

## **Deep Scouting - II**



### **Rapport du Projet :**

## **"Advanced Metrics : Développement de métriques innovantes pour l'analyse de performances (Football)"**

### **Équipe du projet :**

- OBOUNOU ZOLO Elvis
- Dhavalkumar PATEL
- Sadoc Najar
- **Tuteur** : Anuradha KAR

### **Sommaire**

- 1. Contexte du Projet**
- 2. Problématique**
- 3. Présentation des Données Utilisées**
- 4. Travaux Réalisés**
- 5. Résultats Obtenus**
- 6. Discussion et Commentaires sur les Résultats**

# Explication et Justification des Métriques Utilisées pour Évaluer la Performance de la Solution

## 1. Contexte du Projet

Le football, en tant que sport d'équipe, repose sur une multitude d'actions complexes qui, souvent, sont difficiles à analyser de manière précise. Les statistiques traditionnelles telles que les buts, les passes décisives, ou encore le nombre de tirs tentés, ne permettent pas de mesurer la véritable influence d'un joueur sur un match, en particulier en termes de création de menaces ou de transitions de jeu.

Dans ce contexte, l'objectif de ce projet est de développer des métriques avancées permettant une analyse plus fine des performances individuelles et collectives. En particulier, nous avons choisi de mettre au point l'indicateur **Expected Threat (xT)**, qui évalue la menace générée par chaque action sur le terrain, offrant ainsi une nouvelle façon d'évaluer les actions offensives des joueurs.

## 2. Problématique

La problématique que nous abordons ici est celle de l'évaluation de la performance individuelle d'un joueur en football, au-delà des statistiques classiques. Bien que des indicateurs comme les buts ou les passes réussies soient essentiels, ils ne permettent pas d'évaluer la manière dont un joueur influence les phases offensives de son équipe. La capacité d'un joueur à amener le ballon dans des zones dangereuses, ou à créer des opportunités de marquer, reste sous-évaluée.

Notre défi est de développer des métriques permettant de quantifier cette menace générée par un joueur, et d'analyser comment cette menace influe sur le jeu global d'une équipe. La mise en place du calcul de l'**Expected Threat (xT)** constitue donc une réponse à cette problématique, en se basant sur les déplacements du ballon sur le terrain et en mesurant l'impact potentiel de ces actions.

## 3. Présentation des Données Utilisées

Les données utilisées pour ce projet proviennent de la plateforme **StatsBomb**, qui propose des événements détaillés sur chaque match. Nous avons sélectionné les données du fichier **3943077.json**, qui provient du fichier `evebtsl`. Ce fichier contient une série d'événements, chacun décrivant une action sur le terrain (passe, dribble, tir, etc.) et associée à des informations comme la position du joueur, la durée de l'action, le type d'événement, ainsi que des informations sur l'équipe et la phase de jeu.

Les types d'événements traités dans ce projet sont les suivants :

- **Passes** : Passes réussies, passes clés, passes longues, passes latérales, etc.

- **Dribbles** : Actions de dribbles réussies ou échouées.
- **Tirs** : Nombre de tirs, tirs cadrés, Expected Goals (xG).
- **Données sur les joueurs** : Identifiant, position, équipe, rôle tactique, etc.
- **Données sur le match** : Temps, possession, transitions, phases de jeu.

## 4. Travaux Réalisés

### A. Préparation des Données

Les données du fichier **3943077.json** ont été extraites, nettoyées et préparées pour l'analyse. Nous avons utilisé **Python** et la bibliothèque **Pandas** pour importer et manipuler les données. Chaque événement a été analysé en fonction de son type et des coordonnées de départ et d'arrivée.

### B. Calcul de l'Expected Threat (xT)

Pour chaque action (passe, dribble, tir), nous avons calculé un score **xT** basé sur la position du ballon. Le terrain a été divisé en une grille de 8x8 zones (64 zones), chaque zone ayant une valeur de menace associée, plus la zone est proche du but, plus la valeur de menace est élevée. Le calcul du **xT** pour chaque action repose sur la différence de menace entre la zone d'arrivée et la zone de départ.

### C. Visualisation des Résultats

Les scores **xT** ont été visualisés sous forme de graphiques et de cartes de chaleur pour permettre une meilleure compréhension de la manière dont les joueurs utilisent les différentes zones du terrain pour créer des opportunités de but. Les actions les plus menaçantes ont été mises en évidence pour chaque joueur.

## 5. Résultats Obtenus

Les résultats ont montré une variation importante des scores **xT** entre les joueurs, en fonction de leur capacité à amener le ballon dans des zones de haute menace. Parmi les joueurs les plus influents dans la création de menaces, **Lionel Messi** a généré le score **xT** le plus élevé grâce à ses passes et dribbles dans des zones avancées. **Lautaro Martínez**, **Giovani Lo Celso**, et **Enzo Fernandez** ont également montré une capacité à se rapprocher des zones dangereuses, augmentant ainsi leur impact offensif.

Les défenseurs ont, quant à eux, eu des scores **xT** plus faibles, mais ont contribué aux transitions et aux contre-attaques, générant indirectement des occasions dangereuses.

## 6. Discussion et Commentaires sur les Résultats

L'analyse des scores **xT** montre clairement que certaines actions ont un impact beaucoup plus important que d'autres. Par exemple, une passe réussie dans une zone proche du but génère un **xT** beaucoup plus élevé qu'une passe latérale ou vers l'arrière. De même, les dribbles réussis qui déplacent le ballon dans des zones menaçant directement l'adversaire augmentent considérablement le score **xT**.

Les résultats révèlent également que les **milieux de terrain créatifs** et les **attaquants** sont souvent responsables de la majeure partie du **xT** dans une équipe, ce qui confirme leur rôle central dans la création d'occasions. Les **défenseurs** contribuent surtout par la **recupération de balle** et les transitions offensives, bien que leur **xT** individuel reste relativement faible.

## 7. Explication et Justification des Métriques Utilisées pour Évaluer la Performance de la Solution

### Expected Threat (xT)

Le **xT** a été choisi comme métrique principale pour évaluer l'impact des actions sur le terrain. Cette métrique est particulièrement pertinente dans un contexte de football moderne, où la capacité à créer des occasions de but est un facteur clé de succès. Contrairement aux statistiques traditionnelles, le **xT** permet de mesurer non seulement l'intention derrière une action (par exemple, une passe ou un dribble) mais aussi l'impact concret sur la probabilité de marquer.

### Justification des Métriques

- **xT pour les passes** : Les passes sont une composante clé de la création d'occasions. Un **xT** élevé indique qu'un joueur a effectué une passe qui a permis d'amener le ballon dans une zone menaçant de produire un but.

$$xT_{\text{passe}} = xT_{\text{end}} - xT_{\text{start}}$$

**xT\_end** : La valeur **xT** de la zone où se termine la passe (coordonnées dans `pass.end_location`).

**xT\_start** : La valeur **xT** de la zone où commence la passe (coordonnées dans `location`).

Si **xT\_end > xT\_start**, la passe a contribué positivement à créer une menace.

- **xT pour les dribbles** : Les dribbles qui déplacent le ballon dans des zones dangereuses augmentent également le **xT**, ce qui reflète l'importance de la progression du ballon pour créer des occasions

$$xT_{\text{dribble}} = xT_{\text{end}} - xT_{\text{start}}$$

- Les coordonnées de début sont obtenues de location, et celles de fin de carry.end\_location.

- **xT\_total**

Pour un joueur, la contribution totale est simplement la somme des contributions de passes et dribbles :

- $xT_{total} = \sum xT_{passe} + \sum xT_{dribble}$

## **Conclusion :**

Le projet a permis de démontrer la pertinence de l'indicateur **Expected Threat (xT)** pour analyser les performances des joueurs de football. En calculant le **xT** de chaque action, nous avons pu évaluer de manière plus précise l'impact des joueurs dans la création de menaces offensives. Les résultats montrent que les joueurs créatifs (comme **Messi**) ont une influence déterminante sur la progression du jeu et les opportunités de but. Le modèle peut être perfectionné pour inclure des actions défensives et des transitions, offrant ainsi une vue d'ensemble encore plus complète des performances des joueurs.