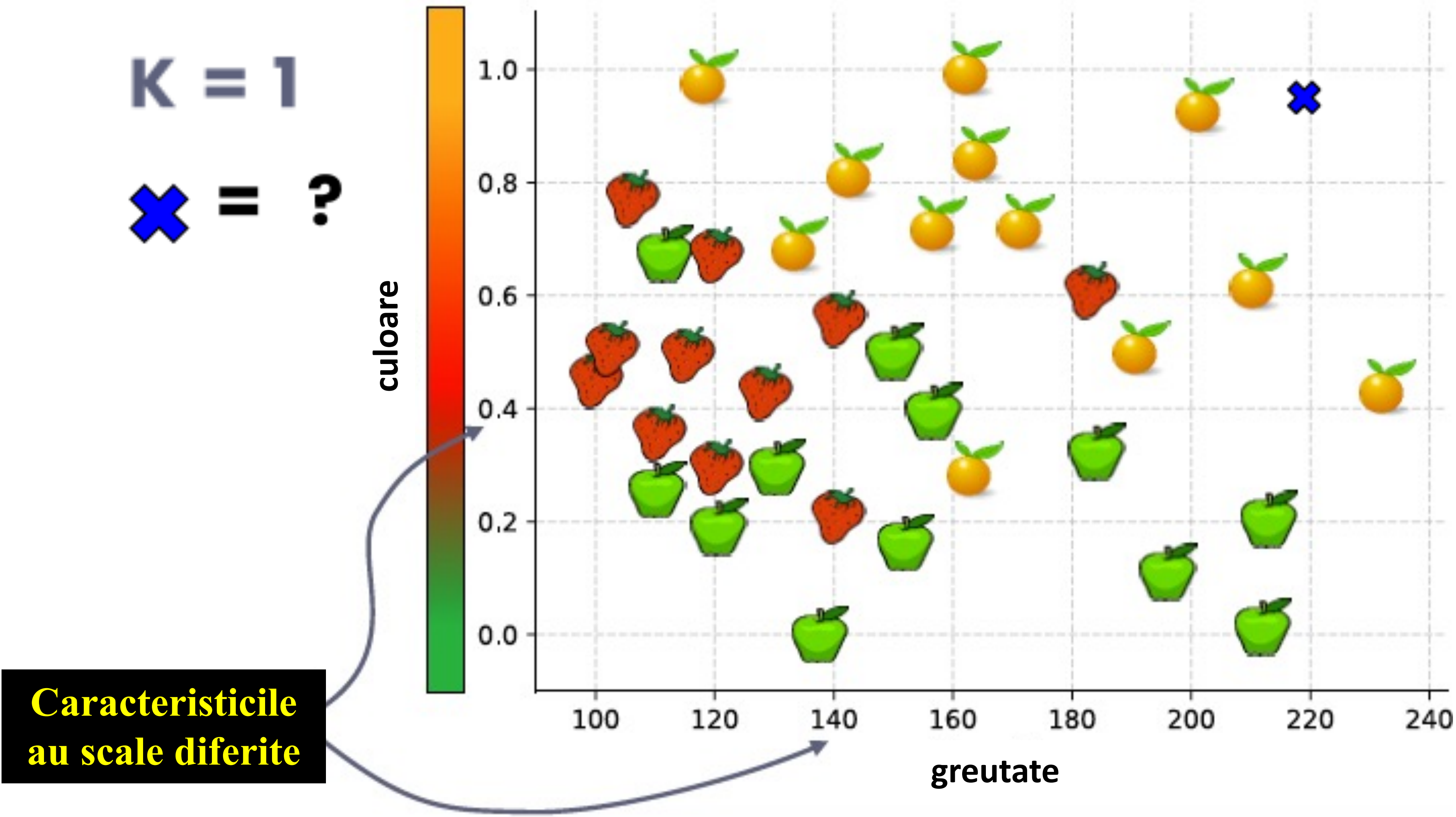
Normalizarea caracteristicilor

Normalizarea caracteristicilor



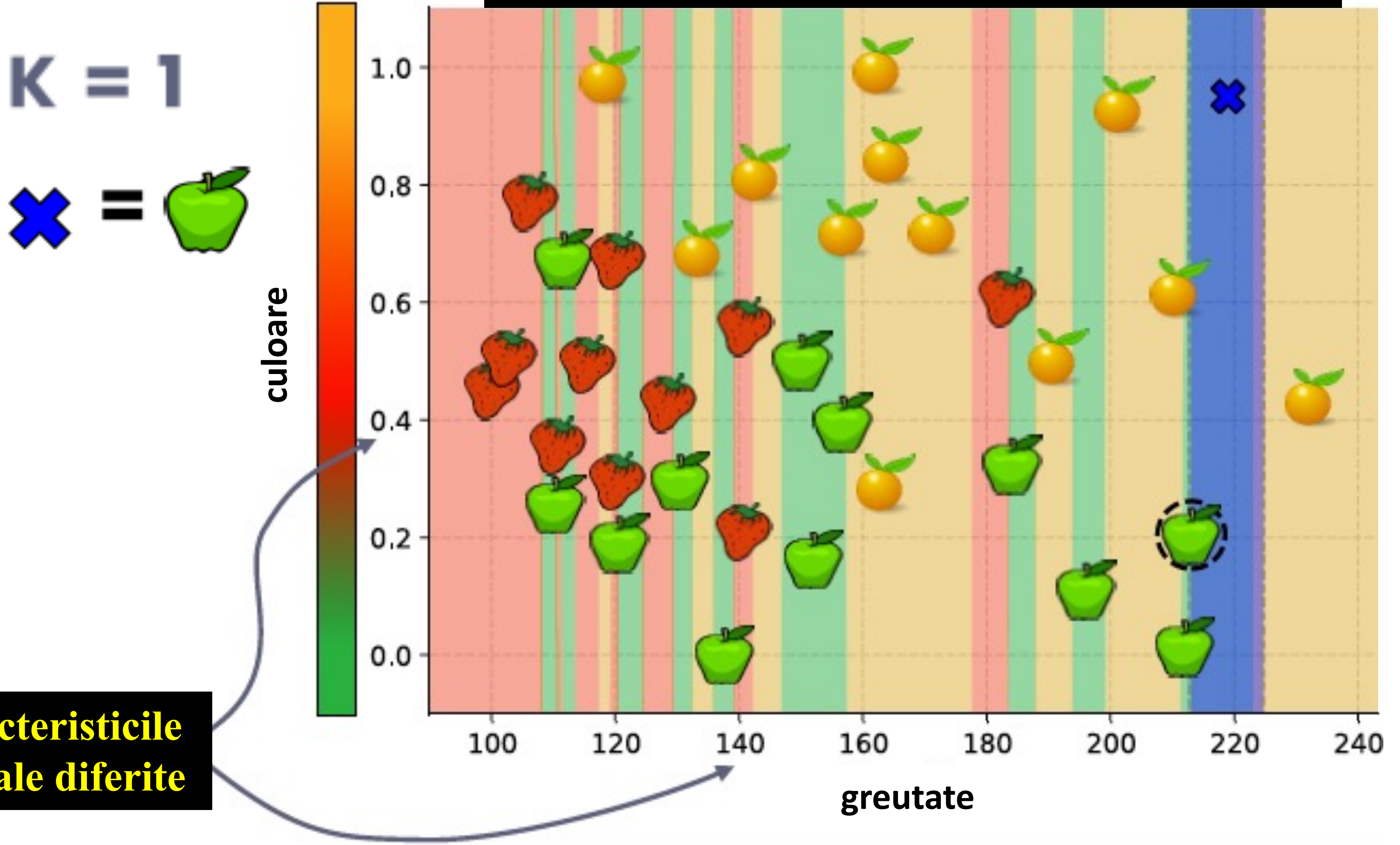
Caracteristicile au scale diferite

Normalizarea caracteristicilor



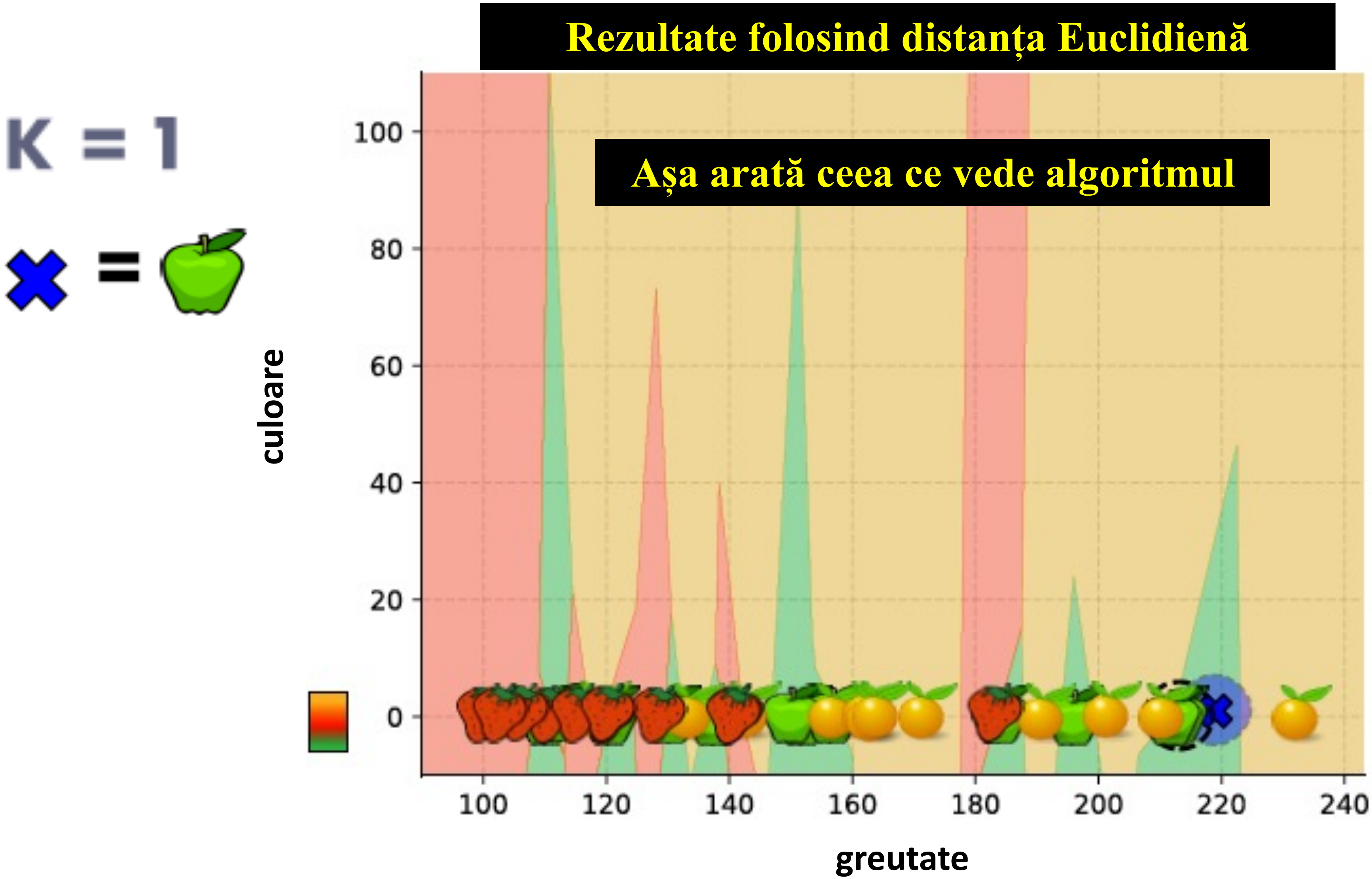
Normalizarea caracteristicilor

Rezultate folosind distanța Euclidienă



Caracteristicile au scale diferite

Normalizarea caracteristicilor



Normalizarea caracteristicilor

- când caracteristicile au scale (mărimi) diferite, unele caracteristici domină pe celelalte în calculul distanței.

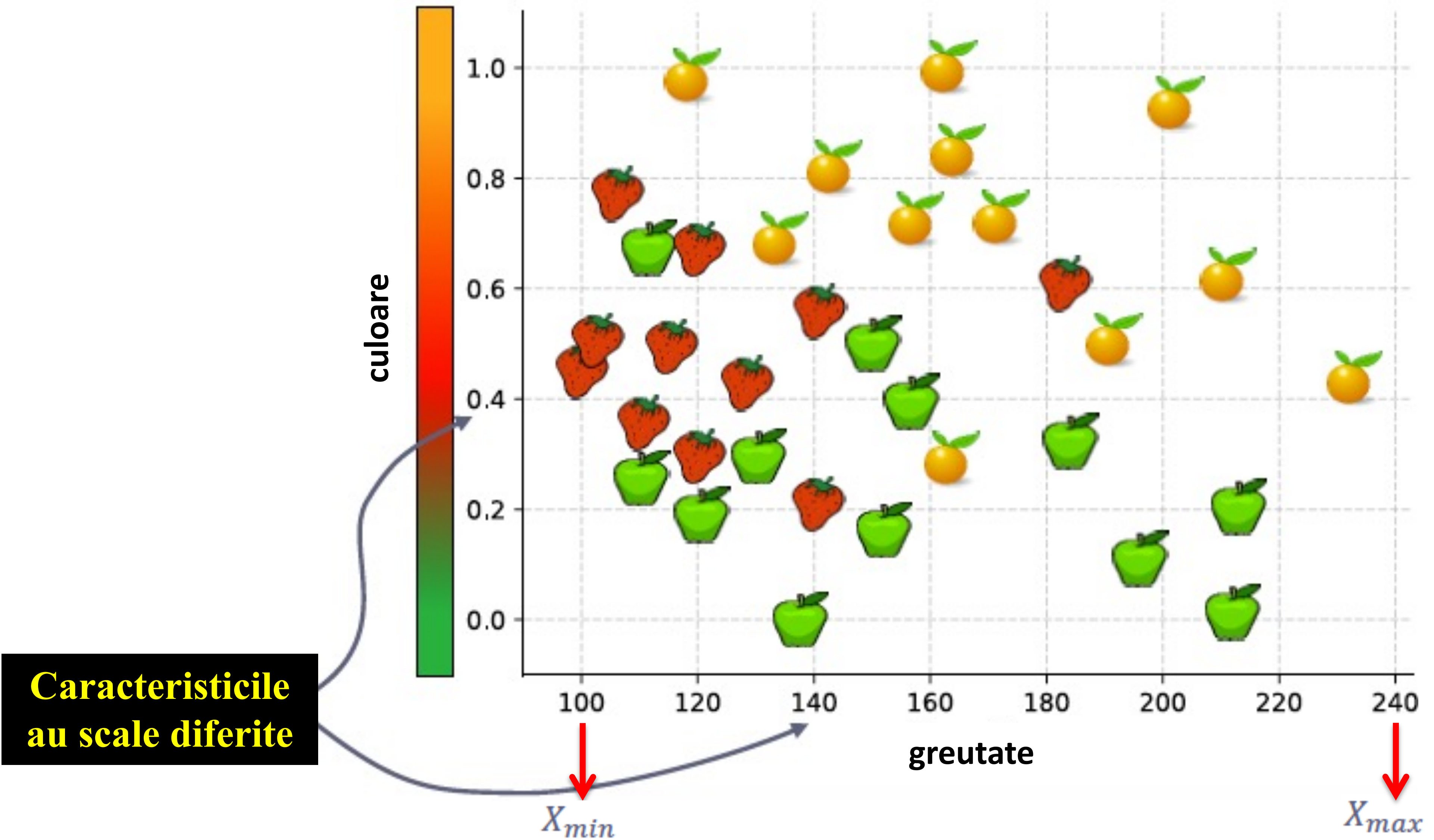
- vrem să avem toate caracteristicile în același interval:
  - ***scalare de tip min-max*** – transformă fiecare caracteristică în parte în intervalul [0, 1]:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

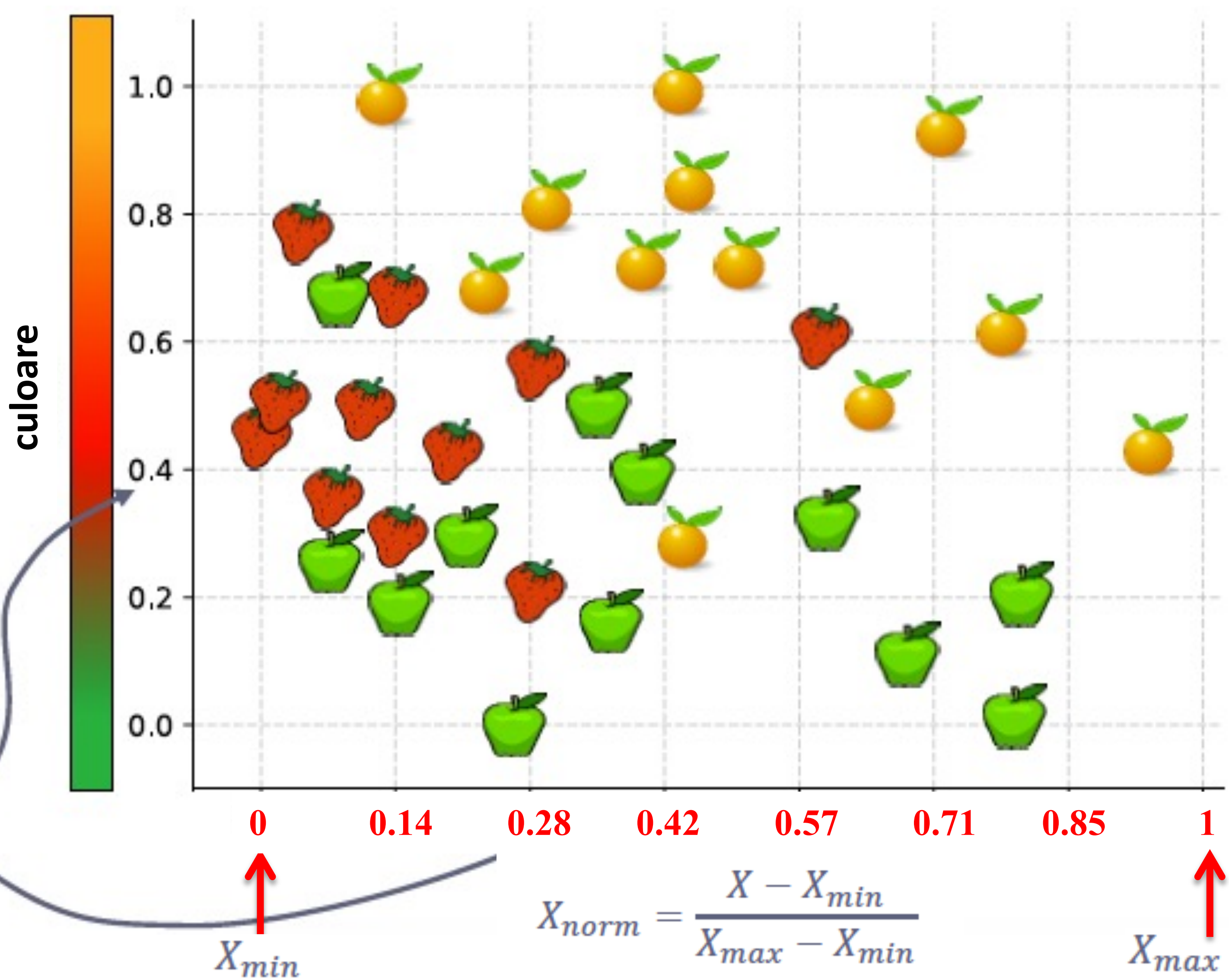
  - ***standardizare*** – transformă fiecare caracteristică în parte astfel încât au proprietăți ale distribuției normale standardizate (medie = 0 și deviație standard = 1):

$$X_{norm} = \frac{X - \mu_X}{\sigma_X}$$

Normalizarea caracteristicilor
Scalare de tip min-max



Normalizarea caracteristicilor
Scalare de tip min-max



Caracteristicile au aceeași scală

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Clasificatorul naïve Bayes

Probabilități - recapitulare

- Avem două zaruri, fiecare având 6 fețe. Amestecăm zarurile și le aruncăm.
- Definim evenimentele următoare:

A: zarul 1 arată fața “3”,

B: zarul 2 arată fața “1”,

C: suma fețelor celor două zaruri este 8.

Calculați probabilitățile următoare:

1) $P(A) = ?$ P(A) = 1/6

P(B) = 1/6

P(C) = 5/36

P(A|B) = 1/6

P(C|A) = 1/6

P(A, B) = 1/36

P(A, C) = 1/36

NU

Probabilități - recapitulare

- Avem două zaruri, fiecare având 6 fețe. Amestecăm zarurile și le aruncăm.
- Definim evenimentele următoare:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A,B) = P(A \cap B) = P(A/B) \times P(B)$$

A: zarul 1 arată fața "3",

B: zarul 2 arată fața "1",

C: suma fețelor celor două zaruri este 8.

Calculați probabilitățile următoare:

1) $P(A) = ?$ P(A) = 1/6
2) $P(B) = ?$ P(B) = 1/6
3) $P(C) = ?$ P(C) = 5/36
4) $P(A|B) = ?$ P(A|B) = 1/6
5) $P(C|A) = ?$ P(C|A) = 1/6
6) $P(A,B) = ?$ P(A, B) = 1/36
7) $P(A,C) = ?$ P(A, C) = 1/36
8) Este P(A,C) = P(A) * P(C)? NU

Ω - spațiu total de evenimente – modelează fenomenul

Ω = {(1,1), (1,2),…., (6,6)} – conține 36 de evenimente elementare, fiecare are probabilitatea 1/36

A = {(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)} P(A) = 1/6

B = {(1,1), (2,1), (3,1), (4,1), (5,1), (6,1)} P(B) = 1/6

C = {(2,6), (3,5), (4,4), (5,3), (6,2)} P(C) = 5/36

Teorema lui Bayes

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \cap A)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

actualizează probabilitea ipotezei A pe baza evidenței B

Exemplu: O boală rară are incidență de 5% asupra populației. Un test medical care se face pentru depistarea bolii are acuratețe 90% (dă răspunsul corect în 90% din cazuri). Dacă o persoană face testul și iese pozitiv care este probabilitatea ca ea să fie bolnavă?

Teorema lui Bayes

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \cap A)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}$$

Exemplu: O boală rară are incidență de 5% asupra populației. Un test medical care se face pentru depistarea bolii are acuratețe 90% (dă răspunsul corect în 90% din cazuri). Dacă o persoană face testul și iese pozitiv care este probabilitatea ca ea să fie bolnavă?

Ω - spațiu total de evenimente – modelează fenomenul

Ω = {(c,t)} – perechi de forma (clasă, test), unde clasa poate fi {bolnav, sanatos} iar testul poate fi {pozitiv, negative}

A = persoana este bolnava, A = {(bolnav, pozitiv), (bolnav, negativ)}
P(A) = 0.05, $P(\bar{A}) = 0.95$

B = testul iese pozitiv, B = {(bolnav,pozitiv), (negativ,pozitiv)}
P(B|A) = 0.9
P($\bar{B}$| $\bar{A}$) $= 0.9$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.9 * 0.05}{P(B)}$$

Teorema lui Bayes

$P(B) = P(B \ \cap \ \Omega) = P\left(B \cap \ (A \cup \overline{A})\right) = P\big((B \cap A) \cup (B \cap \overline{A}\ )\big) = P(B \cap A) + P\big(B \cap \ \overline{A}\big)$

$P(B) = \mathrm{P(B|A)P(A)} + \mathrm{P}\big(\mathrm{B}\big|\overline{A}\big)P(\overline{A})$

A = persoana este bolnava, A = {(bolnav, pozitiv), (bolnav, negativ)}

P(A) = 0.05, $P(\bar{A}) = 0.95$

B = testul iese pozitiv, B = {(bolnav,pozitiv), (negativ,pozitiv)}

P(B|A) = 0.9

P($\bar{B}$| $\bar{A}$) = 0.9, P(B|$\bar{A}$)=0.1

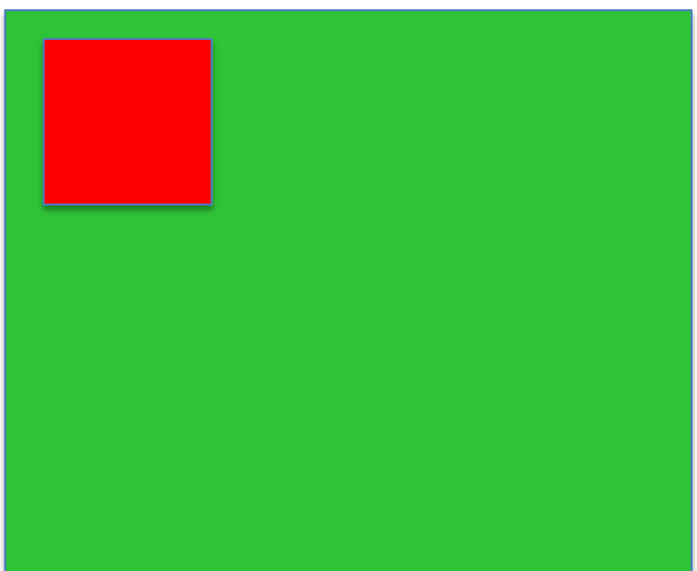
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.9*0.05}{0.9*0.05+0.1*0.95} = \frac{0.045}{0.045+0.095} = \frac{0.045}{0.13} = 34.61\%$$

Teorema lui Bayes

A = persoana este bolnava, A = {(bolnav, pozitiv), (bolnav, negativ)}
P(A) = 0.05, $P(\bar{A}) = 0.95$

B = testul iese pozitiv, B = {(bolnav,pozitiv), (negativ,pozitiv)}
P(B|A) = 0.9
P($\bar{B}$| $\bar{A}$) $= 0.9$, P(B|$\bar{A}$)=0.1

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.9*0.05}{0.9*0.05+0.1*0.95} = \frac{0.045}{0.045+0.095} = \frac{0.045}{0.13} = 34.61\%$$



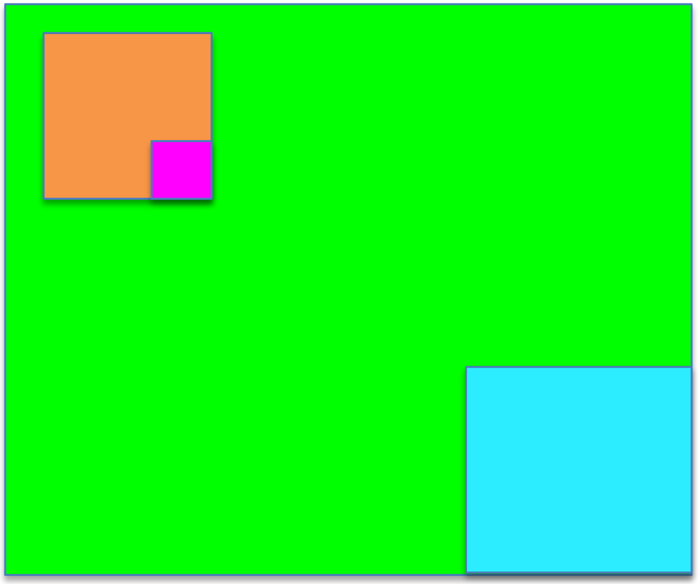
Populatie sanatoasa = 95%
Populatia bolnava = 5%

Teorema lui Bayes

A = persoana este bolnava, A = {(bolnav, pozitiv), (bolnav, negativ)}
P(A) = 0.05, $P(\bar{A}) = 0.95$

B = testul iese pozitiv, B = {(bolnav,pozitiv), (negativ,pozitiv)}
P(B|A) = 0.9
$P(\bar{B}|\ \bar{A}) = 0.9$, P(B|$\bar{A}$)=0.1

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0.9*0.05}{0.9*0.05+0.1*0.95} = \frac{0.045}{0.045+0.095} = \frac{0.045}{0.13} = 34.61\%$$



Populatie sanatoasa, test negativ = 95% * 0.9 = 0.855
Populatie sanatoasa, test pozitiv = 95% * 0.1 = 0.095
Populatia bolnava, test pozitiv = 5% * 0.9 = 0.045
Populatia bolnava, test negativ = 5% * 0.1 = 0.005

Daca testul iese pozitiv, avem ca persoana provine fie din regiunea portocalie (bolnav) fie din regiunea cyan (sanatos): probabilitatea ca persoana sa fie bolnava este 0.045/(0.045+0.095) = 0.3461

Clasificare

- Definirea problemei:
  - date fiind caracteristicile măsurate $X_1, X_2, \ldots, X_n$
  - realizați o predicție a etichetei *c* (clasa)

Clasificare - exemplu

- Definirea problemei:
  - date fiind caracteristicile măsurate $X_1, X_2, \ldots, X_n$
  - realizați o predicție a etichetei *c* (clasa)

- Clasificare de gen :
  - $X_1, X_2, \ldots, X_n$ sunt caracteristici măsurate la oameni: înălțimea, greutatea, etc.
  - $c \in \{B, F\}$ (clasifica dacă o persoană este bărbat sau femeie)

Clasificare - exemplu

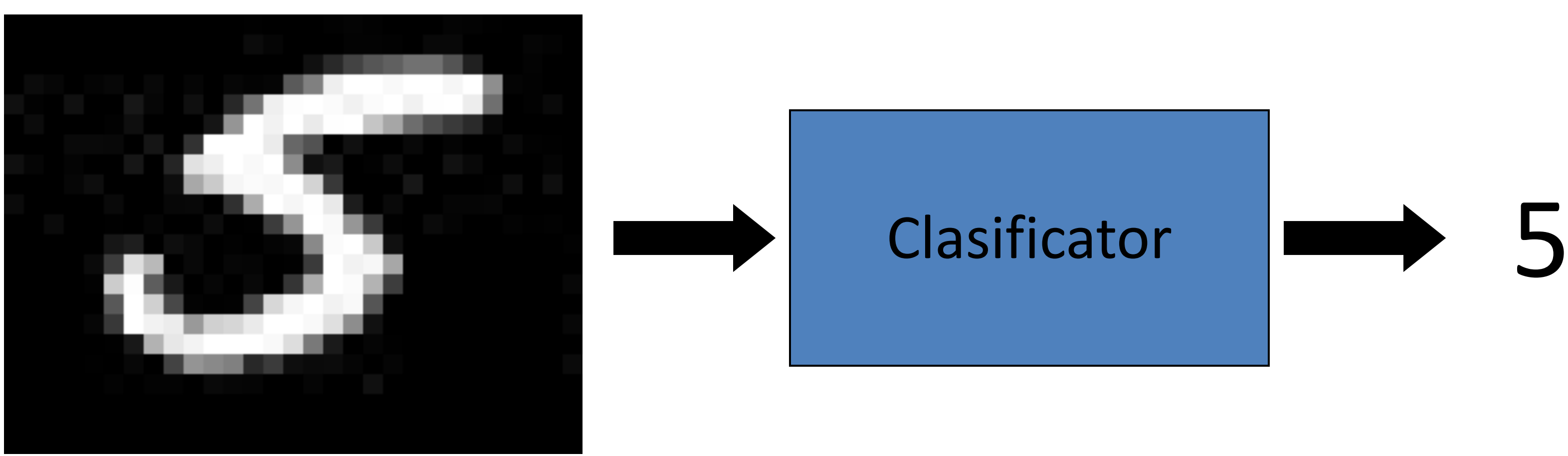
- Definirea problemei:
  - date fiind caracteristicile măsurate $X_1, X_2, \ldots, X_n$
  - realizați o predicție a etichetei *c*

- Clasificare de solvabilitate pentru acordarea de credit :
  - $X_1, X_2, \ldots, X_n$ sunt caracteristici măsurate la oameni de către bănci pentru analiza acordării unui credit: vârstă, sex, venituri, ocupație, stare civilă, număr copii, etc.
  - $c \in \{S, I\}$ (clasifica dacă o persoană este solvabilă sau nu, dacă poate returna creditul la timp)

Clasificare - exemplu

- Definirea problemei:
  - date fiind caracteristicile măsurate $X_1, X_2, \ldots, X_n$
  - realizați o predicție a etichetei *c*

- Clasificarea cifrelor scrise de mână (laborator 3+4):
  - $X_1, X_2, \ldots, X_n$ sunt intensitățile pixelilor dintr-o imagine grayscale
  - $X_1, X_2, \ldots, X_n \in \{0, 1, 2, \ldots 255\}$ – intensități
  - $c \in \{0, 1, 2, \ldots, 9\}$ (clasifică ce cifră este imaginea)



Regula lui Bayes

- Clasificarea cifrelor scrise de mână:
  - $X_1, X_2, \ldots, X_n$ sunt intensitățile pixelilor dintr-o imagine grayscale
  - $X_1, X_2, \ldots, X_n \in \{0, 1, 2, \ldots 255\}$ – intensități
  - $c \in \{0, 1, 2, \ldots, 9\}$ (clasifică ce cifră este imaginea)

- Considerăm X imaginea cu *n* pixeli de intensitate $X_1, X_2, \ldots, X_n$: $X = (X_1, X_2, \ldots, X_n)$

- Regula lui Bayes: $P(c / X) = \dfrac{P(X / c) \times P(c)}{P(X)}$

Ω - spațiu total de evenimente – modelează fenomenul

Ω = {(c,X)} – perechi de forma (clasă, imagine)

Regula lui Bayes

X = 

P(c=0|X) = 0.5

P(c=6|X) = 0.2

P(c=8|X) = 0.3

X = 

P(c=1|X) = 0.05

P(c=3|X) = 0.8

P(c=7|X) = 0.15

X = 

P(c=4|X) = 0.35

P(c=9|X) = 0.65

Imaginea X (din setul MNIST) văzută de algoritm

n = 28 × 28 = 784

X_1,
X_2,
…,
X_{784}

```
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.
    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    5.    5.    0.    0.
    0.    0.    0.    0.    0.    0.  110.  255.  197.    3.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.   27.  214.  230.  107.    0.
    0.    0.    0.    0.    0.    5.  231.  253.  253.   12.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.  161.  253.  253.  229.   48.
    0.    0.    0.    0.    0.   51.  242.  253.  253.  199.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    4.  170.  253.  253.  253.  191.
    0.    0.    0.    0.    0.  112.  253.  253.  253.  241.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.   47.  253.  253.  253.  253.  176.
    0.    0.    0.    0.   73.  242.  253.  253.  253.  129.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.   16.  218.  253.  253.  253.  199.   25.
    0.    0.    0.    0.  217.  253.  253.  253.  253.  208.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.   25.  126.  253.  253.  253.  223.   61.    0.
    0.    0.    0.   79.  241.  253.  253.  253.  253.  232.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    9.  219.  253.  253.  253.  253.  245.   65.    0.
    0.    0.  140.  240.  253.  253.  253.  253.  253.  129.    0.    0.    0.    0.]
[   0.    0.    0.    0.    2.  127.  253.  253.  253.  253.  241.  155.   56.   56.
   56.  153.  239.  253.  253.  253.  253.  253.  253.  129.    0.    0.    0.    0.]
[   0.    0.    0.    0.  128.  253.  253.  253.  253.  253.  253.  253.  253.  253.
  253.  253.  253.  253.  253.  253.  253.  253.  253.   38.    0.    0.    0.    0.]
[   0.    0.    0.    0.  194.  253.  253.  253.  253.  253.  253.  253.  253.  253.
  253.  253.  253.  253.  253.  253.  253.  253.  253.  100.    0.    0.    0.    0.]
[   0.    0.    0.    0.  181.  253.  253.  253.  253.  253.  253.  253.  253.  253.
  253.  253.  253.  253.  253.  253.  253.  253.  253.   93.    0.    0.    0.    0.]
[   0.    0.    0.    0.    2.   55.  172.  172.  182.  241.  172.  172.  172.  192.
  253.  253.  253.  253.  224.  172.   61.   49.   49.    1.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.  157.   82.    0.    0.    0.   63.
  253.  253.  253.  253.   57.    0.    0.    0.    0.    0.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    4.   32.    0.    0.    0.    0.  108.
  253.  253.  253.  176.   23.    0.    0.    0.    0.    0.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.   50.  245.
  253.  253.  222.   26.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.   21.  199.  253.
  253.  253.   61.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.   50.  253.  253.
  253.  253.   49.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.]
[   0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    0.    3.  183.  253.
```

Regula lui Bayes

- Considerăm X imaginea cu pixelii de intensitate X_1, $X_2, \ldots, X_n$: $X = (X_1, X_2, \ldots, X_n)$

- Regula lui Bayes:

Probabilitatea să observăm imaginea X condiționată de faptul că imaginea conține o cifră din clasa c

Probabilitatea a-priori ca o imagine să fie din clasa c (să conțină cifra din clasa c)

$$P(c / X) = \frac{P(X / c) \times P(c)}{P(X)}$$

Probabilitatea să avem cifra din clasa *c* dându-se imaginea X

Probabilitatea să observăm imaginea X

$$Posterior = \frac{Likelihood \times Prior}{Evidence}$$

Regula de clasificare

$$P(c=0\,/\,X)=\frac{P(X\,/\,c=0)\times P(c=0)}{P(X)}$$

$$P(c=1\,/\,X)=\frac{P(X\,/\,c=1)\times P(c=1)}{P(X)}$$

....

$$P(c=9\,/\,X)=\frac{P(X\,/\,c=9)\times P(c=9)}{P(X)}$$

Presupunem că am calculat cele 10 probabilități a-posteriori:
P(c=0|X), P(c=1|X), … P(c=9|X)

Cum arată regula de clasificare?

Regula de clasificare

$$P(c=0/X)=\frac{P(X/c=0)\times P(c=0)}{P(X)}$$

$$P(c=1/X)=\frac{P(X/c=1)\times P(c=1)}{P(X)}$$

....

$$P(c=9/X)=\frac{P(X/c=9)\times P(c=9)}{P(X)}$$

Presupunem că am calculat cele 10 probabilități a-posteriori:
P(c=0|X), P(c=1|X), … P(c=9|X)

Regula de clasificare: $c^{*}=\arg\max_{i=0,1,\ldots,9} P(c=i/X)$

alege clasa care maximizează probabilitatea a-posteriori

Regula de clasificare

$$P(c=0/X)=\frac{P(X/c=0)\times P(c=0)}{P(X)}$$

$$P(c=1/X)=\frac{P(X/c=1)\times P(c=1)}{P(X)}$$

Numitorul este același. Pot renunța la el, nu mă interesează valoarea exactă a probababilității ci care probabilitate e mai mare.

....

$$P(c=9/X)=\frac{P(X/c=9)\times P(c=9)}{P(X)}$$

Presupunem că am calculat cele 10 probabilități a-posteriori: P(c=0|X), P(c=1|X), … P(c=9|X)

Regula de clasificare: $$c^{*}=\arg\max_{i=0,1,\ldots,9} P(c=i/X)$$

alege clasa care maximizează probabilitatea a-posteriori

Regula de clasificare

$\propto$ **- direct proporțional**

$$P(c=0/X) \propto P(X/c=0) \times P(c=0)$$

$$P(c=1/X) \propto P(X/c=1) \times P(c=1)$$

....

$$P(c=9/X) \propto P(X/c=9) \times P(c=9)$$

Presupunem că am calculat cele 10 probabilități:

Regula de clasificare: $c^* = \underset{i=0,1,\ldots,9}{\operatorname{argmax}} P(X/c=i) \times P(c=i)$

alege clasa care maximizează numărătorul

Calculul probabilității a-priori

$$P(c=i\,/\,X) \propto P(X\,/\,c=i) \times \boxed{P(c=i)}$$

Probabilitatea a-priori

- calculăm probabilitatea a-priori numărând câte exemple din mulțimea de antrenare au clasa *I*

- lucrați la laborator cu setul de date MNIST, mulțimea de antrenare are 1000 de exemple etichetate

```python
#compute p_C
p_C = np.zeros(10,'uint8')
for label in train_labels:
    p_C[label] = p_C[label]+1
print(p_C/sum(p_C))
```

```
[0.091 0.12  0.096 0.099 0.105 0.083 0.098 0.104 0.101 0.103]
```

P(c=0) **P(c=9)**

Calculul probabilității likelihood

$$P(c=i\,/\,X) \propto \boxed{P(X\,/\,c=i)} \times P(c=i)$$

Probabilitatea likelihood

- $P(X|c=i) = P(X_1=x_1, X_2=x_2, \ldots, X_{784}=x_{784}|c=i)$

Pentru exemplul de mai devreme vrem să calculăm probabilitatea **comună**:

- $P(X|c=i) = P(X_1=0, X_2=0, \ldots, X_{784}=253|c=i)$

Calculul acestei probabilități este practic imposibil, nu avem așa de multe date de antrenare (avem numai 1000 cu toate clasele, în jur de 100 per clasă) care să cuprindă toate n-uplurile posibile

Calculul probabilității likelihood folosind clasificatorul naïve Bayes

$$P(c=i/X) \propto \boxed{P(X/c=i)} \times P(c=i)$$

Probabilitatea likelihood

- **clasificatorul naïve Bayes consideră caracteristicile independente (nu e întotdeauna adevărat acest lucru)**

$$P(X/c=i) = \prod_{j=1}^{n=784} P(X_j = x_j / c = i)$$

- $P(X|c = i) = P(X_1{=}x_1, X_2{=}x_2, \ldots, X_{784}{=}x_{784}|c = i) =$ (folosind presupunerea de independență între caracteristici) $=$

$= P(X_1{=}x_1|c{=}i) \times P(X_2{=}x_2|c{=}i) \times \ldots P(X_{784}{=}x_{784}|c = i)$

Regula de clasificare pentru clasificatorul naïve Bayes

$$P(c=i\,/\,X) \propto P(X\,/\,c=i) \times P(c=i)$$

Regula de clasificare:

$$c^{*} = \underset{i=0,1,\ldots,9}{\operatorname{argmax}}\, P(X\,/\,c=i) \times P(c=i)$$

Regula de clasificare pentru naïve Bayes:

$$c^{*} = \underset{i=0,1,\ldots,9}{\operatorname{argmax}}\left(\prod_{j=1}^{n=784} P(X_j = x_j\,/\,c=i)\right) \times P(c=i)$$

Regula de clasificare pentru clasificatorul naïve Bayes

$$P(c = i / X) \propto P(X / c = i) \times P(c = i)$$

Regula de clasificare:

$$c^{*} = \underset{i=0,1,\ldots,9}{\operatorname{argmax}} P(X / c = i) \times P(c = i)$$

Regula de clasificare pentru naïve Bayes:

$$c^{*} = \underset{i=0,1,\ldots,9}{\operatorname{argmax}} \left(\prod_{j=1}^{n=784} \underbrace{P(X_j = x_j / c = i)}_{\text{Cum o calculez?}} \right) \times P(c = i)$$

Calculul probabilității individuale $P(X_j = x_j|c=i)$

$$c^* = \underset{i=0,1,\ldots,9}{\operatorname{argmax}}\Big(\prod_{j=1}^{n} \underbrace{P(X_j = x_j / c = i)}_{\text{Cum o calculez?}}\Big) \times P(c = i)$$

$P(X_j = 0|c) = ?$
$P(X_j = 1|c) = ?$
$P(X_j = 2|c) = ?$
…
$P(X_j = 255|c) = ?$

Nu avem date să estimăm corect asemenea probabilități, avem în jur de 100 de puncte X_j în mulțimea de antrenare.

```
[0.091 0.12  0.096 0.099 0.105 0.083 0.098 0.104 0.101 0.103]
```

P(c=0) **P(c=9)**

Calculul probabilității individuale $P(X_j = x_j|c=i)$

Exemplu: consider j = 370, c = 0
(poziția 370 în vectorul de 784 de componente, cifra 0)

```python
#calculeaza p(Xj = xj/c)
c = 0
index = np.ravel(np.where(train_labels == 0))
print(indice)
j = 370
features = train_images[index,j]
print(features.shape)
print(features.size)
print(features)
```

```
[ 15  28  29  33  52  79  86  87  98 113 122 124 141 148 173 176 186 226
 230 232 237 244 247 255 258 265 279 283 287 289 296 315 320 322 332 333
 335 367 378 383 388 397 400 411 446 461 466 479 482 493 498 509 517 531
 544 552 567 569 573 611 640 643 645 660 677 684 689 715 728 732 760 789
 800 806 825 828 835 839 875 880 881 882 899 915 919 932 943 963 972 974
 981]
(91,)
91
[ 25.   0. 120.   0.  86. 253. 243. 218.   0.   0. 253.   0. 252. 229.
   0.   0. 252.   0.   0.   0.   0.   0.   0. 250.   0. 242.   0. 253.
 254.   0. 254. 252.   0.   0.   0.  55.   0.   0.   0. 250.   0.   0.
   0.   0.   0.   0. 254.  11. 255.   0.   0.   0. 254. 252. 253.   0.
 255.   0.   0.   0.   0.   0. 103.  32.  63.  61.  37.   0.   0. 166.
   0.   0.   0.  84. 253.  25. 253.   0.   0.   0. 182.  95. 156. 225.
 191.   0.   0. 112.   0.   0.   0.]
```

Indecșii celor 91 de imagini de antrenare ce conțin cifra 0

Cele 91 de valori ale pixelului 370 pentru exemplele de antrenare cu cifra 0

Calculul probabilității individuale $P(X_j = x_j|c=i)$

Exemplu: consider j = 370, c = 0
(poziția 370 în vectorul de 784 de componente, cifra 0)

```
[ 25.    0. 120.    0.   86. 253. 243. 218.    0.    0. 253.    0. 252. 229.
   0.    0. 252.    0.    0.    0.    0.    0.    0. 250.    0. 242.    0. 253.
 254.    0. 254. 252.    0.    0.    0.   55.    0.    0.    0. 250.    0.    0.
   0.    0.    0.    0. 254.   11. 255.    0.    0.    0. 254. 252. 253.    0.
 255.    0.    0.    0.    0.    0. 103.   32.   63.   61.   37.    0.    0. 166.
   0.    0.    0.   84. 253.   25. 253.    0.    0.    0. 182.   95. 156. 225.
 191.    0.    0. 112.    0.    0.    0.]
```

Cele 91 de valori ale pixelului 370 pentru exemplele de antrenare cu cifra 0

```
[50.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.  0.
  0.  0.  0.  0.  0.  0.  0.  2.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.
  0.  1.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.
  0.  1.  0.  0.  0.  0.  0.  1.  0.  1.  0.  0.  0.  0.  0.  0.  0.  0.
  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  1.  0.  1.  0.  0.  0.
  0.  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.  0.
  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.
  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.
  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.
  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.
  0.  0.  1.  0.  0.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.  0.  0.  0.
  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.
  0.  0.  1.  0.  0.  0.  0.  0.  0.  1.  0.  0.  0.  1.  0.  0.  0.  0.
  0.  0.  0.  0.  0.  0.  0.  0.  1.  1.  0.  0.  0.  0.  0.  0.  2.  0.
  4.  6.  4.  2.]
```

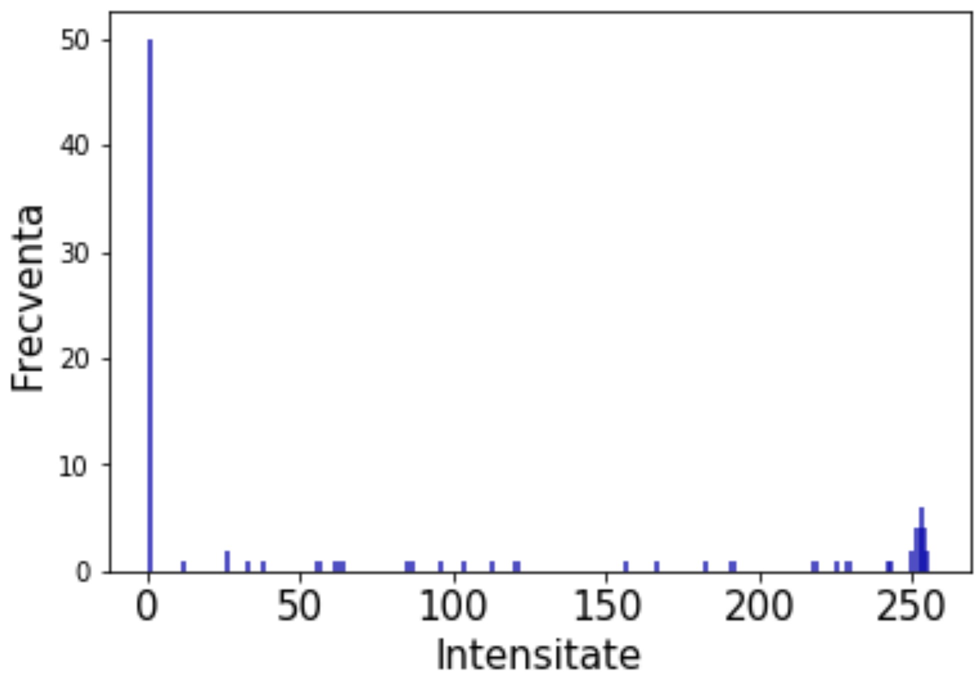
Vector de frecvență pentru fiecare intensitate

Calculul probabilității individuale $P(X_j = x_j|c=i)$

Exemplu: consider j = 370, c = 0
(poziția 370 în vectorul de 784 de componente, cifra 0)

```
[ 25.   0. 120.   0.  86. 253. 243. 218.   0.   0. 253.   0. 252. 229.
   0.   0. 252.   0.   0.   0.   0.   0.   0. 250.   0. 242.   0. 253.
 254.   0. 254. 252.   0.   0.   0.  55.   0.   0.   0. 250.   0.   0.
   0.   0.   0.   0. 254.  11. 255.   0.   0.   0. 254. 252. 253.   0.
 255.   0.   0.   0.   0.   0. 103.  32.  63.  61.  37.   0.   0. 166.
   0.   0.   0.  84. 253.  25. 253.   0.   0.   0. 182.  95. 156. 225.
 191.   0.   0. 112.   0.   0.   0.]
```

Cele 91 de valori ale pixelului 370 pentru exemplele de antrenare cu cifra 0



Vector de frecvență pentru fiecare intensitate

Calculul probabilității individuale $P(X_j = x_j|c=i)$

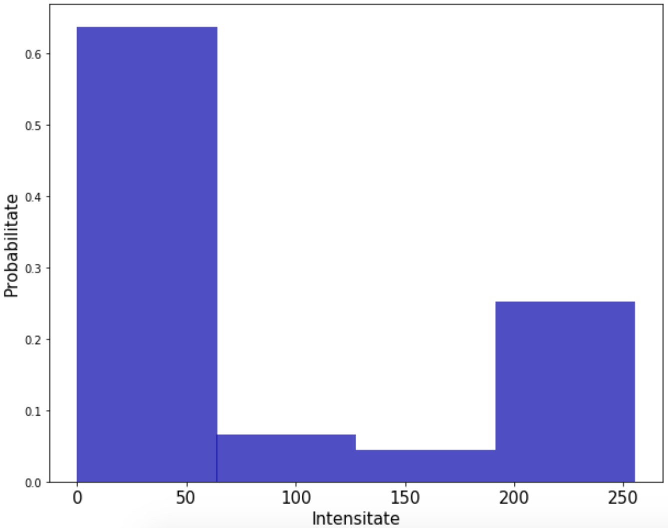
Nu avem date să estimăm corect asemenea probabilități, avem în jur de 100 de puncte X_j în mulțimea de antrenare.

O posibilă soluție:
Împart intervalul de valori posibile [0,255] în p părți egale și estimez care este probabilitatea ca o valoare să apară într-un astfel de interval.

Exemplu: consider p = 4, obțin 4 intervale:
[0, 63], [64, 127], [128, 191], [192,255]

Aproximez:
$P(X_j = 0|c)$ cu $P(X_j \in [0,63]|c)$
$P(X_j = 100|c)$ cu $P(X_j \in [64,127]|c)$

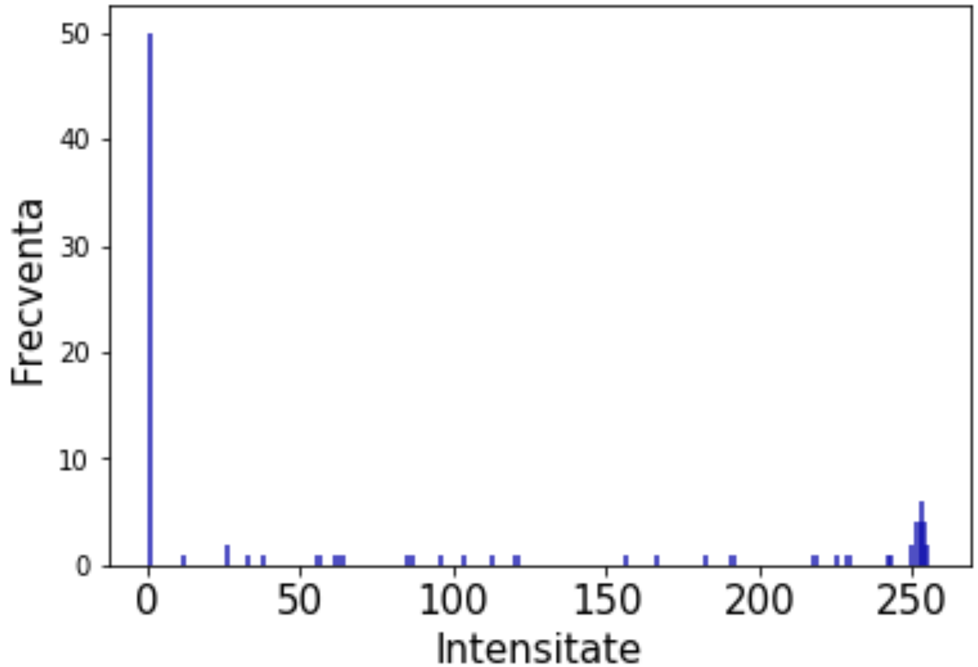
Probabilitatea va arată astfel:



Calculul probabilității individuale $P(X_j = x_j|c=i)$

Nu avem date să estimăm corect asemenea probabilități, avem în jur de 100 de puncte X_j în mulțimea de antrenare.

Soluție alternativă:
Încerc să fitez o distribuție parametrică (distribuție normală, etc.) care să se potrivească cu datele.



Stabilitate numerică

Regula de clasificare pentru naïve Bayes:

$$c^* = \operatorname*{argmax}_{i=0,1,\dots,9}\left(\left(\prod_{j=1}^{n=784} P(X_j = x_j \,/\, c = i)\right) \times P(c = i)\right)$$

înmulțesc n=784 numere subunitare (sunt probabilități), uneori foarte aproape de 0.

Aplic funcția logaritm (e monoton crescătoare), păstrează ierarhia:

$$c^* = \operatorname*{argmax}_{i=0,1,\dots,9}\left(\log\left(\left(\prod_{j=1}^{n=784} P(X_j = x_j \,/\, c = i)\right) \times P(c = i)\right)\right)$$

$$c^* = \operatorname*{argmax}_{i=0,1,\dots,9}\left(\sum_{j=1}^{n=784} \log(P(X_j = x_j \,/\, c = i)) + \log(P(c = i))\right)$$

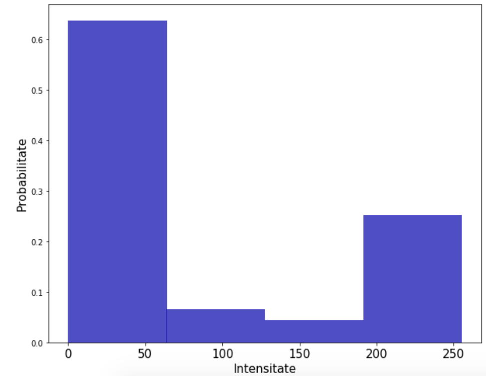
Laborator – Matrice confuzie

	0	1	2	3	4	5	6	7	8	9
0	[[48	0	1	0	0	0	3	0	1	0]
1	[0	46	1	0	0	1	0	0	4	0]
2	[1	0	52	3	1	0	0	1	0	0]
3	[0	0	4	43	1	1	0	0	1	3]
4	[0	0	0	0	40	1	0	0	2	3]
5	[1	0	2	11	1	31	0	0	6	0]
6	[1	0	2	0	2	1	42	0	0	0]
7	[0	0	1	0	2	0	0	38	1	8]
8	[0	0	2	4	2	3	1	1	32	1]
9	[0	0	1	1	12	0	0	1	1	26]]

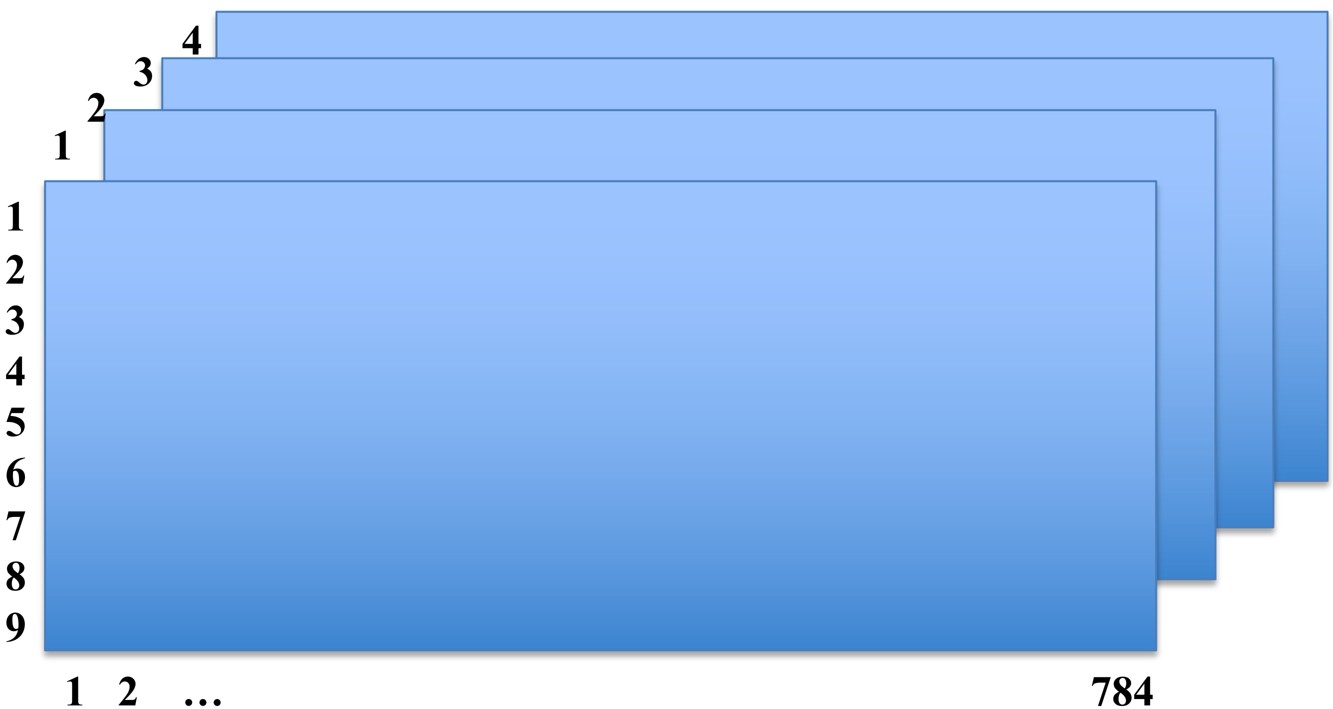
Numărul de exemple de testare cu cifra 9 care au fost clasificate drept cifra 4

Model parametric

Care sunt parametri învățați de către model?



Modelul învăță din datele de antrenare pentru fiecare clasă i (linie) și componentă j (coloană) $P(X_j = x_j|c=i)$ sub forma unui vector de probabilități de dimensiune 4



Numărul de parametri:
784 × 10 × 4 = 31360

De fapt întrucât fiecare vector de dimensiune 4 are suma 1:
784 × 10 × 3 = 24520

Hiperparametrul p = numărul de părți, p = 4

Clasificatorul Bayes pe cazul general

- X vector de n caracteristici $X = (X_1, X_2, \ldots, X_n)$
- c clase (admis/respins, bărbat/femeie,litere, cifre, etc.)
- Regula lui Bayes:

$$P(c / X) = \frac{P(X / c) \times P(c)}{P(X)}$$

$$Posterior = \frac{Likelihood \times Prior}{Evidence}$$

$$P(c = i / X) \propto P(X / c = i) \times P(c = i)$$

- **clasificatorul naïve Bayes consideră caracteristicile independente (nu e întotdeauna adevărat acest lucru)**

$$P(X / c = i) = \prod_{j=1}^{n} P(X_j = x_j / c = i)$$

Clasificatorul Bayes pe cazul general

- X vector de n caracteristici $X = (X_1, X_2, \ldots, X_n)$
- c clase (admis/respins, litere, cifre, etc.)
- Regula de clasificare pentru Naïve Bayes:

$$c^* = \underset{i}{\operatorname{argmax}}\Big(\prod_{j=1}^{n} \underbrace{P(X_j = x_j / c = i)}_{\text{trebuie estimată din date}}\Big) \times P(c = i)$$

- stabilitate numerică:

$$c^* = \underset{i}{\operatorname{argmax}}\Big(\log\Big(\Big(\prod_{j=1}^{n} P(X_j = x_j / c = i)\Big) \times P(c = i)\Big)\Big)$$

$$c^* = \underset{i}{\operatorname{argmax}}\Big(\sum_{j=1}^{n} \log(P(X_j = x_j / c = i)) + \log(P(c = i))\Big)$$