

## Bloc 2

### Aprenentatge Automàtic

## Pràctica 2:

# Aplicació dels algoritmes de Perceptró i Regressió Logística a diversos conjunts de dades

# Objectiu

*Generar i evaluar un classificador sobre datasets (conjunts de dades)*

- Sessions:
  - S1: 22/11
  - S2: 29/11
  - S3: 05/12
  - Examen P2: 13/12

## Sesions de la pràctica 2

### Sessió 1:

- Familiaritzar-se amb l'entorn de treball (Google Colab)
- Analitzar conjunts de dades (datasets): iris, digits, olivetti, openml

### Sessió 2:

- Aplicació de l'algorisme del Perceptró a tasques de classificació: conjunt de dades iris.
- **Exercici**: Aplicar el Perceptró a digits i olivetti.

### Sessió 3:

- Aplicació de la Regressió Logística a tasques de classificació: conjunt de dades iris.

### Exemple d'examen:

- Aplicació de Perceptró i Regressió Logística a un conjunt de dades de OpenML.

### Sessió 4 (**examen**):

- Es demanarà l'aplicació del Regressió Logística per a una tasca diferent d'OpenML.
- Caldrà pujar la solució de **l'Exercici**.

# Sessió 1: entorn de treball i conjunts de dades

- Entorn de treball: Google Colab (<https://colab.research.google.com>)
  - S'hi treballa amb Quaderns "Notebook" (codi + text)
  - Similar a un Notebook de Jupyter
- **Dataset:** col·lecció de dades que s'utilitza per entrenar, provar i validar models d'aprenentatge automàtic.
  - Un conjunt de dades típicament consisteix en diverses "instàncies" o "exemples", cadascun dels quals inclou diverses "característiques" o "atributs".
  - Cada instància en el conjunt de dades també té una "etiqueta" o "objectiu", que és el valor que el model d'aprenentatge automàtic intenta predir.

# Sessió 1: entorn de treball i conjunts de dades

- Ex dataset: iris

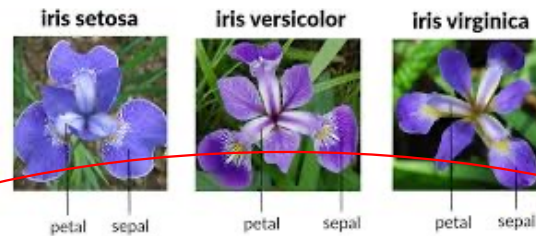


	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...	...	...	...	...	...

species: és la variable que el classificador basat en un model d'aprenentatge automàtic intentarà predir.

# Sessió 1: entorn de treball i conjunts de dades

- Ex dataset: iris



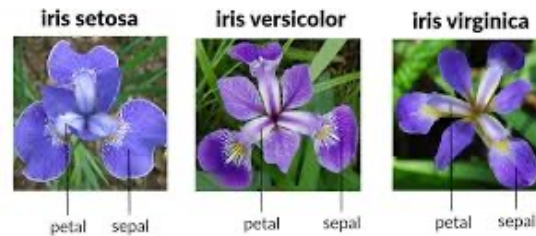
Totes les files (individus) llevat de la columna de classe s'utilitzen per entrenar el model i validar-lo.

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...	...	...	...	...	...

species: és la variable que el classificador basat en un model d'aprenentatge automàtic intentarà predir.

# Sessió 1: entorn de treball i conjunts de dades

- Ex dataset: iris



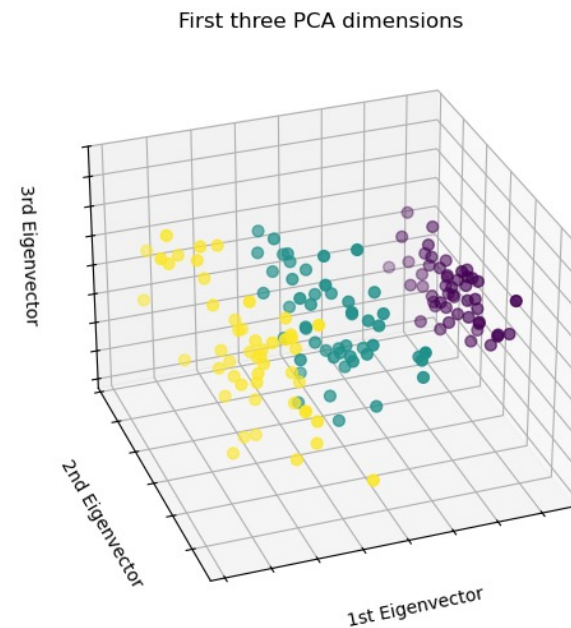
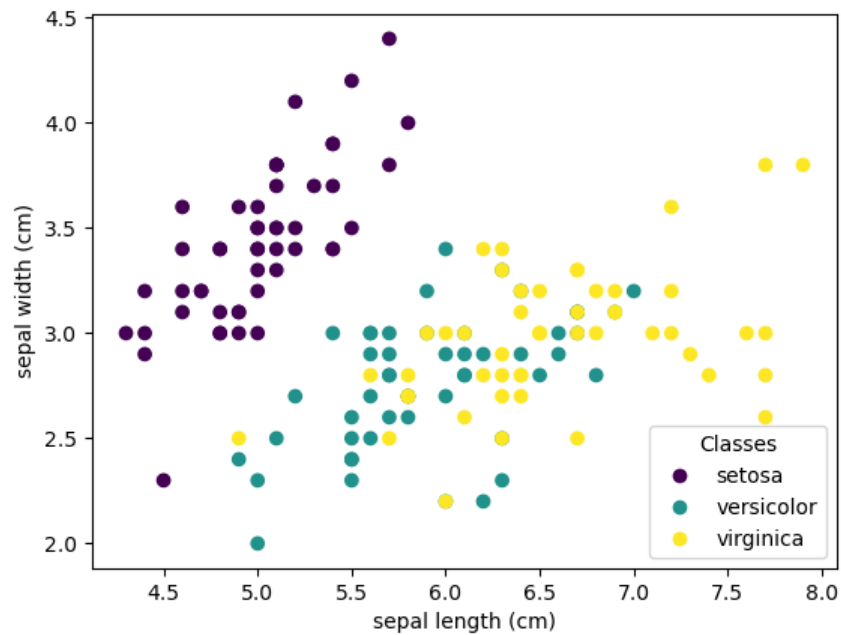
La columna de classe determina el valor a predir pel classificador quan se li dóna un nou individu.

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...	...	...	...	...	...

species: és la variable que el classificador basat en un model d'aprenentatge automàtic intentarà predir.

# Sessió 1: entorn de treball i conjunts de dades

- Ex dataset: iris





# Sessió 1: entorn de treball i conjunts de dades

- Ex dataset: iris

Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	3.6	1.4	0.2	Iris-setosa
6	5.4	3.9	1.7	0.4	Iris-setosa
7	4.6	3.4	1.4	0.3	Iris-setosa
8	5	3.5	1.5	0.2	Iris-setosa
9	4.4	2.9	1.4	0.2	Iris-setosa
10	4.9	3.1	1.5	0.1	Iris-setosa
11	5.4	3.7	1.5	0.2	Iris-setosa
12	4.6	3.4	1.6	0.2	Iris-setosa
13	4.8	3	1.4	0.1	Iris-setosa
14	4.3	3	1.1	0.1	Iris-setosa
15	5.6	4	1.2	0.2	Iris-setosa
16	5.7	4.4	1.5	0.4	Iris-setosa
17	5.4	3.9	1.3	0.4	Iris-setosa
18	5.1	3.5	1.4	0.3	Iris-setosa
19	5.7	3.8	1.7	0.3	Iris-setosa

Entrenament

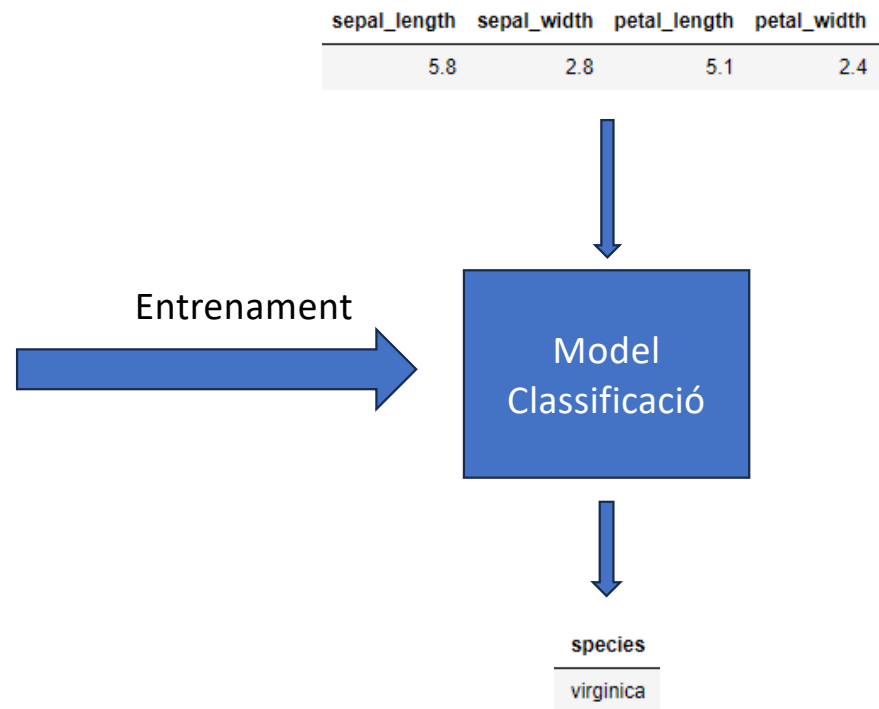


Model  
Classificació

# Sessió 1: entorn de treball i conjunts de dades

- Ex dataset: iris

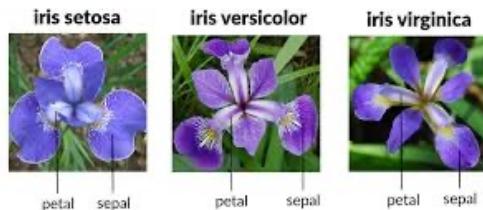
id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
1	5.1	3.5	1.4	0.2	Iris-setosa
2	4.9	3	1.4	0.2	Iris-setosa
3	4.7	3.2	1.3	0.2	Iris-setosa
4	4.6	3.1	1.5	0.2	Iris-setosa
5	5	3.6	1.4	0.2	Iris-setosa
6	5.4	3.9	1.7	0.4	Iris-setosa
7	4.6	3.4	1.4	0.3	Iris-setosa
8	5	3.5	1.5	0.2	Iris-setosa
9	4.4	2.9	1.4	0.2	Iris-setosa
10	4.9	3.1	1.5	0.1	Iris-setosa
11	5.4	3.7	1.5	0.2	Iris-setosa
12	4.6	3.4	1.6	0.2	Iris-setosa
13	4.8	3	1.4	0.1	Iris-setosa
14	4.3	3	1.1	0.1	Iris-setosa
15	5.6	4	1.2	0.2	Iris-setosa
16	5.7	4.4	1.5	0.4	Iris-setosa
17	5.4	3.9	1.3	0.4	Iris-setosa
18	5.1	3.5	1.4	0.3	Iris-setosa
19	5.7	3.8	1.7	0.3	Iris-setosa



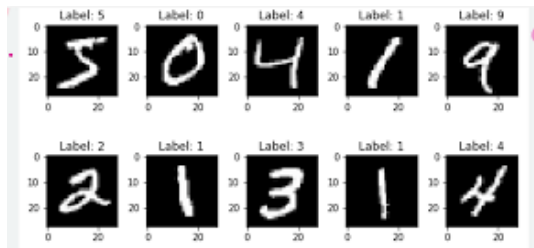
# Sessió 1: entorn de treball i conjunts de dades

- Datasets:

- IRIS



- DIGITS



- OLIVETTI



- OPENML (repositorio)



# Us de Google Colab

Pas 1: Crear un Compte de Google (si encara no en tens un)

- Per utilitzar Google Colab, necessites un compte de Google. Si ja tens un compte de Gmail, YouTube o qualsevol altre servei de Google, pots utilitzar aquest compte. Si no en tens un, segueix aquests passos per crear-ne un de nou:

- 1.Vés a [accounts.google.com](https://accounts.google.com).

- 2.Segueix les instruccions per crear un compte de Google nou.

Pas 2: Accedir a Google Colab

- Un cop tingues el teu compte de Google, pots accedir a Google Colab:

- 1.Vés a [Google Colab](https://colab.google).

- 2.Inicia la sessió amb el teu compte de Google.

# Us de Google Colab

## Pas 3: Crear un Nou Notebook

- Després d'iniciar la sessió a Google Colab, pots començar a crear notebooks:
  1. Fes clic a "Nou Notebook" a la cantonada inferior dreta de la pantalla. Això obrirà un nou notebook en una nova pestanya del navegador.
  2. Pots canviar el nom del notebook fent clic al títol (per defecte, alguna cosa com "Untitled0.ipynb") a la part superior de la pàgina i escrivint el nom que vulgues.

## Pas 4: Escriure i Executar Codi

- Google Colab et permet escriure i executar codi Python de manera interactiva:
  1. Escriu el teu codi Python en una cel·la.
  2. Per executar el codi en aquesta cel·la, prem Shift + Enter o fes clic al botó de reproducció (triangle) a la cantonada superior esquerra de la cel·la.
  3. El resultat de l'execució del codi apareixerà sota de la cel·la.

# Us de Google Colab

## Pas 5: Desar i Compartir el teu Notebook

- Google Colab desa automàticament els teus notebooks a Google Drive:
  1. Per desar manualment, ves a "Fitxer" > "Dezar".
  2. Per compartir el teu notebook, fes clic al botó "Compartir" a la cantonada superior dreta i segueix les instruccions per compartir-lo com ho faries amb qualsevol altre arxiu de Google Drive.

## Pas 6: Importar i Exportar Notebooks

- Pots importar notebooks existents o exportar els teus:
  - Importar: "Fitxer" > "Obre notebook" > tria des de GitHub, Google Drive o puja un arxiu.
  - Exportar: "Fitxer" > "Descarregar" > tria el format que prefereixis (per exemple, .ipynb per a Jupyter Notebook o .py per a un script de Python).

# Sessió 1: entorn de treball i conjunts de dades

Probar els següents notebooks en Google colab:

- iris.ipynb
- digits.ipynb
- olivetti.ipynb
  
- openml.ipynb

# Sessió 1: entorn de treball i conjunts de dades

## **Què cal saber de la sessió de hui?**

- Saber què és i quin format té un dataset.
- Saber interpretar un dataset: n<sup>o</sup> individus, n<sup>o</sup> classes, ...
- Saber carregar un dataset d'openml.
- Saber extraure les característiques i la columna objectiu d'un dataset.