

N9EE39A - Edge Computing

Introduction au projet

Présenté par Adrien THIRION

Email : adrien.thir@gmail.com

Sommaire

1: Présentation du métier de Data Scientist

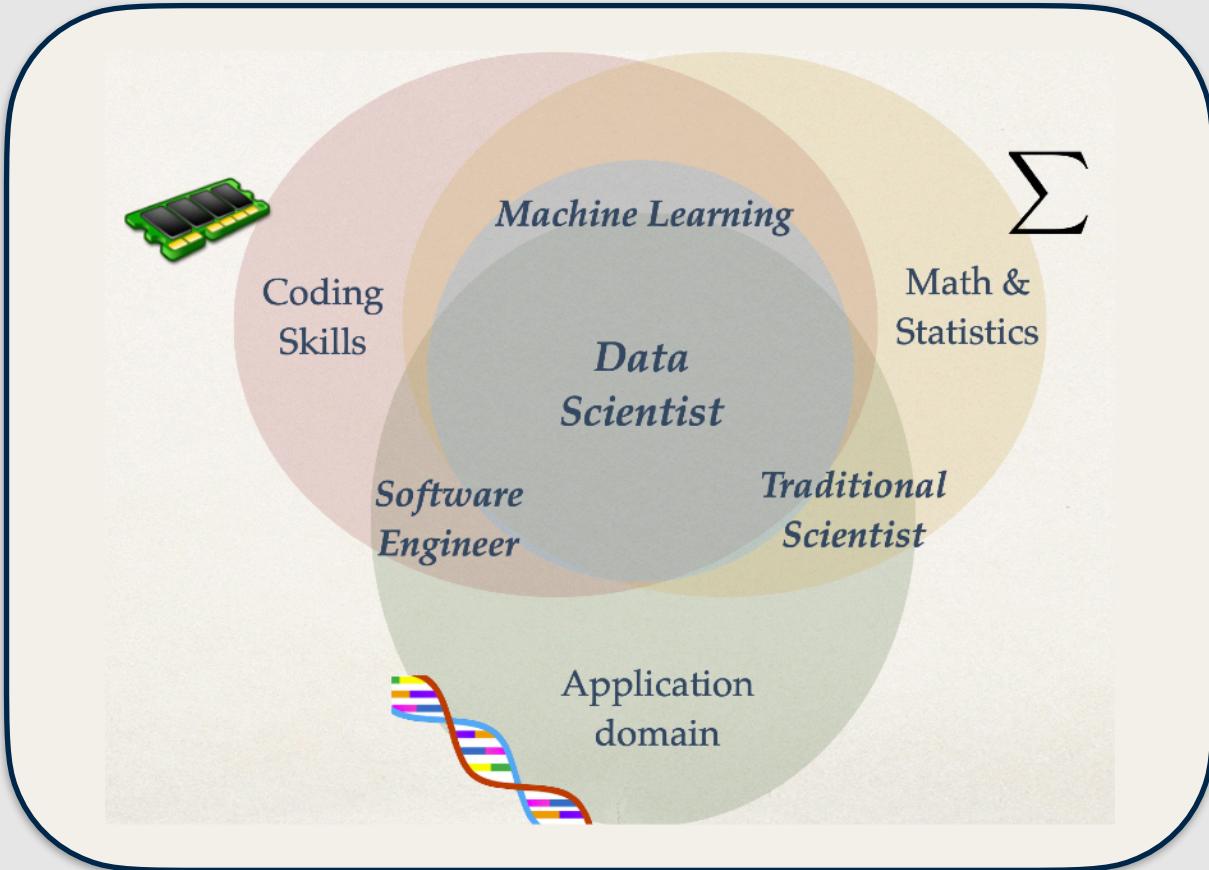
2: Présentation du TP

3: L'Etat de l'art

4: Optimisation d'hyperparamètres

5: Modèles embarqués

Le métier de Data Scientist



Les compétences clés d'un data scientist [1]

[1] Cours EPFL 2019 CS433

Recherche fondamentale

- **Objectif** : Développement de nouvelles méthodes et avancées théoriques.
- **Approche** : Exploratoire, centrée sur des problèmes théoriques.
- **Outils** : Méthodes expérimentales, algorithmes avancés, data sets publics.
- **Résultats** : Mesurés par publications, impact académique.
- **Données** : Simulées ou publiques, moins de contraintes légales.

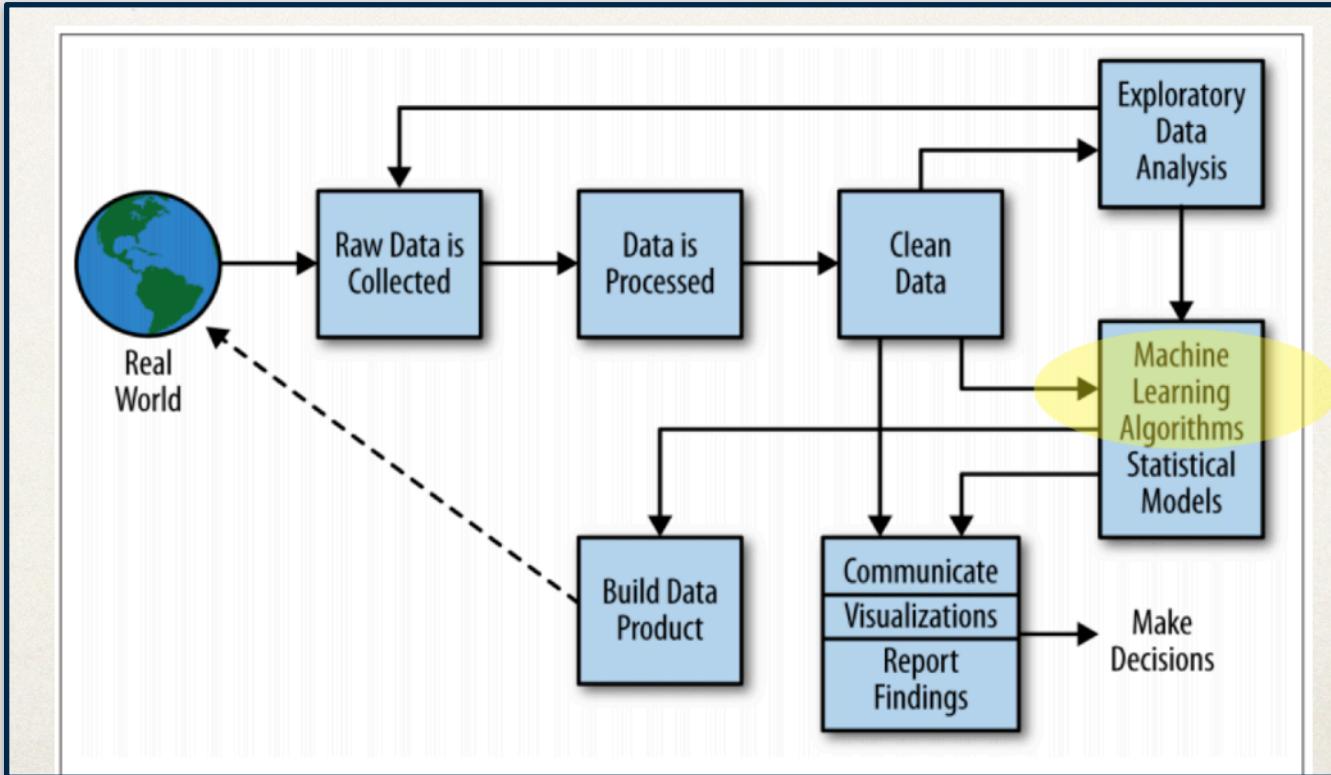
=> Bonne connaissance en math & stats

Recherche « appliquée »

- **Objectif** : Solutions pratiques pour besoins commerciaux immédiats.
- **Approche** : Optimisation orientée ROI (Return of Investments), contraintes de temps et de production.
- **Outils** : Technologies éprouvées, facilement déployables.
- **Résultats** : KPIs (Key Performance Indicators) commerciaux (ventes, coûts, satisfaction client).
- **Données** : Propres à l'entreprise, souvent sensibles et régulées.

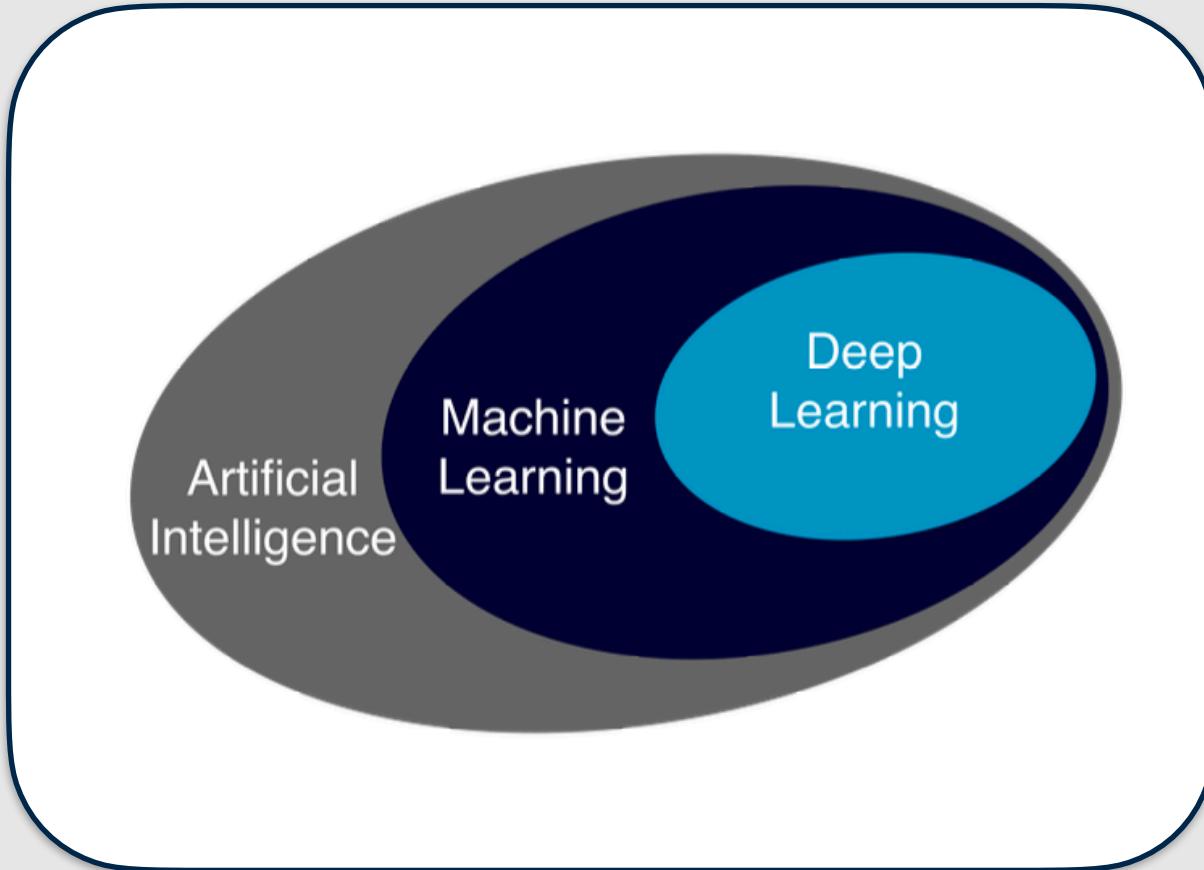
=> Bonne connaissance dans le domaine d'application

Le métier de Data Scientist



Le travail d'un data scientist [1]

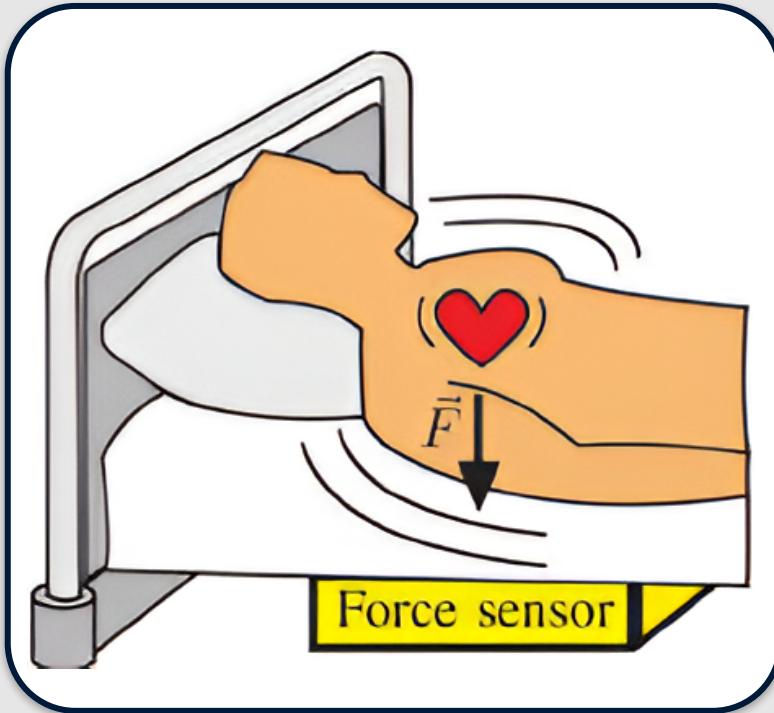
Le métier de Data Scientist



Les compétences clés d'un data scientist

Quand utiliser IA vs ML vs DL ?

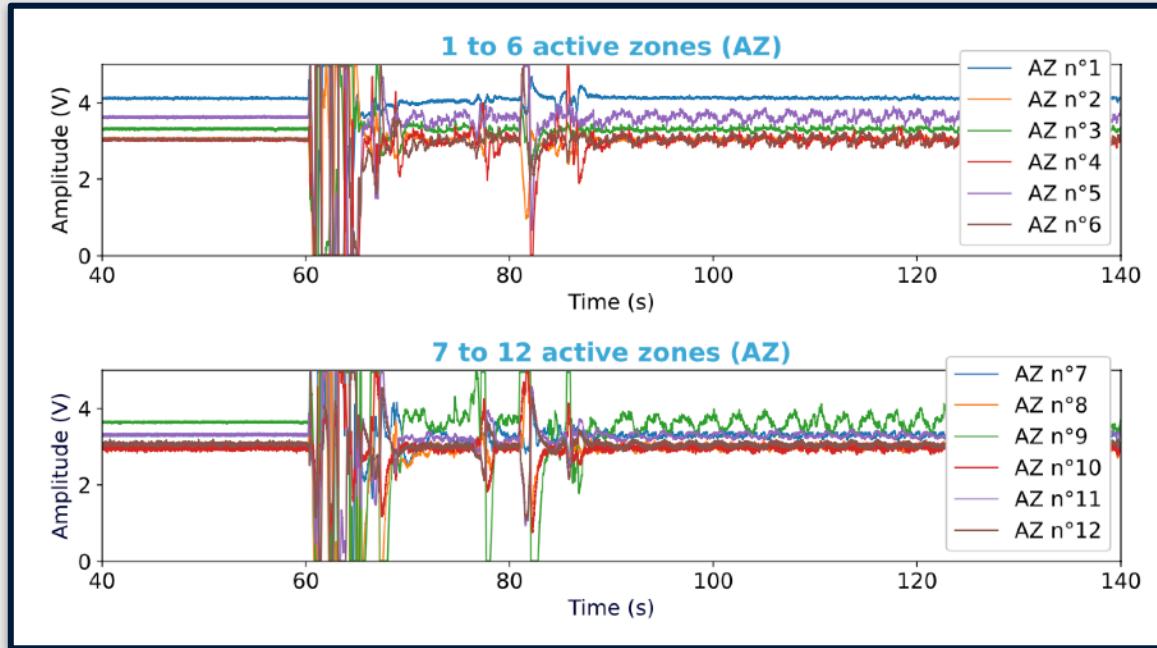
Exemple



La Ballistocardiographie [2]

[2] C. Brüser et al. Adaptive Beat-to-Beat Heart Rate Estimation in Ballistocardiograms. IEEE Engineering in Medicine and Biology Society (March 2011)

Mesure de paramètres vitaux



- Peu de données
- Beaucoup de features
- Grande variation entre les signaux

=> Utilisation d'algorithme d'IA

Détection du stress

- Environ 4000 instances
- 15 features (donnée tabulaire)

=> Utilisation de modèle de ML

Détection des cycle du sommeil

- Environ 50000 instances
- Timeseries à plusieurs dimensions: shape de (20, 30)

=> Utilisation de petit réseau de neurones

Le métier de Data Scientist

Conclusion: si vous voulez être data scientist « appliqué »

- **Etre très curieux:** vous devez apprendre des compétences transverses (domaine d'application)
- **Avoir un esprit critique:** chaque modèle a des limitations, il faut les déceler !
- **Aimer la recherche:** un data scientist doit constamment lire des papiers pour trouver les modèles adaptés

Tips: Lors d'un entretien d'embauche, toujours demander la quantité de données en leur possession

Sommaire

1: Présentation du métier de Data Scientist

2: Présentation du TP

3: L'Etat de l'art

4: L'optimisation d'Hyperparamètres

5: Modèles embarqués

Présentation du projet

Se familiariser avec le métier de Data Scientists, plus particulièrement dans un environnement contraint (système embarqué)

- Le TP va se porter sur un sujet concret: la reconnaissance vocal
- Utilisation de la base de donnée *the TensorFlow Speech Recognition Challenge* réalisé par google.
- Le Rapport devra être fait sur le jupyter notebook du projet
- Il est recommandé d'utiliser google colab durant le tp (nous serons pas responsable des conflits de librairies)
- Ce projet a pour but d'évaluer votre esprit critique, votre analyse d'un problème et votre regard critique sur vos performances: **Le but n'est pas d'avoir les meilleures performances mais comprendre pourquoi vous les avez !**

Présentation du projet

1. L'Analyse de la base de donnée. (2h)
2. L'état de l'art. (2h)
3. Le Preprocessing. (3h)
4. L'apprentissage du modèle. (3h)
5. Les résultats et discussions. (3h)
6. L'allégement du modèle. (4h)
7. La conclusion. (0h15)

=> Lisez bien tous l'énoncé sur le Jupyter notebook

Sommaire

1: Présentation du métier de Data Scientist

2: Présentation du TP

3: L'Etat de l'art

4: L'optimisation d'Hyperparamètres

5: Modèles embarqués

Etat de l'art

- Utiliser des outils de gestion de références (Zotero, Mendeley, EndNote, ...)
- Définir précisément le périmètre: Identifiez des mots-clés spécifiques et généraux pour guider votre recherche
- Priorisez des sources fiables comme Google Scholar, IEEE Xplore, JSTOR, PubMed, ou ScienceDirect. Ne pas oublier les brevets (google Patent).
- Classez les articles par thèmes, méthodologies ou concepts similaires pour une vision structurée (ex. : méthodes, résultats, contextes d'application).
- Comparez les approches, points forts/faiblesses et résultats des études. Cela enrichit votre compréhension et ajoute de la profondeur.
- **Évaluation critique** : Identifiez les limites des études (ex. : biais dans les données, limites des modèles) pour mieux situer votre propre recherche.

Etat de l'art - Quoi lire en première lecture ?

- 1. Titre et résumé:** vérifier que l'article est pertinent.
- 2. Introduction:** comprendre le contexte, la problématique, les objectifs et les motivations du papier.
- 3. Contributions principales** (fin d'introduction ou début de conclusion): comprendre en quoi l'approche est originale par rapport aux recherches existantes.
- 4. Conclusion:** identifiez les limitations et les suggestions de recherches futures, qui peuvent inspirer ou orienter de futur travaux.
- 5. Etat de l'art:** certains papier cité peuvent être plus pertinent que celui-ci
- 6. Figure et tableaux:** une première lecture visuelle des graphiques peut souvent clarifier les résultats principaux.
- 7. Méthodologie (en survol):** comprendre l'approche générale, les données ou techniques principales

Sommaire

1: Présentation du métier de Data Scientist

2: Présentation du TP

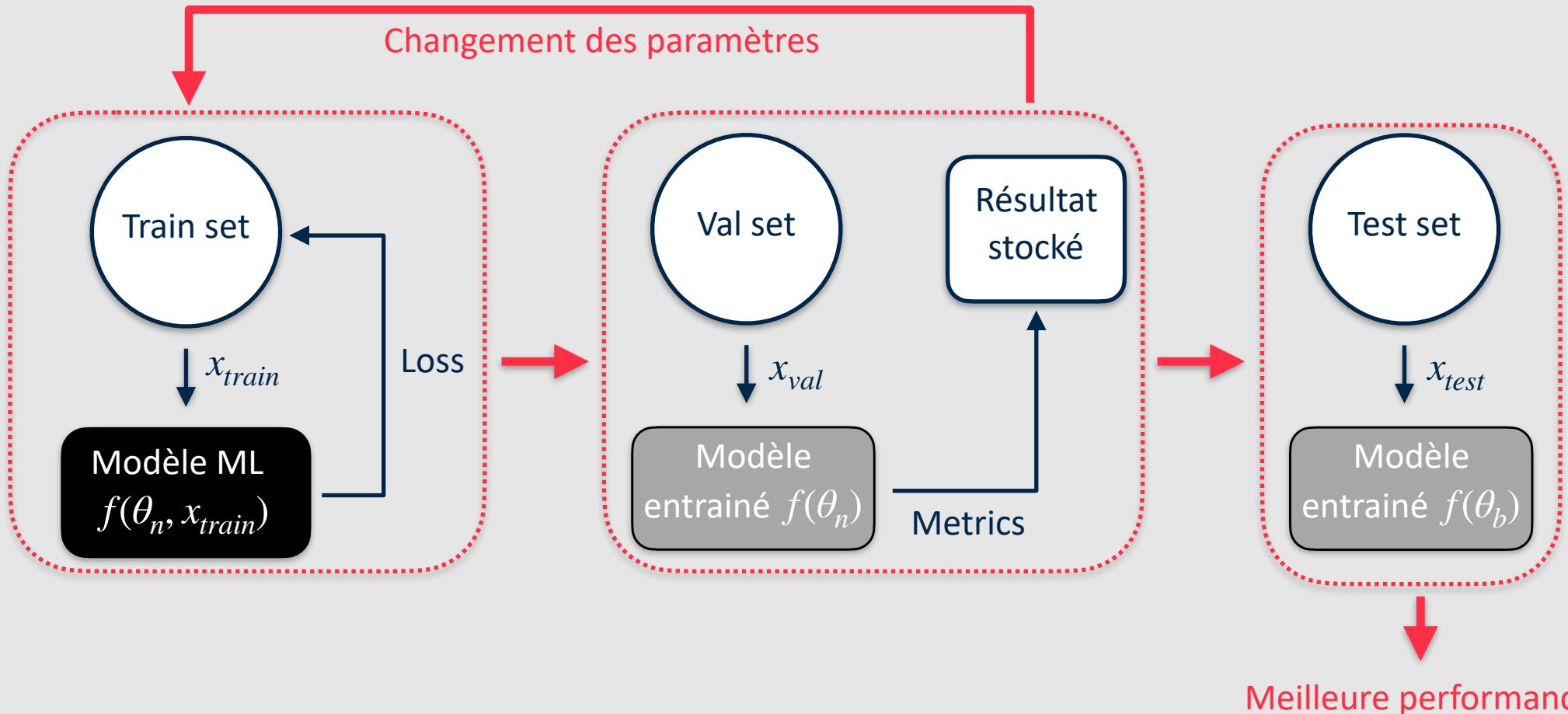
3: L'Etat de l'art

4: L'optimisation d'Hyperparamètres

5: Modèles embarqués

L'optimisation d'Hyperparamètres

Soit θ_n , le n ième jeu de paramètre du modèle (poids, learning rate, nombre de couches, ...) et θ_b , le meilleur jeu de paramètre en fonction de la métrique choisie (accuracy, MSE, F1 score, ...).

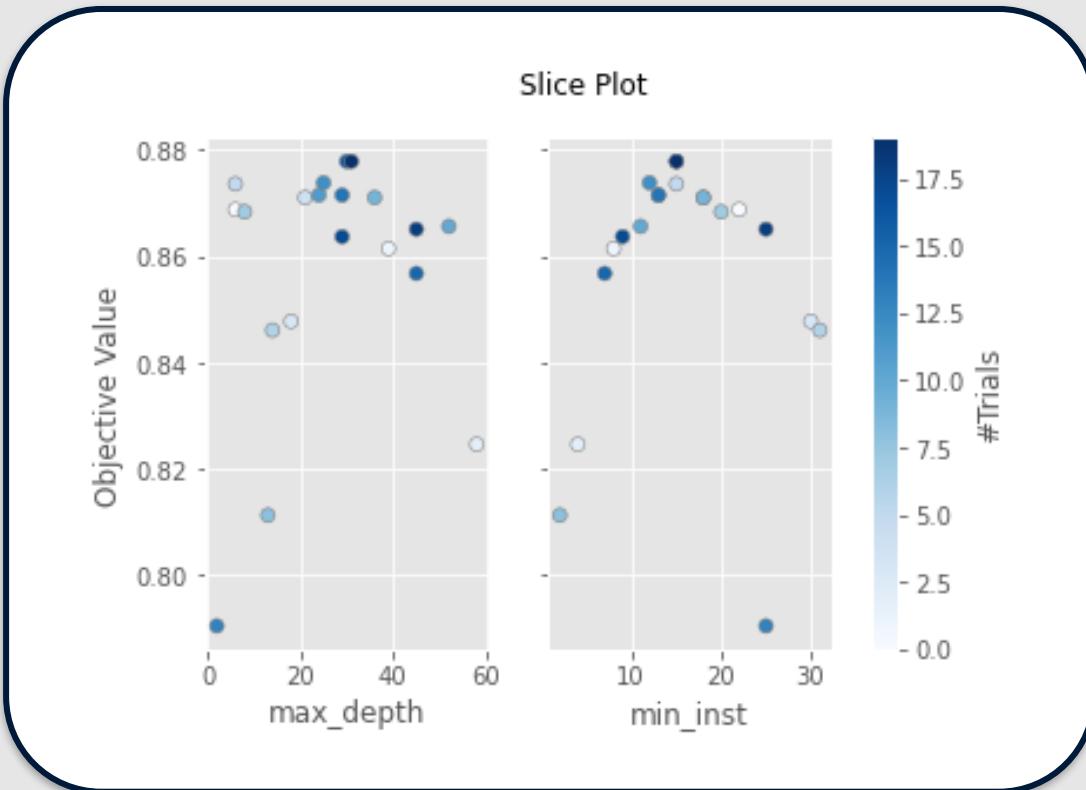


L'optimisation d'Hyperparamètres

Optuna

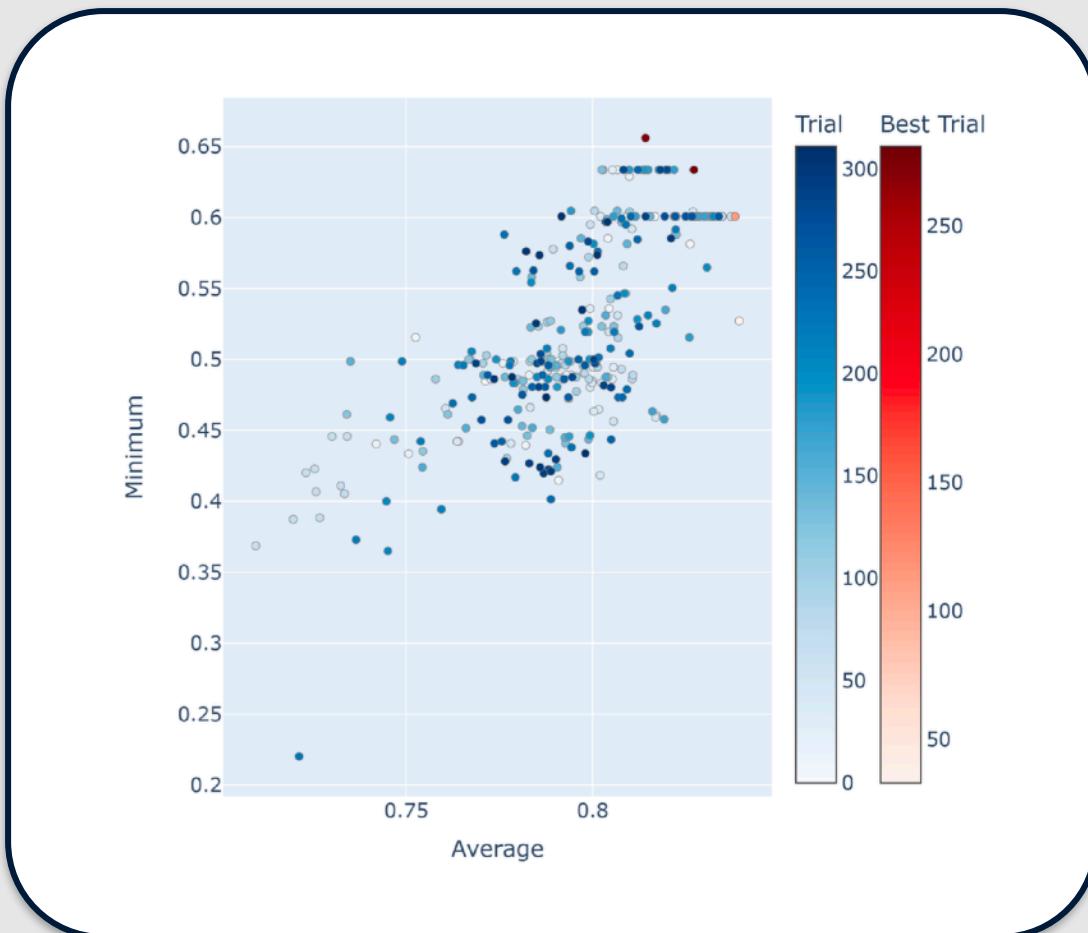
- Aide à tester automatiquement différentes combinaisons d'hyperparamètres, en utilisant une méthode intelligente pour choisir les meilleures options sans tout tester.
- Optuna explore différentes combinaisons d'hyperparamètres, apprend de chaque essai et se concentre sur les combinaisons prometteuses pour économiser du temps et des ressources.
- Il est possible d'optimiser plusieurs métriques simultanément (comme la précision, le temps d'exécution ou la consommation mémoire), en cherchant un compromis entre elles pour répondre aux besoins spécifiques du projet.
- Après optimisation, Optuna propose les configurations qui maximisent les performances selon les différentes métriques, permettant à l'utilisateur de choisir celle qui répond le mieux à ses objectifs (comme un compromis entre précision et temps d'exécution).

L'optimisation d'Hyperparamètres



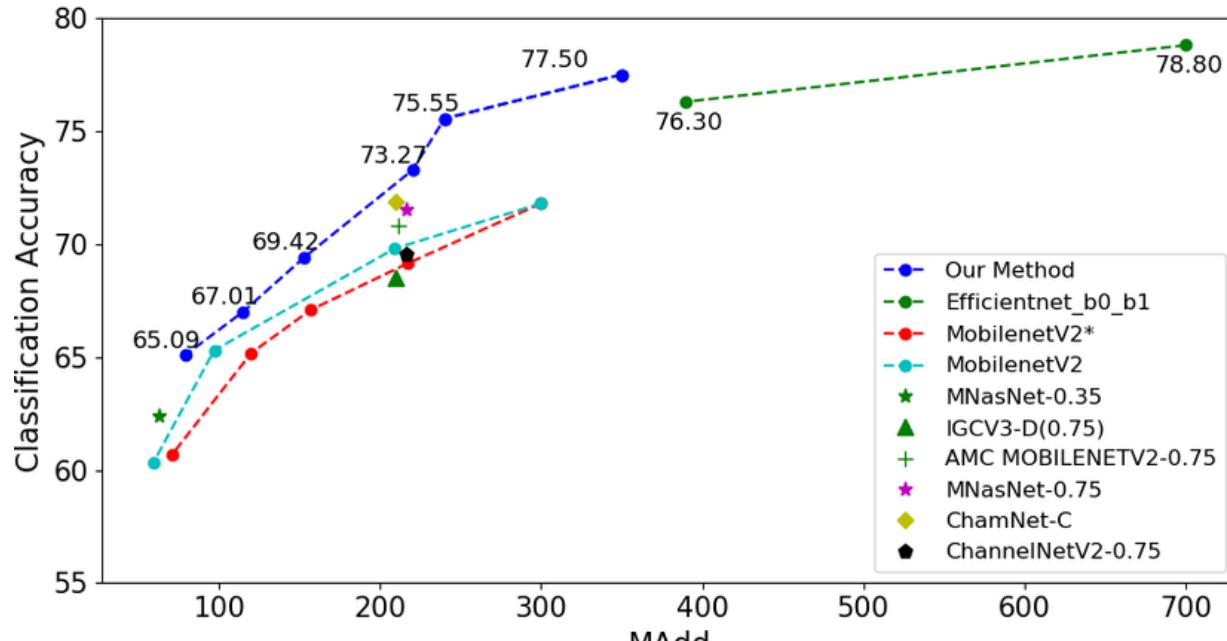
Nuage de performances en fonction des hyperparamètres

L'optimisation d'Hyperparamètres



Nuage de combinaisons possibles

L'optimisation d'hyperparamètres



Frontière de Pareto des différents type d'architecture [3]

[3] Daquan, Zhou & Jin, Xiaojie & Wang, Kaixin & Yang, Jianchao & Feng, Jiashi. (2019). Deep Model Compression via Filter Auto-sampling. 10.48550/arXiv.1907.05642.

Sommaire

1: Présentation du métier de Data Scientist

2: Présentation du TP

3: L'Etat de l'art

4: L'optimisation d'Hyperparamètres

5: Modèles embarqués

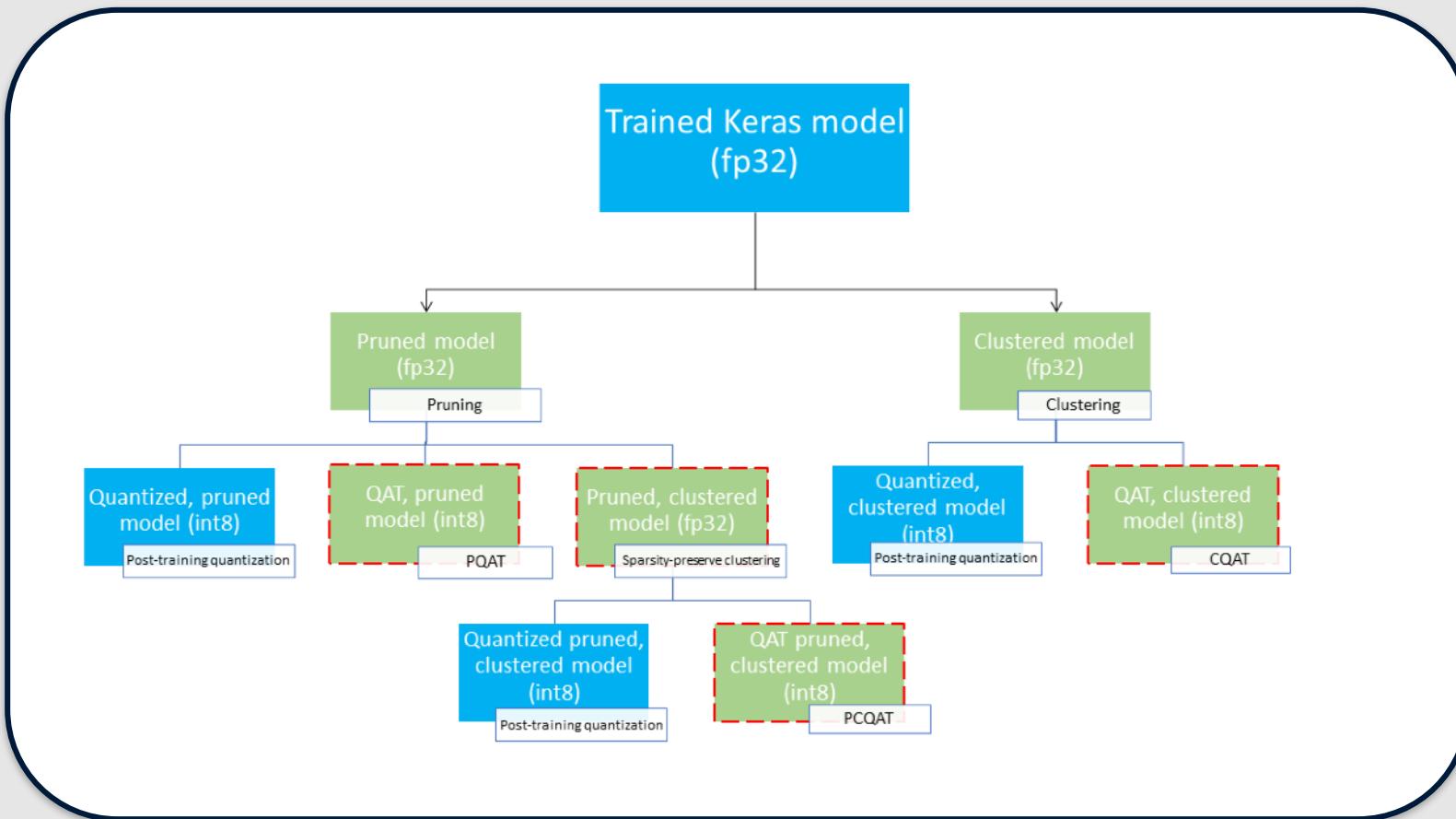
L'optimisation d'Hyperparamètres

- **La quantification:** Réduire le nombre d'octets des paramètres du modèles et les convertir en entiers.
- **L'élagage:** Supprimer les modèles qui ont un impact mineur sur les prédictions.
- **Le clustering:** Regrouper les poids de chaque couche dans un modèle en un nombre prédéfini de clusters, puis en partageant les valeurs centroïdes pour les poids appartenant à chaque cluster individuel.
- **La décimation :** Réduire le nombre d'entrée.

L'optimisation d'Hyperparamètres

- **La quantification:** Réduire le nombre d'octets des paramètres du modèles et les convertir en entiers.
- **L'élagage (pruning):** Supprimer les modèles qui ont un impact mineur sur les prédictions.
- **Le clustering:** Regrouper les poids de chaque couche dans un modèle en un nombre prédéfini de clusters, puis en partageant les valeurs centroïdes pour les poids appartenant à chaque cluster individuel.
- **La décimation :** Réduire le nombre d'entrée.

Optimisation d'hyperparamètres



Optimisation collaborative après entraînement [4]

[4] https://www.tensorflow.org/model_optimization/guide/combine/collaborative_optimization?hl=fr

Thank you for your attention