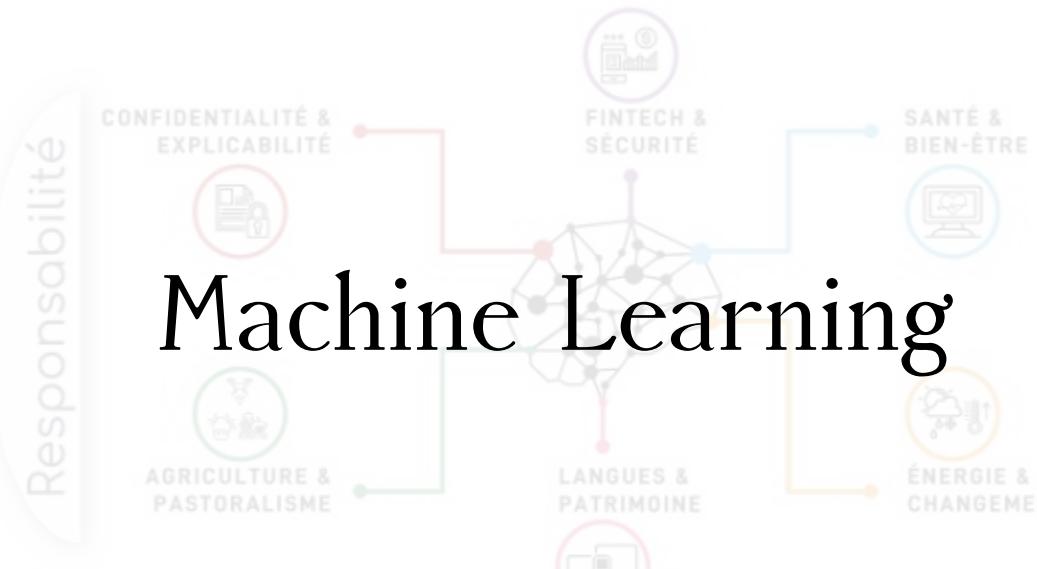


CITADEL AI Summer School 2022



Machine Learning

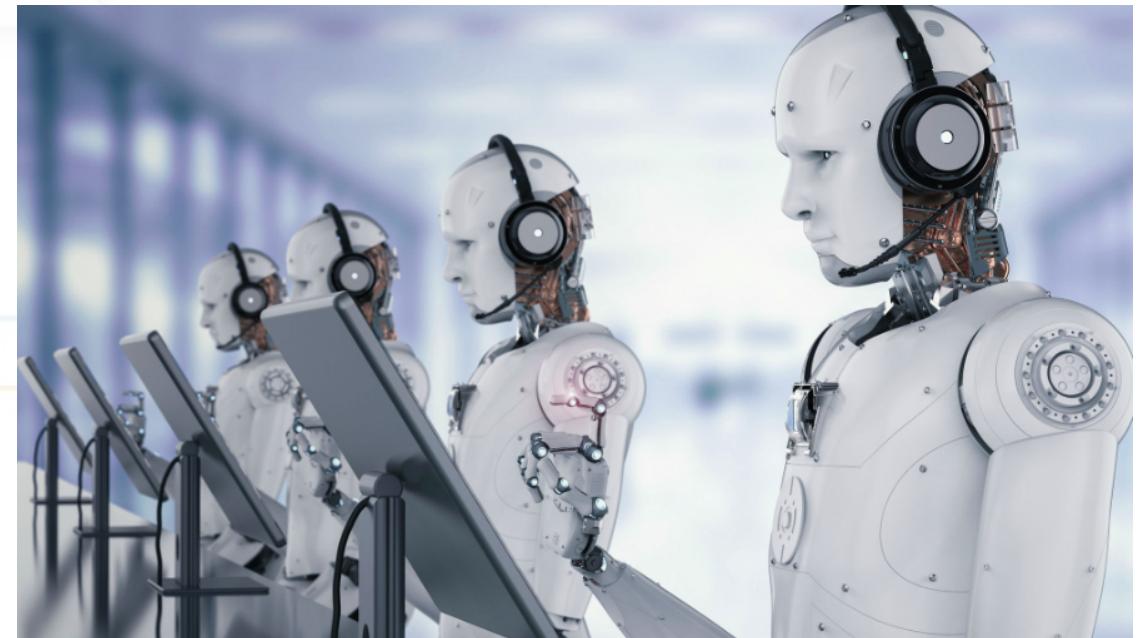
Rodrique Kafando, PhD
rodrique.kafando@citadel.bf



19-25.09.22

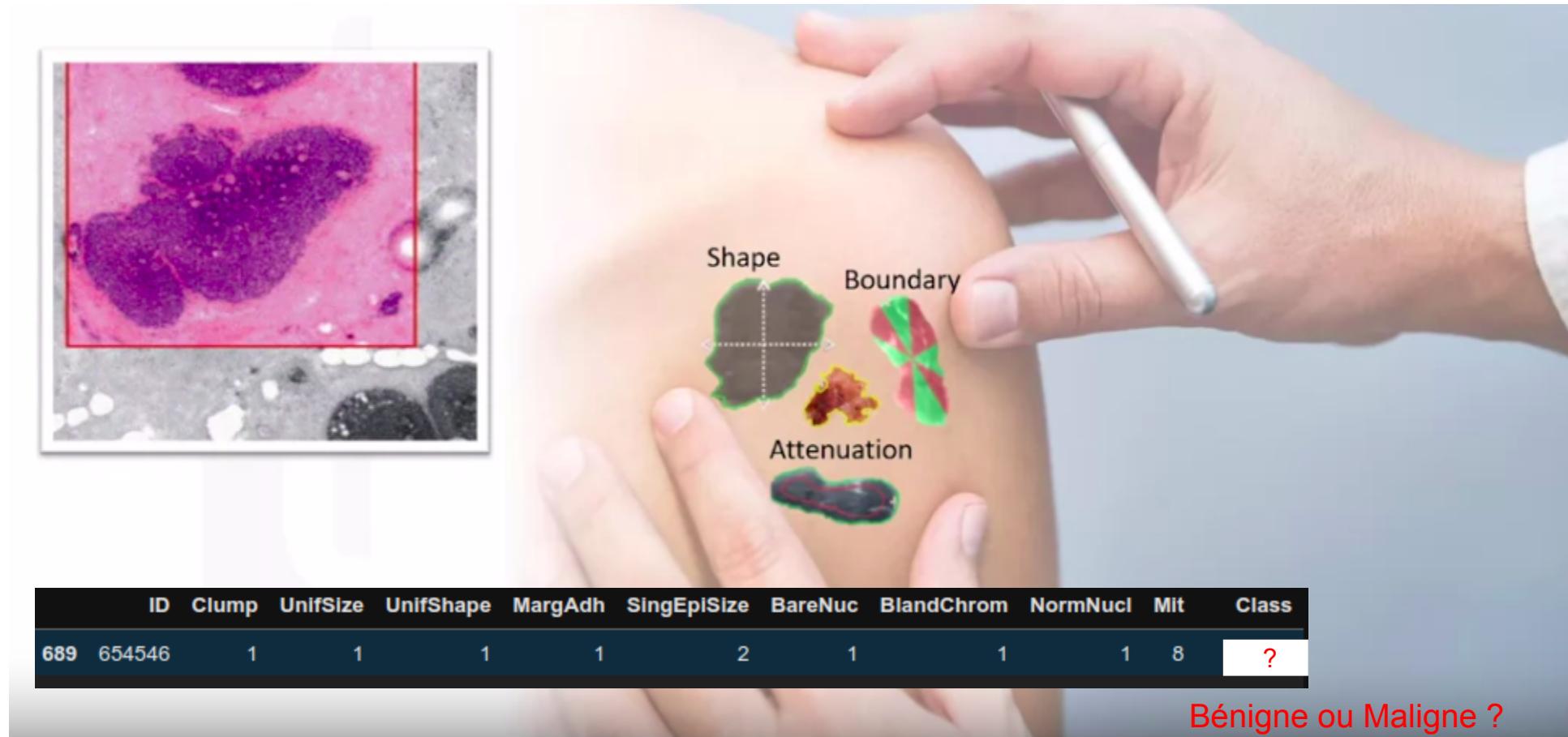


C'est quoi le Machine Learning ?



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

C'est quoi le Machine Learning ?

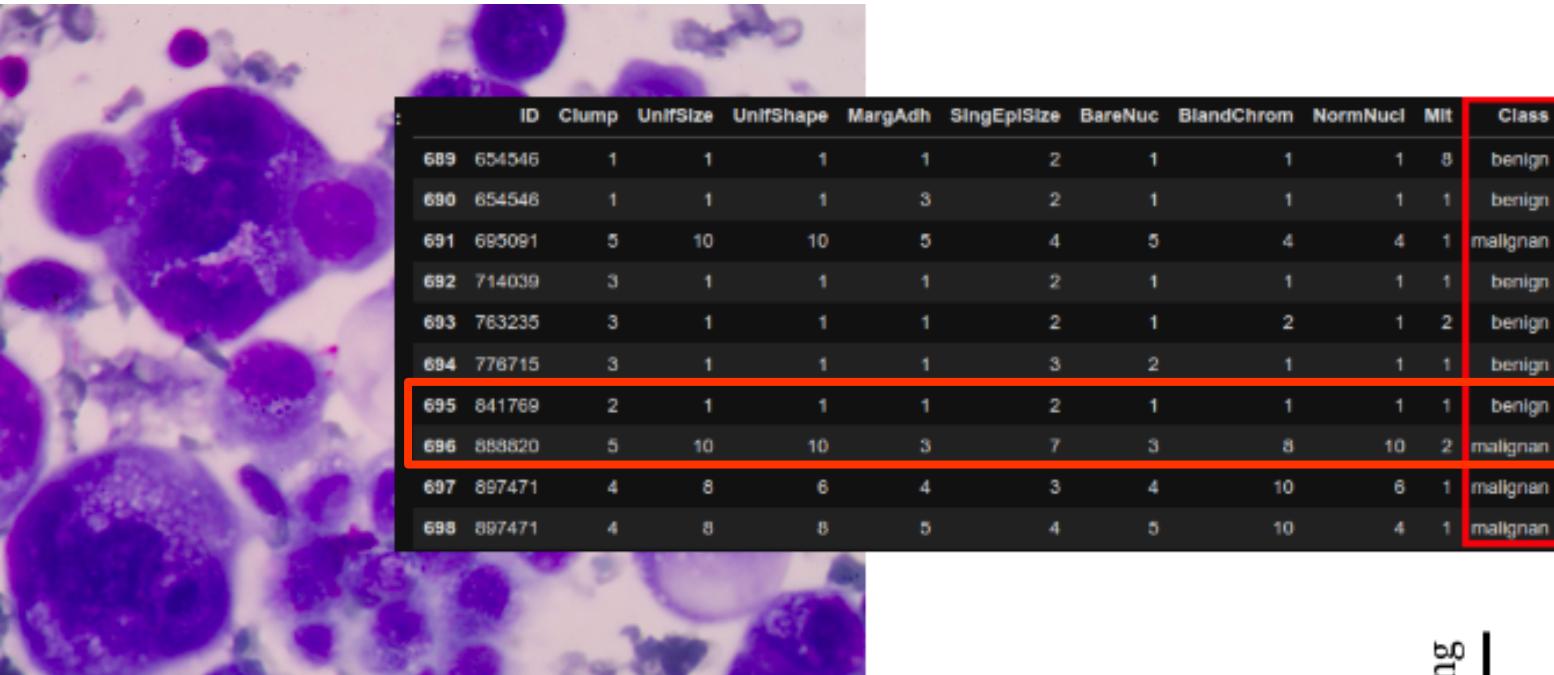


Bénigne ou Maligne ?



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

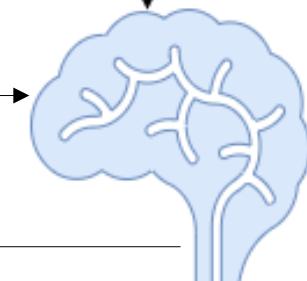
C'est quoi le Machine Learning ?



Caractéristiques distinctes

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
689	654546	1	1	1	1	2	1	1	1	benign

Modeling



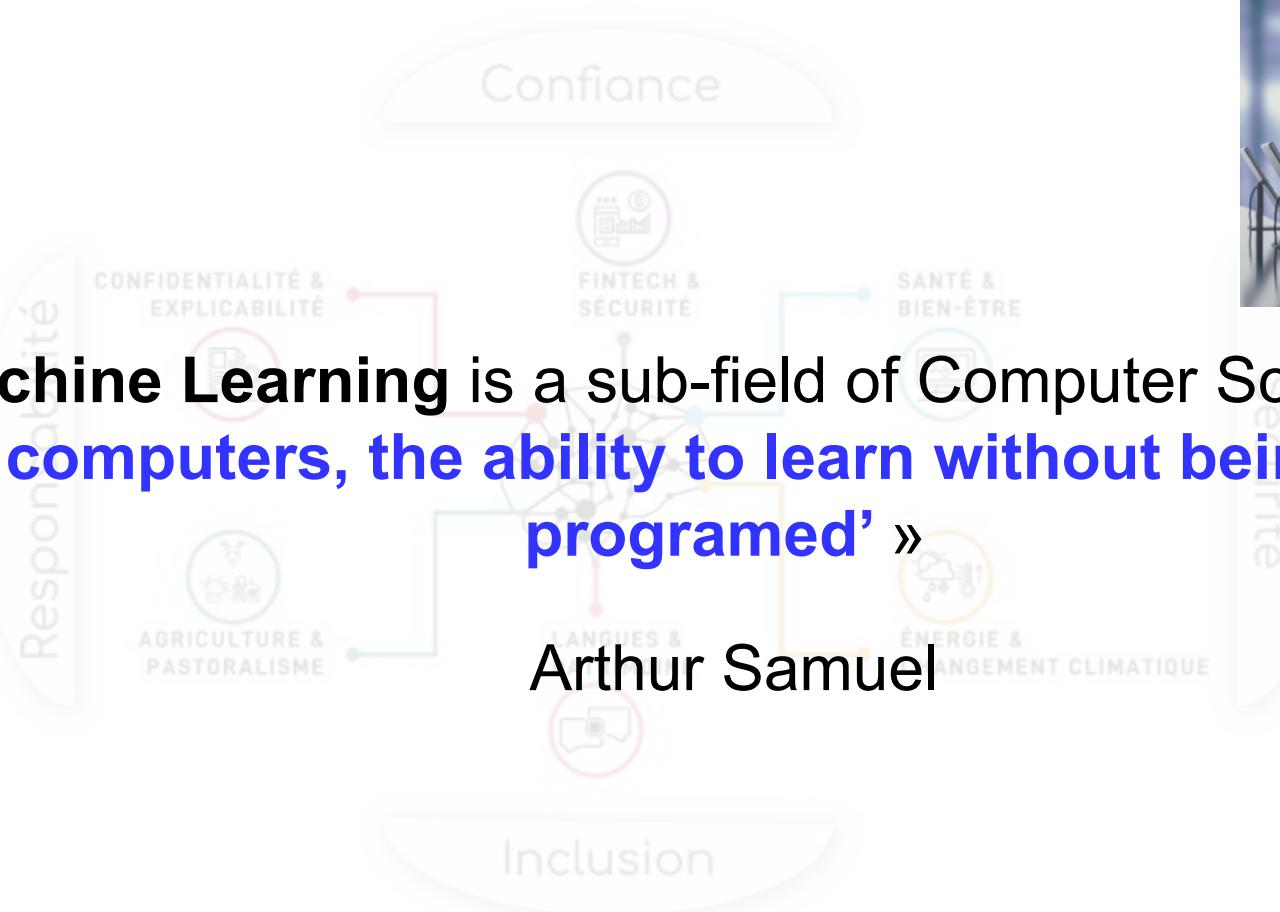
Accuracy =
98 %

... de la prédition...



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

C'est quoi le Machine Learning ?



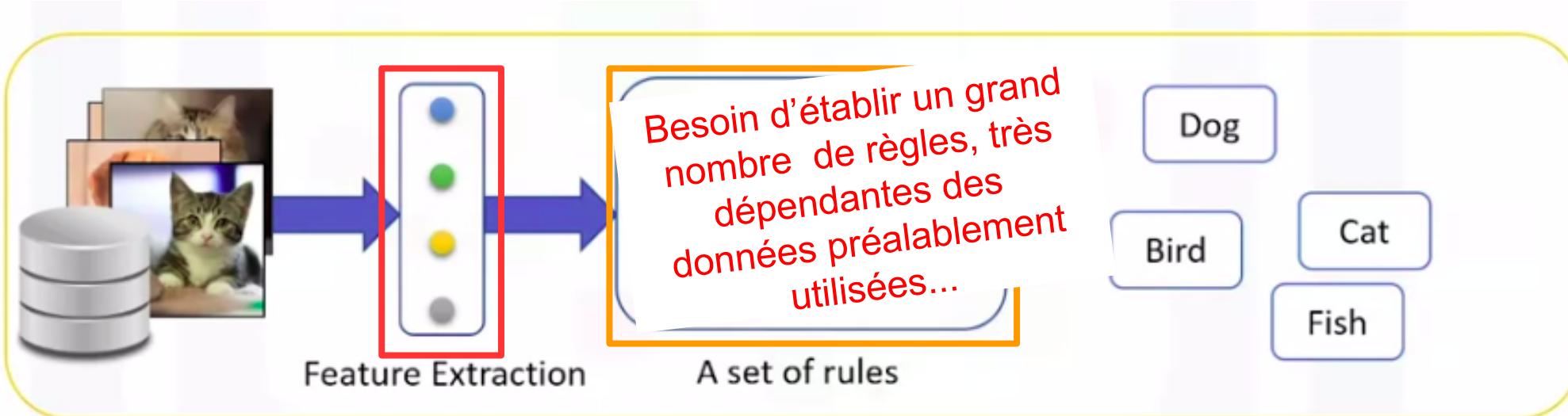
« Machine Learning is a sub-field of Computer Science, **'that gives computers, the ability to learn without being explicitly programed'** »

Arthur Samuel



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Comment ça fonctionne ?

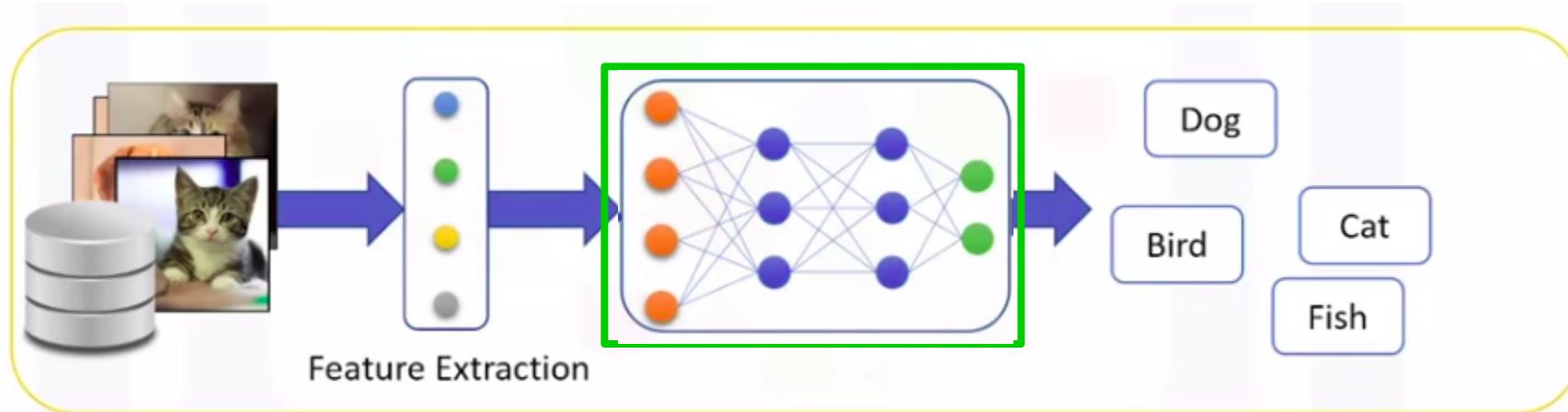


- * Chaque image a ses propres caractéristiques
- * Chaque image sera transformée en un vecteur caractéristique (vector of features)
- * Traditionnellement, chaque ur **FAILURE** gles sera établis pour représenter chacune des caractéristiques



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Comment ça fonctionne ?

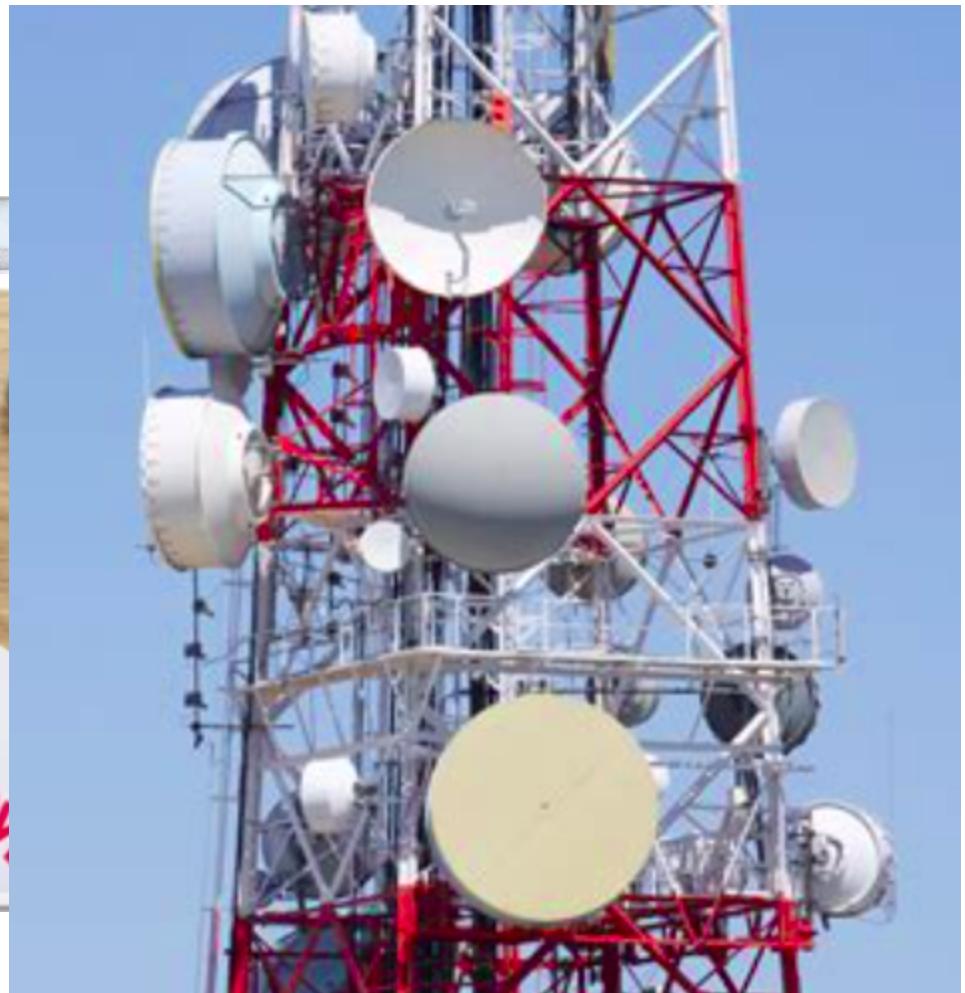
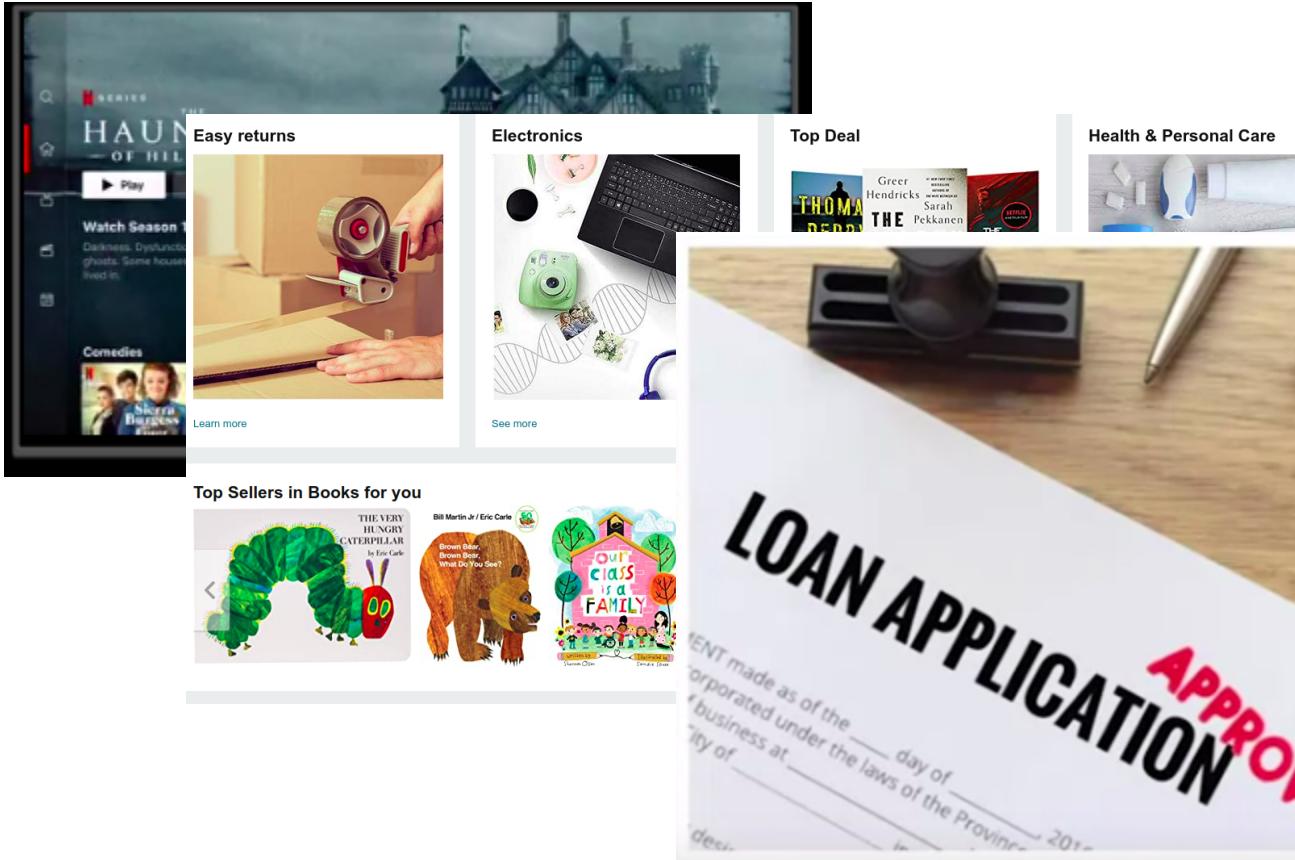


- * Chaque image a ses propres caractéristiques
- * Chaque image sera transformée en un vecteur caractéristique (vector of features)
- * Représentation de toutes les caractéristiques possibles pour chaque objet (animal)



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Quelques exemples...



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Quelques aglos populaires...

Regression/Estimation

Prédire des valeurs continues

Classification

Prédire un item d'une classe/catégorie

Clustering

Trouver des structures cachées/

Associations

Faire des résumés
Associer des patterns co-
occurents / Evenements

Anomaly detection

Déetecter des anomalies /
Comportements inhabituels

Sequence mining

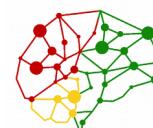
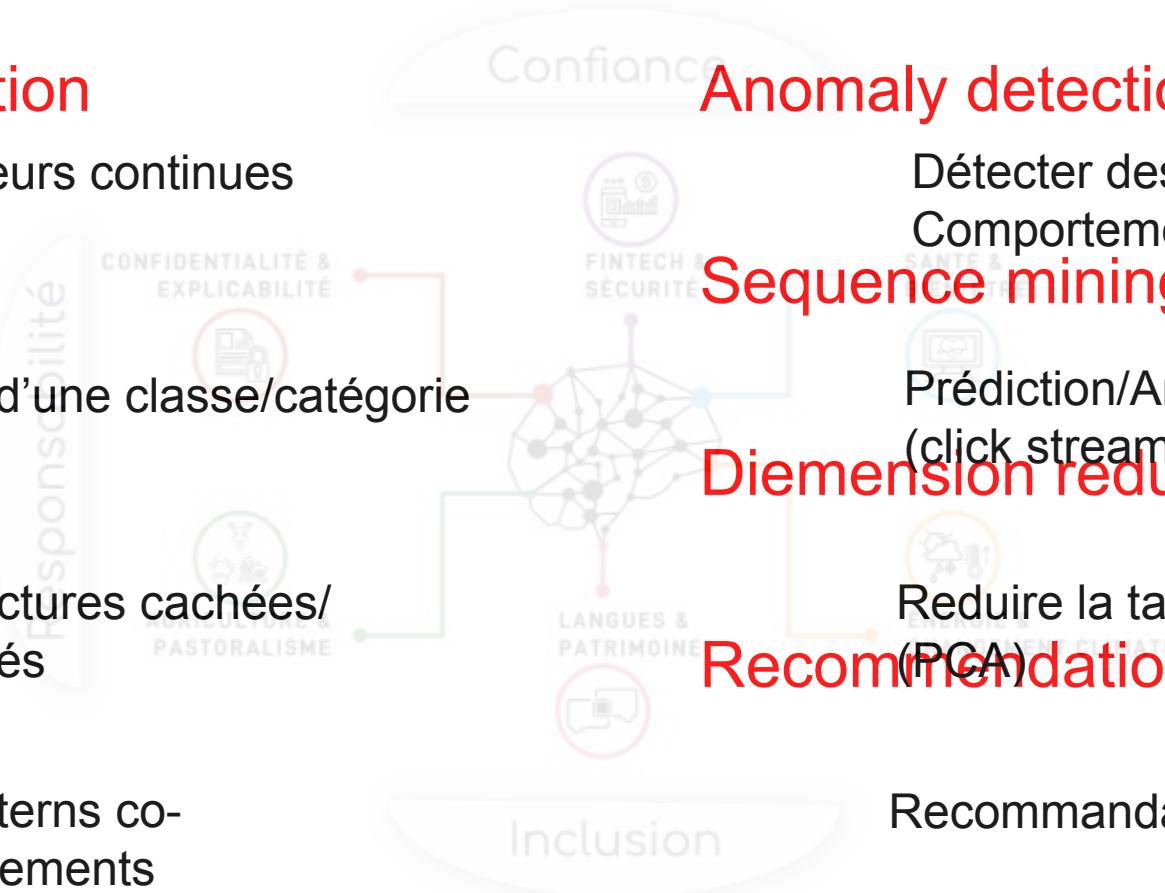
Prédiction/Anticipation d'événements
(click stream with Markov Model)

Dimension reduction

Reducire la taille des données

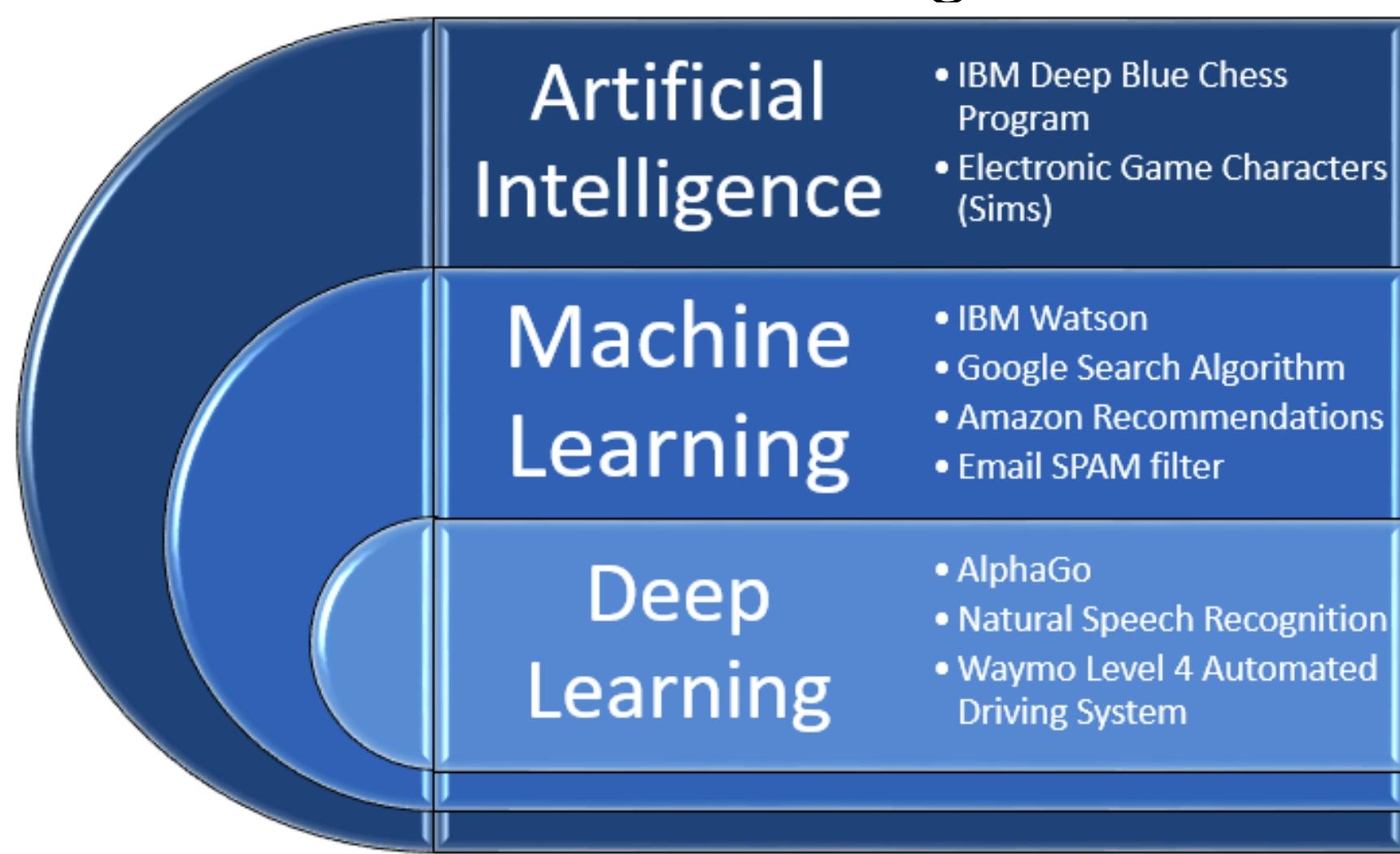
Recommendations Systems

Recommandation d'items

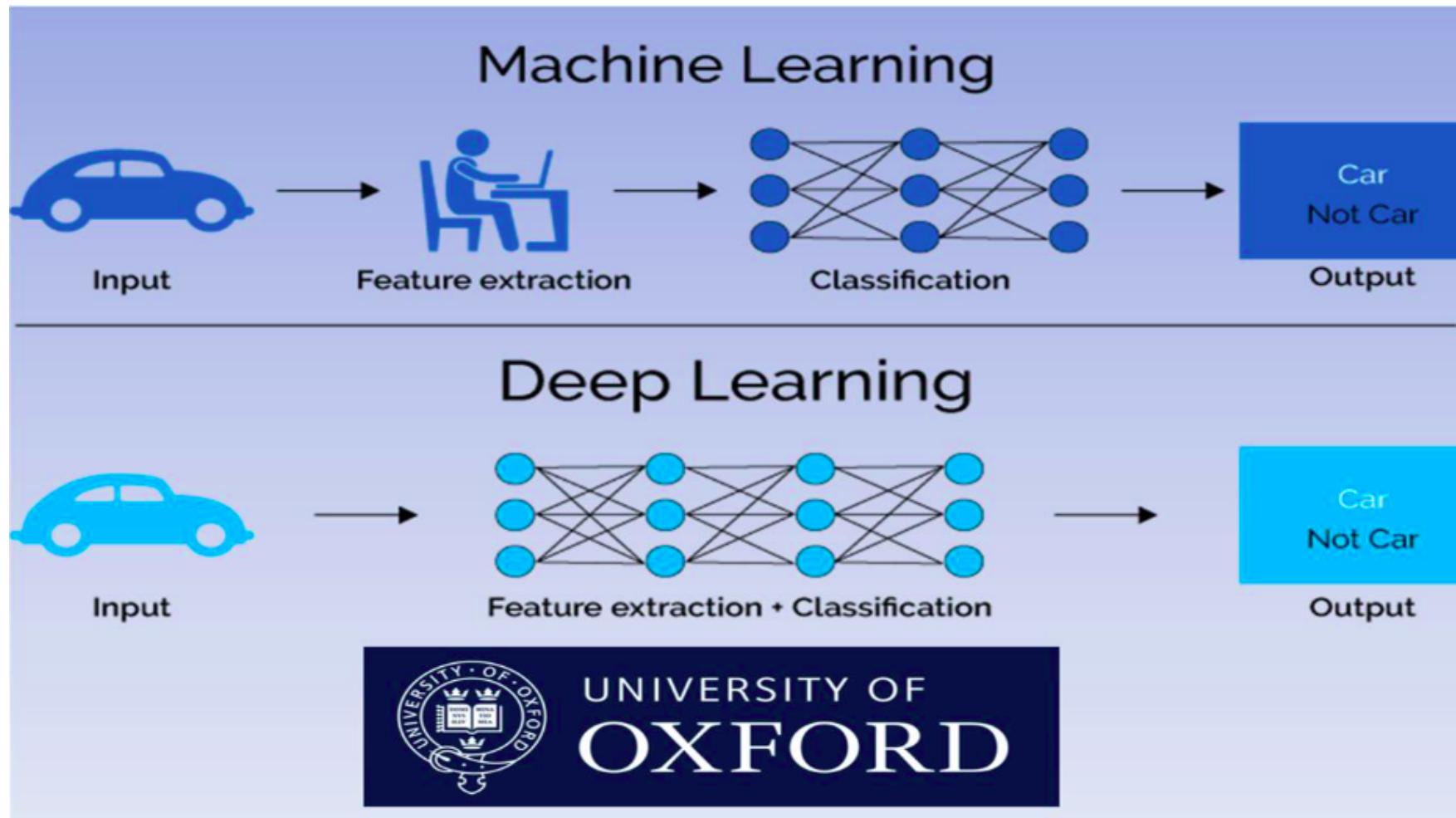


Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Quelles différences entre IA-Machine Learnig-Deep Learning ?



Quelles différences entre IA-Machine Learnig-Deep Learning ?

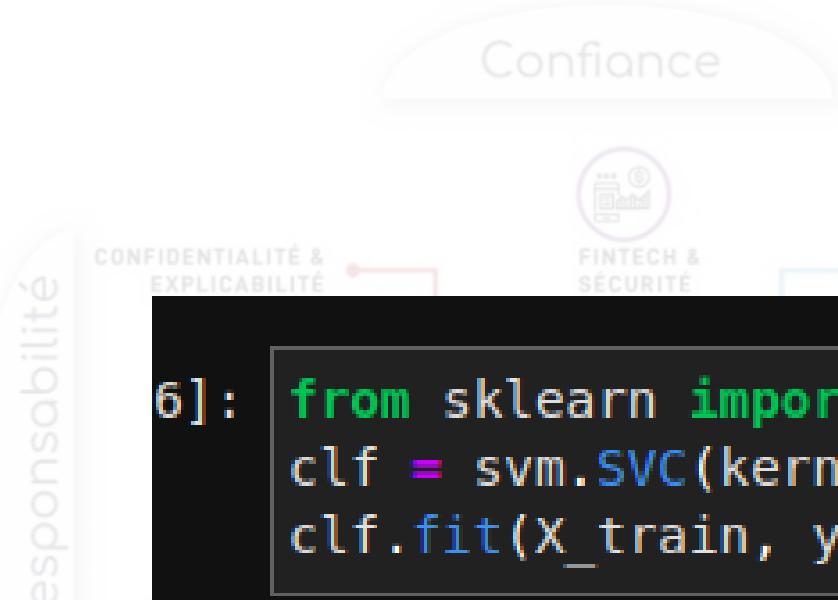


UNIVERSITY OF
OXFORD



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

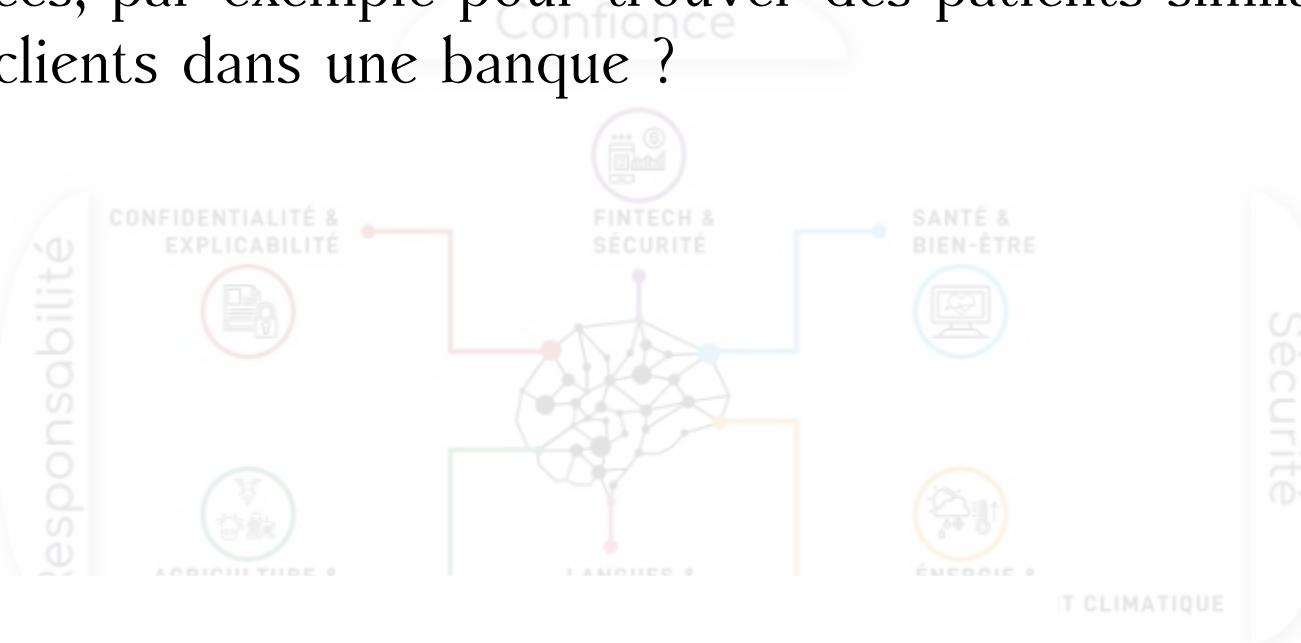
M. L & Python



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Pour terminer...

Quelle technique de M.L est appropriée pour le regroupement de cas similaires dans un ensemble de données, par exemple pour trouver des patients similaires, ou pour la segmentation des clients dans une banque ?



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

M. L & Python

CONFIDENTIALITÉ &
EXPLICABILITÉ

FINTECH &
SÉCURITÉ

SANTÉ &
BIEN-ÊTRE

Responsabilité

Confiance



```
6]: from sklearn import svm  
clf = svm.SVC(kernel='rbf')  
clf.fit(X_train, y_train)  
  
6]: SVC()
```



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

M. L & Python

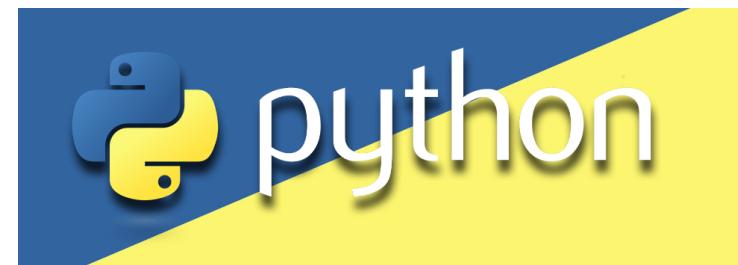
pandas
 $y_i t = \beta' x_{it} + \mu_i + \epsilon_{it}$



NumPy

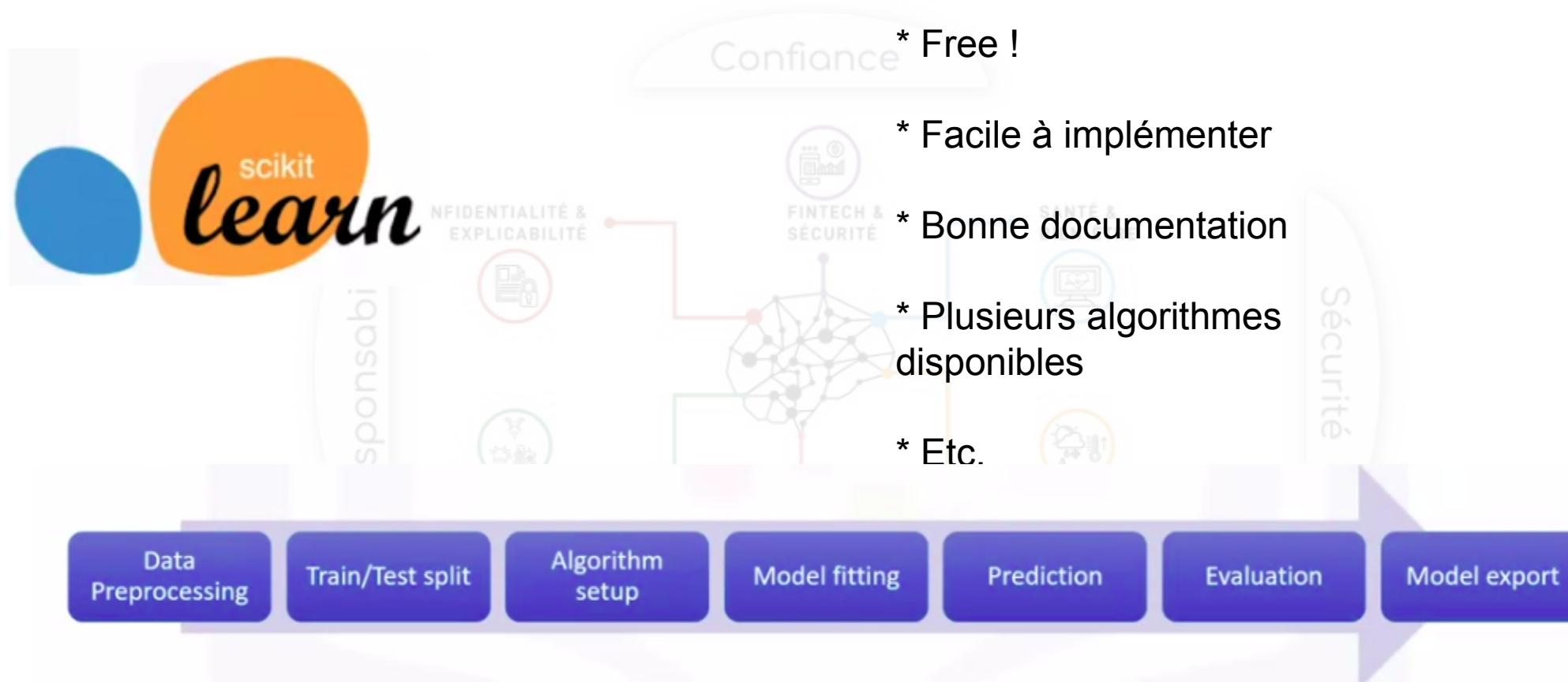
SciPy

matplotlib



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

M. L & Python



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

M. L & Python



```
from sklearn import preprocessing  
X = preprocessing.StandardScaler().fit(X).transform(X)
```

```
from sklearn.model_selection import train_test_split  
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.33)
```

```
from sklearn import svm  
clf = svm.SVC(gamma=0.001, C=100.)
```

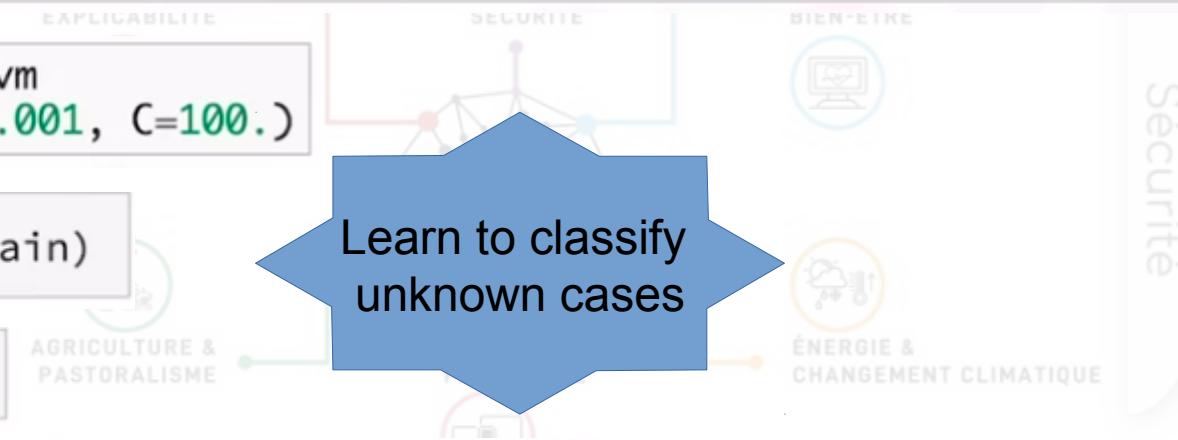
```
clf.fit(X_train, y_train)
```

```
clf.predict(X_test)
```

```
from sklearn.metrics import confusion_matrix  
print(confusion_matrix(y_test, yhat, labels=[1,0]))
```

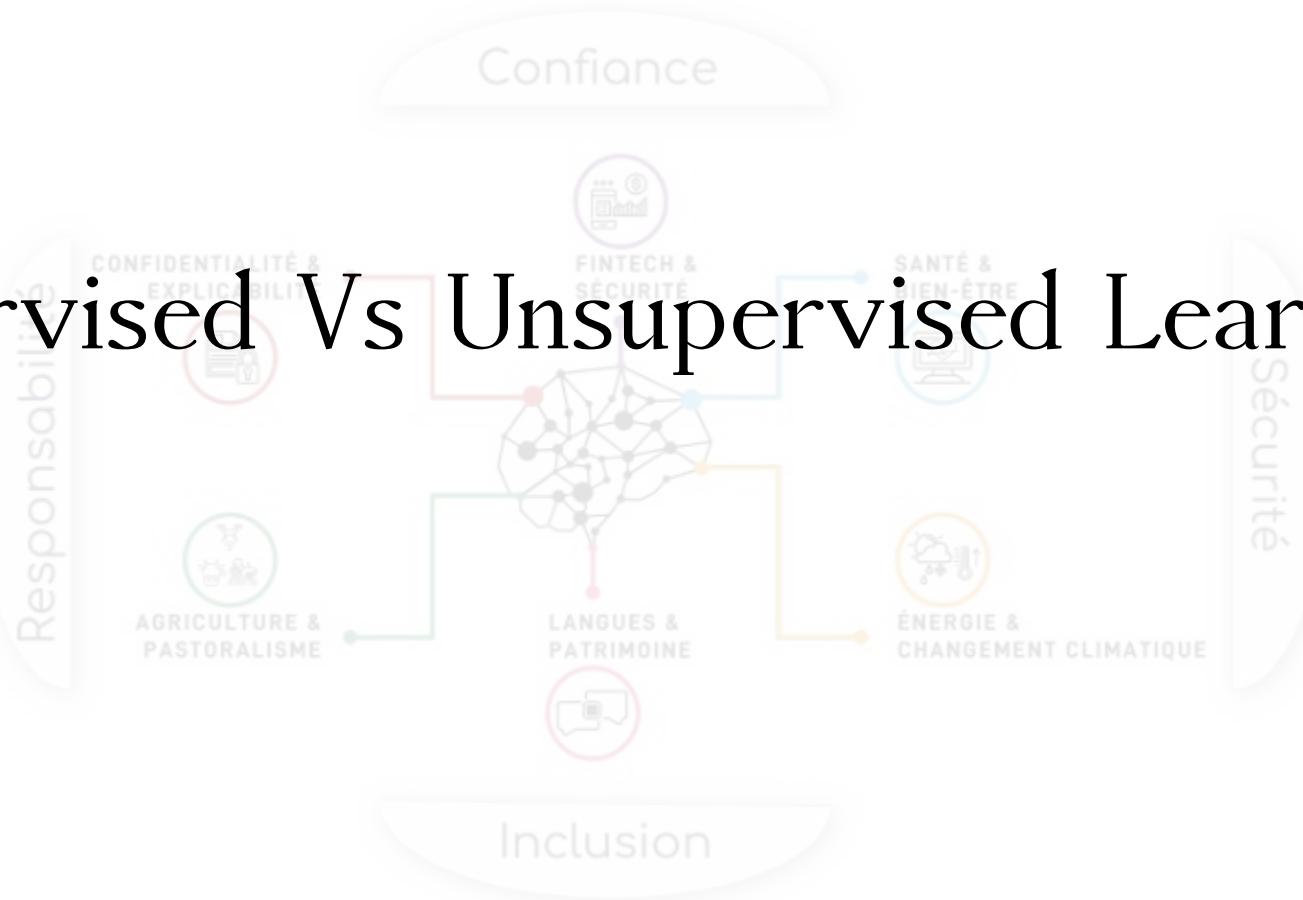
```
import pickle  
s = pickle.dumps(clf)
```

Learn to classify
unknown cases



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Vs Unsupervised Learning



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Learning



Confiance



TECH &
SÉCURITÉ

ES &
MOINE

ision

SANTÉ &
BIEN-ÊTRE



ÉNERGIE
COMMUNICATION

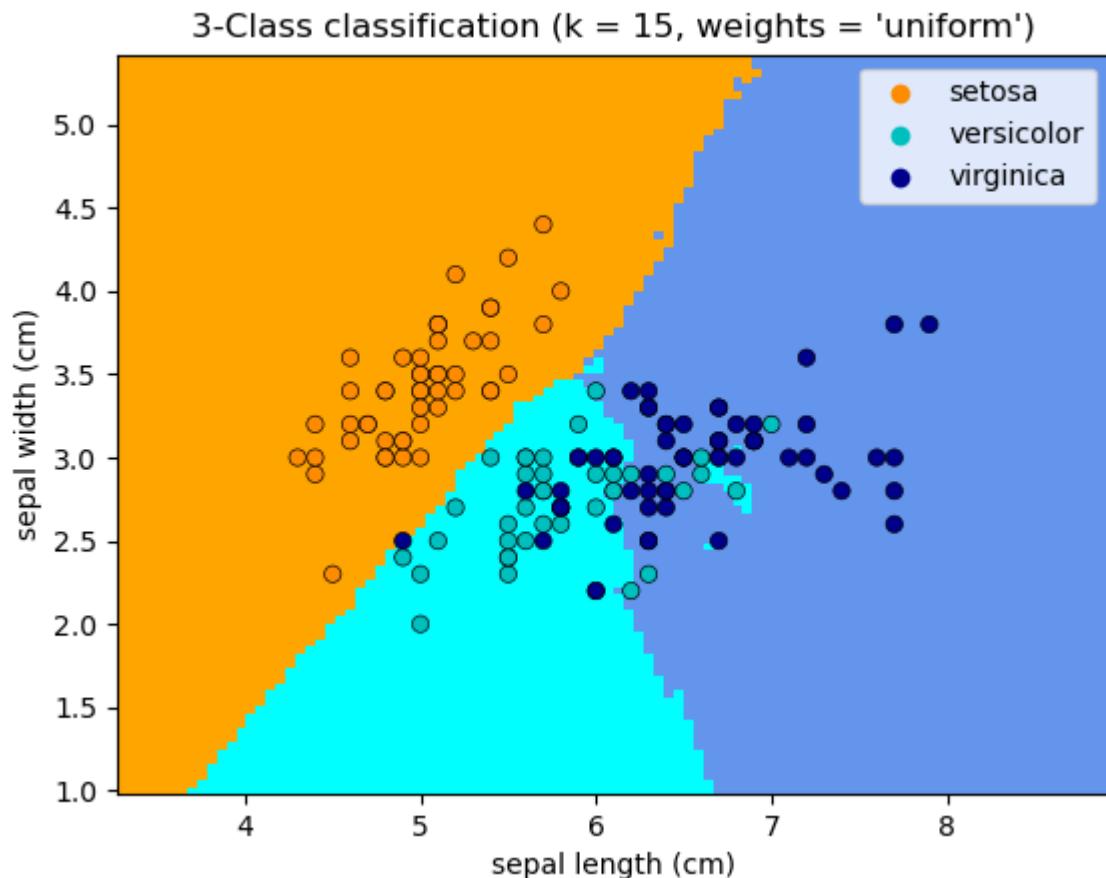
Sécurité

Observer et diriger l'exécution
d'une **tâche**, d'un projet ou d'une
activité.



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Learning



Nous apprenons « **le machine (model)** » et avec cette connaissance, elle pourra prédire ce qui lui est **méconnu** ou carrément des **instances futures**



Centre d'Excellence Interdisciplinaire en Intelligence Artificielle pour le Développement

Supervised Learning

Confiance

[5] :	ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
689	654546	1	1	1	1	2	1	1	1	8	benign
690	654546	1	1	1	3	2	1	1	1	1	benign
691	695091	5	10	10	5	4	5	4	4	1	malignant
692	714039	3	1	1	1	2	1	1	1	1	benign
693	763235	3	1	1	1	2	1	2	1	2	benign
694	776715	3	1	1	1	3	2	1	1	1	benign
695	841769	2	1	1	1	2	1	1	1	1	benign
696	888820	5	10	10	3	7	3	8	10	2	malignant
697	897471	4	8	6	4	3	4	10	6	1	malignant
698	897471	4	8	8	5	4	5	10	4	1	malignant

Inclusion

Attributs

Features

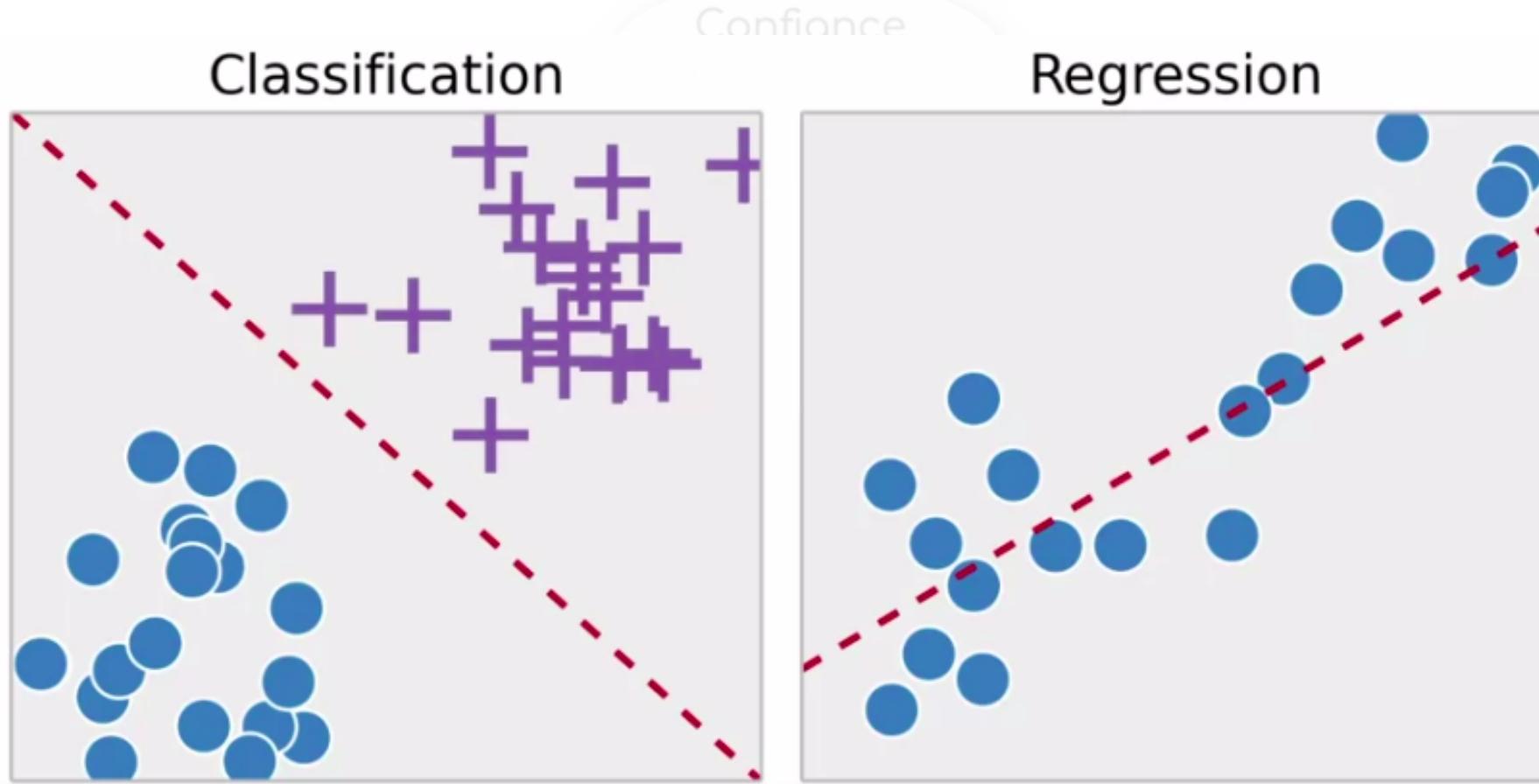
Observation

Class



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

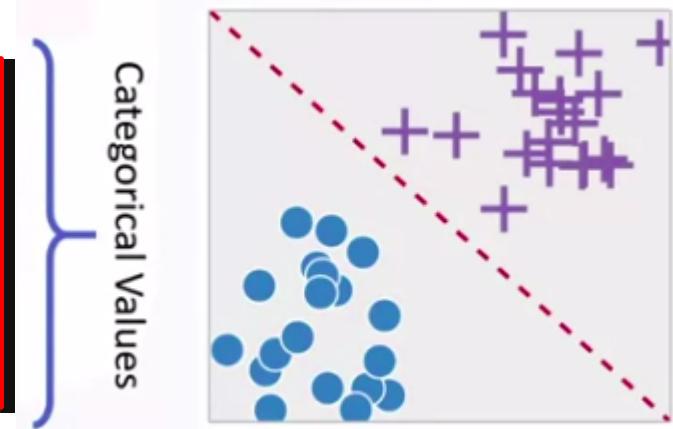
Two types of Supervised Learning



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Learning : Classification

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
689	654546	1	1	1	1	2	1	1	1	benign
690	654546	1	1	1	3	2	1	1	1	benign
691	695091	5	10	10	5	4	5	4	1	malignan
692	714039	3	1	1	1	2	1	1	1	benign
693	763235	3	1	1	1	2	1	2	1	benign



Classification → Processus qui consiste à prédire des classes à valeurs discrètes ou catégorielles



Centre d'Excellence Interdisciplinaire en Intelligence Artificielle pour le Développement

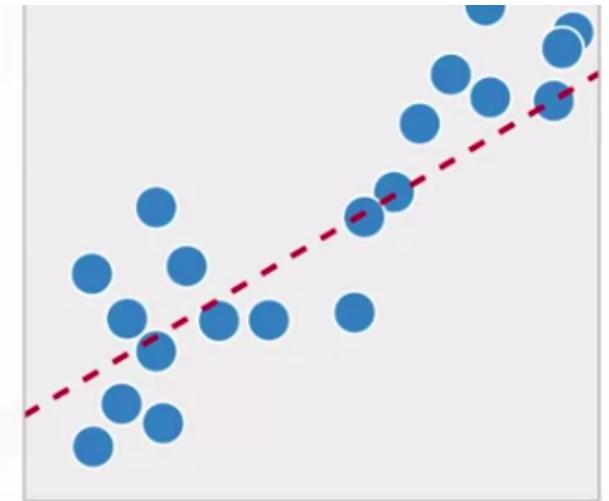
Supervised Learning : Regression

X : independente/explorative variable Y : dependente variable

[8] :

	ENGINESIZE	CYLINDERS	FUELCONSUMPTION_COMB	CO2EMISSIONS
0	2.0	4	8.5	196
1	2.4	4	9.6	221
2	1.5	4	5.9	136
3	3.5	6	11.1	255
4	3.5	6	10.6	244
5	3.5	6	10.0	230
6	3.5	6	10.1	232

Continuous Values



Regression → Processus qui consiste à prédire des classes à
valeurs continues



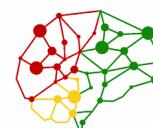
Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Unsupervised Learning

	Customer Id	Age	Edu	Years Employed	Income	Card Debt	Other Debt	Defaulted	DebtIncomeRatio
0	1	41	2	6	19	0.124	1.073	0.0	6.3
1	2	47	1	26	100	4.582	8.218	0.0	12.8
2	3	33	2	10	57	6.111	5.802	1.0	20.9
3	4	29	2	4	19	0.681	0.516	0.0	6.3
4	5	47	1	31	253	9.308	8.908	0.0	7.2



Le modèle **se charge** seul de **découvrir** des **informations**...qui ne sont pas à **priori** détectables visuellement.

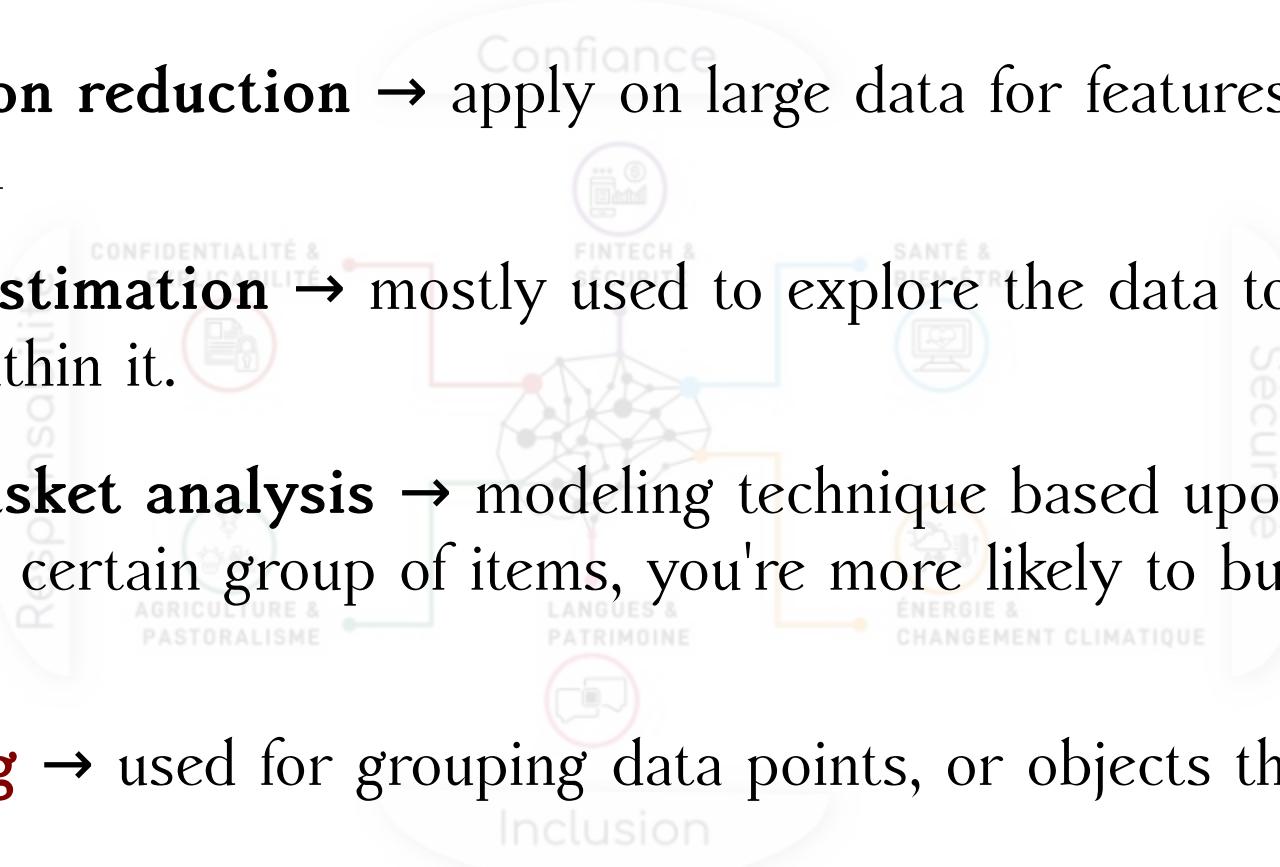


Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Unsupervised Learning

Plusieurs Techniques...

- * **Dimension reduction** → apply on large data for features selection before classification
- * **Density estimation** → mostly used to explore the data to find some structure within it.
- * **Market basket analysis** → modeling technique based upon the theory that if you buy a certain group of items, you're more likely to buy another group of items
- * **Clustering** → used for grouping data points, or objects that are somehow similar.
 - * Anomaly detection
 - * Summarization
 - * Discovering Structure



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Vs Unsupervised Learning

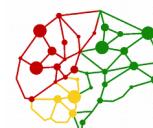
Supervised Learning

- Classification: ★
Classifies labeled data
- Regression: ★
Predicts trends using previous labeled data
- Has more evaluation methods than unsupervised learning ★
- Controlled environment

Unsupervised Learning

- Clustering: ★
Finds patterns and groupings from unlabeled data
- Has fewer evaluation methods than supervised learning ★
- Less controlled environment

la machine crée
des sortis à notre
place



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Vs Unsupervised Learning

Which technique/s is/are considered as Supervised learning?

- * Clustering
- * Regression ★
- * Classification★



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement

Supervised Vs Unsupervised Learning

Which one is a sample application of regression?

- * Predicting whether a patient has cancer or not.
- * Grouping of similar houses in an area.
- * Forecasting rainfall amount for next day. 
- * Predicting if a team will win or not.



Centre d'Excellence Interdisciplinaire en
Intelligence Artificielle pour le Développement