

# Analyse et optimisation de l'expected goal: application au machine learning

TRAVAIL DE BACHELOR HES RÉALISÉ EN VUE DE  
L'OBTENTION DU BACHELOR PAR :

DAVID PAULINO

CONSEILLER AU TRAVAIL DE BACHELOR :  
<PRÉNOM, NOM, TITRE DE LA PERSONNE>

GENÈVE, LE 19 AVRIL 2023

Haute école de Gestion de Genève (HEG-GE)

Filière Informatique de gestion

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Introduction à la problématique . . . . .	2
1.2	Intérêt de la problématique . . . . .	2
1.3	Questions que l'on souhaite répondre dans ce travail . . . . .	3
1.4	Intérêt de ces questions . . . . .	3
1.5	Applications concrètes . . . . .	3
	<b>Références</b>	<b>3</b>

# 1 Introduction

## 1.1 Introduction à la problématique

La problématique de ce travail est de pouvoir analyser et optimiser l'expected goal.

Il est tout d'abord important d'introduire ce qu'est l'expected goal. L'expected goal<sup>1</sup> est une métrique qui permet de déterminer la probabilité qu'un tir soit transformé en but selon les données de ce tir [?]. Un tir qui a un xG de 0.4 a une probabilité de 40% d'être transformé en but. Un tir avec un xG à 1 est la plus grande valeur possible et aurait donc 100% de chance d'être transformé en but.<sup>2</sup> [2]

La problématique m'a semblé cohérente par rapport à ma personnalité puisque je suis passionné par le football et que j'ai toujours été intéressé par les statistiques et les données. J'ai donc décidé de m'intéresser à l'expected goal et de voir comment on pouvait l'optimiser.

## 1.2 Intérêt de la problématique

Cette problématique est très intéressante puisque c'est une donnée qui est dernièrement très populaire dans le monde du football. Elle donne plus d'informations sur le match que les autres statistiques d'un match (possession, nombre de passes, nombre de tirs cadrés, nombre de buts, etc.). Par exemple en ce qui concerne la possession, une équipe peut avoir la possession du ballon pendant 70% du match mais ne pas marquer de buts, ni même être dangereuse avec le ballon. Le nombre de tirs cadrés, s'ils sont effectués tous à l'extérieur de la surface de réparation, ne sont pas forcément dangereux. L'expected goal permet de donner une valeur à chaque tir et donc de déterminer si un tir a une chance d'être transformé en but.

Ensuite, certaines personnes utilisent cette métrique pour leurs paris sportifs. Ils observent la métrique lors des derniers matchs et la compare avec le score réel du match. Si la différence est grande, cela peut permettre de constater un manque de réalisme ou un sur-régime d'une des deux équipes. [3]

---

1. Très souvent réduit par xG

2. Il est important d'indiquer qu'il est très rare qu'un xG d'un tir soit égal à 1 mais il va généralement sans rapprocher fortement selon ses paramètres.

D'autres personnes utilisent cette donnée pour analyser les performances de leurs joueurs et de leurs équipes. Par exemple, la comparaison des xG et du nombre de buts marqués d'un joueur sur une période donnée peut permettre de déterminer la dangerosité et la capacité d'un buteur à terminer des actions. [2] Les xG peuvent aussi être utilisés pour analyser les performances d'une équipe dans des situations bien précises. Une équipe qui possède un haut xG sur des contre-attaques montre qu'elle est très dangereuse sur ce type de situation. [1]

### 1.3 Questions que l'on souhaite répondre dans ce travail

Il y a plusieurs questions qui peuvent se poser dans ce travail. Je vais donc les lister.

- Comment est calculé l'expected goal ?
- Quels sont les paramètres qui influencent le plus l'expected goal ?
- Quel est le meilleur modèle pour prédire l'expected goal ?
- Quels sont les variables qui peuvent être ajoutée pour améliorer le résultat de la prédiction ?
- Quel est le meilleur modèle pour prédire l'expected goal en utilisant les variables ajoutées ?

### 1.4 Intérêt de ces questions

Le but de ces questions est de comprendre comment implémenter un modèle qui permet de prédire l'expected goal. Grâce à cela, on pourra ensuite connaître quels sont les variables obligatoires pour implémenter un tel modèle et quels sont les paramètres qui influencent le plus l'expected goal. Ensuite, on pourra voir quel est le meilleur modèle pour prédire l'expected goal et quels sont les variables qui peuvent être ajoutées pour améliorer le résultat de la prédiction. Enfin, on pourra voir quel est le meilleur modèle pour prédire l'expected goal en utilisant les variables ajoutées.

### 1.5 Applications concrètes

## Références

- [1] xG Explained | FBref.com.

- [2] Luke Petty. What is expected goals ? Expected goals explained.
- [3] David Tennerel. Bien utiliser les expected goals (xg) pour vos paris sportifs.