



WWW.DAUPHINE.PSL.EU

IA For Creative Business

Hands-on Sessions

Alexandre Vérine & Florian Le Bronnec

Présentation des intervenants

Alexandre Vérine

Doctorant MILES à l'Université Paris-Dauphine
Spécialité : Modèles Génératifs d'Image (et de texte)



Florian Le Bronnec

Doctorant MILES à l'Université Paris-Dauphine
Spécialité: Modèles Génératifs de Texte (et d'image)



Présentation du module



Au programme, 4 sessions:

Session 1 : Avant l'IA, la data

Session 2 : Le Machine Learning en pratique

Session 3 : IA pour le texte

Session 4 : IA pour l'image (et le son)



Objectif:

**De la théorie à
la pratique**

Maintenant, à votre tour !

- *30s sur votre parcours*
- *30s sur votre vécu par rapport à l'IA*



WWW.DAUPHINE.PSL.EU

01/02/204

Avant l'IA, la Data



Qui construit l'IA ?

Les métiers

Qui construit l'IA ?



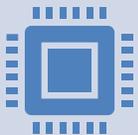
Data Scientist

Extrait des informations à partir de données en utilisant des techniques analytiques et statistiques pour résoudre des problèmes commerciaux ou scientifiques.



Data Engineer

Conçoit et construit l'infrastructure et les systèmes nécessaires à la collecte, au stockage et au traitement efficaces des données.



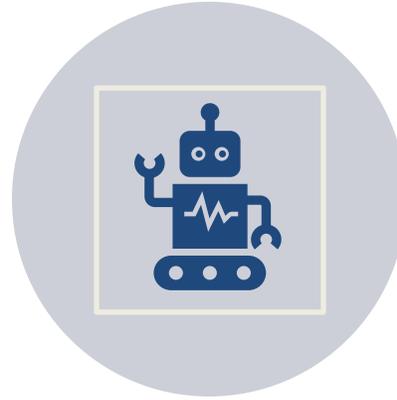
Software Engineer

Conçoit, développe et teste des logiciels et des systèmes informatiques en utilisant des pratiques d'ingénierie logicielle pour répondre aux besoins fonctionnels et aux exigences de qualité.

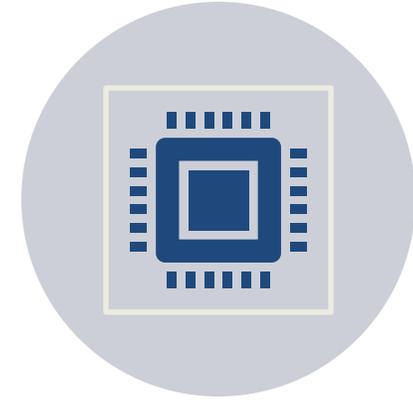
Qui construit l'IA ?



DATA SCIENTIST
30%



DATA ENGINEER
30%



SOFTWARE ENGINEER
30%

Derniers 10% ?

Research Scientist !

**Quelle est la brique
fondamentale d'une IA ?**

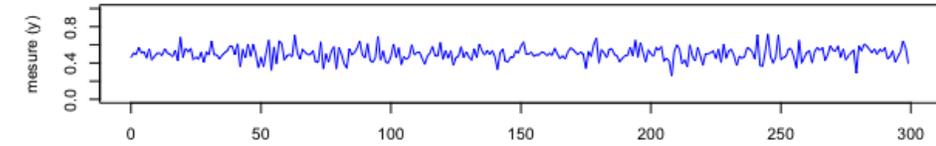
La data



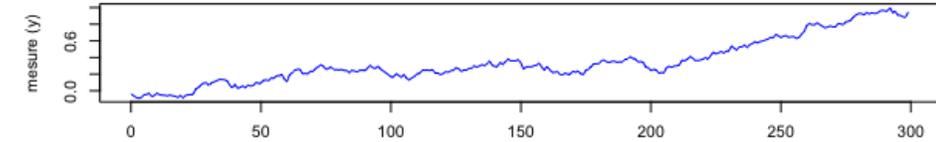
La donnée : le nerf de la guerre



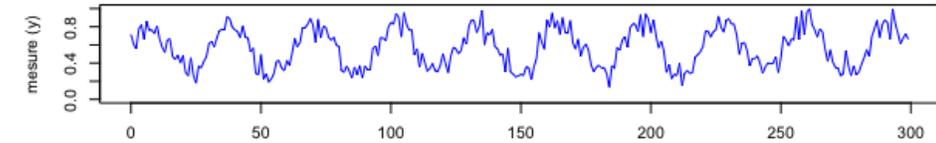
Série 1: stationnaire



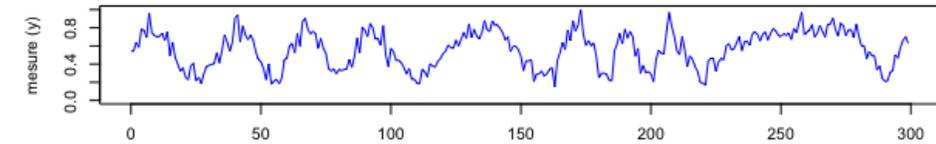
Série 2: évolutive (non-stationnaire)



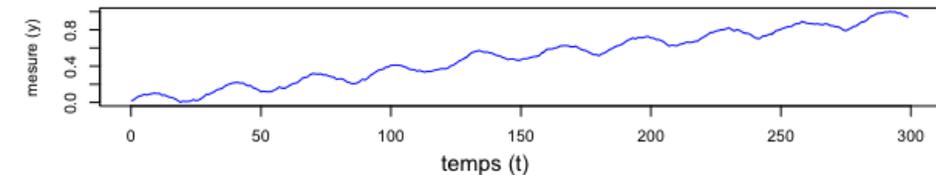
Série 3: stationnaire en tendance et cyclique, cycle régulier



Série 4: stationnaire en tendance et cyclique, cycle irrégulier

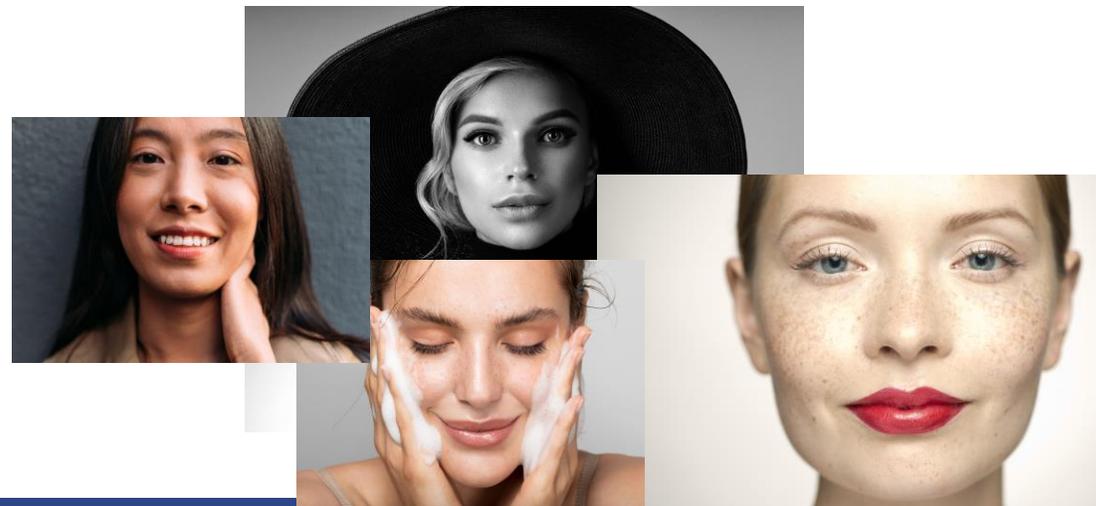
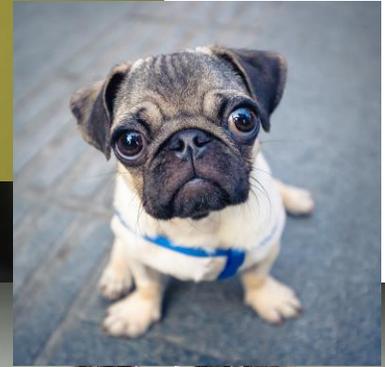


Série 5: évolutive (non-stationnaire) et cyclique

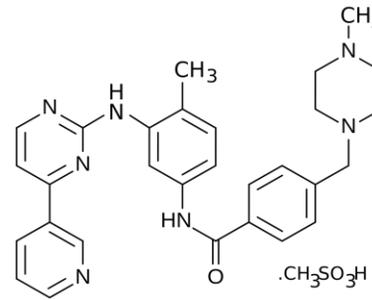
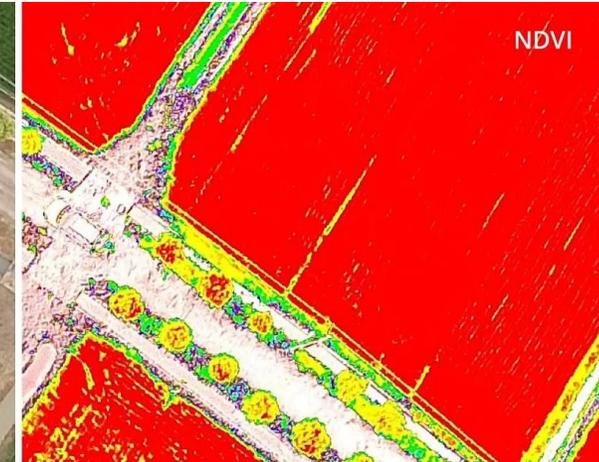
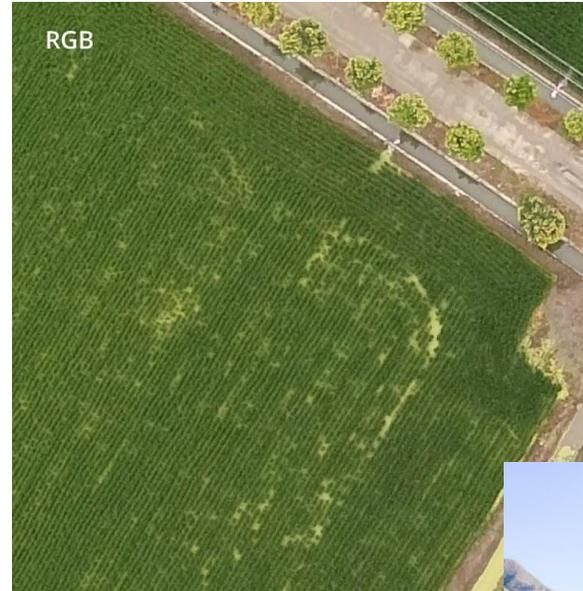
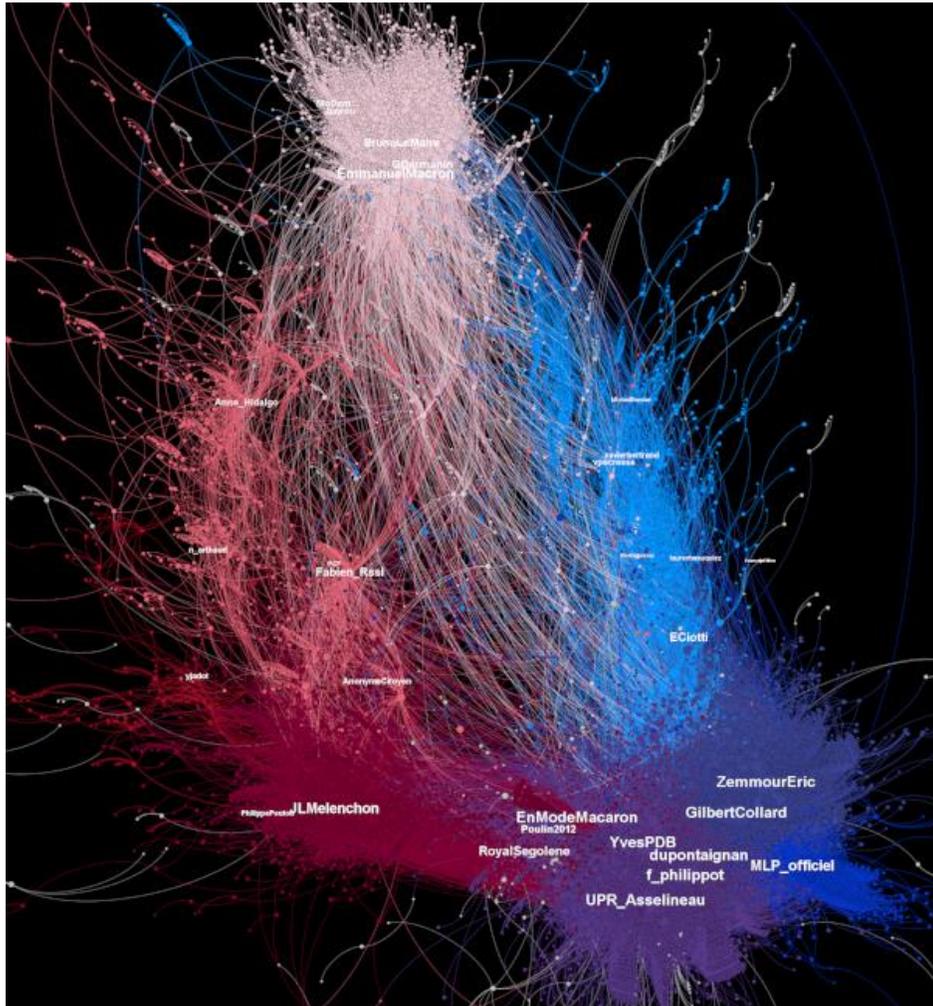


Crédit: <https://ourednik.info/maps/2023/05/21/analysestemporellesavecrl/>

La donnée : le nerf de la guerre



La donnée : le nerf de la guerre



**À quel type de données
êtes-vous confrontés ?**

Modalité de données : résumé

- Grande variété de types de données
- Algorithmes adaptés à chaque modalité
- Quelques types de données à avoir en tête : ***données tabulaires, texte, image, séries temporelles***
- Variations internes dans ces catégories (image médicale vs instagram, code vs post twitter, ...)
- La **multi-modalité** est un vrai sujet et rajoute une difficulté supplémentaire

Etude de cas : attribution de crédit

- **Objectif:** analyser les attributions de prêts effectués par une banque.
- **Les différentes étapes que l'on va suivre:**
 1. Réflexions pré-collecte de données
 2. Post-traitement des données (familiarisation avec les données collectées)
 3. Analyse descriptive

À vous de jouer !
Remplissez le formulaire
à l'adresse:

alexverine.com





**data
privacy**

Inspection des réponses

Collecte de données : constat

Confidentialité

- La collecte de données peut vite contenir des informations sensibles
- Nécessité de définir un protocole pour le traitement et le stockage
- Obligations légale et morale

Collecte

- Format incohérent
- Données globalement peu exploitables telles quelles
- Actions de nettoyage requises

Collecte de données : résumé

- Processus à part entière, nécessitant une vraie réflexion
- Point d'attention : confidentialité (RGPD)
- Structure de la donnée :  Structure  Nettoyage
- Phase de nettoyage nécessaire
- Avoir en tête la finalité : réflexions, nettoyage, structure doivent être guidés par l'analyse finale

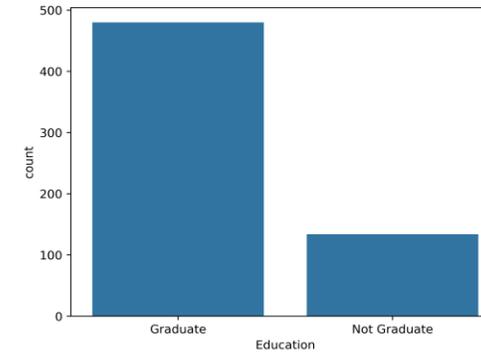
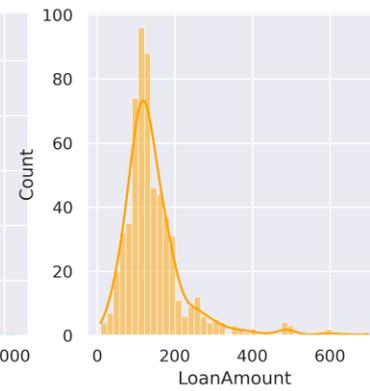
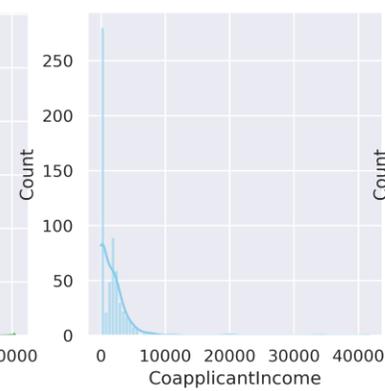
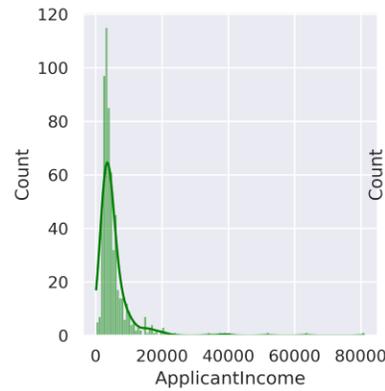
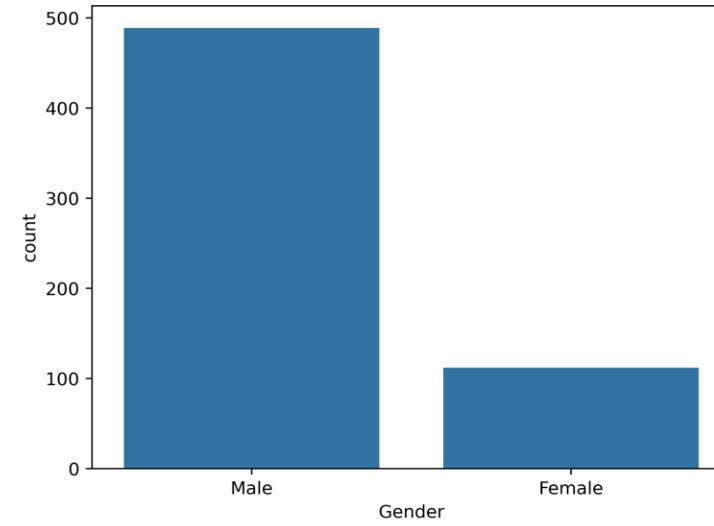
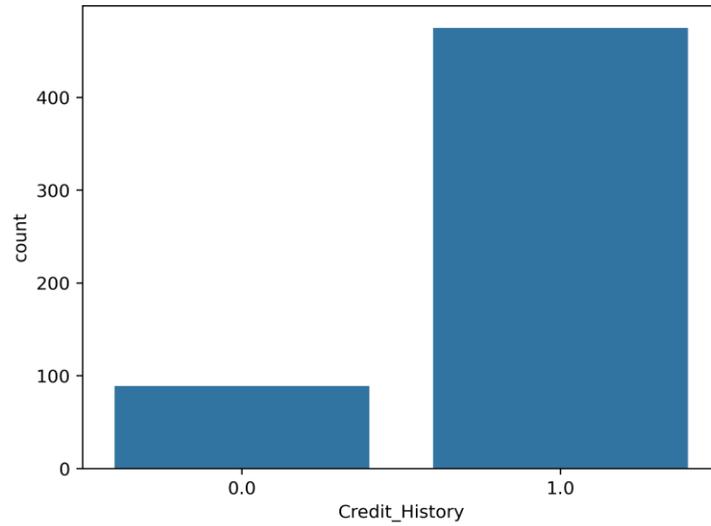
**Prochaine étape:
Que faire avec nos
données collectées ?**

Data visualisation et analyse descriptive

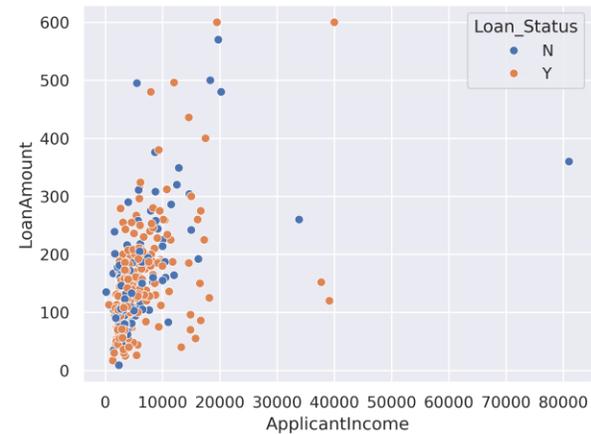
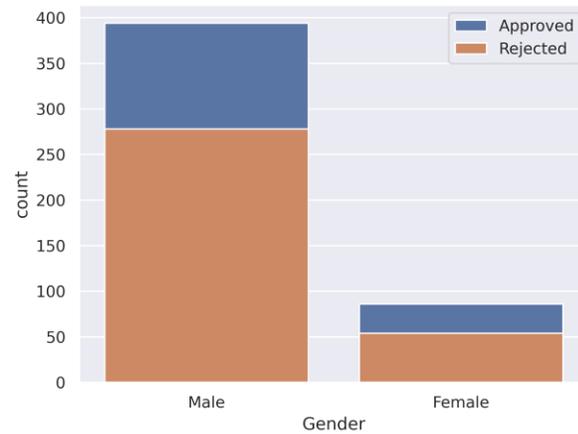
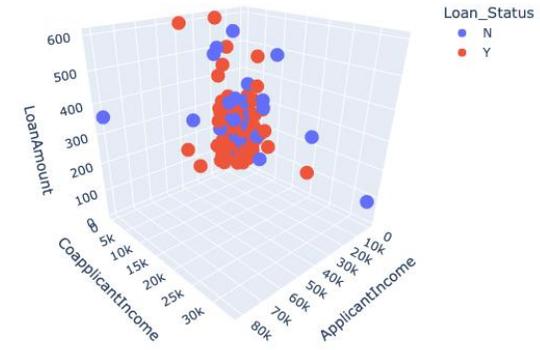
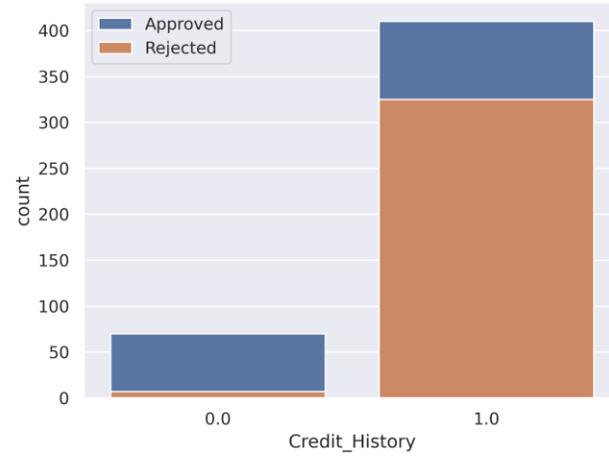
- Mener des analyses simples, permettant de quantifier ce que l'on a recueilli
- Les analyses se font à l'échelle de l'ensemble des données
- Les analyses doivent être quantitatives
- Manière synthétique de faire cela : data visualisation

Tutoriel Python !

Visualiser les features



Visualiser les répartitions des classes



Nettoyer nos données

- Nombre de valeurs manquantes :

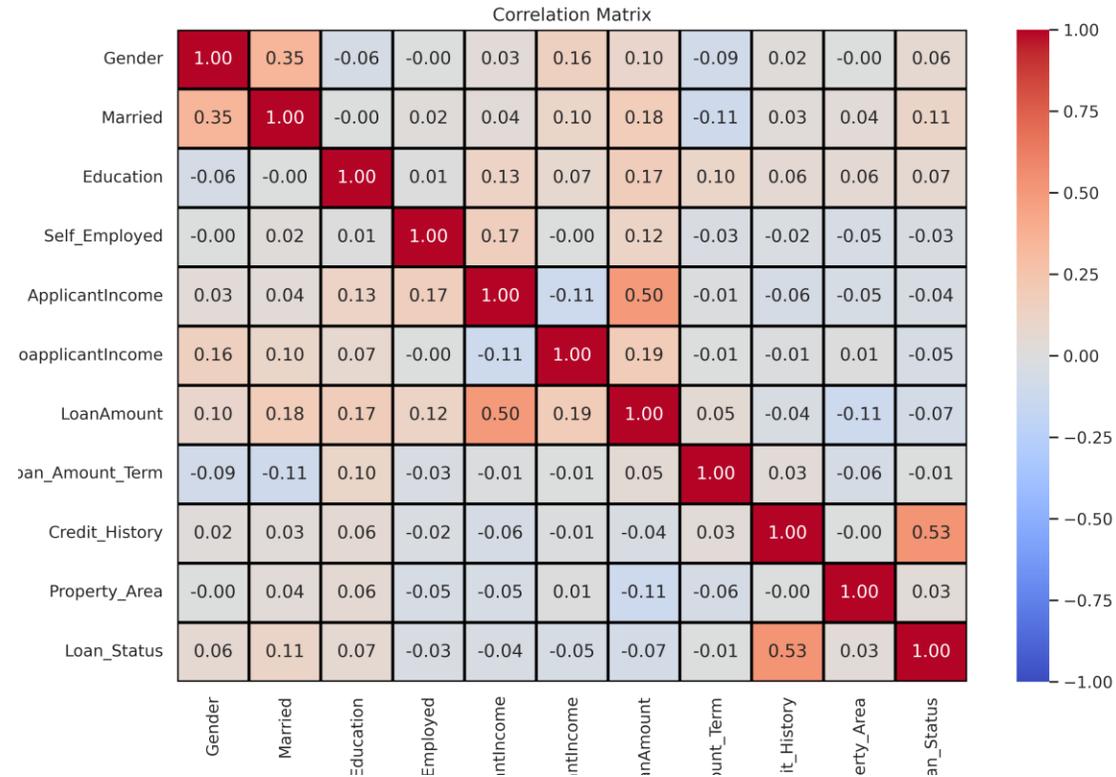
Gender	13
Married	3
Dependents	15
Self_Employed	32
LoanAmount	22
Loan_Amount_Term	14
Credit_History	50

Nombre de données : 614

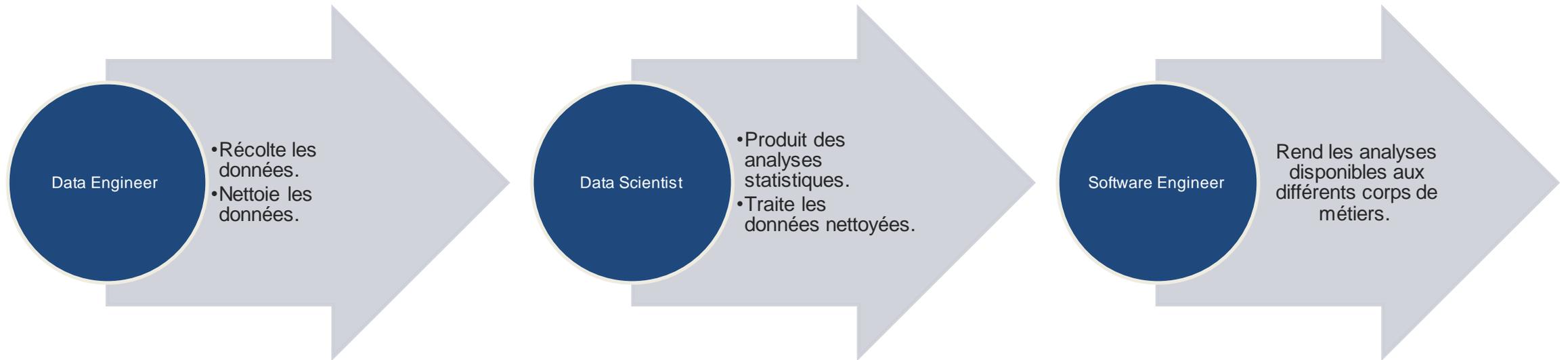


Nombre de données complètes : 480

Analyses statistiques



Quelle est la part de chaque métier dans notre étude de cas ?



Visualisation : résumé

- Permet d'avoir un aperçu synthétique et visuel des données
- Peut donner des idées d'algorithmes à appliquer
- Permet d'éliminer certains types d'algorithmes

Prochaine séance :

15/03/2024





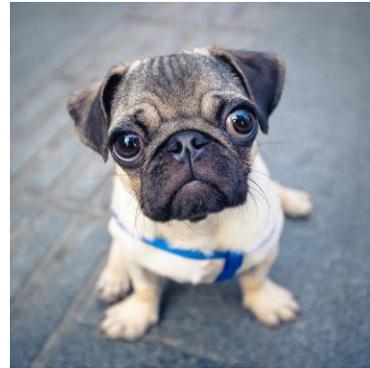
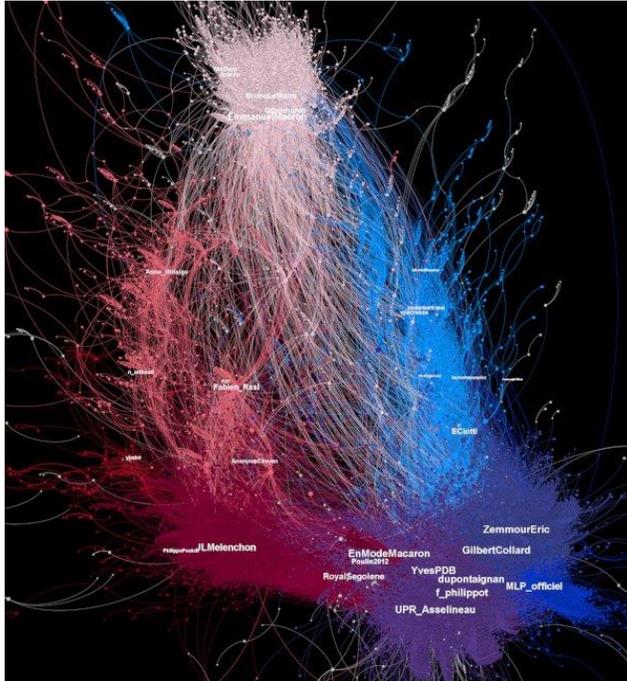
WWW.DAUPHINE.PSL.EU

15/03/2024

Après la Data, l'IA



Rappels de la dernière séance



	A	B	C	D	E	F	G	H	I	
1		Nombre	Último	Máximo	Mínimo	Var.	Var. %	Vol.	Hora	
2		Abertis	12,76	12,76	12,58	0,12	0,95%	379,39K	11:47:15	
3		Acciona	63,965	64,18	62,775	1,135	1,81%	27,19K	12:02:11	
4		Acerinox	12,16	12,275	12,04	0,355	3,01%	596,81K	11:47:39	
5		ACS	28,4	28,61	28,06	0,3	1,07%	99,12K	11:47:25	
6		Aena	125,9	126,6	125,4	0,25	0,20%	22,75K	11:43:53	
7		Amadeus	41,785	41,97	41,545	-0,035	-0,08%	73,23K	12:03:25	
8		ArcelorMittal	7,5	7,6	7,491	-0,026	-0,35%	905,85K	11:48:11	
9		Banco Popular	0,932	0,939	0,911	0,012	1,30%	11,14M	12:04:11	
10		Banco Sabadell	1,25	1,254	1,225	0,022	1,79%	9,14M	12:02:37	
11		Bankia	0,864	0,869	0,859	0,005	0,58%	2,90M	11:48:19	
12		Bankinter	7,256	7,276	7,176	0,058	0,81%	666,66K	11:45:54	
13		BBVA	5,961	5,964	5,835	0,114	1,95%	5,35M	12:04:09	
14		Caixabank	2,923	2,938	2,858	0,063	2,20%	5,62M	11:48:23	
15		Cellnex Telecom	13,32	13,36	13,1	0,22	1,64%	278,38K	11:48:23	
					.27	4,23	0,004	0,09%	995,17K	11:47:27
					.52	73,045	0,58	2,53%	237,01K	12:03:49

amazon.com

Customer Reviews

789 of 845 people found the following review helpful:

★★★★★ **Gripping but amazingly incomplete,**
October 27, 2011

By
David Dennis (West Palm Beach, FL United States) - [See all my reviews](#)

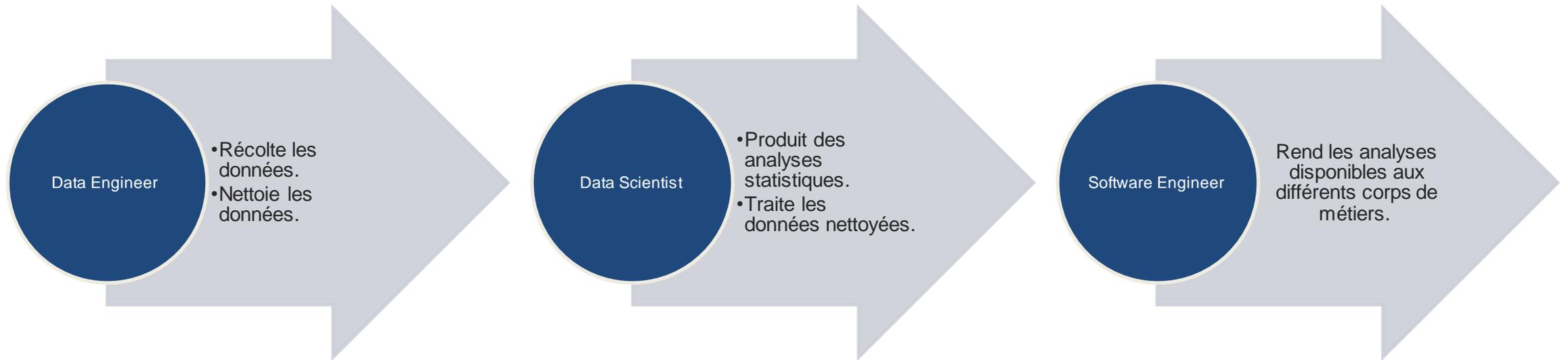
REAL NAME™

This review is from: Steve Jobs (Hardcover)

This is a gripping journey into the life of an amazing individual. Despite its girth of nearly 600 pages, the book zips along at a torrid pace.

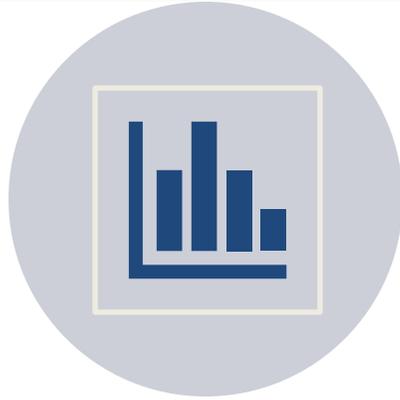
Rappels

Data Engineer, Data Scientist, Software Engineer

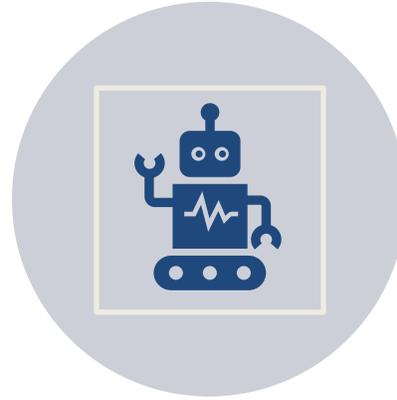


**Des questions, des
remarques ?**

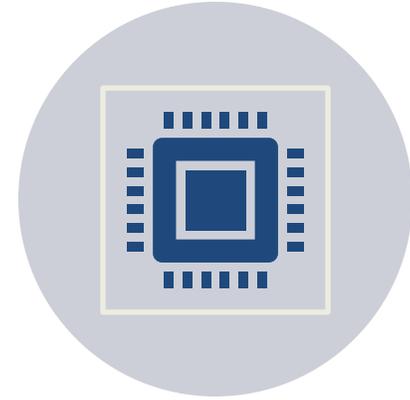
Aujourd'hui



DATA SCIENTIST



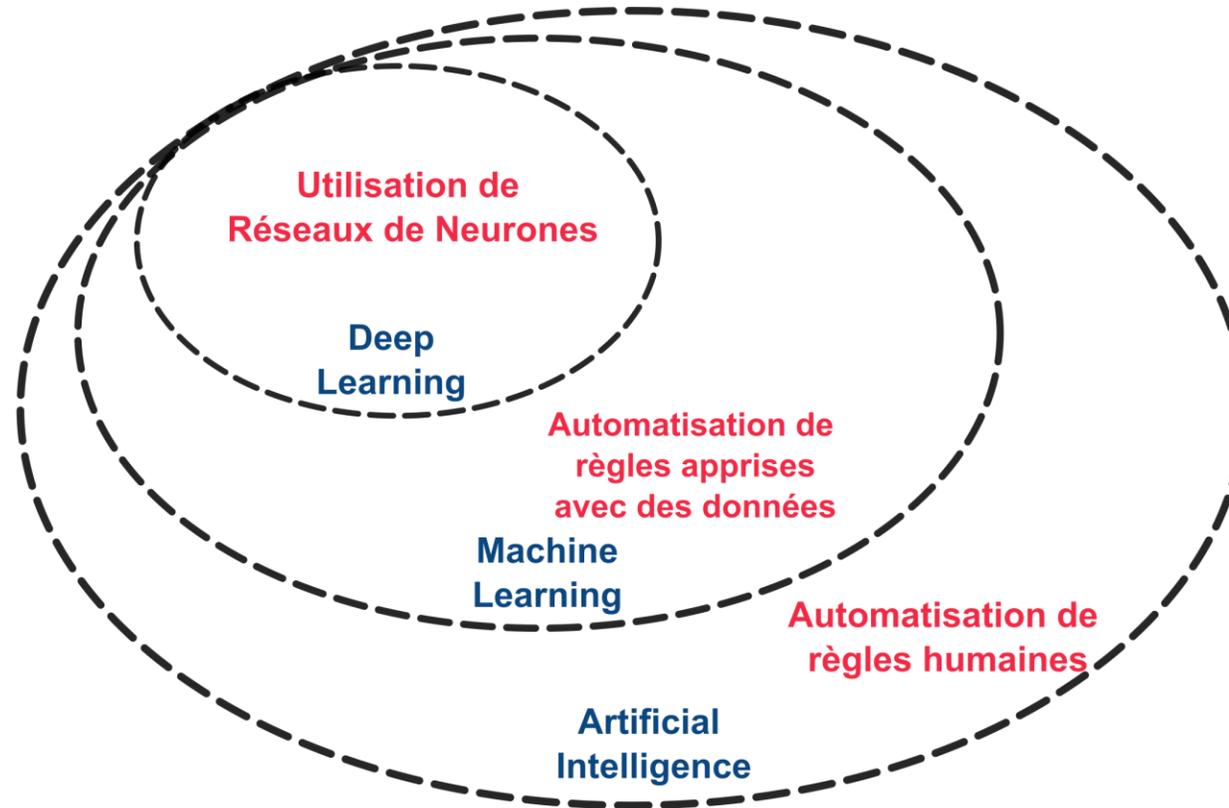
DATA ENGINEER



SOFTWARE ENGINEER

➔ Découvertes et prises en main des outils du Data Scientist

Terminologie



Quelles sont les 3 étapes que va suivre le Data Scientist ?

Travail du Data Scientist



Feature engineering : sélection et transformation des variables d'un jeu de données



Construction et entraînement d'un modèle : calibration d'un modèle sur les données



Évaluation : mesure les performances du modèle sur des données différentes de celles d'entraînement

Travail du Data Scientist



Feature engineering



Construction et entraînement d'un modèle



Évaluation



Répéter

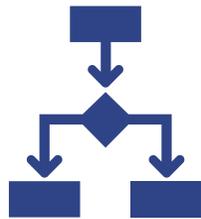
Premiers algorithmes de Machine Learning



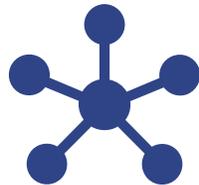
Modèles linéaires



Plus proches voisins



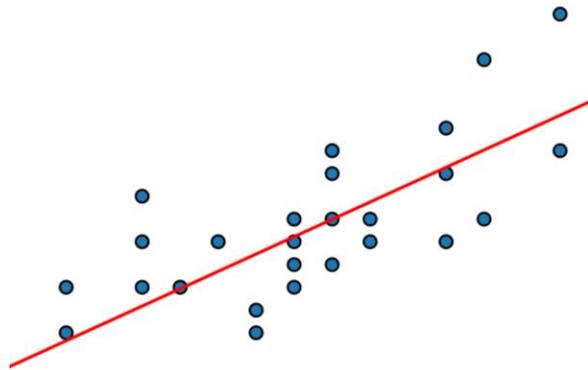
Arbres de décision



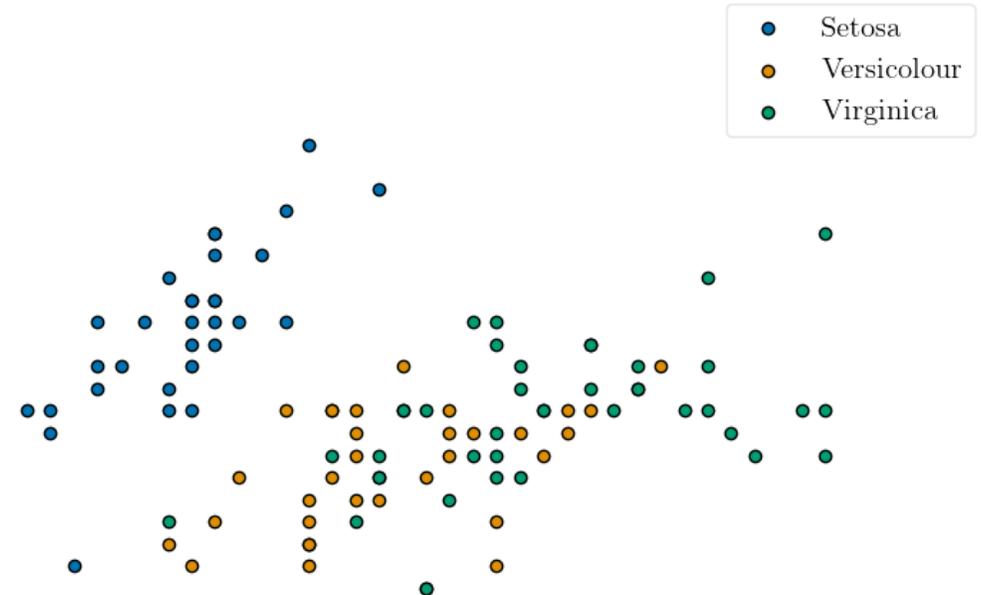
Réseaux de neurones

Régression et Classification

Régression: vise à prédire une valeur numérique continue



Classification: prédire des étiquettes **catégoriques**



Modèles linéaires



Objectif : Apprendre des règles de décisions basées sur **des relations linéaires** (c'est à dire des **lignes droites**) entre différentes variables d'un problème.

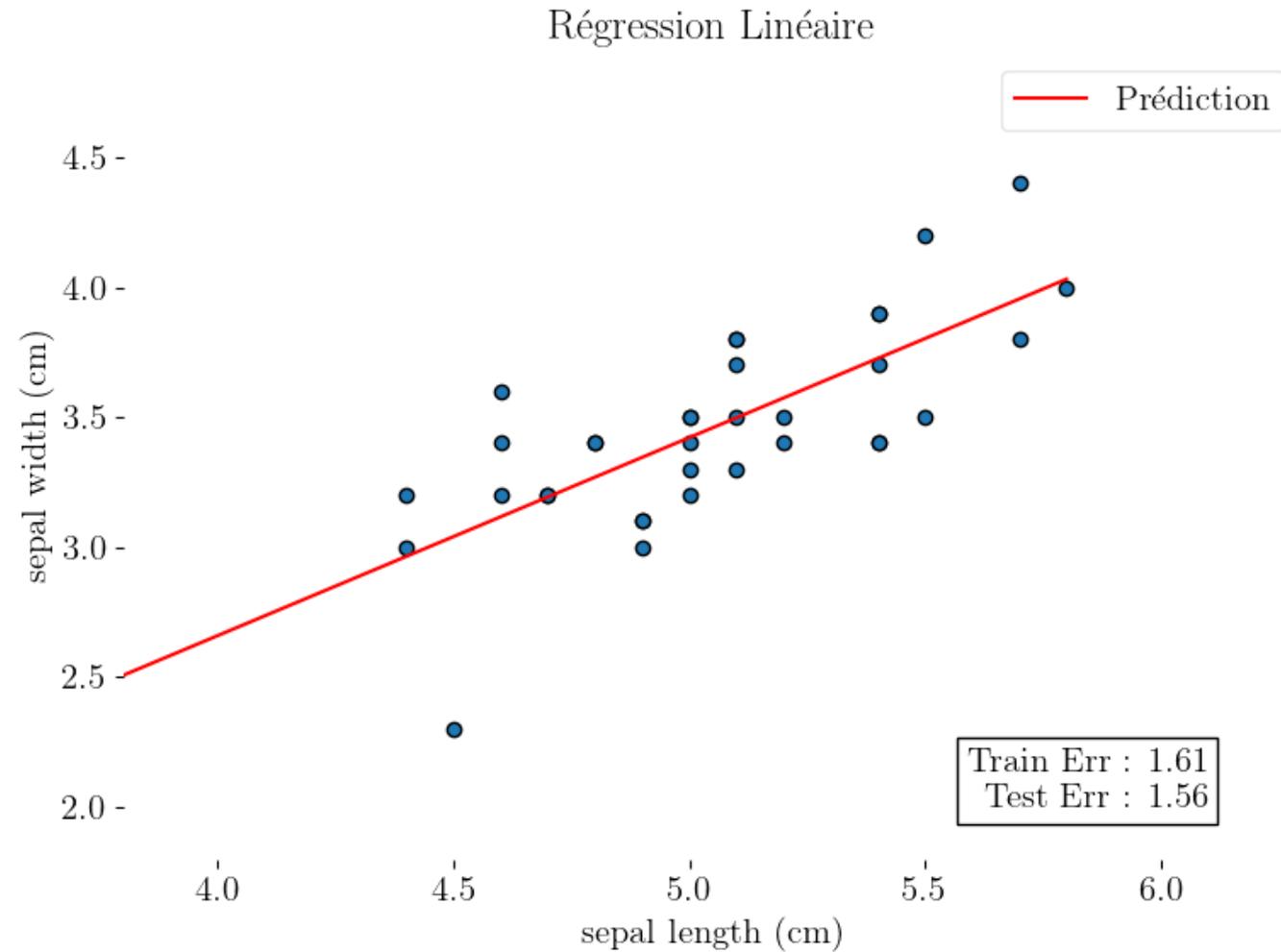


Classification : Tracer une **ligne qui sépare** les classes d'un problème.



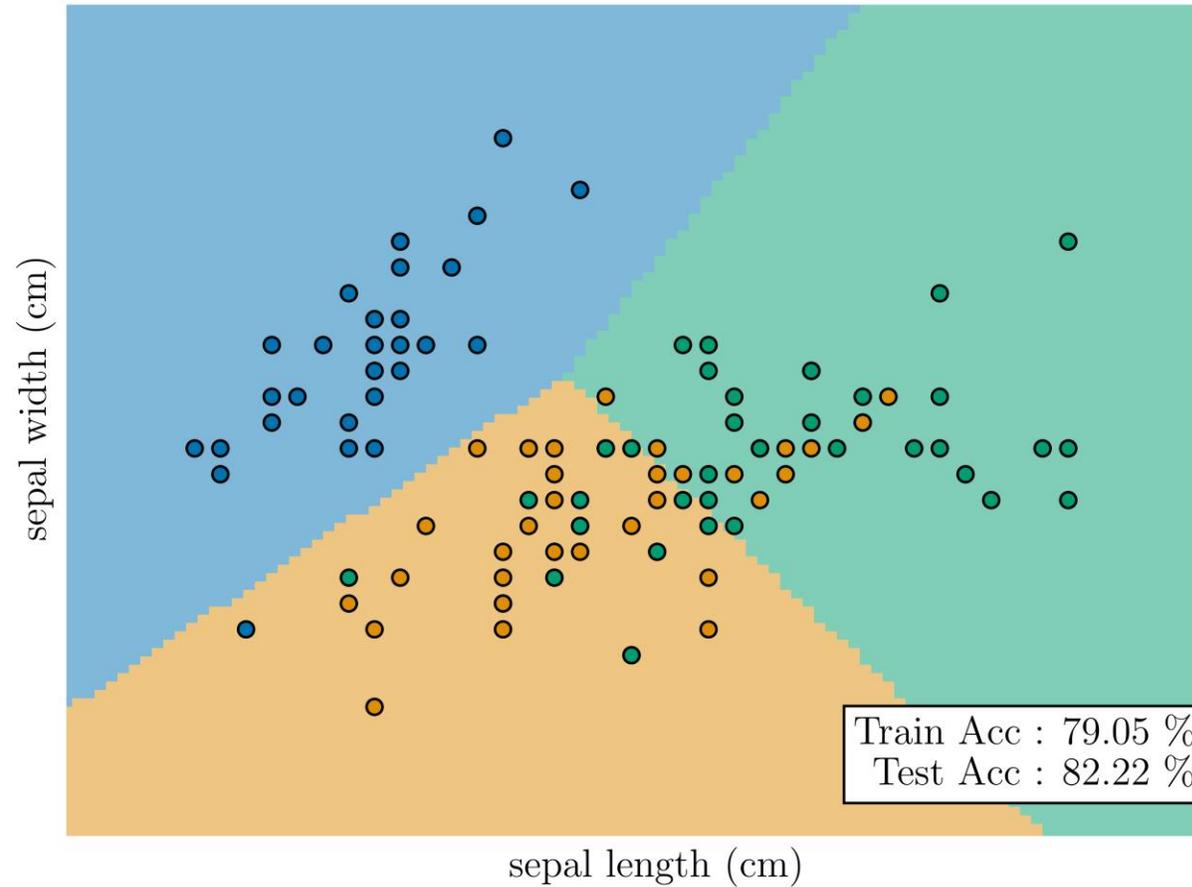
Régression : Trouver une **droite d'interpolation** de la relation entre mes données.

Modèle linéaire - Régression



Modèle linéaire - Classification

Classification linéaire



Avantages / Inconvénients

- **Avantages :**

- Modèles **très simples** à comprendre et intuitifs
- Peu de paramètres, donc **moins sujets à l'overfitting**
- Interprétation de **l'importance des features**
- Très **rapides à entraîner**, les algorithmes d'apprentissage sont très optimisés.
- Efficaces sur une **grande variété de problèmes**

- **Inconvénients :**

- Hypothèse mathématiques de linéarité peut être **trop simples pour certains cas réels complexes**

Plus proches voisins



Objectif : Assigner un label en fonction des données connues les plus similaires.



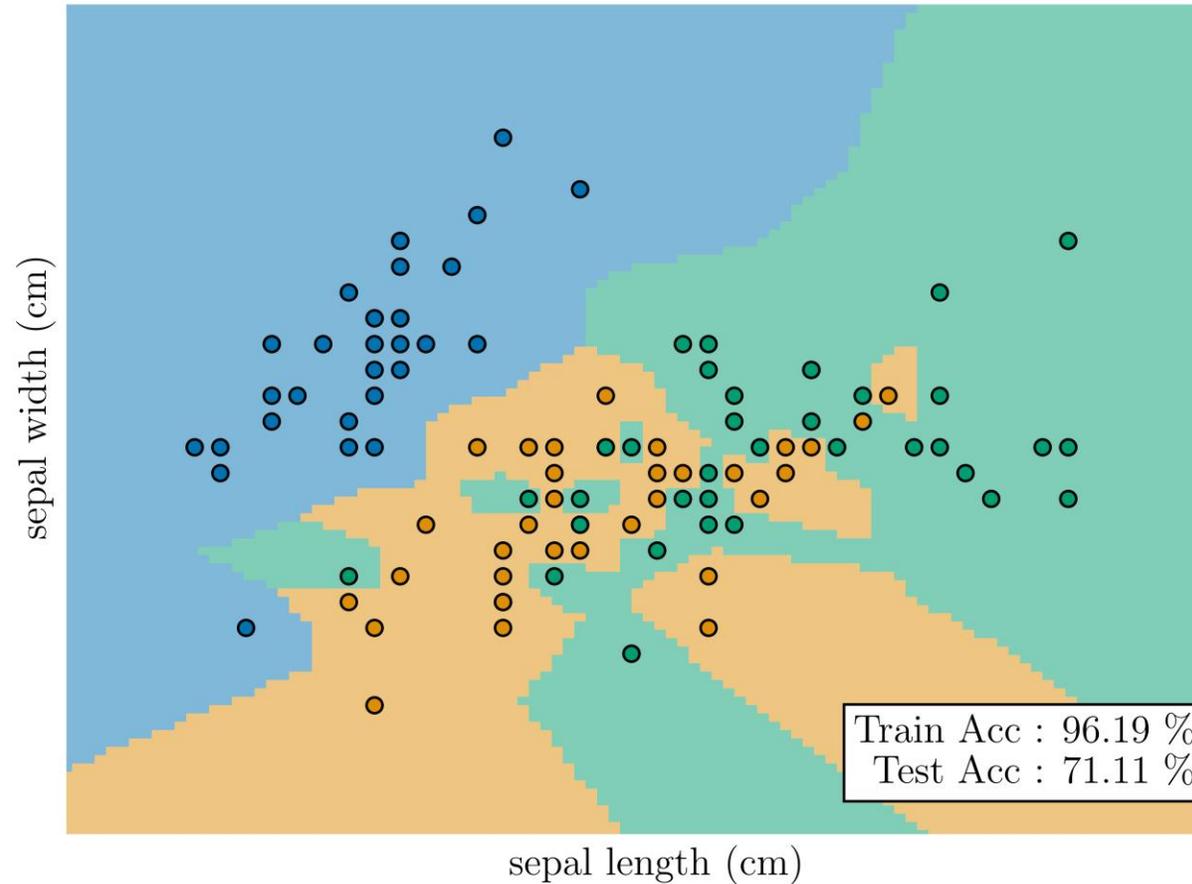
Classification : On assigne la classe majoritaire parmi les données les plus similaires.



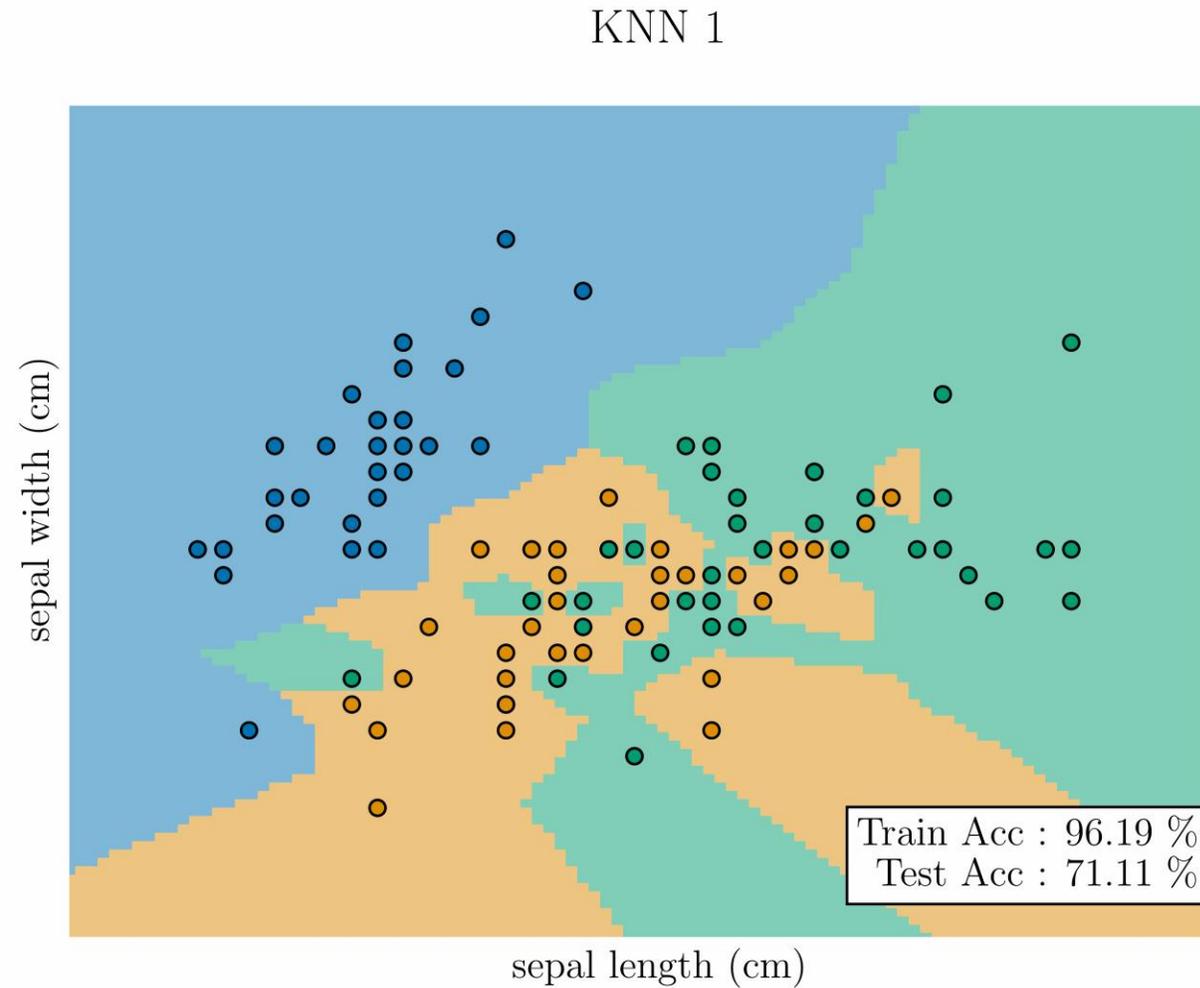
Régression : On assigne la valeur moyenne des observations similaires.

Plus proches voisins (k-NN) - Classification

KNN 1

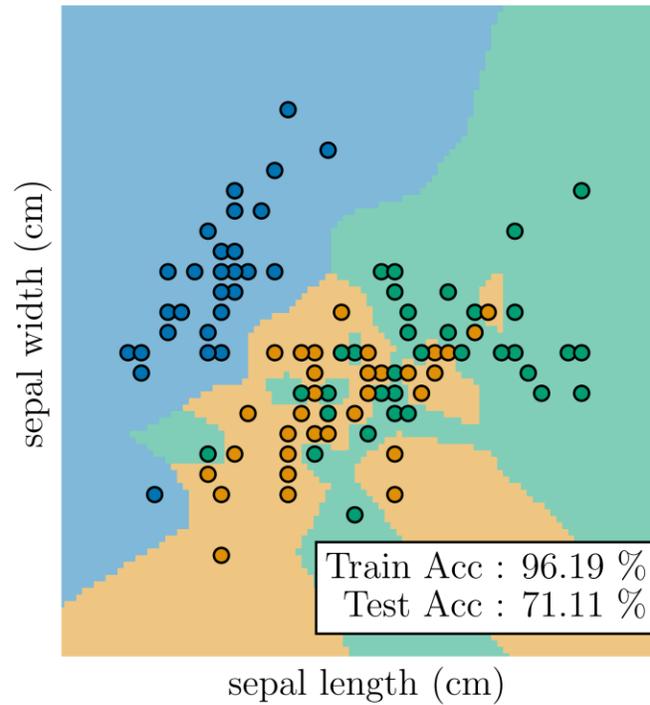


Plus proches voisins (k-NN) - Classification

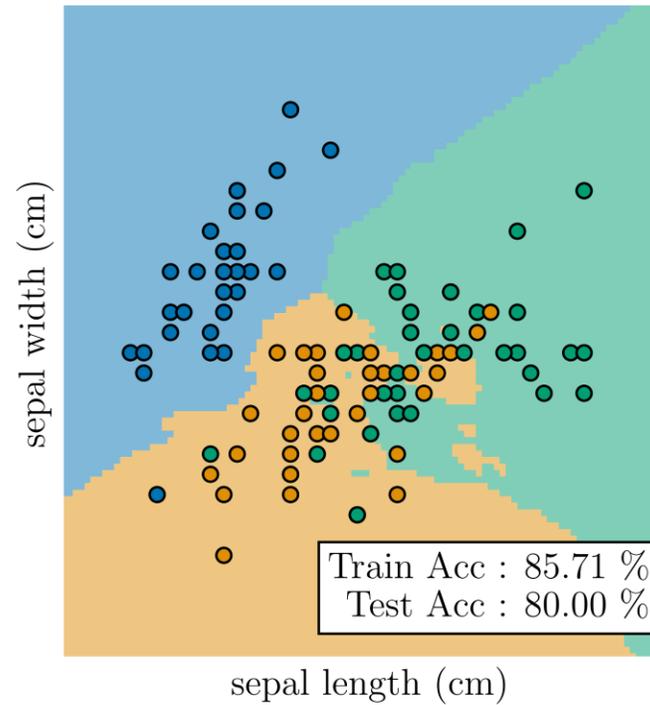


Plus proches voisins (k-NN)

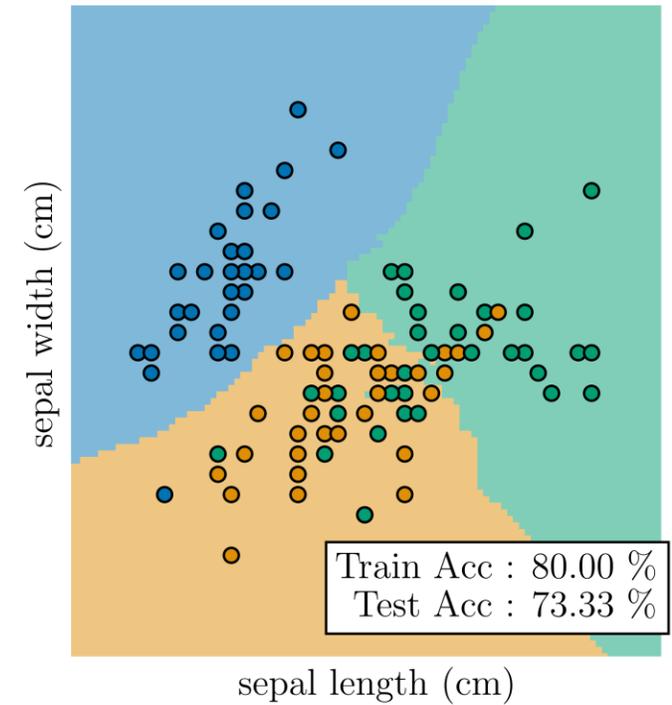
KNN 1



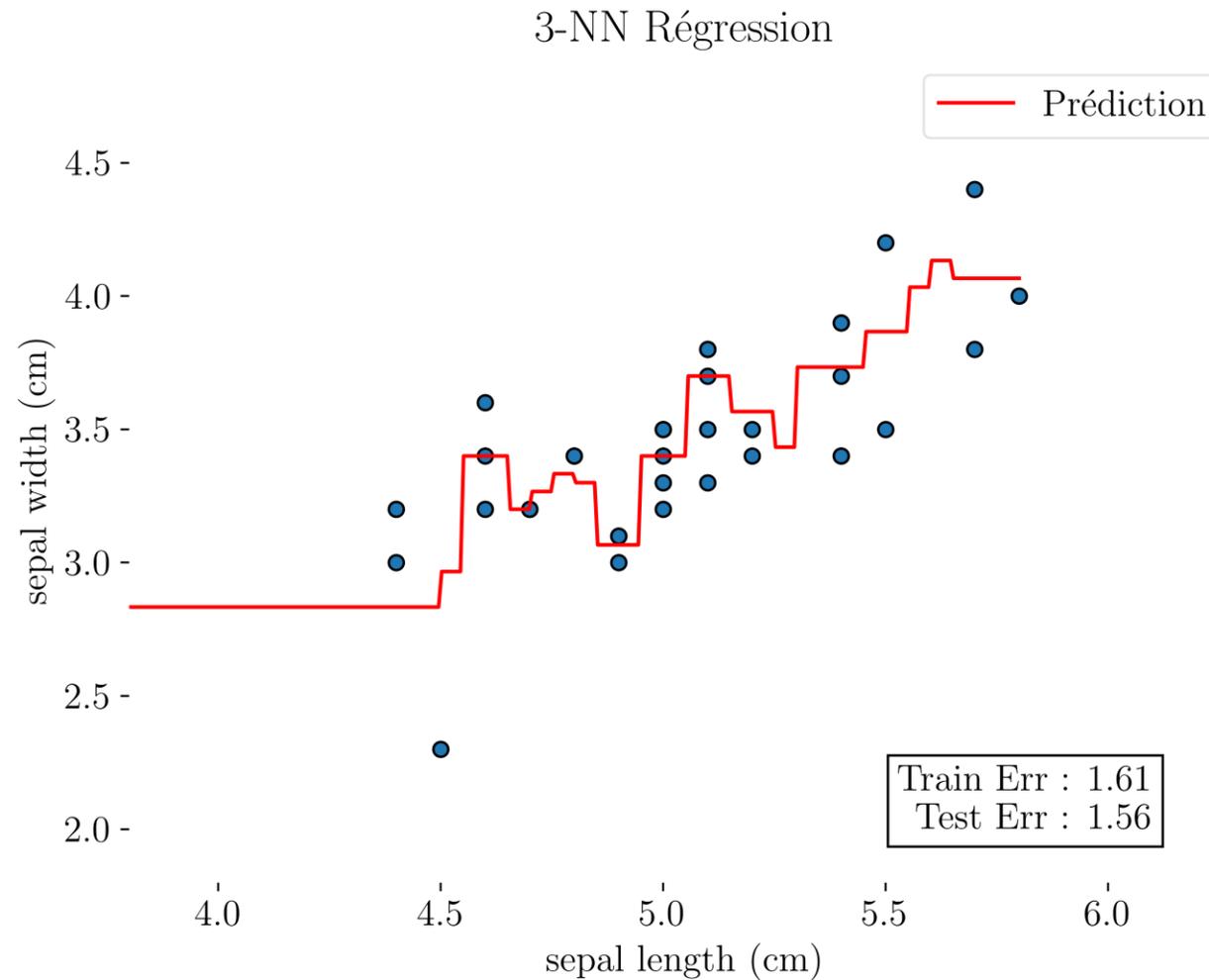
KNN 5



KNN 20



Plus proches voisins (kNN) - Régression



Avantages / Inconvénients

- **Avantages :**

- Modèles **très simples** à comprendre et intuitifs
- Gestion des données non-linéaires
- Basé **uniquement sur les données**, pas de modélisation

- **Inconvénients :**

- Sensibles à la définition de **similarité** retenue
- Ne fais **pas de sélection de features**
- Ne passe pas à l'échelle pour des jeux de données avec beaucoup de features

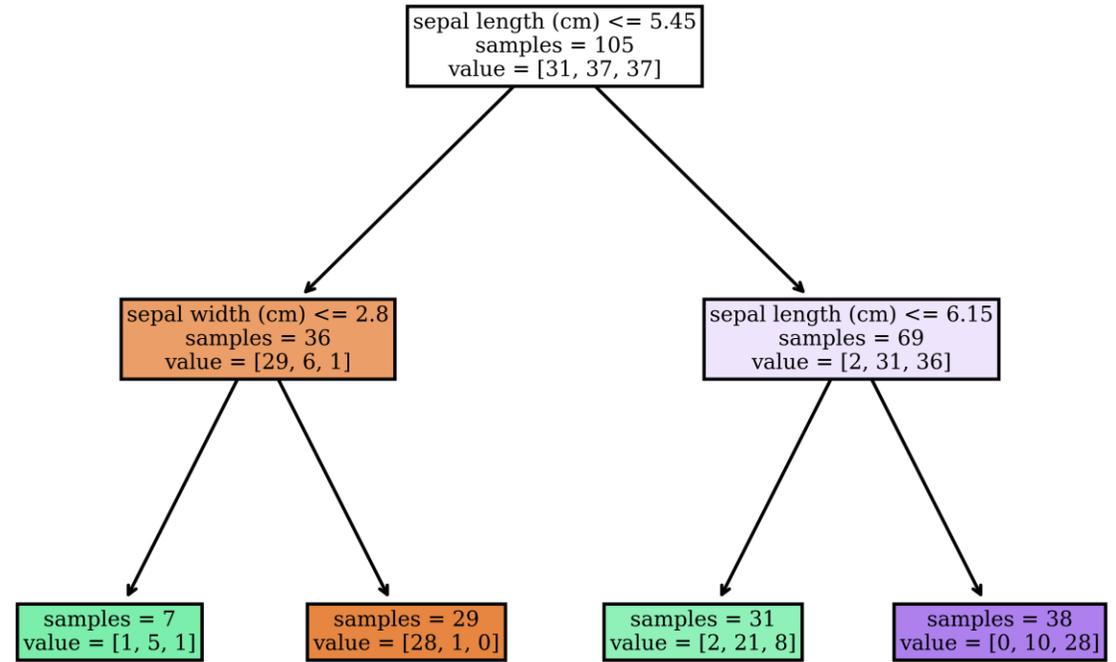
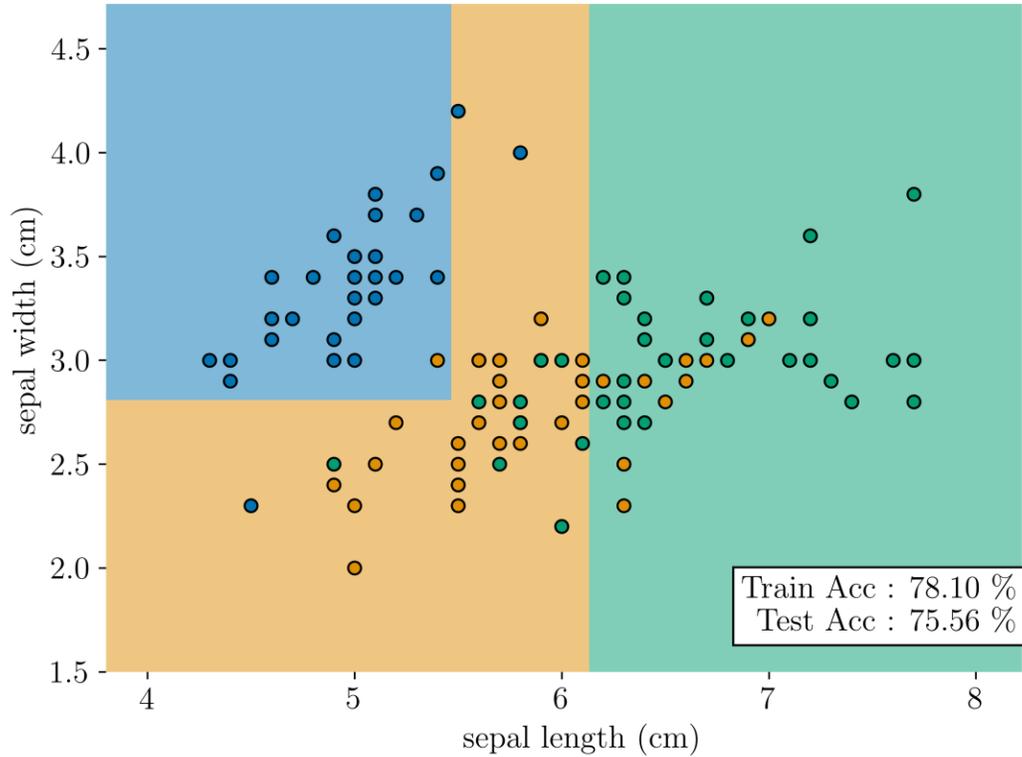
Arbres de décision



Objectif : Séparer les données selon les valeurs de leurs features.

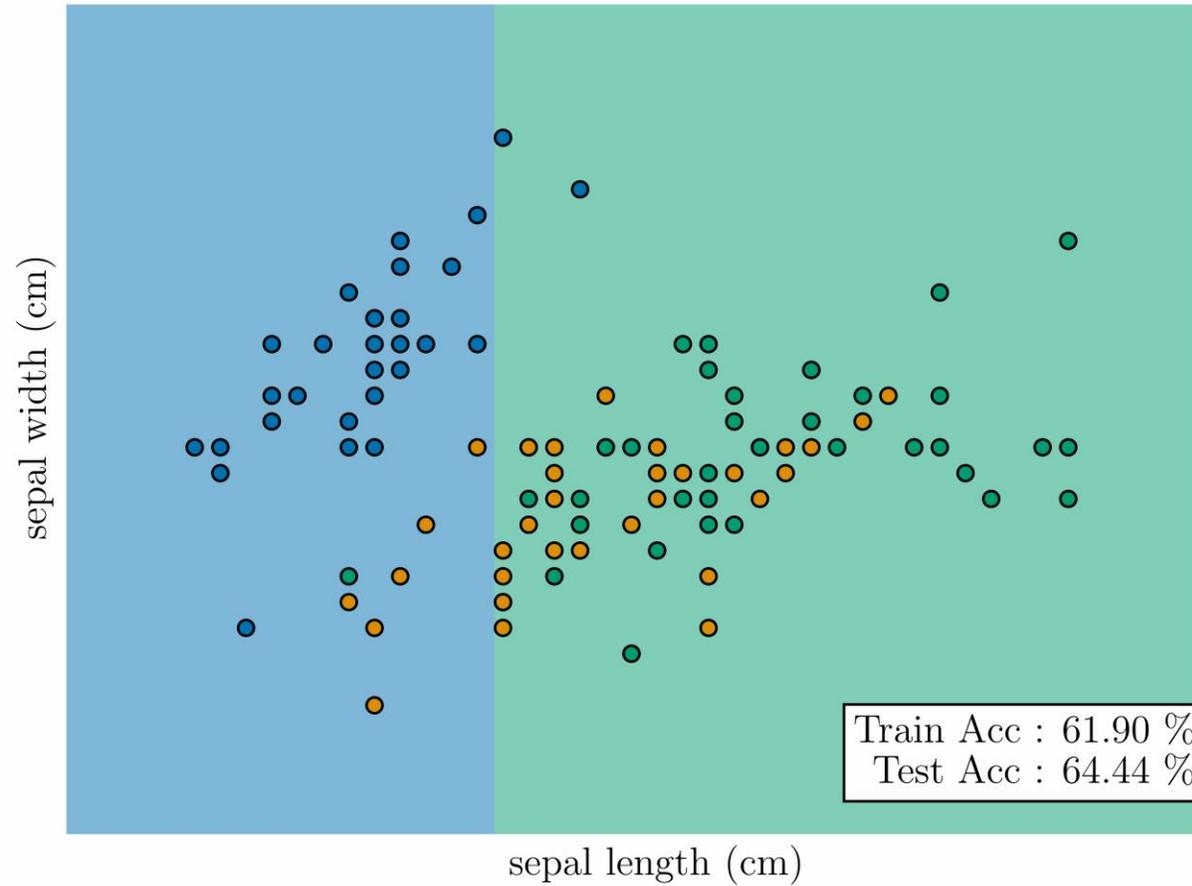
Arbres de décision

Decision Tree with depth 2



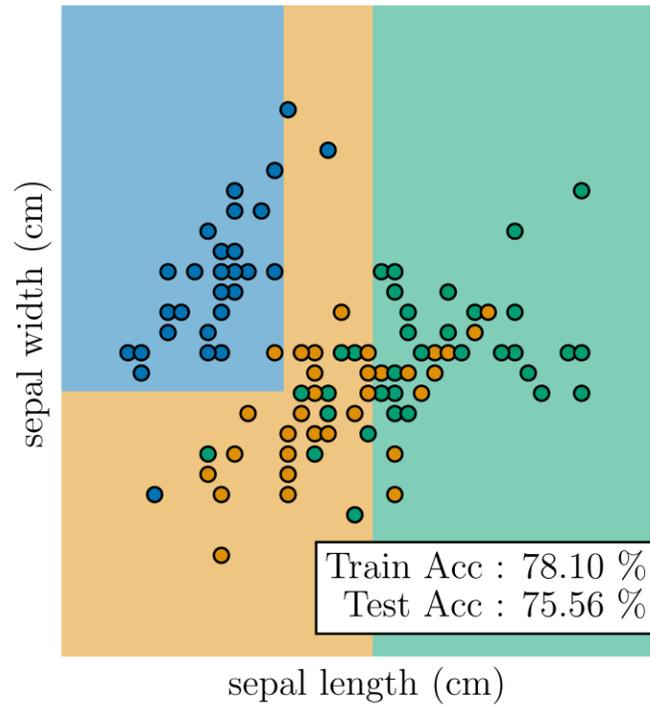
Arbres de décision

Decision Tree with depth 1

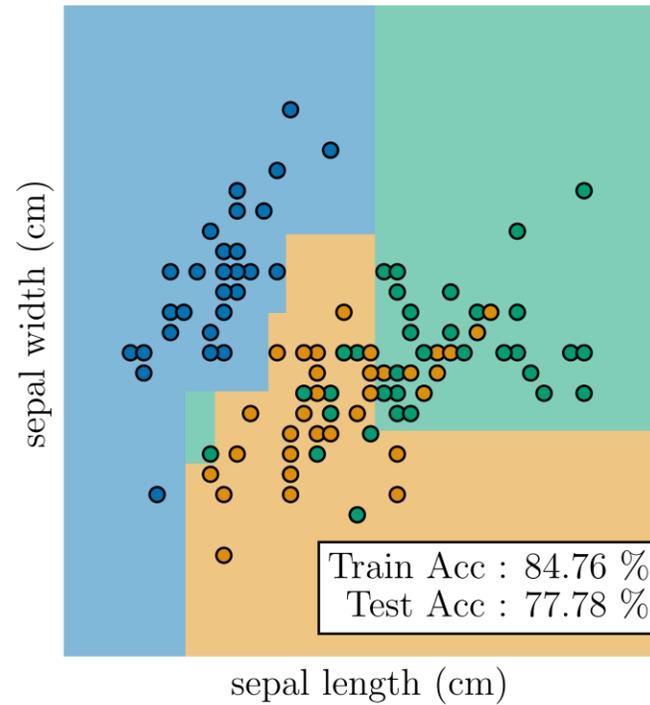


Arbres de décision

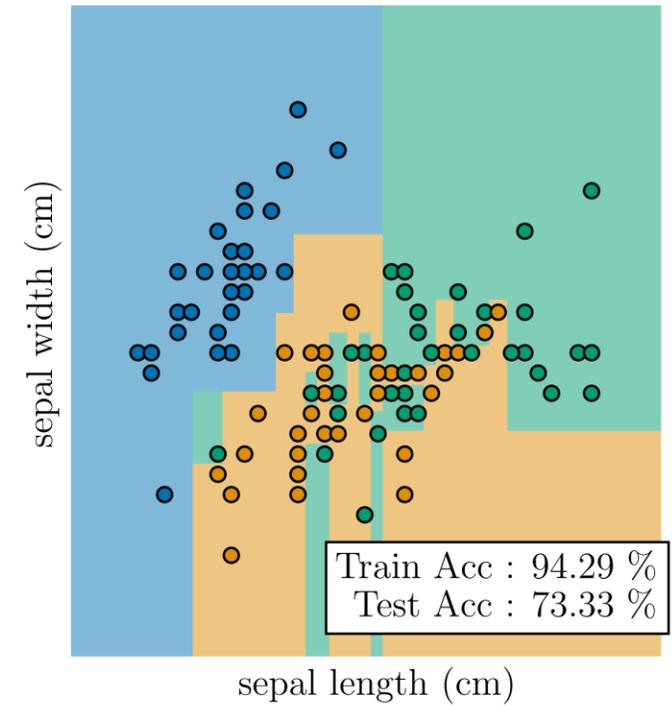
Depth 2



Depth 5



Depth 10



Avantages / Inconvénients

- **Avantages :**

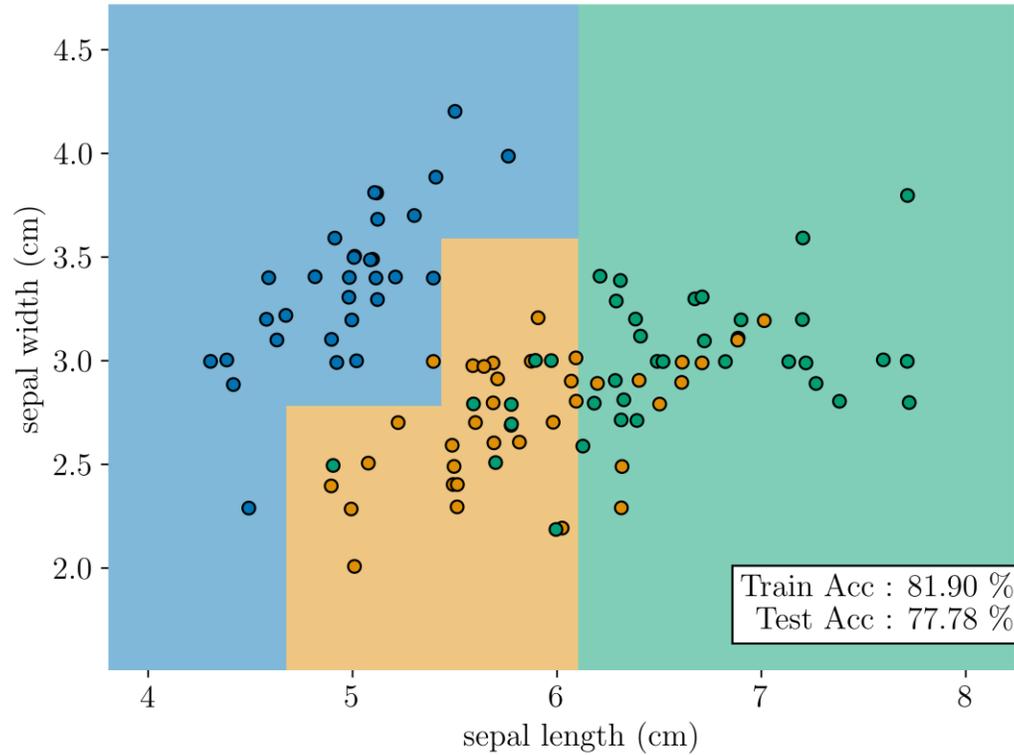
- Modèles interprétables
- **Sélection de features**
- Gestion des données non-linéaires

- **Inconvénients :**

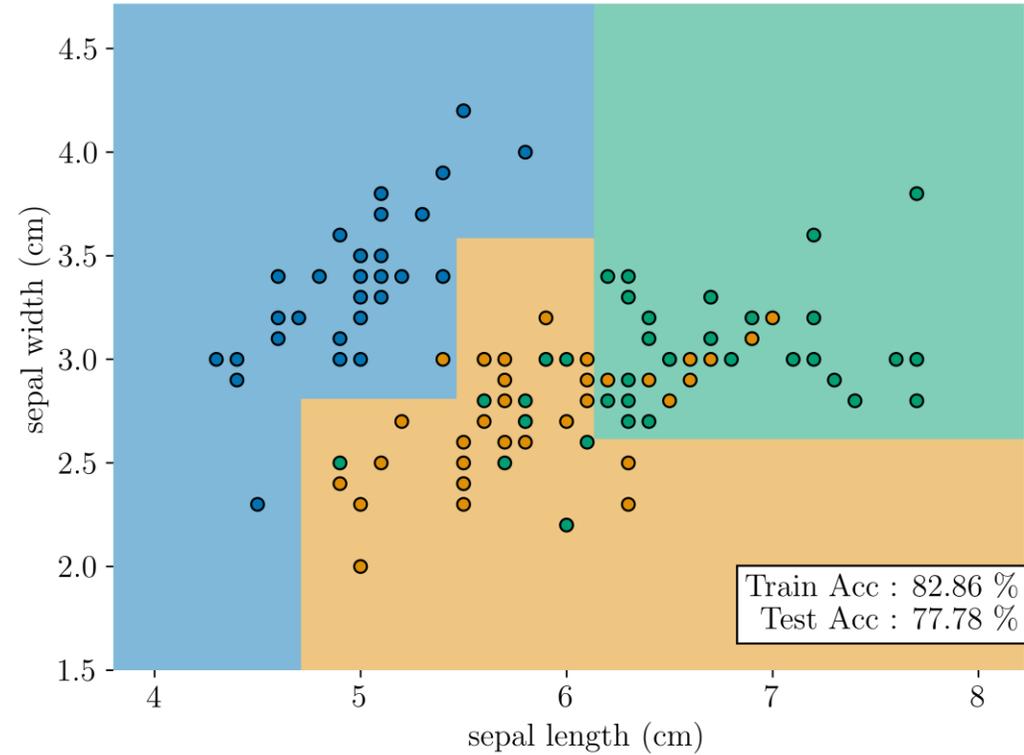
- Overfitting lorsque la profondeur est trop grande
- **Sensibles aux bruits** (la structure peut-être complètement modifiée si les données changent un petit peu)

Sensibilité au bruit

Decision Tree with depth 3

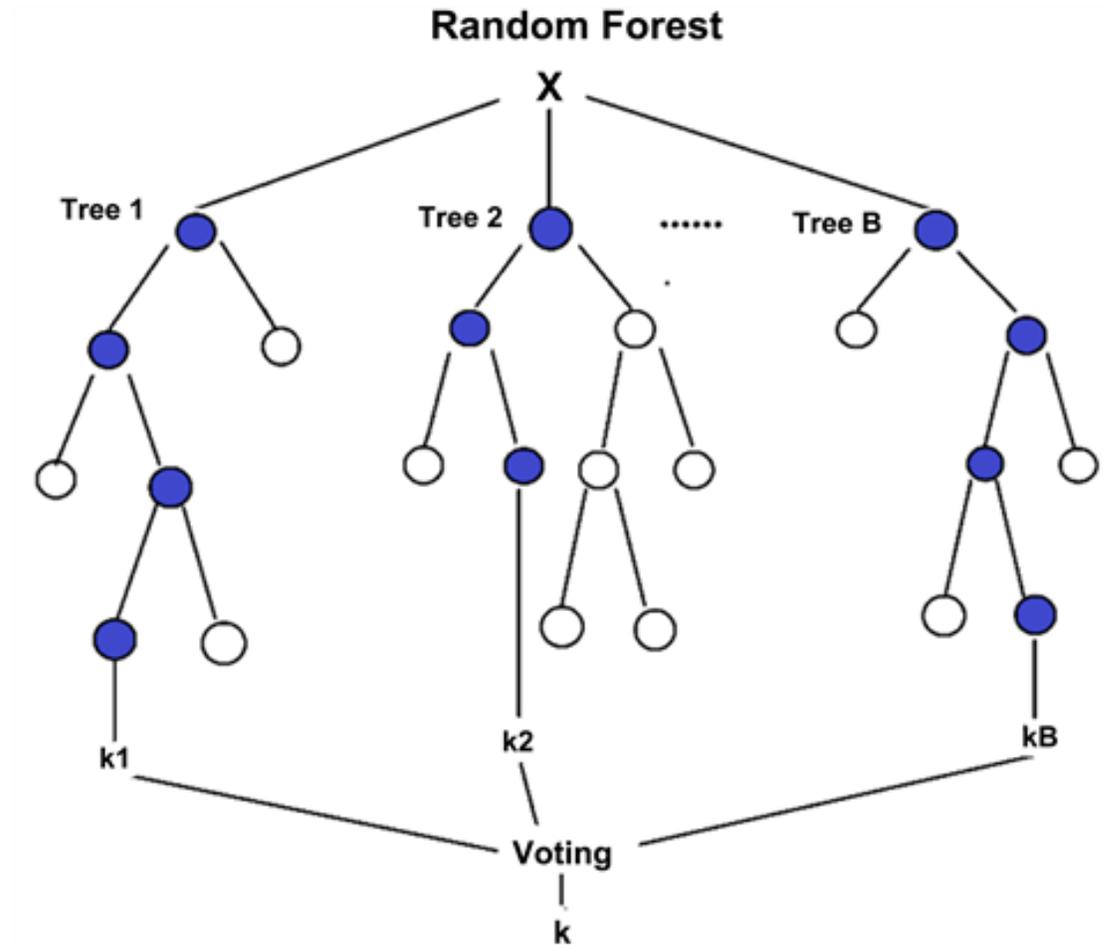


Decision Tree with depth 3



Forêts aléatoires

- La forêt aléatoire est **modèle ensembliste**.
- Il est composé d'un ensemble d'arbre de décision.
- Un modèle d'ensemble agrège la prédiction de plusieurs **modèles différents**.



Avantages / Inconvénients

- **Avantages :**

- **Sélection de features**
- Gestion des données non-linéaires
- **Robuste au bruit**
- Réduit l'overfitting

- **Inconvénients :**

- Plus de paramètres que les arbres simples
- Plus complexes à bien optimiser (plusieurs hyperparamètres)
- Moins interprétables que les arbres de décision

Réseaux de neurones



Objectif : Apprendre des règles de décision continues grâce à un modèle complexes.

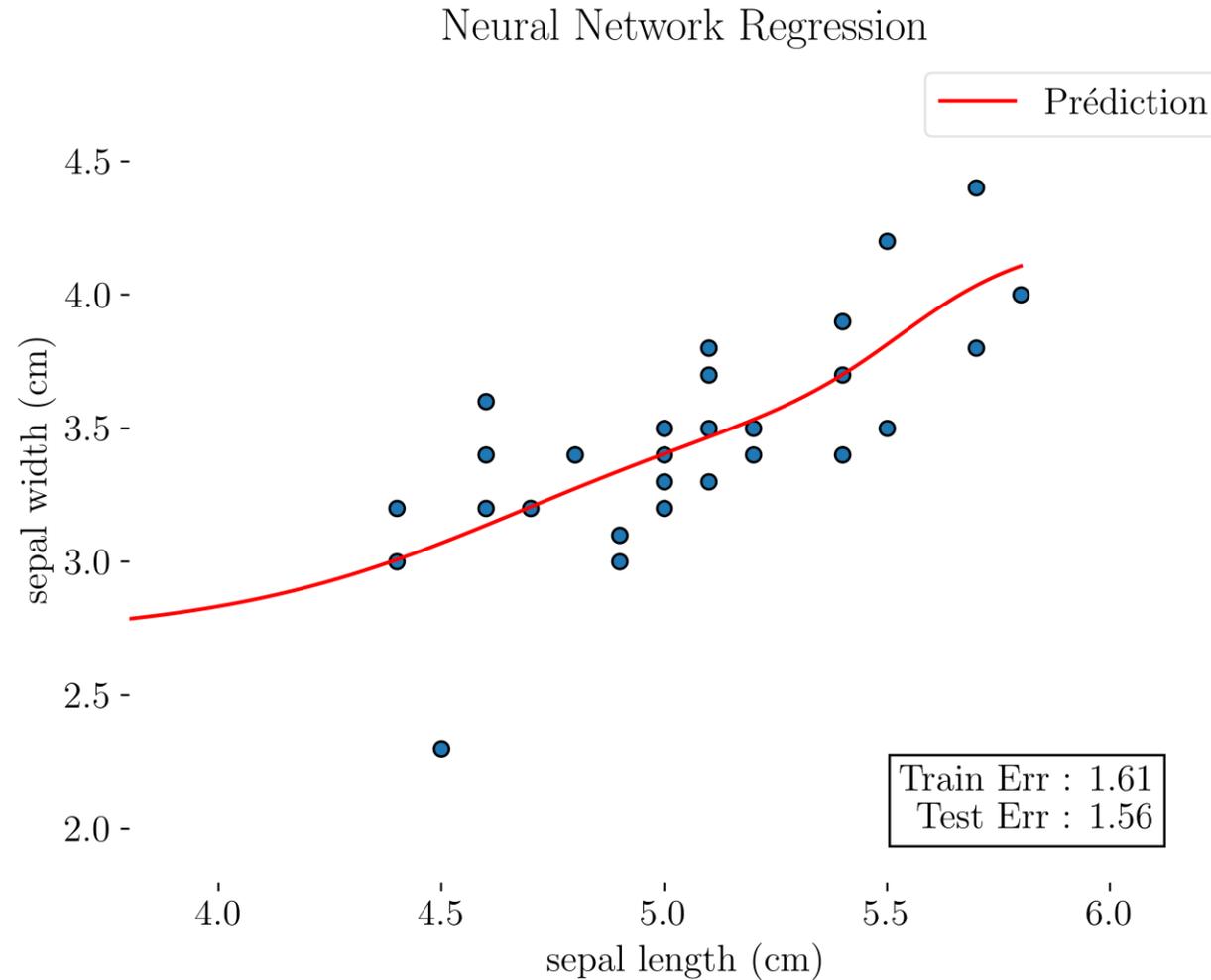


Classification : Trouver une séparation non-linéaire des données.



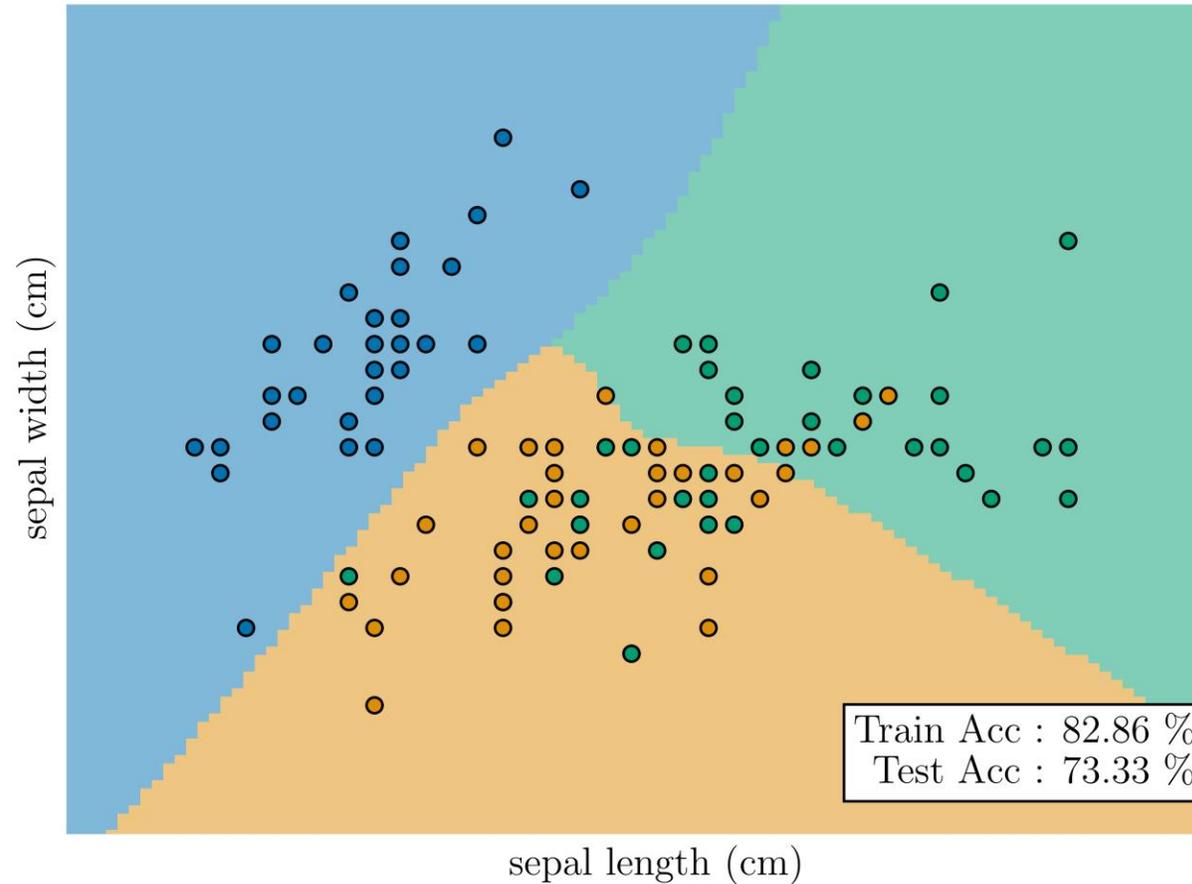
Régression : Interpoler de manière non-linéaire les données.

Réseaux de neurones - Régression



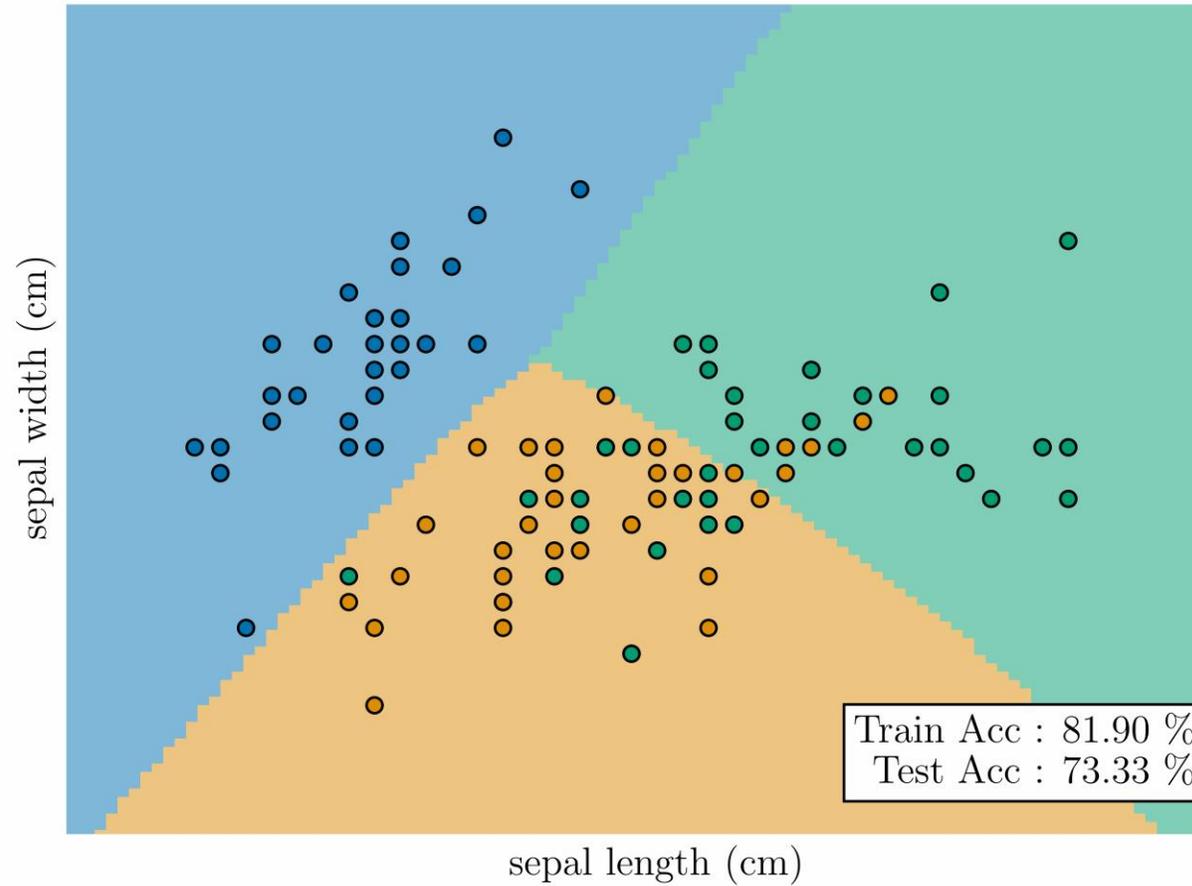
Réseaux de neurones - Classification

Neural Network Classification



Réseaux de neurones - Classification

NN - 100 iterations



Avantages / Inconvénients

- **Avantages :**

- Capacité à apprendre des **relations complexes non-linéaires**
- Capables de s'entraîner sur des **jeux de données très volumineux**

- **Inconvénients :**

- Besoin d'une grande quantité de données
- **Non interprétables**
- Sensibles à l'**overfitting**

Algorithmes de Machine Learning : résumé

- 3 algorithmes très simples à toujours essayer : **modèles linéaires, k-NN, arbres de décision**
- Les **forêts aléatoires** sont des algorithmes très puissants, meilleures que les réseaux de neurones dans certaines de tâches
- Les modèles linéaires, mais surtout les arbres et forêts aléatoires sont capable de faire de la **sélection de features**

Use-Cases

- Prédiction d'attribution de crédit
- Prédiction d'utilisation de vélo en libre-service
- Prédiction de résiliation d'abonnement
- Classification d'images de chiffres

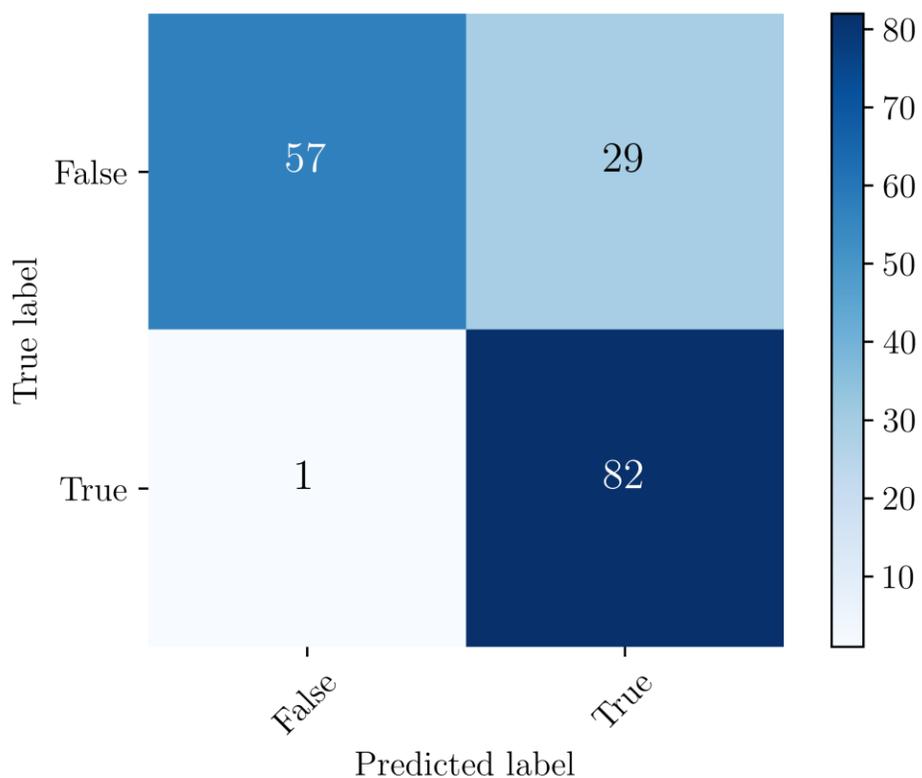
Attribution de Crédit

Aperçu du jeu données :

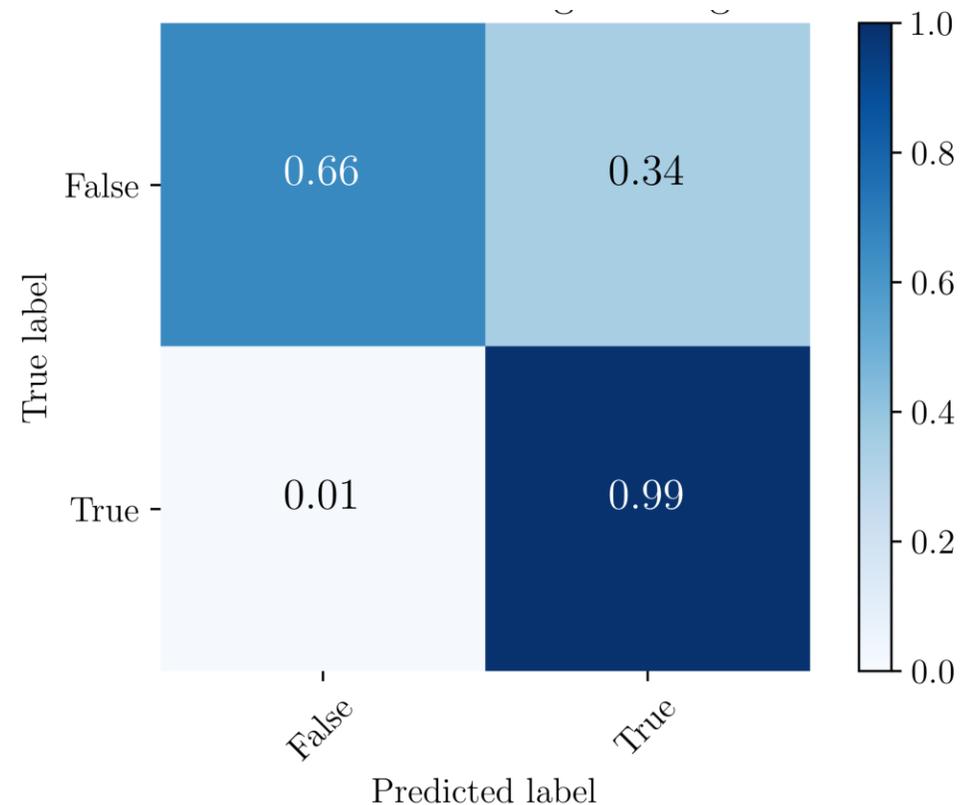
	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Status
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0	Urban	Y
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0	Rural	N
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0	Urban	Y
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0	1.0	Urban	Y
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0	Urban	Y
5	LP001011	Male	Yes	2	Graduate	Yes	5417	4196.0	267.0	360.0	1.0	Urban	Y
6	LP001013	Male	Yes	0	Not Graduate	No	2333	1516.0	95.0	360.0	1.0	Urban	Y
7	LP001014	Male	Yes	3+	Graduate	No	3036	2504.0	158.0	360.0	0.0	Semiurban	N
8	LP001018	Male	Yes	2	Graduate	No	4006	1526.0	168.0	360.0	1.0	Urban	Y
9	LP001020	Male	Yes	1	Graduate	No	12841	10968.0	349.0	360.0	1.0	Semiurban	N

Modèle linéaire

Classifieur Linéaire - Matrice de Confusion

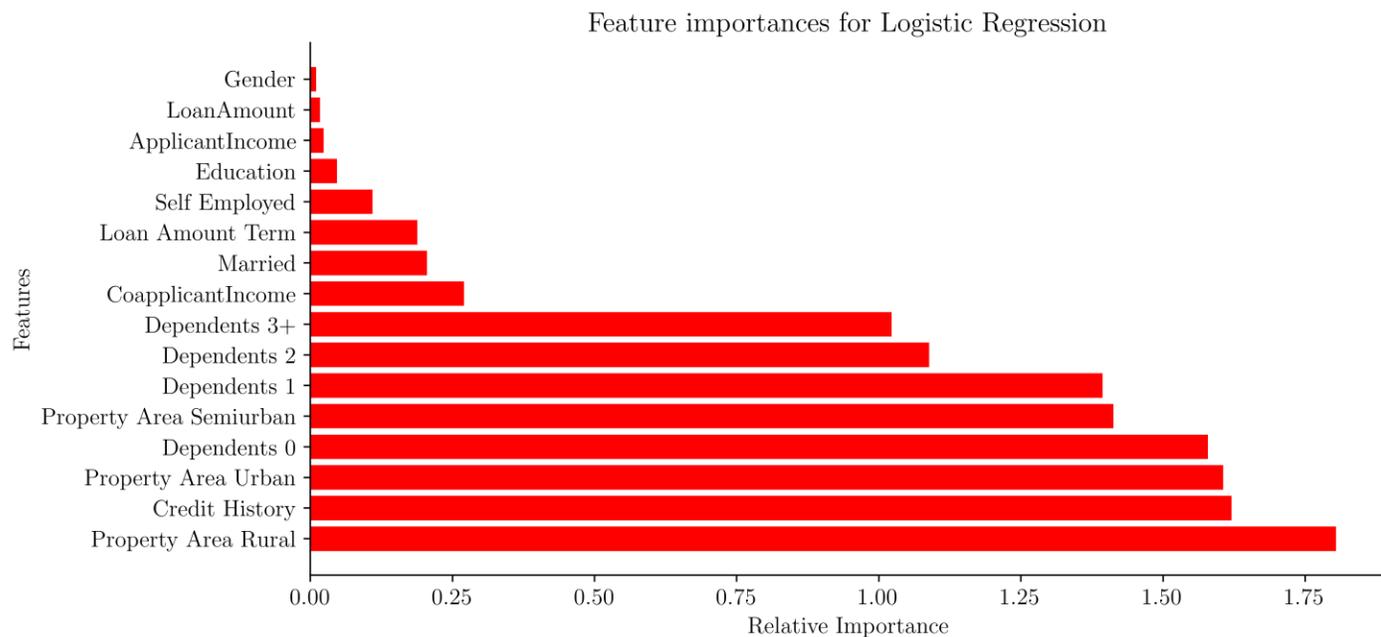


Matrice de confusion – valeurs brutes

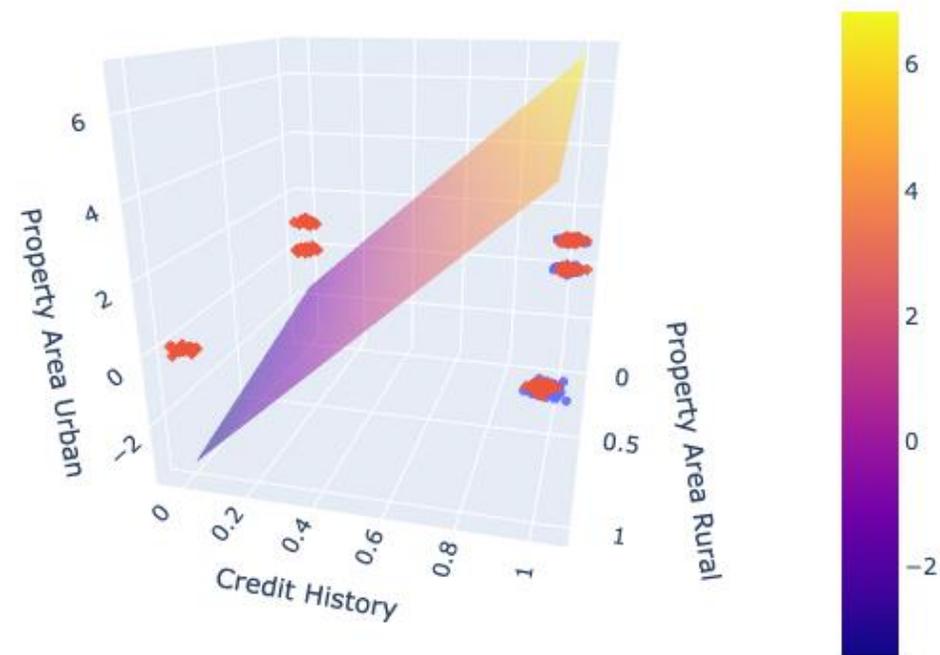


Matrice de confusion – Pourcentages

Régression Linéaire



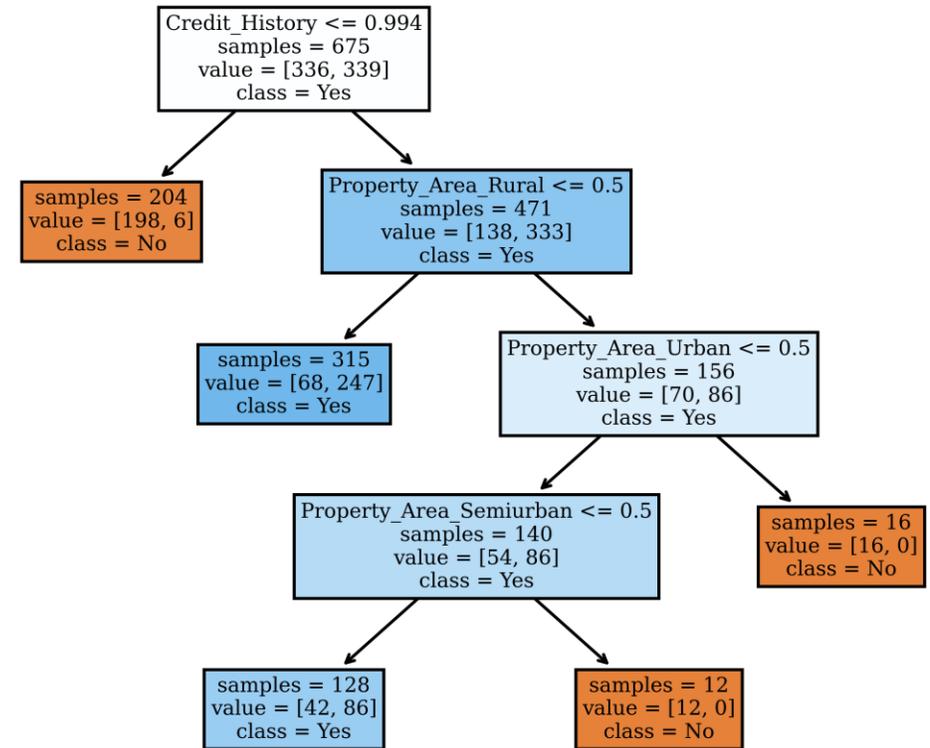
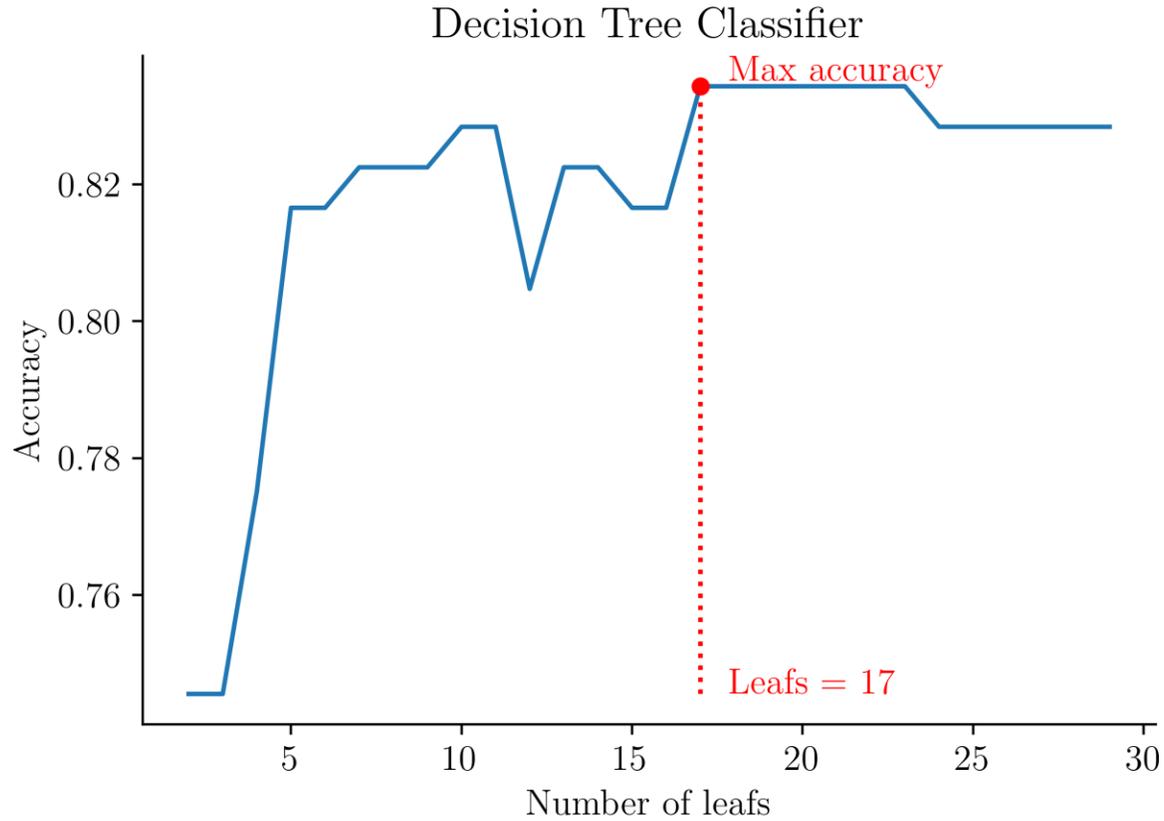
Importances des attributs dans l'attribution du crédit



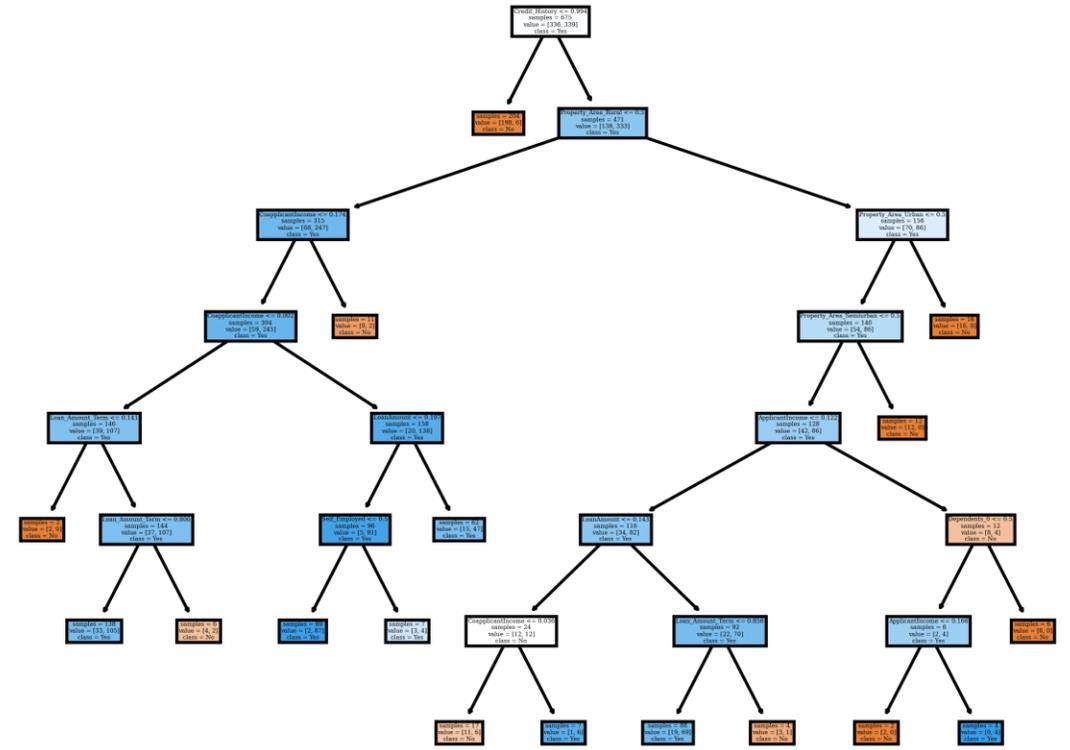
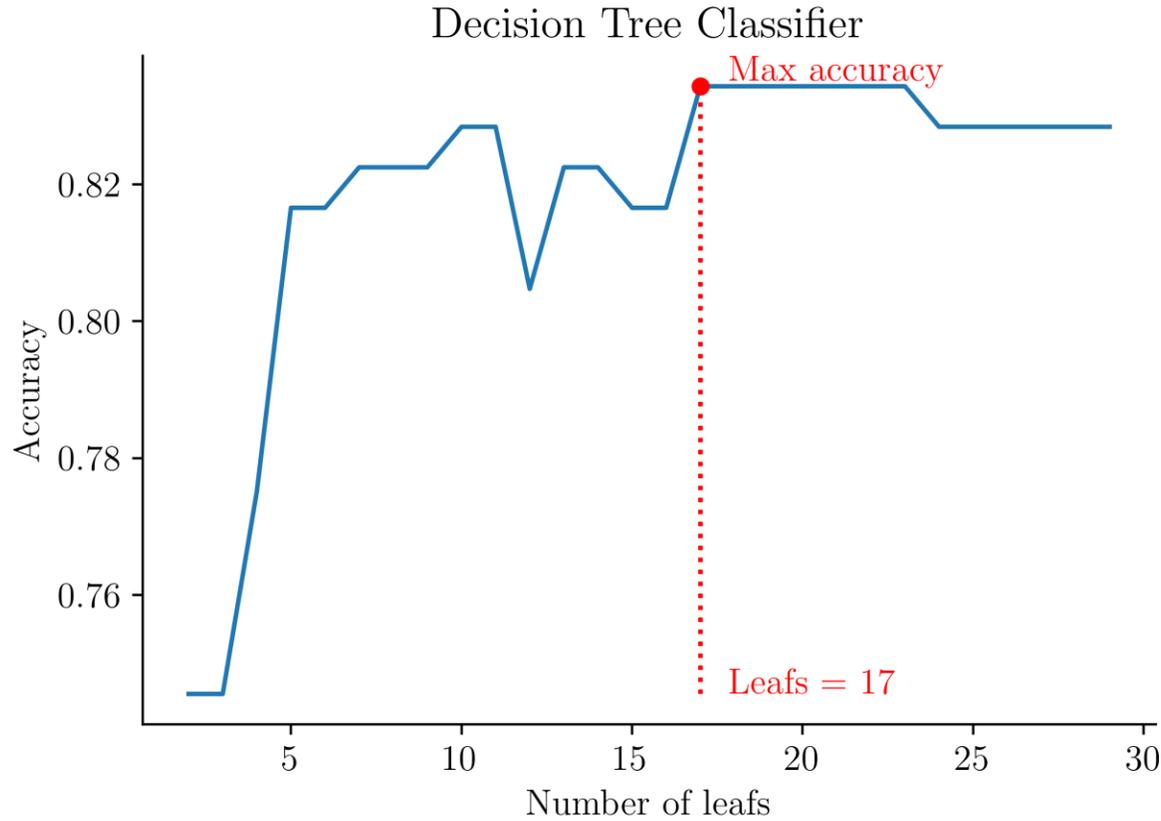
Frontière de séparation sur les 3 attributs les plus discriminants

Arbre de décision

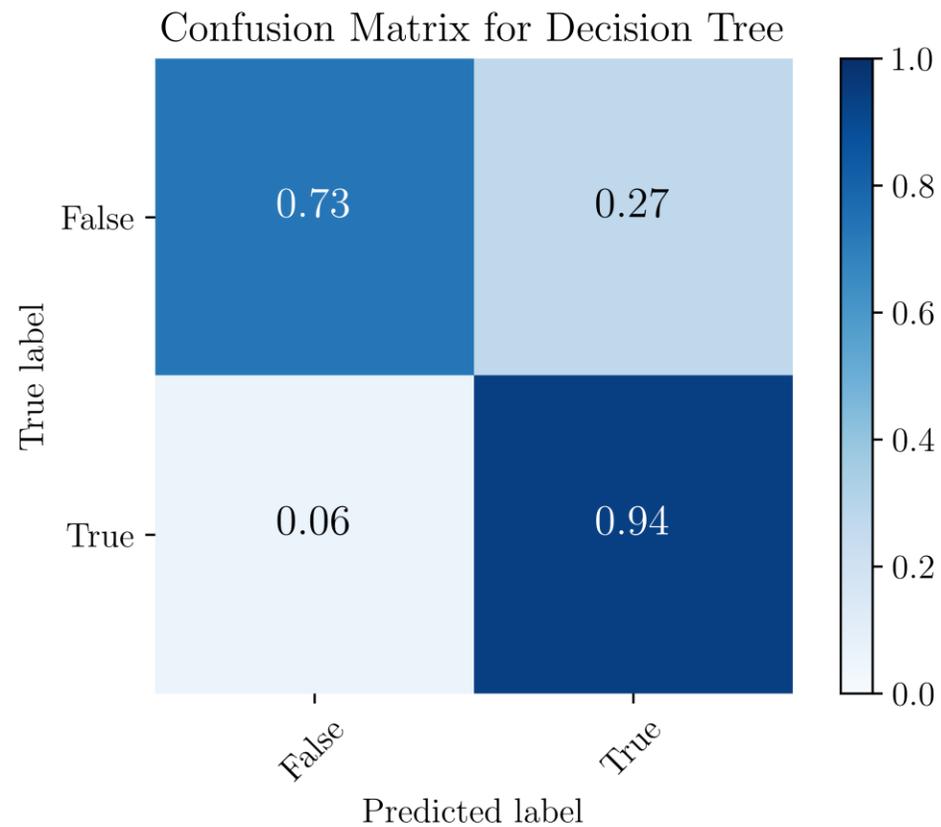
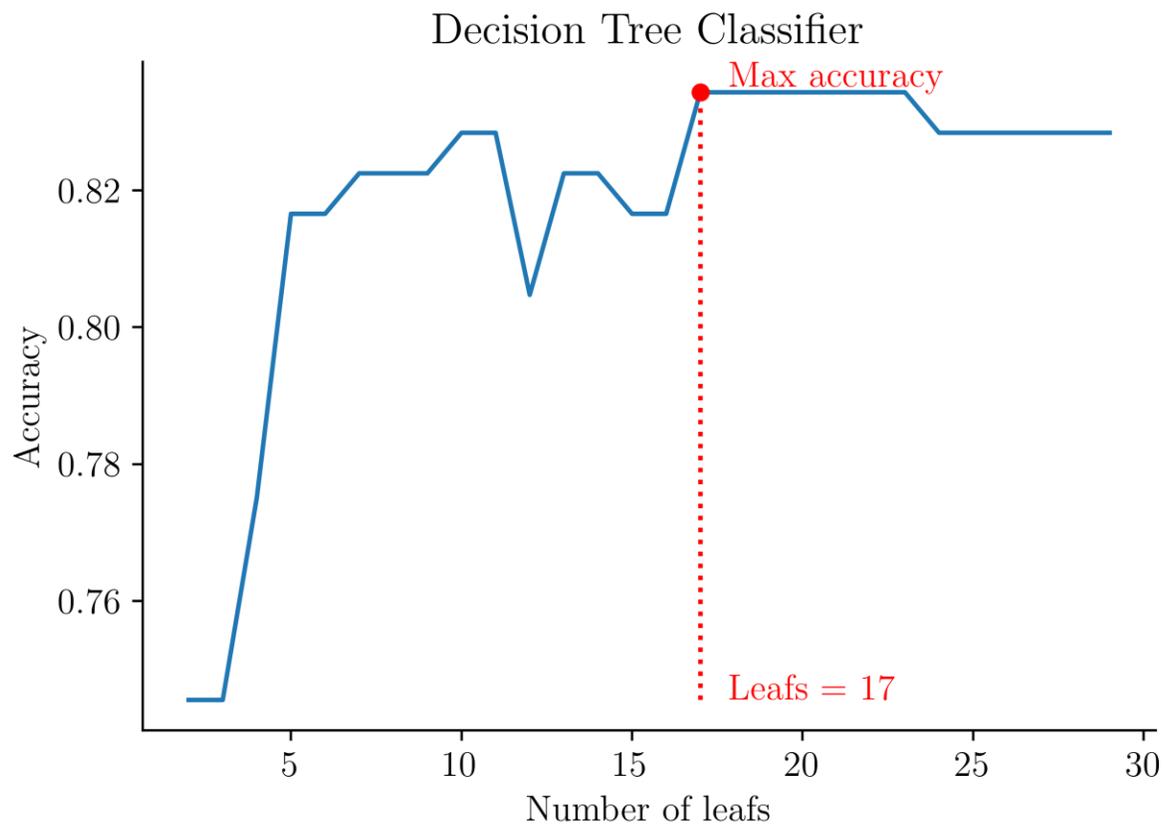
Arbre de Décision



Arbre de Décision

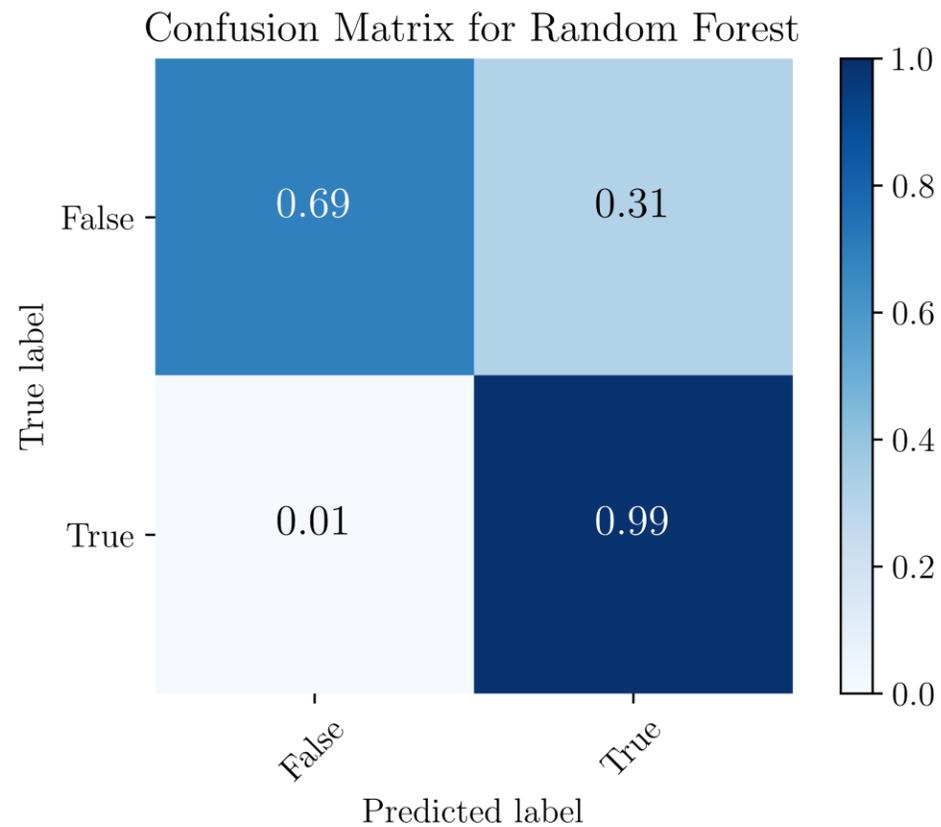
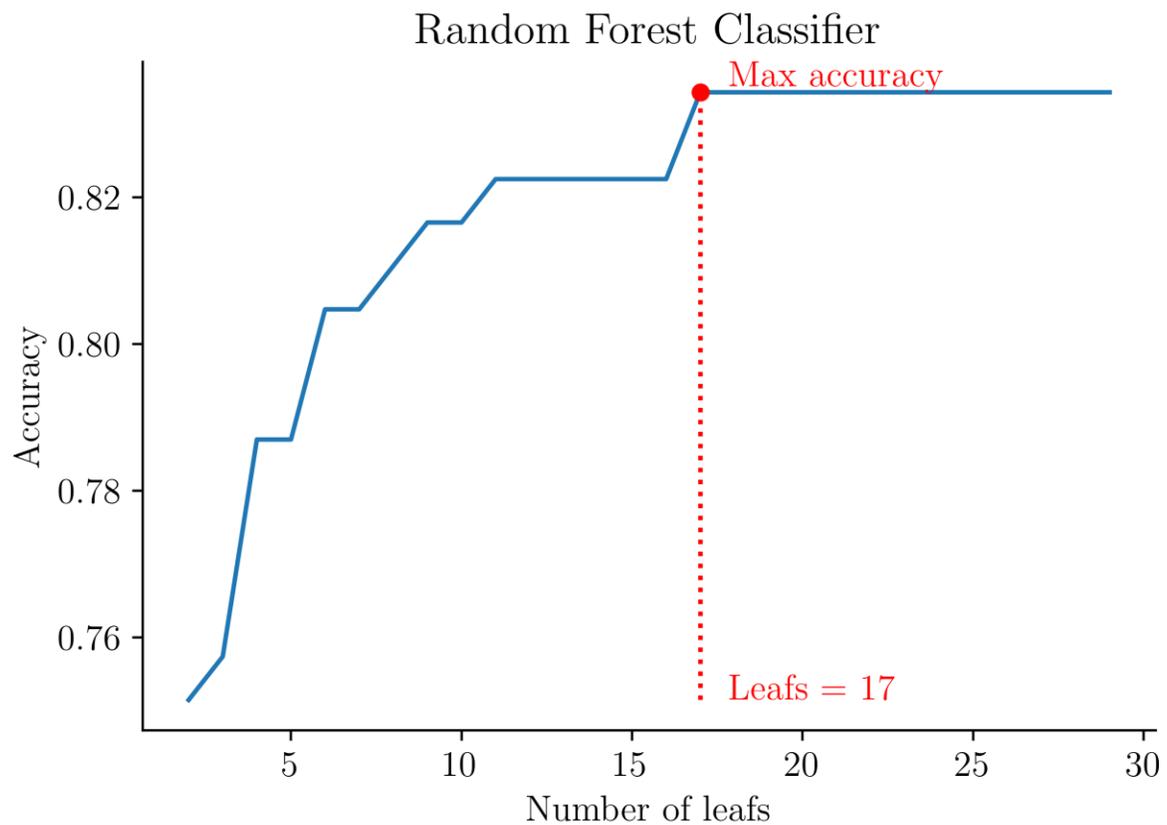


Arbre de Décision



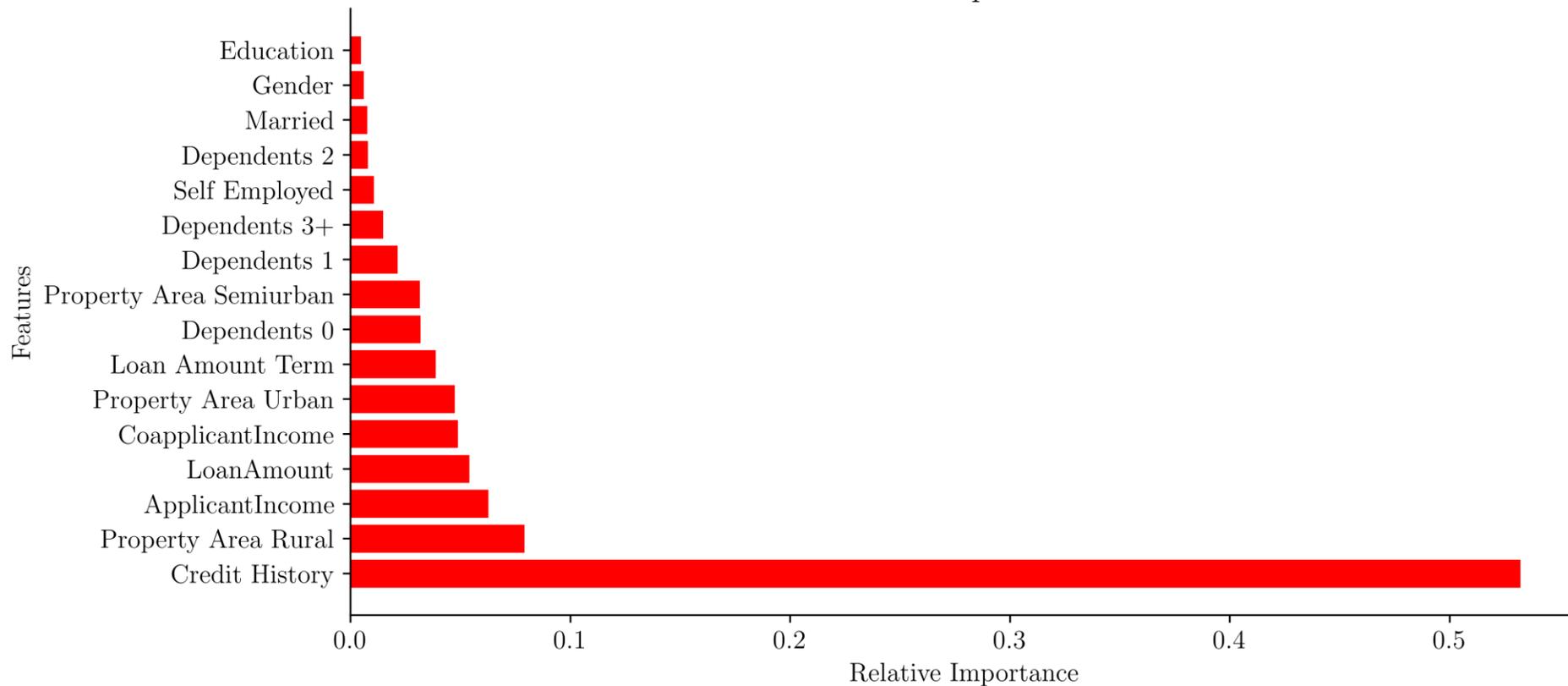
Forêt Aléatoire

Forêt Aléatoire



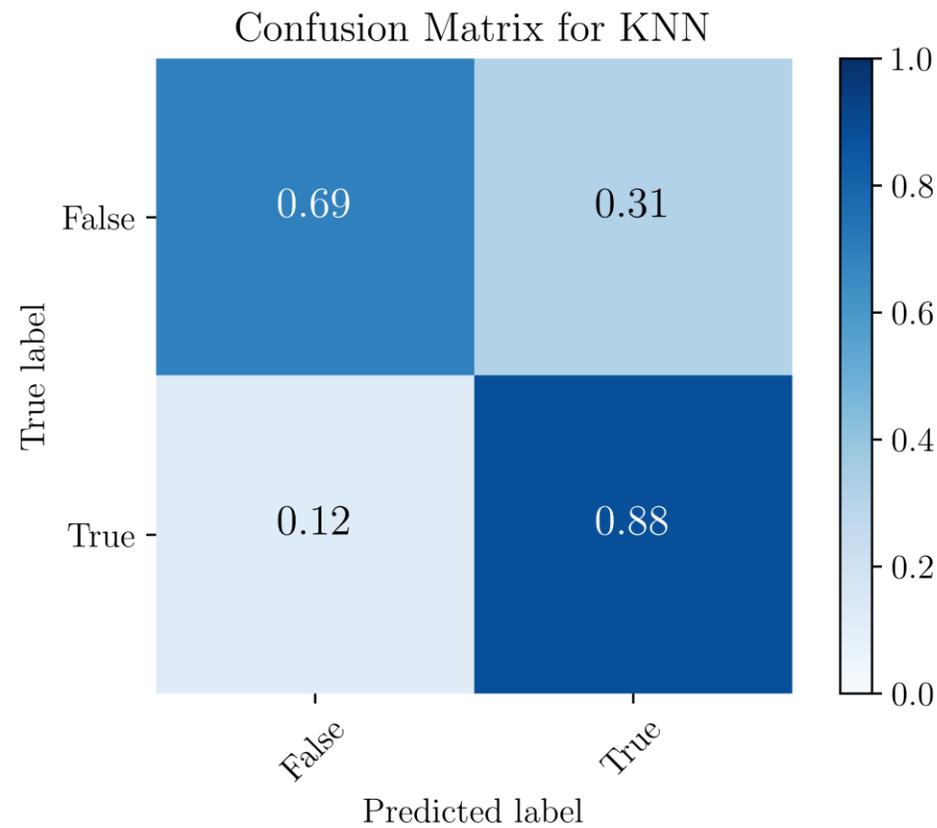
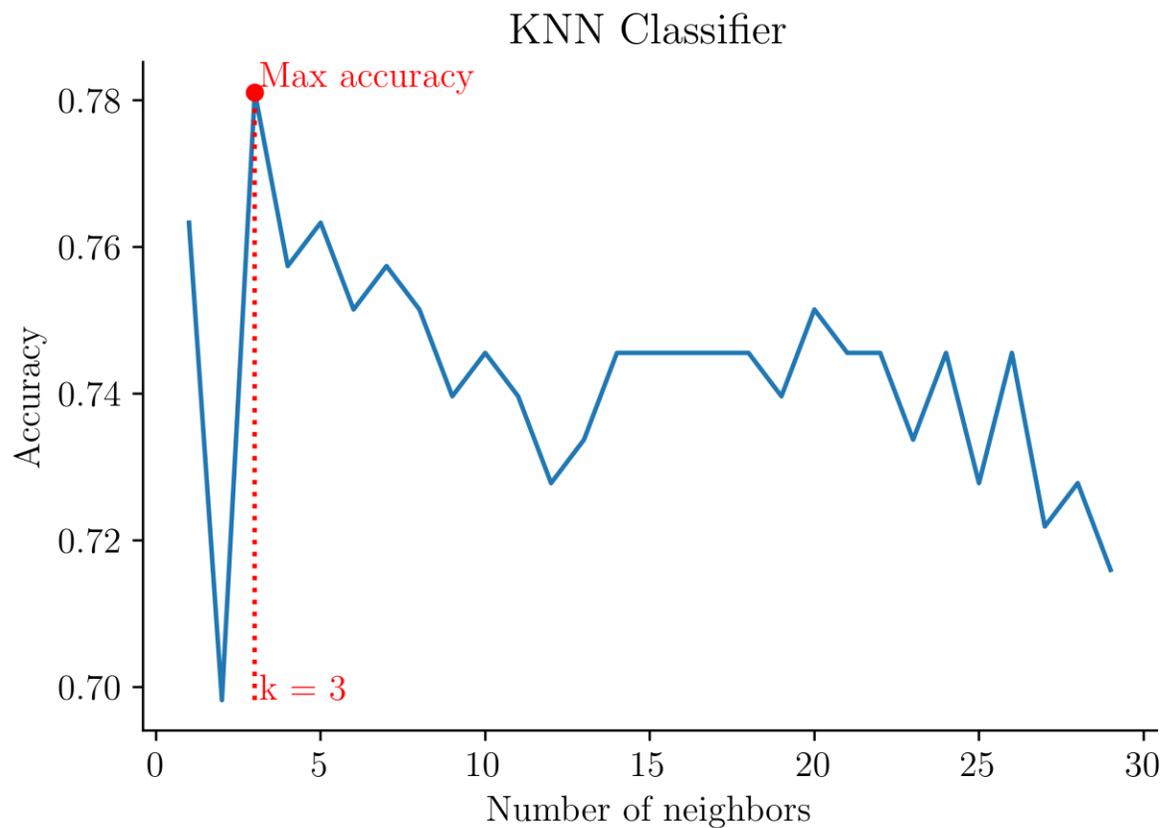
Forêt Aléatoire

Feature importances



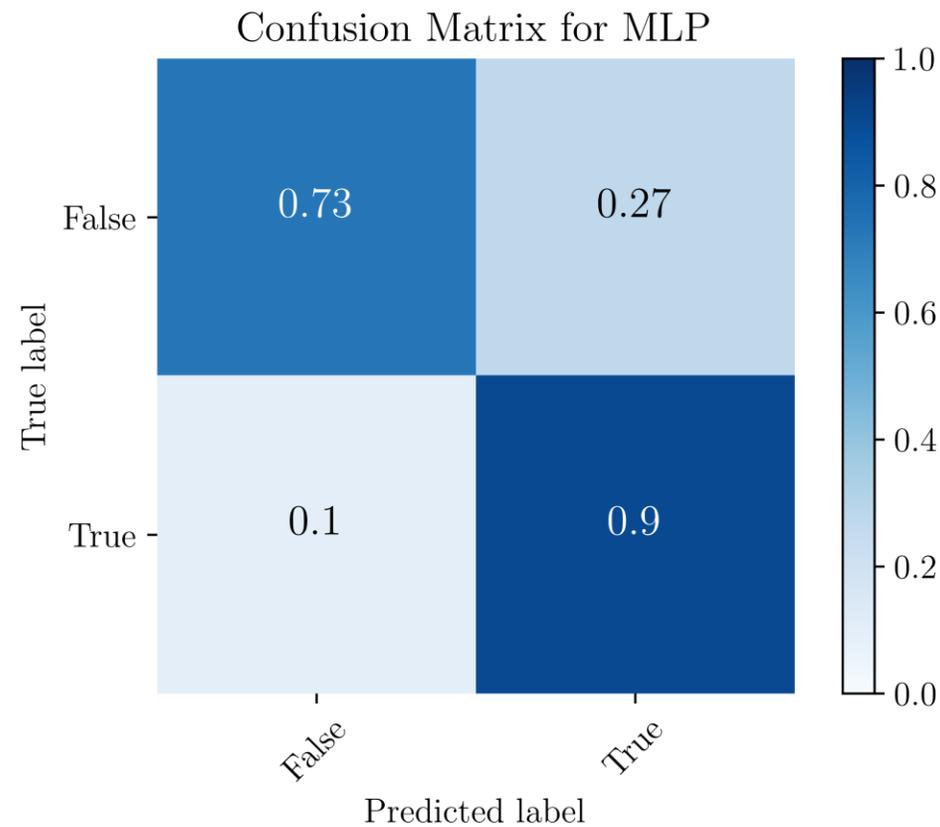
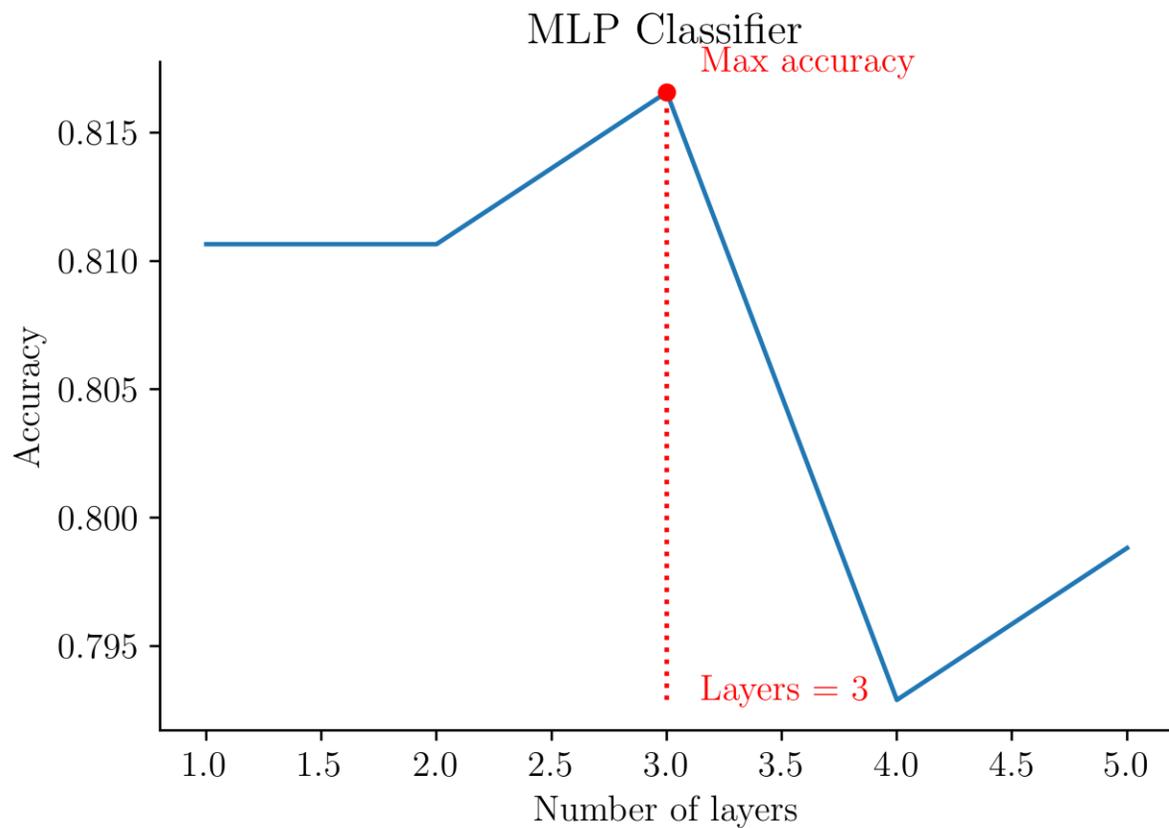
Plus proches voisins

K plus proche voisin

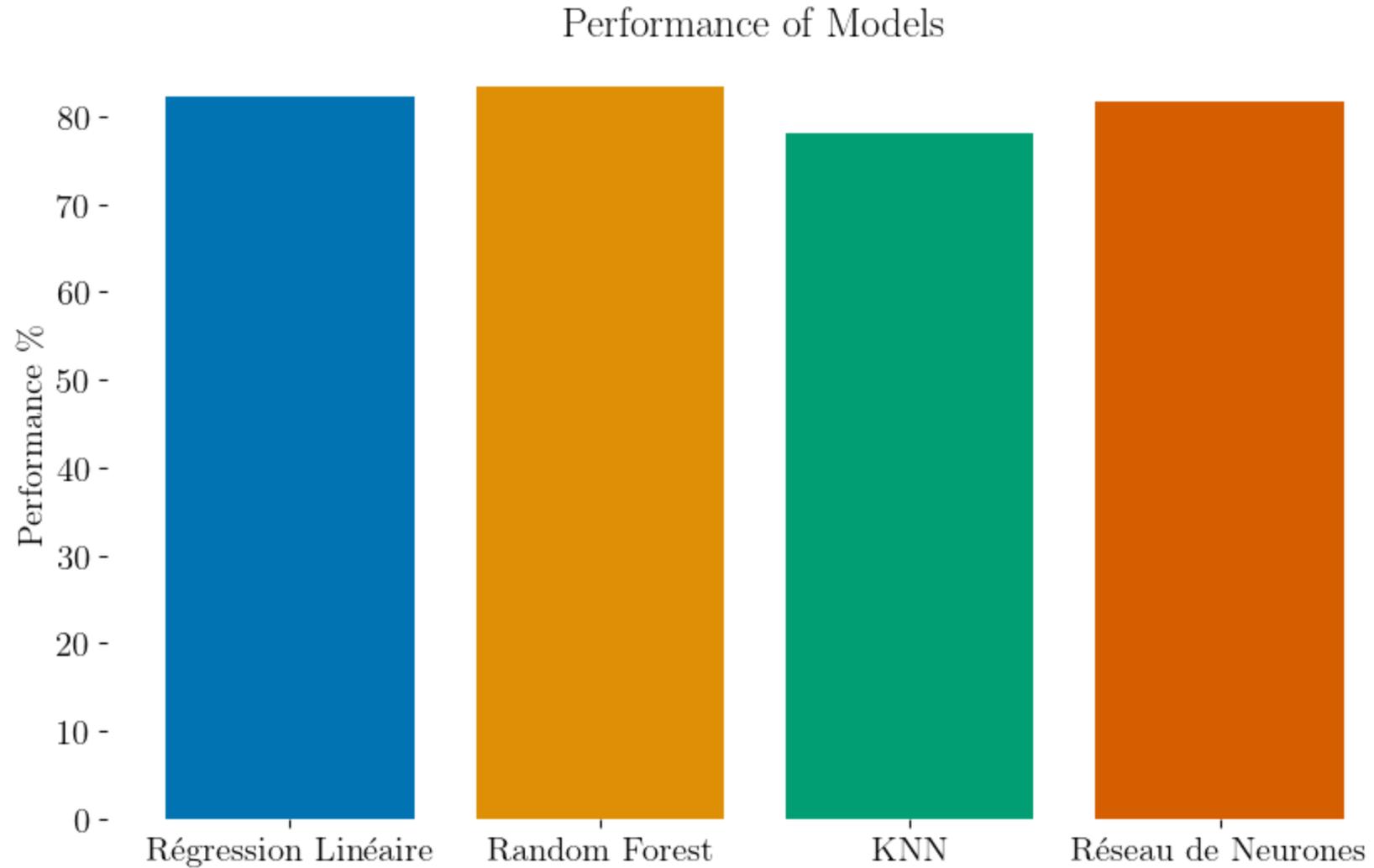


Réseau de Neurones

Réseau de Neurones

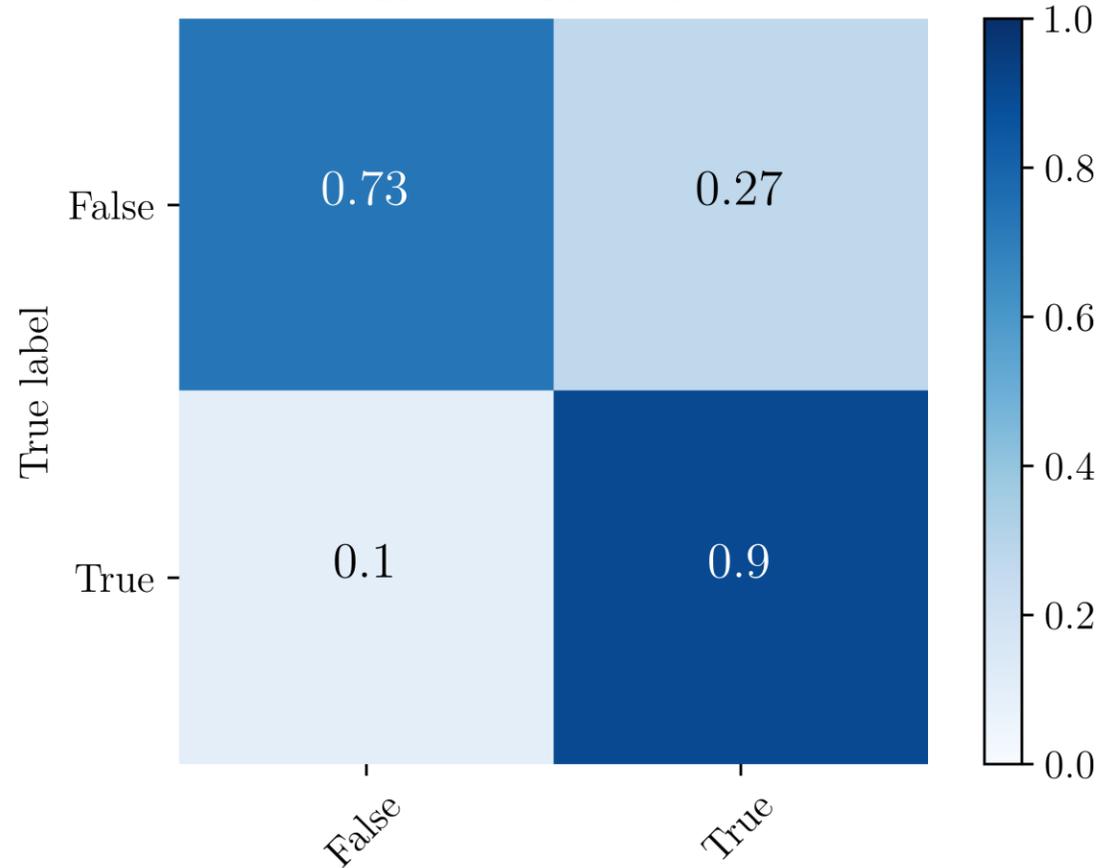


Résumé

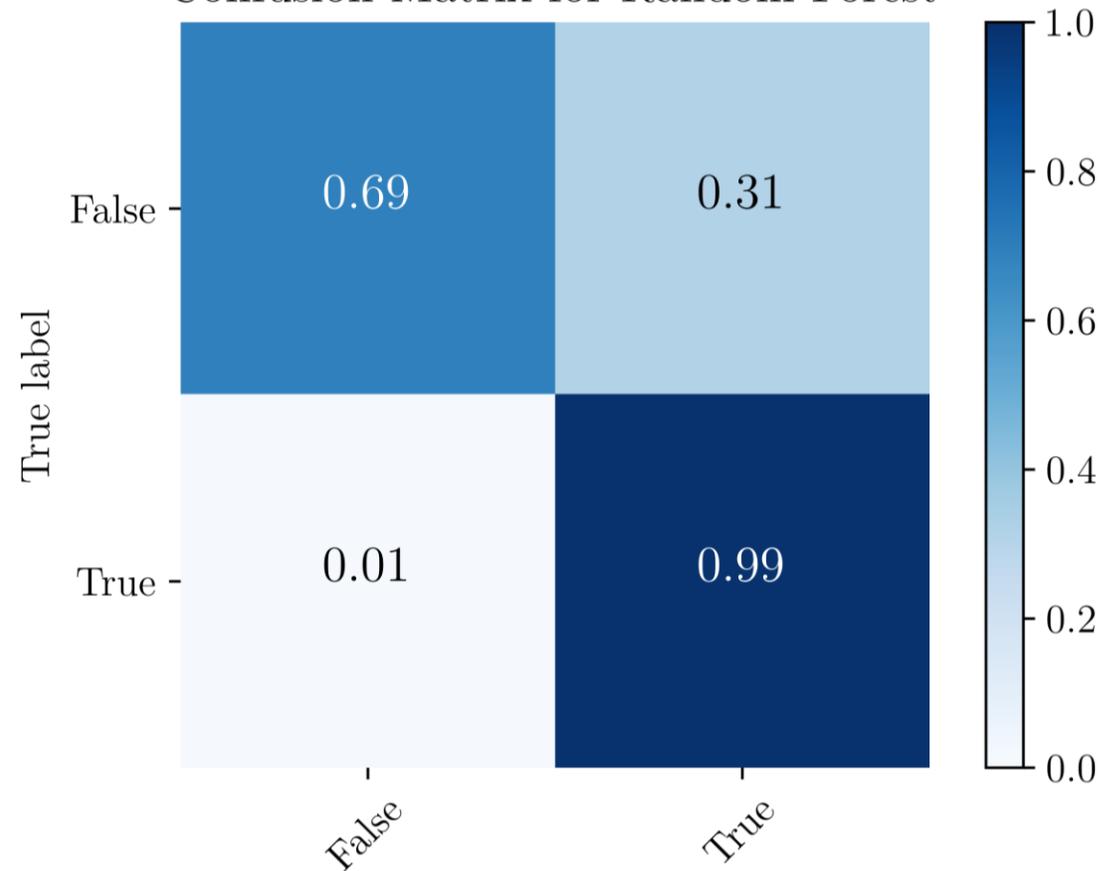


Résumé

Confusion Matrix for MLP

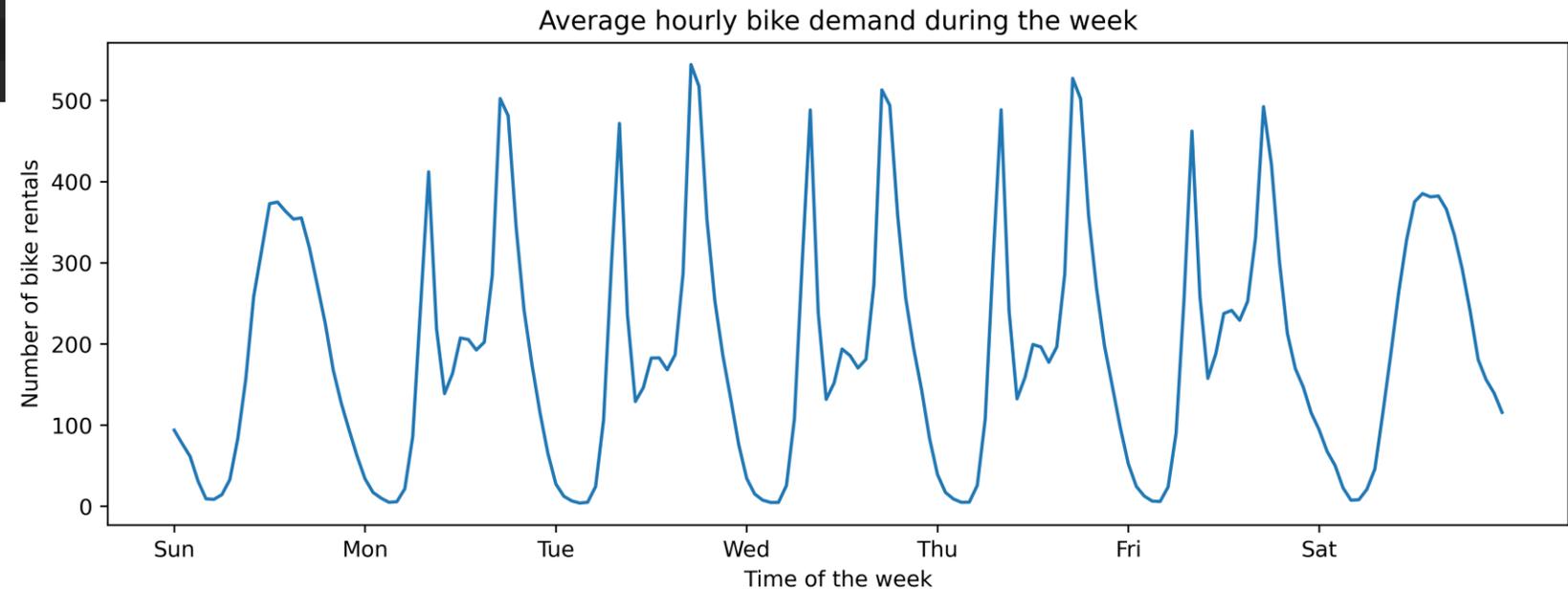


Confusion Matrix for Random Forest



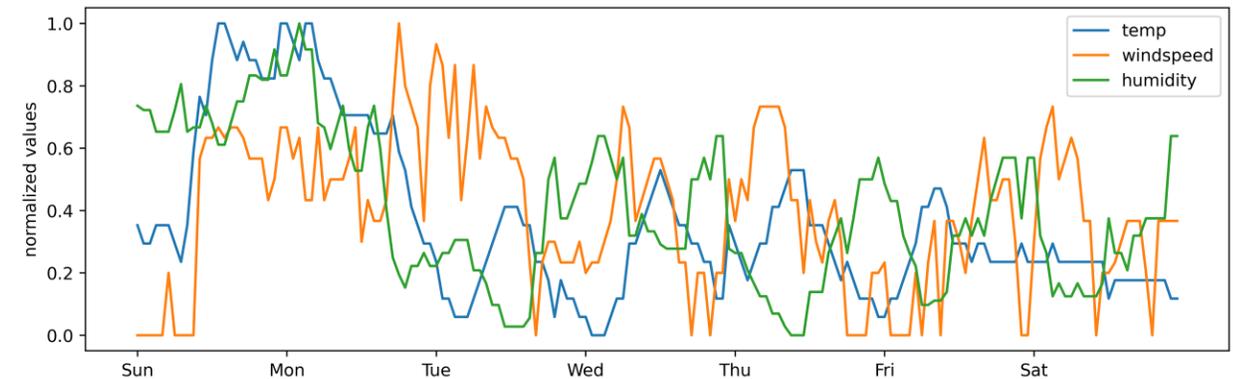
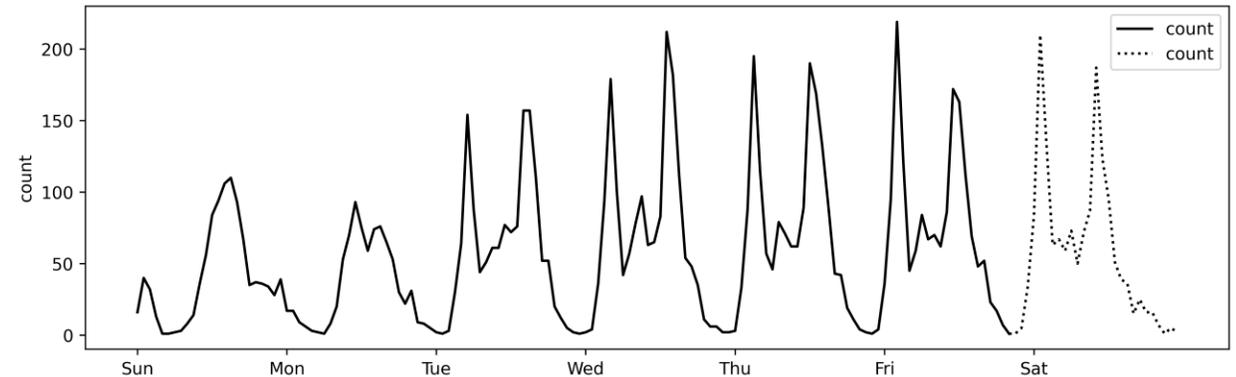
Prédiction d'utilisation de vélos en libre-service

	season	year	month	hour	holiday	weekday	workingday	weather	temp	feel_temp	humidity	windspeed	count
0	spring	0	1	0	False	6	False	clear	9.84	14.395	0.81	0.0000	16
1	spring	0	1	1	False	6	False	clear	9.02	13.635	0.80	0.0000	40
2	spring	0	1	2	False	6	False	clear	9.02	13.635	0.80	0.0000	32
3	spring	0	1	3	False	6	False	clear	9.84	14.395	0.75	0.0000	13
4	spring	0	1	4	False	6	False	clear	9.84	14.395	0.75	0.0000	1
5	spring	0	1	5	False	6	False	misty	9.84	12.880	0.75	6.0032	1
6	spring	0	1	6	False	6	False	clear	9.02	13.635	0.80	0.0000	2
7	spring	0	1	7	False	6	False	clear	9.84	14.395	0.75	0.0000	1
8	spring	0	1	8	False	6	False	clear	9.02	13.635	0.80	0.0000	2
9	spring	0	1	9	False	6	False	clear	9.84	14.395	0.75	0.0000	1



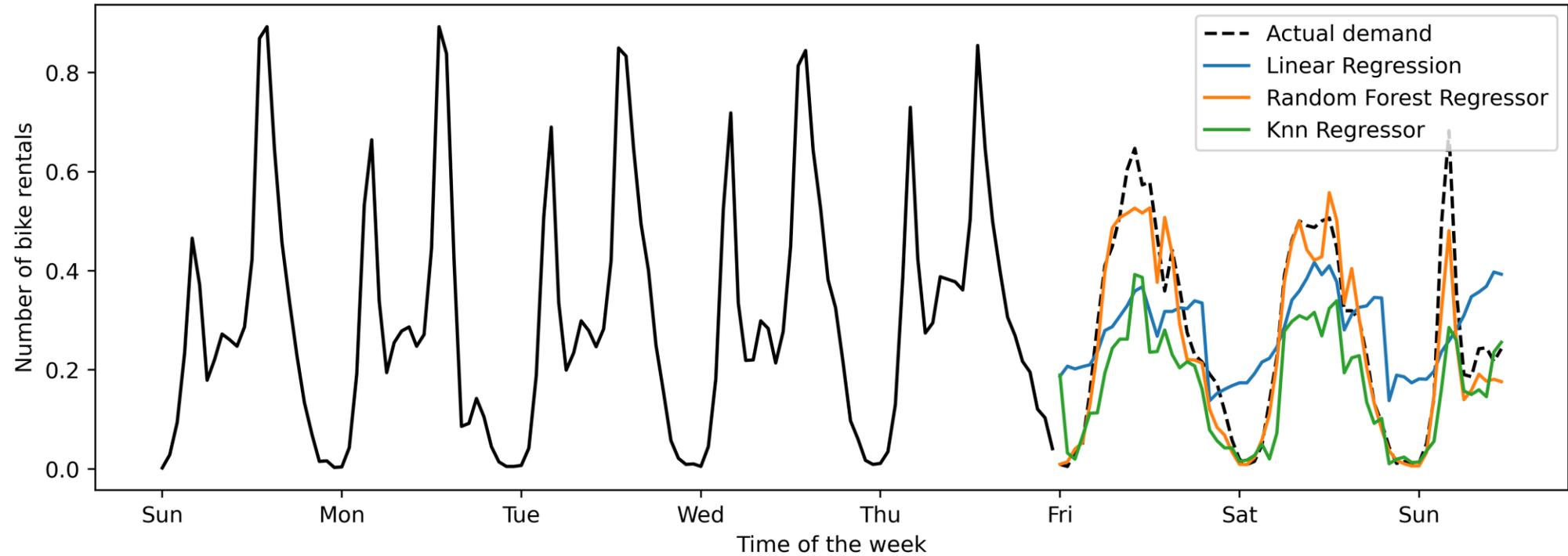
Prédiction d'utilisation de vélo en libre-service

- **Objectif** : Prédire l'utilisation de vélo sur les **prochains jours** en ayant accès aux futures données de température, vent, etc....



Prédiction d'utilisation de vélo en libre-service

Predictions by linear models



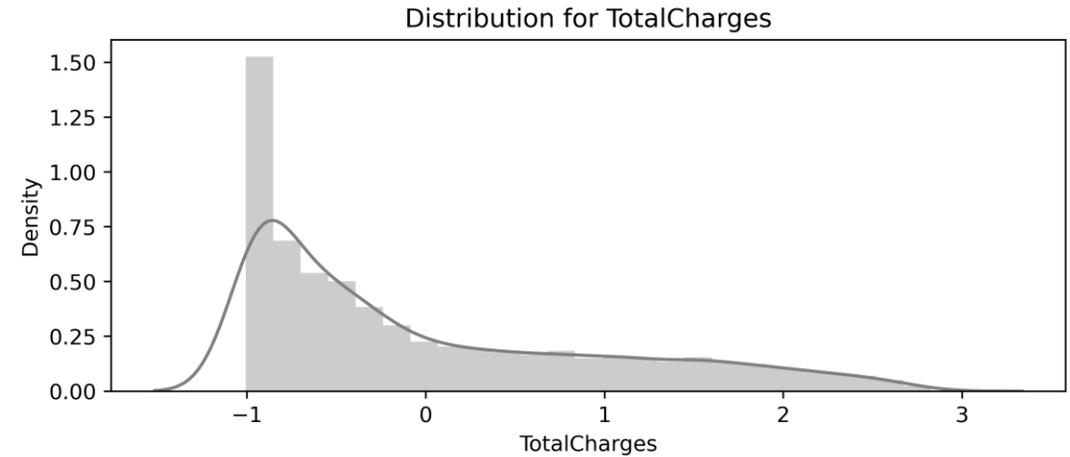
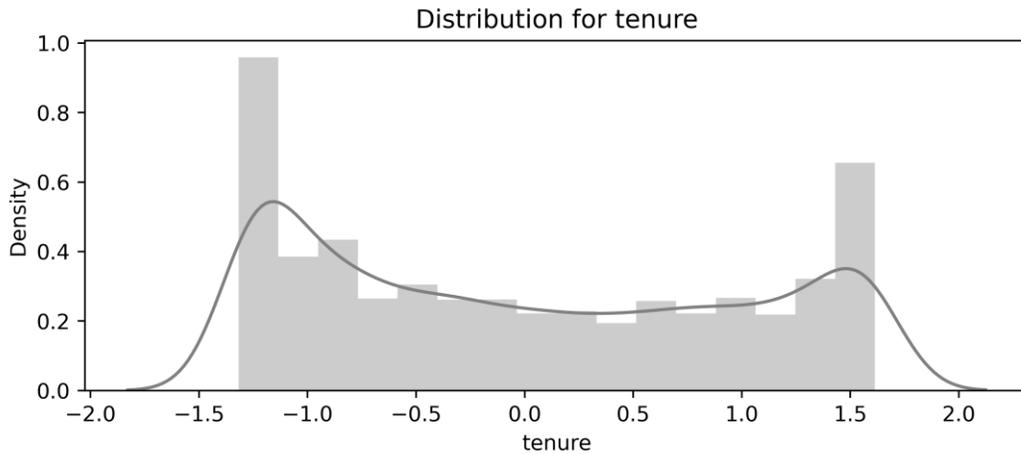
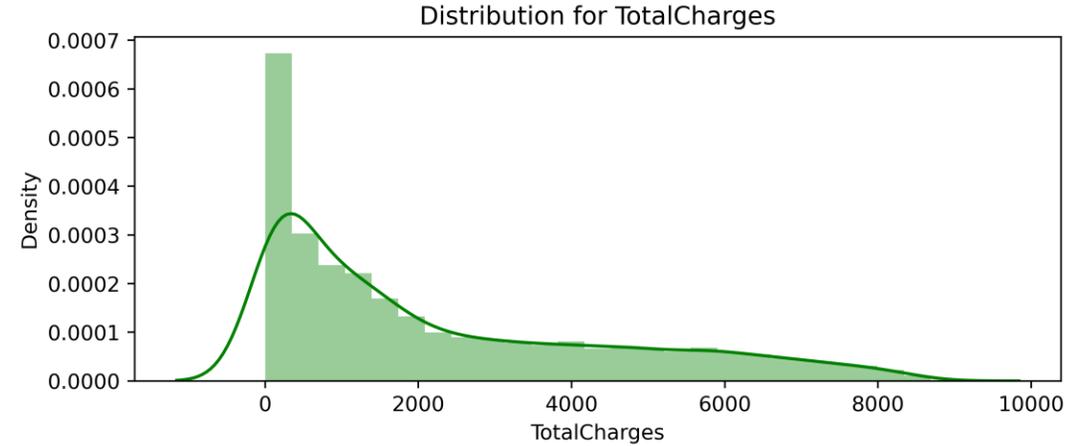
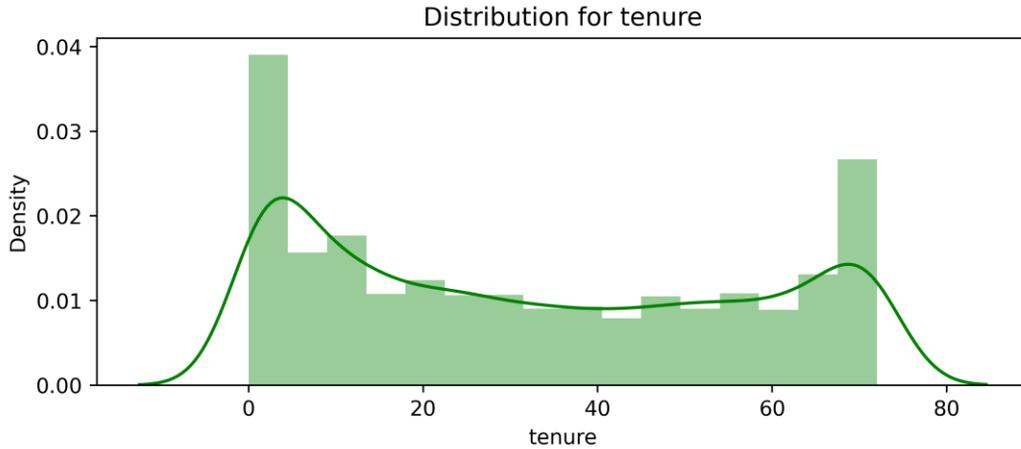
**Que faire en
Feature Engineering ?**

Présentation du problème

Jeu de données de clients d'un opérateur internet.

- customerID
- gender
- SeniorCitizen
- Partner
- Dependents
- tenure
- PhoneService
- MultipleLines
- InternetService
- OnlineSecurity
- OnlineBackup
- DeviceProtection
- TechSupport
- StreamingTV
- StreamingMovies
- Contract
- PaperlessBilling
- PaymentMethod
- MonthlyCharges
- TotalCharges
- **Churn**

Etape (presque) obligatoire : Mise à l'échelle des données



Approche Naïve : Prédiction de résiliation

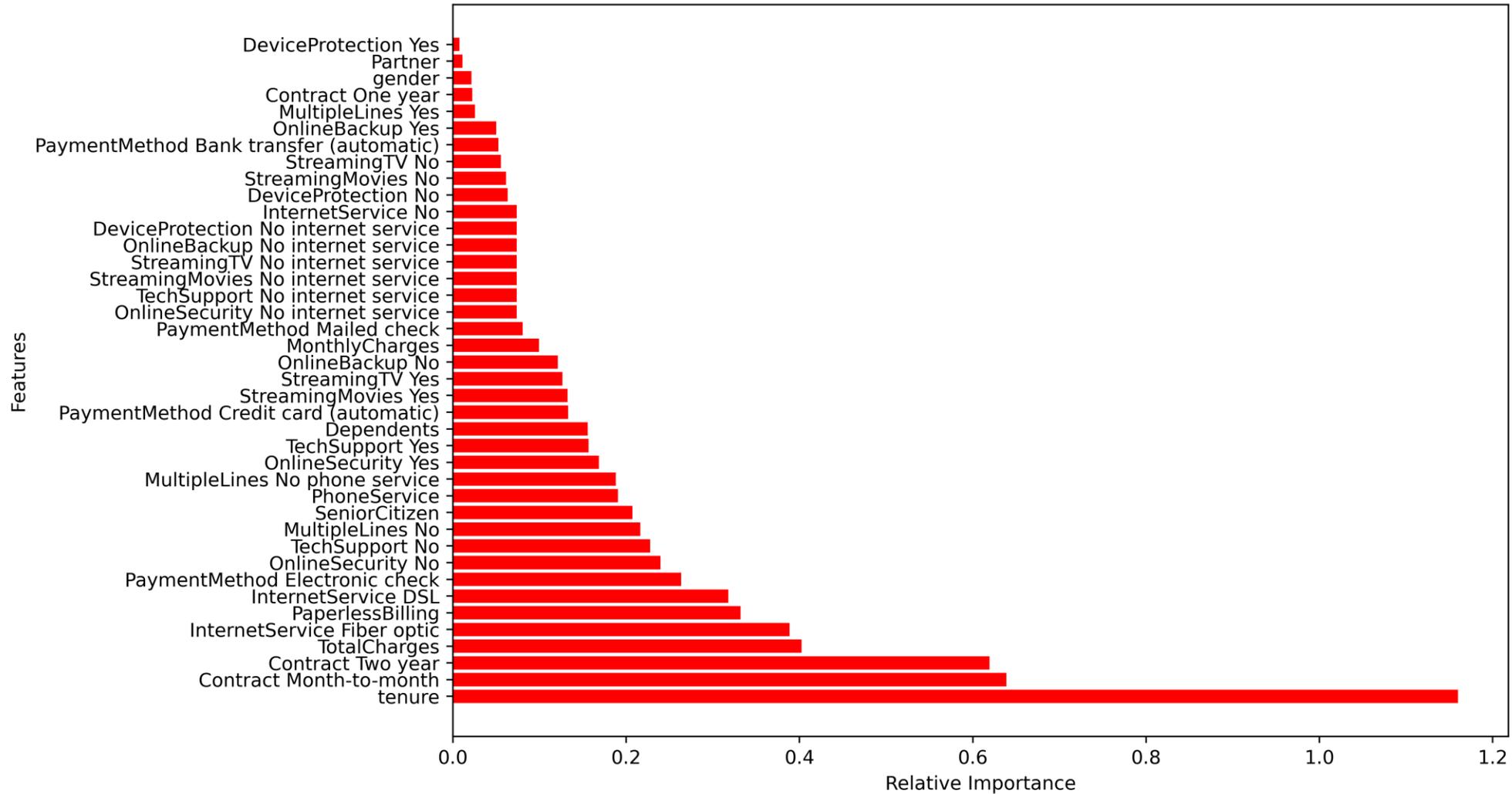
Modèle	Régression Linéaire
Performance	80.21%

Approche Naïve : Prédiction de résiliation

Modèle	Régression Linéaire	Knn	Forêt Aléatoire
Performance	80.21%	79.13%	79.60%

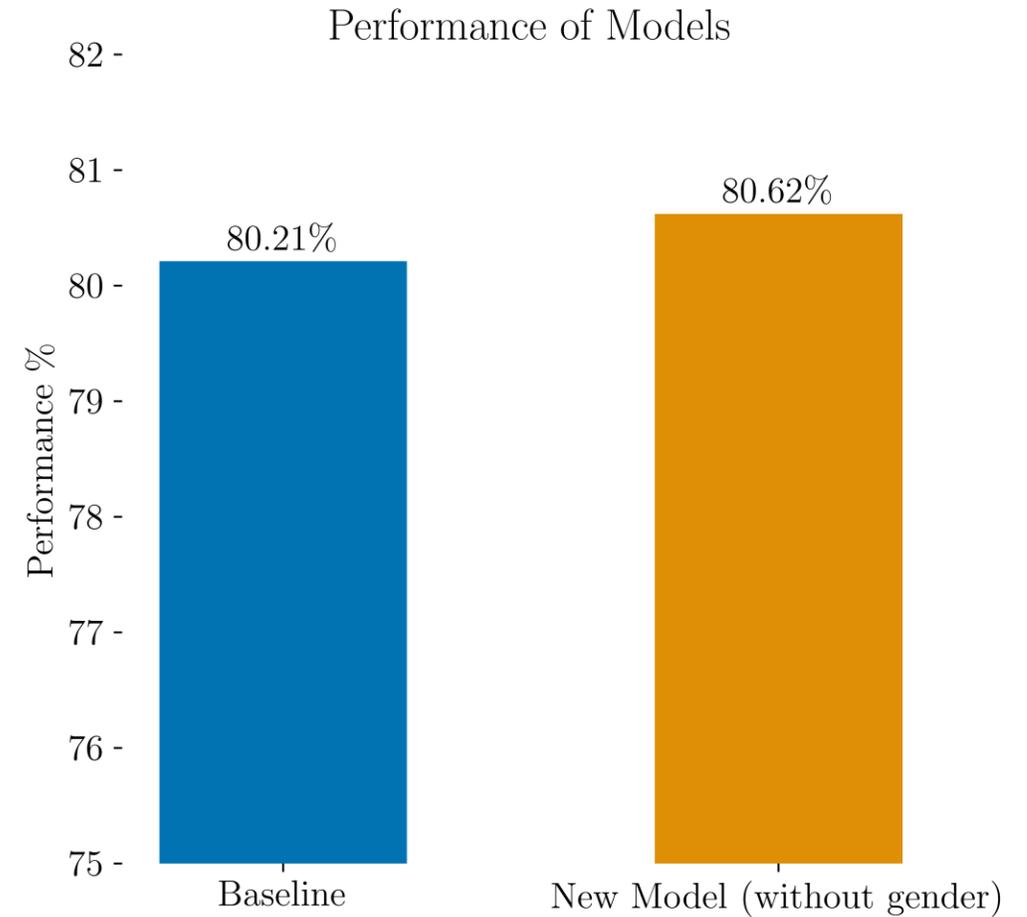
Importance des features pour le modèle linéaire

Feature importances for Logistic Classification



Feature Engineering : Enlever les attributs inutiles

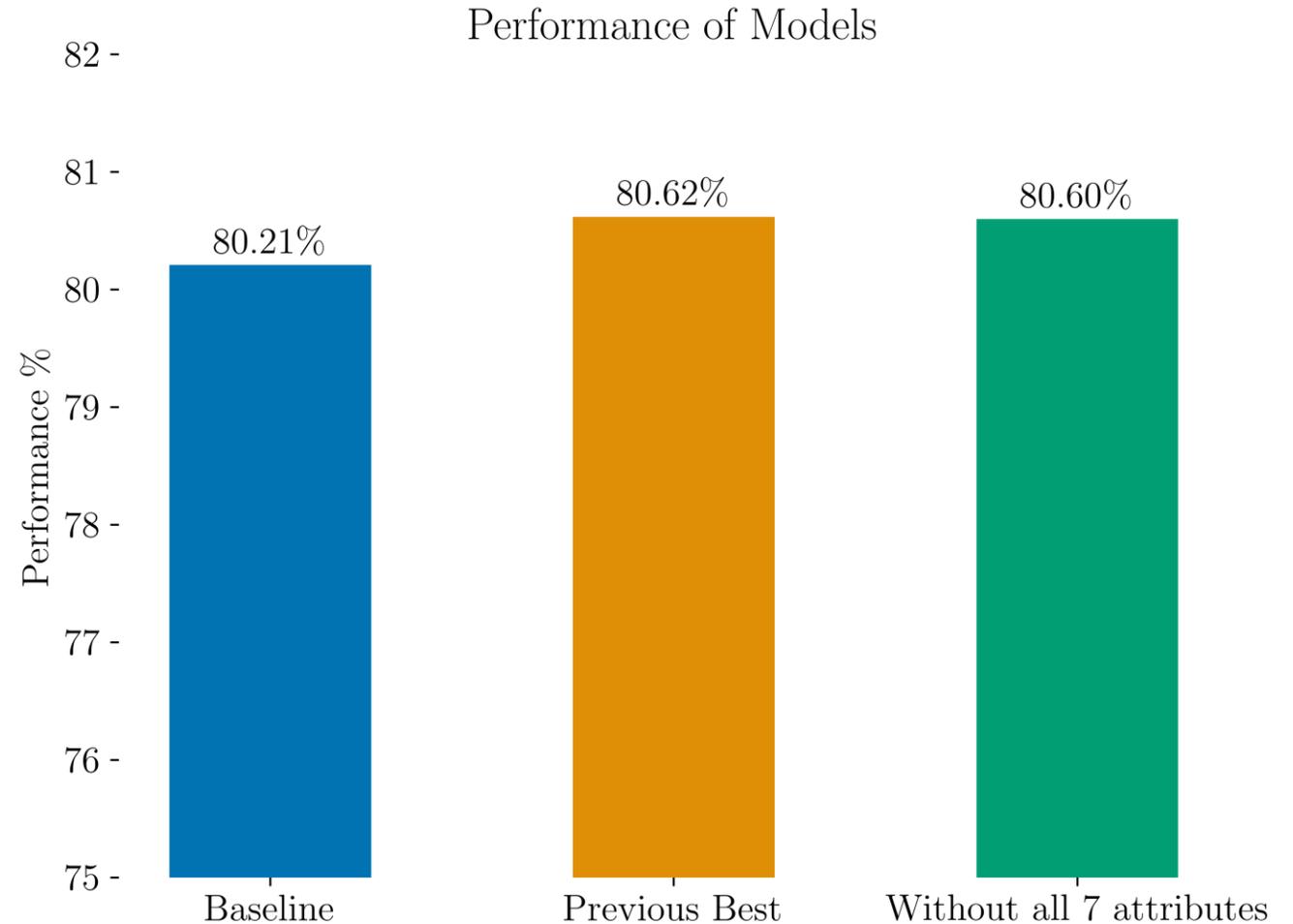
- Suppression de l'attribut **gender**.



Feature Engineering : Enlever les attributs redondants

Suppression des attributs :

- 'OnlineSecurity_No'
- 'OnlineBackup_No'
- 'DeviceProtection_No'
- 'TechSupport_No'
- 'StreamingTV_No'
- 'StreamingMovies_No'
- 'PhoneService'

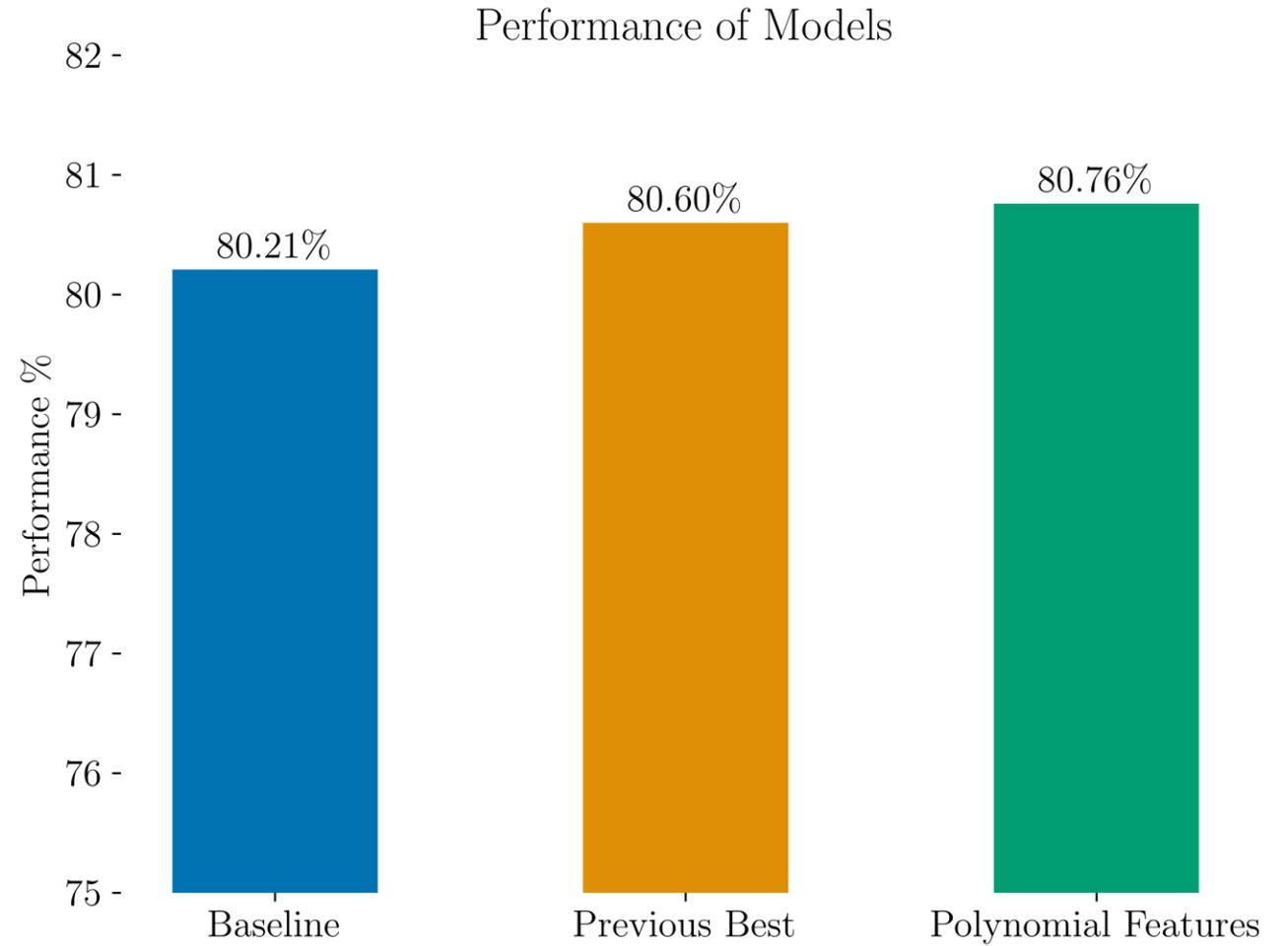


**Comment ajouter de
"nouvelles" features ?**

Feature Engineering : Ajouter des features automatiquement

Ajout de 9 features créées à partir de :

- 'tenure'
- 'MonthlyCharges'
- 'TotalCharges'



**Une autre manière
d'ajouter des features ?**

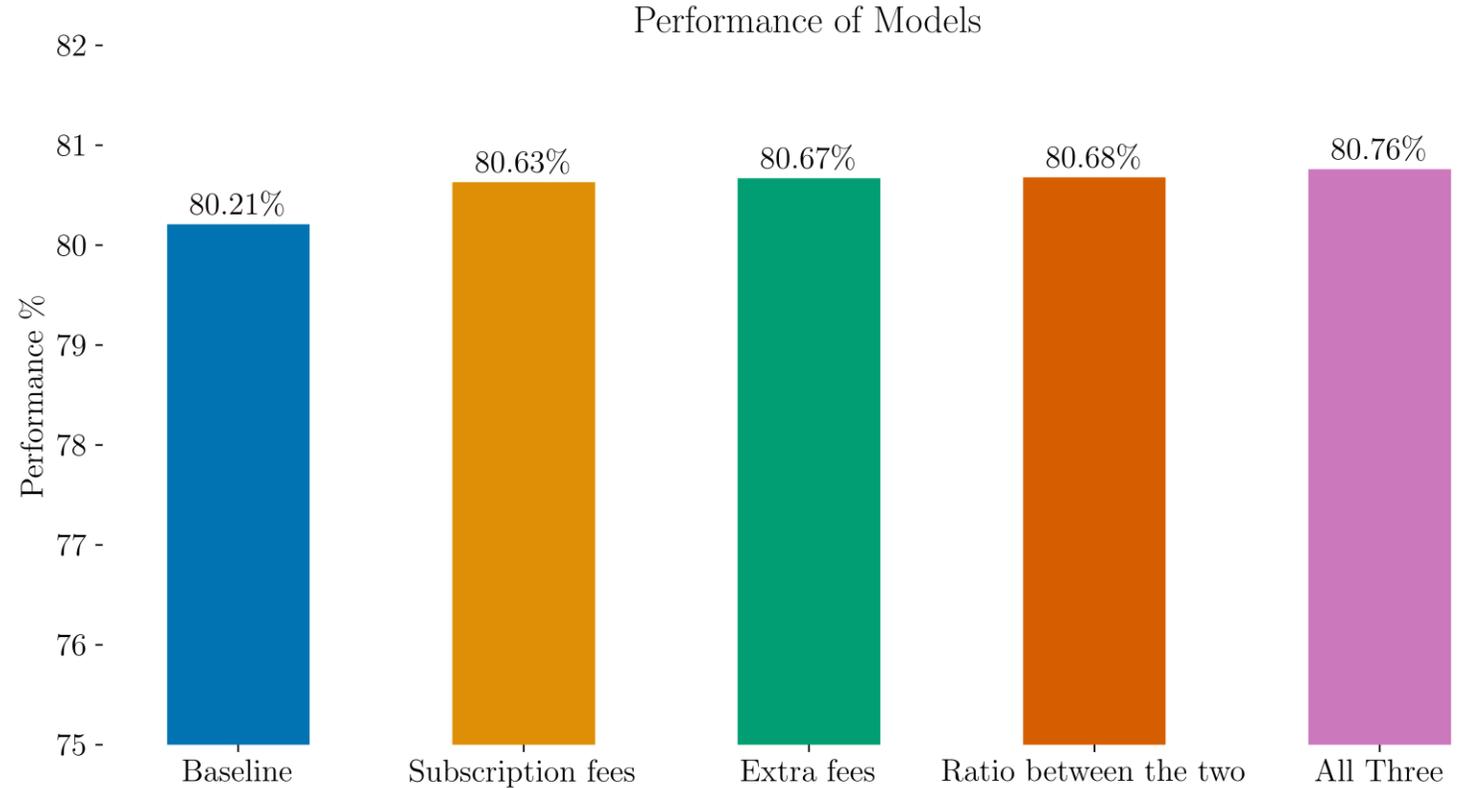
Feature Engineering : Ajouter des features "métier"

Ajout de :

$\text{tenure} \times \text{MonthlyCharges}$

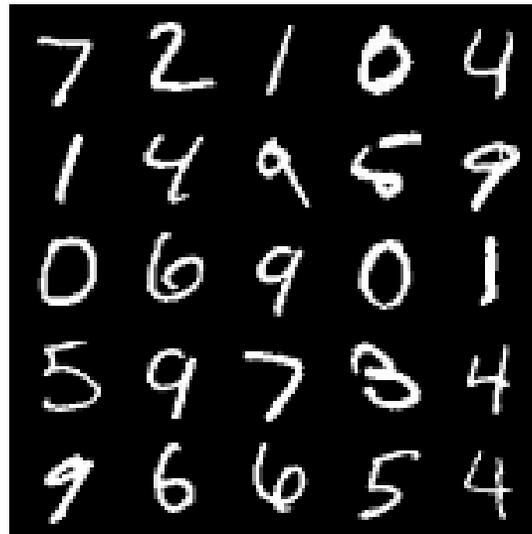
$\frac{\text{tenure} \times \text{MonthlyCharges}}{\text{TotalCharges}}$

$\text{TotalCharges} - \text{tenure} \times \text{MonthlyCharges}$



Classification d'images de chiffres

Classification d'images de chiffres : MNIST



Classification d'images de chiffres : MNIST

Predicted: 4



Predicted: 2



Predicted: 4



Predicted: 4



Predicted: 4



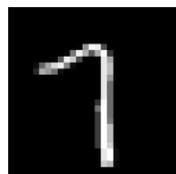
Predicted: 4



Predicted: 4



Predicted: 1



Predicted: 5



Predicted: 4



Predicted: 8



Predicted: 4



Predicted: 9



Predicted: 4



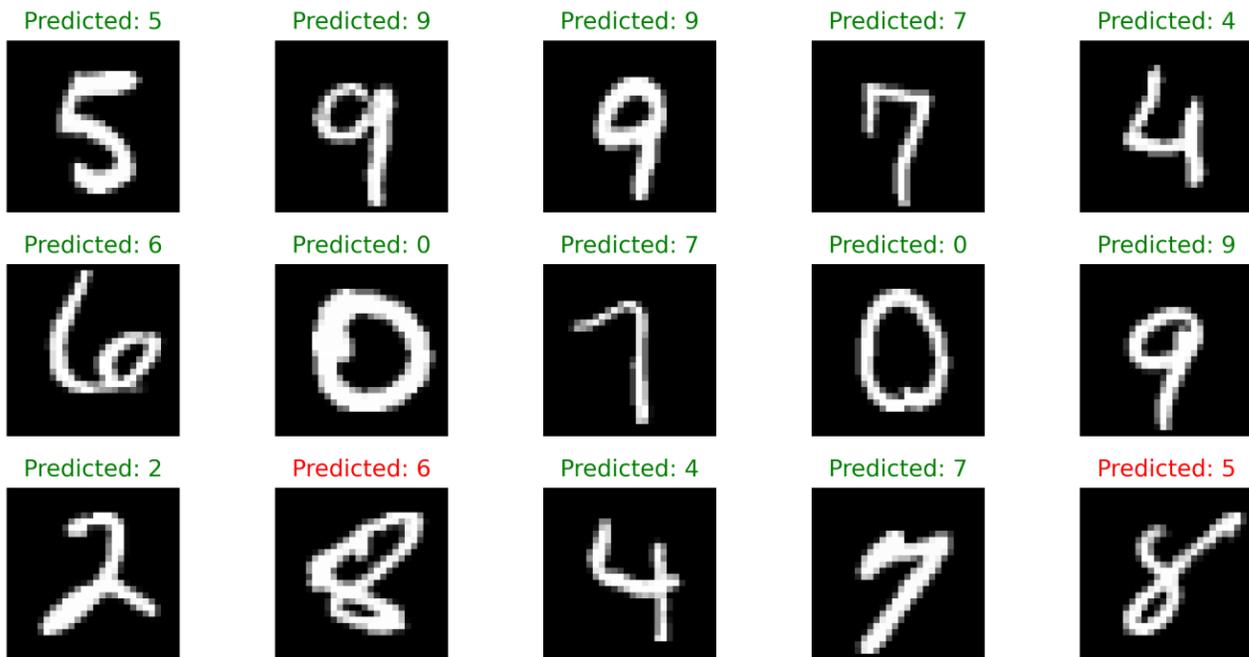
Predicted: 5



Modèle naïf : 10% de précision

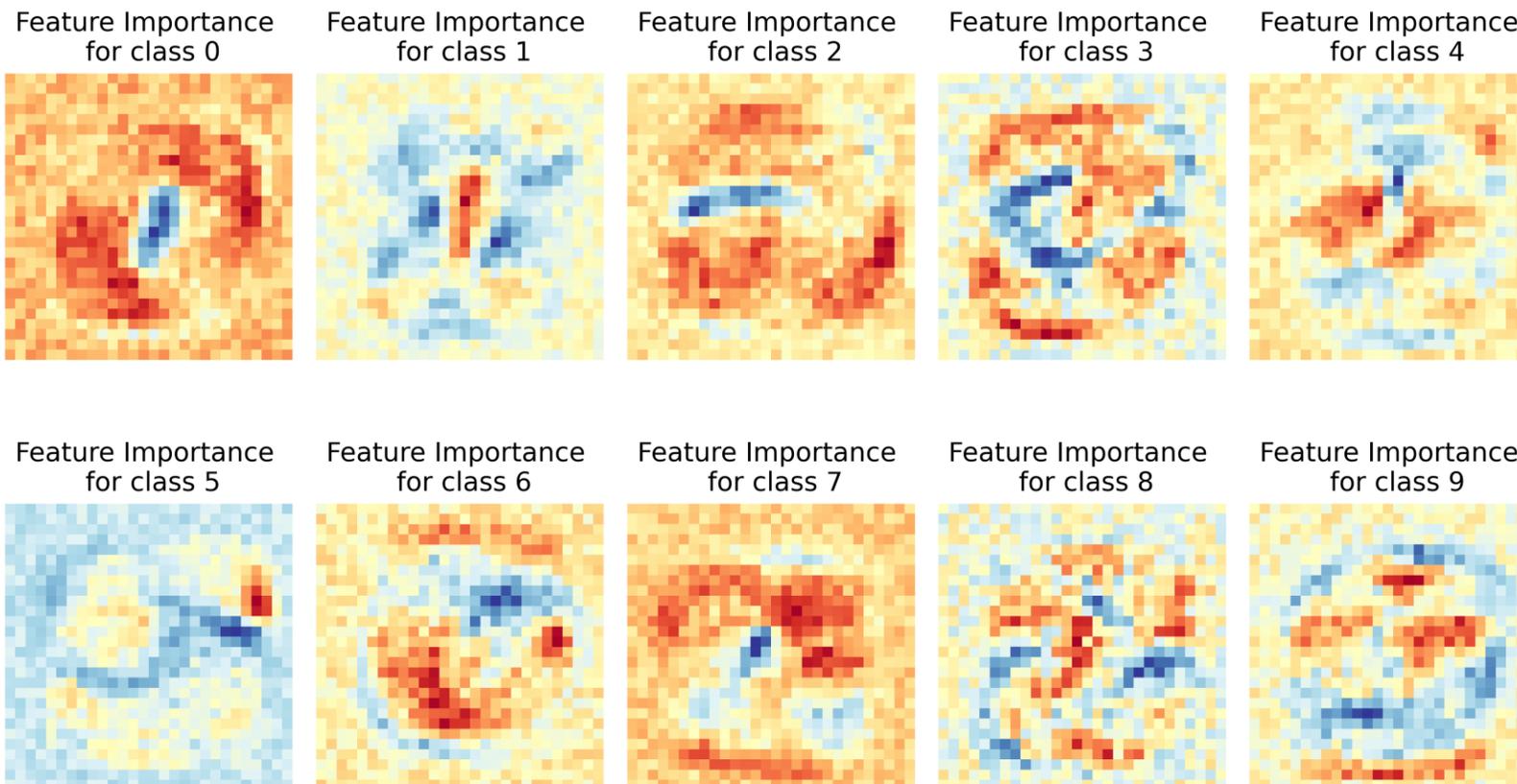
**Comment interpréter les
prédictions des
modèles pour des images ?**

Classification d'images de chiffres : MNIST



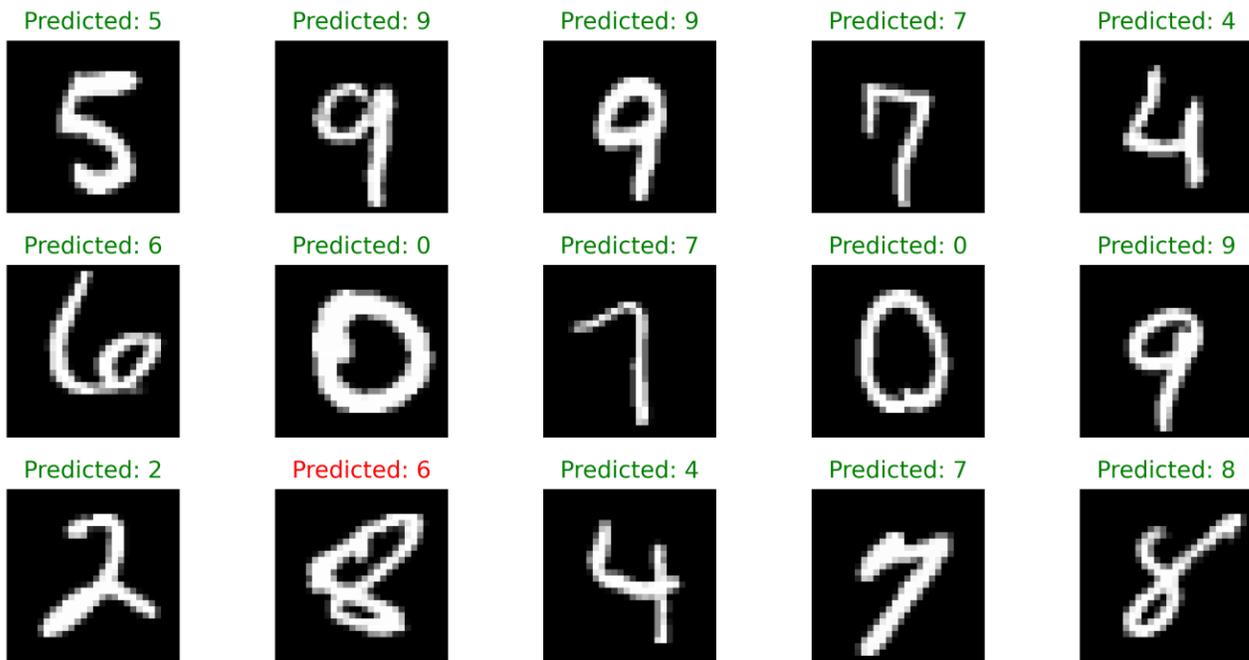
Régression Linéaire : 93.41%

Interprétation d'un modèle de régression linéaire



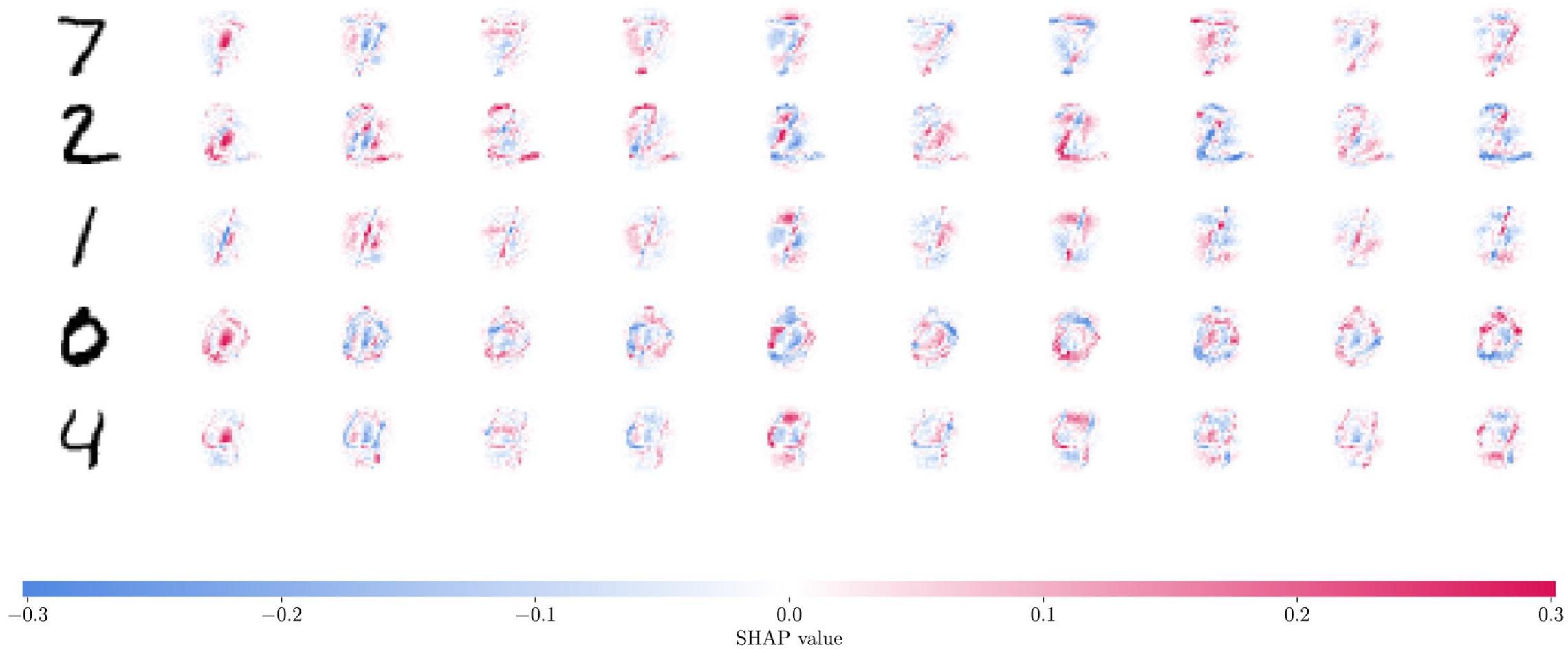
Modèle linéaires : interprétation peut se faire directement au niveau des paramètres

Classification d'images de chiffres : MNIST



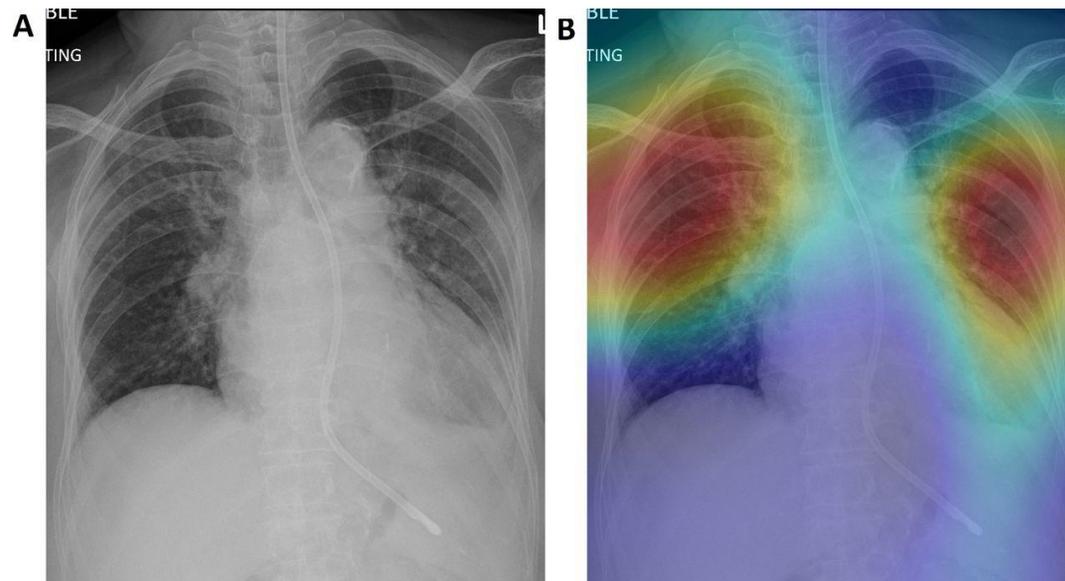
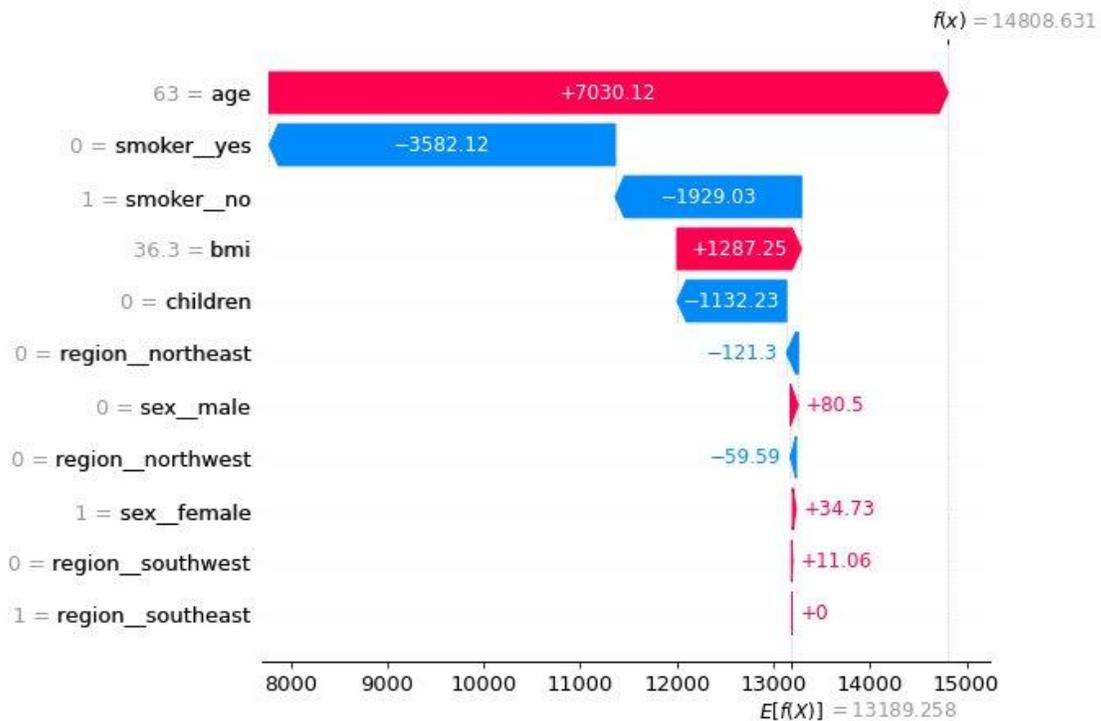
Réseau de neurones : 97.44%

Interpréter une décision



Classification d'images de chiffres

Interprétation d'une décision



Etudes de cas : résumé

- L'**évaluation** des modèles est une phase subtile : bien respecter les processus d'évaluation (**séparer les données** pour ne pas évaluer sur ce qui a permis d'entraîner le modèle)
- Bien définir ses **métriques** et être conscients de ce qu'elles mesurent et surtout de **ce qu'elles ne mesurent pas**
- Le **feature engineering** peut drastiquement booster les performances d'un modèle et bénéficiera toujours de retours d'experts
- L'interprétabilité des modèles a de la valeur, mais va dépendre :
 - Du modèle utilisé (plus difficile avec un réseau de neurone)
 - Des données (différences entre une image et des données tabulaires)

Prochaine séance :

22/03/2024





WWW.DAUPHINE.PSL.EU

22/03/2024

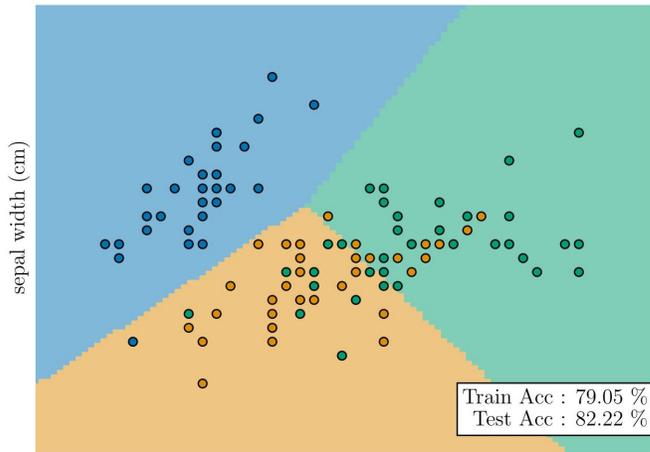
AI for texts



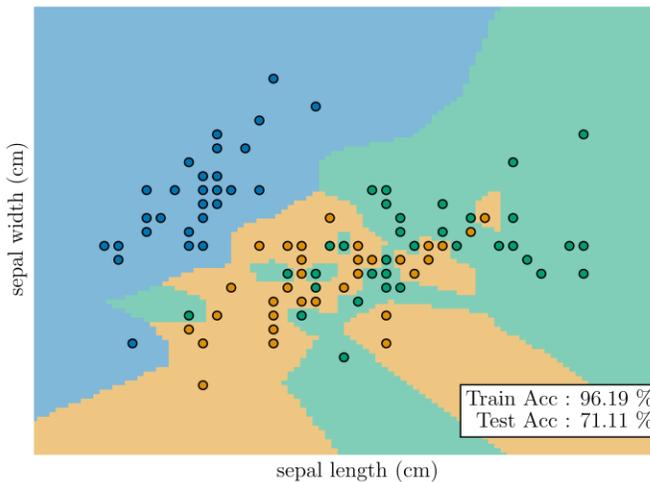
Rappels

Algorithmes de Machine Learning

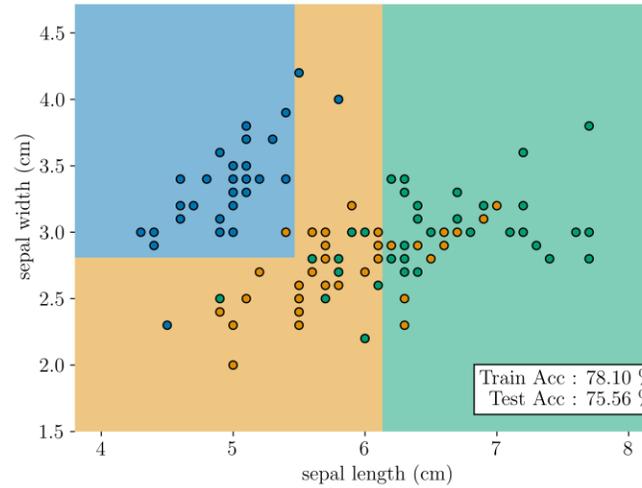
Classification linéaire



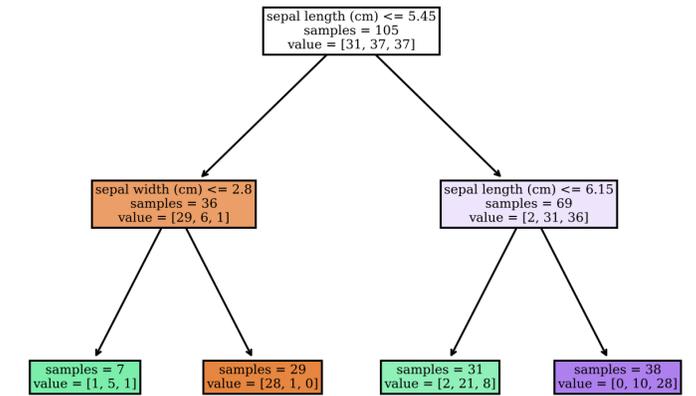
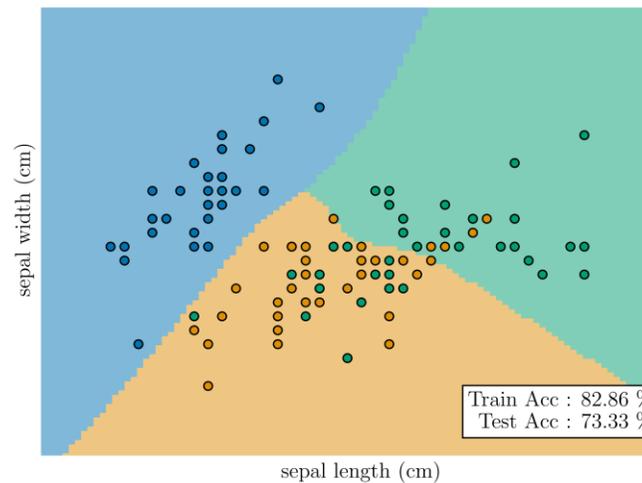
KNN 1



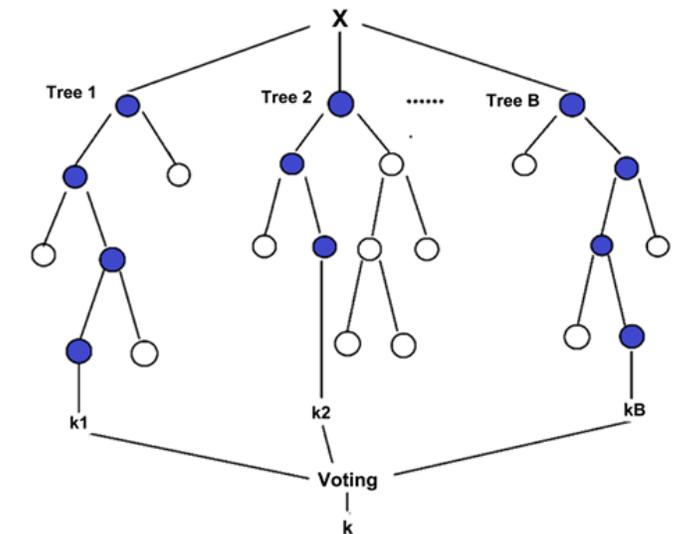
Decision Tree with depth 2



Neural Network Classification

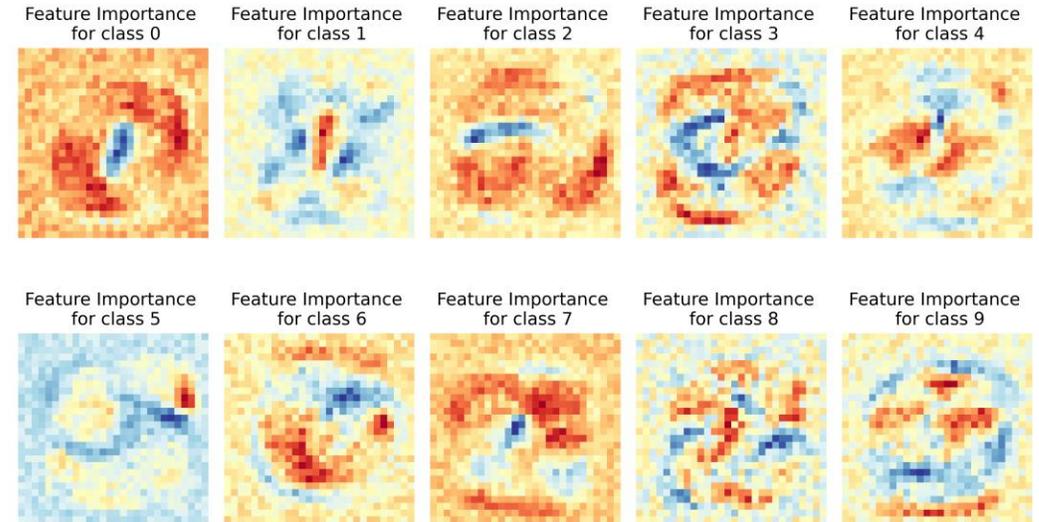
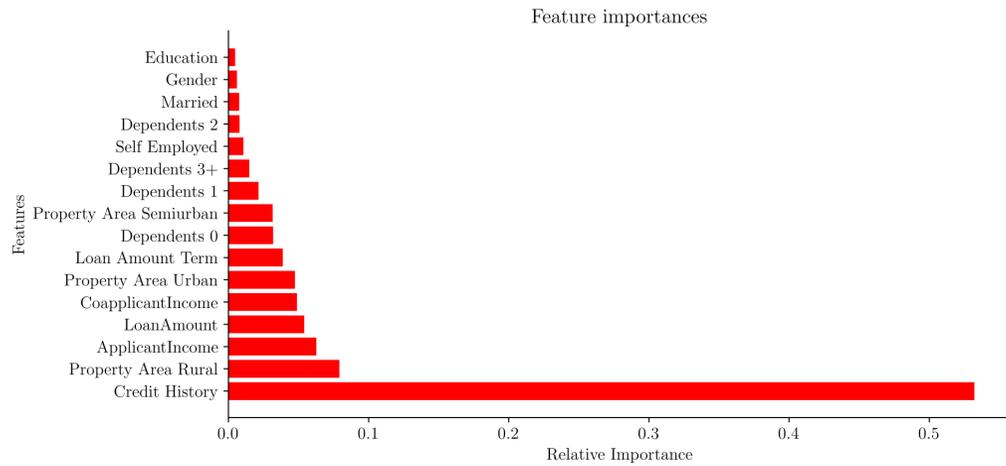


Random Forest



Rappels

Algorithmes de Machine Learning



Algorithmes de Machine Learning



Evaluation : séparation des données



Choix des **métriques** d'évaluation



Interprétabilité : sélection de features, explication des décisions

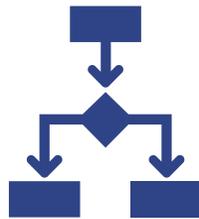
Aujourd'hui



Modèles linéaires



Plus proches voisins



Arbres de décision



Réseaux de neurones

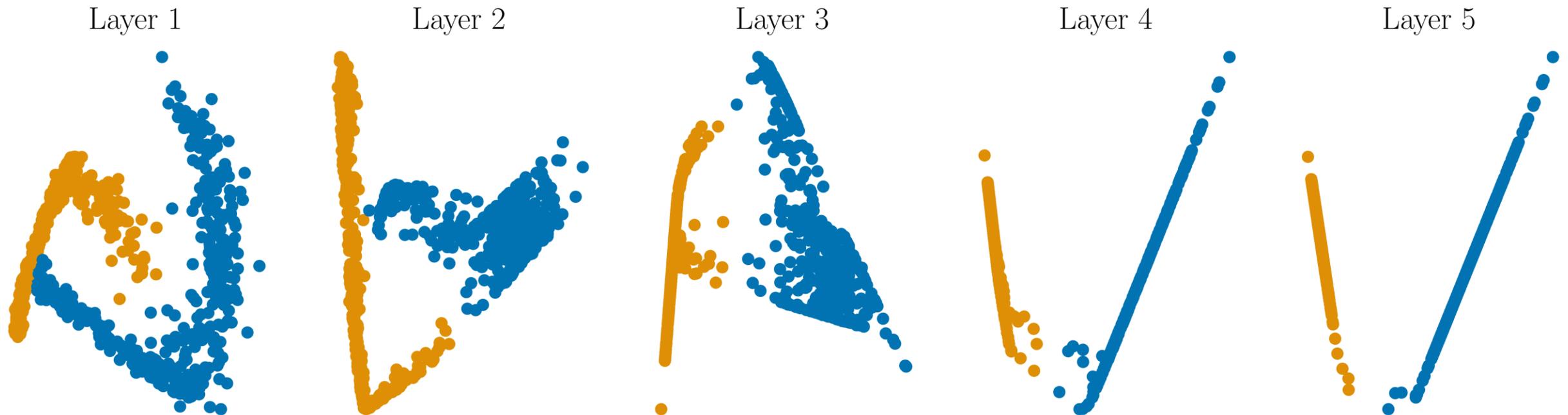
Illustration d'un réseau de neurones

- **Objectif** : être capable de séparer ces deux classes



Illustration d'un réseau de neurones

- Modification de l'espace de représentations des données par le réseau



Modèles discriminatifs vs génératifs



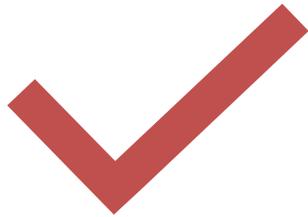
Modèles discriminatifs : associent des valeurs aux données d'entrée (étiquette catégorique, valeur continue)



Modèles génératifs : génèrent des données

**Quels usages pour des
modèles discriminatifs ?**

Modèles discriminatifs en TAL / NLP



Classification de textes :

- Prédire si un avis client est positif ou négatif
- Détection de spams
- Détection de faux avis
- Détection de thèmes



Classification de mots ou de portion de documents :

- Reconnaissance d'entité nommée
- Etiquetage grammatical

**Quels usages pour des
modèles génératifs ?**

Modèles génératifs en TAL / NLP



Traduction de texte



Réponse à des questions



Résumé de texte



Système conversationnel

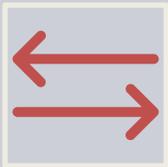
Depuis 2018 : évolution rapide des modèles de TAL / NLP



Nouvelles architectures de réseaux de neurones (transformers)



Augmentation de la taille des jeux données avec l'apprentissage non-supervisé



Développement de l'apprentissage par transfert

Modèles de TAL / NLP

- **Modèles discriminatifs** très utilisés historiquement
- Arrivée récente des **modèles génératifs** (premiers modèles pour la traduction vers 2014)
- **Evolution rapide** depuis 2018 (pour les modèles discriminatifs et génératifs)

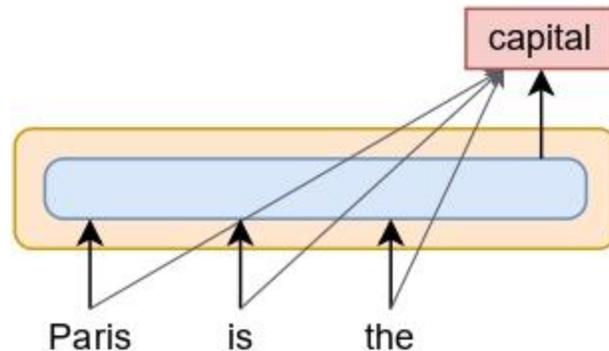
En général :

- **Modèles discriminatifs** : plus petits et plus simples à entraîner
Fonctionnement similaire aux algorithmes de ML étudiés la dernière fois
- **Modèles génératifs** : modèles de "grande taille", complexes à entraîner et évaluer mais capables de couvrir plus d'usages (au moins pour les plus récents)

Modèles génératifs type GPT : Principe de fonctionnement

Rappels des principes de fonctionnement des modèles génératifs

- Modèles **auto-regressifs**
- Génèrent **un mot après l'autre**
- Le mot généré à chaque étape **dépend uniquement de ce qui précède**
- Ce qui précède la réponse est appelé le **contexte**



Analyse de l'interface de ChatGPT

Interface ChatGPT

The screenshot displays the ChatGPT interface. On the left is a dark sidebar with navigation options: ChatGPT, DALL-E, Data Analyst, Web Browser, and See all. Below these are sections for 'Yesterday' (with a red box around the 'Feedback on short story' item), 'Previous 7 Days', and a list of recent prompts. At the bottom of the sidebar are 'Refer a friend' and 'Me' (Personal account). The main chat area is titled 'ChatGPT 3.5' and contains a user message (2) asking for feedback on a short story. The ChatGPT response (3) provides a detailed critique with three numbered points: 1. Punctuation and grammar, 2. Clarity, and 3. Show, don't tell. A red box highlights the input field (1) at the bottom of the chat area. A red box also highlights the entire chat content area. At the bottom of the main area, there is a disclaimer: 'ChatGPT can make mistakes. Consider checking important information.'

ChatGPT



Une interface d'entrée utilisateur



La conversation (avec les réponses du modèle)



Un historique des conversations

**De quoi est composé le
contexte ?**

Le contexte



Instruction utilisateur



Questions / réponses précédentes (de la même conversation)



Instruction système



Reformulations cachées

Prise en main

L'instruction système

- Lien d'accès à la démo sur le site de la formation : **alexverine** sur Google

L'instruction système

- Impose un **contexte global**, présent tout au long de la conversation
- Utilisée pour la **modération et la "safety"**
- **Pas toujours modifiable** par l'utilisateur (dans ce cas, il faut l'émuler au cours de la conversation)

Utilisation des modèles conversationnels : résumé

- La **contextualisation** de la conversation va déterminer la qualité de l'interaction
- **Plus le contexte contient d'information**, plus le modèle va s'aligner avec ce qu'on attend de lui
- Quelques interactions cachées à l'utilisateur
- Le contexte du modèle est constitué du **prompt**, de l'**instruction système** et de l'**historique de la conversation**.

Quelques conseils

 Séparer la **requête** en sous-tâches plus simples

 En cas de **refus** du modèle, lui dire que c'est un jeu, de faire "comme si"

 Attention aux problèmes de **logiques / de maths**

 Compromis **temps d'écriture de l'instruction** / qualité de la réponse

 Se rapprocher de ce sur quoi le modèle a été entraîné (forum internet, Wikipedia, etc)

 Ecrire en **anglais**

 Se renseigner sur les **dates d'entraînement** (demander au modèle)

Grandes organisations

- **OpenAI** : GPT3, GPT4, **ChatGPT**
- **Microsoft** : **Copilot** (intégré dans le navigateur)
- **Anthropic** : **Claude**
- **Meta** : Llama
- **Google** : Bard, **Gemini**
- **Mistral** : Mistral, Mixtral, **Le Chat**
- **xAI** : Grok
- **HuggingFace** : BLOOM (partenariat avec le CNRS)
- (LLM Leaderboard)

Hiérarchie des modèles

- **Modèles pré-entraînés : modèles généraux** entraînés sur de gros jeux de données très divers
 - Développement uniquement par les plus **grosses organisations**
 - Mise à disposition optionnelle de leurs modèles à la communauté (attention aux licenses)
- **Modèles spécialisés : modèles répondant à une application spécifique**
 - Développement interne avec des ressources "standards" et des données internes
 - Peuvent se baser sur des modèles pré-entraînés
 - Modèles de recherche souvent mis en "open-source" mais les applications industrielles sont propriétaires

**Comment utiliser ces
modèles dans les
organisations ?**

Un mot sur l'open-source

- Des laboratoires, des entreprises entraînent des modèles
- Ils les mettent "**librement**" à **disposition** de la communauté (= n'importe qui peut venir les télécharger)
- Fonctionnement similaire pour certains jeux de données
- **HuggingFace** est l'acteur majeur de cette mise à disposition des modèles

Avantages / inconvénient

Avantages :

- Gratuit
- Beaucoup de contrôle sur l'utilisation des modèles
- Développement interne

Inconvénients :

- Compétences techniques en IA
- Installations pour le développement des algorithmes de ML

Les APIs

- Certains modèles ne sont accessibles qu'à travers des APIs (ou des interfaces webs) : GPT3, GPT4, Mistral Large.

Avantages :

- Compétences de développement web
- Pas d'installations techniques

Inconvénients :

- Moins de contrôle
- Confidentialité des données
- Dépendance par rapport au fournisseur

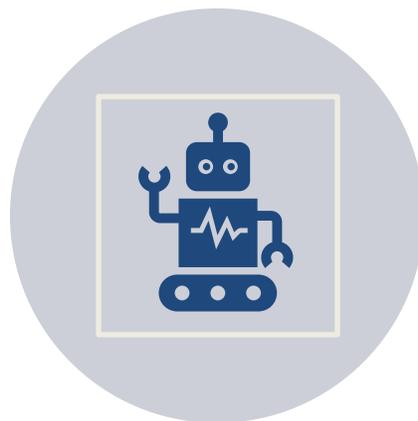
Utilisation des modèles à grande échelle : résumé

- L'open-source peut être un allié puissant
- Mais demande plus de compétences techniques : à privilégier pour les petits modèles
- Utilisation des plus gros modèles par des APIs propriétaires

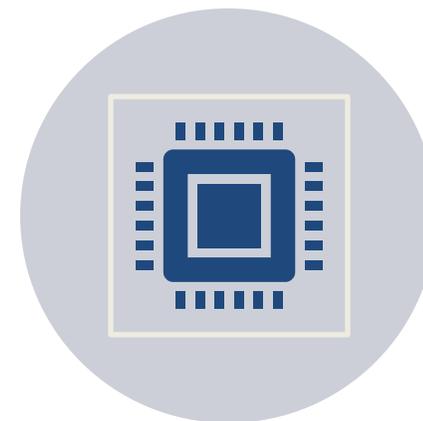
Que sont ces autres outils?



DATA SCIENTIST



DATA ENGINEER



SOFTWARE ENGINEER

Research Scientist

Fonctionnement des autres outils

- Propose des applications très spécifiques
- Des produits dérivés de modèles généralistes pré-entraînés

Listes d'autres outils

- Explosion du nombre d'outils d'IA : <https://theresanaiforthat.com/> (attention, liste non filtrée)

Code

- **ChatGPT**
- **Copilot** : Assistant dynamique de code
- Ecriture de documentation automatique : <https://stenography.dev/>

Transcription de textes

- "Hacky" : <https://whisper.ggerganov.com/>
- Moins cher / compliqué : télécharger sur youtube et utiliser la transcription automatique
- [Speechmatics](#) (multilingues)
- [Otter.ai](#) (meeting assistant)

Autres

- [Otter.ai](#) : Assistant de réunion. Transcription, résumé, follow-up actions etc.
- [Notion ai](#) : Prise de notes organisées et structurée (disponible gratuitement sans AI, déjà très intéressant).
- [Guidde](#) : Tutoriel étape par étape automatique (enregistrement de certaines actions à l'écran, puis création d'une page de tutoriel avec les principales étapes).

Résumé sur les autres outils

- Basés sur des modèles généralistes
- Leur valeur vient d'une forte automatisation et de l'utilisation de données spécialisées
- Grand nombre d'outils : définir le besoin avant l'outil
- Chercher des signaux de fiabilité (essais gratuits)
- S'assurer que nos outils ne permettent pas déjà de le faire (Suite Office, Google, CRM Salesforce, etc)

Prochaine séance :

29/03/2024





WWW.DAUPHINE.PSL.EU

29/03/2024

Modèles Génératifs pour l'image



Au programme du jour



Historique des modèles génératifs



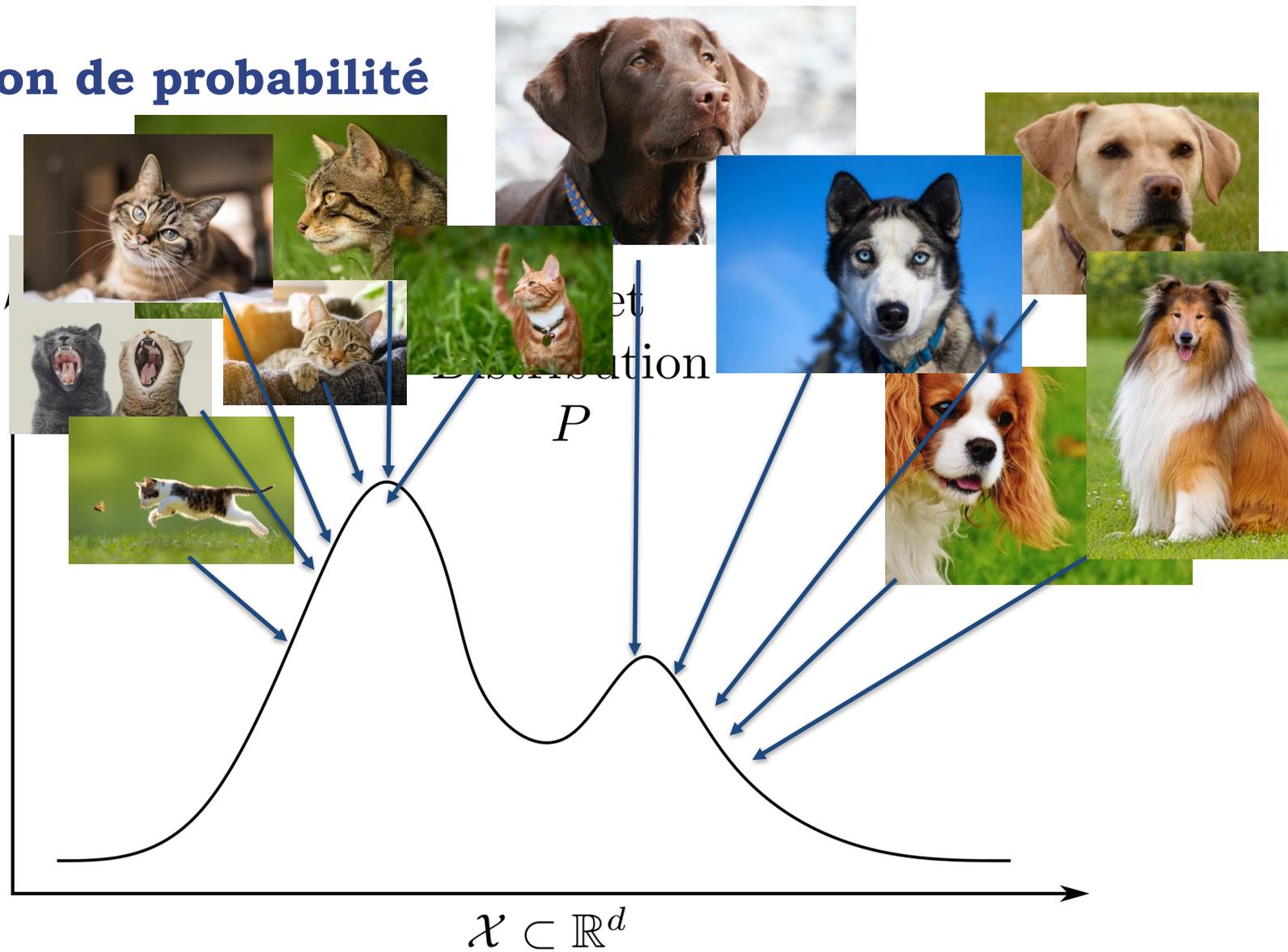
Les modèles de diffusion en détails



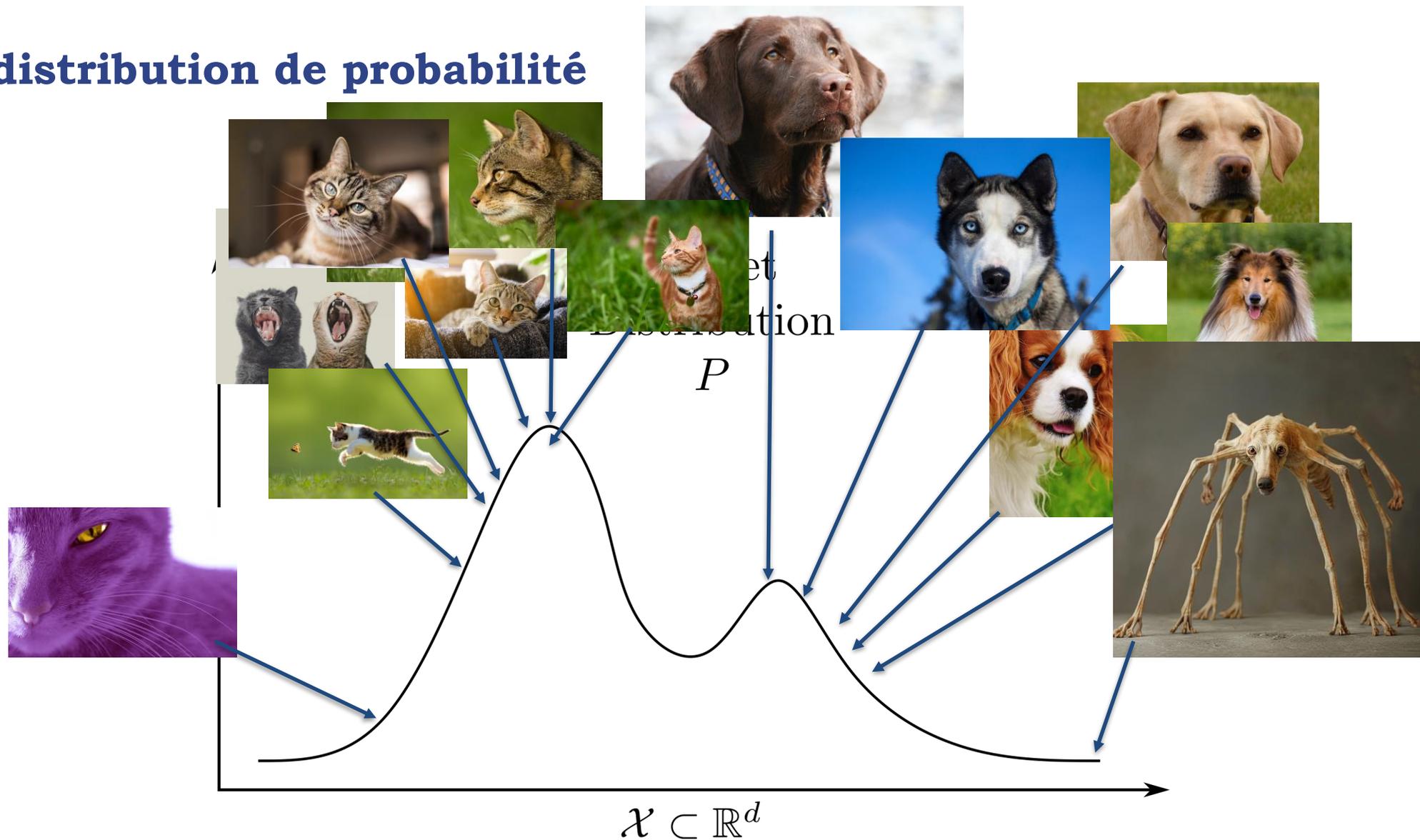
Leurs applications

**Qu'est-ce qu'un
"modèle génératif" ?**

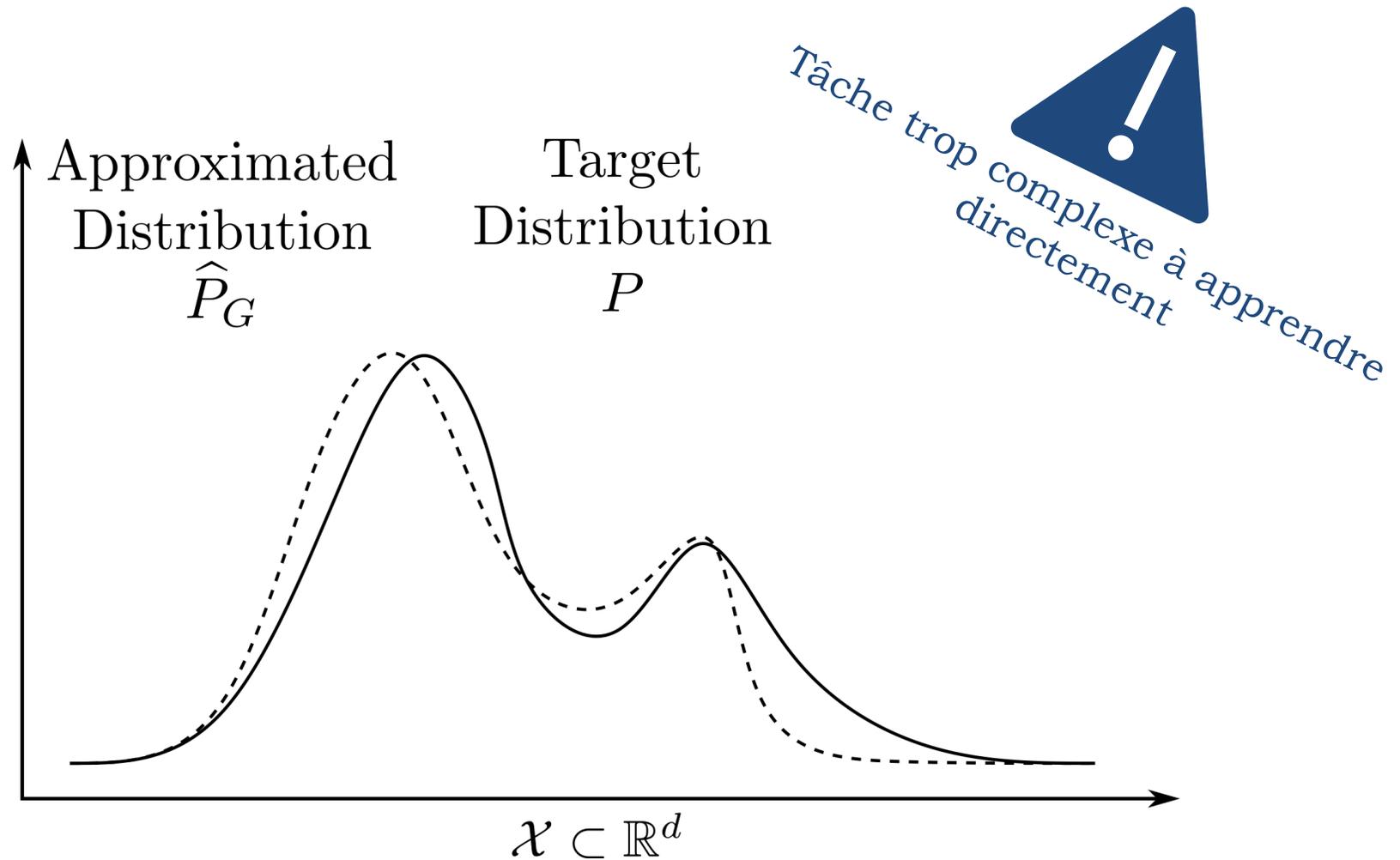
Une distribution de probabilité



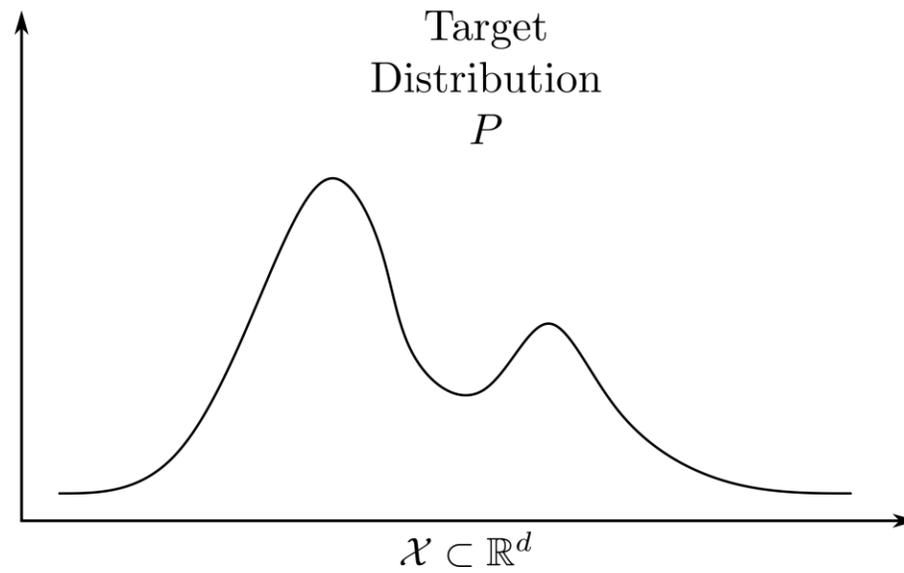
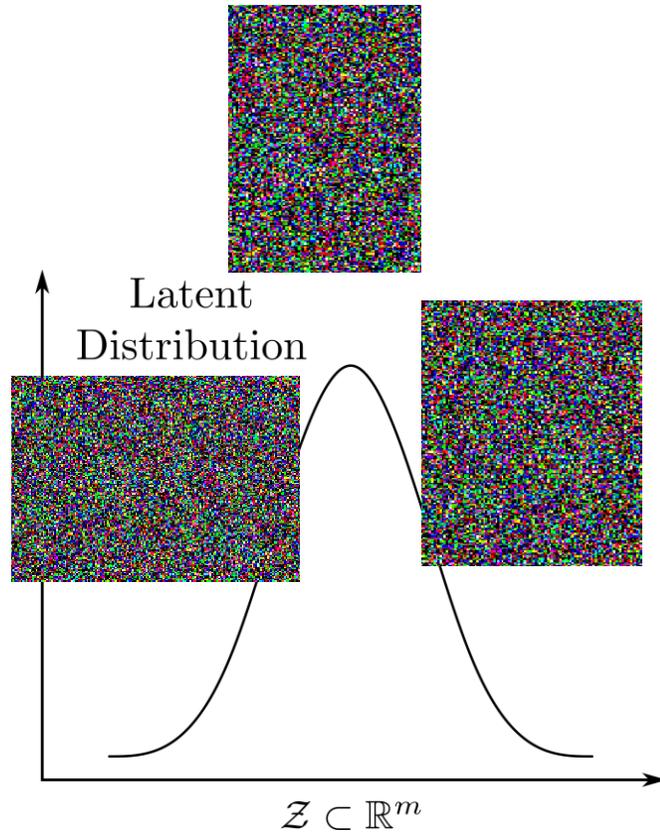
Une distribution de probabilité



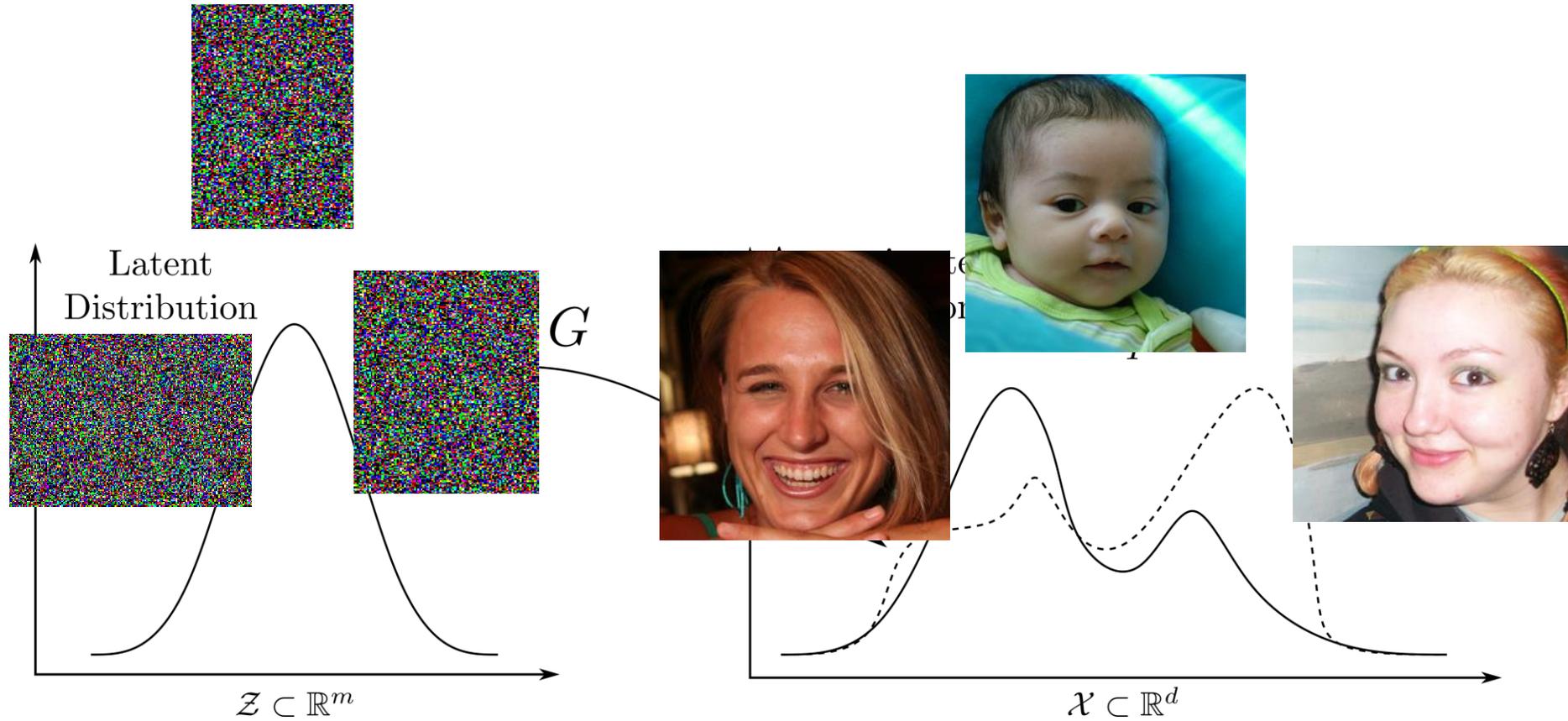
L'objectif : Apprendre une distribution de probabilité



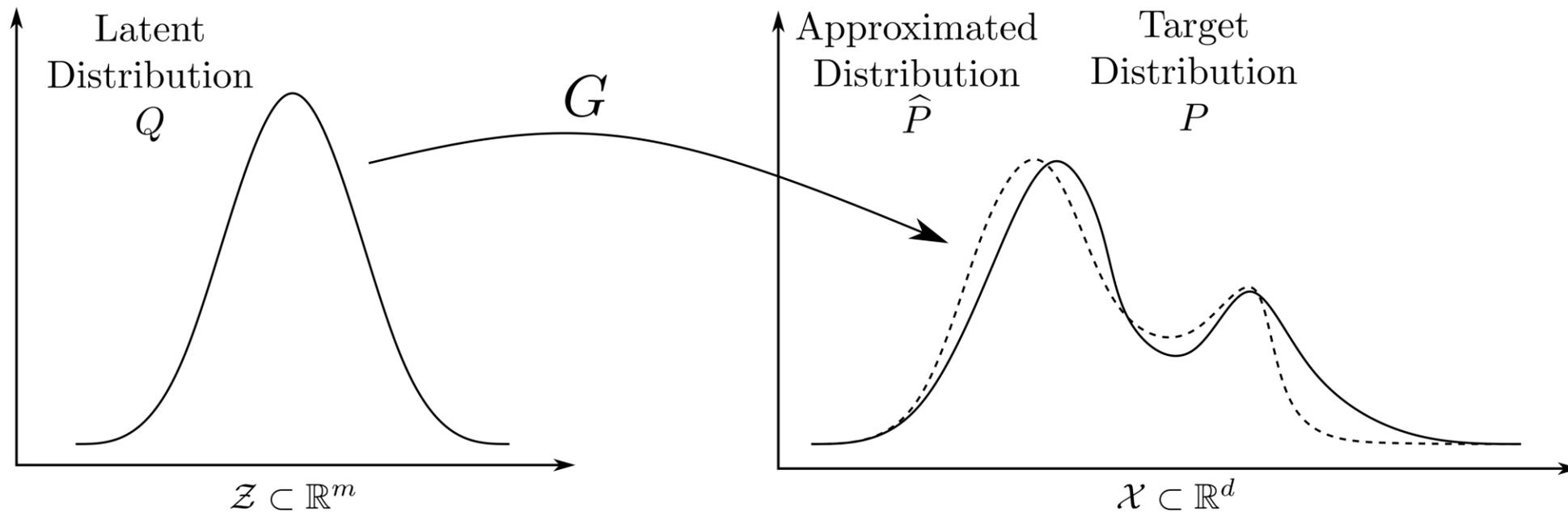
Une distribution plus simple



Une distribution plus simple et une transformation

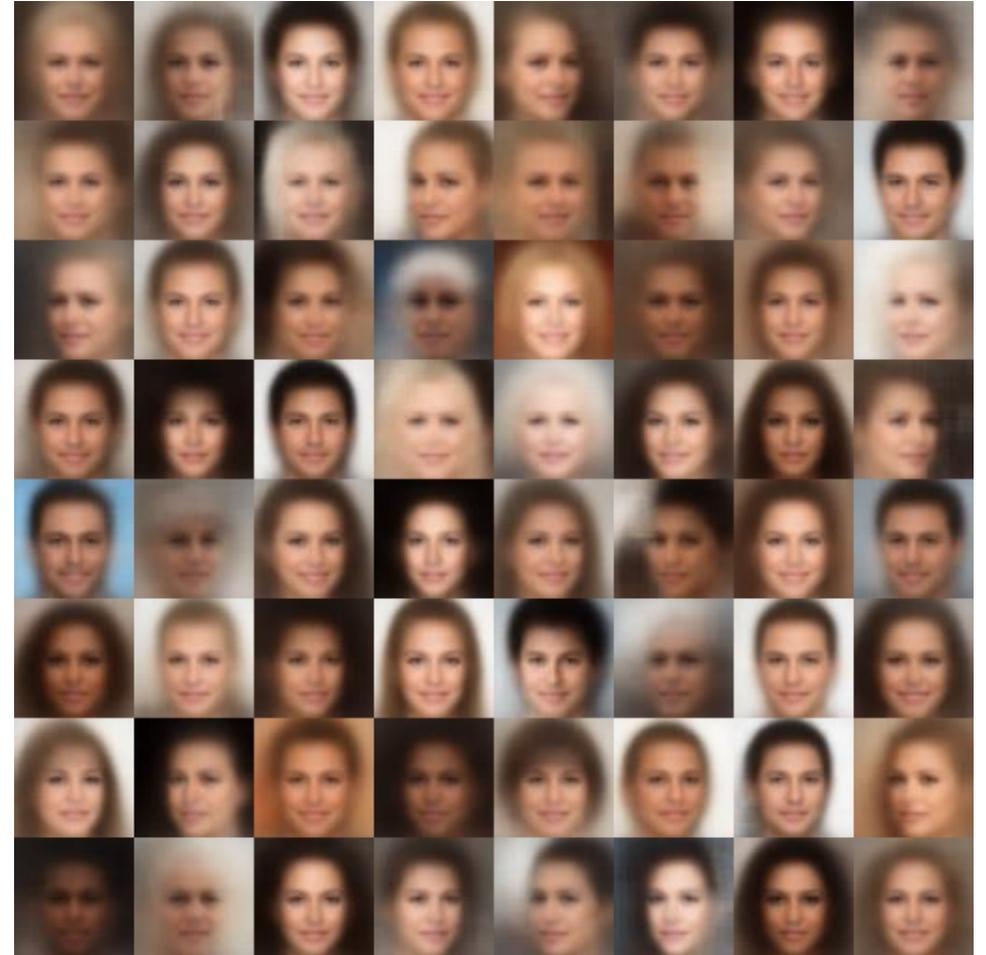
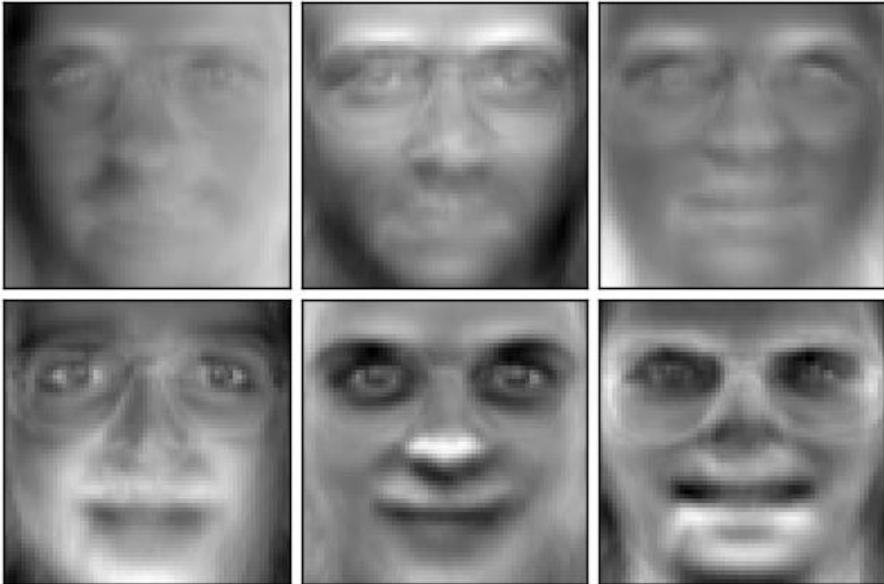


Un modèle génératif : Apprendre à générer des points d'une distribution

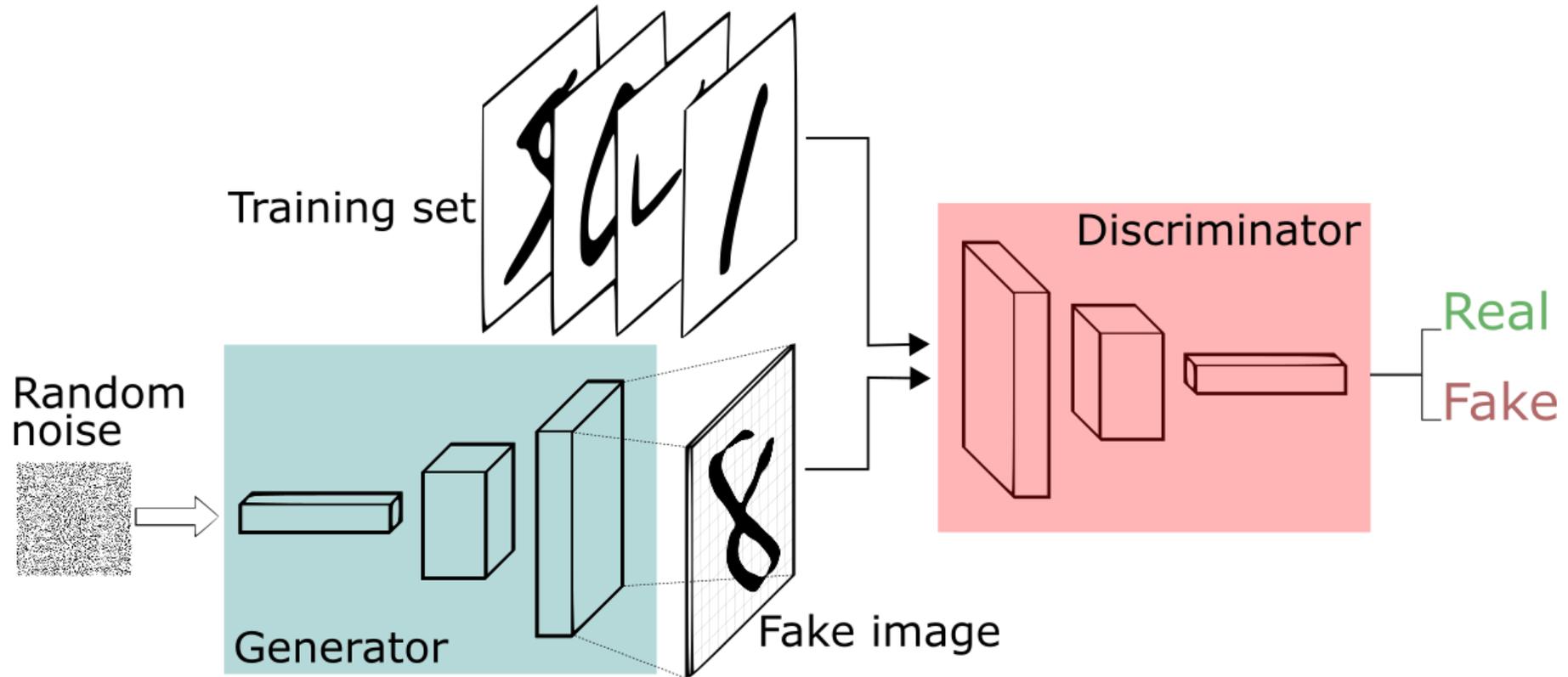


1996 à 2013 : Des Eigens face aux VAE

Eigenfaces - RandomizedPCA - Train time 0.2s



2014 : Generative Adversarial Networks

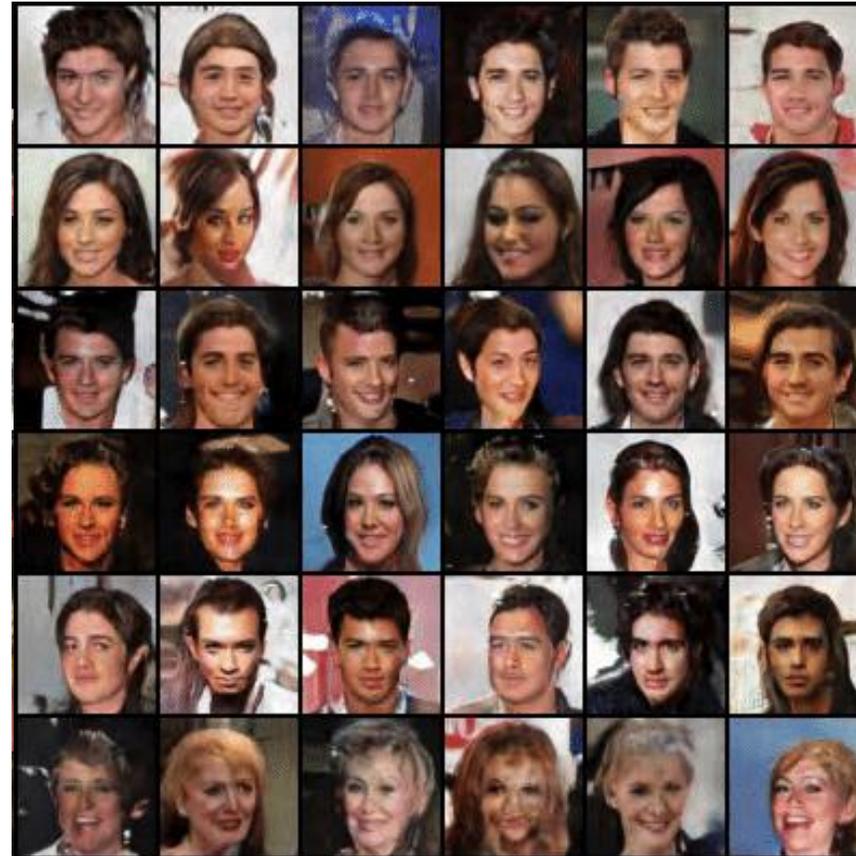


2014 : Generative Adversarial Networks



2014

2014 : Generative Adversarial Networks



2016

2014 : Generative Adversarial Networks



2018

2014 : Generative Adversarial Networks



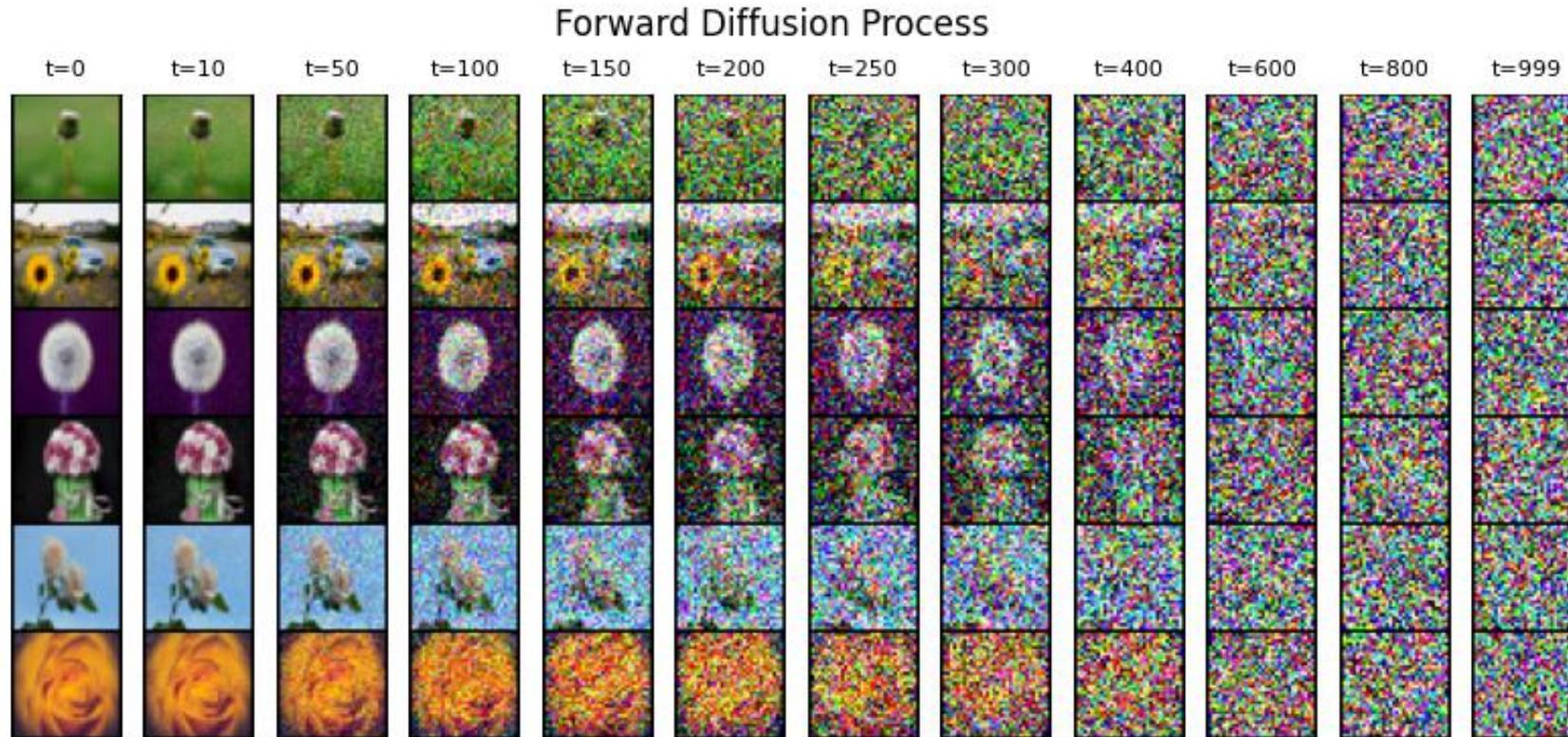
2019

[This person does not exist](#)

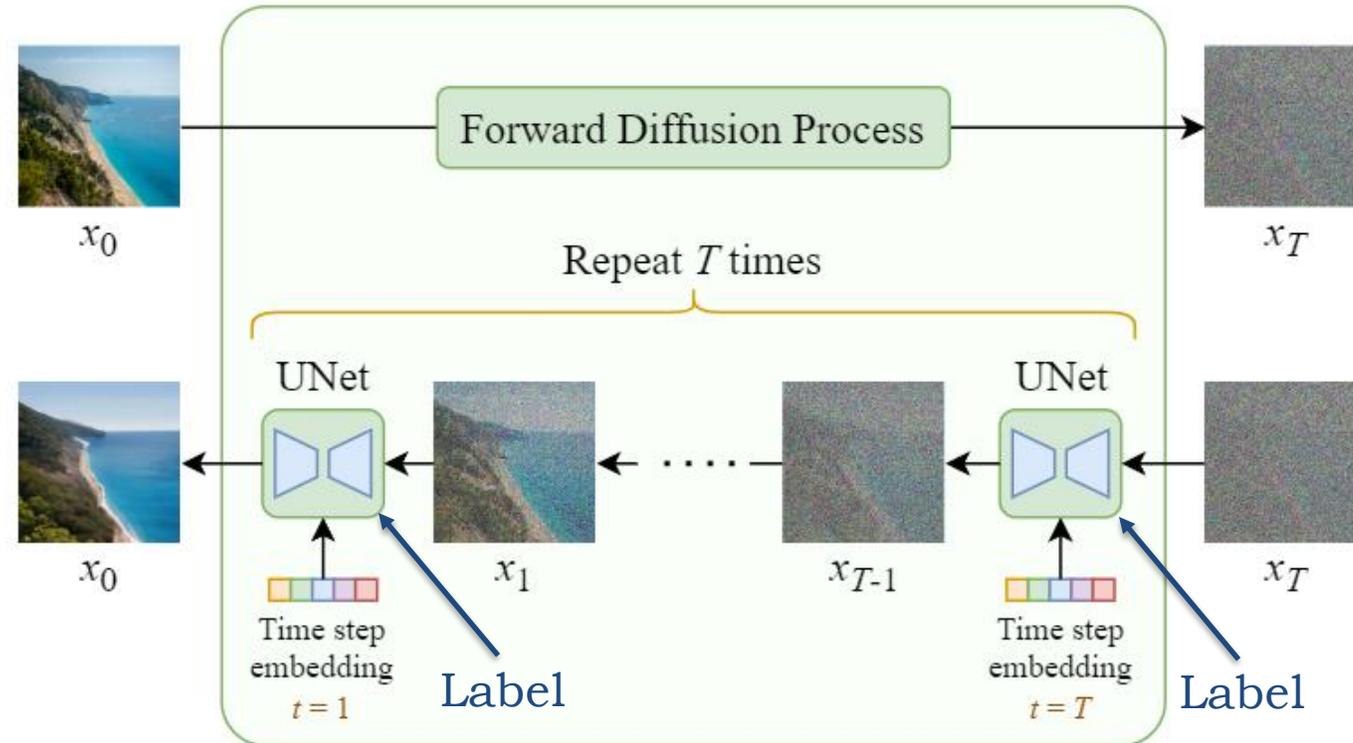
2019 : Modèles de Diffusion



Modèles de Diffusion ?



Modèle de diffusion : Apprendre à débruiter



Modèles de diffusion

Avantages:

- Plus stable à entraîner
- Plus expressif

Inconvénients:

- Plus long à entraîner
- Plus long à l'inférence

2021 : Guided Diffusion Models

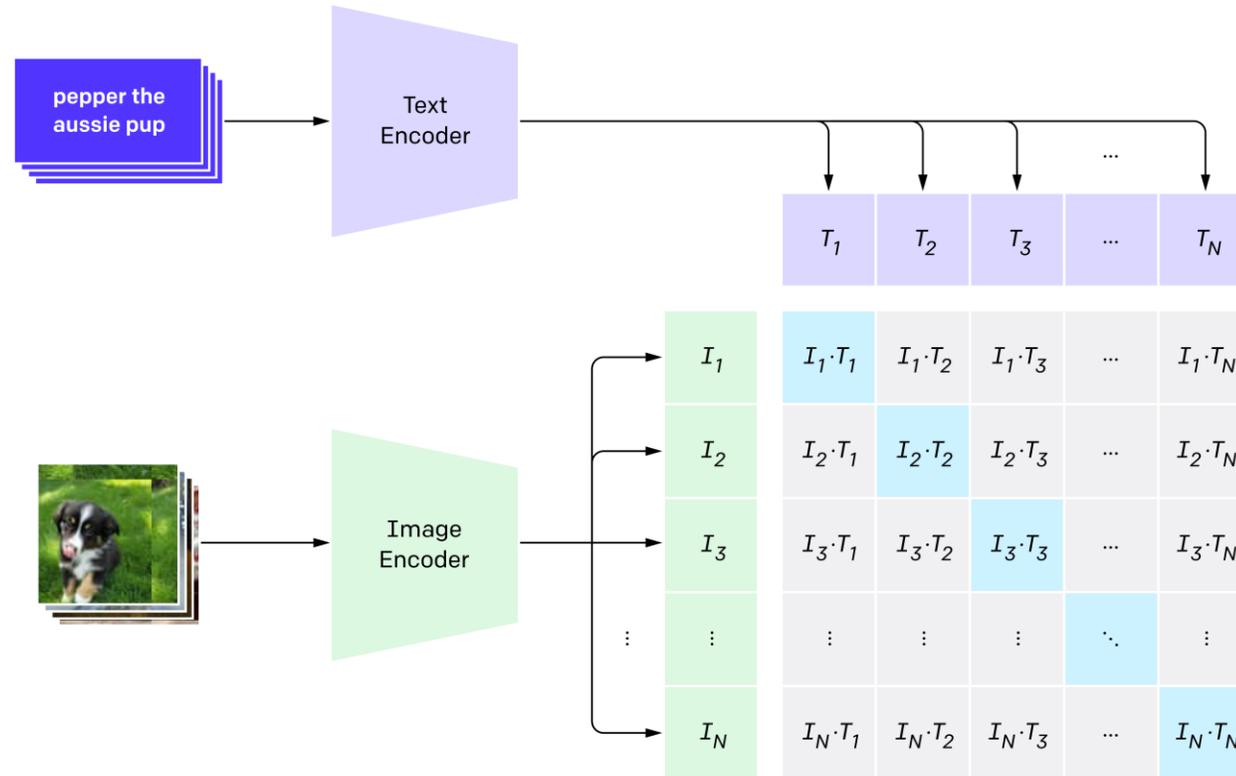


Sans Label

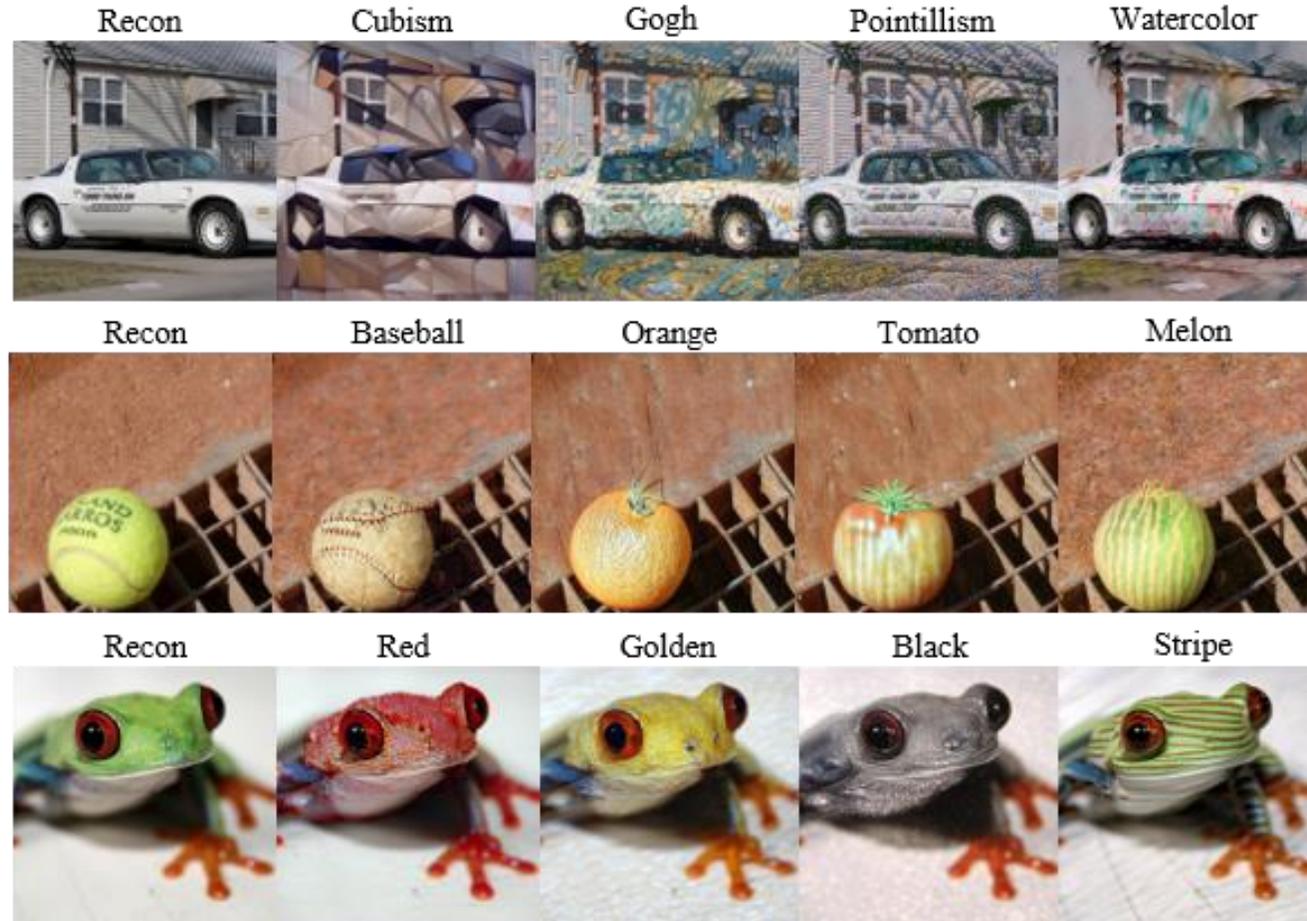


Avec Label

2021 : CLIP

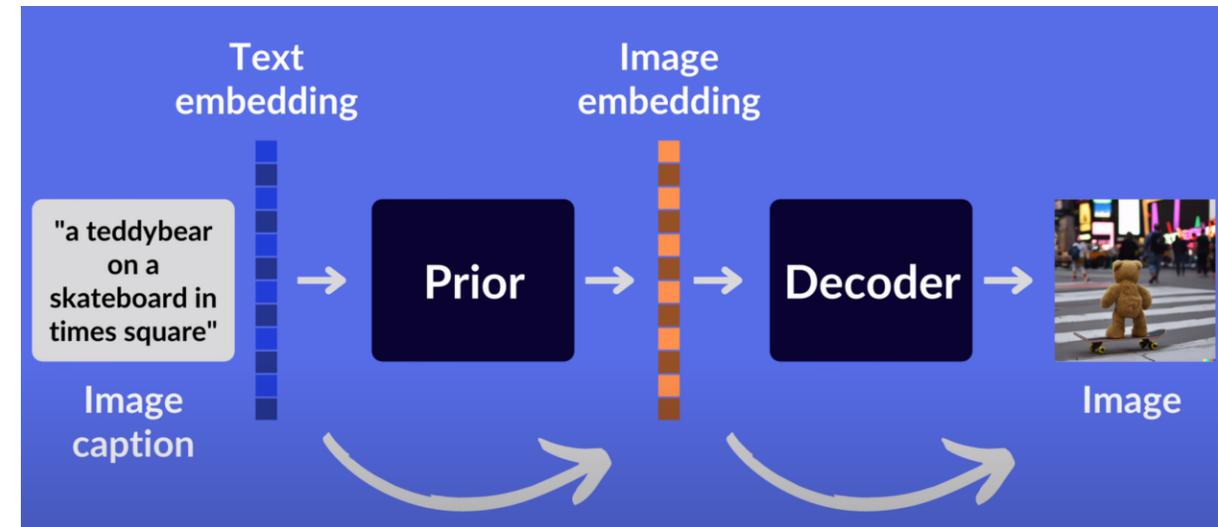
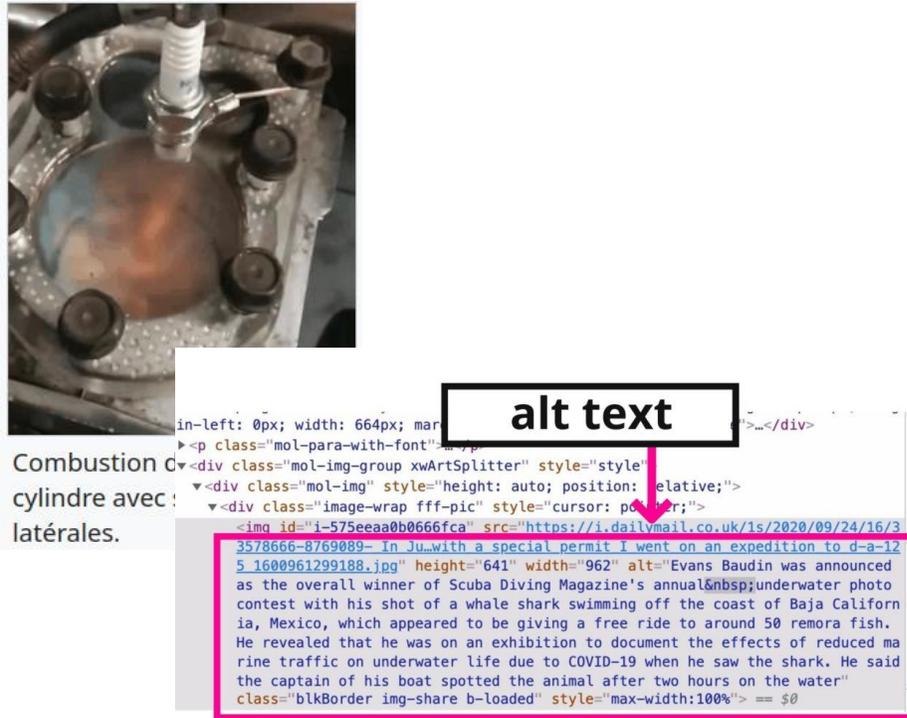


2021 : CLIP Diffusion

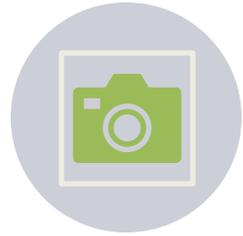


Comment ces modèles d'aujourd'hui ont vu le jour ?

- Nouvelles architectures multi-modale
- Forte disponibilité d'images et de textes alignés



Les modèles de diffusion en 2023



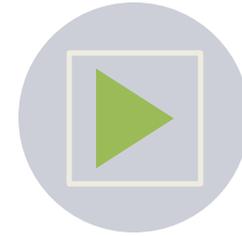
TEXTE VERS IMAGE



IMAGE VERS IMAGE



IMAGE VERS VIDÉO



TEXTE VERS VIDÉO



SONS....

Texte vers Image

- Principaux modèles :
 - Dall-E (OpenAI)
 - MidJourney (Independent)
 - Stable Diffusion (Open Source)
- Objectif : **générer des images à partir de textes.**

Prompt: Photo of a banana dog



Une histoire de compromis

MidJourney V5



Dall-E 2



Dall-E 3

- Développé par OpenAI
- Accessible via ChatGPT Premium
- API disponible

- Avantage :
 - Intégré à ChatGPT
 - Facile à prendre en main
- Principal défaut :
 - Passer par l'interface de chat gpt
 - Très peu interactif

F You
generate an image which is a photo of a banana dog

ChatGPT



Here's the whimsical photo of a banana transformed to resemble a dog, sitting in a natural setting and exuding a playful and adorable character.

🔊 📄 ↺ 🗑️

Dall-E 3

Use You
add a baby pacifier in the dog's mouth

ChatGPT

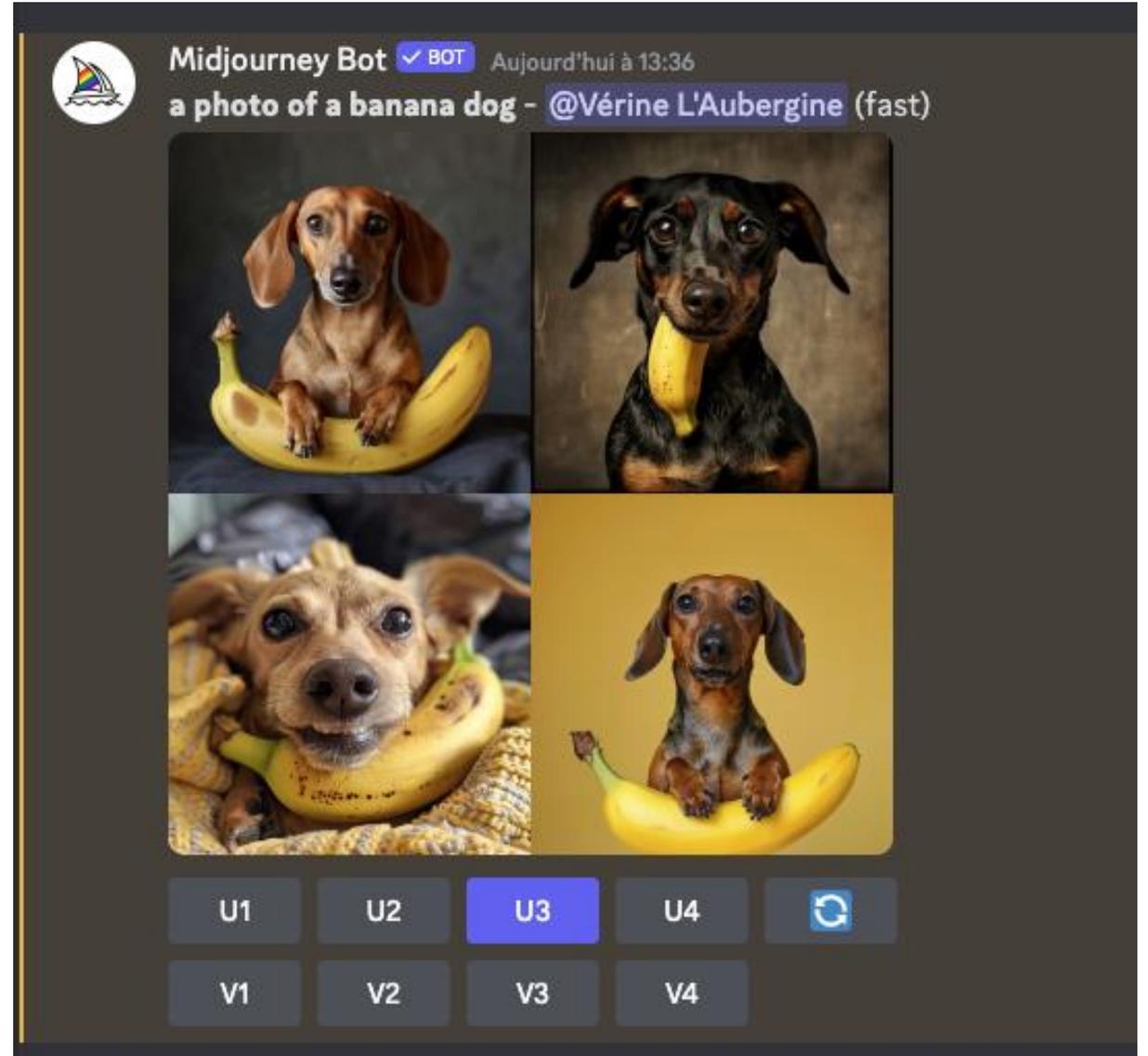


I've added a baby pacifier to the banana dog's mouth, enhancing its whimsical and adorable appearance.

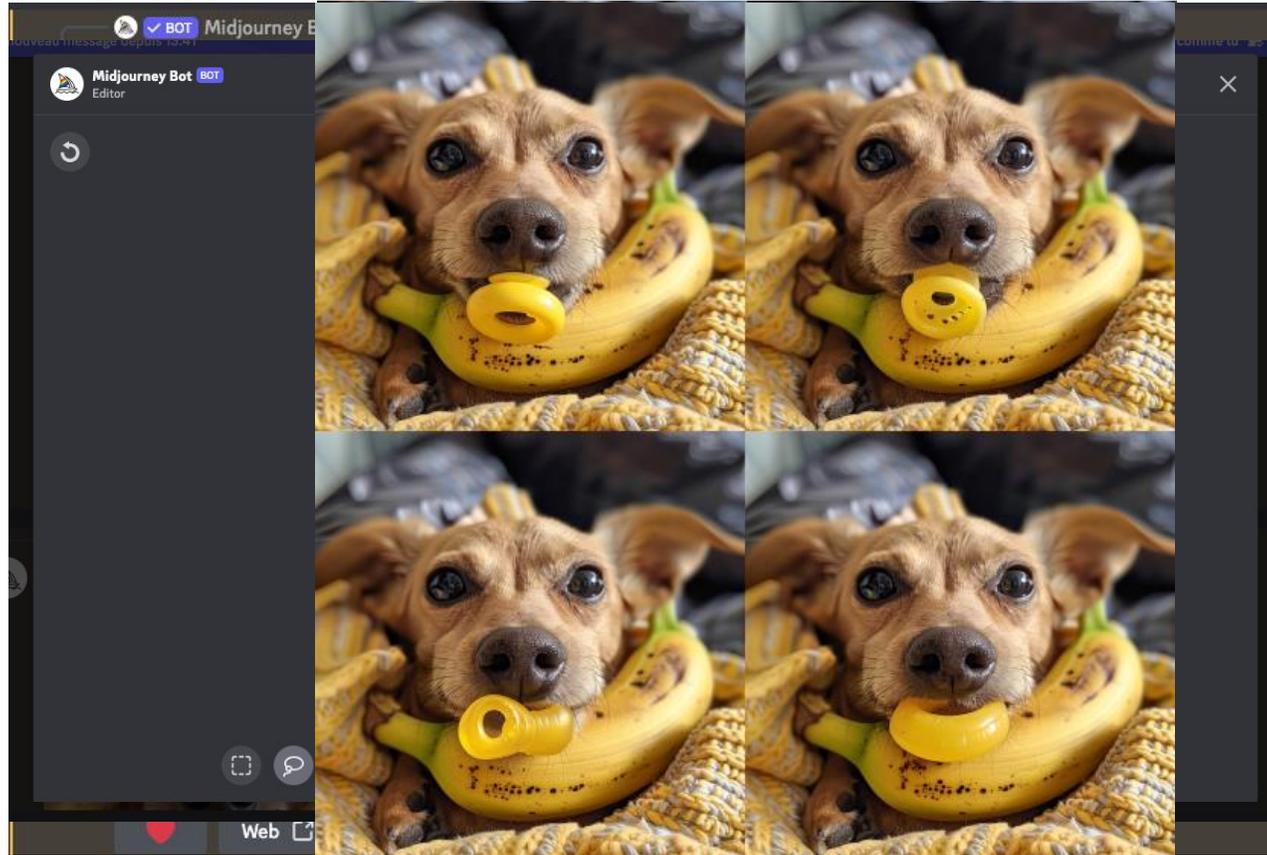
🔊 📄 ↻ 🗑️

MidJourney

- Développé par Midjourney
- Accessible via Discord (payant)
- (A mon avis) Le plus complet



MidJourney



Fonctionnement des modèles : prompt engineering

- Modèles très dépendants du prompt (peut-être encore plus qu'en texte)
- Difficile d'avoir un rendu satisfaisant du premier coup
- Conseils pour un bon prompt (jusqu'à 350 mots):
 - Détailler les objets, leur emplacement, leur couleur
 - Donner le ratio de l'image et des détails photographiques (Ouverture, longueur de focal, ISO, etc...)
 - Donner des détails d'ambiance, de lumière, etc...
 - Utiliser des mots spécifiques.
 - Utiliser des références de styles : [Midlibrary.io](https://www.midlibrary.io)
- Utiliser des archives de prompt (curated)
 - [PromptHero](https://www.prompthero.com)
 - [PromptBase](https://www.promptbase.com)

Texte vers Image (mais aussi Image vers Texte)



/describe



1. A banana in the shape of an animal, sitting on its back legs like it's playing with someone. The background is grassy and sunny. A cute dog made from bananas sits next to him. High resolution photography. Highly detailed sharp focus --ar 100:99
2. A cute banana dog made of bananas sitting on the ground, sunlight, grass in background, photo realistic, hyper detailed --ar 100:99
3. A cute banana dog made of bananas, sitting on the ground in front of green grass, sunny day, bright light, realistic photography style, creative character design, cute and colorful, closeup shots, simple background, high definition --ar 100:99
4. A cute banana dog made of bananas sitting on the ground, green grass in background, sunlight, high resolution photography, insanely detailed --ar 100:99

Texte vers Image

1. A banana in the shape of an animal, sitting on its back legs like it's playing with someone. The background is grassy and sunny. A cute dog made from bananas sits next to him. High resolution photography. Highly detailed sharp focus --ar 100:99
2. A cute banana dog made of bananas sitting on the ground, sunlight, grass in background, photo realistic, hyper detailed --ar 100:99
3. A cute banana dog made of bananas, sitting on the ground in front of green grass, sunny day, bright light, realistic photography style, creative character design, cute and colorful, closeup shots, simple background, high definition --ar 100:99
4. A cute banana dog made of bananas sitting on the ground, green grass in background, sunlight, high resolution photography, insanely detailed --ar 100:99



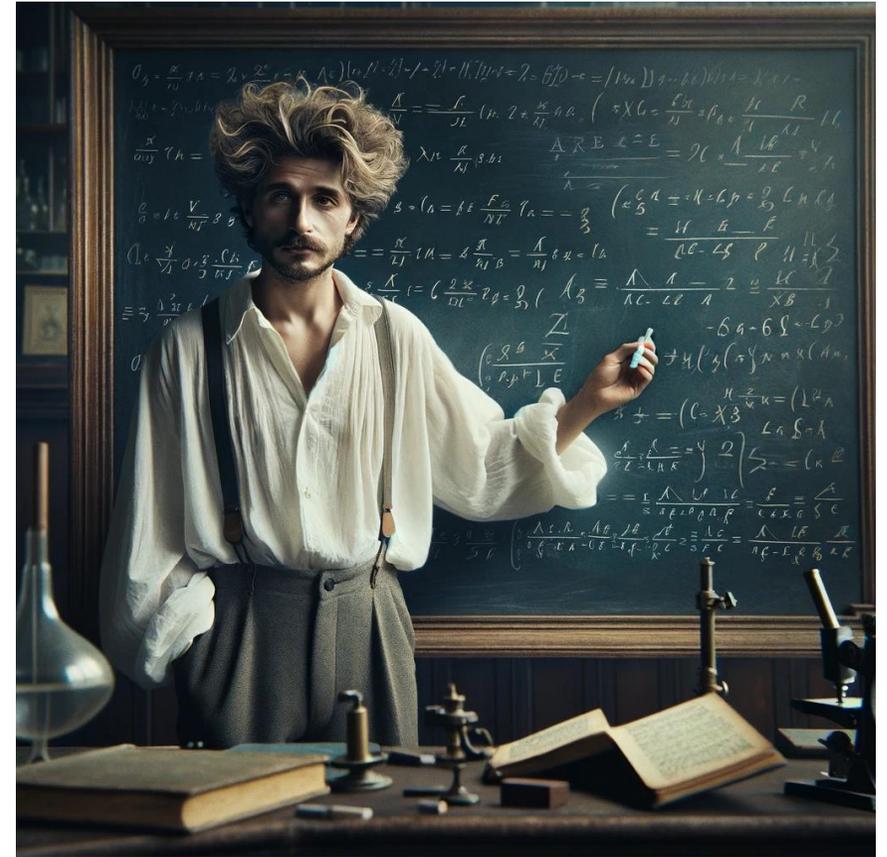
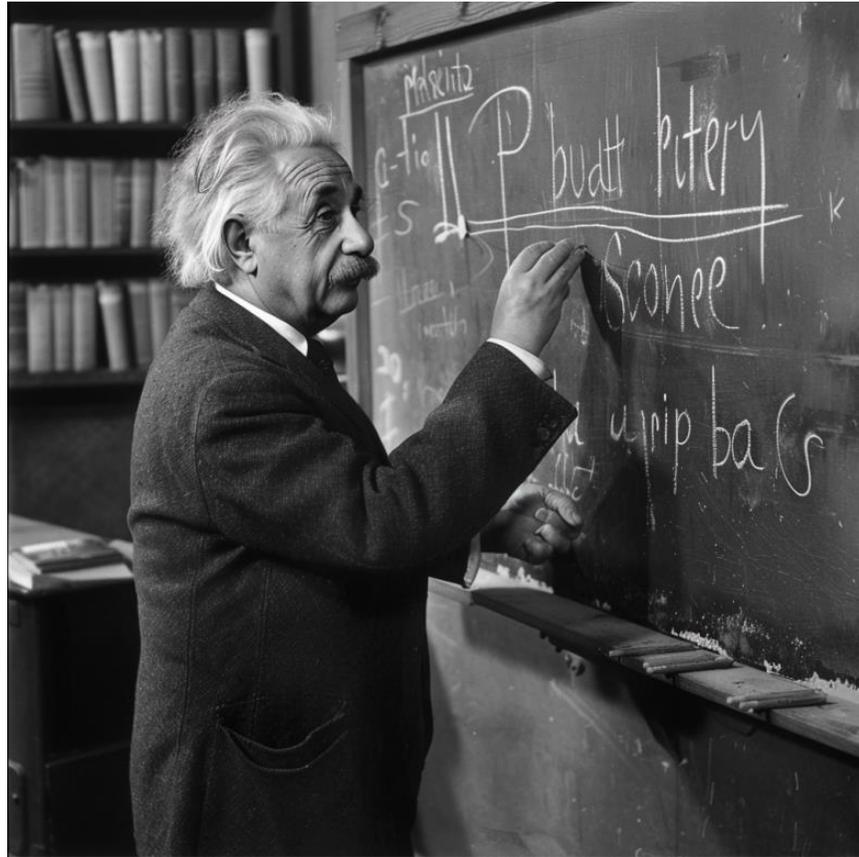
Texte (+ exemple) vers Image



A cute banana dog made of bananas sitting on the ground, green grass in background, sunlight, high resolution photography, insanely detailed --ar 100:99 --sref (l'image du dessus)

Texte vers Image (+ Texte)

Albert Einstein writing his name on the blackboard



ling
ic
ting.
ge
k

Texte vers Image (+ Texte)

The first page of the first book of William Shakespeare



 You
Generate a photo a the first page of the first book of William Shakespear

 ChatGPT
 Error creating image

I'm unable to generate an image of the first page of William Shakespeare's first book due to content policy restrictions. If there's anything else I can help you with, please let me know!

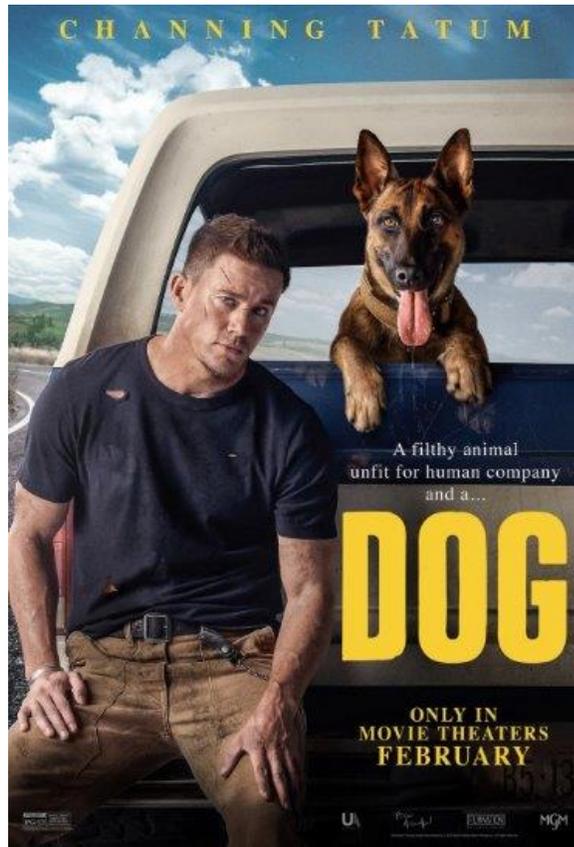


Image vers Image

- Principaux modèles/acteurs :
 - Midjourney
 - Runway ML
 - FireFly (Adobe)
- Modèle Génératif pour le Business
 - Flair.ai
 - Mokker.ai
 - Uizard.io
 - MS Designer

Image vers Image

- Objectif : changer le chien par un chat

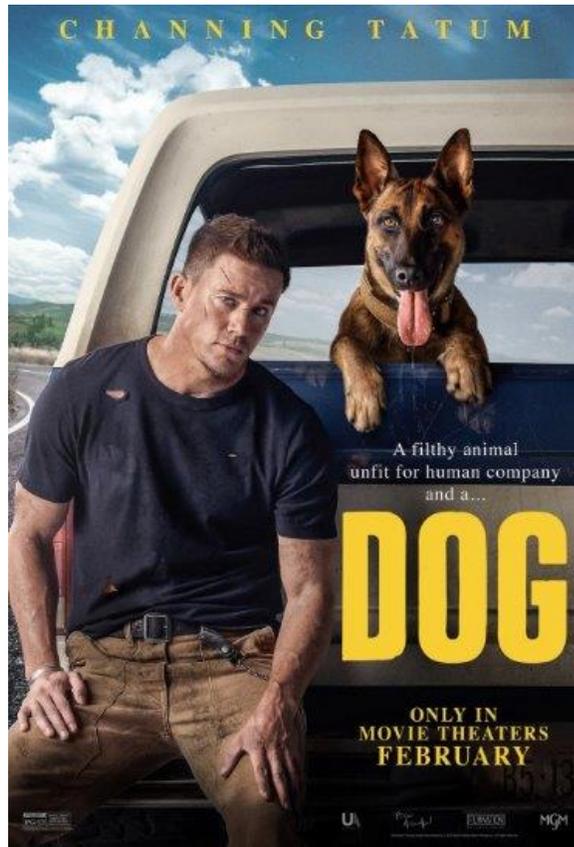


MidJourney



Image vers Image

- Objectif : changer le chien par un chat



DALL-3



Image vers Image

- Objectif : changer le chien par un chat

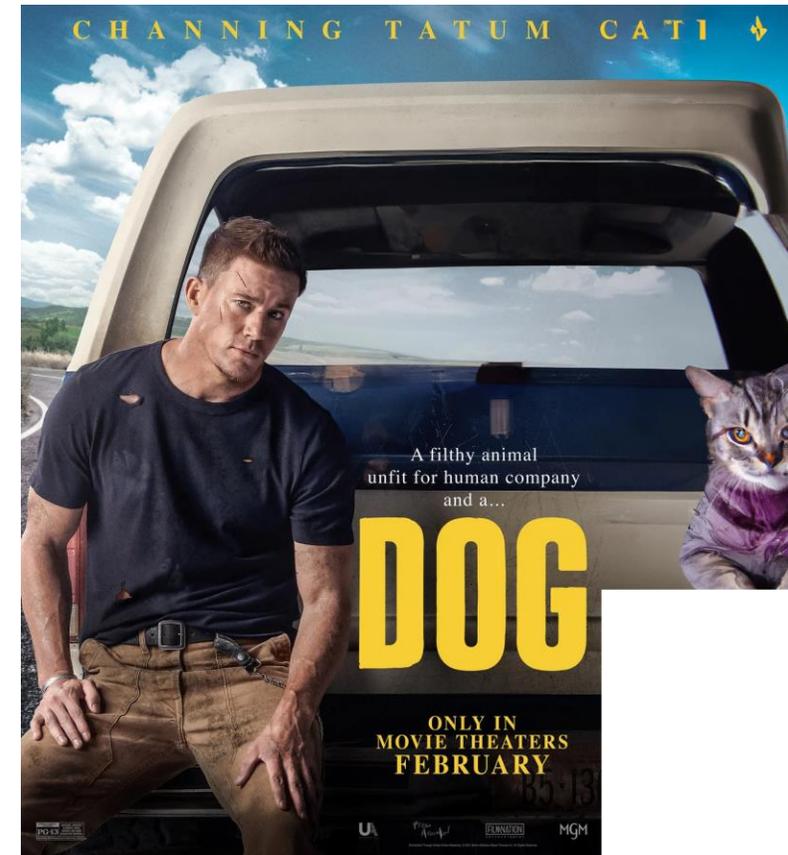
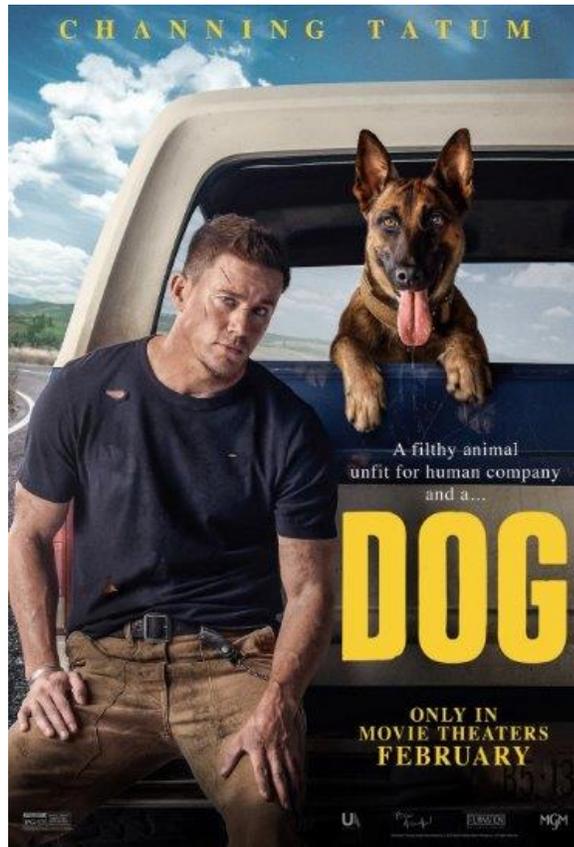


Image vers Image

- Objectif : changer le chien par un chat

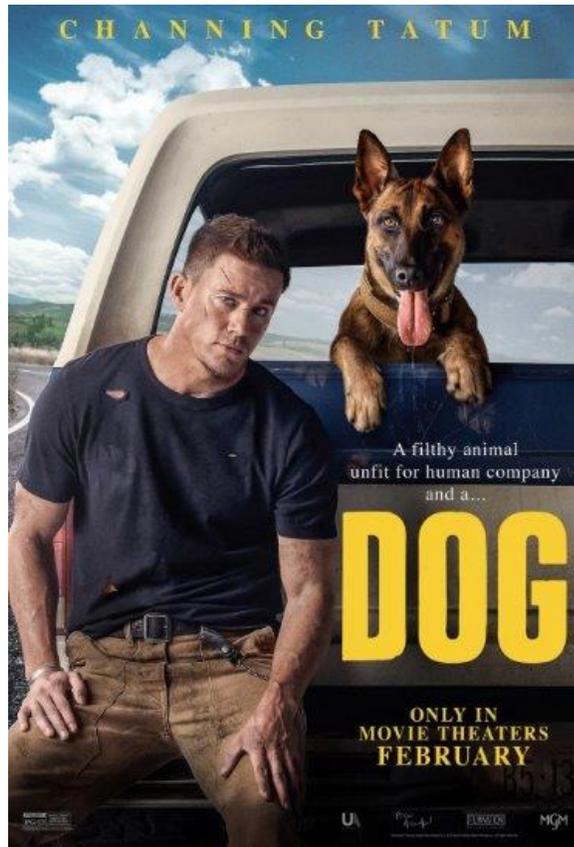
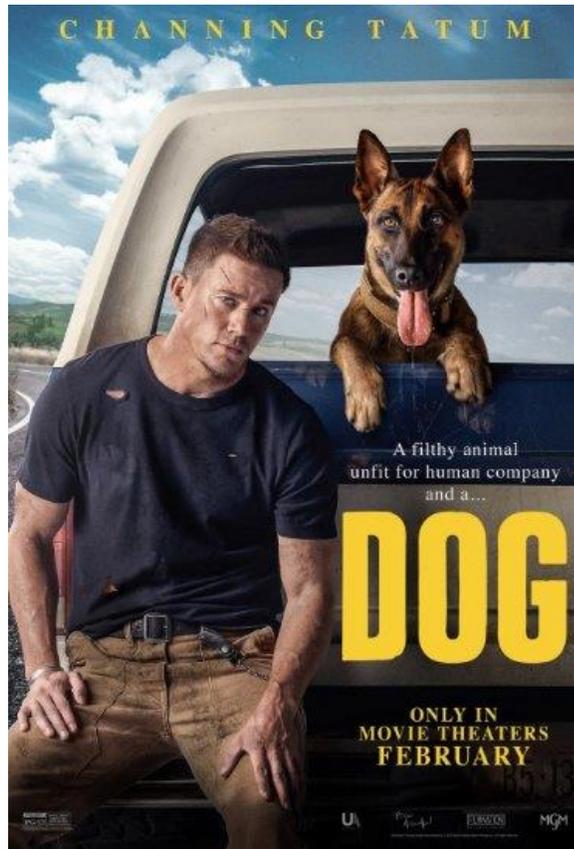


Image vers Image

- Objectif : changer le chien par un chat



FireFly

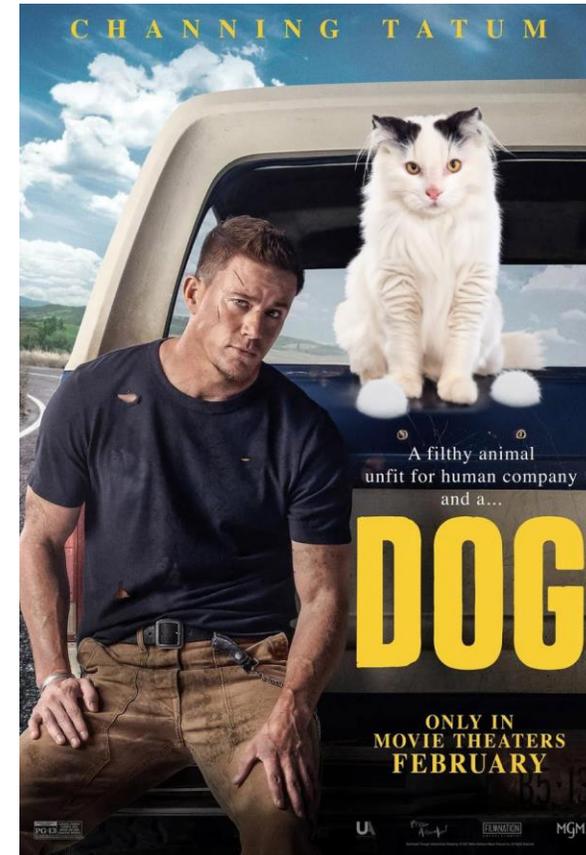


Image vers Image

- Objectif : changer le chien par un chat

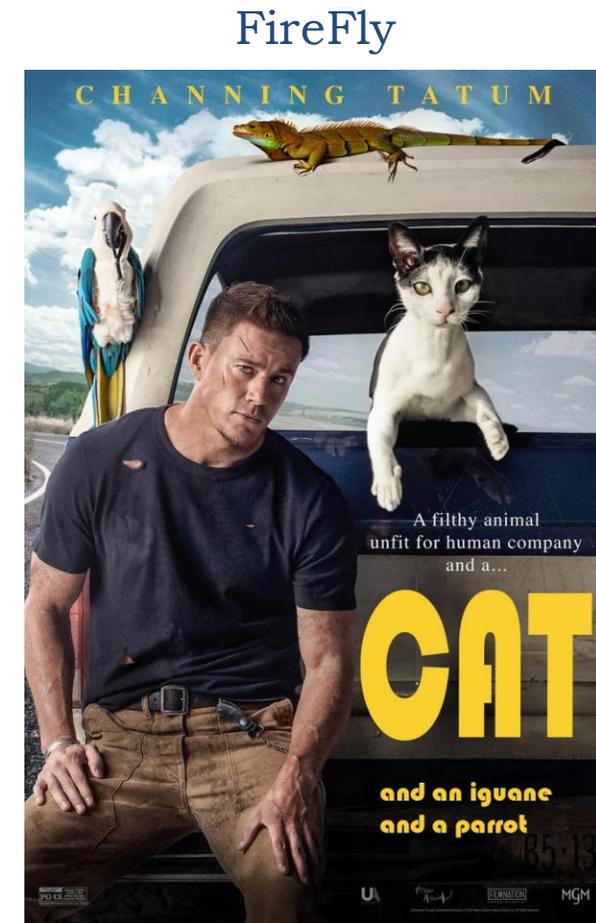
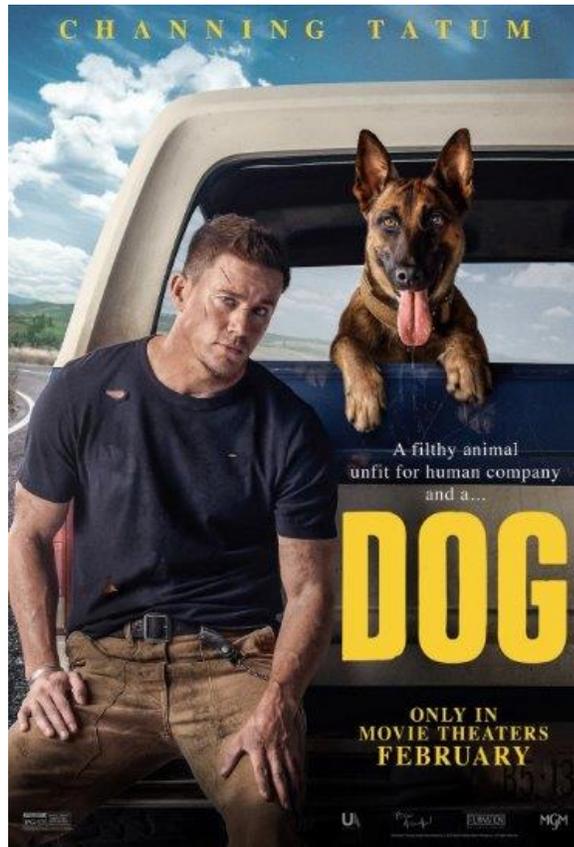


Image vers Image

FireFly

=

La référence en
In Painting



Images vers Image

- Mixer des images



DALL-3



Images vers Image

- Mixer des images



MidJourney



Images vers Image

- Mixer des images



MidJourney

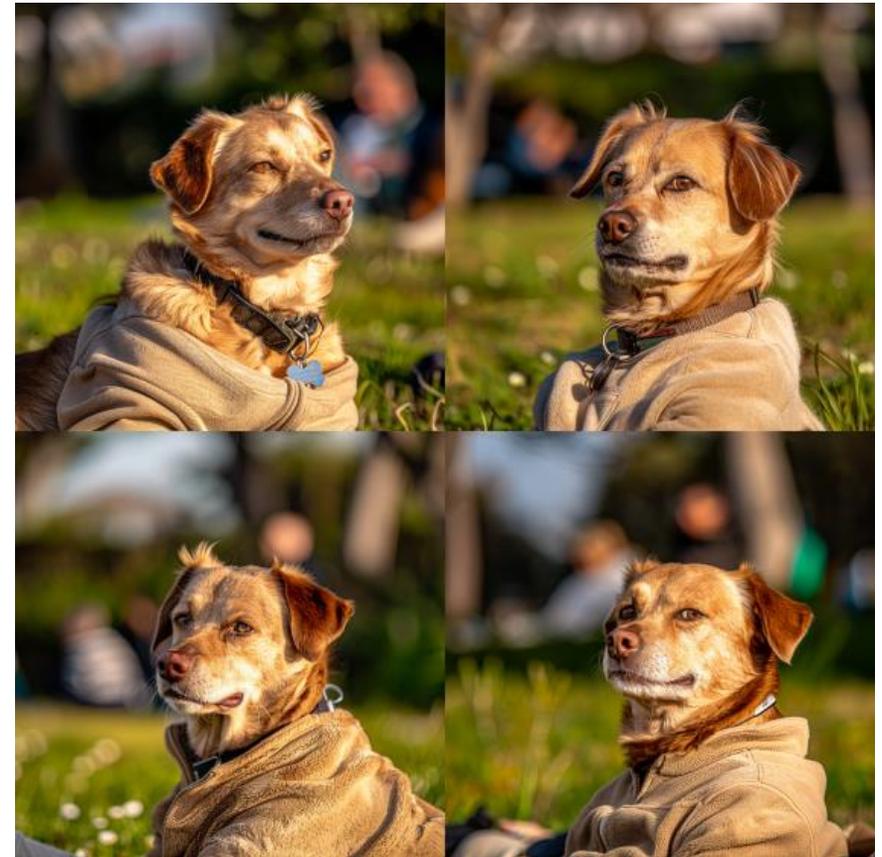


Images vers Image

- Mixer des images



MidJourney



Images vers Image

- Mixer des images



MidJourney



Images vers Image

- Mixer des images



MidJourney



Images vers Image

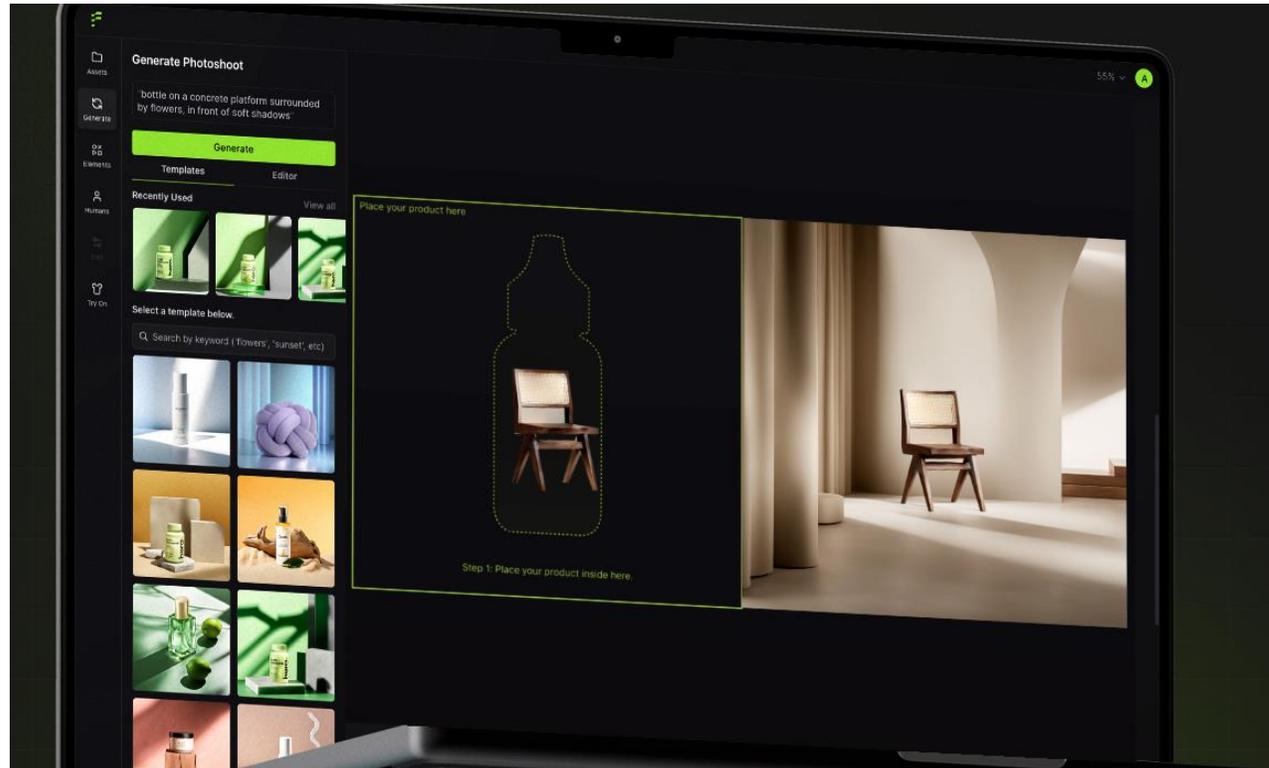
- Mixer des images



MidJourney



Image vers Image : [Flair.ai](https://www.flair.ai)



NEW FEATURE

Fashion photoshoots

Fit your cloth onto AI generated models. Create high-quality on-model imagery at scale.

Prompt

photo of a woman wearing a long-sleeve navy top in front of white background.

Generate

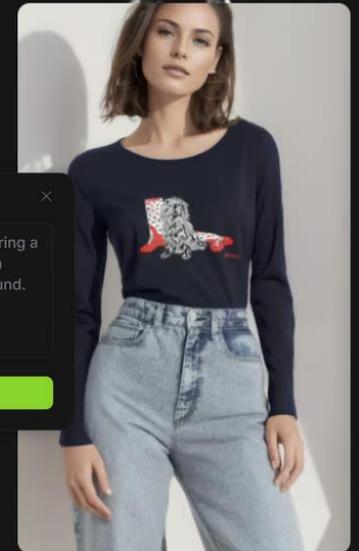
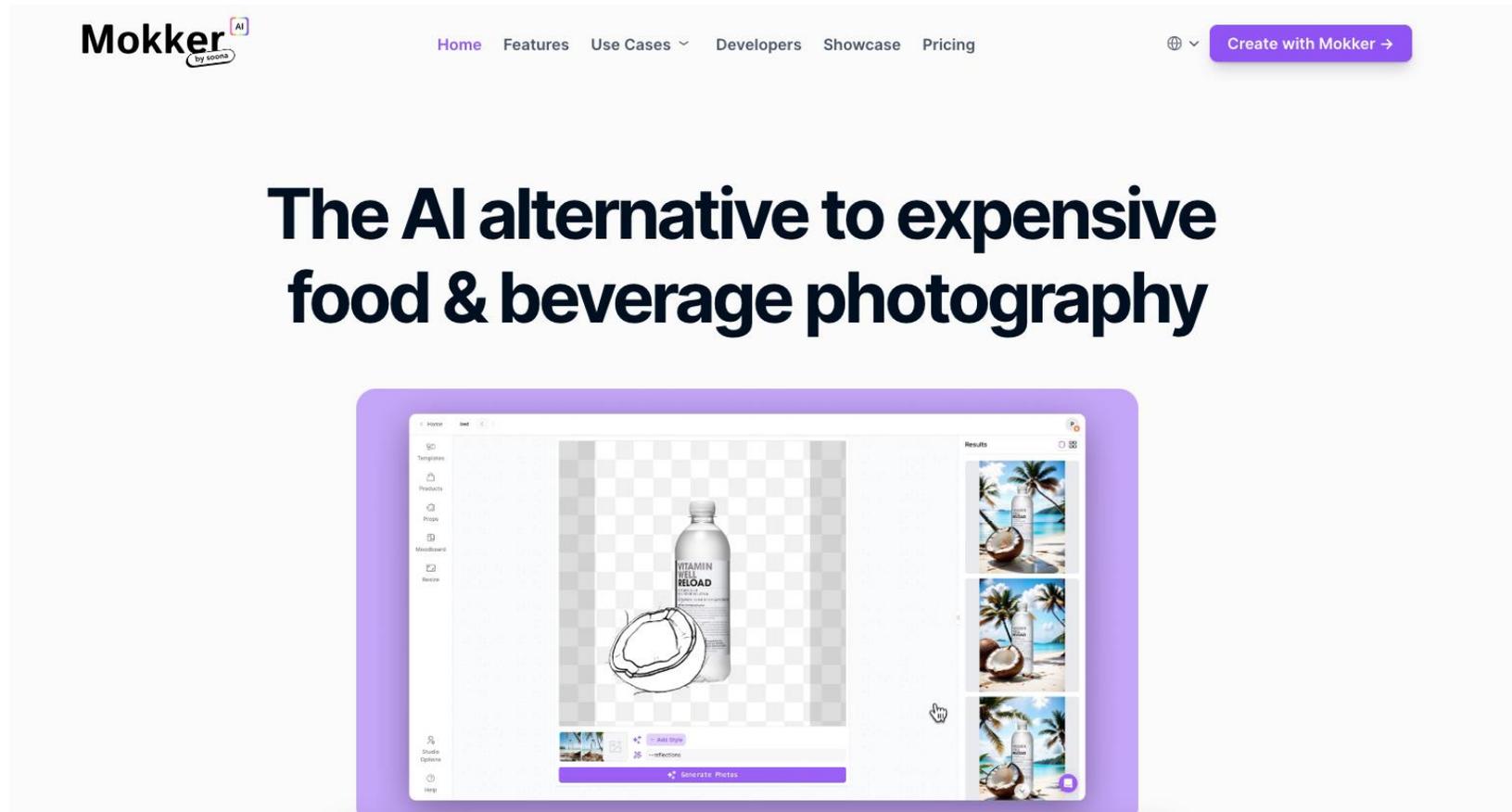


Image vers Image: [Mokker.ai](https://www.mokker.ai)

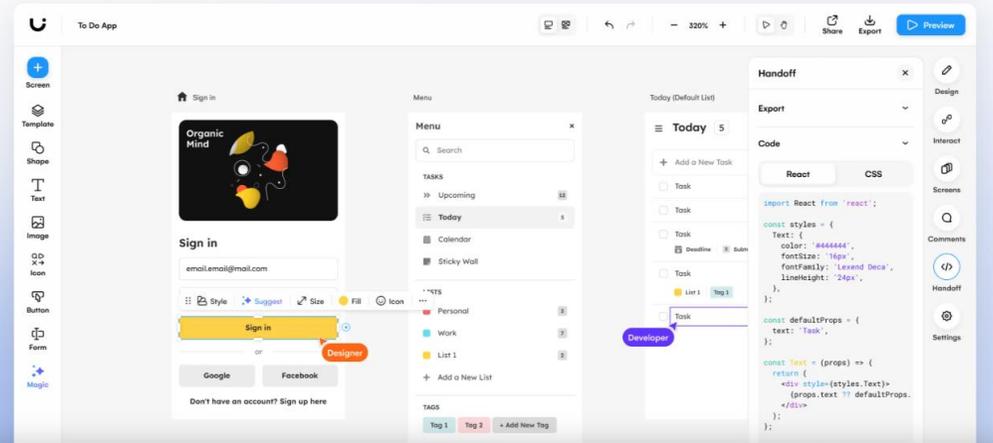


Texte (et Image) vers UI : [Uizard.io](https://uizard.io)

Visualize product ideas fast and easy with AI

Create, validate, and communicate design concepts in minutes, not days.
Empower your product team today!

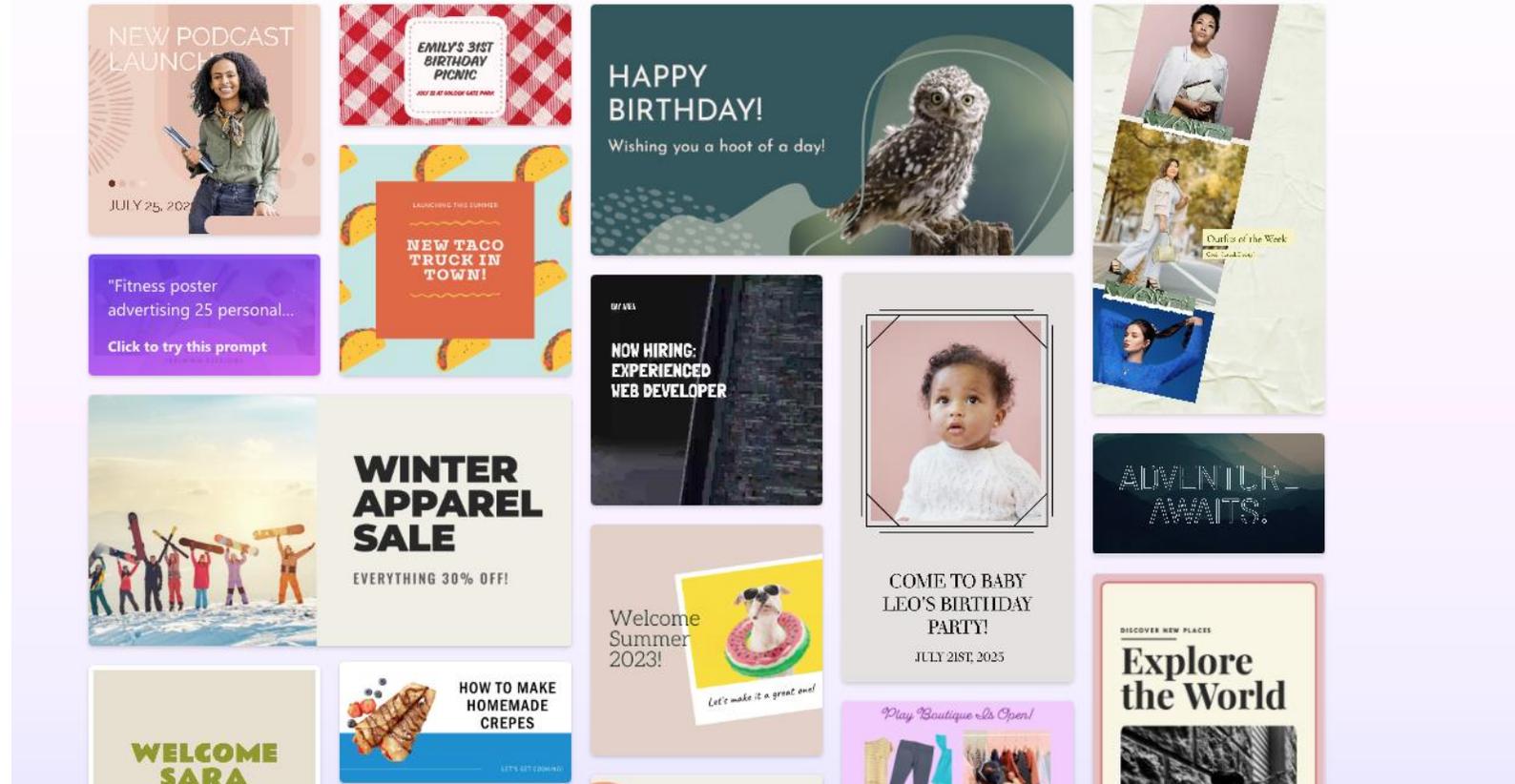
An e-commerce website design for a clothing brand Generate



The screenshot displays the Uizard.io design tool interface. On the left, there's a sidebar with various design tools like Screen, Templates, Shape, Text, Image, Icon, Button, Form, and Magic. The main workspace shows a 'Sign in' form for 'Organic Mind' with a search bar, a menu, and a task list. A 'Developer' role is highlighted on the task list. The right sidebar shows a 'Code' editor with React and CSS code.

```
import React from 'react';  
  
const styles = {  
  text: {  
    color: '#666666',  
    fontSize: '16px',  
    fontFamily: 'Legend Deca',  
    lineHeight: '24px',  
  },  
};  
  
const defaultProps = {  
  text: 'Task',  
};  
  
const Text = (props) => {  
  return (  
    <div style={styles.Text}>  
      {props.text || defaultProps.  
    </div>  
  );  
};
```

Image et Texte vers Image: MS Designer



Et la vidéo ?

- Principaux acteurs:
 - Runway ML
 - Sora

Image (+Texte) vers Vidéo: Runwayml.com



Make the people dancing

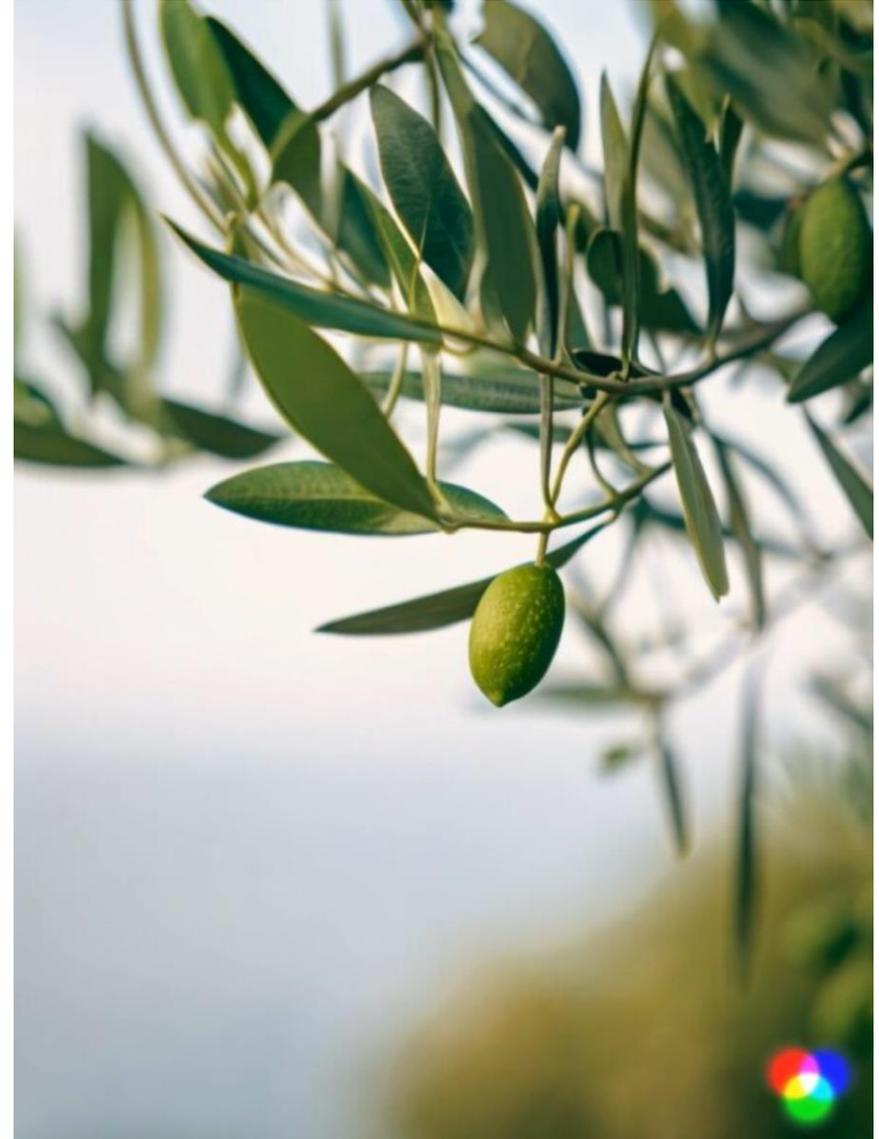


Image Vers Vidéo



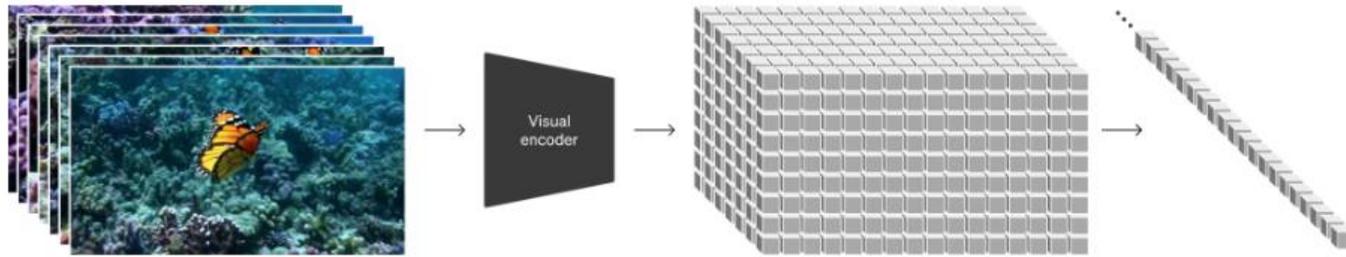
+ Movement brush

Image Vers Vidéo



+ Camera Movement

Vivement Sora !



Et le son ?

- Modèle pour la voix :
 - [Eleven Labs](#)
 - [TTS](#) (OpenAI)
- Modèle pour le doublage ou Voix+Video
 - [HeyGen](#)
- Modèle pour la musique :
 - [MusicML](#) (Google)
 - [StabilityAI](#)

Merci à tous !



UNIVERSITÉ PARIS DAUPHINE - PSL
Place du Maréchal de Lattre de Tassigny – 75775 Paris cedex 16