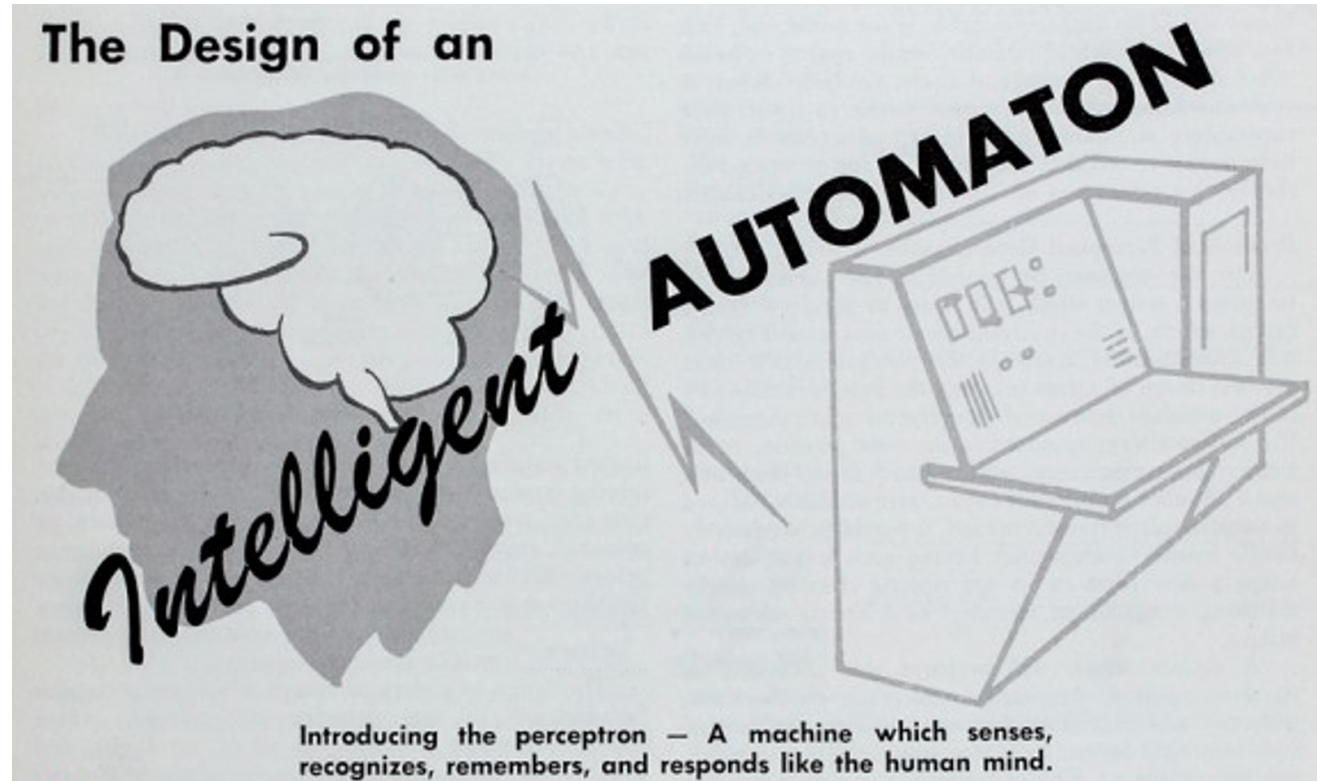


Des données à l'intelligence artificielle



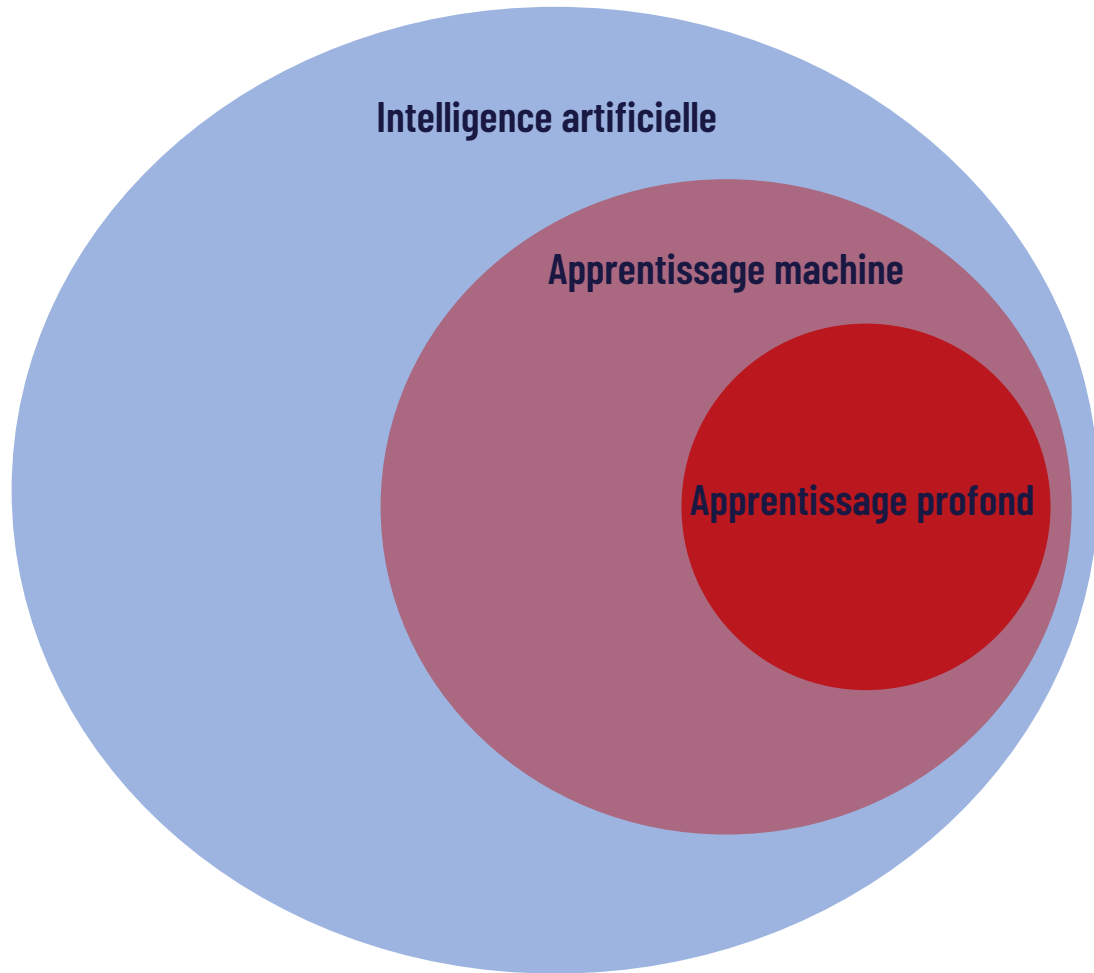
<https://news.cornell.edu/stories/2019/09/professors-perceptron-paved-way-ai-60-years-too-soon>



Yves Terrat PhD

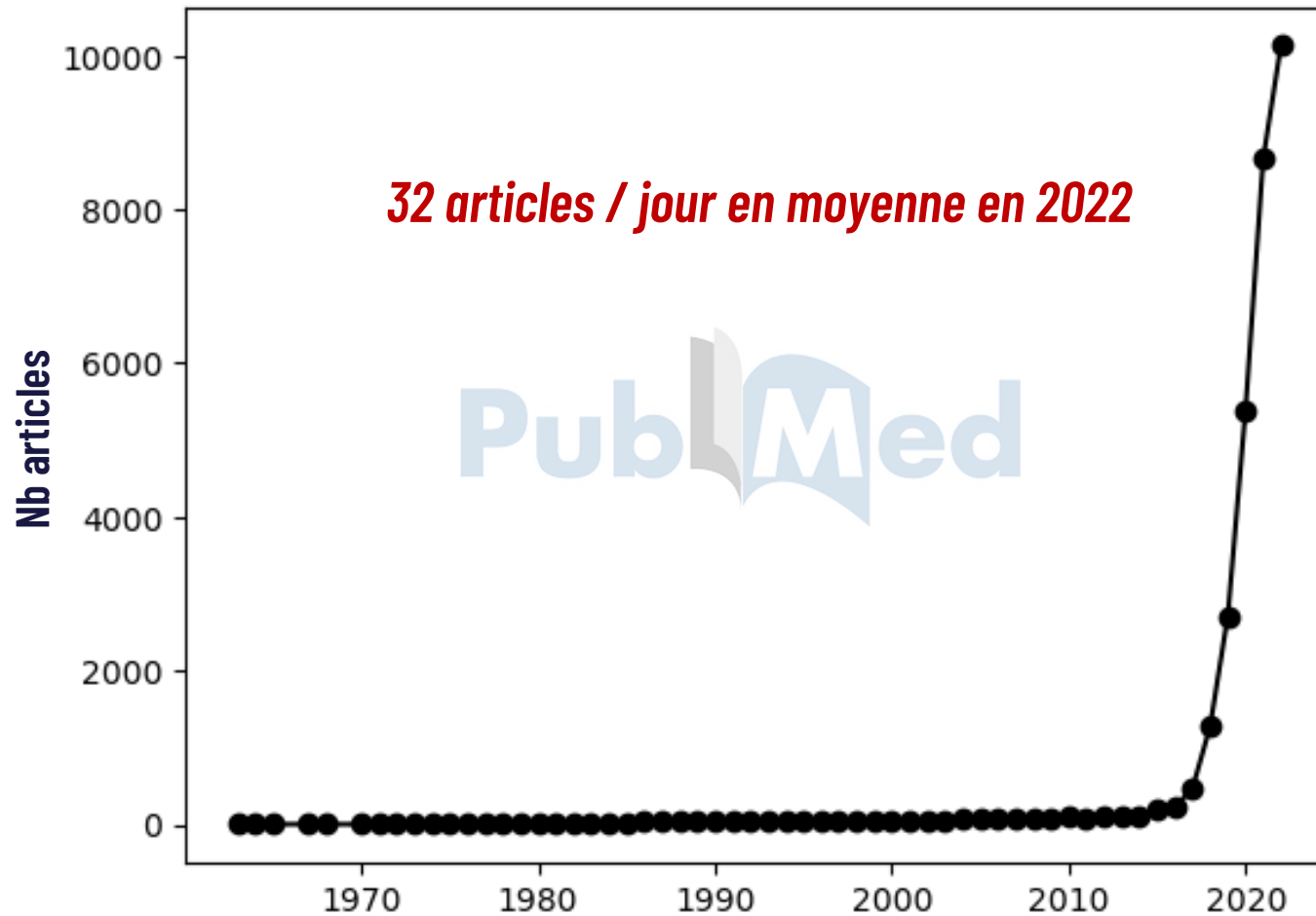
Consortium Santé Numérique, UdeM-IVADO

Qu'est-ce que l'intelligence artificielle ?



*« Algorithmes qui sont
capables d'apprendre des
taches complexes
sans avoir été
explicitement programmés »*

Le « big bang » de l'intelligence artificielle en santé



Source : PubMed, keyword : (Artificial intelligence) [Title/Abstract]

A-t-on besoin de l'IA ?

[Published: 03 April 2018](#)

Points of Significance

Statistics versus machine learning

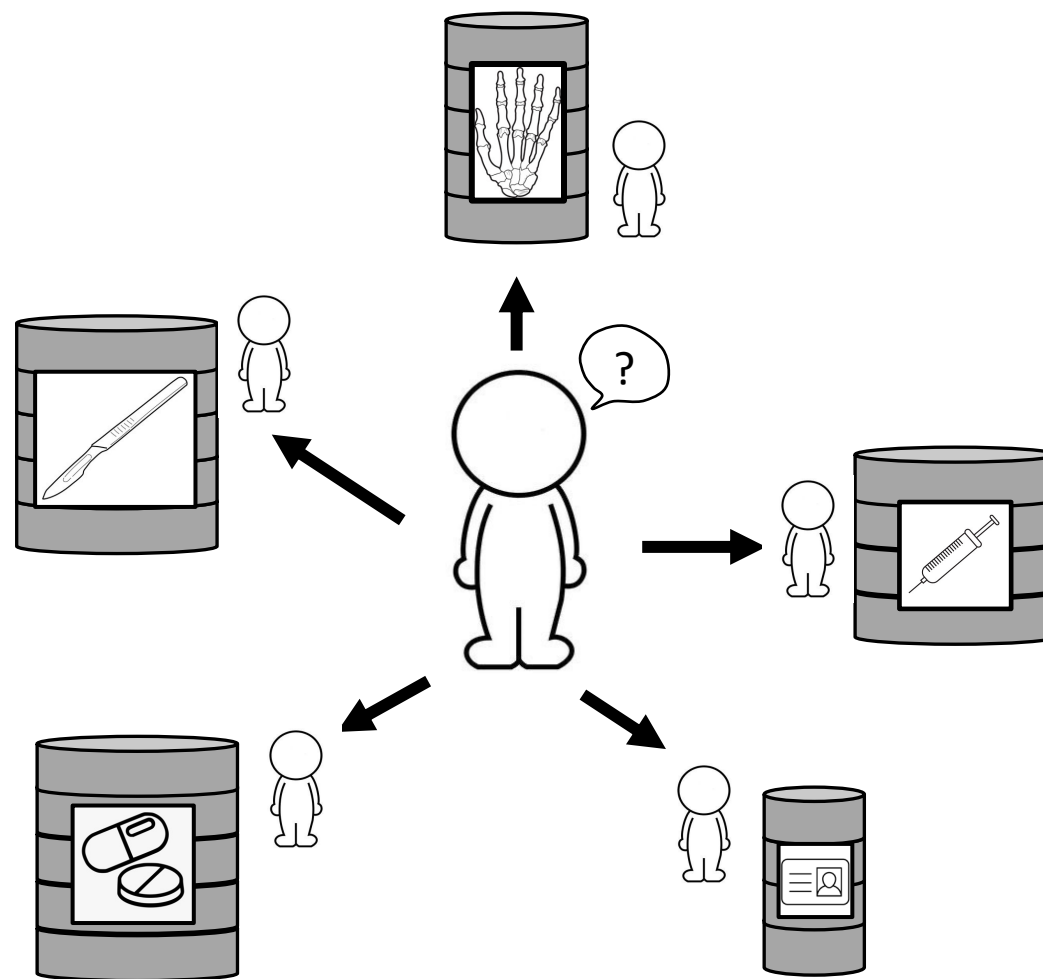
[Danilo Bzdok](#), [Naomi Altman](#) & [Martin Krzywinski](#)

[Nature Methods](#) **15**, 233–234 (2018) | [Cite this article](#)

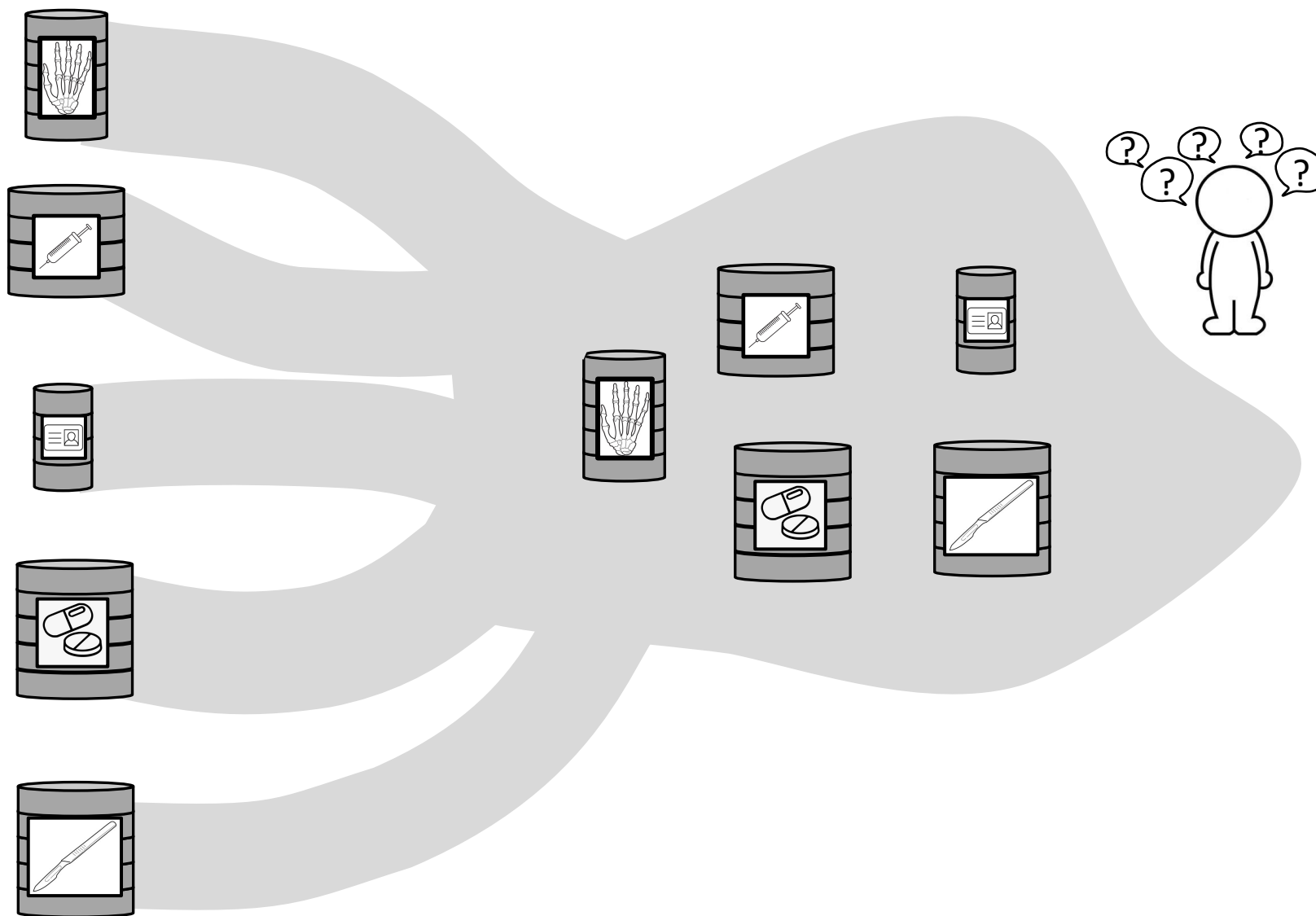
95k Accesses | **462** Citations | **360** Altmetric | [Metrics](#)

- IA adapté aux données massives
- IA n'a pas d'a priori sur la distribution des données
- IA peut gérer des milliers de variables et des millions d'exemples

Disponibilité des données massives en santé : organisation en silos



Disponibilité des données massives en santé : les lacs de données





Définir son modèle d'IA

Fonction (*question*

*domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)*



Définir son modèle d'IA

Fonction (*question,
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...*)



Définir son modèle d'IA

Fonction (*question,*
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)



Définir son modèle d'IA

Fonction (*question,*
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)



Définir son modèle d'IA

Fonction (*question,*
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)



Définir son modèle d'IA

Fonction (*question,*
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)



Définir son modèle d'IA

Fonction (*question,*
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)



Définir son modèle d'IA

Fonction (*question,
domaine d'étude,
taille des données,
qualité des données,
accès à une infrastructure de calcul,
temps de calcul,
performance,
interprétabilité,
...)*)



Préparation des données

	Patient.e 1	Patient.e 2	Patient.e 3	Patient.e 4	Patient.e 5	Patient.e 6
Éosinophiles	1.3	0.55	0. 57	0.2	0.5	0.8
Age	54	37	27	64	73	40
Mutation	Oui	Oui	Non	Non	Non	Oui
Statut	Cancer	Non cancer	Cancer	Non cancer	Non cancer	Cancer

NB : on ne parlera pas ici du pré-traitement des données

Préparation des données

Variables
Résultat

	Patient.e 1	Patient.e 2	Patient.e 3	Patient.e 4	Patient.e 5	Patient.e 6
Éosinophiles	1.3	0.55	0. 57	0.2	0.5	0.8
Age	54	37	27	64	73	40
Mutation	Oui	Oui	Non	Non	Non	Oui
Statut	Cancer	Non cancer	Cancer	Non cancer	Non cancer	Cancer

Préparation des données

Patient.e 5	Patient.e 6	Patient.e 5	Patient.e 6	Patient.e 5
0.5	0.8	0.5	0.8	0.5
73	40	73	40	73
Non	Oui	Non	Oui	Non
Non cancer	Cancer	Non cancer	Cancer	Non cancer

Entrainement

Patient.e 5	Patient.e 6
0.5	0.8
73	40
Non	Oui
Non cancer	Cancer

Test

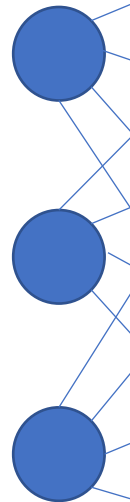
Entraîner un réseau de neurones artificiels

entrées

Éosino

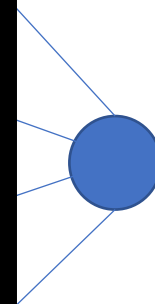
Age

Mutation

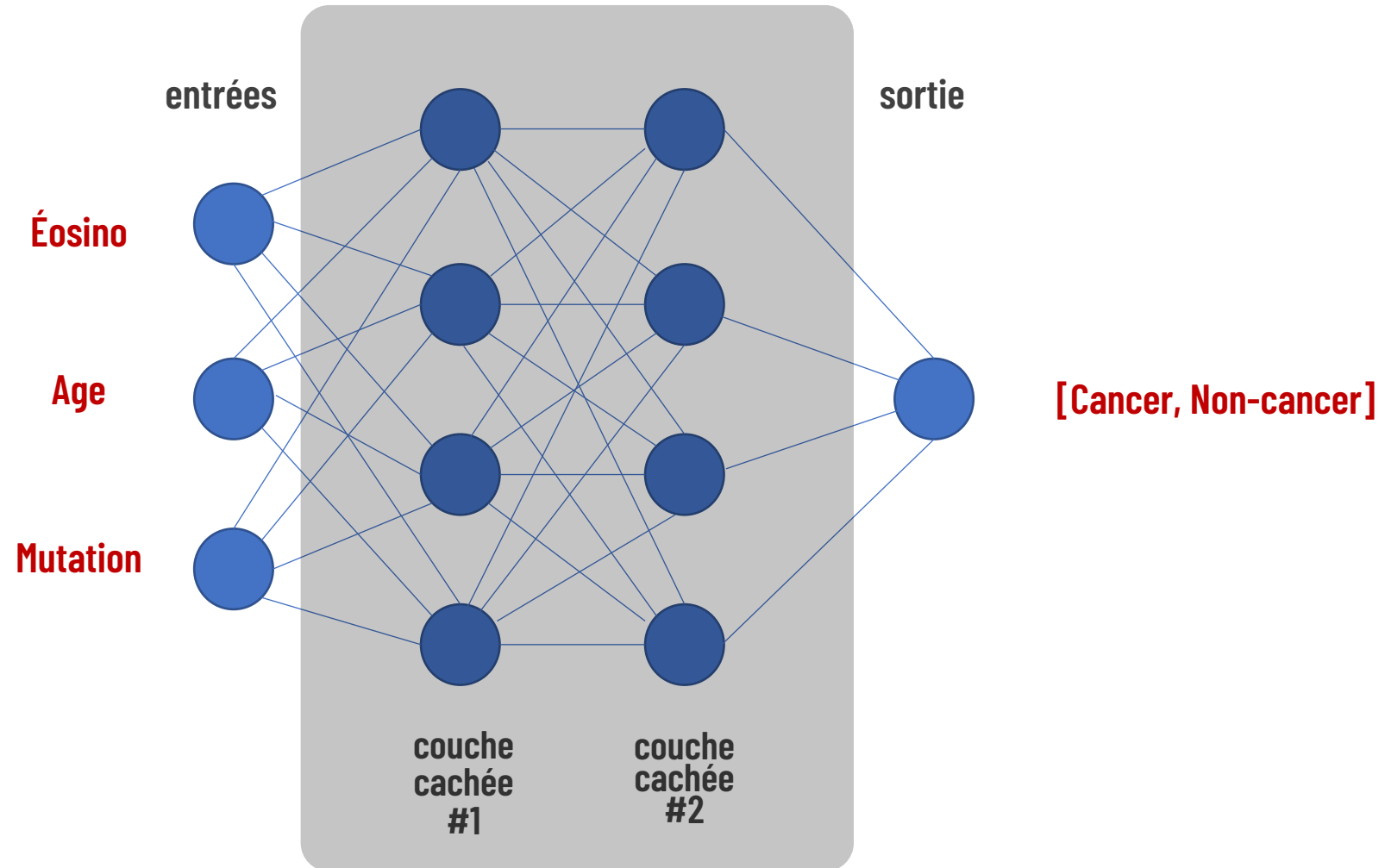


sortie

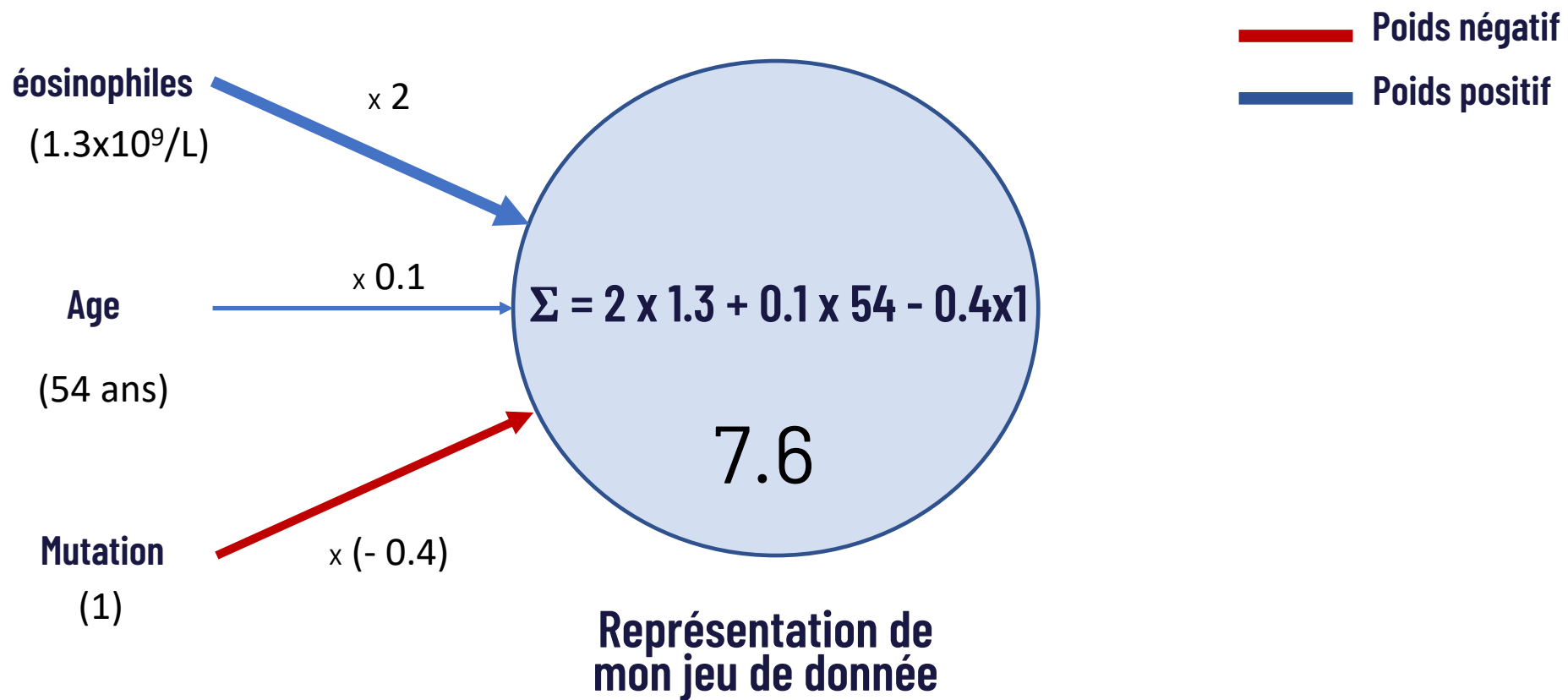
[Cancer, Non-cancer]



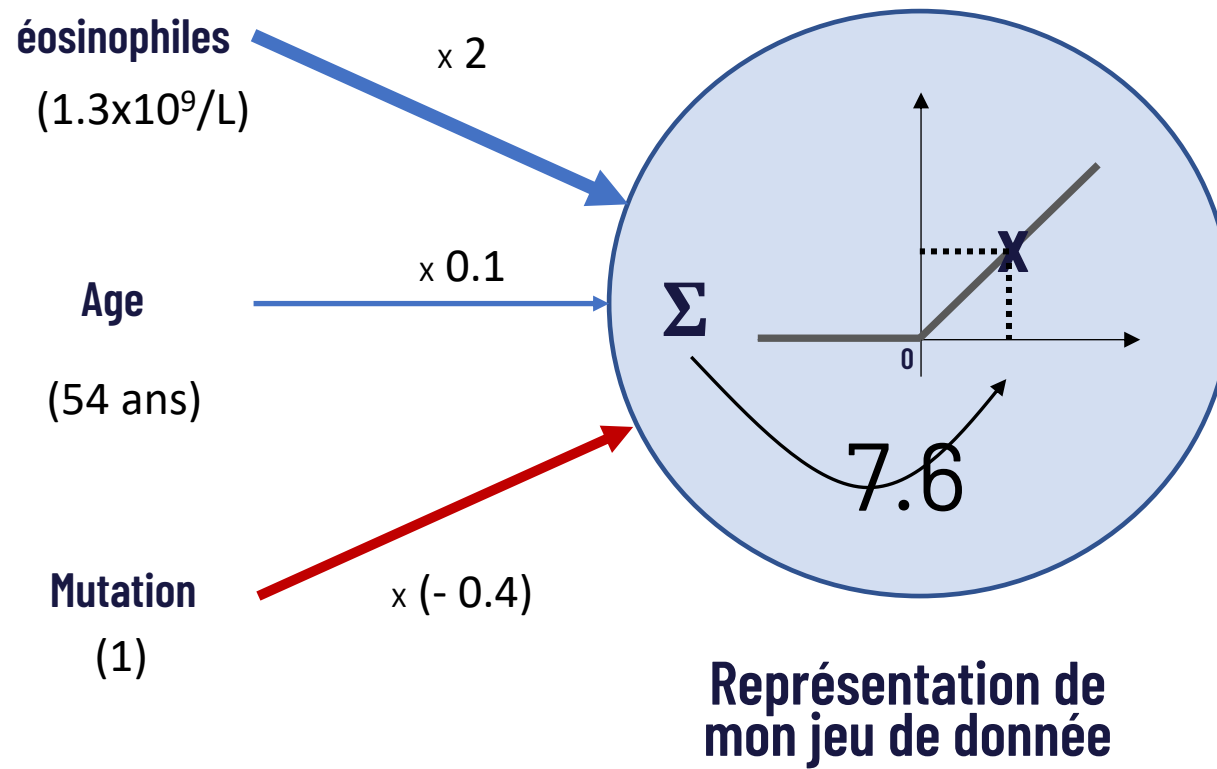
Entraîner un réseau de neurones artificiels



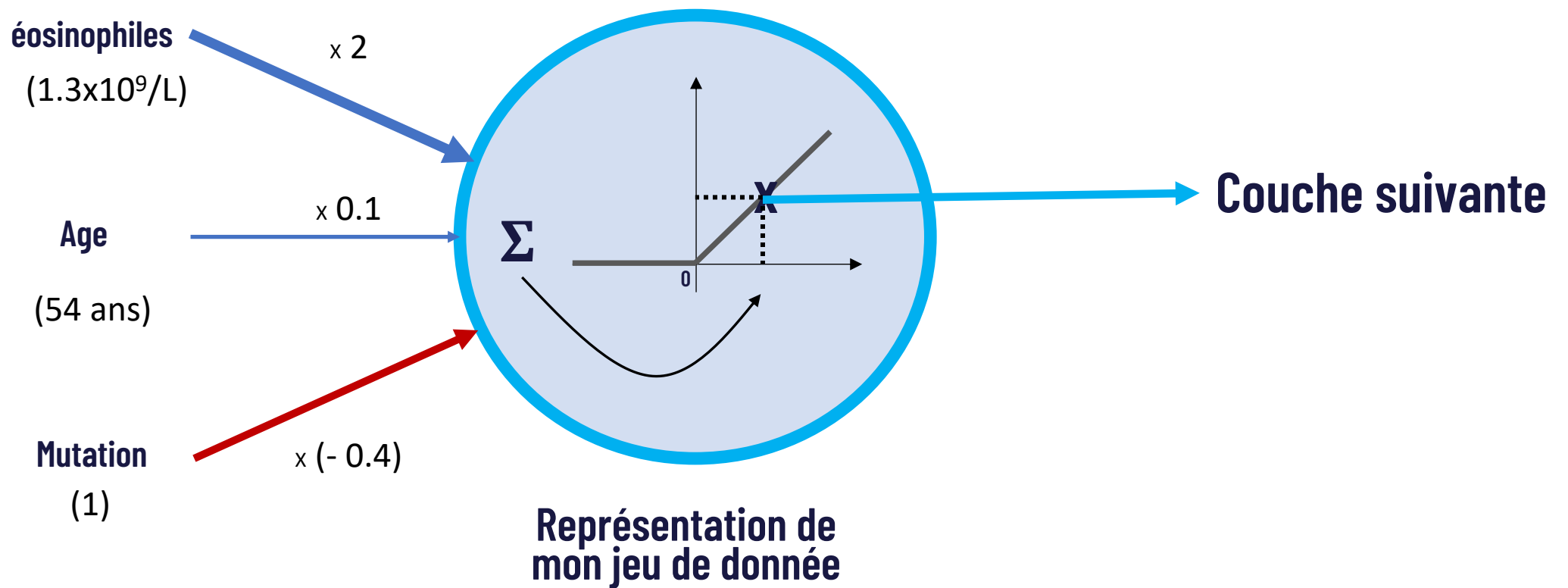
Entraîner un réseau de neurones artificiels



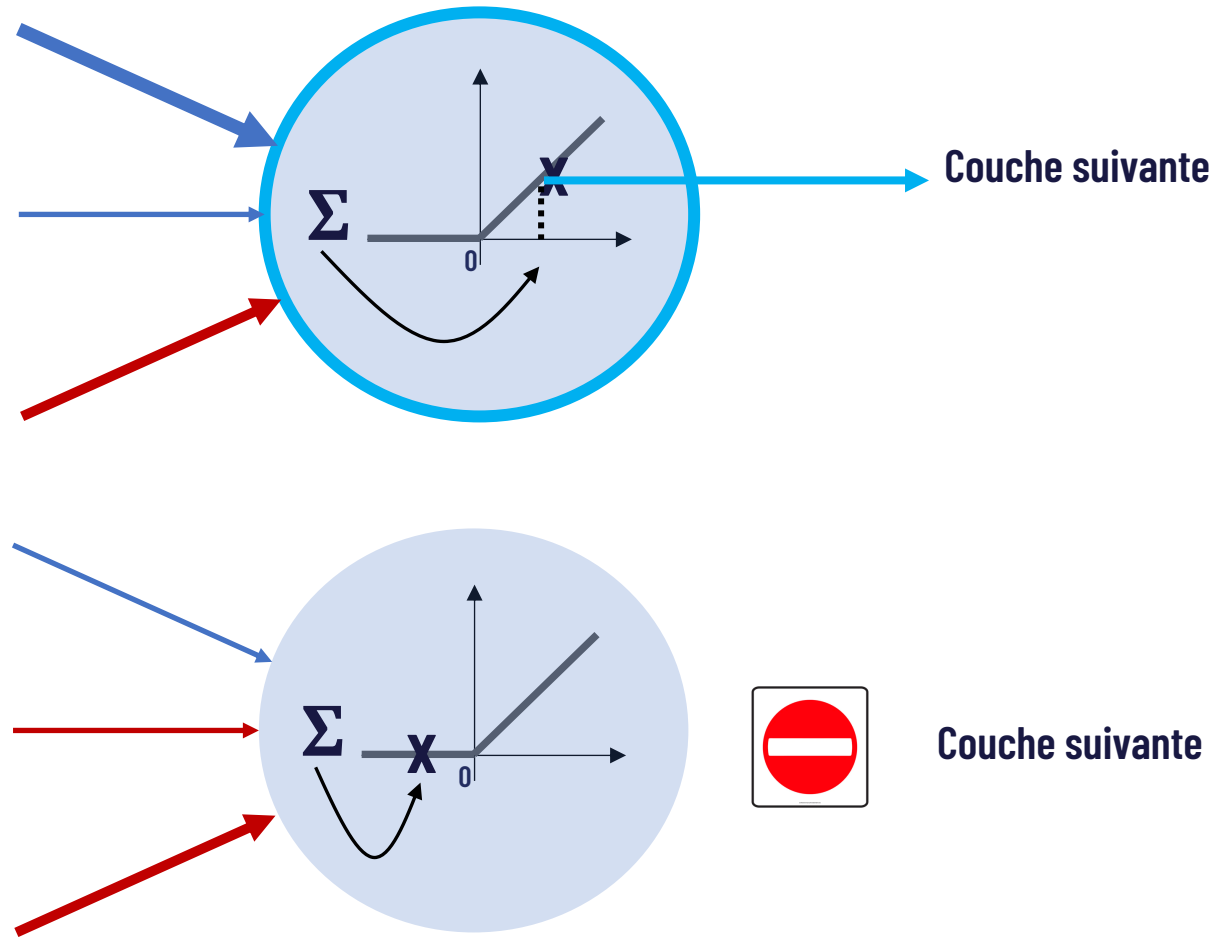
Entraîner un réseau de neurones artificiels



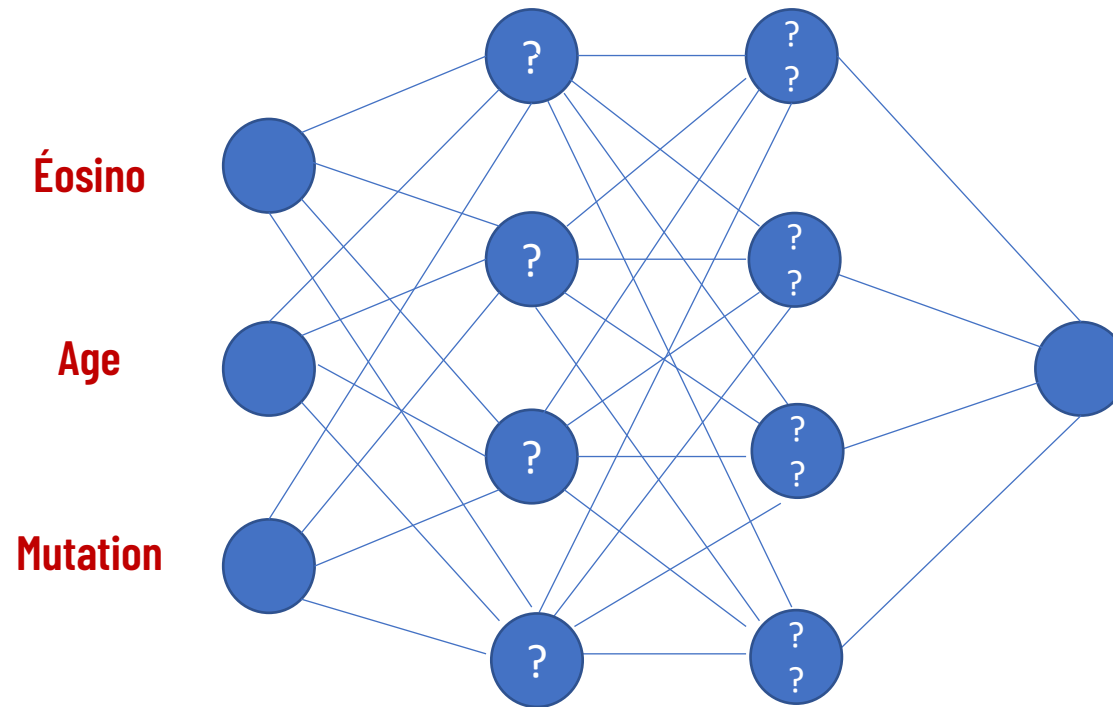
Entraîner un réseau de neurones artificiels



Entraîner un réseau de neurones artificiels



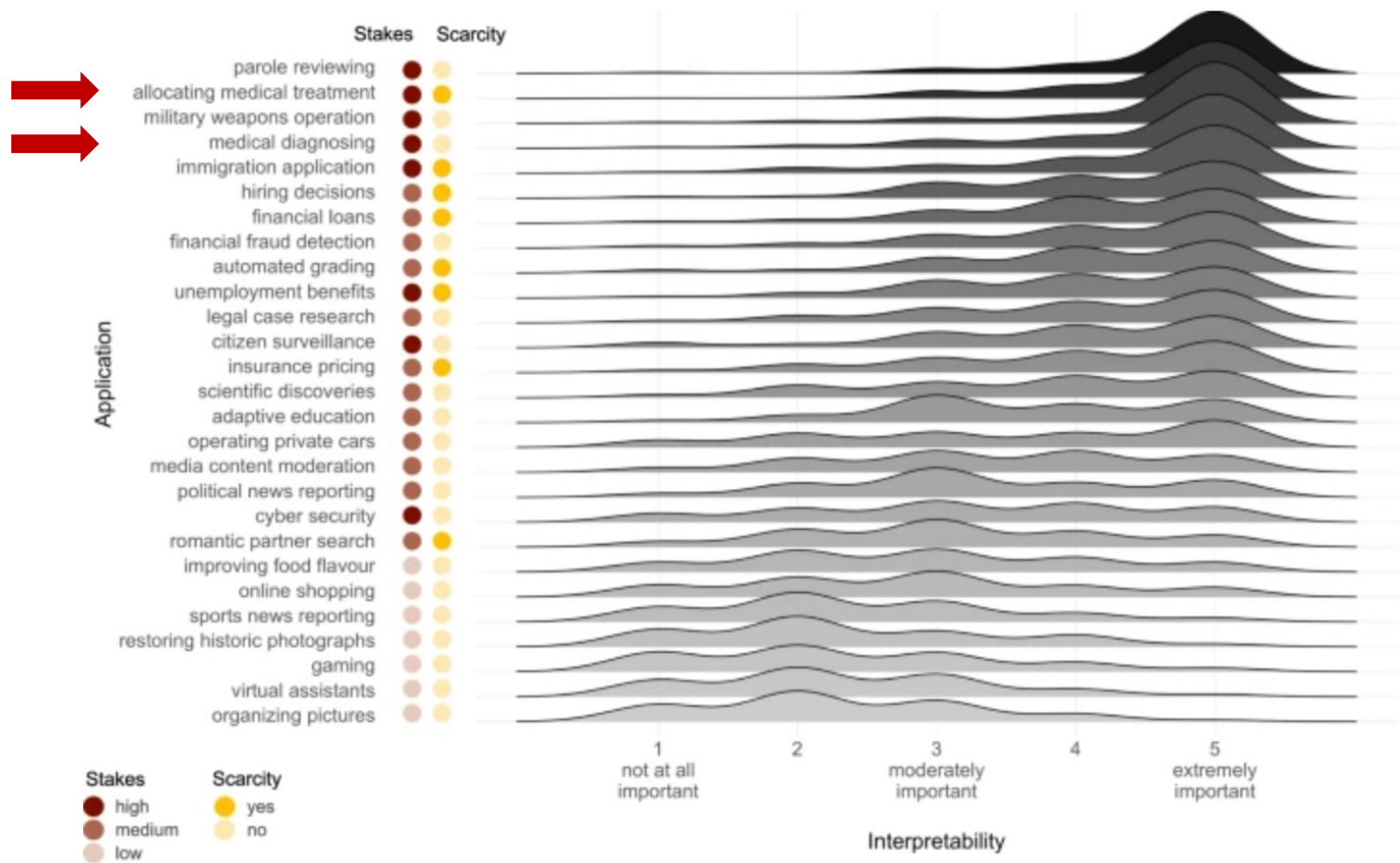
Entraîner un réseau de neurones artificiels



Interprétabilité ???

Une IA explicable

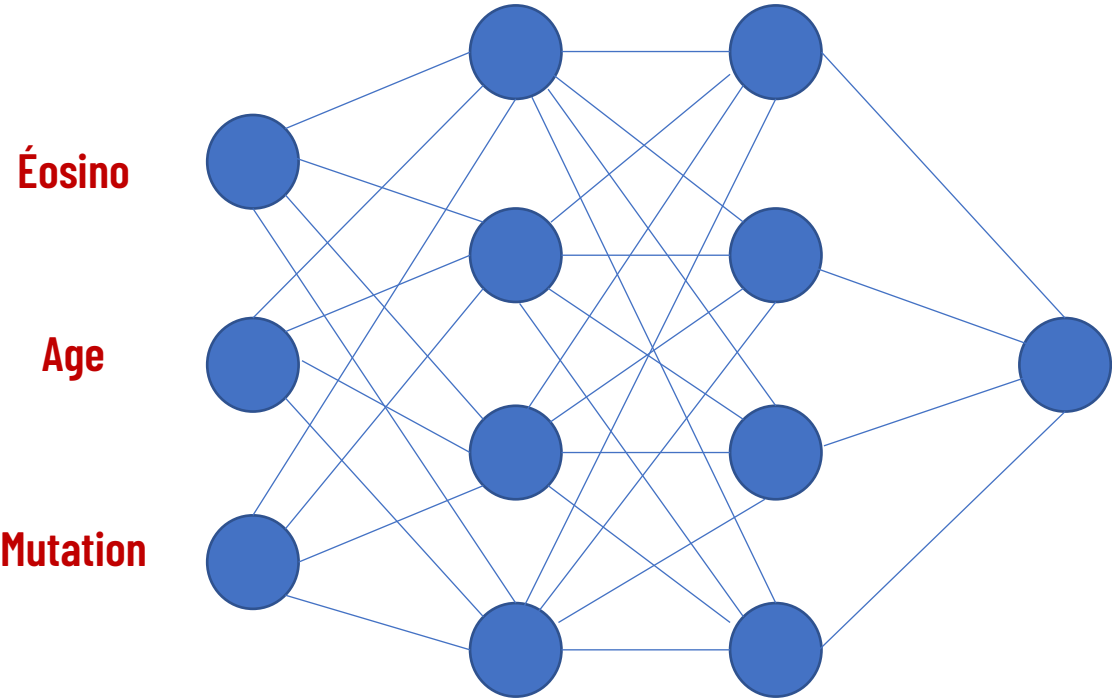
Fig. 1: Attitudes towards interpretability across real-world AI applications.



<https://www.nature.com/articles/s41467-022-33417-3>

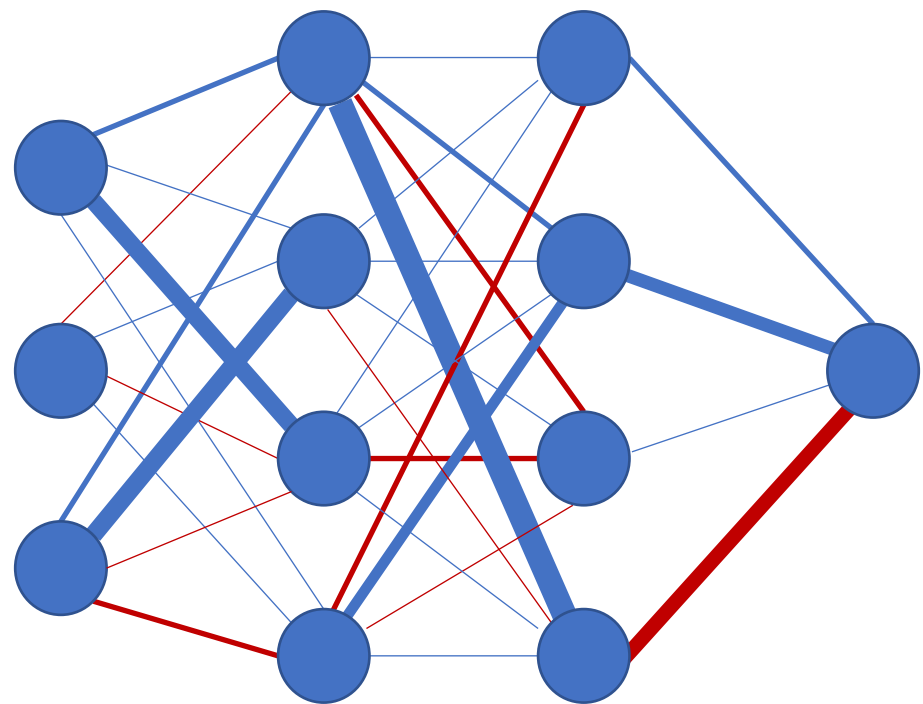


Entraîner un réseau de neurones artificiels





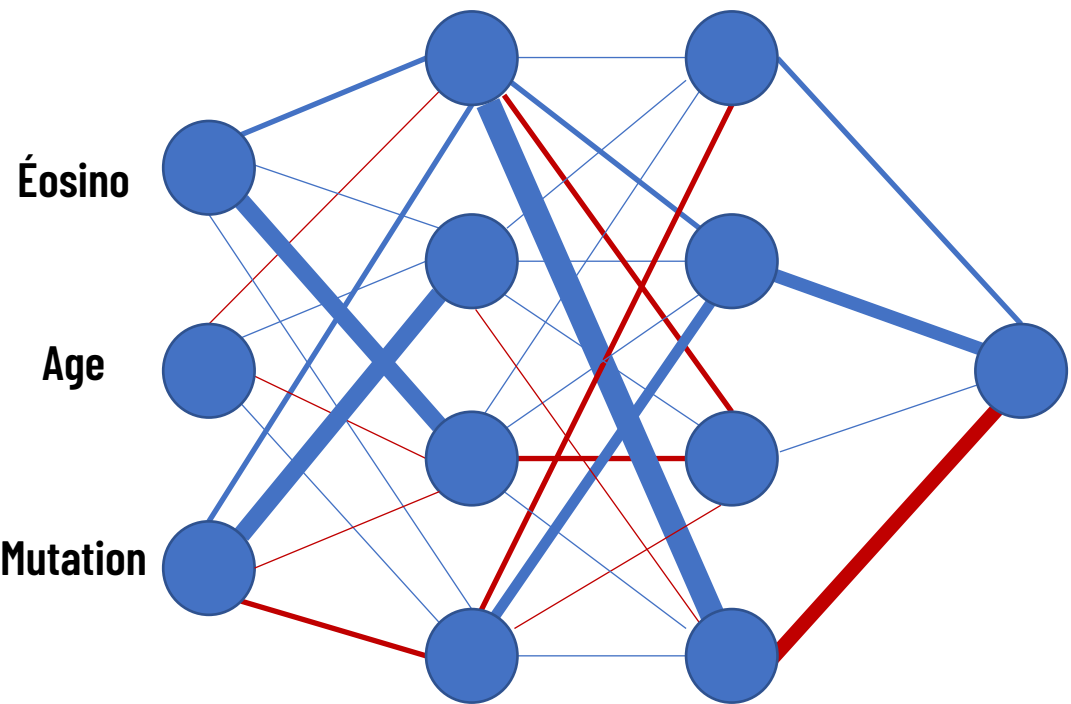
Entrainer un réseau de neurones artificiels



Entraîner un réseau de neurones artificiels

P1	P2	P3	P4
1.3	0.55	0.57	0.2
54	37	27	64
Oui	Oui	Non	Non

Cancer	Non Cancer	Cancer	Non Cancer
--------	------------	--------	------------

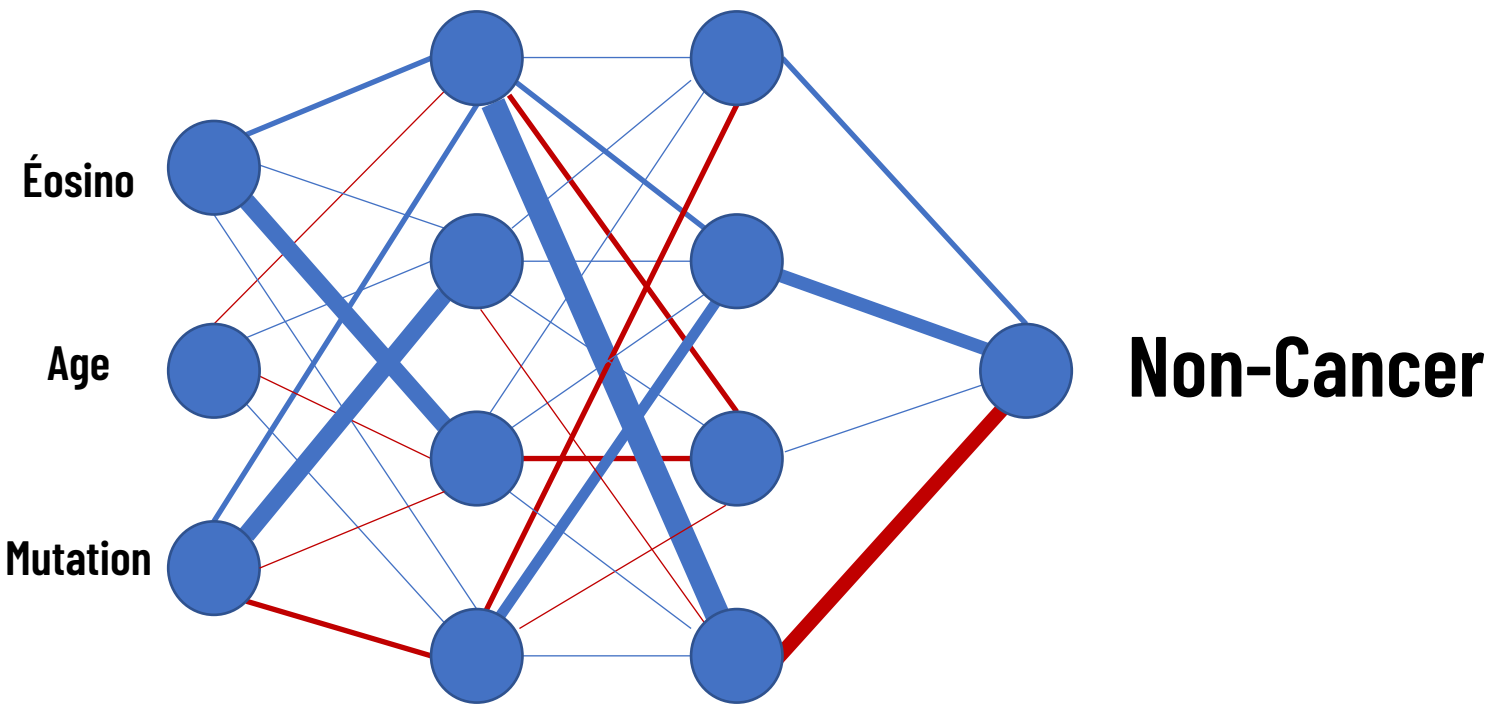


Entraînement

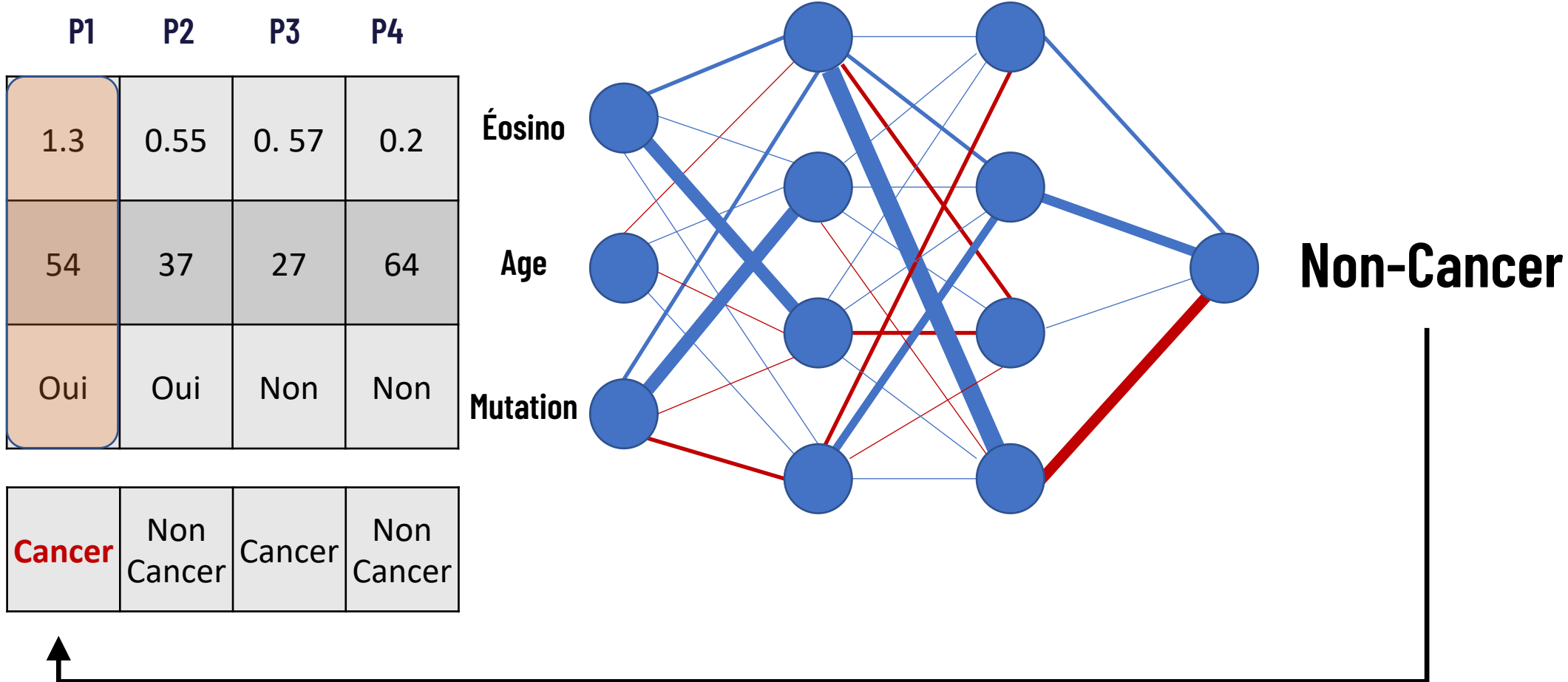
Entraîner un réseau de neurones artificiels

P1	P2	P3	P4
1.3	0.55	0.57	0.2
54	37	27	64
Oui	Oui	Non	Non

Cancer	Non Cancer	Cancer	Non Cancer
--------	------------	--------	------------



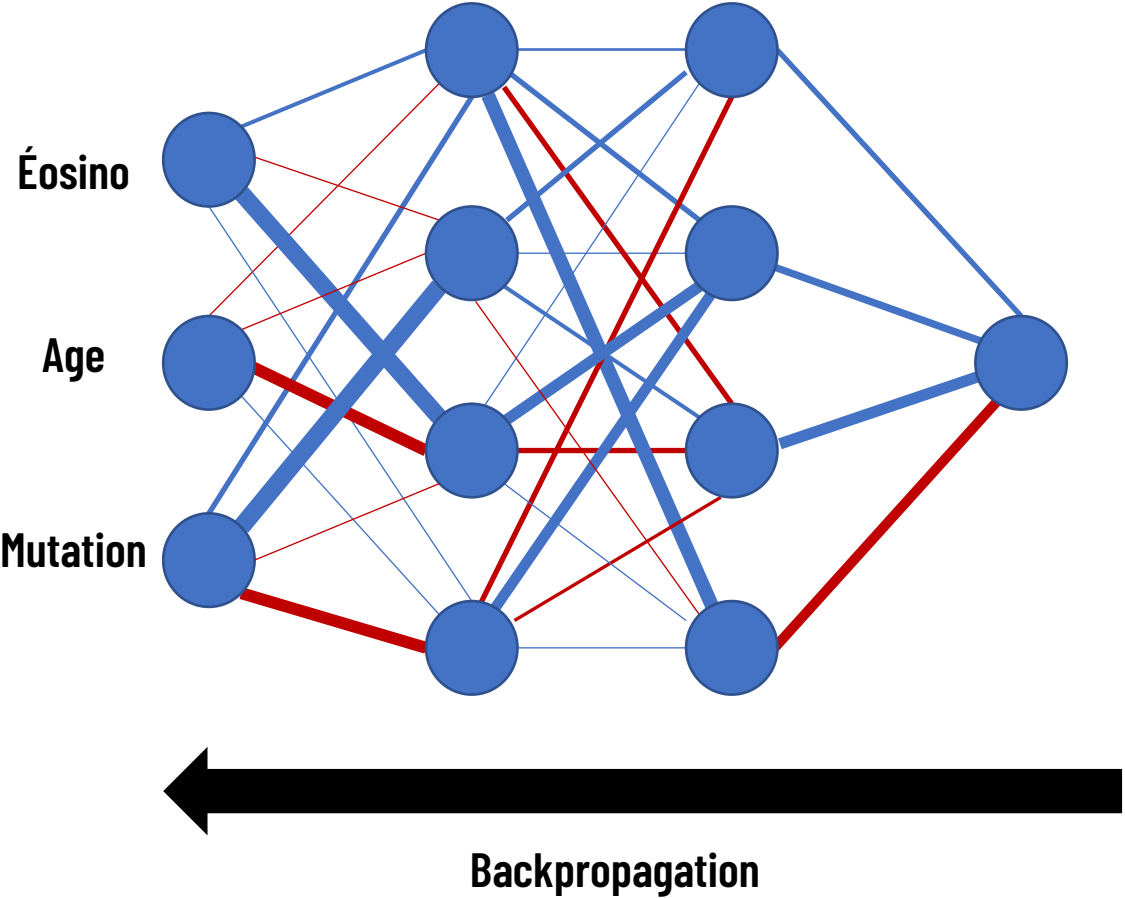
Entraîner un réseau de neurones artificiels



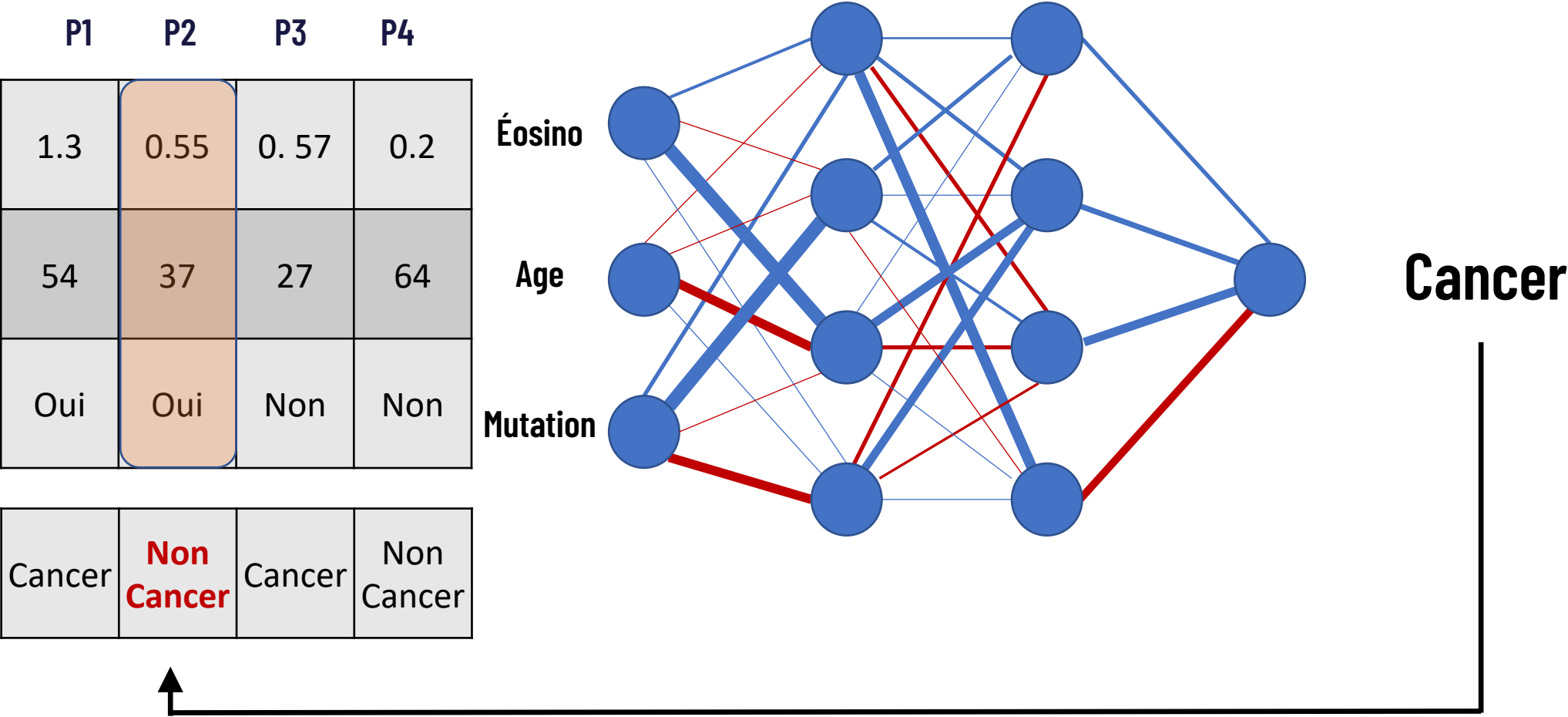
Entraîner un réseau de neurones artificiels

P1	P2	P3	P4
1.3	0.55	0.57	0.2
54	37	27	64
Oui	Oui	Non	Non

Cancer	Non Cancer	Cancer	Non Cancer
--------	------------	--------	------------



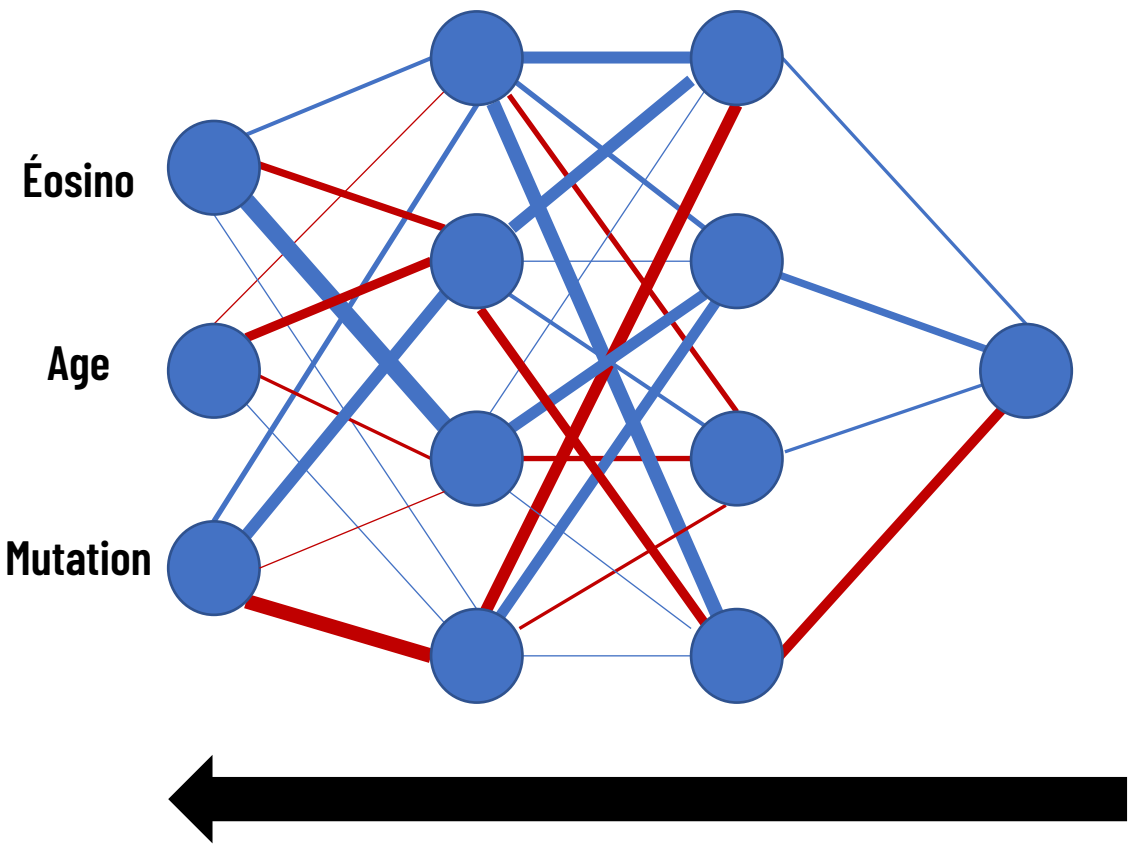
Entraîner un réseau de neurones artificiels



Entraîner un réseau de neurones artificiels

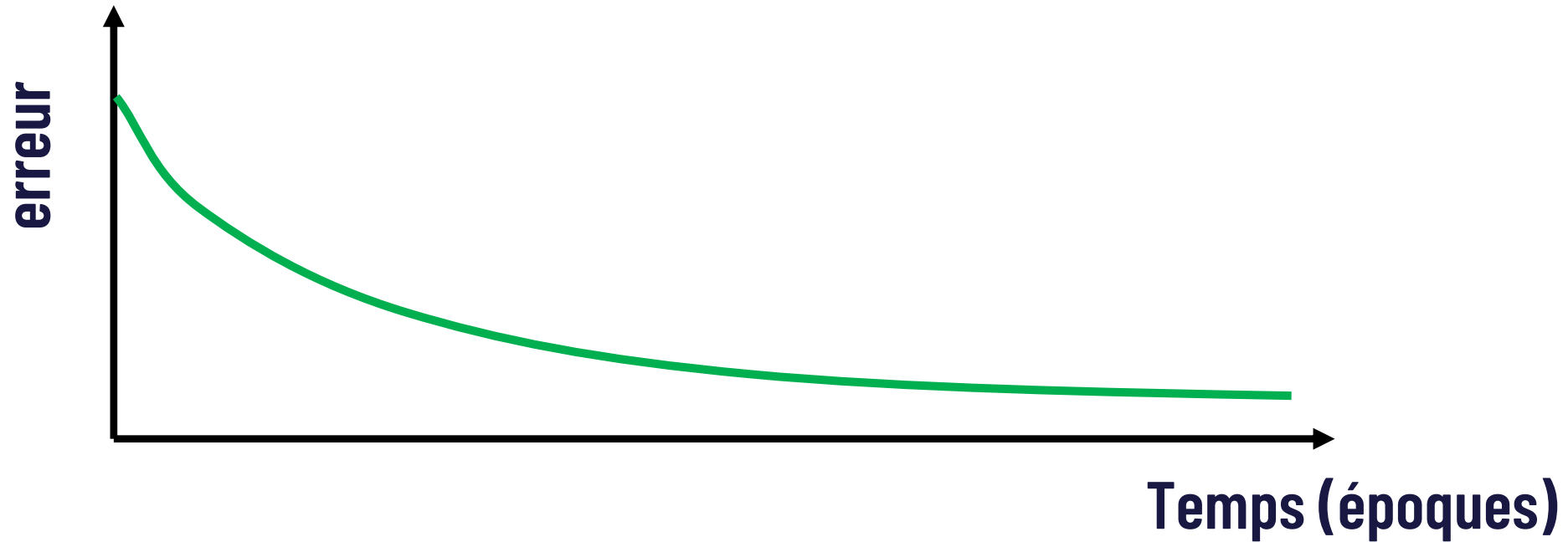
P1	P2	P3	P4
1.3	0.55	0.57	0.2
54	37	27	64
Oui	Oui	Non	Non

Cancer	Non Cancer	Cancer	Non Cancer
--------	---------------	--------	---------------



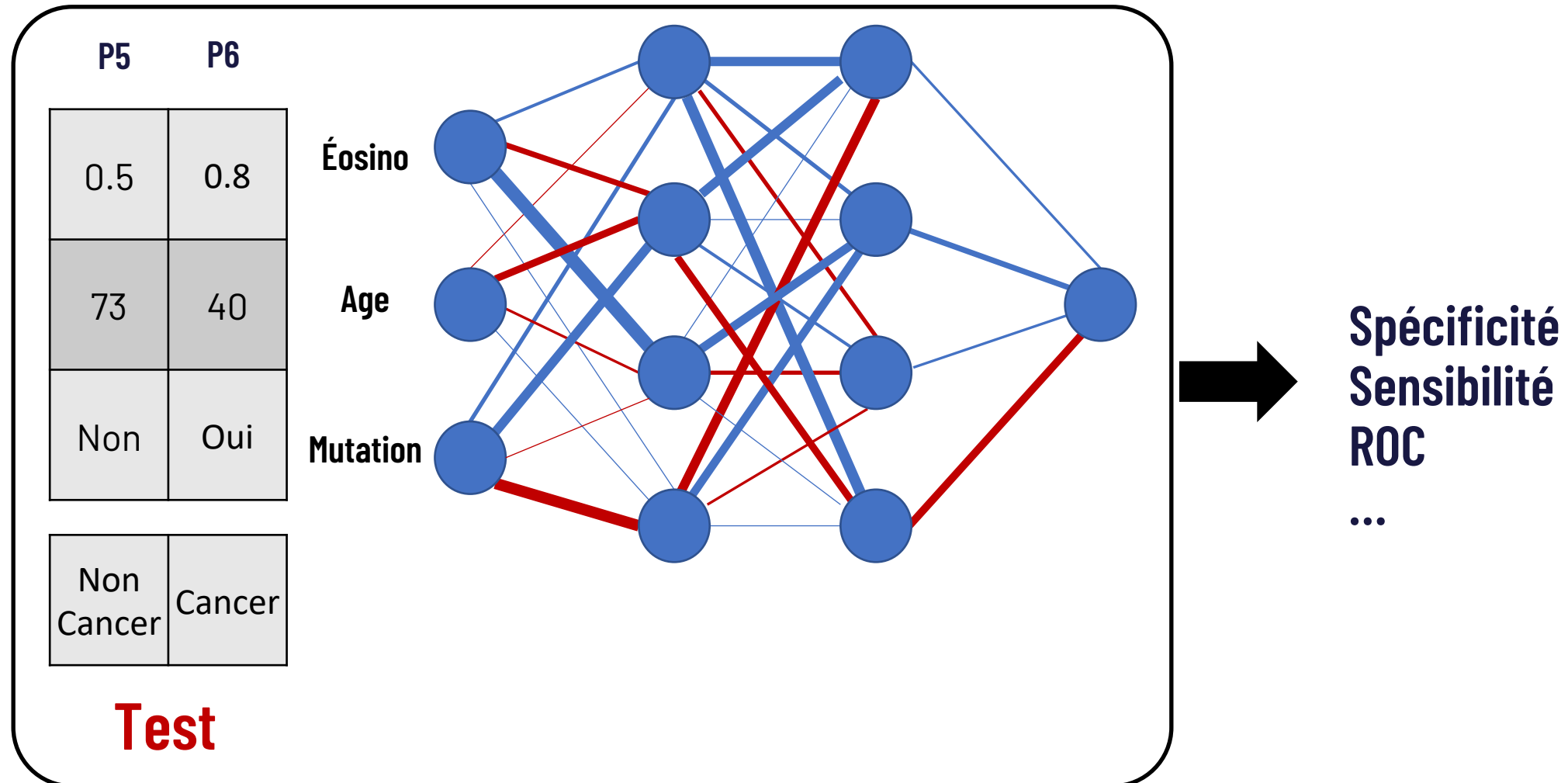
Backpropagation

Entraîner un réseau de neurones artificiels

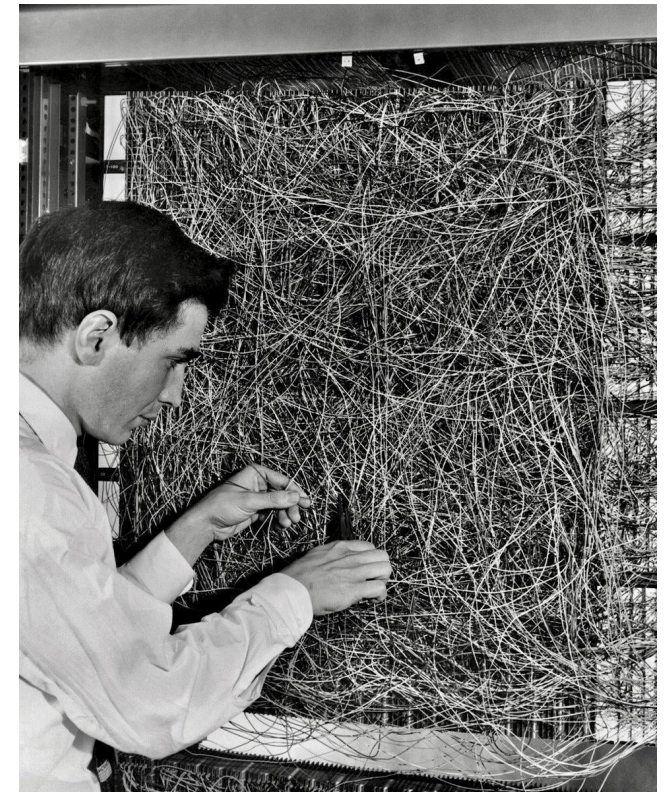
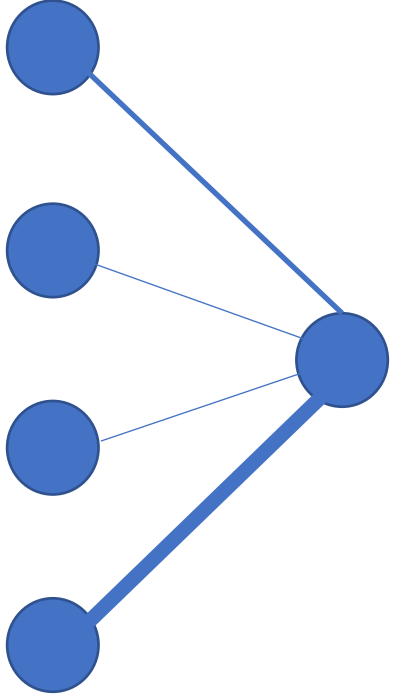


Un réseau de neurones devient « intelligent » car il absorbe de la donnée

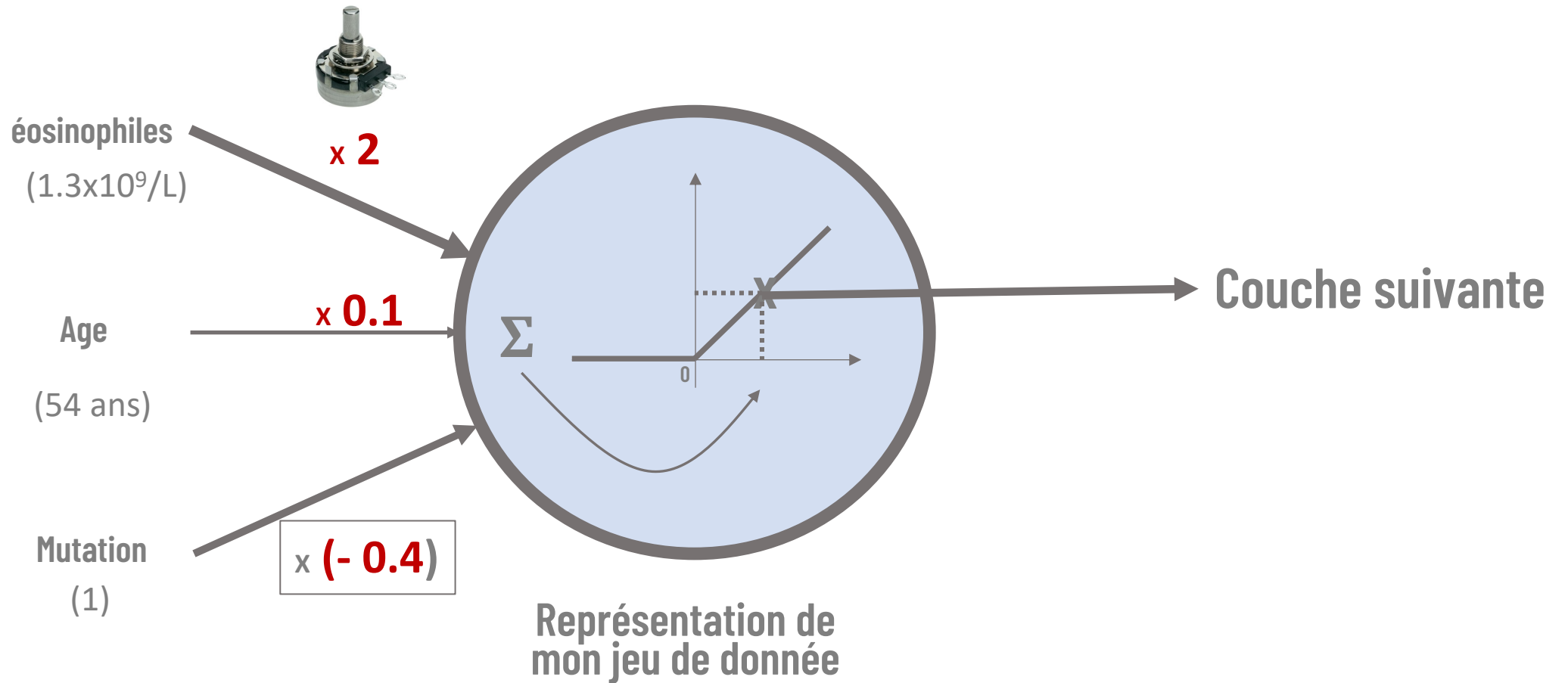
Valider l'entraînement : notre modèle généralise-t-il bien ?



Le perceptron, Rosenblatt, 1959



Poids numériques : potentiomètres analogiques



Une intelligence artificielle à la portée de (presque) toutes et tous

```
from sklearn.neural_network import MLPClassifier
steps = [('scaler', StandardScaler()),
         ('KB', SelectKBest(score_func=f_classif, k = 4)),
         ('MLP', MLPClassifier(random_state=rs))]
pipeline = Pipeline(steps)
diabetesDF = pd.read_csv("/content/drive/My Drive/RanDonnees/diabetes.csv")
X = diabetesDF[diabetesDF.columns[0:8]]
y = diabetesDF[diabetesDF.columns[-1]]
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.1, random_state=rs, stratify=y)
params = [
    {'MLP__solver': ['adam'],
     'MLP__learning_rate_init': [0.0001],
     'MLP__max_iter': [10000],
     'MLP__hidden_layer_sizes': [(10,20,20,10), (20,50,50,20), (50,100,100,50)],
     'MLP__activation': ['tanh'],
     'MLP__alpha': [0.0001, 0.001, 0.05],
     'MLP__early_stopping': [False]}
]
grid = GridSearchCV(pipeline, param_grid=params, cv=5,n_jobs=-1)
grid.fit(X_train, y_train)
print('Test score = {}'.format(grid.score(X_test,y_test)))
print(grid.best_params_)
res_df = pd.DataFrame({'mean_test_score': grid.cv_results_['mean_test_score']})
```



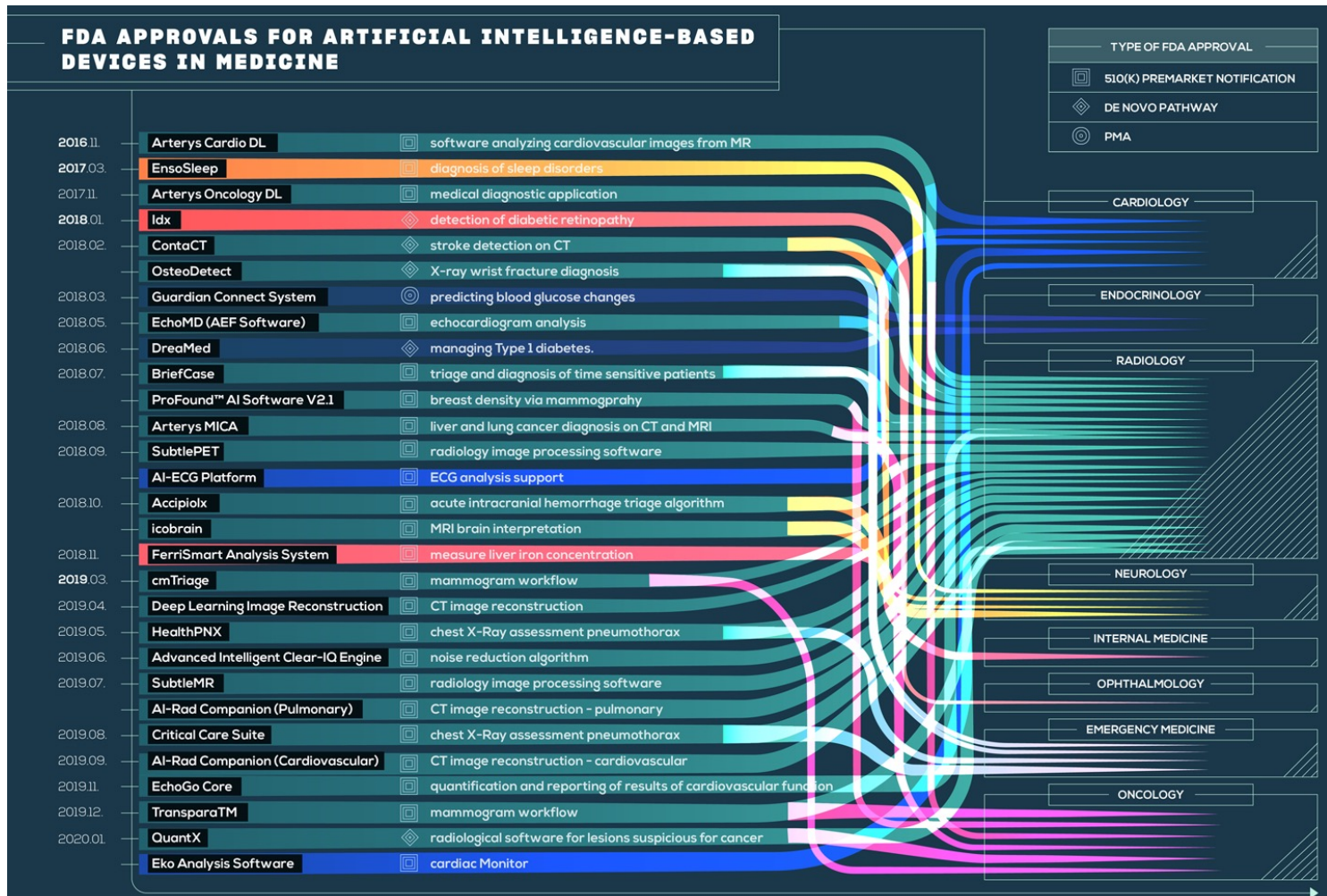
Processing < 20 lignes



Forces et faiblesse des réseaux de neurones

- Puissants et flexibles
- Gourmands en données
- Simple série de calculs algébriques
- « garbage in, garbage out » / « garbage in, carnage out »
- Peu interprétables (« Black box »)
- Mono-taches (« Catastrophe forgetting »)

Quels domaines d'application en santé ?

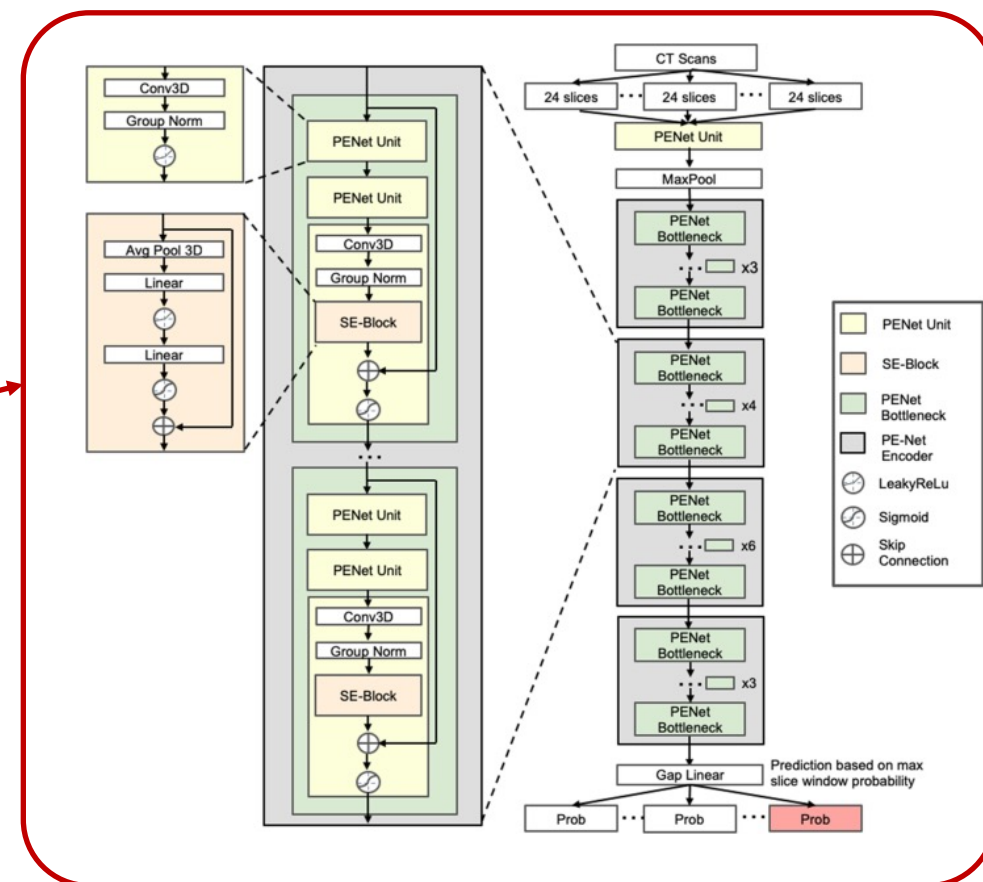
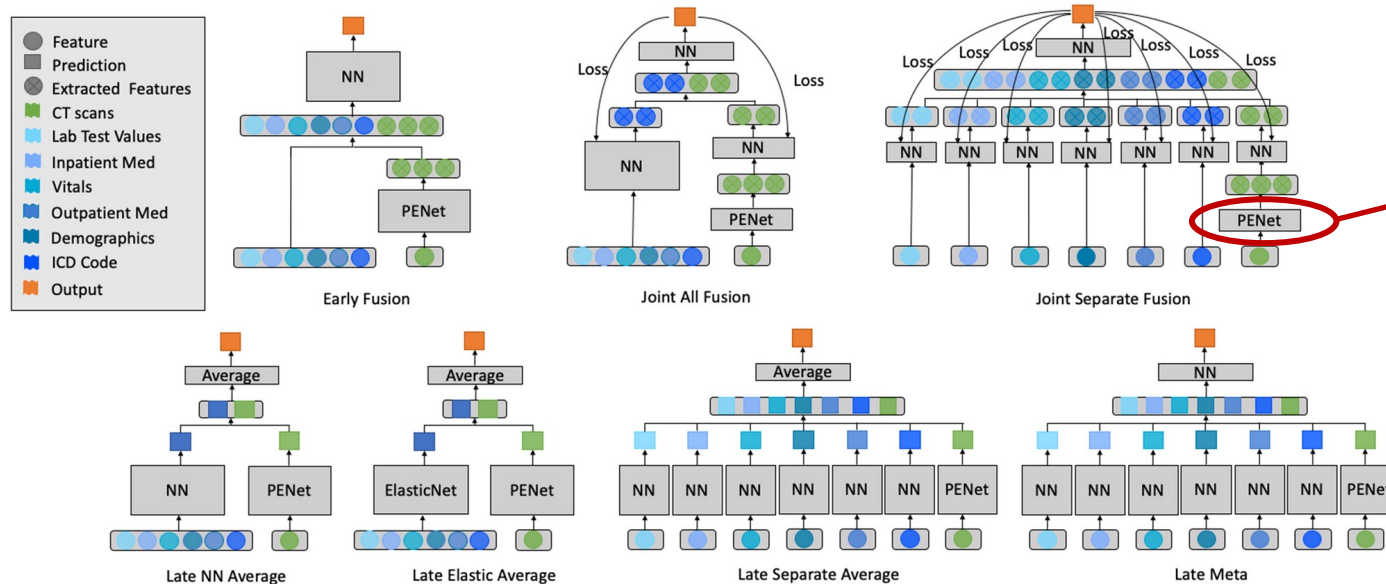


Des architectures très complexes pour l'analyse multimodale

Article | [Open Access](#) | [Published: 17 December 2020](#)

Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection

[Shih-Cheng Huang](#) , [Anuj Pareek](#), [Roham Zamanian](#), [Imon Banerjee](#) & [Matthew P. Lungren](#)





IA et données : promesses et désillusions

Editorial

June 21, 2021

The Epic Sepsis Model Falls Short—The Importance of External Validation

Anand R. Habib, MD, MPhil^{1,2}; Anthony L. Lin, MD¹; Richard W. Grant, MD, MPH^{3,4}

[» Author Affiliations](#) | [Article Information](#)

JAMA Intern Med. 2021;181(8):1040-1041. doi:10.1001/jamainternmed.2021.3333

JAMA Internal Medicine | [Original Investigation](#)

External Validation of a Widely Implemented Proprietary Sepsis Prediction Model in Hospitalized Patients

Andrew Wong, MD; Erkin Otles, MEng; John P. Donnelly, PhD; Andrew Krumm, PhD; Jeffrey McCullough, PhD; Olivia DeTroyer-Cooley, BSE; Justin Pestrue, MEcon; Marie Phillips, BA; Judy Konye, MSN, RN; Carleen Penozza, MHSA, RN; Muhammad Ghous, MBBS; Karandeep Singh, MD, MMSc

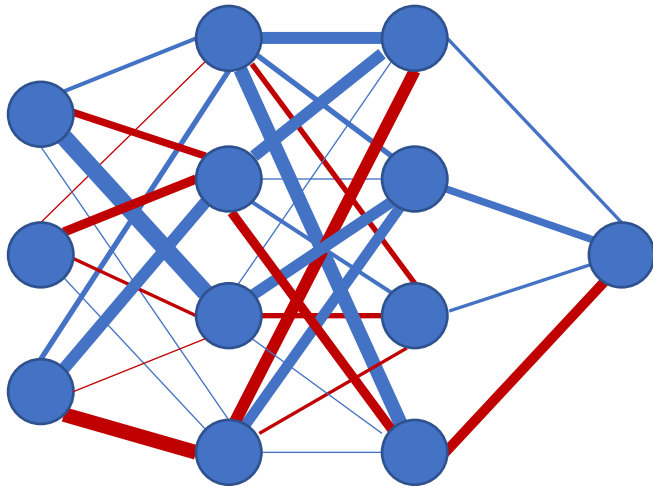
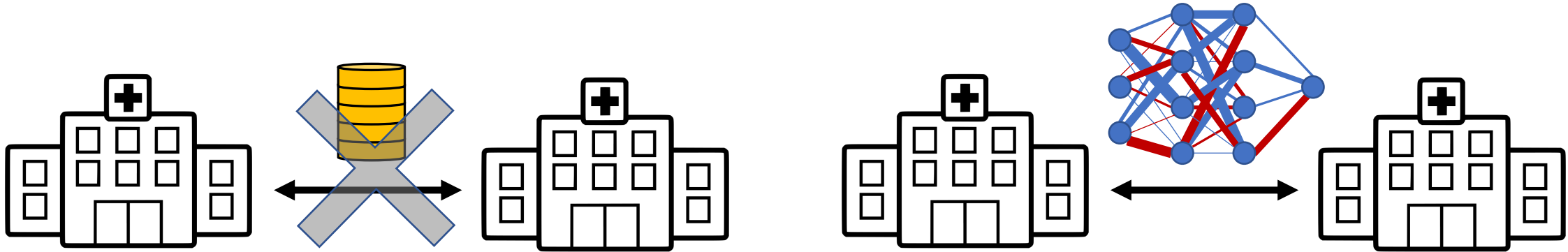


IA et données : promesses et désillusions dans le modèle sepsis

- **Identifie 7% de sepsis non identifiés par des clinicien.ne.s**
- **N'identifie pas 67 % des sepsis identifiés par des clinicien.ne.s**
- **Génère des alertes pour 18% des patient.e.s. alors que le sepsis concerne seulement 7% des cas**

Le cœur du problème reste l'accès aux données

Comment faire une IA plus généralisable ? L'apprentissage fédéré



=

```
from sklearn.neural_network import MLPClassifier
steps = [
    ('scaler', StandardScaler()),
    ('KB', SelectKBest(score_func=f_classif, k=4)),
    ('MLP', MLPClassifier(random_state=rs))
]
pipeline = Pipeline(steps)
diabetesDF = pd.read_csv('/content/drive/My Drive/RanDonnees/diabetes.csv')
X = diabetesDF[diabetesDF.columns[0:8]]
y = diabetesDF[diabetesDF.columns[-1]]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.1, random_state=rs, stratify=y)
params = [
    {'MLP__solver': ['adam'],
     'MLP__learning_rate_init': [0.0001],
     'MLP__max_iter': [10000],
     'MLP__hidden_layer_sizes': [(10,20,20,10), (20,50,50,20), (50,100,100,50)],
     'MLP__activation': ['tanh'],
     'MLP__alpha': [0.0001, 0.001, 0.05],
     'MLP__early_stopping': [False]}
]
grid = GridSearchCV(pipeline, param_grid=params, cv=5, n_jobs=-1)
grid.fit(X_train, y_train)
print('Test score = {}'.format(grid.score(X_test, y_test)))
print(grid.best_params_)
res_df = pd.DataFrame({'mean_test_score': grid.cv_results_['mean_test_score']})
```

+ [1.3, -0.7, 3.43, ..., 1.2]

Développer des modèles d'IA plus robustes

- Accès aux modèles
 - Accès aux données (...)
 - Accès aux métadonnées
 - Publier des résultats négatifs
 - Développer des méthodes d'évaluation robustes
 - Rendre les modèles plus explicables
- Science Ouverte**