

QUAND LES DONNÉES PERMETTENT DE TOUT SAVOIR SUR SON ADVERSAIRE

L'Université de Rouen Normandie est partenaire de The Conversation, média en ligne proposant du contenu d'actualité élaboré avec des universitaires. À travers cette rubrique, retrouvez les articles de nos collègues.

Lorsque l'on s'intéresse à l'analyse de la performance sportive, peu importe le sport auquel nous nous attachons, il s'agit de répondre à deux questions fondamentales :

- Pour quelles raisons un joueur ou une équipe a perdu son précédent match ?
- Comment faire pour gagner le match à venir ?

Un analyste de la performance sportive va donc s'intéresser à ces deux problèmes, collecter des données sur le jeu et les analyser pour trouver des réponses. Après avoir été analyste de données et consultant dans le football anglais, je m'occupe depuis quelques années de la formation d'étudiants et doctorants à l'utilisation de données pour l'analyse de la performance sportive. Entre les sciences du sport et les statistiques, ces compétences sont en effet de plus en plus recherchées par les clubs professionnels afin de répondre à l'inflation de la donnée sportive qui est de plus en plus accessible et comprise pour le grand public.

Utiliser les données pour analyser la performance sportive

Pour prendre un exemple spécifique en badminton, il s'agit donc d'un côté d'analyser sa propre performance passée, et d'y déceler ses propres points faibles mais également ses points forts. Dans ce cas, l'analyste de la performance sportive va analyser l'activité du joueur pendant ce dernier match, pour en tirer des conclusions.

Par exemple [il a été montré](#) depuis peu que les joueurs d'origine chinoise ou japonaise ont tendance à utiliser la fatigue de l'adversaire avant de leur mettre la pression, à savoir qu'ils ont tendance à faire durer les échanges et ne cherchent pas à mener au score dans la première moitié du match. Les joueurs d'origine indienne ou taiwanaise ont eux plus

souvent des stratégies opposées. Il existe donc de multiples profils de joueurs et connaître précisément le profil d'un futur adversaire permet d'éviter de tomber dans le piège.

Or le problème majeur qui apparaît ici réside dans la comparaison, quelle est donc la valeur « normale » ? Imaginez que vous devez faire une prise de sang, et lorsque vous recevez les résultats il n'y a que les valeurs brutes mais pas les barèmes, il est impossible de savoir si ces analyses sont classiques, normales, ou bien si elles sortent de l'ordinaire. Pour pallier ce problème, l'analyste sportif va avoir besoin d'une grande base de données qui servira de référence, de barème pour qualifier le style de jeu d'un joueur spécifique par rapport à une norme, à un profil standard.

D'un autre côté, il s'agit d'analyser l'adversaire à venir et en quelque sorte de chercher à prédire ce qu'il va faire, la manière dont il va jouer ce match à venir. À ce jeu-là, [nous avons récemment mis en avant](#) la nécessité d'utiliser des données les plus récentes possibles afin d'avoir une prédiction la plus proche de ce qu'il va se passer. En fait, pour prédire l'activité de l'adversaire sur le prochain match, utiliser moins de données du passé (trois ou quatre matchs) mais des données très récentes (vraiment les derniers matchs) est la combinaison qui apporte la meilleure qualité de prédiction. Ceci s'explique par le caractère très dynamique de la performance sportive qui dépend de nombreux facteurs : état de forme du joueur, fatigue, blessure légère, motivation, etc., qui tous peuvent modifier la performance très rapidement. Ici, le problème pour l'analyste n'est pas tant la quantité de données, mais l'actualisation de ces données qui doivent constamment être mises à jour.

Analyser la performance sportive nécessite donc d'avoir une grande base de données, qui sert de référence ou de norme, mais également de données continuellement actualisées. Même si dans certains sports (principalement en football et basketball) des données sont accessibles pour certains matchs professionnels, dans 99 % des cas il faut collecter les données pour pouvoir les analyser, et c'est là que l'intelligence artificielle devient le plus grand atout de l'analyste.

L'intelligence artificielle dans la collecte et l'analyse des données sportives

Depuis bien longtemps les sciences du sport peuvent collecter des données sur les sportifs pour analyser leur performance. Or très souvent cela nécessite de rencontrer les sportifs, de les faire venir dans un laboratoire d'analyse du mouvement et le plus souvent de les équiper avec du matériel de mesure.

Analyse scientifique des mouvements d'un tennisman.

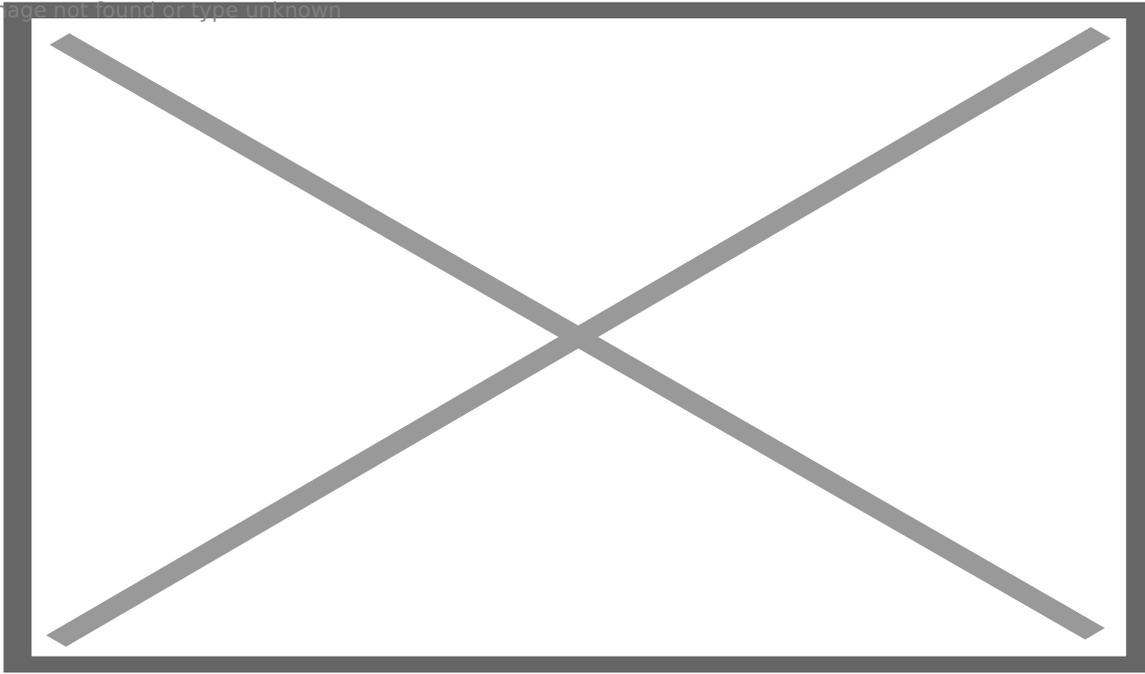
Ces dernières années, les avancées dans le domaine de l'analyse vidéo par ordinateur a rendu ces processus de collecte de données bien plus simple, facile à organiser, et surtout ne nécessite plus de matériel très avancé ni de faire venir les joueurs dans un endroit dédié à l'analyse de la performance. Par exemple, à partir d'une simple vidéo (d'une caméra, d'un smartphone) il est aujourd'hui possible d'analyser toute la biomécanique d'un joueur.

À partir d'une simple vidéo d'un match de badminton amateur, l'IA peut reconnaître les joueurs, le court, les déplacements des joueurs sur ce court ainsi que les mouvements et la technique des joueurs. L'avantage indéniable de ces nouvelles techniques réside dans la possible exploitation de toutes les vidéos existantes afin de créer une base de données quasi infinie.

Vidéo d'un match de badminton où l'IA reconnaît les joueurs et le court de badminton (droite), afin de collecter les déplacements des deux joueurs (gauche) ainsi que leurs mouvements/techniques (centre)

Les données pour analyser la performance peuvent maintenant être collectées n'importe où avec une simple caméra, à l'entraînement ou dans un tournoi amateur. Également, tout match diffusé à la télévision peut maintenant être analysé et il n'est plus nécessaire de connaître personnellement Roger Federer ou le champion du monde de badminton pour collecter des données sur leurs performances. Dans tous les cas, si votre prochain adversaire est Roger Federer, pourquoi accepterait-il de venir collecter des données pour que vous puissiez analyser ses points faibles ? Aucun joueur ou équipe ne ferait ça. Ainsi, pouvoir simplement utiliser des vidéos existantes pour collecter les données ouvre des opportunités d'analyse de n'importe quel joueur au monde.

Image not found or type unknown



Application de l'IA pour collecter les données sur un match télévisé disponible en ligne.

L'application de ces nouvelles techniques, en cours de développement voire déjà bien développées dans le football, la natation ou encore le tennis de table permet de régler les problèmes de collecte de données en grandes quantités (c'est un ordinateur qui travaille), sur des données toujours actualisées (pendant n'importe quelle compétition ou entraînement) mais également cela permet de collecter des données dans l'environnement réel de compétition. En effet, une des limites les plus importantes de l'analyse de la performance sportive en laboratoire est le manque de validité écologique de la mesure.

Autrement dit, les sportifs, parce qu'ils savent qu'ils sont observés, ou tout simplement par le matériel qu'ils doivent porter et qui n'est pas naturel, n'agissent pas exactement de la même manière qu'ils le feraient en compétition. L'intelligence artificielle permet donc une avancée sans précédent dans la collecte de « big data » dans le sport, mais plus il y a de données, plus il y a de « bruit » et plus il est difficile de trouver les informations réellement pertinentes dans cette immense masse de données.

Que faire de toutes ces données ?

Le champ d'application de l'IA dans le sport où il reste énormément à développer reste probablement l'analyse des données des données collectées, avec pour objectif de

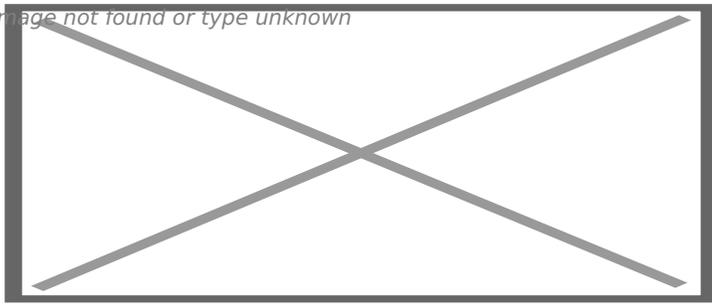
trouver la (petite) information qui fera la différence lors d'un match ou d'une compétition importante.

Dans toutes les données qui peuvent être collectées aujourd'hui pendant le jeu, une grande majorité est en fait inutile ou du moins ne permet pas de réellement impacter la performance. Est-ce qu'un plus grand nombre de passes effectuées dans un match de foot permet de gagner ce match ? Pas vraiment. Est-ce qu'une plus importante possession du ballon permet de gagner un match de football ? Pas vraiment. Est-ce que courir de plus longue distance par match permettent de gagner ce match ? Est-ce que des échanges plus longs en tennis permettent de gagner un match ? Peut-être pour certains joueurs. Une infinité de questions similaires peuvent ainsi être posées et la réponse se trouve dans les données collectées, encore faut-il pouvoir trouver cette réponse.

L'application de diverses techniques d'apprentissage par ordinateur ; réseaux de neurones, arbres décisionnels, clustering, etc., permettent de trouver ces informations « cachées » dans la masse de données sous forme de règles. Par exemple au tennis, le joueur X multiplie par quatre ses chances de gagner son échange du moment où son adversaire Y a couru plus de 90 mètres durant l'échange et qu'il joue sur son revers. En football, si une contre-attaque implique à minima un attaquant de plus que le nombre de défenseurs présents, fait progresser le ballon vers l'avant à une vitesse supérieure à 24 km/h, et atteint la surface de réparation en moins de 17 secondes après la récupération du ballon, elle aura 75 % de chance d'aboutir à un tir cadré.

Ce type d'informations, qui est « repéré » par l'IA au sein de grandes masses de données permet ainsi de connaître les points forts d'une équipe ou d'un joueur, reconnaître des schémas de mouvements qui sont souvent observés et mènent à marquer des points. Il est également possible de reconnaître des points faibles, donc lorsqu'un point est perdu par un joueur, que s'est-il passé avant et peut-on y trouver des récurrences ? Exploitées efficacement, ces informations peuvent alors permettre d'adapter une stratégie face à un adversaire bien spécifique, d'adapter un style de jeu, voire d'adapter des programmes d'entraînement pour par exemple réaliser une contre-attaque d'une manière où elle sera la plus efficace.

Image not found or type unknown



Cet article est publié dans le cadre de la

Fête de la science (qui a lieu du 6 au 16 octobre 2023 en métropole et du 10 au 27 novembre 2023

en outre-mer et à l'international), et dont The Conversation France est partenaire. Cette nouvelle édition porte sur la thématique « sport et science ».

Retrouvez tous les événements de votre région sur le site [Fetedelascience.fr](https://fetedelascience.fr).

Auteur

[John Komar](#), Maître de conférences en sciences du sport, [Université de Rouen Normandie](#)

Cet article est republié à partir de [The Conversation](#) sous licence Creative Commons. Lire l'[article original](#).

Publié le : 2023-10-11 10:36:45