



Image générée par IA



Image générée par IA

Arnaud Ferré, Chargé de Recherche en *BioNLP*



Image générée par des humains

Extrait de la BD « Sciences en bulles »

Arnaud Ferré : Des robots qui apprennent à lire
Peb&Fox/Syndicat national de l'édition, CC BY-NC-ND2

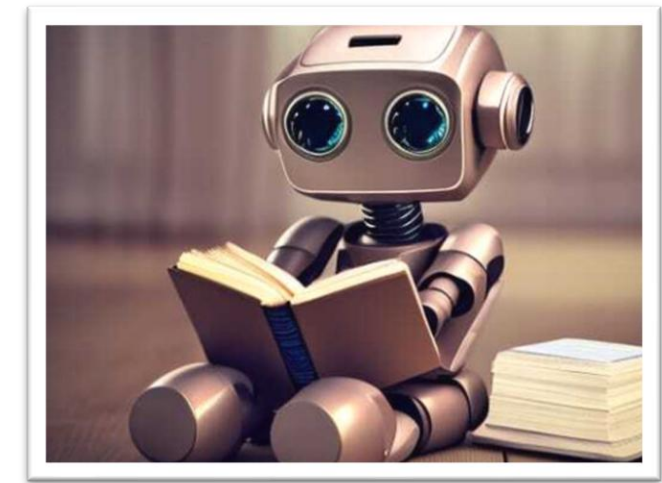


Image générée par IA

Un peu d'histoire de l'IA



L'IA : deux écoles principales

Intelligence Artificielle

Apprentissage automatique

(machine learning)
(IA connexionniste)



Induction	
Platon est un humain	Aristote est un humain
Platon est mortel	Aristote est mortel

⇒ Tous les humains sont mortels

IA symbolique

Déduction

Tous les humains sont mortels
Socrate est un humain

⇒ Socrate est mortel



L'IA : deux écoles principales

Intelligence Artificielle

Apprentissage automatique (machine learning)



Induction	
Platon est un humain	Aristote est un humain
Platon est mortel	Aristote est mortel

⇒ Tous les humains sont mortels

IA symbolique

Déduction
Tous les humains sont mortels
Socrate est un humain



⇒ Socrate est mortel

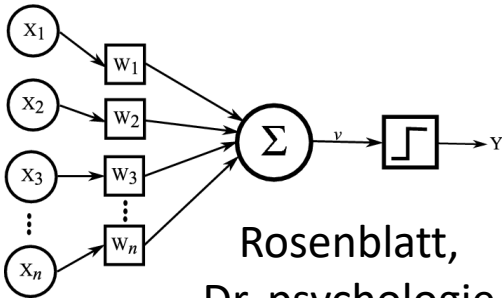
Apprentissage (neuronal) profond (deep learning)



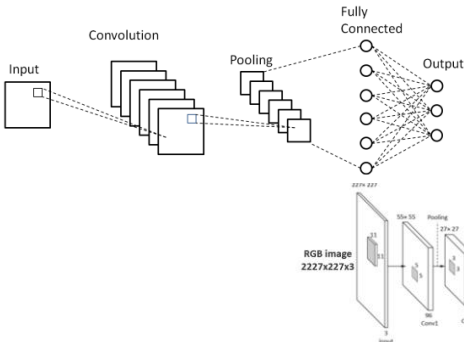
Depuis 2012

Un peu d'histoire...

McCulloch et Pitts,
Dr. neurobio

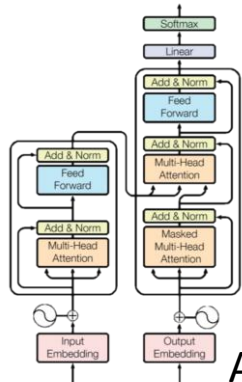


Rosenblatt,
Dr. psychologie



Paul Warbos,
Dr. sciences sociales

2010 : GPU
GTX 580



Ashish Vaswani,
Dr. informatique



Midjourney



2022

30s : naissance
de l'informatique

1943 : neurone formel
1957 : perceptron



1974-1986 : perceptron
multicouche &
rétropropagation



1997-1998 : RNN/CNN

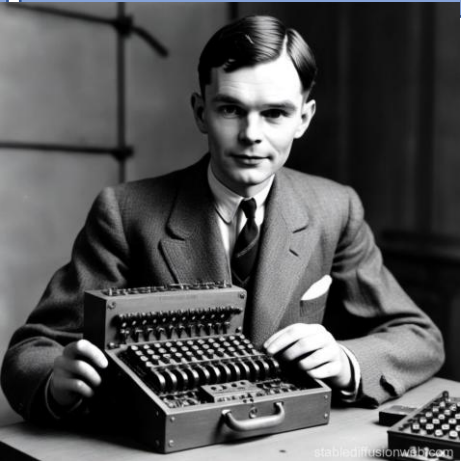
2012 : AlexNet

2014 : Word2Vec

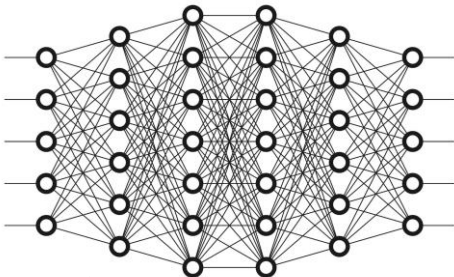
2017 : transformeur

2018 : GPT1/BERT

Ère des IA génératives



Turing et la
machine
Enigma



World's top two cited scientists

Geoffrey Hinton,
Dr. informatique

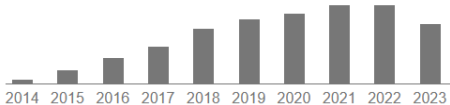
prix Nobel de
Physique 2024

Yoshua Bengio,
Dr. informatique

Yann Le Cun,
Dr. informatique

Tomáš Mikolov,
Dr. informatique

Cité 38279 fois



Jacob Devlin,
Ing. informatique

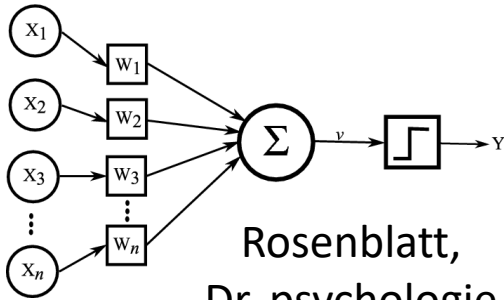
prix Turing
2018

(pour le deep learning)

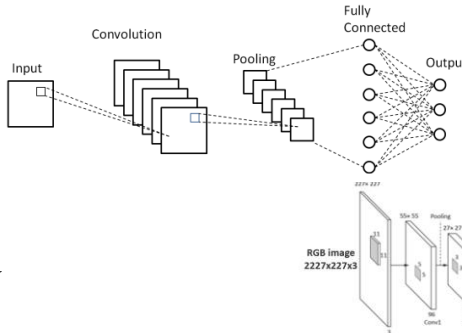


Un peu d'histoire...

McCulloch et Pitts,
Dr. neurobio

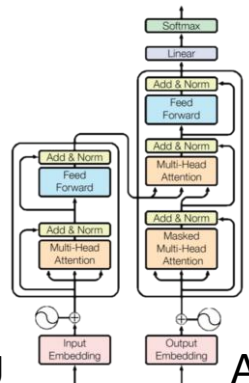


Rosenblatt,
Dr. psychologie



Paul Warbos,
Dr. sciences sociales

2010 : GPU
GTX 580



Ashish Vaswani,
Dr. informatique

2022

30s : naissance
de l'informatique

1943 : neurone formel
1957 : perceptron



1974-1986 : perceptron
multicouche &
rétropropagation



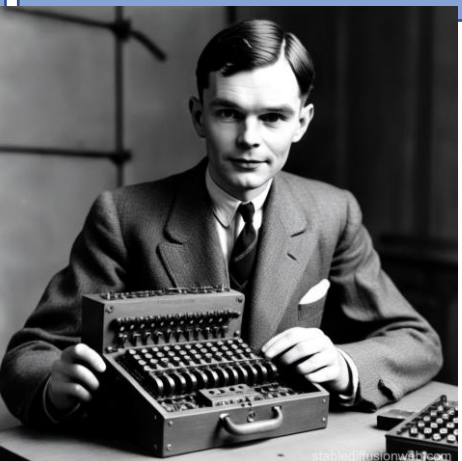
1997-1998 : RNN/CNN

2012 : AlexNet

2017 : transformeur

2014 : Word2Vec

2018 : GPT1/BERT



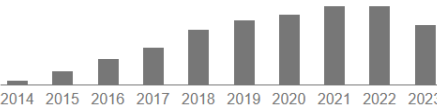
Turing et la
machine
Enigma

1986 : création de la section
informatique au CNRS
(scission de la section math)

1989 : Création de l' AFIA

Tomáš Mikolov,
Dr. informatique

Cité 38279 fois



Jacob Devlin,
Ing. informatique

World's top two cited scientists

Geoffrey Hinton,
Dr. informatique

Yoshua Bengio,
Dr. informatique

Yann Le Cun,
Dr. informatique

prix Nobel de
Physique 2024

prix Turing
2018
(pour le deep learning)

prix Nobel de
Chimie 2024



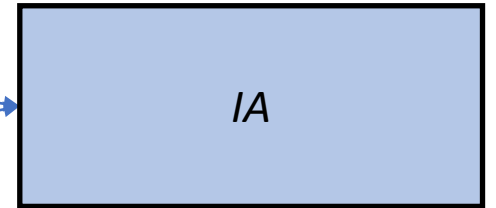
Comment les données brutes (texte, image, etc.) sont-elles transformées pour être utilisables par une machine ?

Comment les données brutes (texte, image, etc.) sont-elles transformées pour être utilisables par une machine ?



« Turing et la machine Enigma »

?



Comment les données brutes (texte, image, etc.) sont-elles transformées pour être utilisables par une machine ?



« Turing et la machine Enigma »

Traduire en nombre



IA

Les cas simples : valeurs numériques

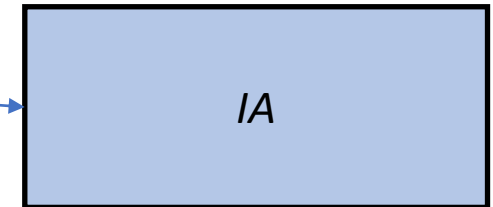
- Observation n°1 :
 - circonférence=**53**cm,
 - hauteur=**14**m
- Observation n°2 :
 - circonférence=**71**cm
 - hauteur=**23**m
- ...



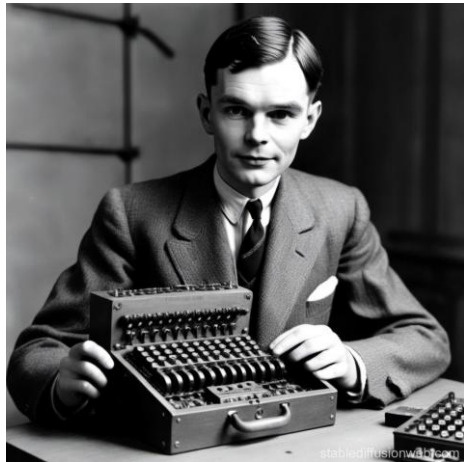
Image générée par IA

Les cas simples : valeurs numériques

- Observation n°1 :
 - circonférence=**53**cm,
 - hauteur=**14**m
- Observation n°2 :
 - circonférence=**71**cm
 - hauteur=**23**m
- ...



La représentation d'images



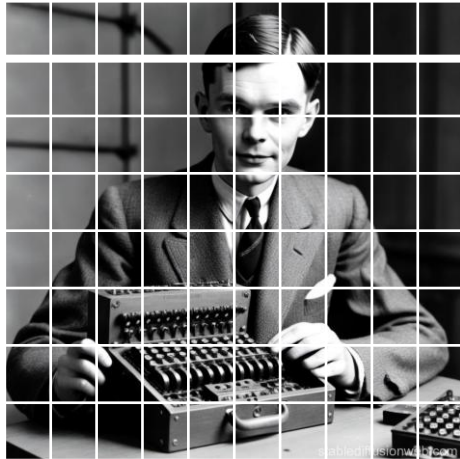
?

Traduire en nombre



IA

La représentation d'images : pixels = valeurs



2	9	7	1	2	3	8	4	7	5
6	2	7	9	5	1	3	5	7	7
9	1	3	5	0	4	7	0	5	9
6	3	1	0	4	8	6	2	2	9
3	0	3	4	7	8	9	6	2	0
0	2	4	5	8	7	4	8	7	0
2	3	4	7	8	9	6	1	0	2
4	8	6	3	7	4	5	8	9	3

Valeurs des pixels

Traduire en nombre

IA

Plus complexe : le texte

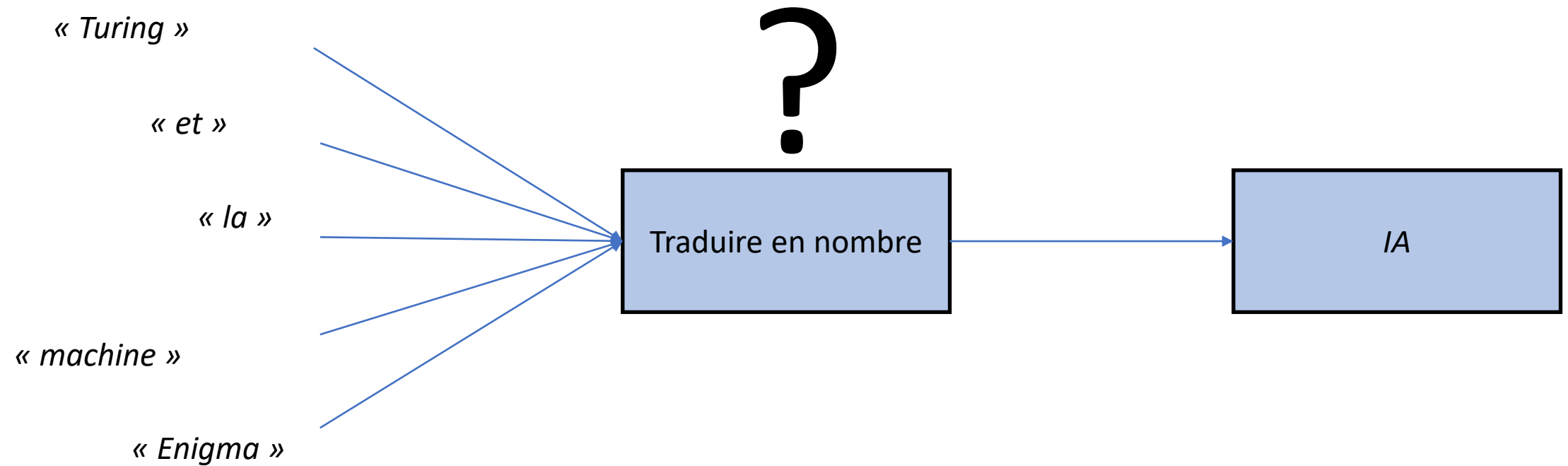
?

Traduire en nombre

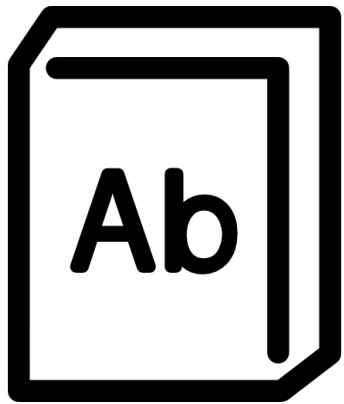
IA

*« Turing et la
machine Enigma »*

Plus complexe : le texte



La représentation de texte 1-parmi-N (*one-hot*)



Dictionnaire

N mots

abaisser

abandon

abas

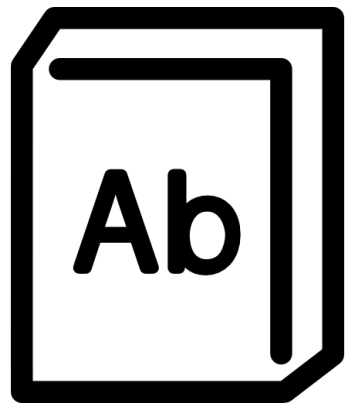
abordage

aborder

...

zythum

La représentation de texte 1-parmi-N (*one-hot*)



Dictionnaire

N mots

abaisser

abandon

abas

abordage

aborder

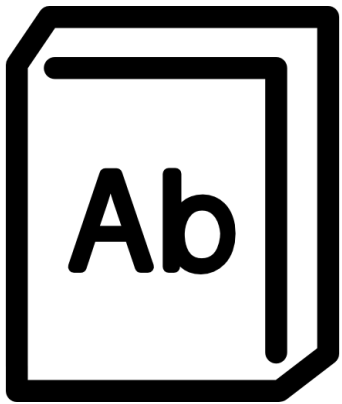
...

zythum

N valeurs numériques

(1, 0, 0, 0, 0, 0, ..., 0)

La représentation de texte 1-parmi-N (*one-hot*)



Dictionnaire

abaisser

abandon

abas

abordage

aborder

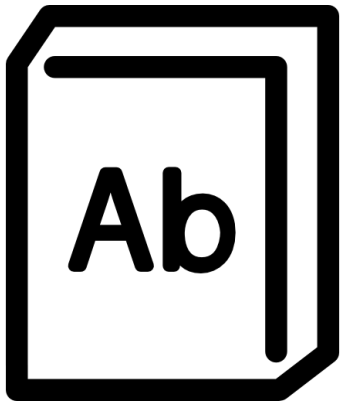
...

zythum

(1, 0, 0, 0, 0, 0, ..., 0)

(0, 1, 0, 0, 0, 0, ..., 0)

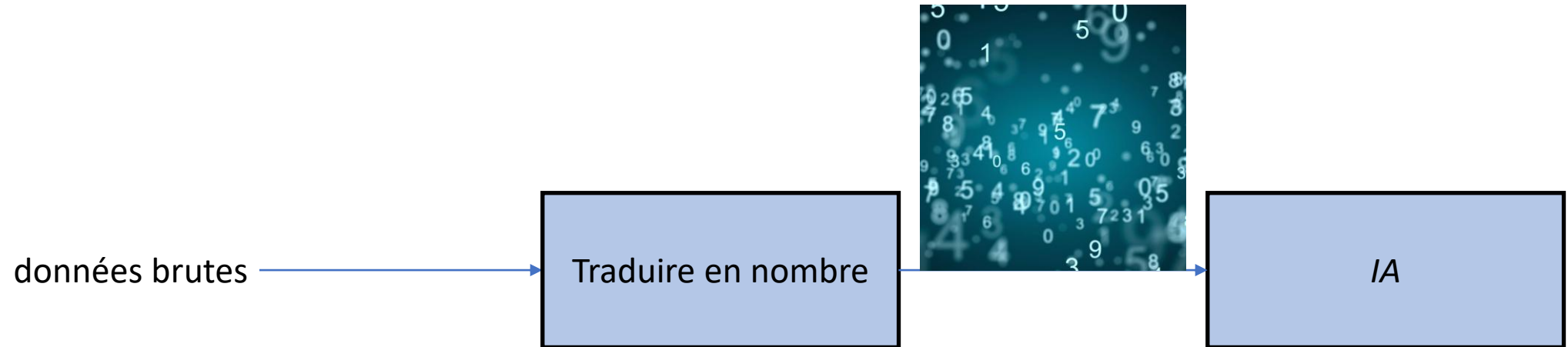
La représentation de texte 1-parmi-N (*one-hot*)



Dictionnaire

abaisser	→	(1, 0, 0, 0, 0, 0, ..., 0)
abandon	→	(0, 1, 0, 0, 0, 0, ..., 0)
abas	→	(0, 0, 1, 0, 0, 0, ..., 0)
abordage	→	(0, 0, 0, 1, 0, 0, ..., 0)
aborder	→	(0, 0, 0, 0, 1, 0, ..., 0)
...		...
zythum	→	(0, 0, 0, 0, 0, 0, ..., 1)

Bilan : une IA manipulent des vecteurs de données

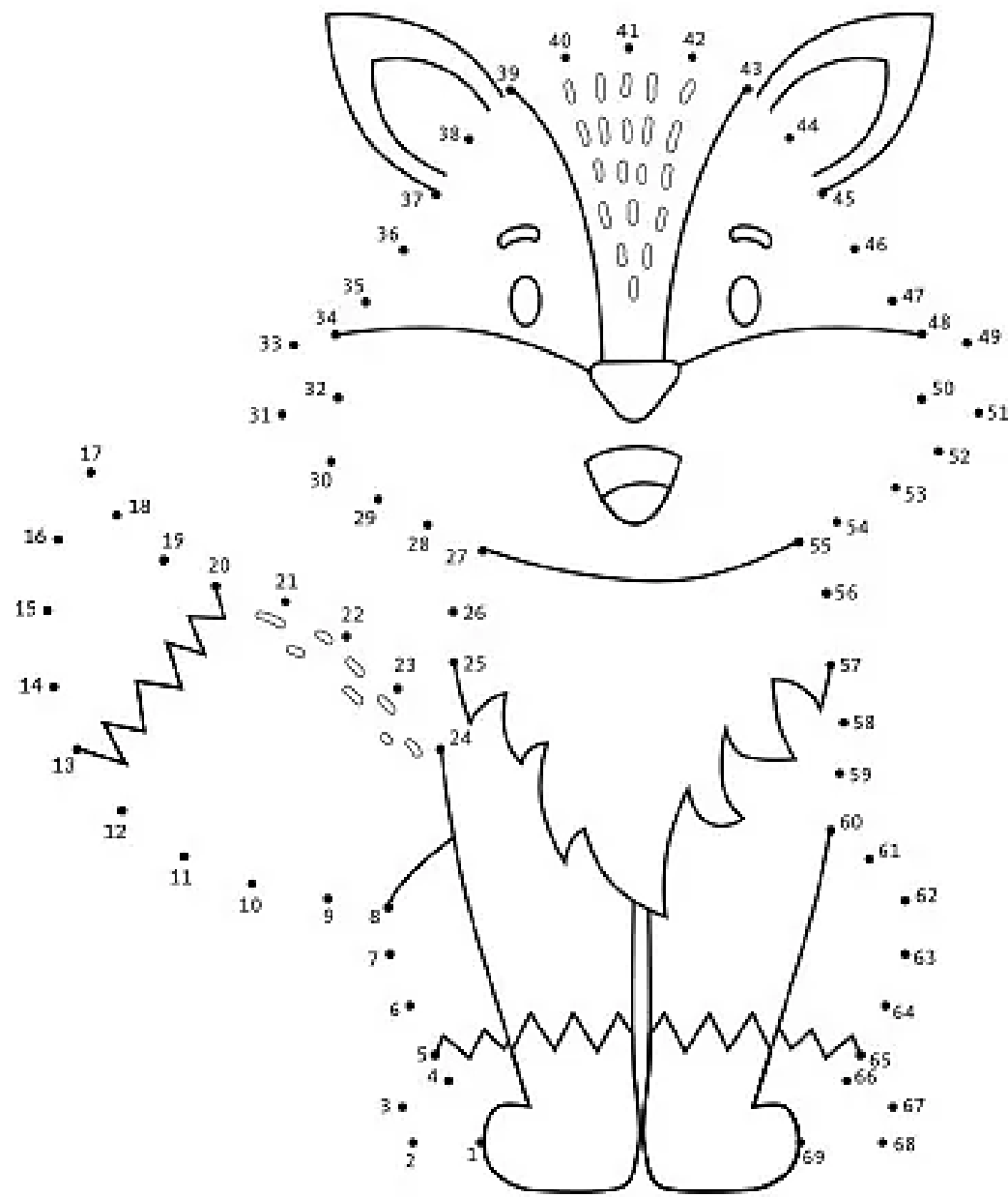


Si données représentées par : (nombre 1, nombre 2, ..., nombre N) ,
on appelle ça un **vecteur** réel (de taille N).

Les IAs modernes fonctionnent quasiment toutes avec des **vecteurs**.

Présentation de l'apprentissage supervisé : Comment apprendre à partir d'observations ?

Apprendre : un jeu d'enfant ?



Apprendre : un jeu d'enfant ?



Apprendre : un jeu d'enfant ?

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?



Image générée par IA

Apprentissage : la régression supervisée

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?



Image générée par IA

Apprentissage : la régression supervisée

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?



Image générée par IA

Apprentissage : la régression supervisée

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?

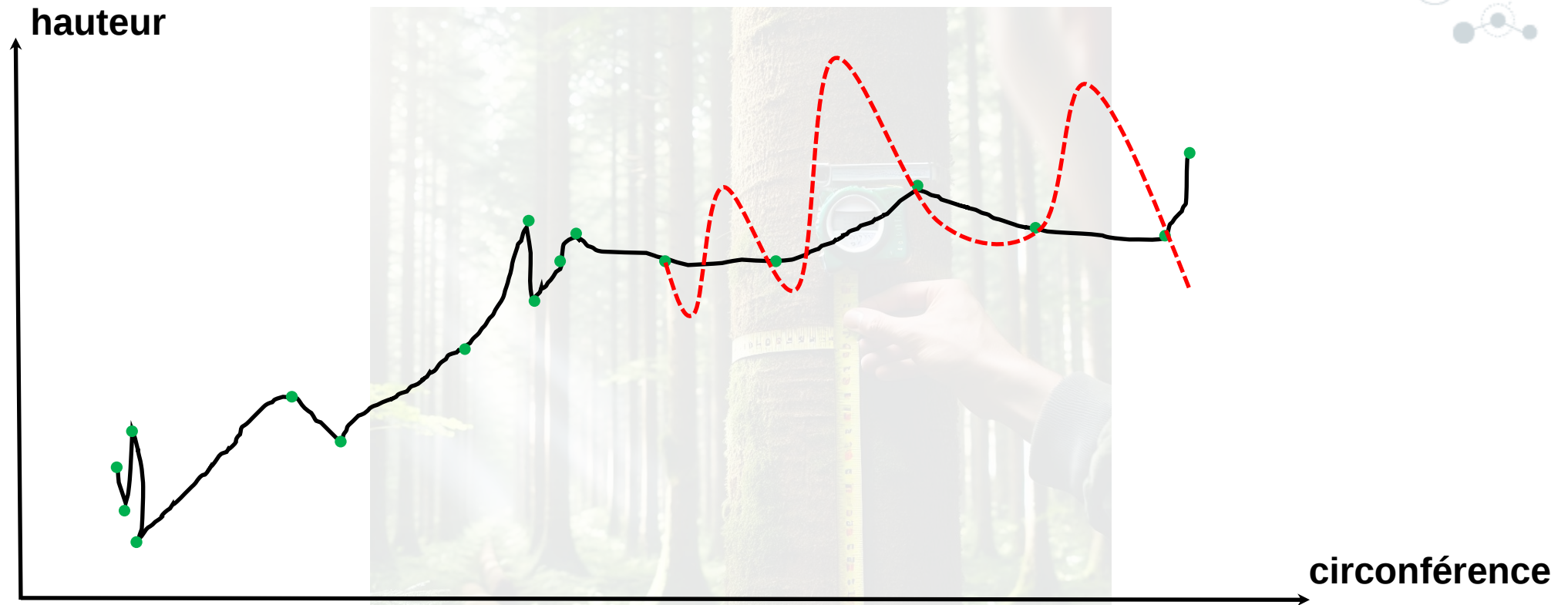


Image générée par IA

Apprentissage : la régression supervisée

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?

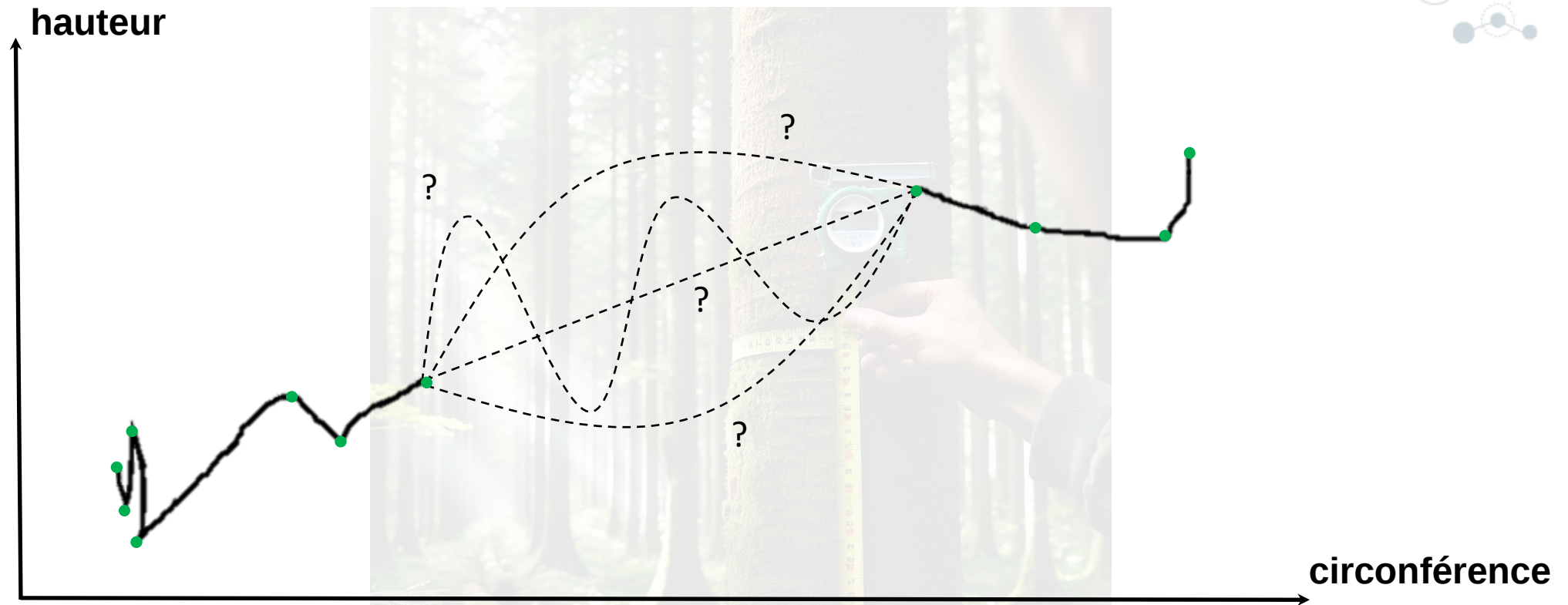


Image générée par IA

Apprentissage : la régression supervisée

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?

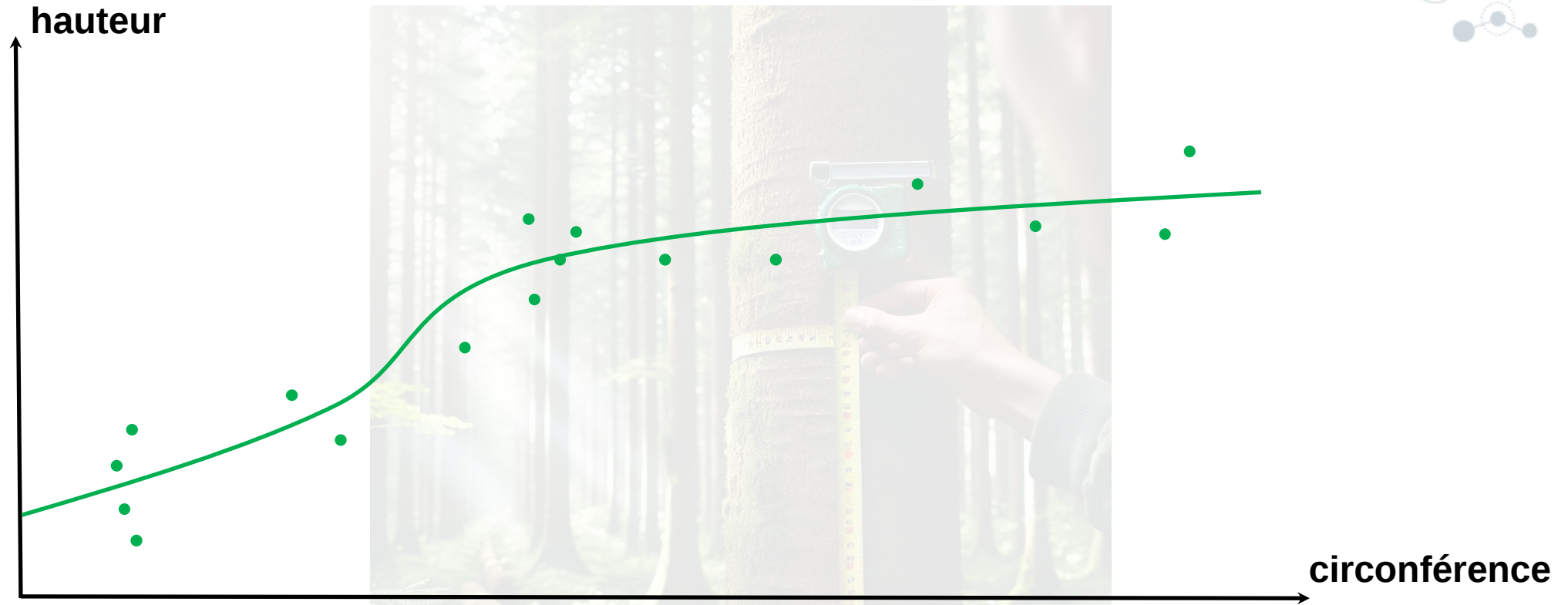


Image générée par IA

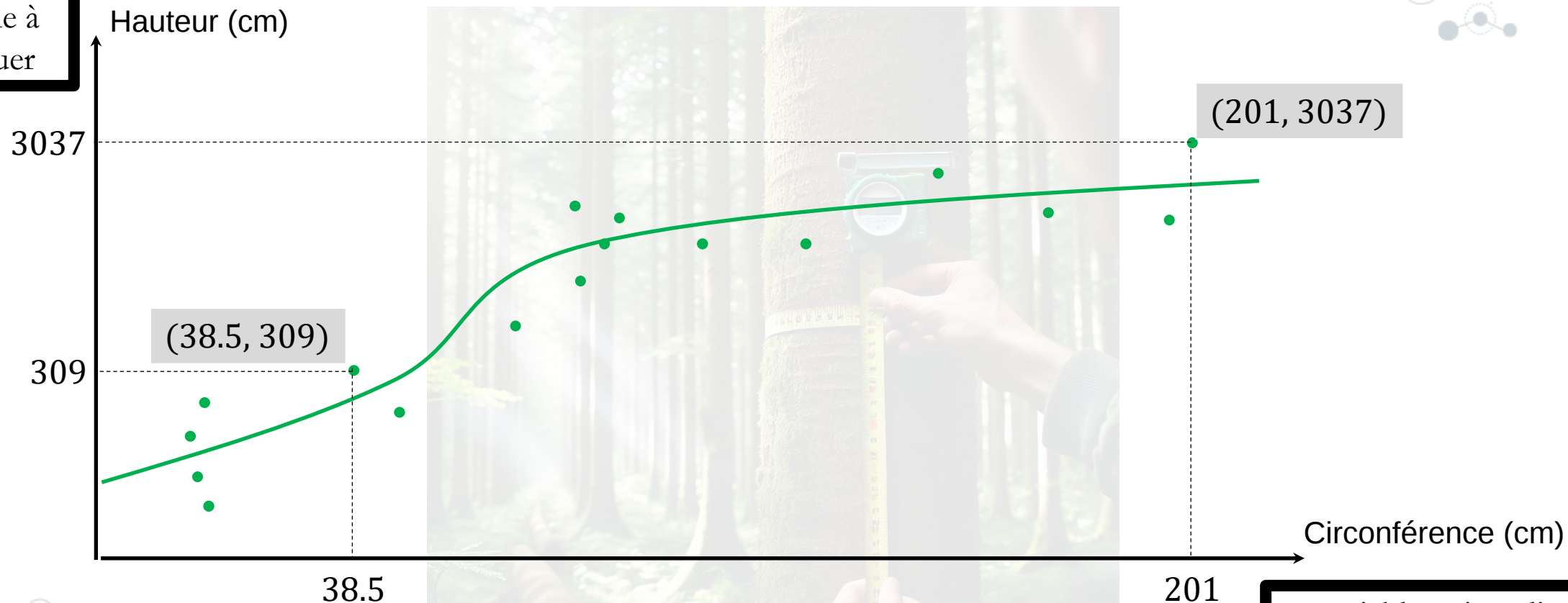
Apprendre à partir de vecteurs

($\approx$ des points dans un espace)

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?

Variable à
expliquer

Hauteur (cm)



Circonférence (cm)

Variable qui explique

Image générée par IA

Apprendre à partir de vecteurs

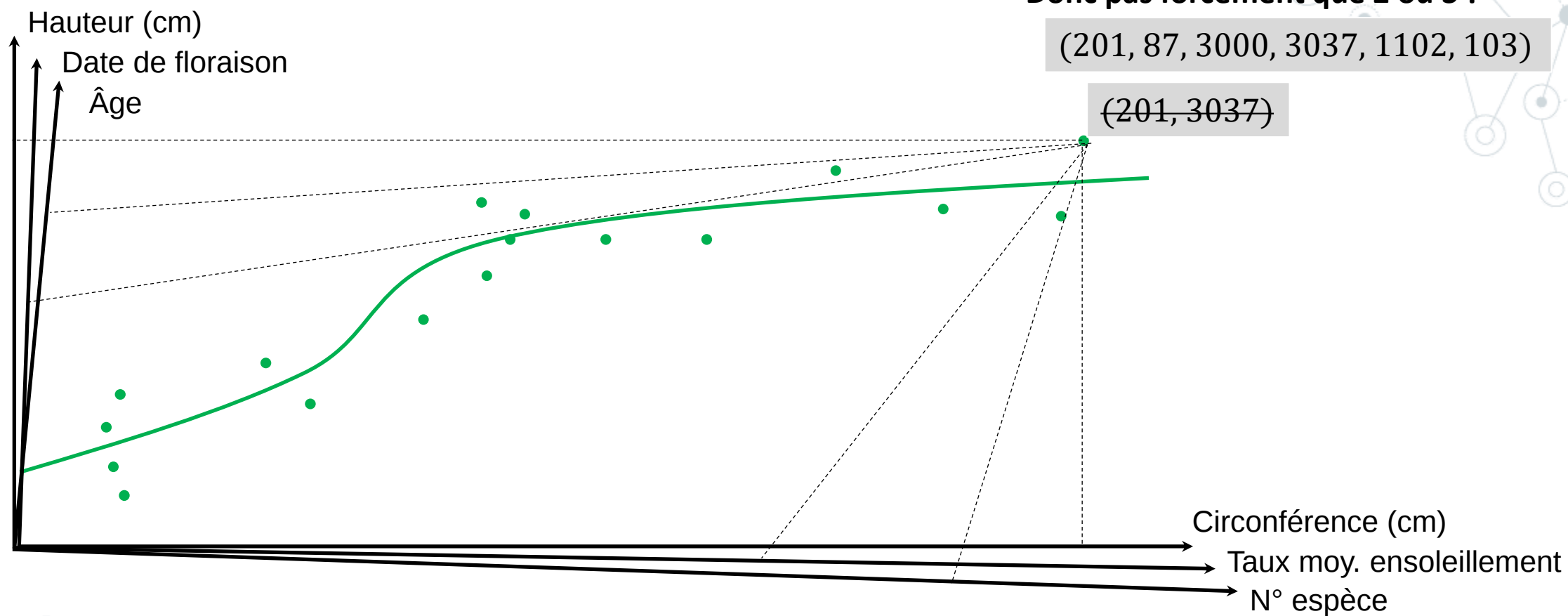
(≈ des points dans un espace)

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?

Vecteur = une liste (ordonnée) de nombres
Donc pas forcément que 2 ou 3 !

(201, 87, 3000, 3037, 1102, 103)

(201, 3037)



Variables à
expliquer
= vecteur en
sortie

Variables qui expliquent
= vecteur en entrée

Apprentissage : Pourquoi ? Comment ?

Quelle est la hauteur d'un arbre en fonction de sa circonférence ?

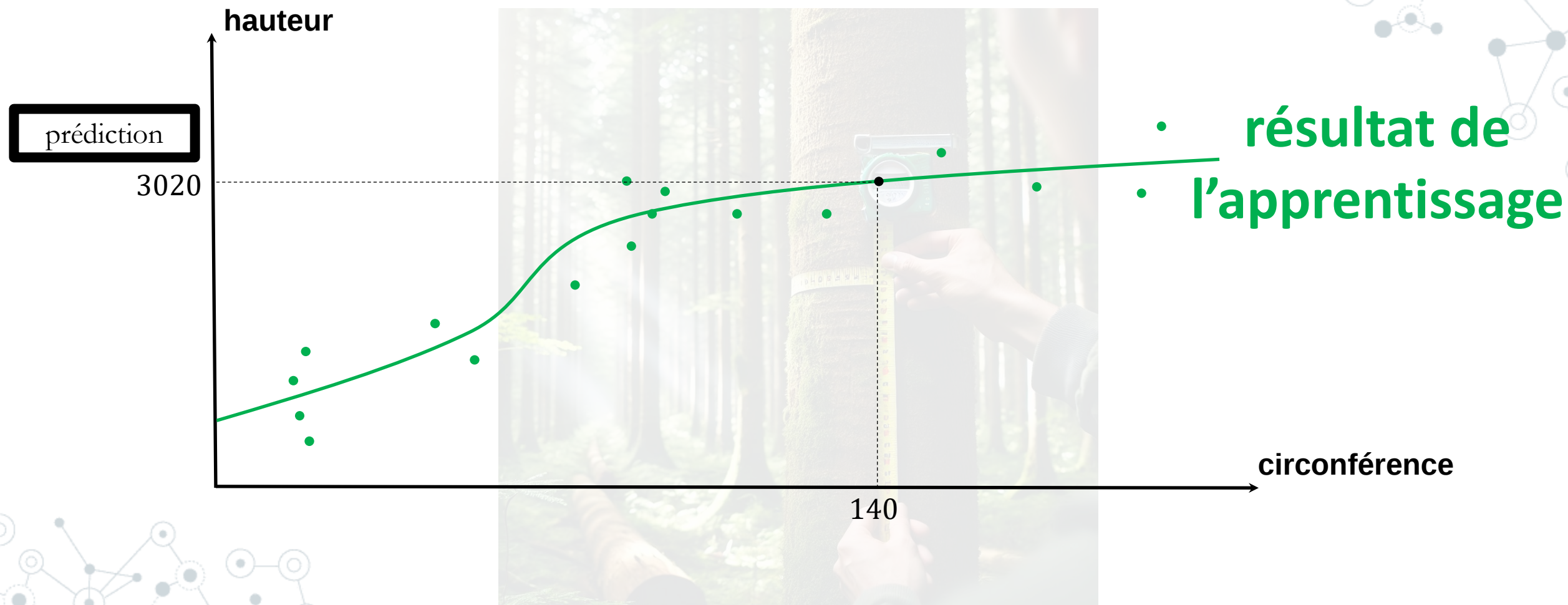


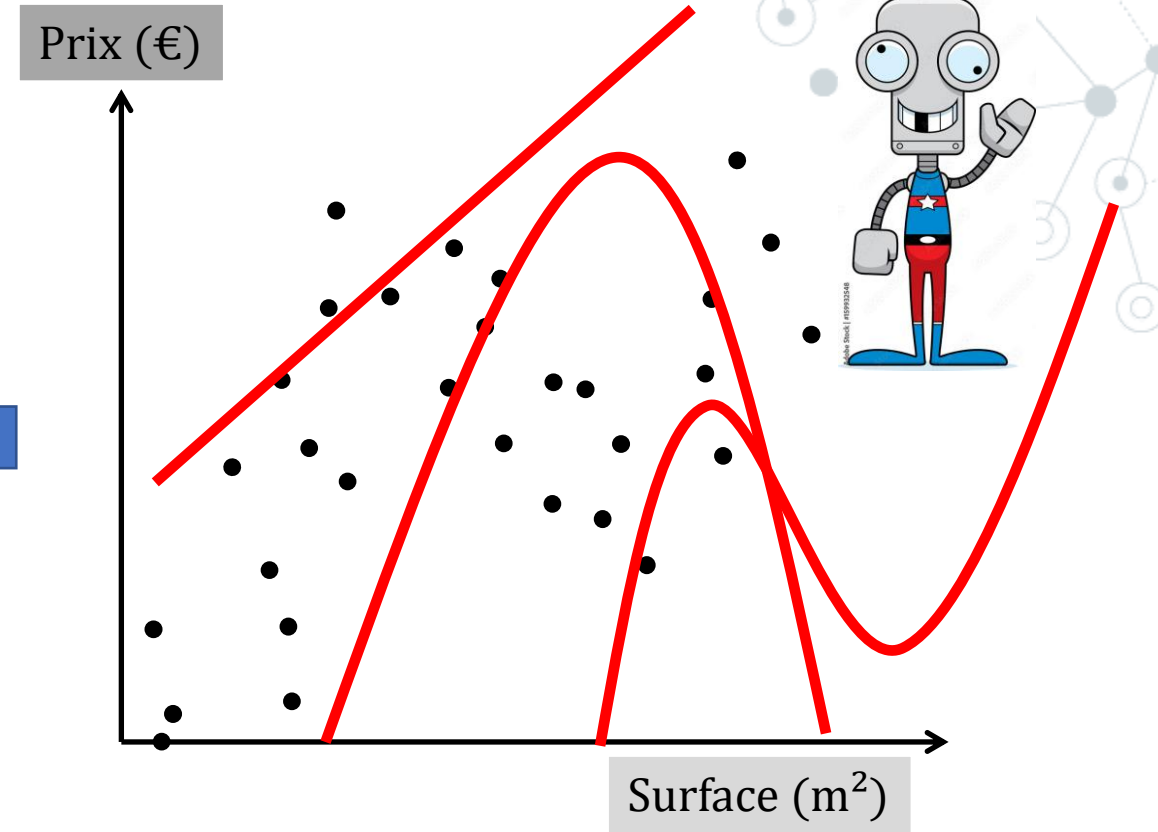
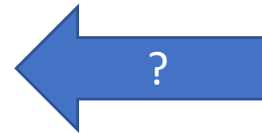
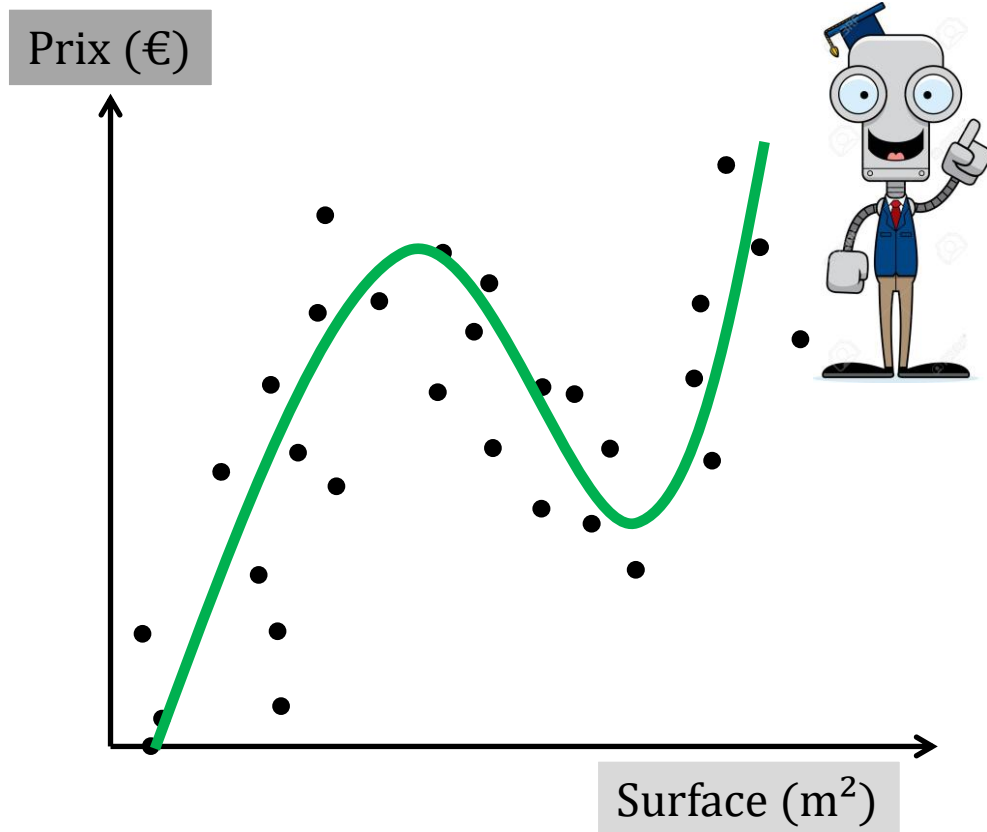
Image générée par IA

Les bases de l'apprentissage neuronal profond : Comment fonctionnent les réseaux de neurones artificiels ?

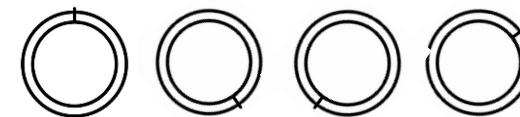
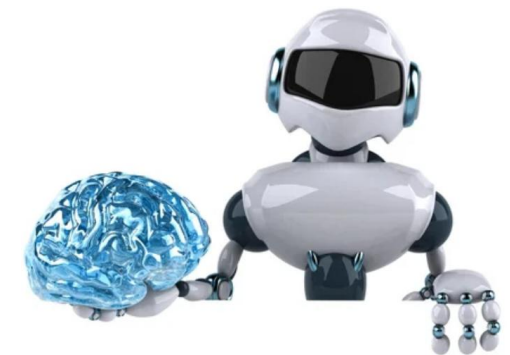
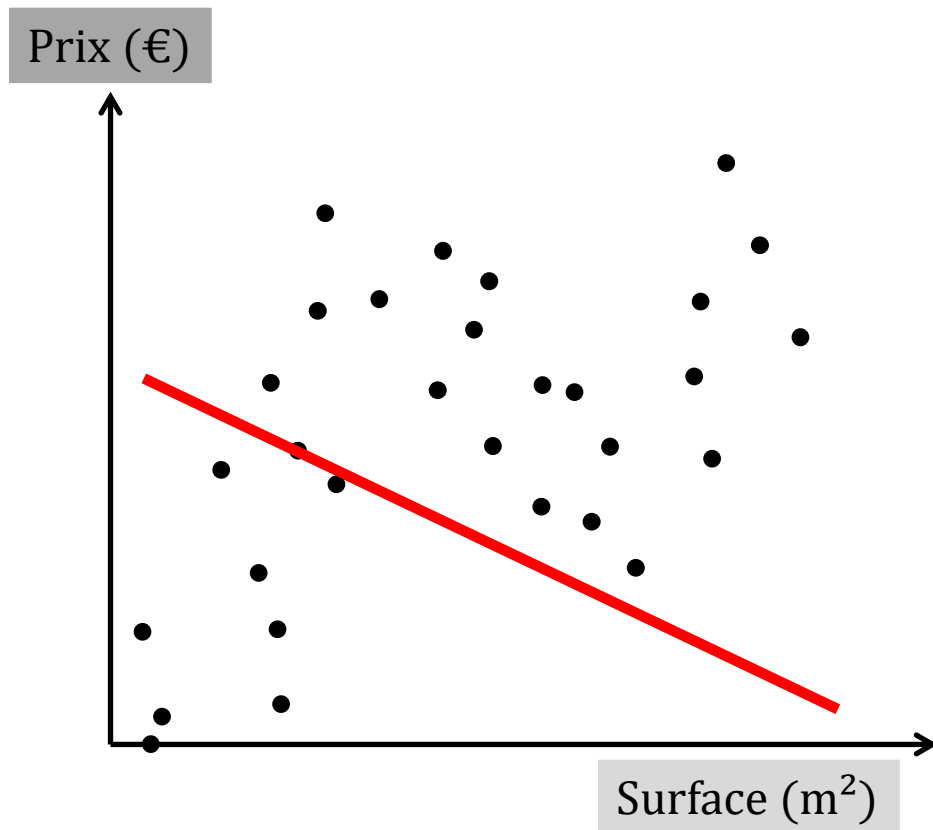
**Geoff Hinton after writing the
paper on backprop in 1986**



Objectif et initialisation

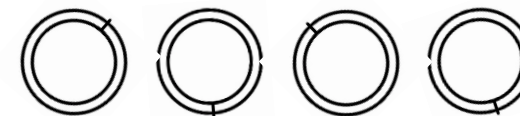
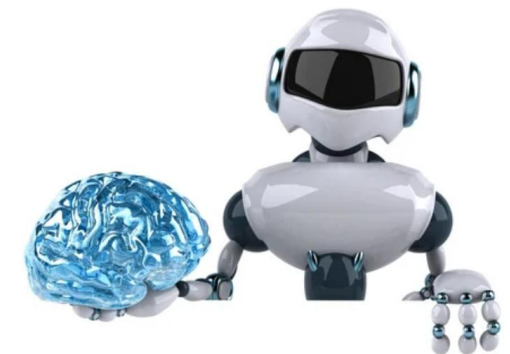
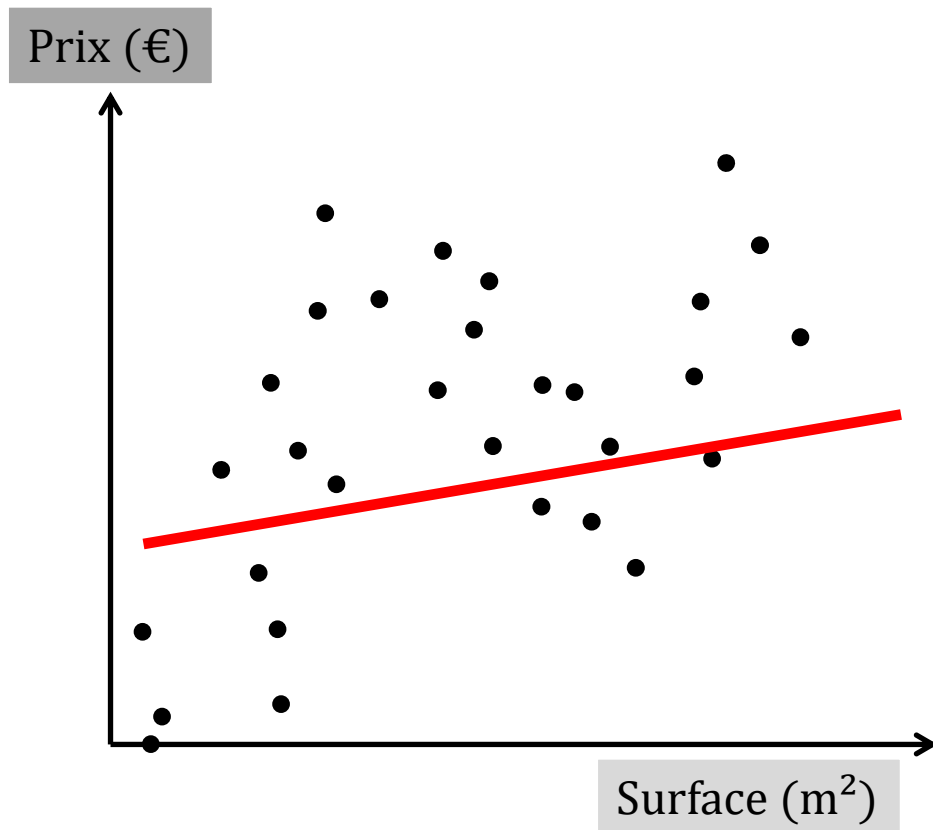


Entraînement : optimisation



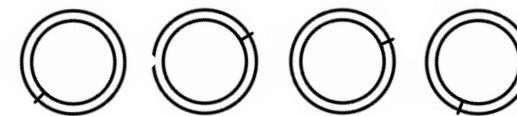
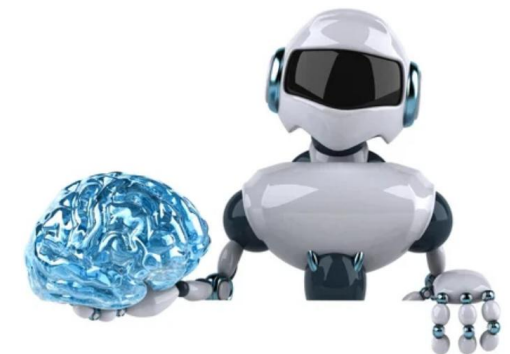
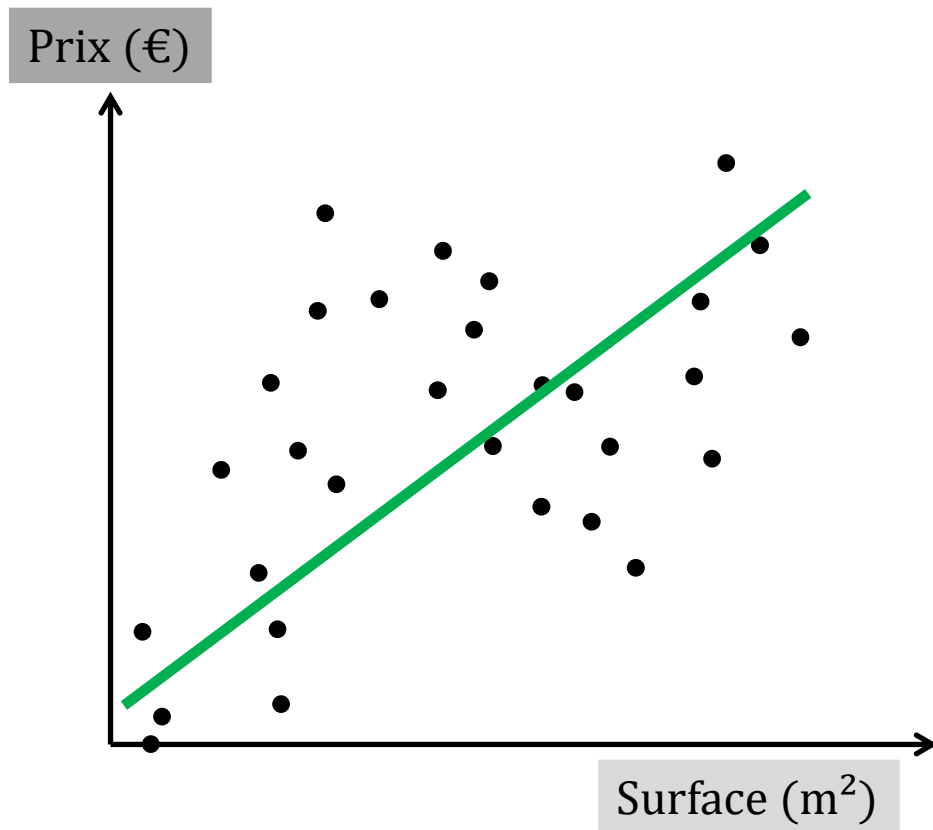
Paramètres de l'IA

Entraînement : optimisation



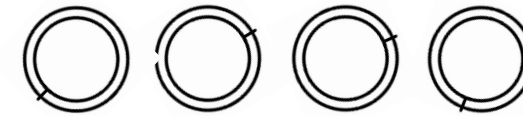
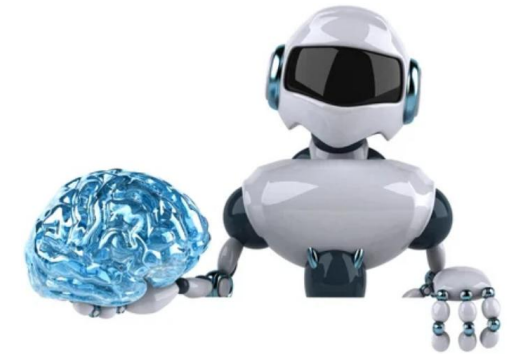
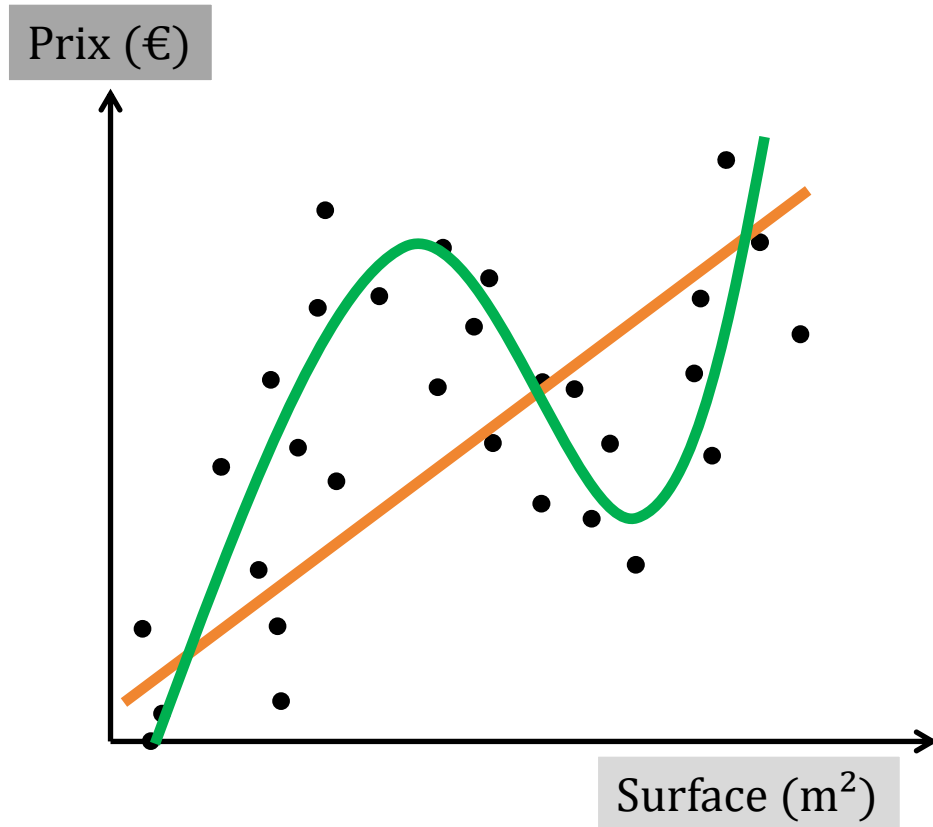
Paramètres de l'IA

Entraînement : optimisation



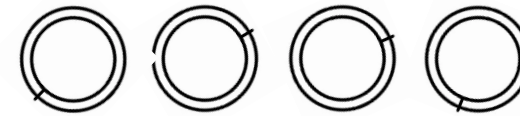
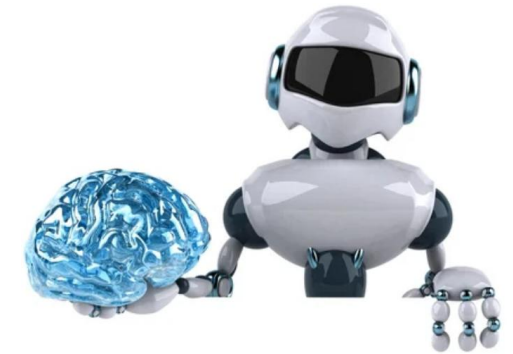
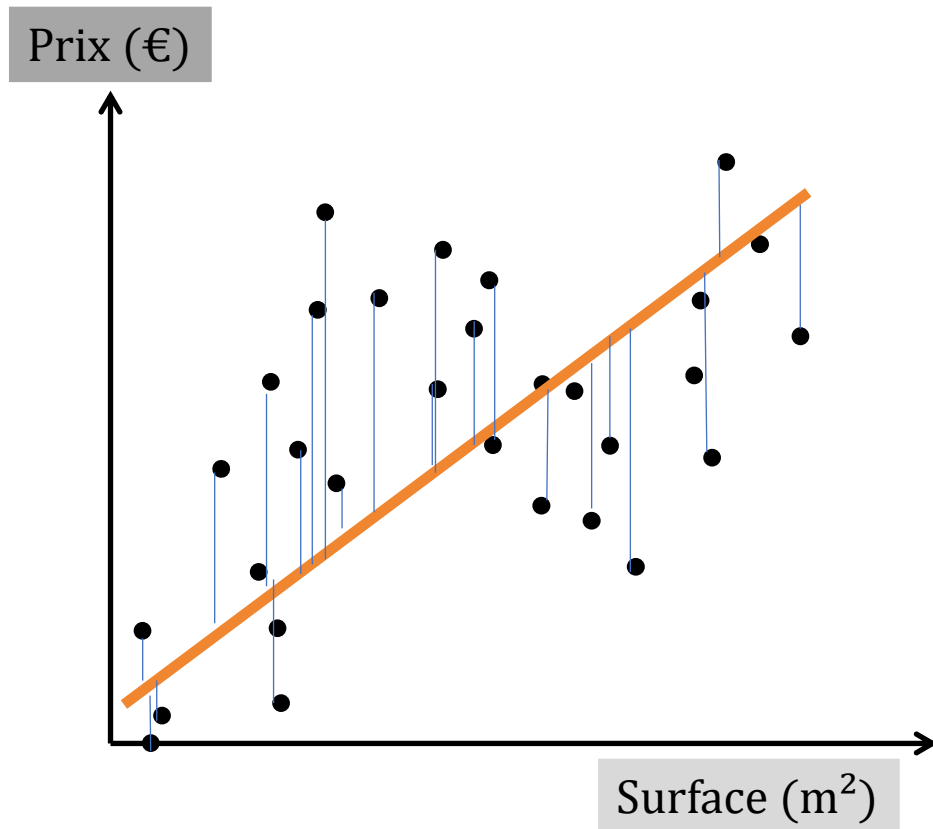
Paramètres de l'IA

Entraînement : choix de l'architecture (hyperparamètre)



Paramètres de l'IA

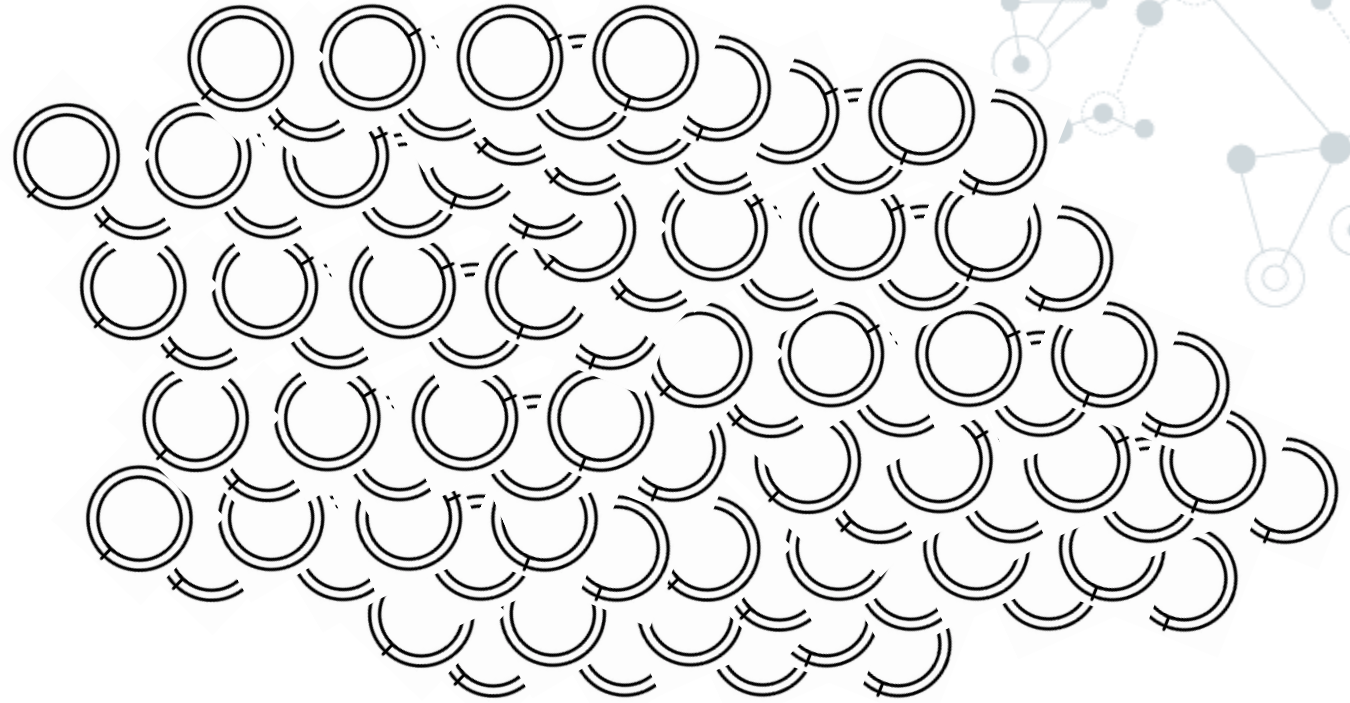
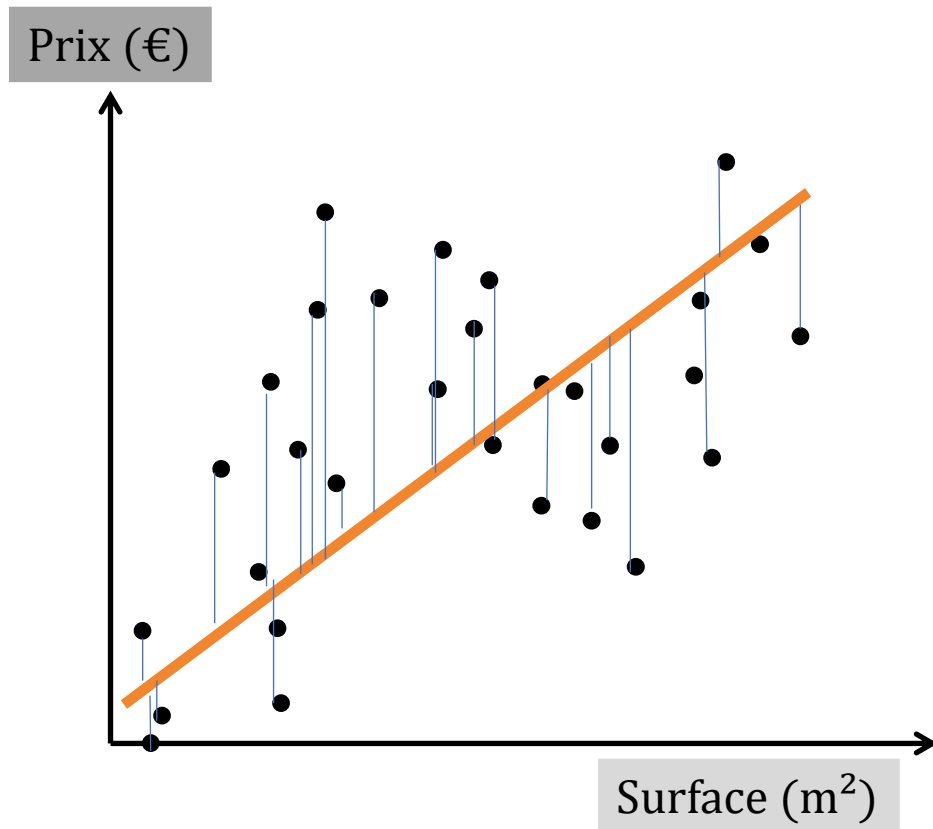
Entraînement : fonction objectif



Paramètres de l'IA

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Entraînement : fonction objectif

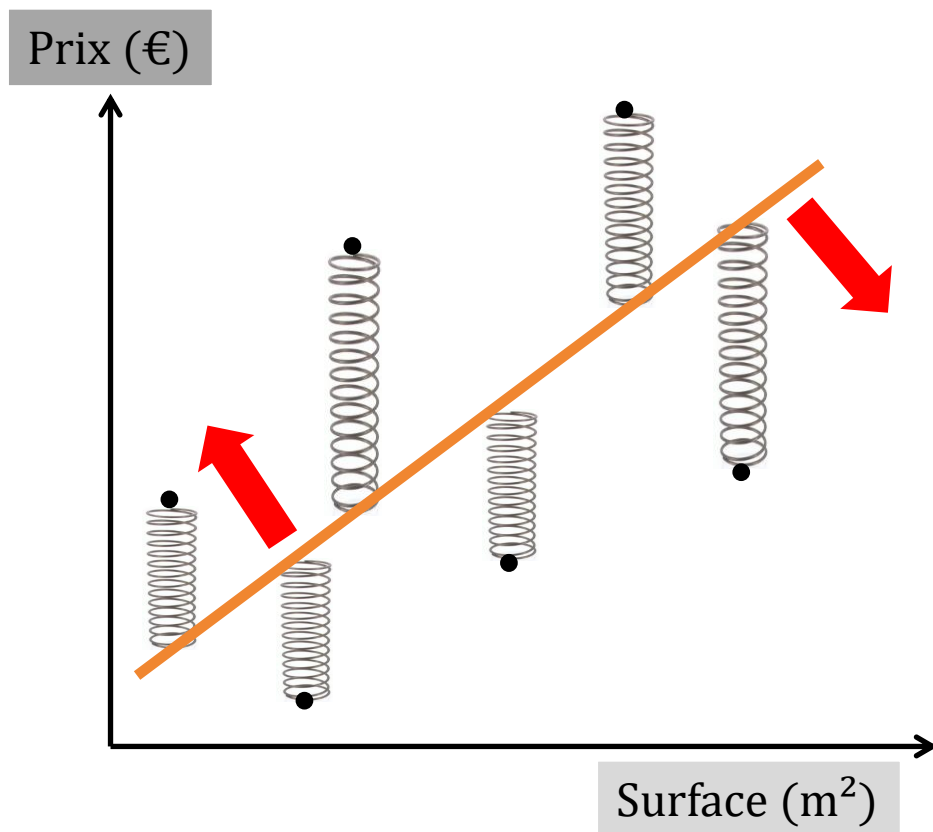


Paramètres de l'IA

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Entraînement : descente de gradient

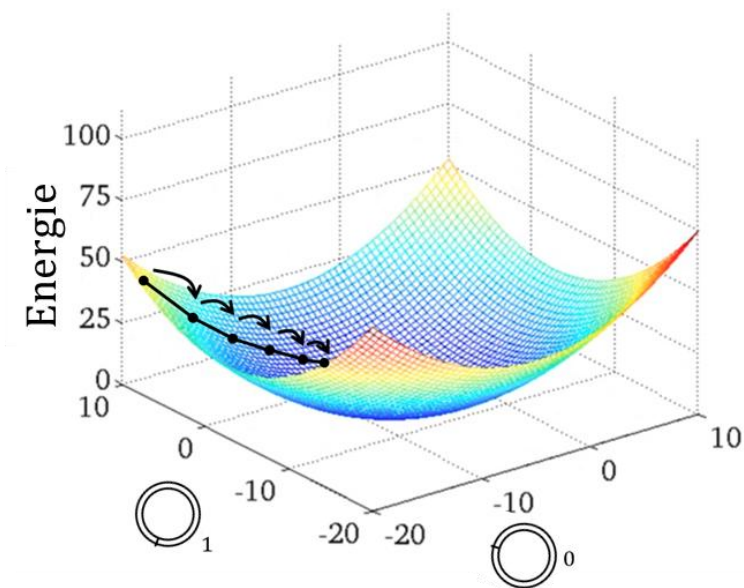
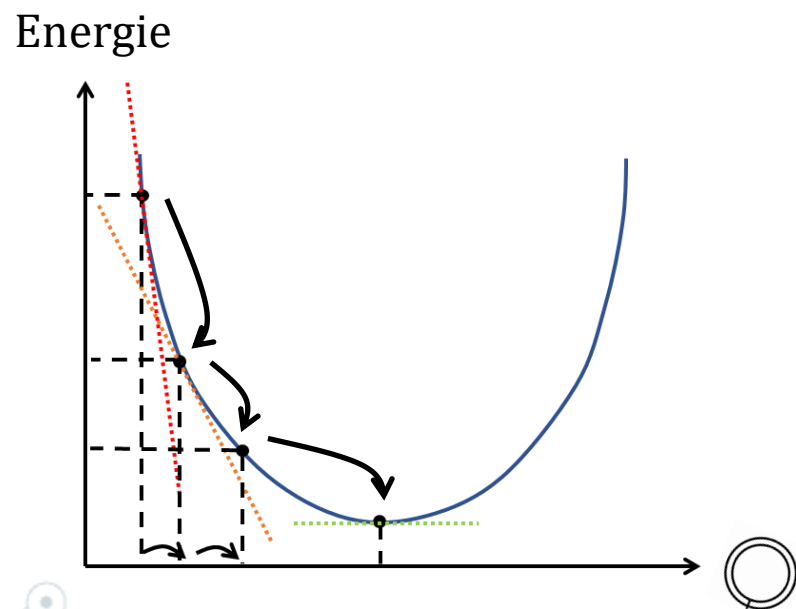
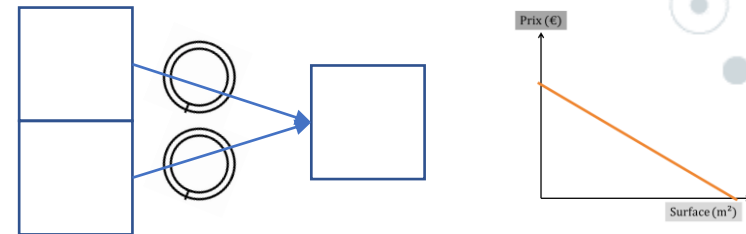
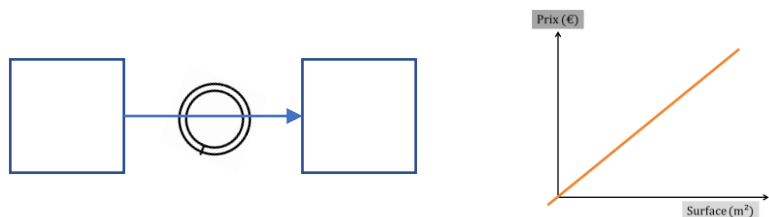
= minimiser « l'énergie »



$$E_{\text{totale}} = \frac{1}{2}k \sum_{i=1}^n (h - y_i)^2 \quad \approx \quad J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Entraînement : descente de gradient

= minimiser « l'énergie »

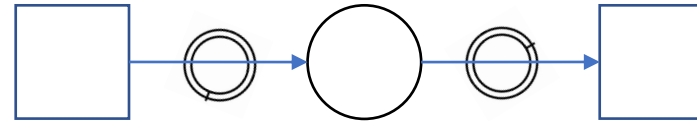


$$\frac{\partial J}{\partial \odot_i}$$



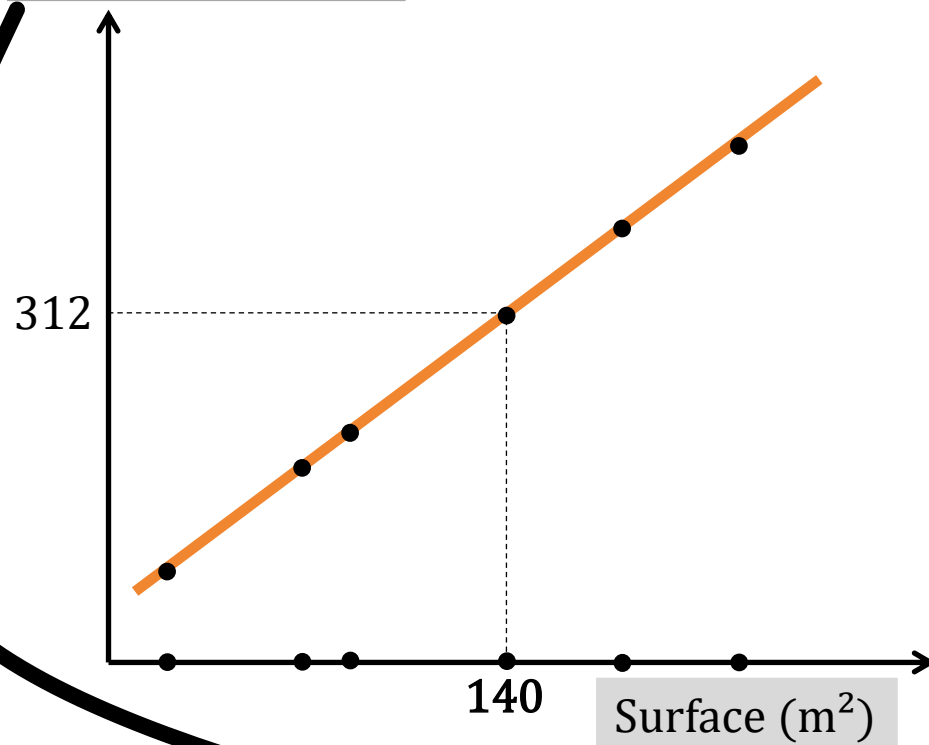
$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Perceptron multicouche

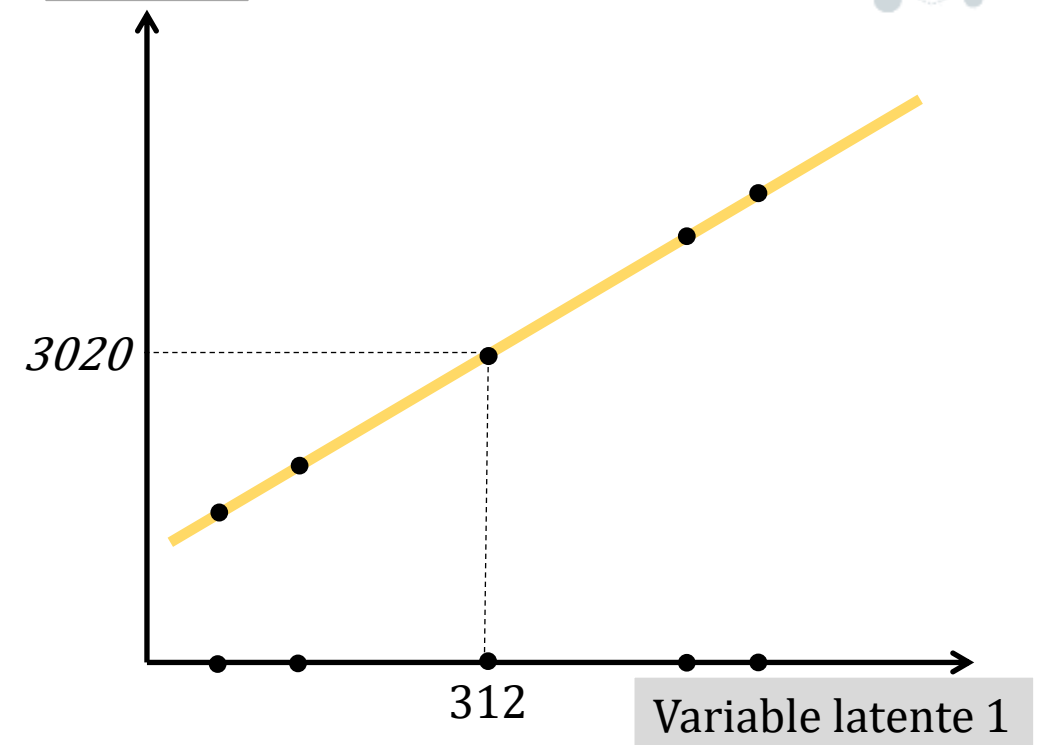


Variables à
expliquer

Variable latente 1

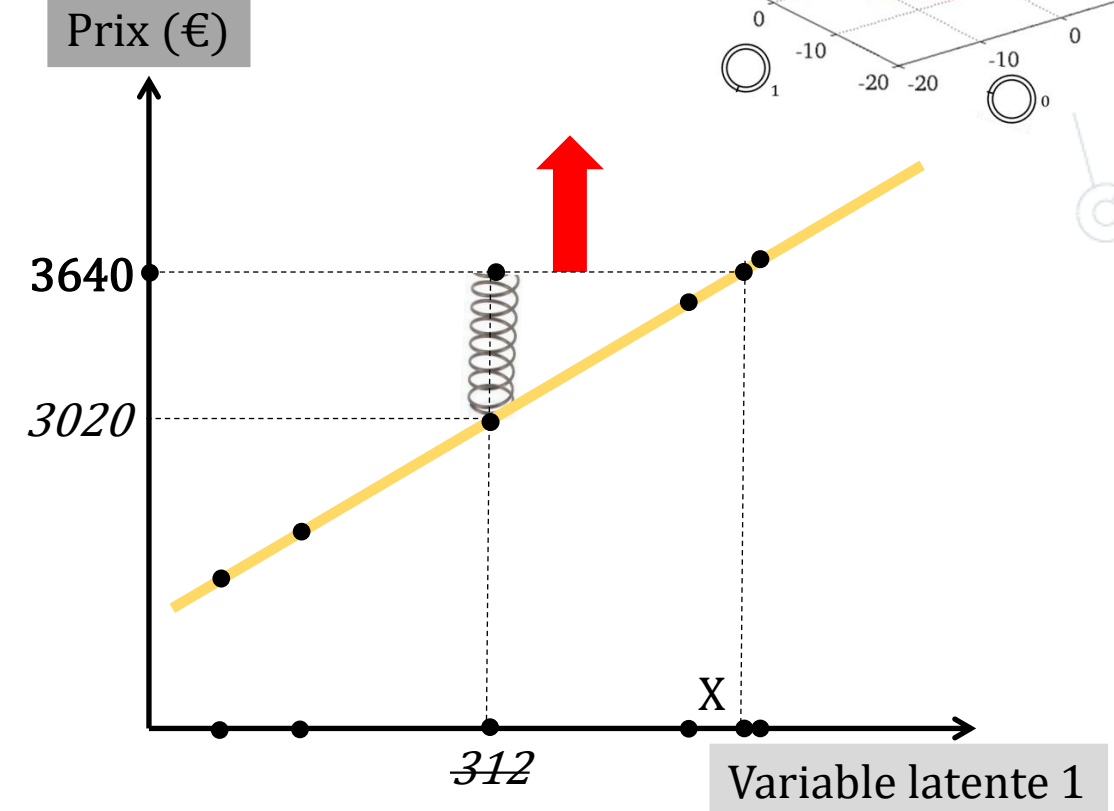
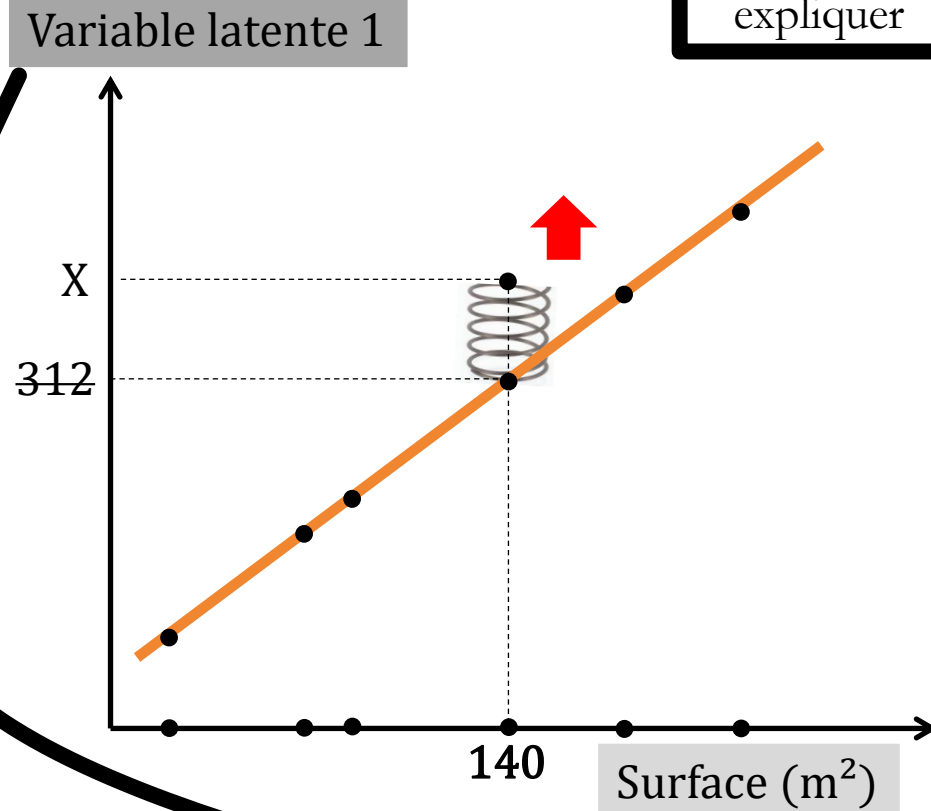
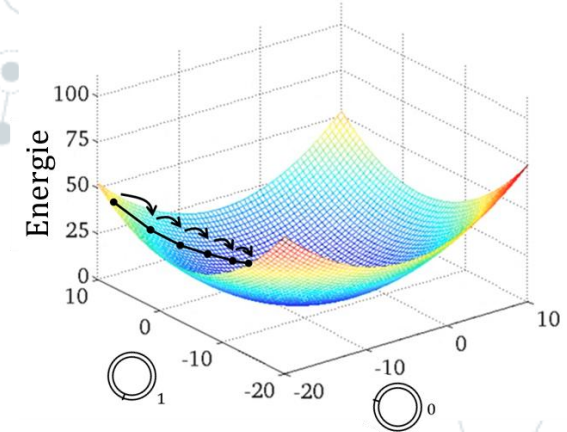
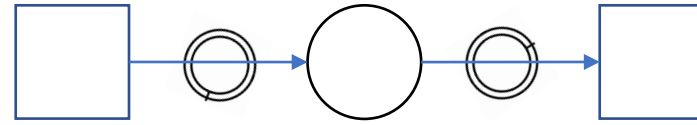


Prix (€)



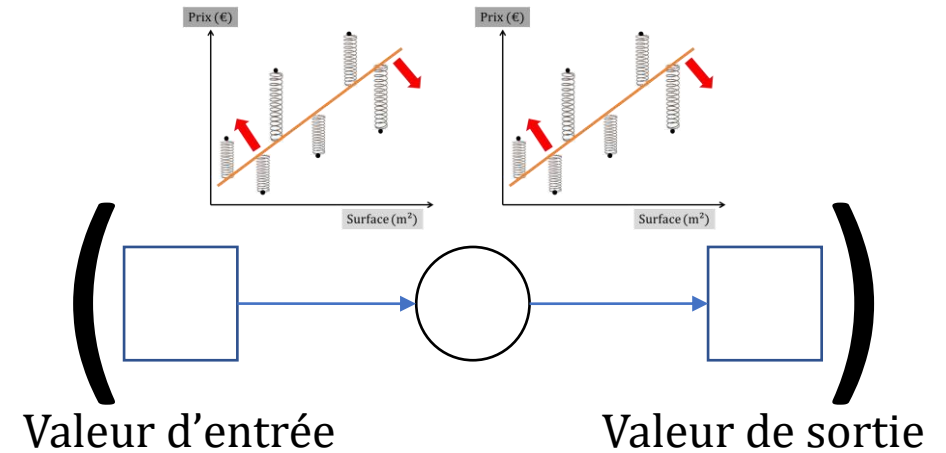
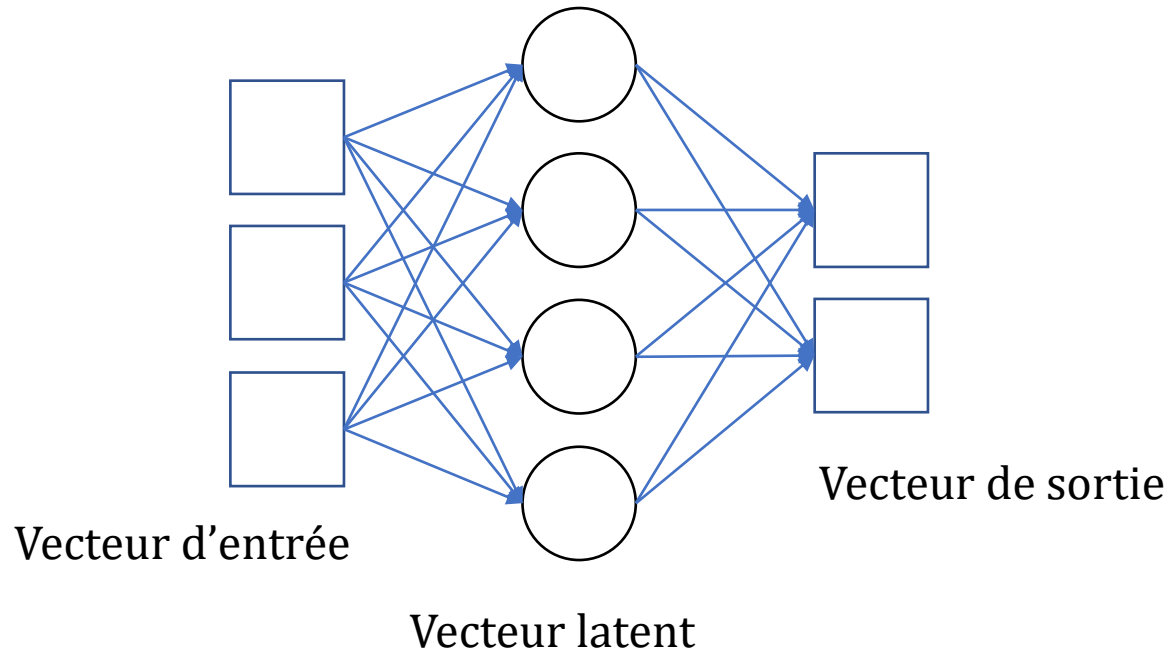
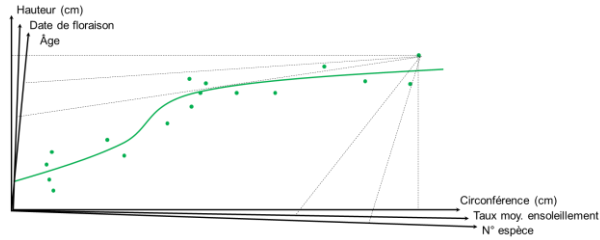
Variables qui expliquent

Perceptron multicouche

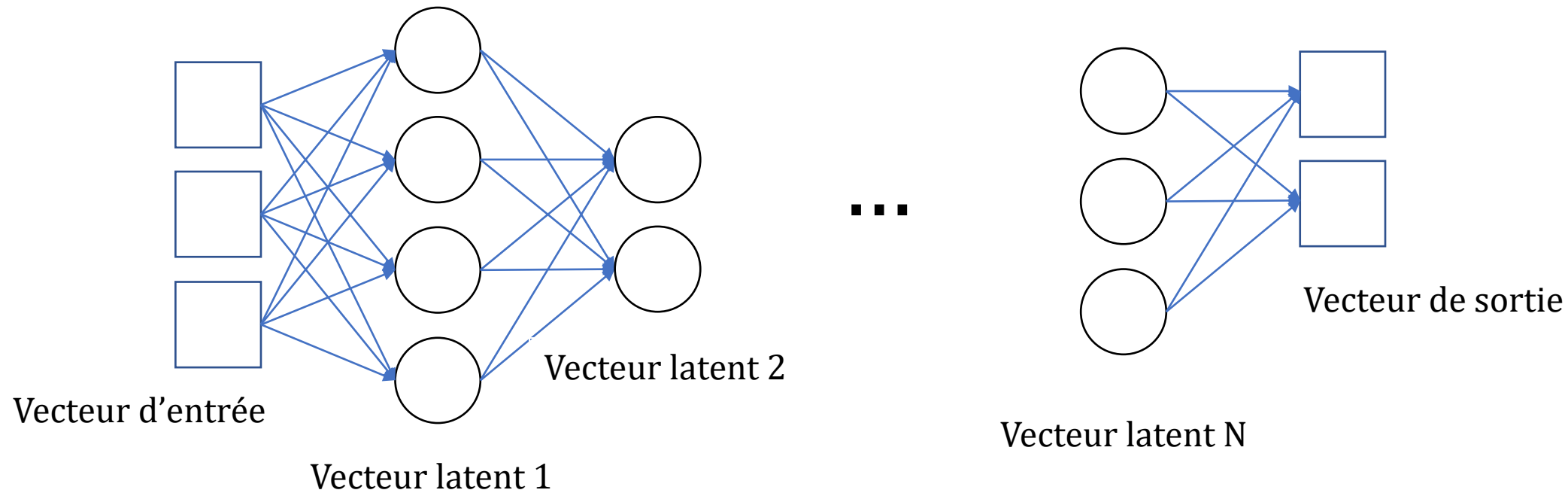


Variables qui expliquent

Perceptron multicouche (à 1 couche cachée)

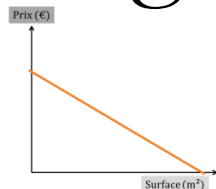


L'apprentissage neuronal profond (aka *deep learning*)



En pratique, plus N augmente, plus on apprend des choses compliquées, mais...

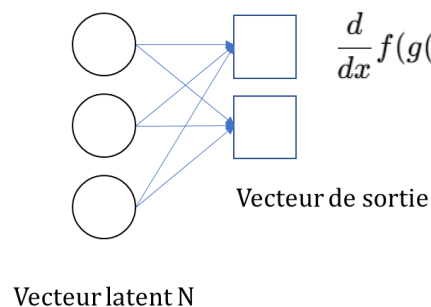
(aka *deep learning*)



$$\frac{\partial J}{\partial \bigcirc_i}$$

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

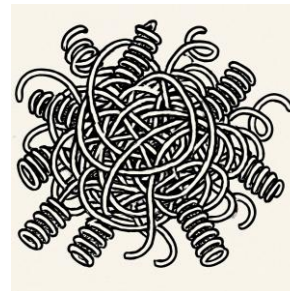
$$\Rightarrow \frac{\partial J}{\partial w_1}(w_1, w_0) = \frac{2 \cdot w_1}{N} \sum_{i=1}^N \left[\frac{e^{-w_1 x_i - w_0}}{(1 + e^{-w_1 x_i - w_0})^2} \cdot \left(y_i - \frac{1}{1 + e^{-w_1 x_i - w_0}} \right) \right]$$



$$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x)$$



[illegible]



La rétropropagation du gradient

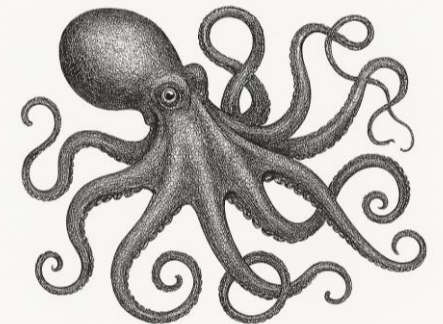


- $e_{N+1}^{(j)} = g'_{N+1} \left(\frac{1}{N_e} \sum_{e=1}^{N_e} (y_p - g_{N+1}(z_{N+1}^{(j)})) \right)$
- $e_k^{(i)} = g'_{k+1} \left(\frac{1}{N_e} \sum_{e=1}^{N_e} (y_p - g_{k+1}(z_{k+1}^{(j)})) \right)$
- $\frac{\partial L}{\partial w_{i,j}^{(k)}} = e_{k+1}^{(j)} \cdot e_k^{(i)}$

« Pour savoir dans quel sens il faut tourner un bouton dans une couche, on a simplement besoin de connaître l'erreur dans la couche suivante. »

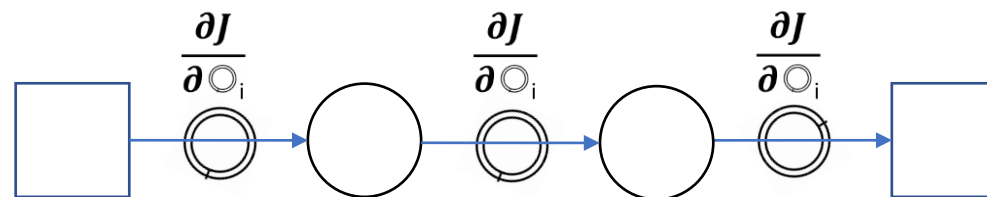
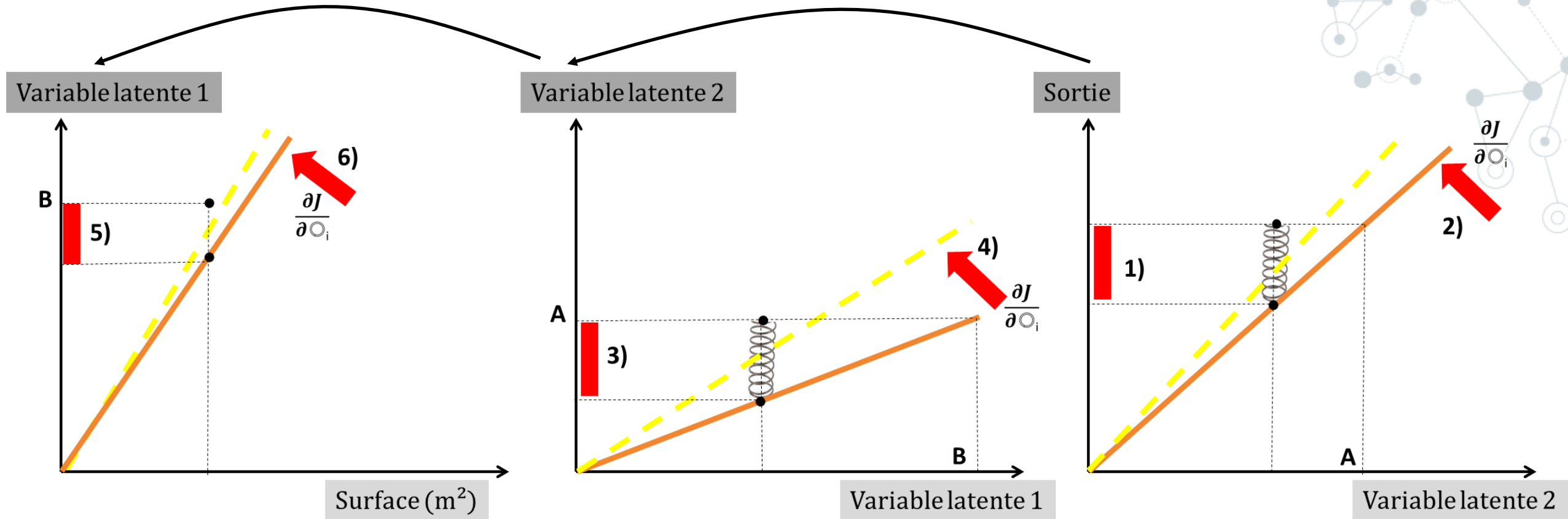
Deep Learning

How Machines Learn



Arnaud Ferré

La rétropropagation du gradient



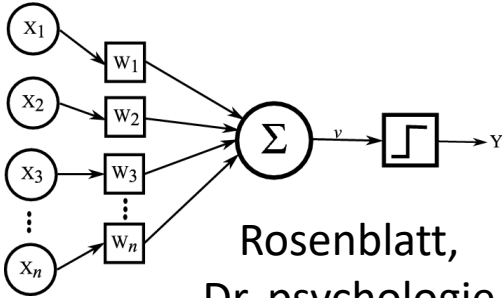
Résumé

Comment les machines apprennent ?

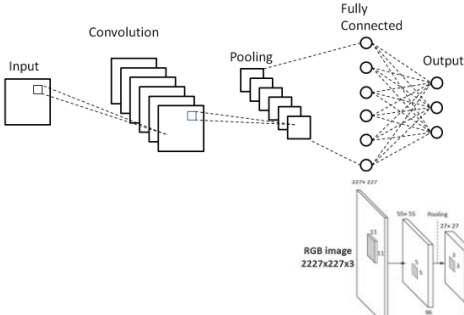
- 1) Grâce à des données/observations (= **exemples d'entraînement**).
- 2) Données traduites en **vecteurs**.
- 3) Vecteurs utilisés pour **entraîner** (votre modèle neuronal)
= trouver **la courbe qui approxime le mieux** vos données vectorisées.
 - Ce qui veut dire optimiser « l'énergie » de votre modèle (coût/objectif/perte).
 - Ce qui est réalisable grâce à la rétropropagation (dans le cas du *deep learning*).
- 4) Si tout s'est bien passé, vous pouvez alors faire des **prédictions** pertinentes pour de nouvelles données.

Conclusion

McCulloch et Pitts,
Dr. neurobio

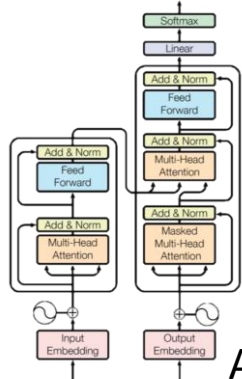


Rosenblatt,
Dr. psychologie



Paul Warbos,
Dr. sciences sociales

2010 : GPU
GTX 580



Ashish Vaswani,
Dr. informatique



Midjourney



ChatGPT

2022

30s : naissance
de l'informatique

1943 : neurone formel
1957 : perceptron



1974-1986 : perceptron
multicouche &
rétropropagation



1997-1998 : RNN/CNN

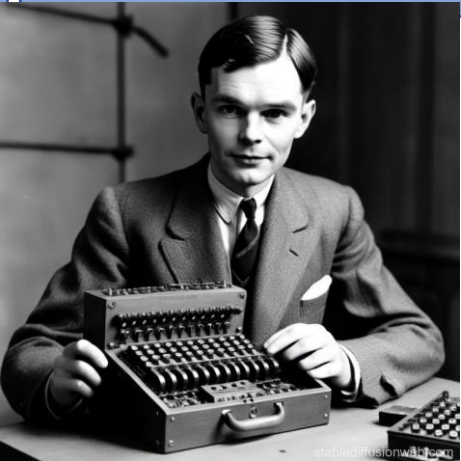
2012 : AlexNet

2014 : Word2Vec

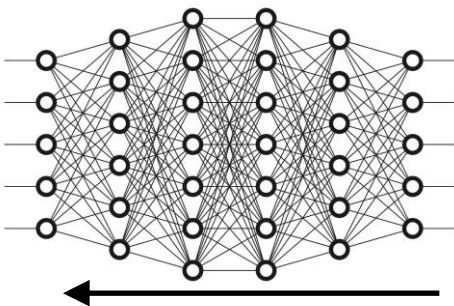
2017 : transformeur

2018 : GPT1/BERT

Ère des IA génératives

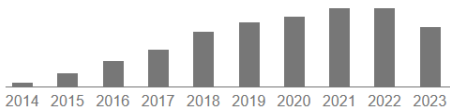


Turing et la
machine
Enigma



Tomáš Mikolov,
Dr. informatique

Cité 38279 fois



Jacob Devlin,
Ing. informatique

Geoffrey Hinton,
Dr. informatique

prix Nobel de
Physique 2024

Yoshua Bengio,
Dr. informatique

Yann Le Cun,
Dr. informatique

prix Turing
2018
(pour le deep learning)

