

TacStic: Analyse des situations de jeu au football en vue d'une classification pour l'évaluation et l'entraînement de la perception du footballeur.



Figure 1 - Outil de réalité virtuelle du laboratoire m2s

- Nawfel Seghrouchni – Stagiaire DIGISPORT
- Alexandre Vu – Doctorant au laboratoire M2S Mouvement, Sport, Santé (Univ. Rennes 2, ENS Rennes)

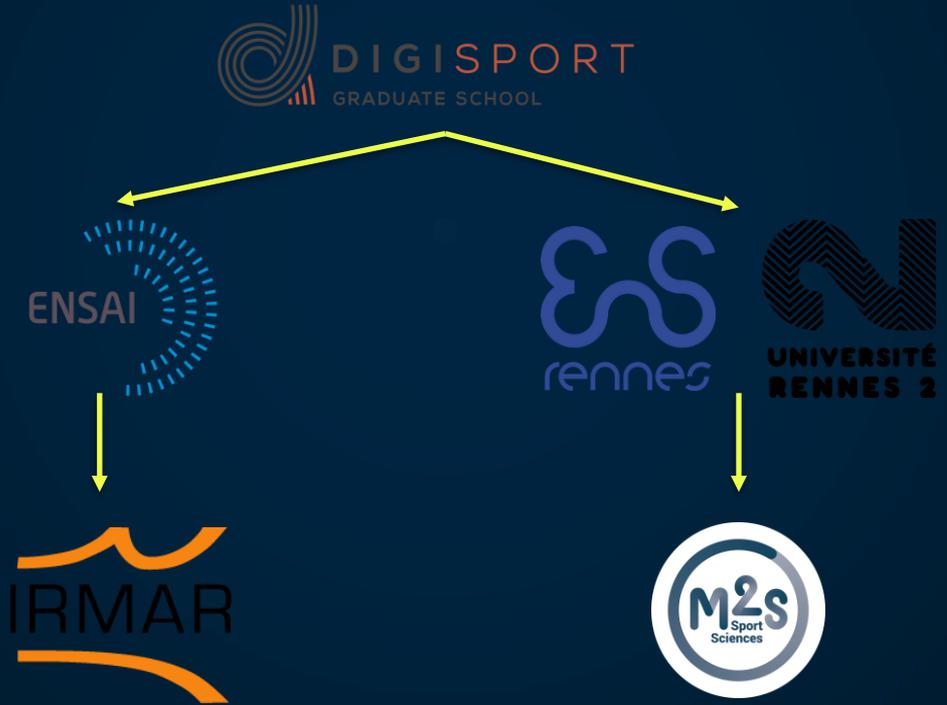
Tuteurs :

- Brigitte Gelein - Enseignante en statistique à l'Ensaï, Chercheuse IRMAR (UMR CNRS 6074)
- Anthony Sorel - Ingénieur chercheur au laboratoire M2S (Univ. Rennes 2, ENS Rennes)

Mars → Septembre 2021



STRUCTURE D'ACCEUIL



SITUATION



Figure 2 - Team Training LOSC vs SRFC Ligue
1 Uber Eats, Round 1, Saturday 22 August
2020 source L'Équipe

- Match LOSC vs SRFC
- 4-2-3-1 ?
- Comment moderniser la compréhension du football ?

OBJECTIF :

- Proposer une méthodologie pour classifier ces situations de jeu par similarité.

DONNEES DE TRACKING

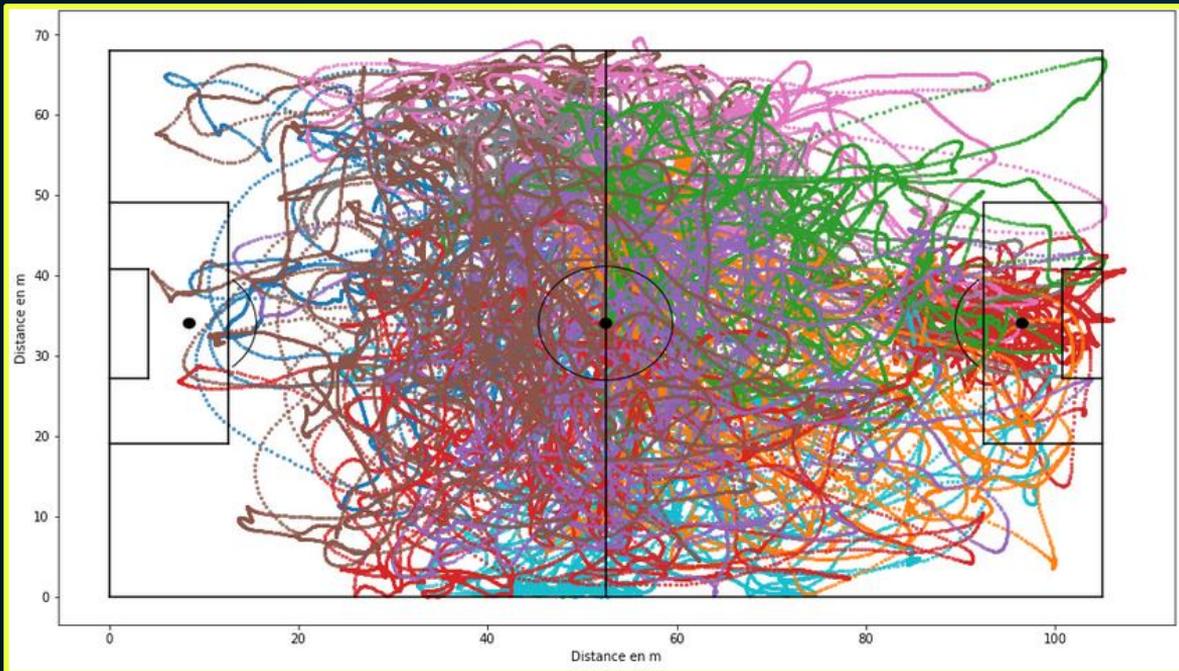


Figure 3 – Traces des joueurs sur 18 min du match.

- Amélioration constante des systèmes de capteurs.
- Position (x_i, y_i) des joueurs et du ballon toutes les 0.1 secondes. (10 Hz)
- Séries temporelles multivariées.
- Difficile d'interpréter, manipuler les données brutes → un pré processing est nécessaire.

VISUALISATION DES DONNEES



Figure 4 – Vidéo but LOSC image canal+.

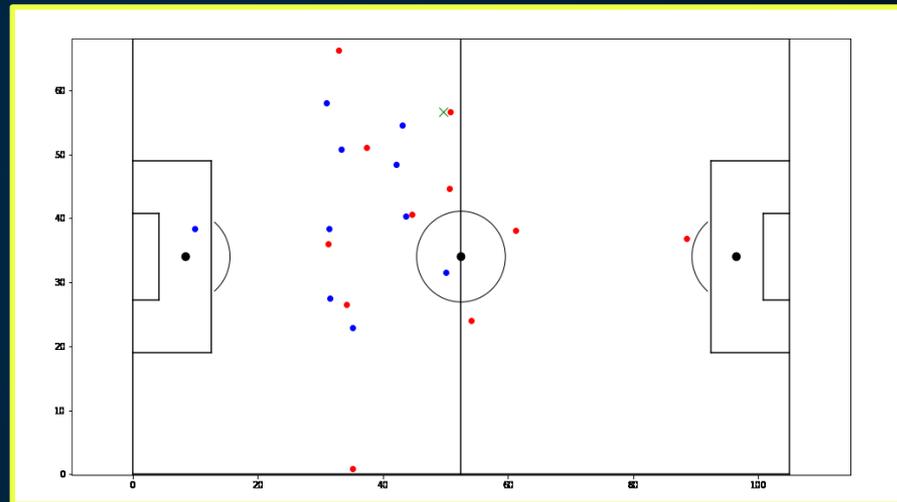


Figure 5 – Gif du but du LOSC en représentation 2D.

CREATION DE DESCRIPTEURS



ZONE D'ACTION

Permet d'avoir une représentation de l'architecture d'un bloc d'équipe.



DYNAMIQUE DES JOUEURS

Fournit des informations sur la vitesse d'une action. (contre-attaque, utilisation de tactiques,... etc.)



RÔLE DES JOUEURS

Permet d'anticiper les changements de rôle et d'organiser notre jeu de données.

VITESSE DES JOUEURS

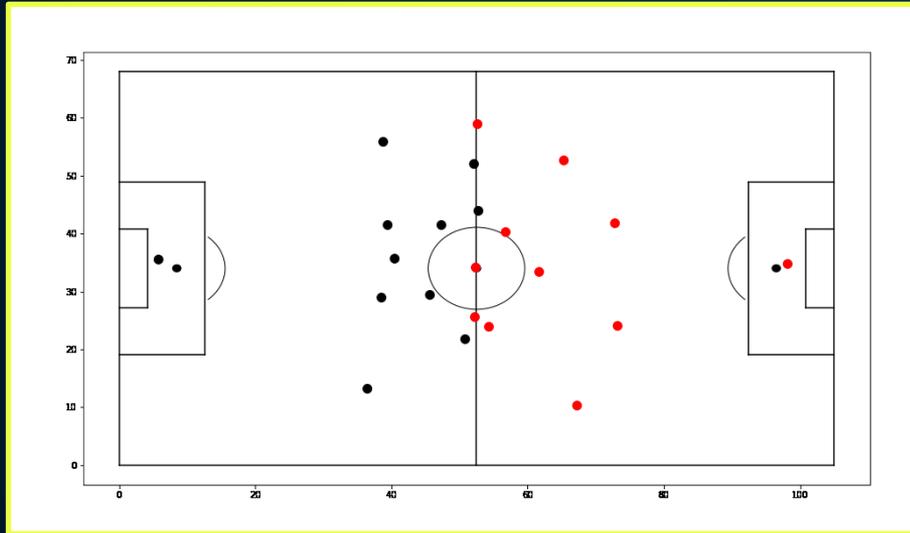


Figure 9 – Gif des représentations des vitesses de chaque joueur, représentation 2D.

- Vitesse des joueurs entre deux prises de position.
- Fournit un indicateur de l'orientation d'un joueur.
- Fournit un appui visuel.
- Indicateur de la vitesse d'une action et de la façon de se projeter.

ZONE D'ACTION

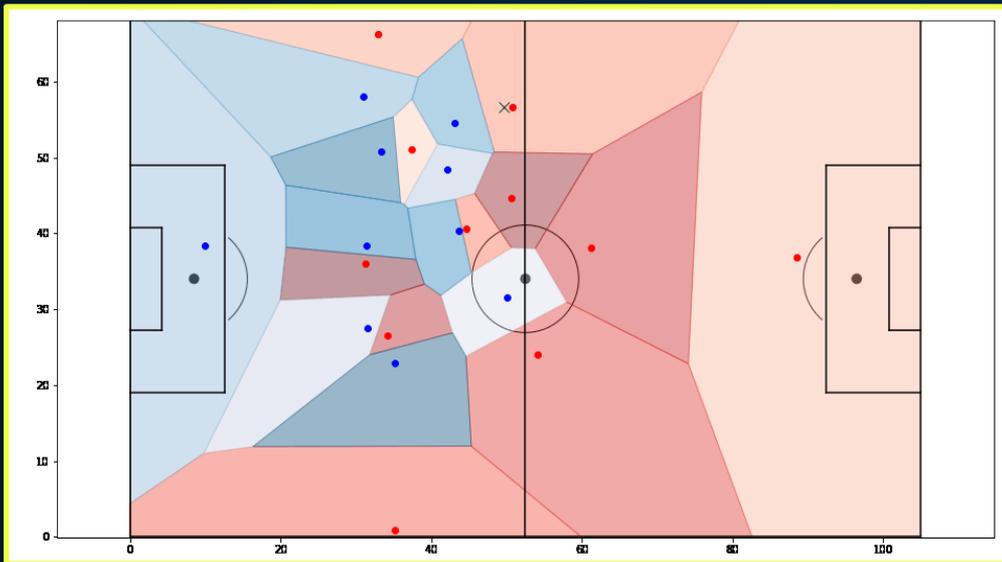


Figure 6 – Gif des cellules de Voronoï sur le but, représentation 2D.

- « Zone d'influence » du joueur.
- La frontière se situe sur la médiatrice qui sépare deux joueurs.
- Une équipe dans le camp adverse couvrira une plus grande surface.
- La vitesse et l'orientation d'un joueur ne sont pas prises en compte dans le calcul de ces zones.

ZONE D'ACTION

	Away_10 _x	Away_11 _x	Away_14 _x	Away_15 _x	Away_16 _x	Away_17 _x	Away_20 _x	Away_21 _x	Away_22 _x	Away_28 _x
1,6E+12	Outside	58,13	73,33	77,13	1102,89	590,12	Outside	102,07	Outside	66,93
1,6E+12	Outside	56,93	75,1	73,87	1102,57	588,41	Outside	103,96	Outside	65,77
1,6E+12	Outside	56,14	76,87	72,27	1102,96	586,28	Outside	105,58	Outside	64,4

Figure 7 – Jeu de données des surfaces de Voronoï par joueur.

ZONE D'ACTION

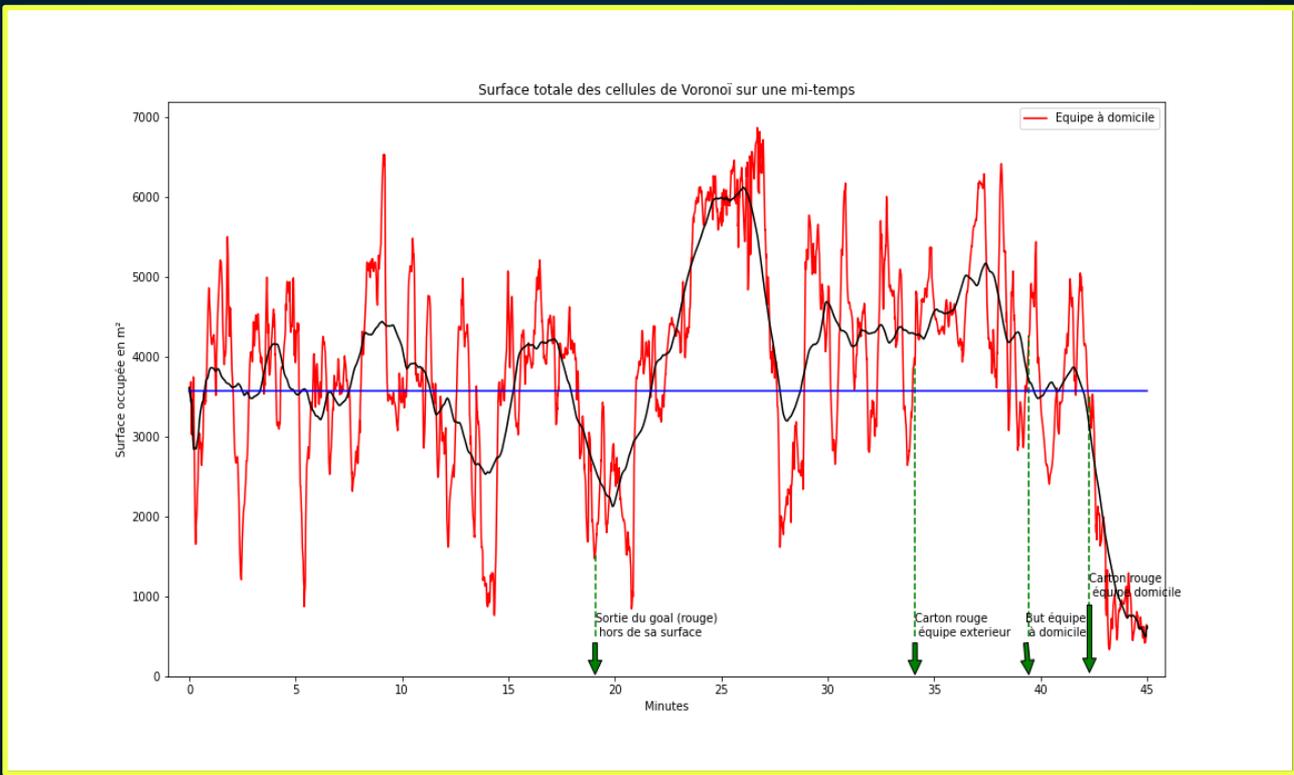
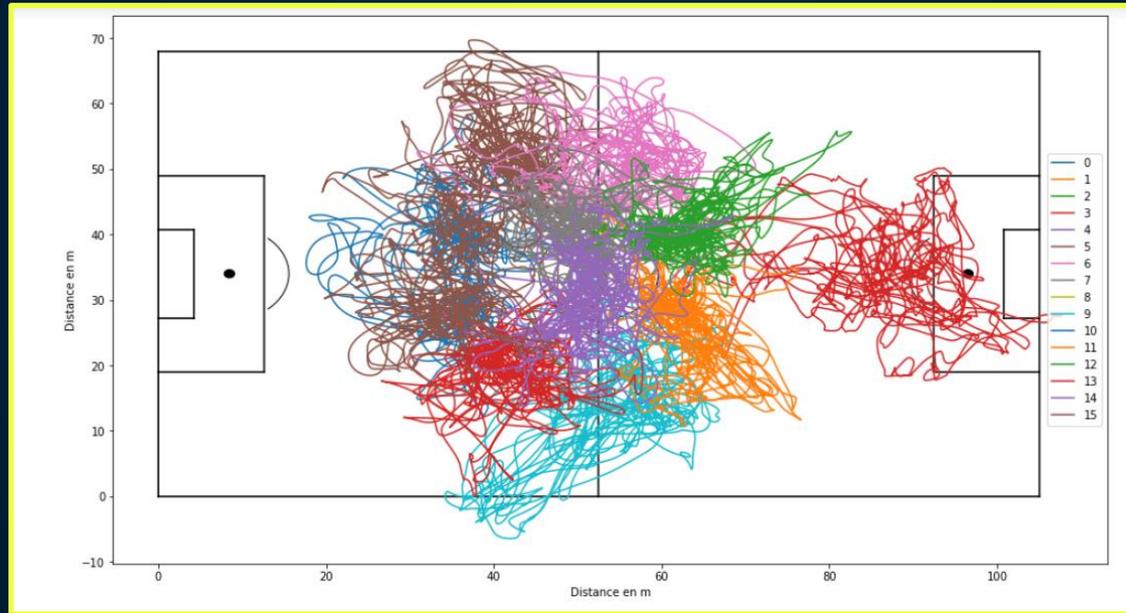


Figure 8 – Surface totale des cellules de Voronoï sur une mi-temps avec ajout de données événements.

RÔLE DES JOUEURS



Traces des joueurs sur 18 min du match centrées.

RÔLE DES JOUEURS

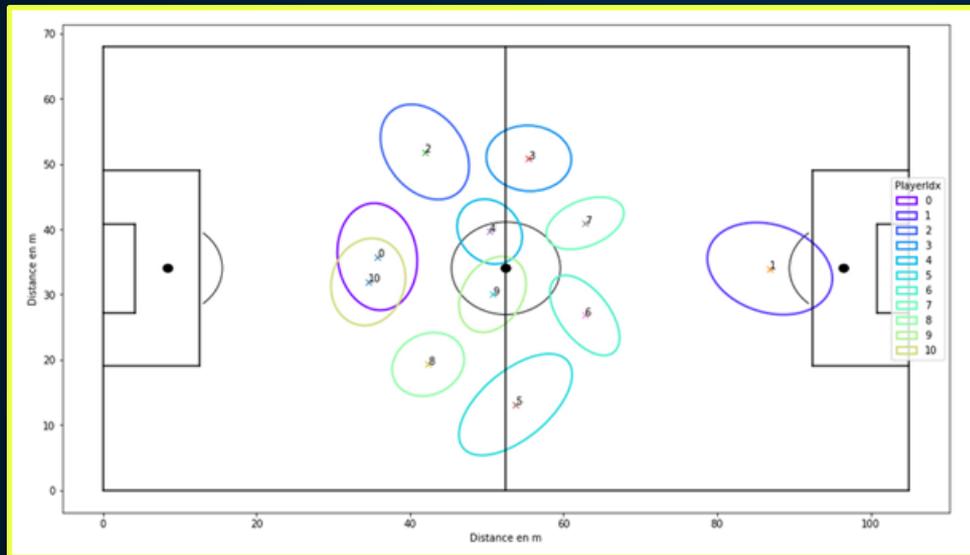


Figure 11 – Distribution elliptique des joueurs sur 18 min du match.

- On peut tracer les distributions des positions 2D des joueurs.
- Différentes façons de calculer la distribution.
- Privilégier les méthodes non paramétriques.
- OBJECTIF :

Avoir en temps réel le rôle associé au joueur à chaque instant du match.

RÔLE DES JOUEURS : utilisation d'un noyau gaussien

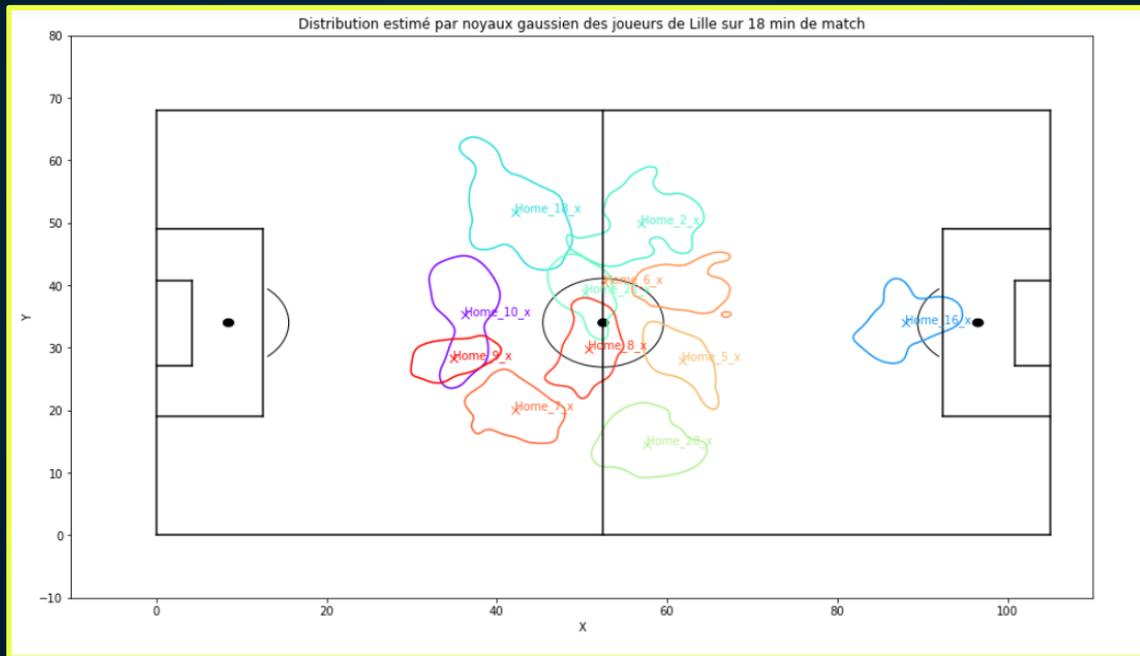


Figure 12 – Distribution estimée par noyau gaussien sur 18 min du match.

ATTRIBUTION DES RÔLES

```
7275
['rcb', 'lcb', 'rb', 'ram', 'rdm', 'rcf', 'lam', 'ldm', 'lb', 'lcf']
['rcb', 'lcb', 'rb', 'lb', 'rdm', 'rcf', 'lam', 'ldm', 'ram', 'lcf']

7290
['rcb', 'lcb', 'rb', 'lb', 'rdm', 'rcf', 'lam', 'ldm', 'ram', 'lcf']
['rcb', 'lcb', 'rb', 'ram', 'rdm', 'rcf', 'lam', 'ldm', 'lb', 'lcf']

8655
['rcb', 'lcb', 'rb', 'ram', 'rdm', 'rcf', 'lam', 'ldm', 'lb', 'lcf']
['lcf', 'lcb', 'rb', 'rcb', 'rdm', 'rcf', 'lam', 'lb', 'ram', 'ldm']

8670
['lcf', 'lcb', 'rb', 'rcb', 'rdm', 'rcf', 'lam', 'lb', 'ram', 'ldm']
['ldm', 'lcb', 'rb', 'lcf', 'rdm', 'rcf', 'lam', 'ram', 'rcb', 'lb']

8685
['ldm', 'lcb', 'rb', 'lcf', 'rdm', 'rcf', 'lam', 'ram', 'rcb', 'lb']
['lb', 'lcb', 'rb', 'ldm', 'rdm', 'rcf', 'lam', 'rcb', 'lcf', 'ram']
```

Figure 13 – Sortie attribution des rôles.

- Utilisation de l'organisation des rôles au temps initial.
- Utilisation de l'Algorithme de recherche de chemin minimal Hongrois.
- Réduire le nombre de changements d'organisation :
 - Augmenter le pas.
 - Lisser les données.
 - Supprimer certains rôles.

RÔLES DES JOUEURS

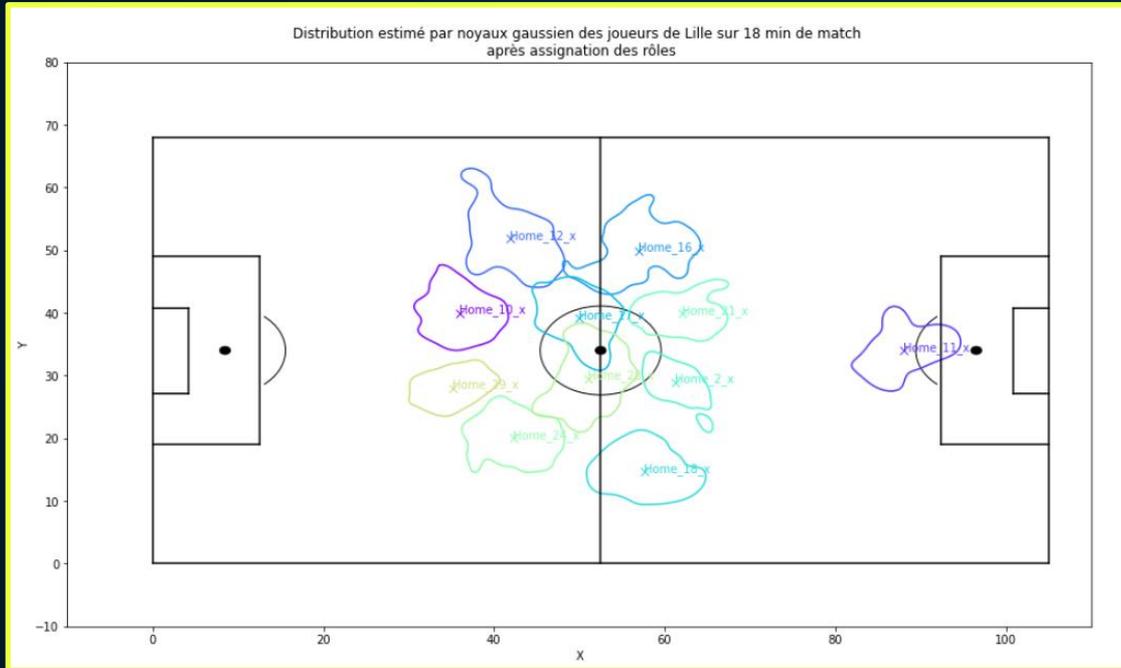
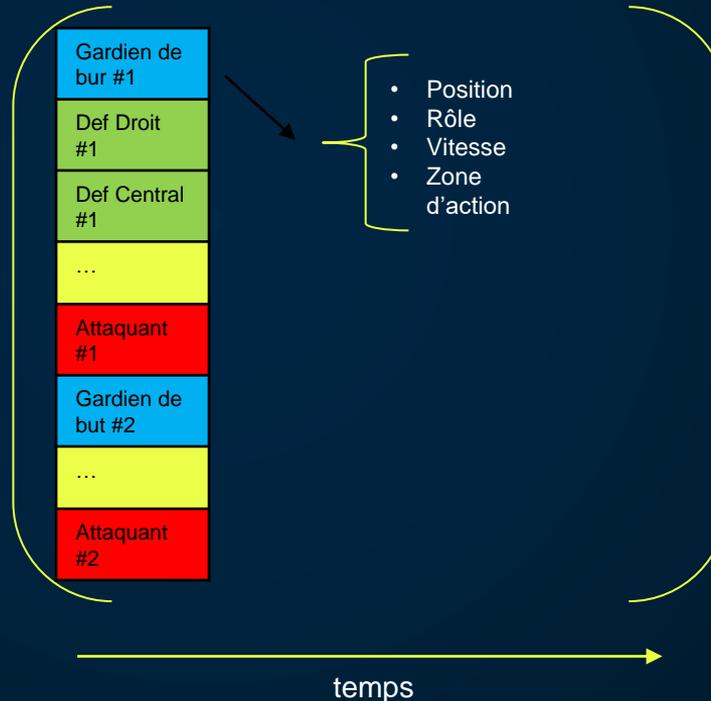


Figure 14 – Distribution estimée par noyau gaussien après attribution des rôles.



ALGORITHME DE MACHINE LEARNING

- Algorithme non supervisé.
- Unité statistique :
Joueur
- Aide d'un professionnel pour étiqueter le clustering et vérifier sa qualité.



SEPARATION EN POSSESSION

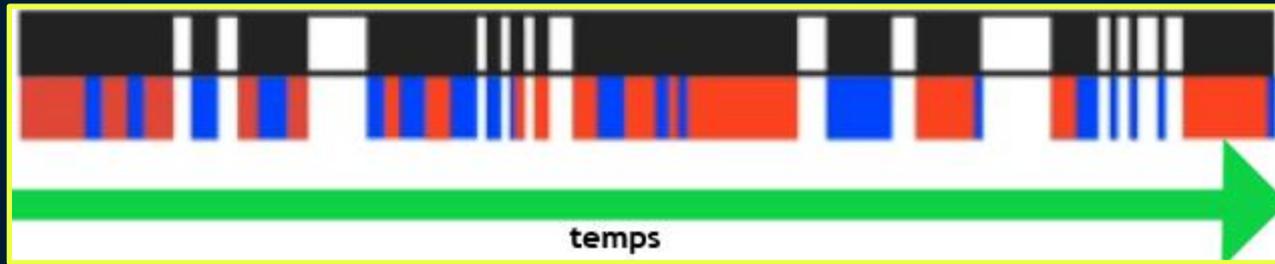


Figure 15 – Segmentation d'un match en possession.

DETECTION DE RUPTURES

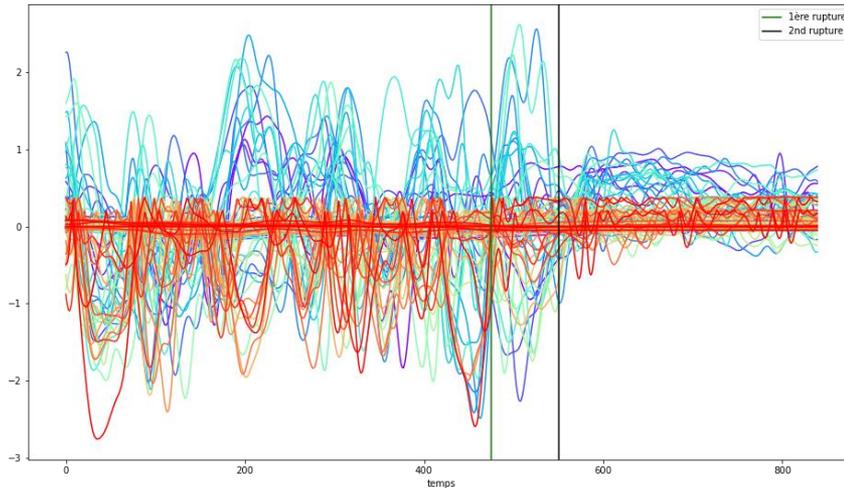


Figure 16 – Ruptures au sein d’une possession.

- Possession de 840 frames, 1min 40s.
- [475, 550, 840] sont de temps de rupture.
- Permet de réduire les phases de possession longues et d’identifier des sous possessions.



METHODE 1 : CLUSTERING

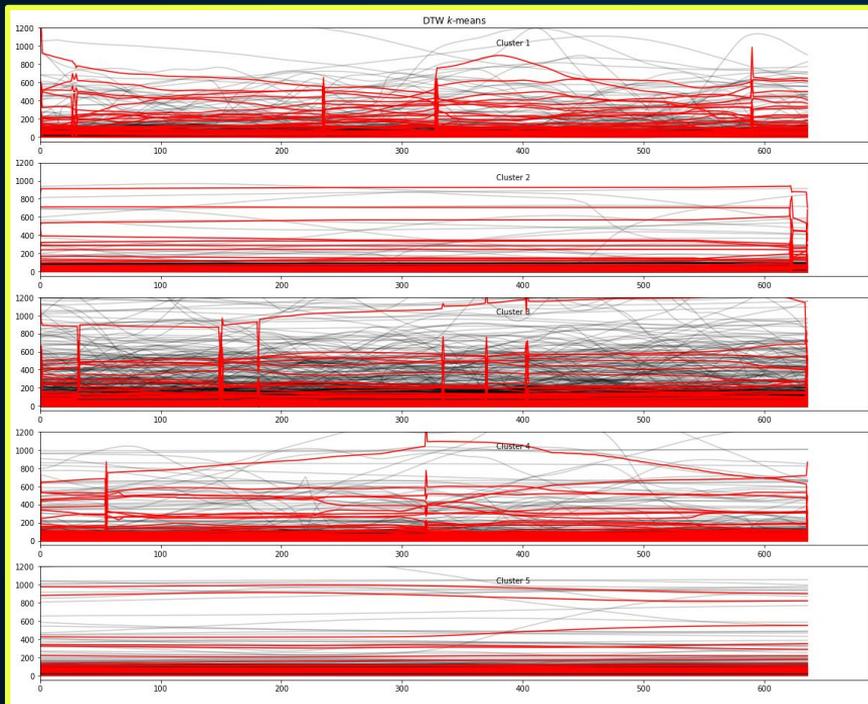


Figure 17 – Représentation des barycentres (lignes rouges) des clusters d'un Kmeans DTW.

- Dynamical Time Warping est une mesure de similarité élastique.
- C'est une mesure invariante aux décalages temporels.
- On va chercher le meilleur alignement entre deux séries temporelles.
- Coûteux en temps pour créer des clusters.

PERTINENCE DES RESULTATS

- Cluster 2, distance euclidienne
- Cluster 2, mesure de similarité DTW
- Cluster 4, distance euclidienne
- Cluster 2, mesure de similarité DTW



Figure 18 – Séquence de possession 9, image canal+.



Figure 19 – Séquence de possession 37, image canal+.

REDUIRE LES DIMENSIONS

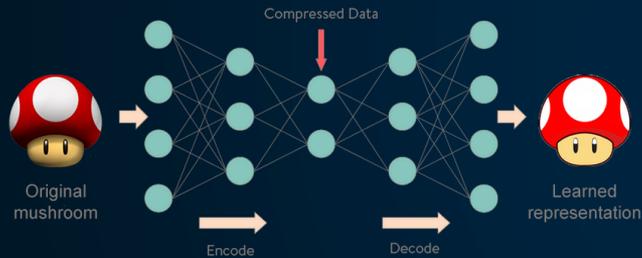


Schéma Auto encodeur pour réduire les dimensions.

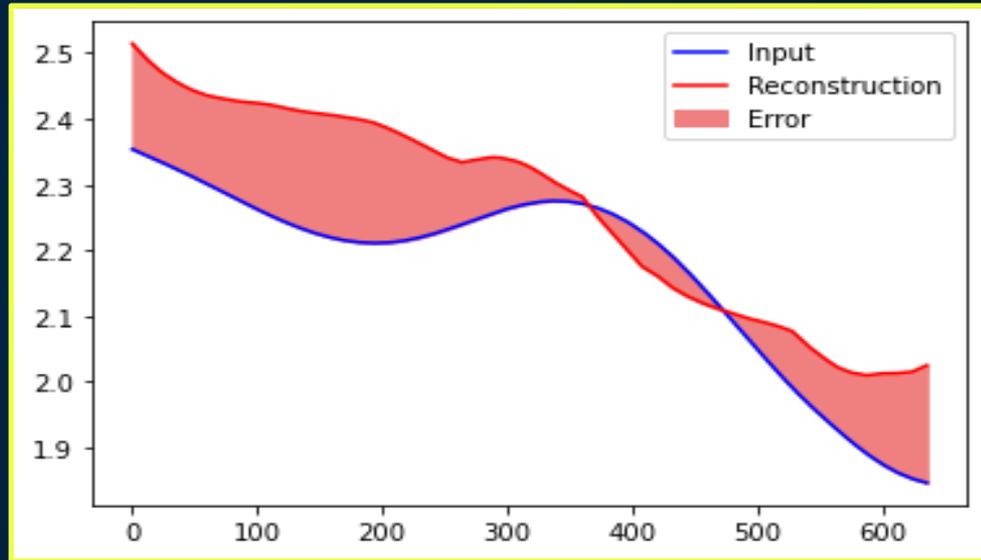


Figure 20 - Visualisation de l'erreur de reconstruction entre deux séries.

METHODE 2 : CREATION D'UN DENDROGRAMME

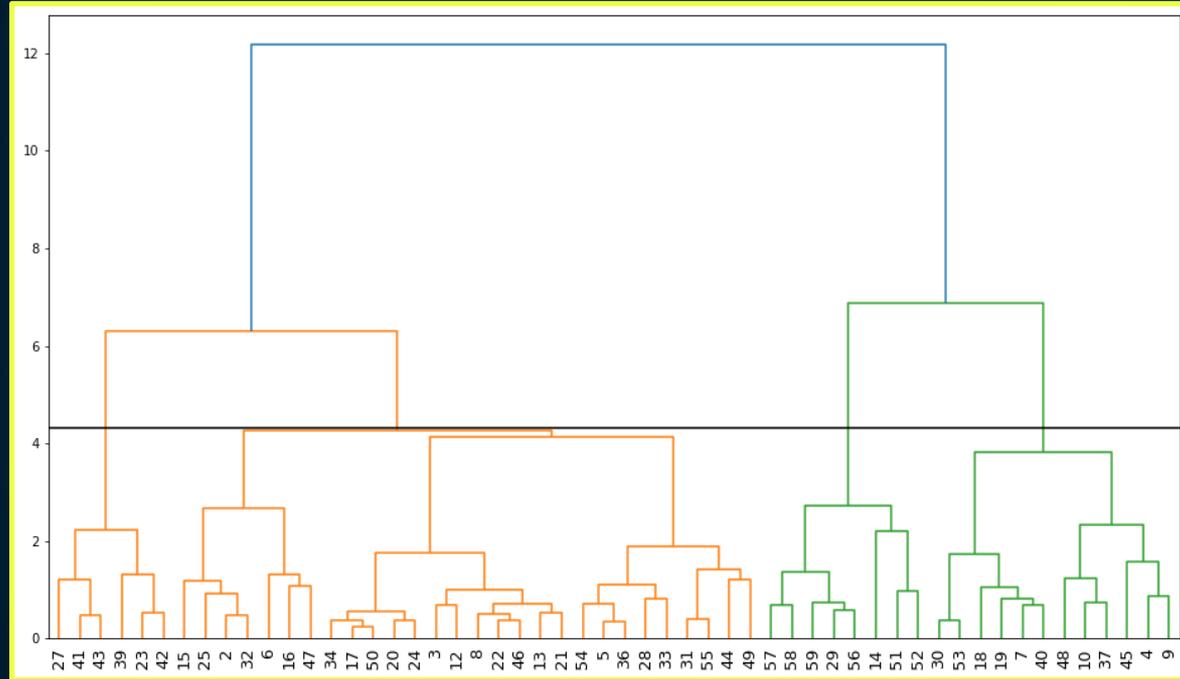


Figure 21 - Classification hiérarchique des possessions à partir des données reconstruite.

PERTINENCE DES RESULTATS



Séquence de possession 5, image canal+.



Séquence de possession 22, image canal+.



Séquence de possession 9, image canal+.

- Possessions 5 et 22 similaires d'après dendrogramme et 9 plus éloigné des deux.

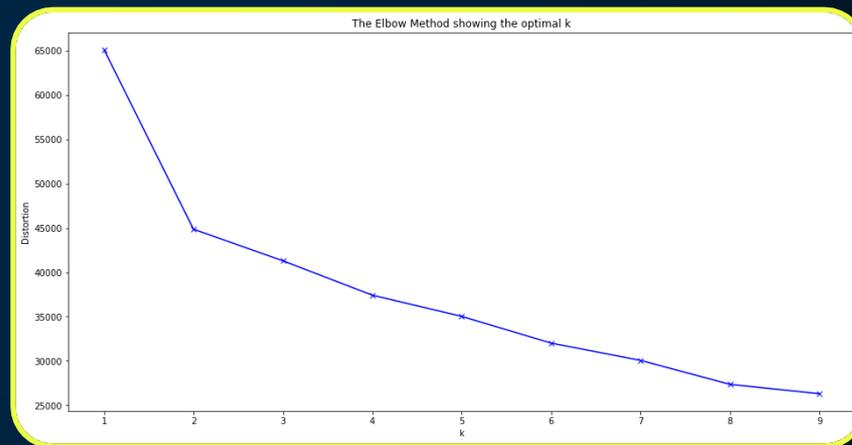
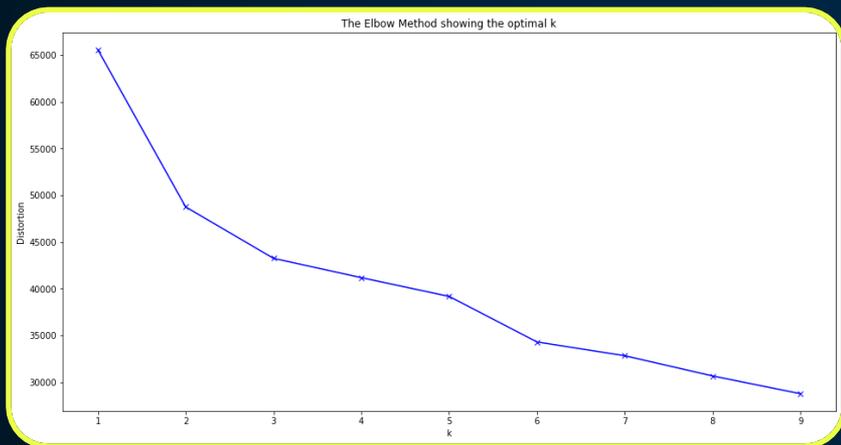
Perspectives

- Ajuster l'influence des variables.
- Envisager d'autres unités statistiques d'observation.
- Caractériser les clusters.
- Réaliser un K-means avec les séries réduites.
- Augmenter la taille de la base de données d'apprentissage.

**Merci de m'avoir
écouté !**

METHODE 1 : CLUSTERING

Package tslearn
Romain Tavenard
(2020)



RÔLE DES JOUEURS : utilisation d'un noyau gaussien

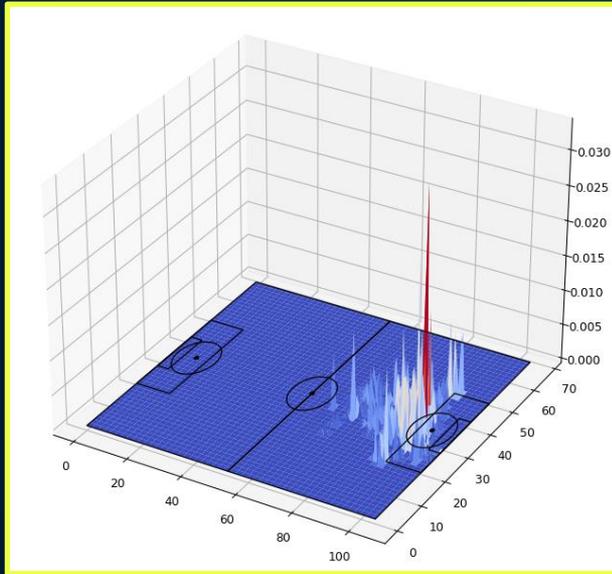


Figure 13 - Distribution par noyau gaussien, largeur de bande $h=0,01$

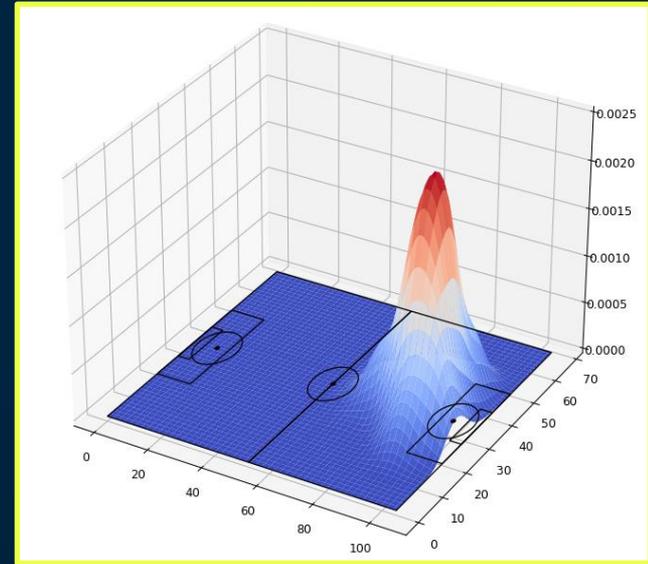


Figure 14 - Distribution par noyau gaussien, largeur de bande $h=0,5$