

Hackathon

Approches d'apprentissage machine multimodal
pour prédire l'intensité de l'activité physique
perçue par les athlètes professionnels.

ANIZON Thibault
BRIVARY Guillaume
MARANDIN Tristan
MAUGER Mika
TROGNON Jean-Baptiste
WITKOWICZ Nathan



SOMMAIRE



01

INTRODUCTION

02

ETAT DE L'ART

03

DATASET ET PRETRAITEMENT

04

ARCHITECTURE

05

ENTRAINEMENT

06

CONCLUSION

INTRODUCTION



Objectif

Développer des modèles prédictifs du Taux de Perception de l'Effort (RPE) chez les footballeurs professionnels



RPE : une métrique subjective qui évalue la charge et l'intensité de l'entraînement chez les athlètes.

But du projet : Prédire le RPE grâce aux données des équipes de Montpellier.

Les données : des données **anthropométriques**, des données **GPS et accéléromètre**, les **conditions météorologiques**, et les **scores historiques de RPE**.

RPE SCALE	RATE OF PERCEIVED EXERTION
10 /	MAX EFFORT ACTIVITY Feels almost impossible to keep going. Completely out of breath, unable to talk. Cannot maintain for more than a very short time.
9 /	VERY HARD ACTIVITY Very difficult to maintain exercise intensity. Can barely breathe and speak only a few words.
7-8 /	VIGOROUS ACTIVITY Borderline uncomfortable. Short of breath, can speak a sentence.
4-6 /	MODERATE ACTIVITY Breathing heavily, can hold a short conversation. Still somewhat comfortable, but becoming noticeably more challenging.
2-3 /	LIGHT ACTIVITY Feels like you can maintain for hours. Easy to breathe and carry a conversation.
1 /	VERY LIGHT ACTIVITY Hardly any exertion, but more than sleeping, watching TV, etc.

ETAT DE L'ART

JEUX DE DONNÉES DIVERS

Intégration multi-sources
(GPS, biométrie, météo).

MACHINE LEARNING AVANCÉ

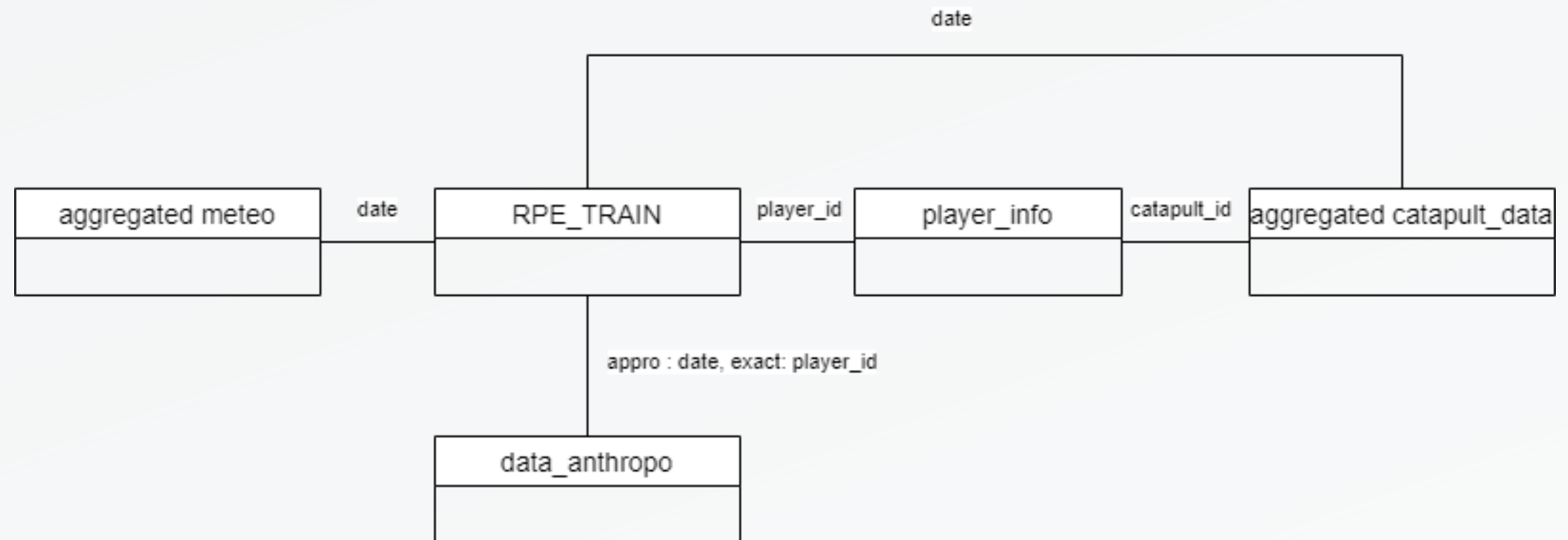
- RNN : Réseaux de neurones récurrents, adaptés aux données séquentielles.
- LSTM : Version avancée des RNN, meilleure gestion de la mémoire à long terme.

KIM, J. ET AL. (2022) 'A DEEP LEARNING APPROACH FOR FATIGUE PREDICTION IN SPORTS USING GPS DATA AND RATE OF PERCEIVED EXERTION

DATASET ET PRE-TRAITEMENT

Problématiques :

- Données réparties dans différentes tables -> jointures
- Différentes échelles de temps -> agrégations
- Valeurs manquantes : compléter avec moyennes, zéros, valeurs proches selon le cas





RANDOM

FOREST

RESULTAT RANDOM FOREST

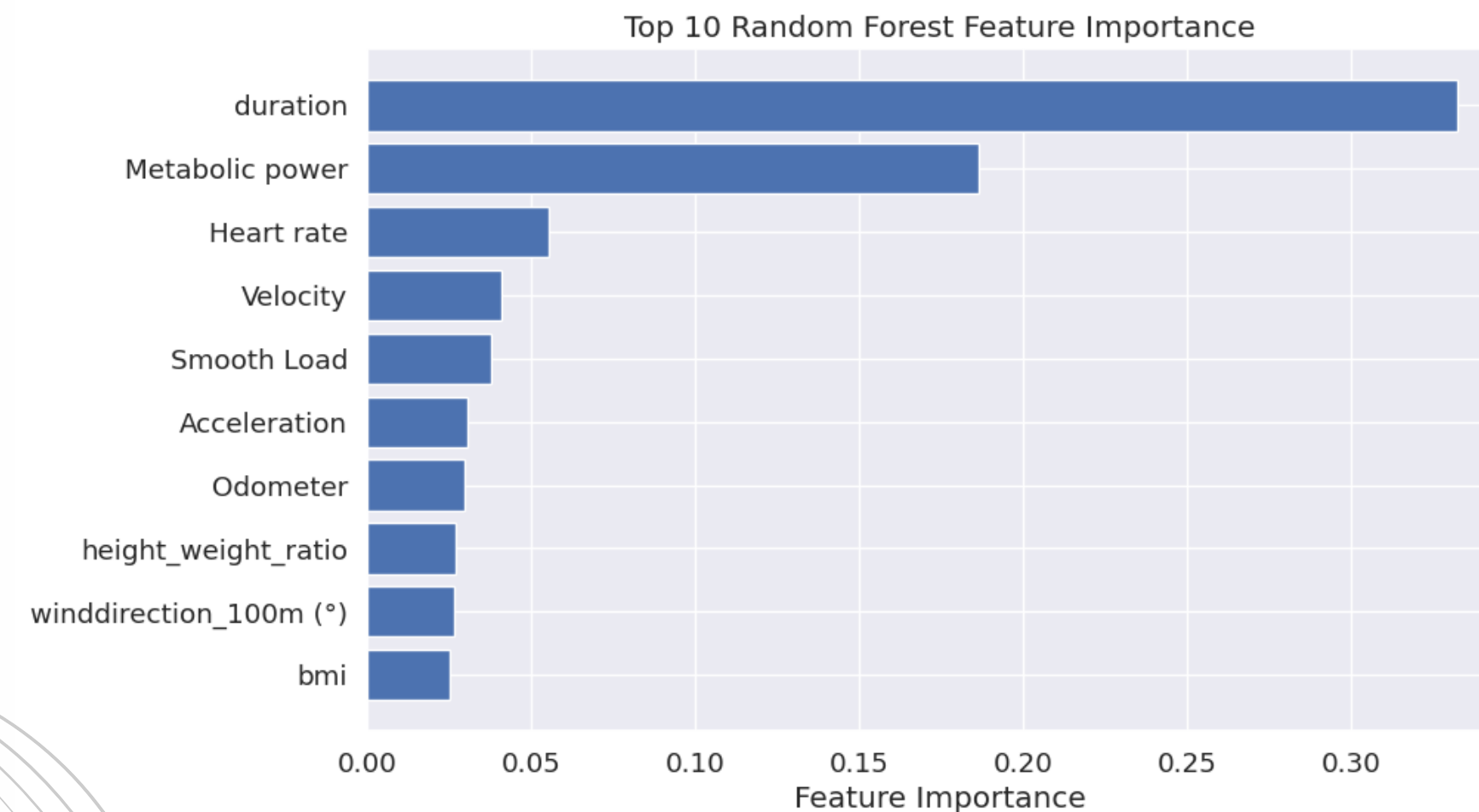
Indicateurs

Les 10 indicateurs les plus importants selon Random Forest :

- Pas mal de Nan pour la duration
- *BMI* et *Height Weight Ratio* très corrélés.

Résultats

- MAE de 0.92 : Précis malgré les valeurs manquantes.
- Permet d'affiner le modèle de prévision du RPE.



Résultats avec Random Forest

Méthode d'évaluation

Calcul de l'Erreur Absolue Moyenne (MAE) pour fixer une référence.

Nos résultats

- Modèle Random Forest
MAE = 0.92
- Modèle Random Forest
Joueur par Joueur
MAE = 0.94



FATIGUE

NETWORK

Données joueur-météo

Retirer les dates incohérentes

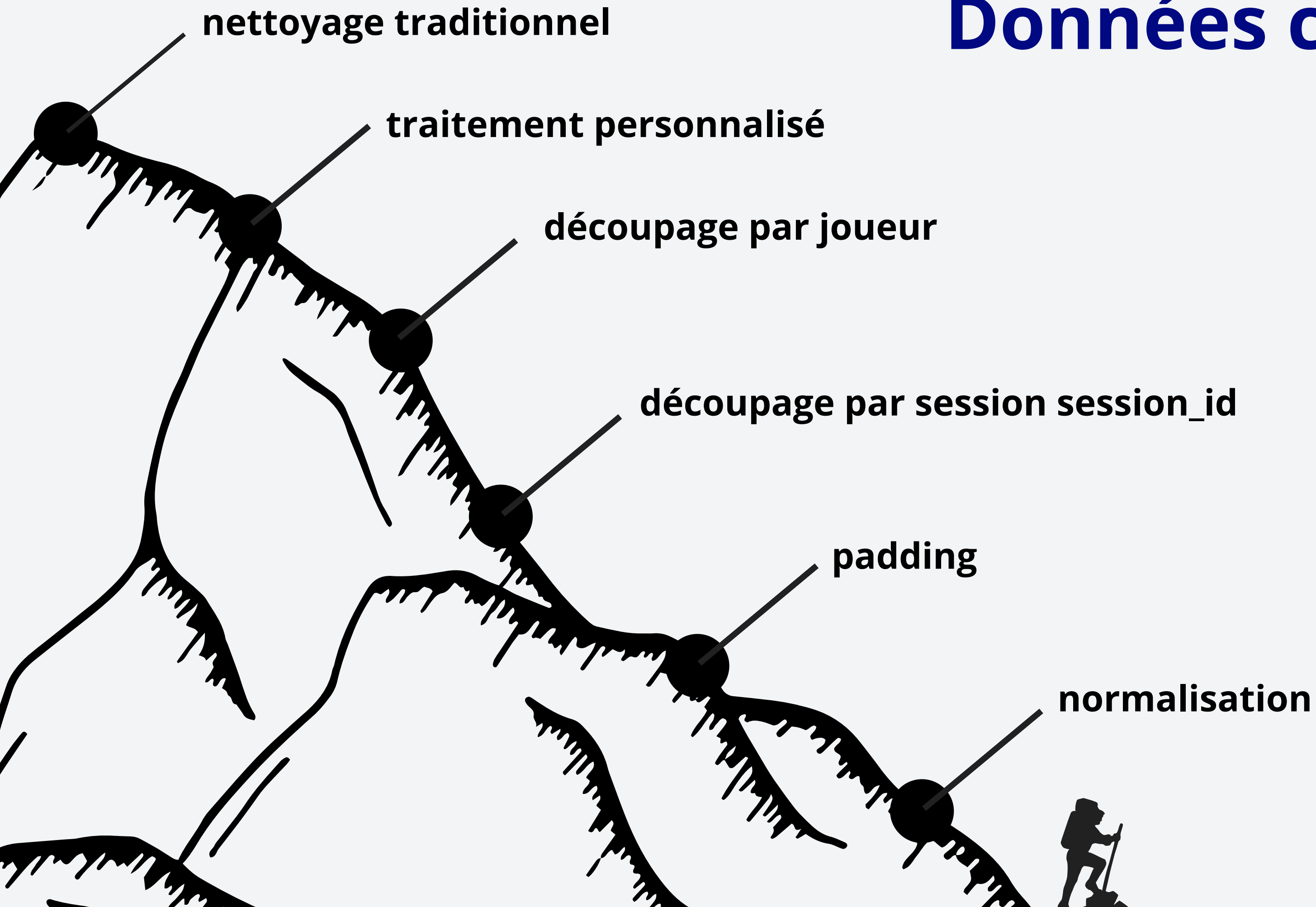
Split par player_id

Normalisation

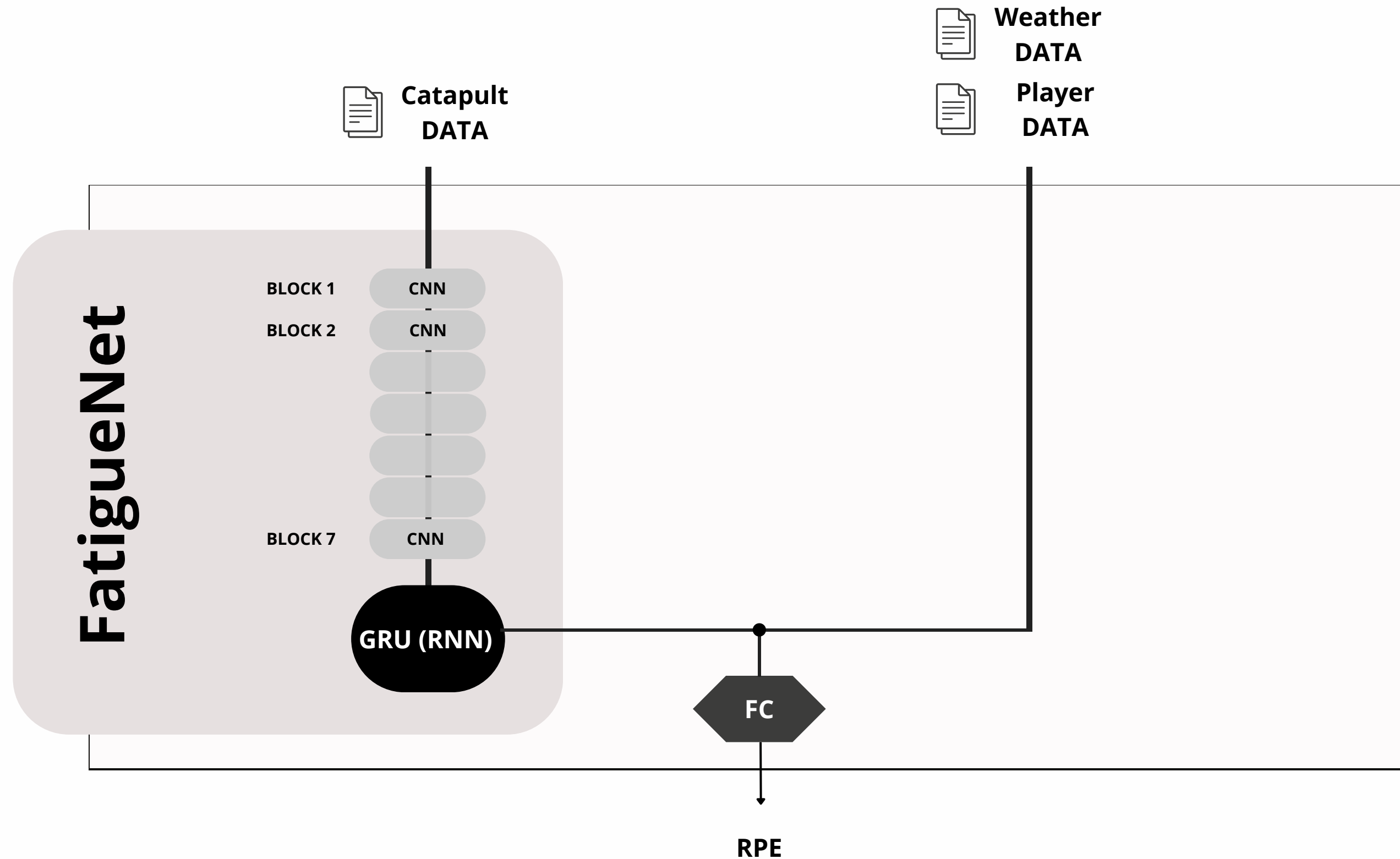
Merging



Données catapultes



Architecture Deep Learning



Résultats avec FatigueNet

Résultats de J. Kim et al.

- Modèle Random Forest
MAE = 1.0537
- Modèle FatigueNet
MAE = 0.8494

Nos résultats

- Modèle Random Forest
MAE = 0.92
- Modèle par joueur
MAE = 1.7521
- Modèle global
MAE = 3.1936

DISCUSSION

Comparaison des résultats

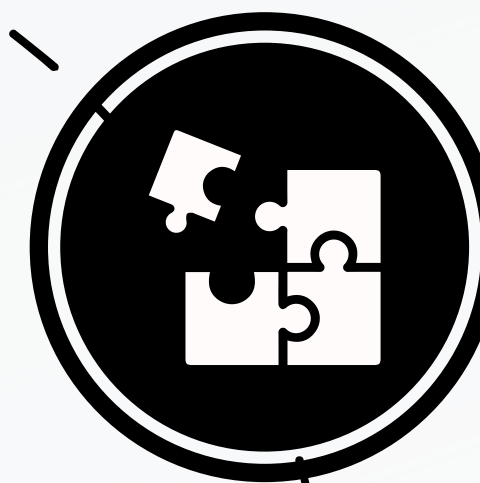
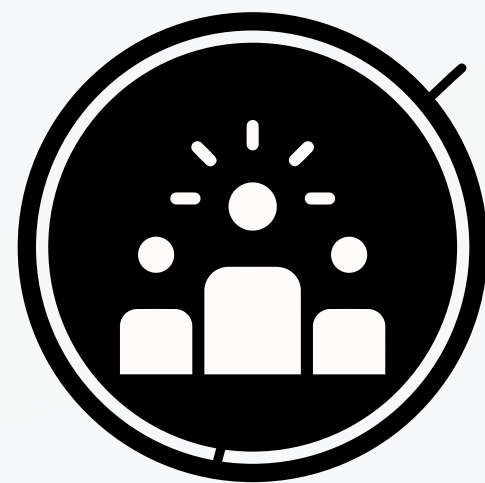
- **Scores honorables** au vu du peu de données
- **Incohérent** avec la littérature

Discussion

Faible traitement des données **météorologiques** et relatives aux **joueurs**

Pistes d'amélioration

- Donner plus de **pertinence** aux **features** de **catapult**
- Ajouter **une couche de RNN** prenant en compte les sessions précédentes
- Ajout de features pertinentes (**variables cycliques**)



Conclusion

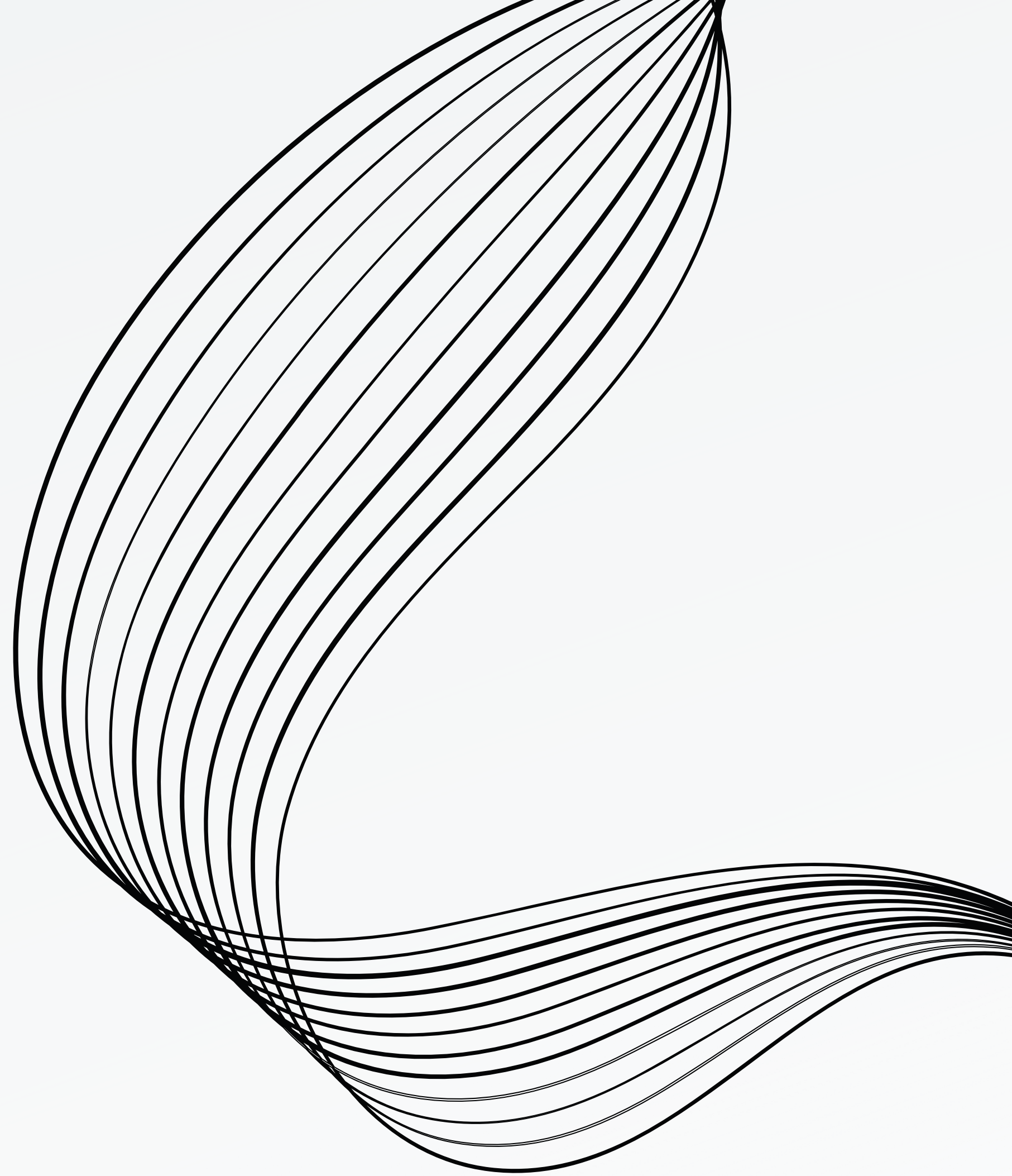
Expérience positive

- Appréhension de problème complexes
- Travail en groupe multi-compétences
- Gestion d'équipe
- Développement de nouvelles compétences

Futurs travaux

- Collecte de data par session
- Prendre en compte les pistes d'amélioration

**MERCI POUR
VOTRE ECOUTE**



Annexe 1 : RNNs - GRUs

GRUs (Gated Recurrent Units) :

- Similaire aux LSTMs mais avec une structure simplifiée, ce qui rend les GRUs moins onéreux en termes de calcul.
- Combine les portes d'entrée et de sortie en une seule porte de mise à jour, simplifiant le modèle tout en conservant une performance robuste.

Unité récurrente fermée,
version entièrement fermée

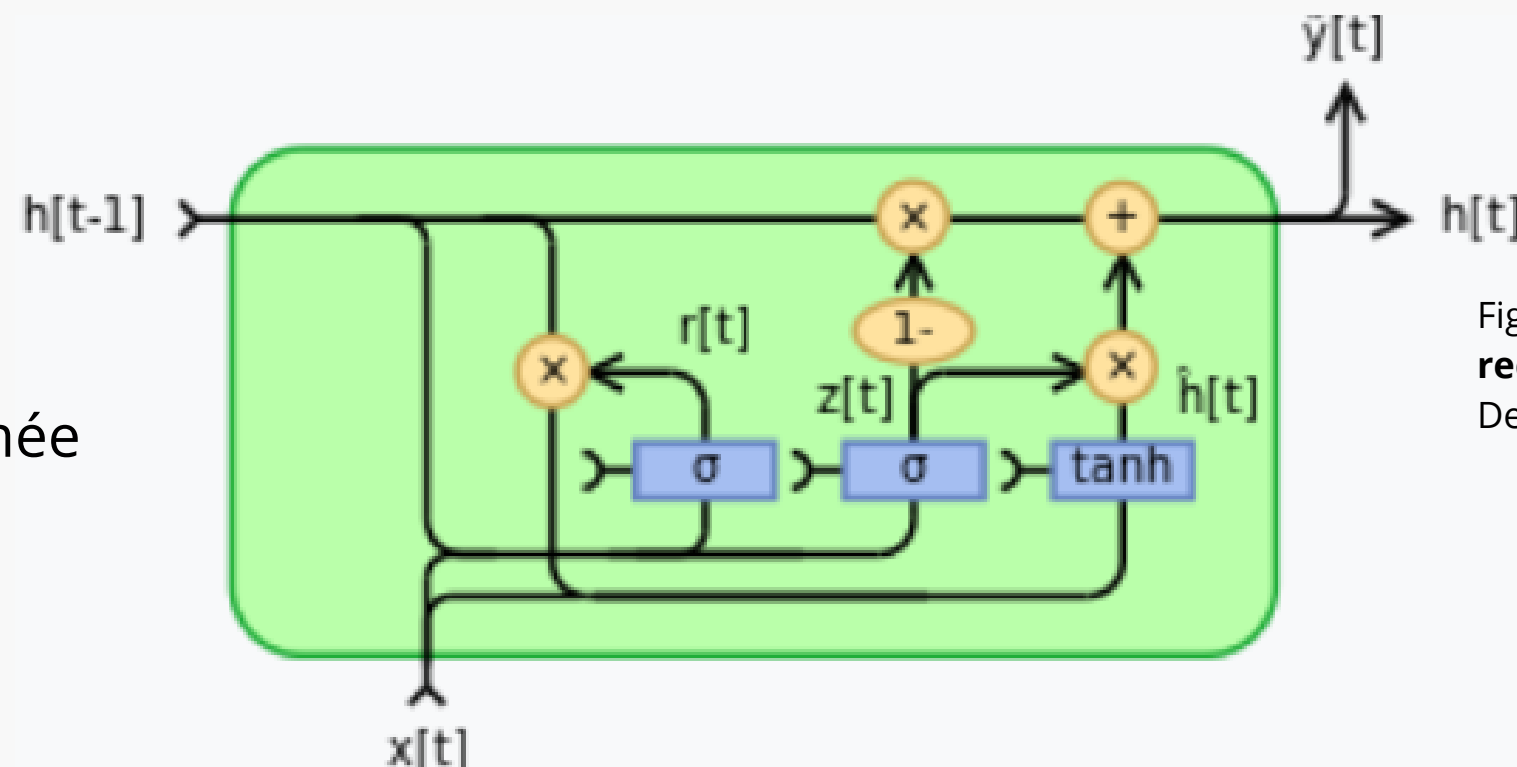


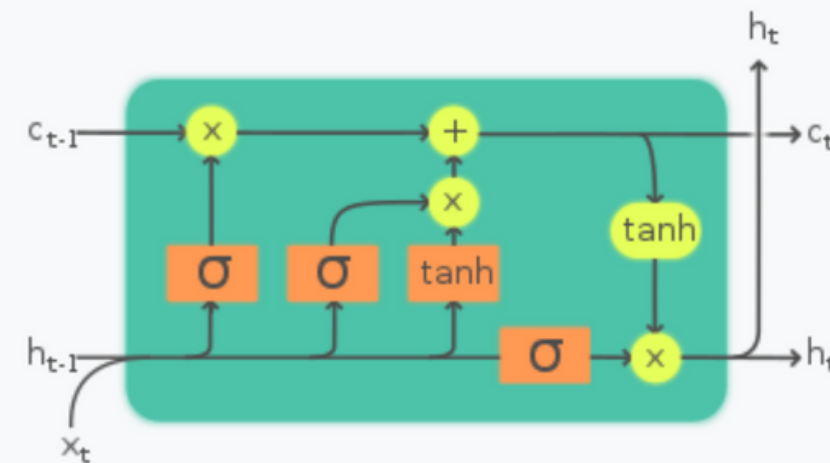
Figure from : Wikipedia contributors. (2023, October 27). **Gated recurrent unit**. In Wikipedia, The Free Encyclopedia. Retrieved 10:04, December 20, 2023, from [Gated recurrent unit Wikipedia page](#)

Annexe 2 : RNNs - LSTMs

LSTMs (Long Short-Term Memory) :

- Variante avancée des RNNs, spécialement conçue pour pallier le problème de la disparition du gradient.
- Comprend des portes qui régulent le flux d'informations, permettant la conservation sur de longues périodes et l'oubli sélectif.

La cellule peut traiter les données de manière séquentielle et maintenir son état caché dans le temps.



Legend: Layer ComponentwiseCopy Concatenate

Figure from : Wikipedia contributors. (2023, November 17). **Long short-term memory**. In Wikipedia, The Free Encyclopedia. Retrieved 10:10, December 20, 2023, from [Long short-term memory Wikipedia page](#)

Appendice 3 : qsdf