



## Une histoire de l'intelligence artificielle : de la logique formelle à l'IA générative

Sorbonne Université - Faculté de Santé

1.4.1 : connaître les grands enjeux liés à l'intelligence artificielle, aux algorithmes, aux biais et aux systèmes d'aide à la décision ainsi que les principes éthiques associés aux traitements des données de santé

Licence CC BY-NC-ND 4.0

<https://creativecommons.org/licenses/by-nc-nd/4.0/>



03/04/2026

Intelligence artificielle, test de Turing, approches symboliques de l'IA, règles de décision, moteur d'inférences, bases de connaissances, *evidence-based medicine*, approches numériques de l'IA, apprentissage machine (*machine learning*), apprentissage profond (*deep learning*), IA générative.

# Une histoire de l'intelligence artificielle : de la logique formelle à l'IA générative

Pr Brigitte SEROUSSI

Sorbonne Université, Faculté de Santé

2025 - 2026



# OBJECTIFS PÉDAGOGIQUES

- Connaître le test de Turing
- Connaître les approches symboliques de l'IA
- Connaître les approches numériques de l'IA



# PLAN



**1 Naissance de l'IA**



**2 IA symbolique et systèmes experts**



**3 IA numérique et Machine learning / Deep learning**



**4 IA générative**

# Définitions et principes généraux

- L'IA est un ensemble de techniques informatiques visant à reproduire des fonctions cognitives humaines, telles que :
  - le raisonnement,
  - l'apprentissage,
  - la perception (vision, langage),
  - la prise de décision
- Elle peut se diviser en deux grandes approches :
  - IA symbolique (ou à base de connaissances),
  - IA numérique / connexionniste (ou à base de données) comme le *machine learning* ou le *deep learning*.



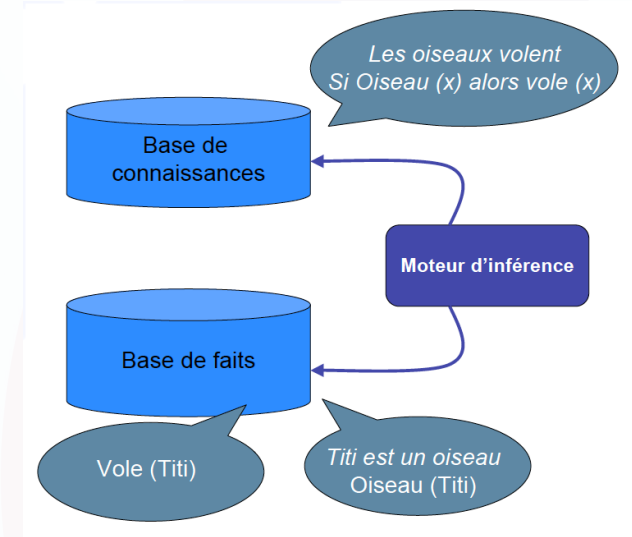
# Les origines de l'IA : peut-on faire penser une machine ?

- En 1950, Alan Turing, mathématicien britannique, publie un article devenu célèbre : “Computing Machinery and Intelligence”. Il y propose ce qu'on appellera plus tard le test de Turing : si une machine peut dialoguer de façon à être confondue avec un humain, peut-on dire qu'elle pense ?
- Six ans plus tard, en 1956, lors d'un colloque d'été au *Dartmouth College* (États-Unis), un petit groupe de chercheurs — dont John McCarthy, Marvin Minsky, Claude Shannon — pose les bases de ce nouveau champ de recherche qu'ils nomment “Artificial Intelligence”. C'est la naissance officielle de l'IA.



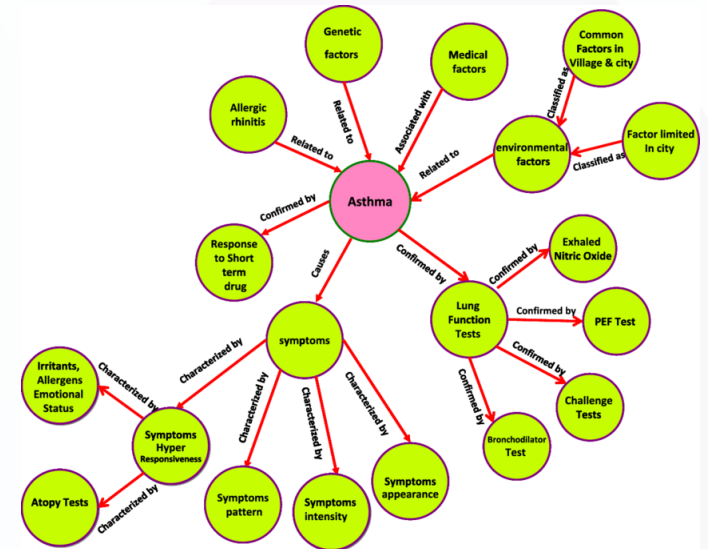
# Les premiers pas : des règles, pas encore des neurones

- Dans les années 1960–1970, l'IA progresse grâce à l'approche symbolique. C'est l'époque des systèmes experts, capables de reproduire le raisonnement d'un médecin ou d'un expert engagé dans une démarche cognitive.
- 3 modules :
  - Base de connaissances : représentation explicite et déclarative de connaissances symboliques
  - Moteur d'inférences : articulation logique des connaissances (déduction, induction, abduction)
  - Base de faits : ensemble de critères patient



# L'IA à base de connaissances

- L'IA à base de connaissances :
  - Une représentation formelle du savoir médical.
- Formalismes de représentation des connaissances
  - Des règles logiques du type "SI symptômes A et B ALORS maladie X" utilisant la logique des prédicats (ex. "Fièvre(x)  $\wedge$  Toux(x)  $\rightarrow$  Grippe(x)")
  - Réseaux sémantiques : structures en graphes reliant des concepts.
  - Des ontologies (organisations hiérarchiques des concepts médicaux, ex : SNOMED, MeSH),



# L'exemple de Mycin

- 1972 : un des premiers systèmes capables d'aider à diagnostiquer des infections bactériennes et à proposer un traitement
- Représentation des connaissances sous la forme de règles logiques "SI... ALORS..."

```
      param  context  operation value
      -----  -
      (defrule 52
        if (site culture is blood )
           (gram organism is neg )
           (morphl organism is rod )
           (burn patient is serious)
        then 0.4
           (identity organism is pseudomonas )
      )
```

<--- rule number  
<--- premise/condition  
<--- confidence factor (cf)  
<--- conclusion

Patient's name: **SYLVIA\_FISHER**

Sex: **FEMALE**

Age: **27**

----- culture-1 -----

From what site was specimen CULTURE-1 taken? **BLOOD**

How many days ago was this culture (CULTURE-1) obtained? **3**

----- organism-1 -----

Enter the identity (genus) of ORGANISM-1? **UNKNOWN**

The gram strain of ORGANISM-1? **?**

Must be one of: acid-fast, pos, neg

The gram strain of ORGANISM-1? **NEG**

Is ORGANISM-1 a rod or coccus? **ROD**

Is PATIENT-1 a burn patient? If so, mild or serious? **WHY**

[Why is the value of burn being asked for?]

It is known that:

The site of the culture is blood

The gram of the organism is neg

The morphology of the organism is rod

Therefore,

Rule 52:

If The burn of the patient is serious

Then with certainty of 0.4

The identity of the organism is pseudomonas

Is PATIENT-1 a burn patient? If so, mild or serious? **SERIOUS**

What is the aerobicity of ORGANISM-1? **AEROBIC**

Is PATIENT-1 a compromised host? **YES**

Findings for ORGANISM-1

for these goals: identity

IDENTITY: ENTERO 0.8, PSEUDOMONAS 0.76

# Evaluation de l'IA à base de connaissances

- Séduction initiale des médecins
  - IA explicable, justification qualitative des raisonnements
  - Technoscience « magique »
- Déception en pratique
  - Faibles performances en dehors des cas jouets
  - Critique des bases de connaissances
  - Subjectivité des experts mise en question
- Médecins exclus du raisonnement médical



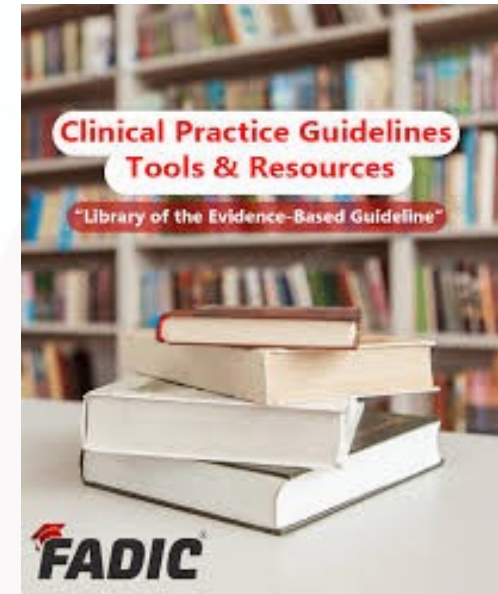
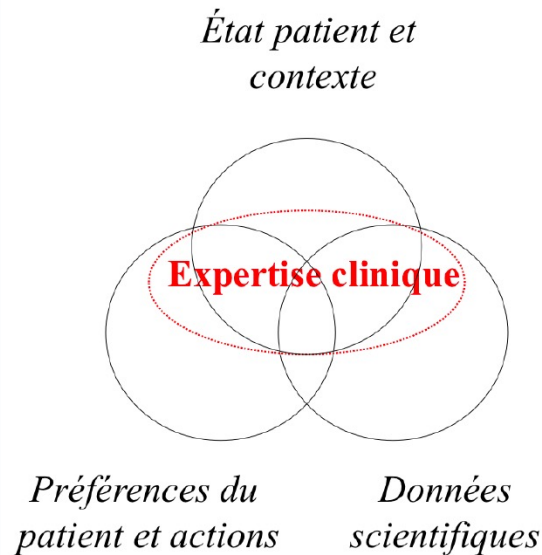
# L'hiver de l'IA : quand les promesses ne sont pas tenues ...

- Les années 1980–1990 voient l'enthousiasme retomber
- Les promesses de l'IA ne se concrétisent pas
  - Les systèmes experts sont limités par des règles considérées comme « rigides », et ne savent pas apprendre par eux-mêmes.
  - L'IA symbolique est efficace dans des domaines bien formalisés, mais limitée face à l'incertitude et la complexité du monde réel.
- Faute de résultats pratiques et de puissance de calcul, les financements s'effondrent.
- On parle alors d'« hiver de l'IA ».



# Le retour de l'IA symbolique : l'émergence de l'evidence-based medicine

- Evidence-Based Medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.
- The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research.

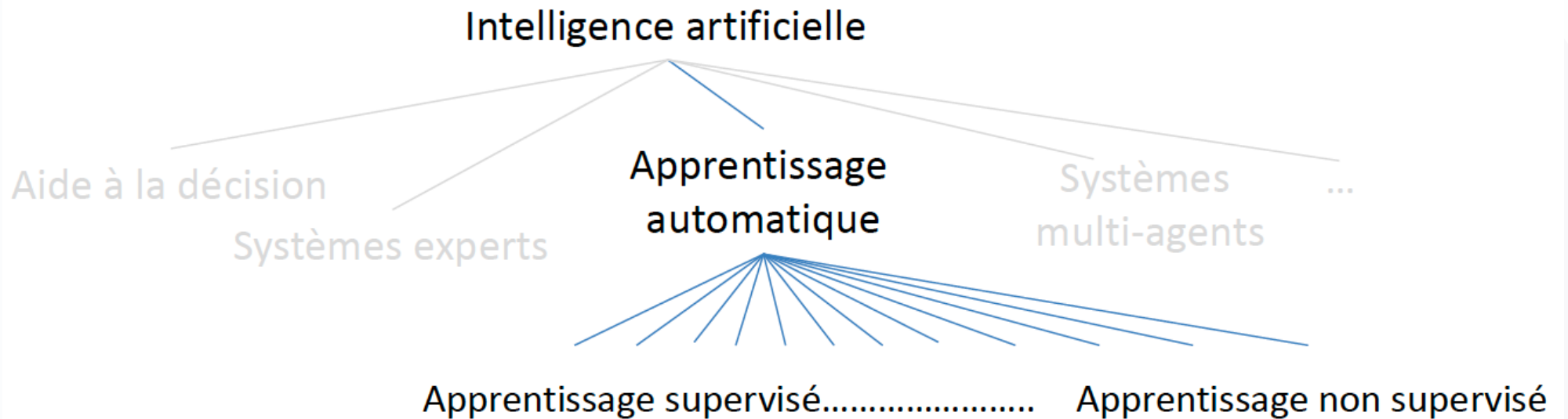


# L'IA numérique : l'apprentissage par les données

- Dans les années 1990, une nouvelle approche émerge : le *machine learning* (ou apprentissage machine).
- Le ML apprend des modèles à partir de données (plutôt que de suivre des règles programmées).
- Deux grandes familles :
  - Apprentissage supervisé : on apprend à partir d'exemples annotés (ex. "cancer" vs "pas cancer"),
  - Apprentissage non supervisé : on cherche des motifs cachés (clustering, réduction de dimensions).

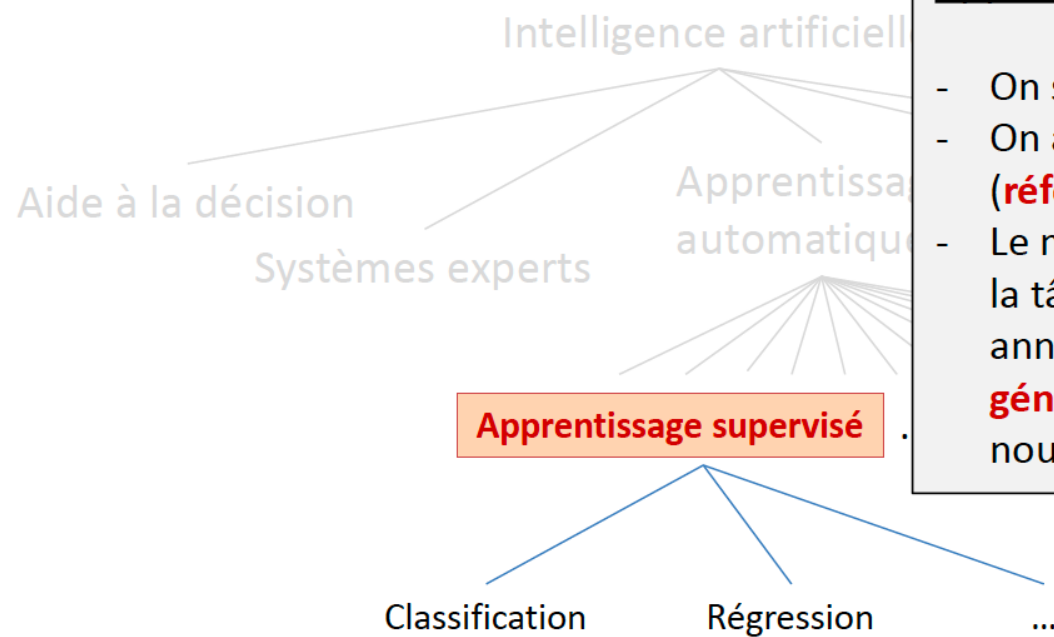


# L'apprentissage automatique : rappels méthodologiques



Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques



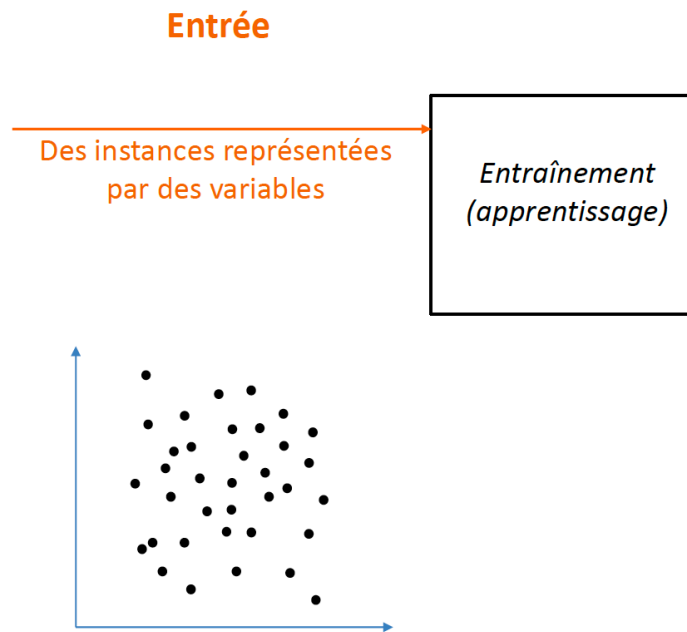
## Apprentissage supervisé

- On sait ce que l'on cherche
- On a des données « annotées »  
**(référence, gold standard)**
- Le modèle doit apprendre à résoudre la tâche en utilisant les instances annotées, pour être capable de **généraliser** ensuite sur des instances nouvelles

Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

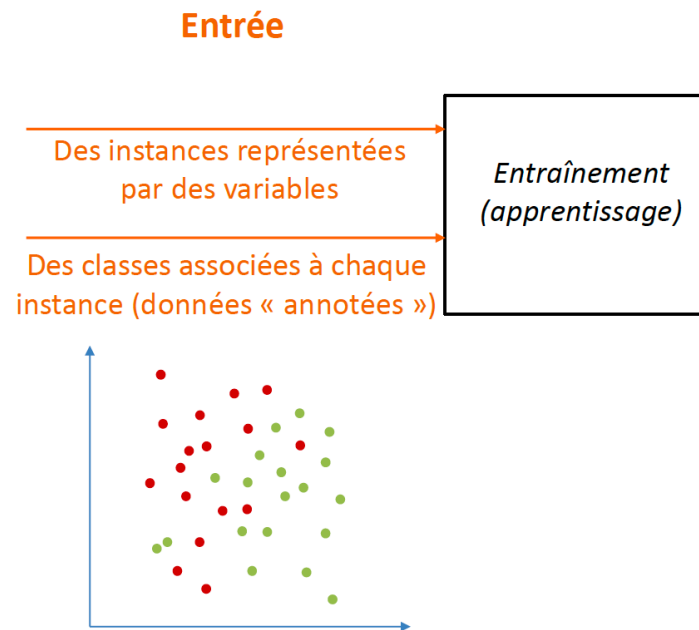
## Classification supervisée



Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

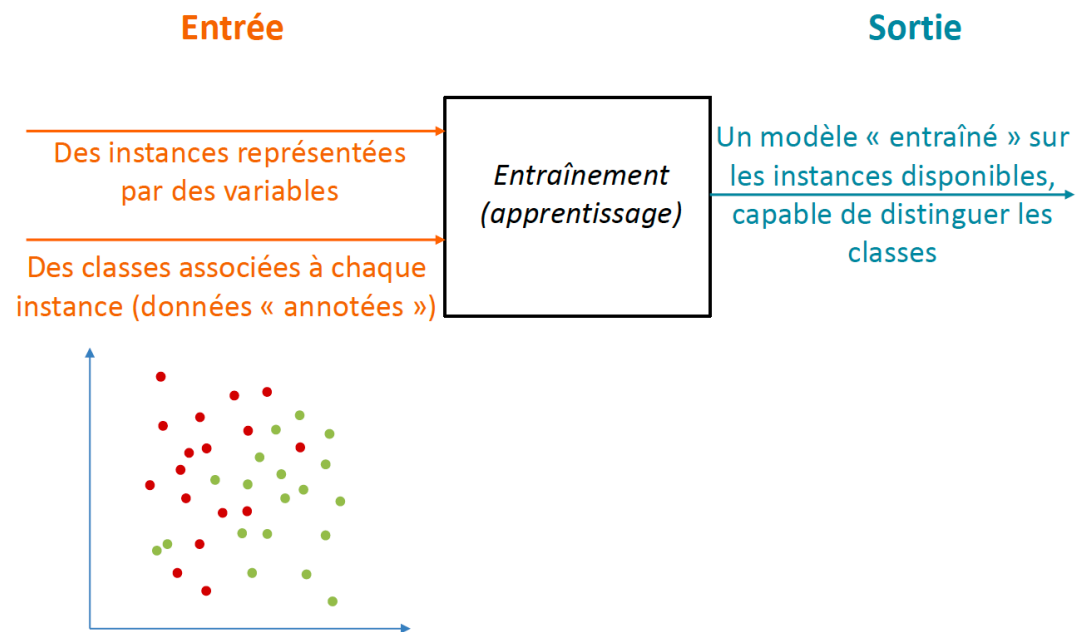
## Classification supervisée



Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

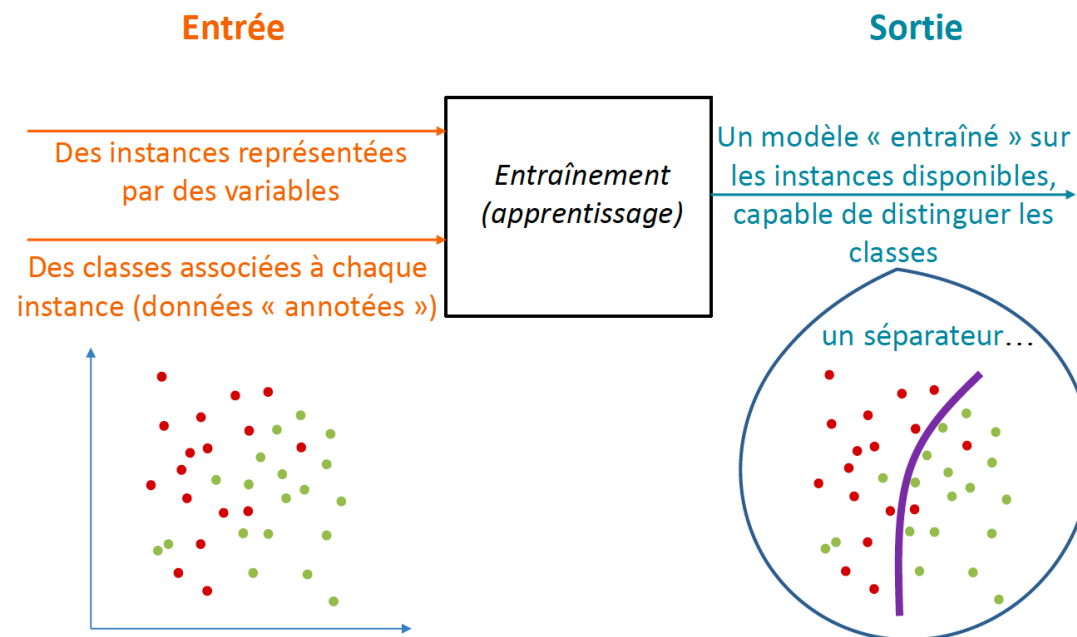
## Classification supervisée



Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée

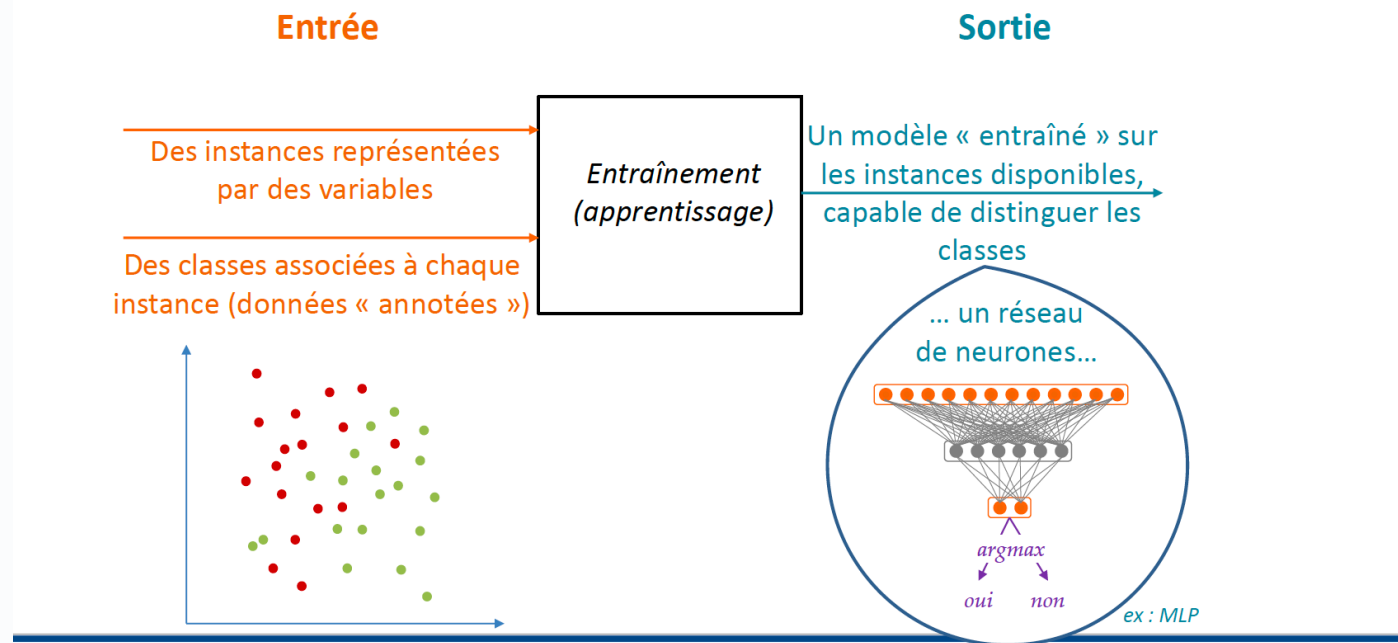


Slides mises à disposition par le Pr Xavier Tannier



# L'apprentissage automatique : rappels méthodologiques

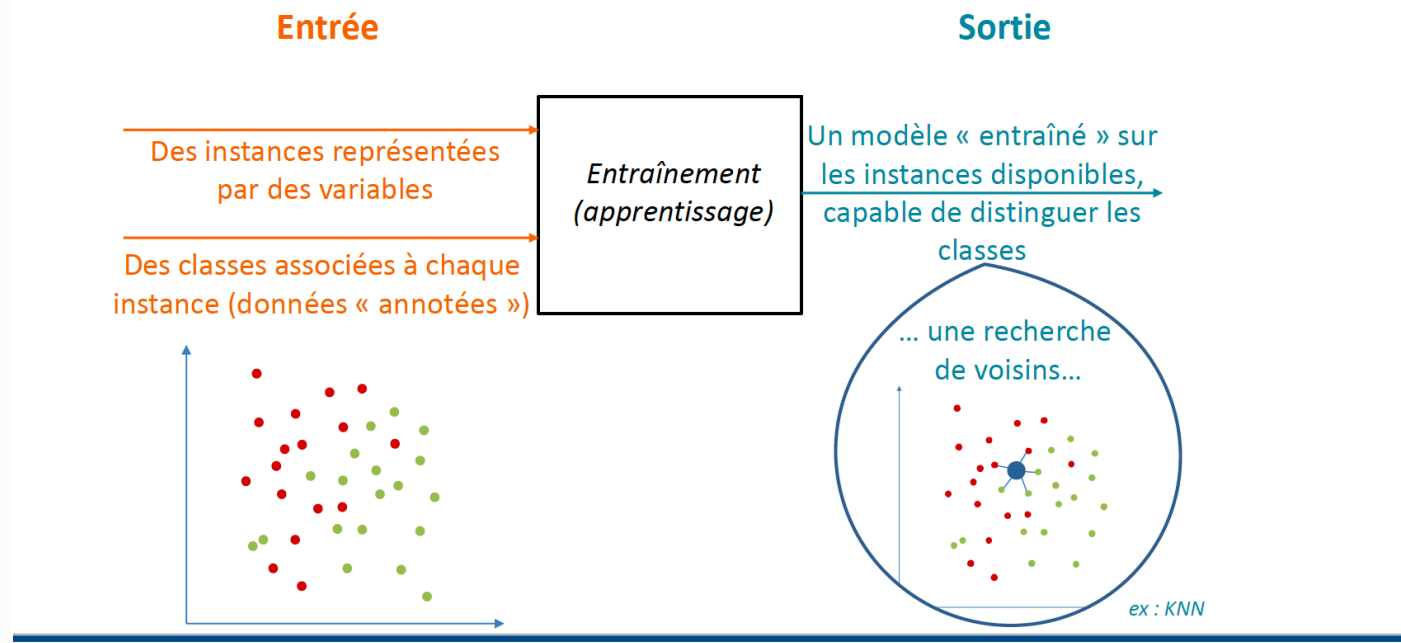
## Classification supervisée



Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée

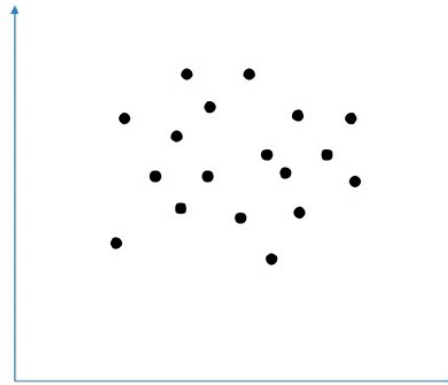


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : inférence

- On applique ensuite le modèle sur de nouvelles données, « non annotées »

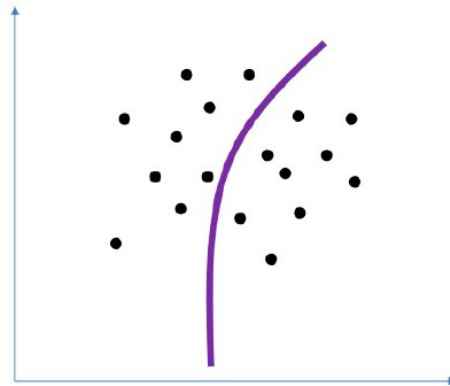


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : inférence

- On applique ensuite le modèle sur de nouvelles données, « non annotées »

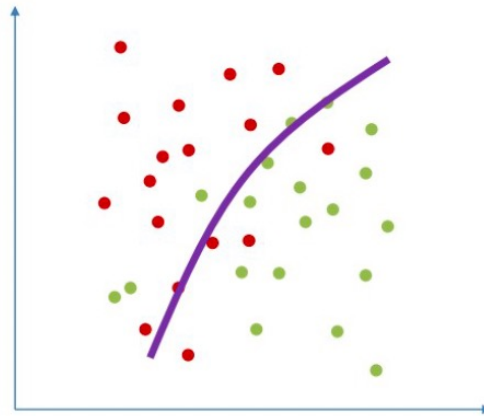


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : évaluation

Jeu de test : prédiction du système

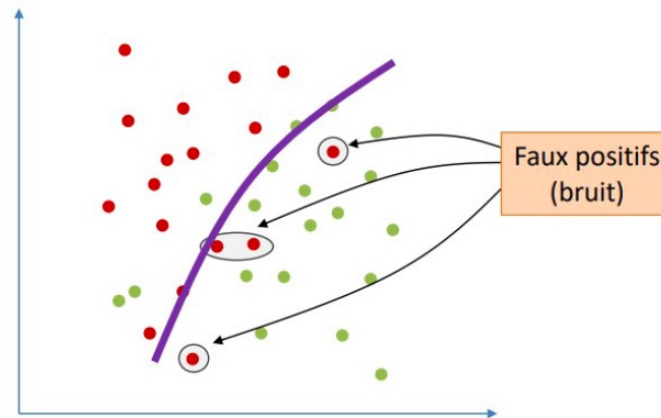


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : évaluation

Jeu de test : classe « VERT »

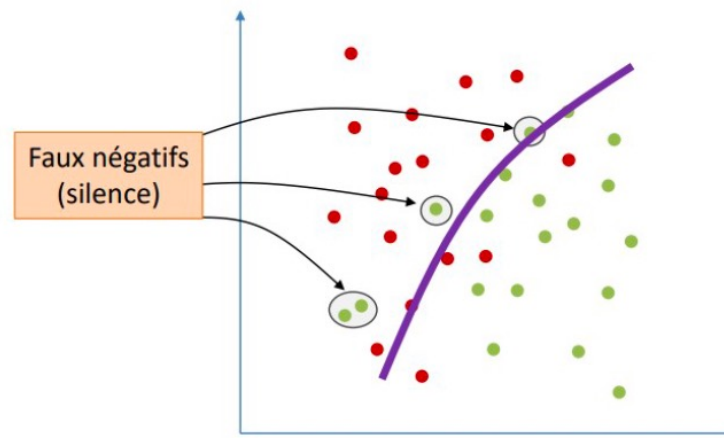


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : évaluation

Jeu de test : classe « VERT »

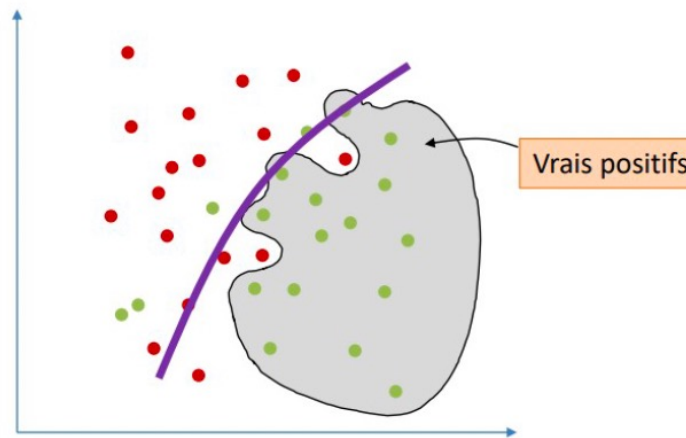


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : évaluation

Jeu de test : classe « VERT »

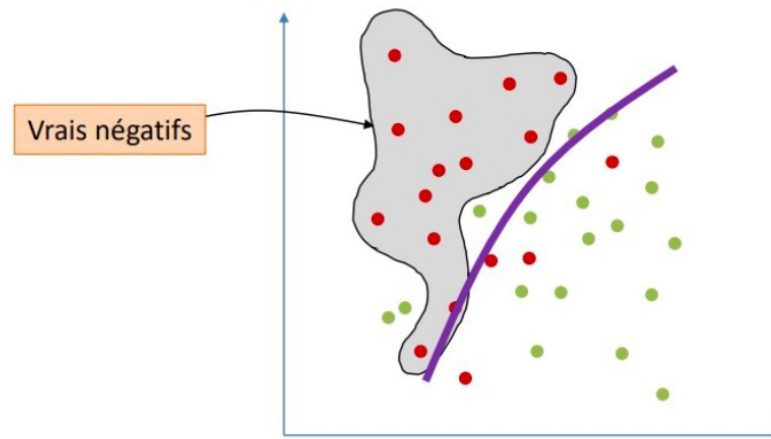


Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : évaluation

Jeu de test : classe « VERT »



Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## Classification supervisée : évaluation

- **Précision** (ou valeur prédictive positive)  
le système a-t-il raison quand il annonce la classe ?
- **Rappel** (ou sensibilité) : le système manque-t-il beaucoup d'éléments de la classe ?
- **Spécificité** =  $\frac{\text{Vrais négatifs}}{\text{Vrais négatifs} + \text{Faux positifs}}$  le système évite-t-il les fausses alarmes ?
- **F-mesure** = moyenne harmonique de précision et rappel  
$$= \frac{(1+\beta^2) \times (\text{Précision} \times \text{Rappel})}{(\beta^2 \times \text{Précision} + \text{Rappel})}$$
- **F<sub>1</sub>-mesure** =  $2 \frac{\text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$

Slides mises à disposition par le Pr Xavier Tannier

# L'apprentissage automatique : rappels méthodologiques

## L'interprétabilité

Le système sait-il « expliquer » ses prédictions ?

- Quelle partie de l'entrée a conduit à la prédiction ?
  - Saillance et perturbations adversariales
  - Support aux réponses à des questions



- Quelles règles de décision ont conduit à la prédiction ?
- Quels exemples du jeu d'entraînement ont causé la prédiction ?
  
- Intégration de la notion d'interprétabilité dans les modèles (mise en abyme)

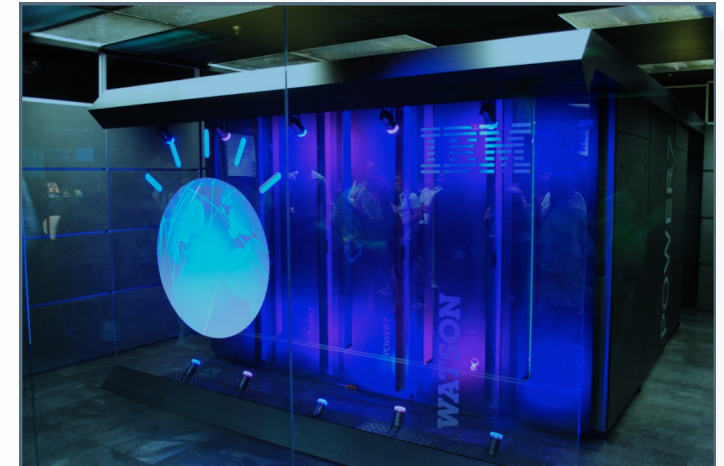
Yonatan Belinkov, Sebastian Gehrmann, Ellie Pavlick  
Interpretability and Analysis in Neural NLP  
ACL 2020

Eric Wallace, Matt Gardner, Sameer Singh  
Tutorial on Interpreting Predictions of NLP Models  
EMNLP 2020

Slides mises à disposition par le Pr Xavier Tannier

# L'IA numérique : le système Watson

- Grâce à des algorithmes statistiques et à l'augmentation des données disponibles, les performances s'améliorent peu à peu. En 2011, le système Watson d'IBM bat deux champions humains au jeu Jeopardy! en répondant à des questions complexes en langage naturel.



# La révolution du deep learning : l'IA apprend à voir, entendre, parler

- 2012 : AlexNet, un modèle de *deep learning* (ou apprentissage profond) gagne une compétition de reconnaissance d'images. Il utilise un réseau de neurones artificiels très profond pour apprendre à identifier des objets... et il surpasse de loin tous les autres modèles.
- Le *deep learning* explose dans les années qui suivent :
  - Lecture d'images médicales (radiologie, pathologie),
  - Interprétation de signaux biologiques,
  - Génération de texte ou de la parole.
- Le *deep learning*, une sous-branche du ML, est fondé sur les réseaux de neurones profonds (ex : CNN pour images, LSTM pour données séquentielles).



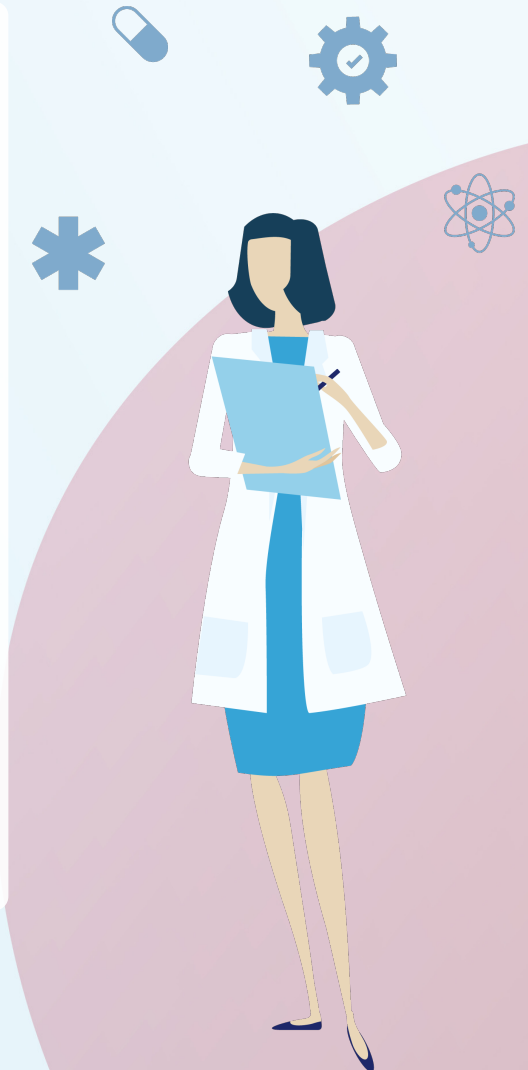
## Émergence de l'IA générative (depuis 2022)

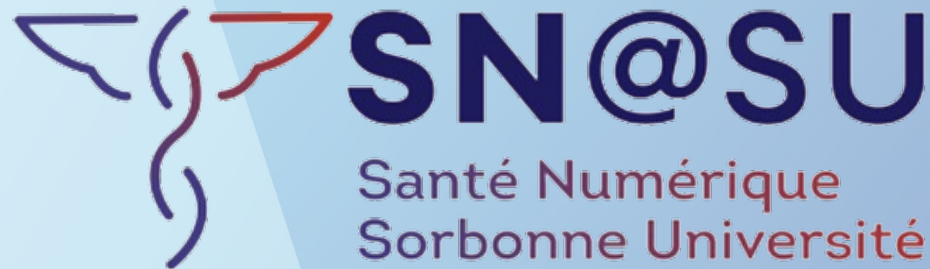
- Nouveaux modèles capables de générer du texte, des images, du code, voire des diagnostics.
- Basés sur les grands modèles de langage (LLMs) comme :
  - GPT (OpenAI),
  - BERT (Google),
  - Med-PaLM ou BioGPT pour les applications médicales.
- L'IA générative permet :
  - La rédaction automatique de comptes rendus médicaux,
  - L'aide à la documentation des dossiers patients,
  - La génération d'images médicales synthétiques pour l'entraînement de modèles.



## TAKE HOME MESSAGES

- L'IA a plus de 70 ans d'histoire, avec des phases d'enthousiasme et de stagnation.
- Elle a évolué de la logique symbolique à l'apprentissage par les données, puis à la génération de contenus.
- En médecine, l'IA moderne est déjà en pratique.





*Ce travail a bénéficié d'une aide de l'État gérée par l'Agence Nationale de la Recherche au titre de France 2030 portant la référence ANR-23-CMAS-0001*

*Cette ressource pédagogique est placée sous la licence CC-BY-NC-ND 4.0*