

Grau en Estadística

Títol: Quantificació, optimització i causalitat de l'espai en el futbol

Autor: Xavier Martínez Pla

Director: Montserrat Guillén i Estany

Departament: Departament d'Econometria, Estadística i Economia Aplicada

Convocatòria: Juny 2021



Resum

El futbol és un dels esports més aleatoris. Un gol succeeix poques vegades durant els noranta minuts d'un partit, la qual cosa fa que l'atzar estigui present en el resultat final. A més, segons Johan Cruyff un jugador sols té la pilota uns tres/cinc minuts en el transcurs del partit. Com a conseqüència, és important quantificar el rendiment d'un equip independentment del resultat i també el rendiment d'un jugador quan no disposa de la pilota. Per tal d'assolir aquest objectiu, en aquest treball final de grau es presenta un algorisme que quantifica una part del rendiment, l'espai generat en fase ofensiva durant el partit, aquest espai es pot diferenciar tant per jugadors com per parts del camp.

A partir d'aquest indicador es pot avaluar el posicionament d'un equip durant el partit, a més identificar bons i mals moments. També es pot identificar aquells jugadors que no s'han ubicat de forma correcta i per tant han generat menys perill. A partir de la diferenciació de l'espai en zones del camp completat amb un mapa de calor, es pot veure en quines s'ha disposat de més espai i si ha estat o no utilitzat.

Els dos avantatges principals del mètode que es presenta en el treball són:

1. L'alta eficiència computacional del mètode, la qual cosa permet que sigui interactiu.
2. Proporciona informació sobre tots els jugadors que estan en el camp, independentment de si tenen o no la pilota.

Tot i que l'enfocament de l'algorisme ha estat dissenyat com a mesura ofensiva, també es podria aplicar com a mesura defensiva, per identificar els defenses que han minimitzat els espais rivals.

Paraules claus: Diagrama de Voronoi, Futbol, Espai futbolístic, Punts fantasmes, Mapa de Calor, Posicionament en el futbol.

“Quantification, optimization and causality of the space in football”

Football is one of the most random sports. A goal rarely happens during the ninety minutes of a match, which means that luck is present in the result. According to Johan Cruyff, a player only has the ball for about three/five minutes during the match. As a result, it is important to quantify the performance of a team regardless of the result and the performance of a player when he does not have the ball.

In order to achieve this goal, in this final degree project I present an algorithm that quantifies a part of the performance, specifically the space generated in the offensive phase during the match; this space can be differentiated both by players and by areas of the field.

From this indicator, the position of a team on the field during the match can be assessed, as well as which moments have been the best and which moments have been the worst. It is also possible to identify those players who have not been positioned correctly and have therefore created less danger.

By differentiating the space according to the zones of the field, it can be seen in which zone more space has been available and whether it has been used.

The different results obtained with the method presented in this work are completed with a heat map in order to reach more consistent conclusions about the development of a match.

The two main advantage of the method presented in the work are:

1. The high computation efficiency of the method, which allows it to be interactive.
2. The space indicator provides information on all the players who are on the field, regardless of whether or not they have the ball.

While the approach of the algorithm has been designed as an offensive measure, it could also be applied as a defensive measure, with the goal of identifying the defense players that have minimized rival spaces.

Keywords: Voronoi diagram, Football, Football space, Ghost points, Heat map, Football positioning.

Classificació AMS (MSC2010):

28-04	Explicit machine computation and programs
62J05	Linear regression
68U05	Computer graphics; computational geometry
97R40	Artificial intelligence

A la Montse, per tot el temps dedicat i per l'atenció que he rebut en tot moment. A la meva família pel suport incondicional i al F.C Barcelona per fer-me estimar el futbol.

ÍNDEX

1. INTRODUCCIÓ	1
2. RECERCA PRÈVIA	3
3. METODOLOGIA	6
3.1. Diagrama de Voronoi	6
3.2. Models estadístics.....	9
3.2.1. Regressió Lineal.....	9
3.2.2. Gradient Tree Boosting	9
4. ALGORISME PEL CÀLCUL D'ESPAI	11
4.1. Pas 1: Acotació del diagrama de Voronoi.....	11
4.2. Pas 2: Càlcul de l'àrea	15
4.3. Eficiència computacional de l'algorisme:	16
5. SOFTWARE UTILITZAT	17
5.1. Per què Python i no R?	17
6. DADES	18
6.1. Origen	18
6.2. Descripció	18
7. RESULTATS	21
7.1. Descriptiva general.....	21
7.2. Quantificació de l'espai.....	27
7.2.1. Espai pel conjunt de l'equip.....	27
7.2.2. Espai porters	30
7.2.3. Espai jugadors.....	31
7.3. Optimització (Utilització de l'espai).....	37
7.4. Modelització	45
8. DISCUSSIÓ DELS RESULTATS	51
9. CONCLUSIONS	53
10. BIBLIOGRAFIA	54
11. ANNEX	56
Annex 1.....	56
Annex 2.....	66

Capítol 1: INTRODUCCIÓ

L'objectiu d'aquest treball és el de presentar un estadístic, una metodologia, per poder calcular l'espai que genera un equip en el terreny de joc durant qualsevol moment d'un partit, a més que es pugui repartir tant pels jugadors com per les zones del camp que siguin d'interès.

Un cop realitzat el primer pas, el següent és buscar si aquest espai generat ha estat utilitzat o no i esbrinar com optimitzar al màxim l'espai creat i alhora minimitzar l'espai que s'ha permès crear.

Ja com a darrera part, mitjançant mètodes de modelització, intentar buscar alguna causalitat (relació) entre l'espai creat i diferents variables que intentin explicar l'espai generat.

La principal contribució d'aquest treball és aplicar el diagrama de Voronoi al món del futbol.

Un cop realitzada la selectivitat, em vaig apuntar al grau d'Estadística de forma "aleatòria", ja que no tenia gens clar quin grau realitzar. Durant primer i segon, el meu interès per l'Estadística va anar augmentant de forma progressiva, les diferents assignatures cada cop em semblaven més interessants. A les acaballes de segon, vaig descobrir el món de l'estadística esportiva, el qual em va atreure especialment ja que relacionava el grau que estava estudiant amb l'esport, que sempre ha atret la meua atenció.

L'any passat, sobretot entre Març i Juny, vaig començar a llegir força llibres sobre l'estadística aplicada al món del futbol, com per exemple: Fútbol y Matemáticas de David Sumpter, Football Hackers de Christoph Biermann... En aquest moment, vaig sumar coneixement a la meua curiositat sobre l'estadística en el futbol.

A l'hora d'escollir tema sobre el treball final de grau tenia molt clar que el tema central seria el futbol, però havia de buscar quin era l'enfoc més interessant i alhora quin era realista.

Un dia d'estiu, vaig llegir un article (Javier Fernandez & Luke Bornn, 2018) que parlava de l'espai i de les metodologies que estaven sorgint per quantificar-lo, aquestes es basaven en mapes de calor i en metodologies avançades que sortien del nivell del grau, tot i això la idea de l'espai i la seva quantificació em va semblar interessant i aplicable, ja que coneixia que l'empresa Metrica-Sports oferia gratuïtament dades d'un partit complet on s'identifica la posició en el camp de cada jugador (coordenades) i l'explicació de cada acció (event) que succeeix durant els 90 minuts de joc.

Un cop decidit que dins del tema futbolístic s'analitzaria l'espai, vaig decidir que el primer punt i el més important era el de quantificar aquest espai, en altres paraules crear un estadístic que permeti assignar un número que mesuri l'espai generat durant el partit.

Creat l'estadístic el següent pas havia de ser aplicar-lo al partit de diferents formes i per buscar resultats i poder identificar un bon o mal rendiment.

Resumint, el problema a tractar en aquest treball és mesurar i analitzar l'espai que es genera en cada acció en un partit de futbol. Una definició d'espai en termes futbolístics, seria: part del terreny de joc en la que un jugador amb pilota, o que té opcions de rebre-la, pot desenvolupar la seva activitat sense o amb una mínima oposició.

En un futur, m'agradaria dedicar-me al camp de l'estadística esportiva, particularment aplicada al futbol. Per tant, estic molt motivat i animat en realitzar el treball final de grau, ja que penso que em pot servir com a carta de presentació en un futur. A més escollir un tema que m'atrau, em permetrà gaudir desenvolupant el treball.

Per desenvolupar tota aquesta feina, he utilitzat principalment el llenguatge de programació Python, ja que així he pogut ampliar el meu coneixement en software estadístic. A més és molt adient per resoldre el problema plantejat degut a les llibreries de les quals podem fer ús.

Durant el treball s'utilitzen tres paraules aplicades de l'àmbit futbolístic que podrien semblar sinònims pel lector, per evitar confusions es defineixen abans de començar:

- Moment: Fa referència a cada 0,04 segons, cada instant del que es disposa coordenades dels jugadors.
- Jugada: Conjunt de moments, fins que la pilota canvia d'un jugador a un altre, essent els dos del mateix equip.
- Acció: Conjunt de jugades, fins que la pilota canvia d'equip.

Capítol 2: RECERCA PRÈVIA

En la taula següent es mostren un seguit de treballs relacionats amb la temàtica que ens ocupa en aquest TFG, els quatre primers. Mentre que els altres quatre presenten diverses aplicacions del Diagrama de Voronoi.

Títol, Autor i Any	Resum	Comentaris
Wide Open Spaces: A statistical technique for measuring space creation in professional soccer. Fernandez, J. i Bornn, L. (2018)	Proposa un model innovador per quantificar el nivell d'influència en una zona del terreny de joc en cada acció, és un model que incorpora informació dels moviments, distància entre jugadors i pilota i la velocitat dels jugadors. A partir d'aquest model quantifica l'espai ocupat (SOG) i l'espai generat (SGG).	- Parla de l'aplicació del Voronoi al futbol. Comenta les bones propietats que té, però no li convenç, ja que en aplicar el Voronoi se suposa que l'espai és discret, quan realment és continu. A més la distància entre jugadors i pilota no es té en compte. - Personalment crec que l'enfocament d'aquest article és excel·lent, ja que té en compte factors que jo no tinc en compte, tot i això penso que la seva alta complexitat limiten la seva aplicació a equips d'elit.
Evaluation of changes in space control due to passing behavior in elite soccer using Voronoi-cells. Rein, R. et al. (2016)	Els autors proposen un enfocament per avaluar el comportament de les passades, mitjançant el diagrama de Voronoi investiguen si la passada realitzada ha tingut canvi en l'espai. L'assumpció que fan en aquest mètode és que una passada efectiva dona un avantatge en el control de l'espai. Una passada que augmenti l'espai, la categoritzen com una passada efectiva.	- Tal com es veu en l'apartat de resultats, aquest enfocament sols té sentit en passades que es realitzen del mig del camp a zona ofensiva. Tot i això em sembla interessant avaluar una acció en funció de l'espai generat. - La debilitat del mètode proposat penso que recau en les passades que tenen lloc en la part ofensiva, en espais reduïts.
Classification of Passes in Football Matches using Spatiotemporal Data. Horton, M. et al. (2014)	A partir de mètodes de ML, els autors proposen un sistema automatitzat per tal de classificar les passades entre bones, normals i dolentes. En el model actuen diversos predictors, entre els quals un grup d'ells es calculen a partir de geometria computacional, com l'aplicació del Diagrama de Voronoi. En aquest cas, concretament els utilitza com a mesura d'aproximació de la posició estratègica de l'equip.	- L'enfocament d'utilitzar el Diagrama de Voronoi com a mitjà per obtenir una variable explicativa sembla molt interessant. Ja que tal com s'ha vist l'aplicació del Voronoi en un partit de futbol recull informació sobre el domini posicional. En entrar com a variable explicativa junt amb altres variables, els possibles inconvenients de l'aplicació de Voronoi es veuen reduïts.
Spatial dynamics of team sports	L'objectiu de recerca en aquest cas tracta d'investigar el comportament de les dinàmiques espacials dels	- El primer punt interessant és l'aplicació del Voronoi en el futbol-sala, en comptes del futbol. La

<p>exposed by Voronoi diagrams.</p> <p>Fonesca, S. et al. (2012)</p>	<p>jugadors de futbol-sala en una situació de superioritat numèrica, de 5 contra 4.</p> <p>Per caracteritzar la interacció entre els dos equips, utilitza els Diagrama de Voronoi, i calcula l'àrea que domina cada equip, demostrant que l'equip que ataca amb superioritat numèrica disposa de més àrea.</p> <p>Per caracteritzar la interacció entre els jugadors d'un mateix equip, utilitza la distància mínima entre jugadors d'un mateix equip. La qual és mínima sempre en l'equip que defensa, ja que tracta de jugar el més junt possible.</p>	<p>qual cosa demostra la seva diversitat en aplicar-lo a diferents esports.</p> <p>- La combinació del Voronoi i de la distància entre els jugadors d'un mateix equip, em sembla molt interessant per entendre com es mou un equip. A més que afegir una distància d'aquest tipus pot ajudar a contrarestar alguns dels inconvenients del Diagrama de Voronoi.</p>
<p>Segmentation of Page Images Using the Area Voronoi Diagram.</p> <p>Kise, K. et al. (1997)</p>	<p>Els autors proposen una aplicació del Diagrama de Voronoi en un camp al qual no estem acostumats a veure'l, el de les dades textuais.</p> <p>A partir del Diagrama de Voronoi els autors segmenten una pàgina de text, extraient les diferents regions amb l'objectiu d'identificar quina part del text representa cada regió.</p> <p>Aquest mètode és molt eficient computacionalment i mostra bons resultats independentment del format, disseny i angle d'inclinació del text. Tot i això, mostra algunes limitacions a l'hora de segmentar taules i figures.</p>	<p>- L'aplicació innovadora del Voronoi en el camp de les dades textuais, és el que fa més interessant aquest article.</p> <p>- Un altre punt alternatiu d'aquest mètode és que mostra un altre enfocament en termes de la distància que s'utilitza en calcular el Diagrama de Voronoi.</p>
<p>Computational analysis of RNA-protein interaction interface via the Voronoi diagram.</p> <p>Mahdavi, S. et al. (2011)</p>	<p>La geometria computacional juga un rol clau en l'estudi d'estructures biològiques complexes, és un primer pas per entendre els perfils d'unió. En aquest article els autors proposen un mètode innovador en l'estudi de la superfície d'interacció entre l'RNA-Proteïna. En aquest cas, com a tècnica geomètrica apliquen el Diagrama de Voronoi ponderat per l'estructura atòmica.</p> <p>Com aproximació sobre la grandària de la interfície RNA-Proteïna s'utilitza el nombre de tesselles de Voronoi.</p>	<p>- El punt principal d'aquest article és que és el primer que aplica tècniques computacionals geomètriques en la interfície complexa d'RNA-Proteïna. A més mostra molt bons resultats en comparació dels mètodes clàssics.</p> <p>- L'altre punt interessant a comentar és que utilitza un Diagrama de Voronoi alternatiu, el ponderat.</p>
<p>Application of Voronoi diagrams in contemporary architecture and town planning.</p> <p>Nowak, A. (2015)</p>	<p>Actualment, els mètodes de disseny modern es basen en bona part en comprendre la naturalesa dels processos i en els principis de les estructures biològiques. Cada cop s'utilitza més els models matemàtics per la seva representació.</p> <p>El desenvolupament tecnològic dels ordinadors permet crear estructures complexes inspirades en formes naturals.</p>	<p>-La finalitat de l'aplicació de Voronoi en l'arquitectura és molt interessant i alhora diferent del que s'ha vist anteriorment, ja que en aquest cas no s'utilitza per calcular, sinó que s'utilitza com a font d'inspiració i de discretització de les estructures.</p> <p>- L'article permet observar com la divisió de l'espai és un problema que fascina a parts iguals a matemàtics que a arquitectes. A</p>

	<p>L'autora exposa com es discretitza la superfície a partir del Diagrama de Voronoi, aquesta tècnica s'utilitza sobretot en alçats d'edificis contemporanis. A més es veu com el Voronoi és una eina geomètrica que els arquitectes la utilitzen per donar forma a les estructures, a més a més és una font d'inspiració.</p>	<p>més veiem com la geometria computacional, concretament el Diagrama de Voronoi, crea noves oportunitats per a projectes arquitectònics i urbanístics.</p>
<p>A Voronoi diagram based population model for social species of wildlife. Stewart, C. et al. (2010)</p>	<p>Els autors utilitzen el Diagrama de Voronoi per la creació d'un model de població en espècies de vida salvatge, que expliqui l'estructura espacial dins de l'hàbitat. A partir del model es veu com els grups socials veïns estan relacionats amb camins més curts, aquest enfocament ofereix una molt bona representació visual del repartiment de la població. A partir del model creat se simulen dinàmiques de població.</p>	<p>-El punt més interessant d'aquest enfocament és la sortida que s'obté amb el model, mapes de densitat. Ja que a banda de demostrar que el Voronoi és una bona eina per quantificar també es veu com és una eina per visualitzar de forma ràpida l'estructura poblacional i detectar efectes d'espai. A més, es veu la utilitat de combinar el Diagrama de Voronoi amb models de dinàmica poblacional.</p>

Capítol 3: METODOLOGIA

En aquesta secció, es revisaran els mètodes teòrics utilitzats al llarg del treball.

Per quantificar l'espai que es genera en el terreny de joc en cada posició s'utilitza una estructura coneguda en el món de la geometria, els diagrames de Voronoi, que permeten calcular l'àrea que ocupa cada jugador. Un cop calculada s'estudia quin és el pes adient que hem de donar en cada zona de joc, segons la ubicació del jugador. Això permet instaurar un estadístic per quantificar l'espai aprofitable que disposa un equip en una determinada acció.

Diagrama de Voronoi

El diagrama de Voronoi és una de les tècniques més populars dins la geometria computacional per fer una partició de l'espai.

Donat un conjunt de punts $\{p_1, \dots, p_n\}$ en el pla, el diagrama de Voronoi es defineix com la subdivisió del pla en subregions de manera que hi ha una regió $V(p_i)$ associada a cada punt p_i tal que la regió $V(p_i)$ està formada pel conjunt de punts que són més propers a p_i que a qualsevol dels p_j ($j \neq i$). Aquesta regió $V(p_i)$ s'anomena tessella de Voronoi o regió de Voronoi associada al punt p_i .

Aquesta definició es pot fer per qualsevol funció distància, tot i això quan parlem de diagrama de Voronoi, s'entén que es considera la distància euclidiana.

Tal i com es pot veure en la Figura 1, fixats p_i i p_j ($j \neq i$), la recta r_{ij} està formada pels punts que equidisten de p_i i p_j .

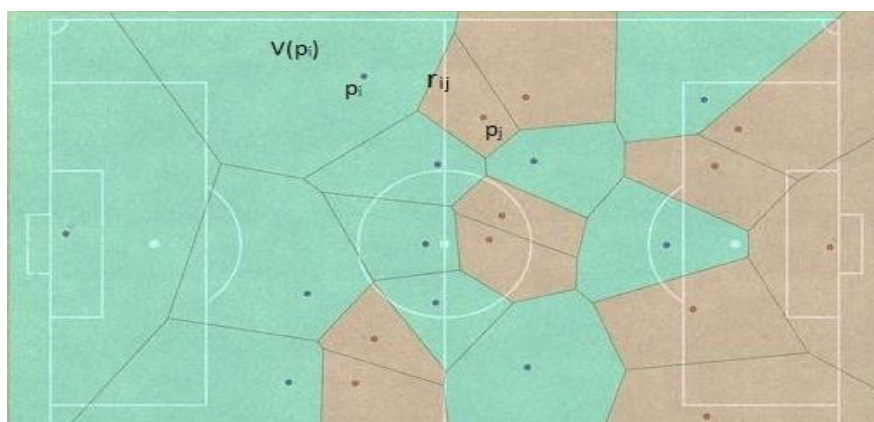


FIGURA 1: Exemple del Diagrama de Voronoi

La finalitat de l'aplicació del Diagrama de Voronoi és obtenir informació referent a la proximitat d'un punt qualsevol del pla als punts d'un conjunt donat.

Les propietats del diagrama de Voronoi són:

- Les regions de Voronoi són conjunts convexos. Les regions de Voronoi vénen limitades per arestes de Voronoi.
- Una regió de Voronoi és no afitada si i sols si el seu generador es troba en la frontera de l'envolupant convexa definida pels punts. (En el nostre cas, la regió sempre és acotada, ja que aquesta vindrà retallada per l'àrea del camp de futbol)
- Les arestes de Voronoi limiten dues regions de Voronoi veïnes. Cada punt d'una aresta de Voronoi és equidistant als dos punts que defineixen les regions de Voronoi que limita.
- El vèrtex en el qual tres regions de Voronoi ($V(p_i), V(p_j), i V(p_k)$) s'intersequen s'anomena un vèrtex de Voronoi, té la propietat de ser equidistant als punts $p_i, p_j, i p_k$. És el centre del cercle que passa per aquests tres punts.
- El diagrama de Voronoi d'un conjunt de punts és únic i està definit pels punts.
- L'estructura de dades dual que en resulta del Diagrama de Voronoi és la triangulació de Delaunay. És a dir, és la triangulació dels punts donats de manera que cada punt s'uneix mitjançant una aresta amb els seus veïns de Voronoi. En matèria futbolística, aquesta propietat s'utilitza com a suport a l'hora de visualitzar les associacions entre jugadors, que existeixen dins del terreny de joc.

Les aplicacions del diagrama de Voronoi són molt diverses, tot i que es destaca en l'anàlisi de recursos, estudiant les àrees d'influència i també molt característic en l'estudi de la naturalesa.

Els algorismes més populars per obtenir el Diagrama de Voronoi són:

- Intersecció de semiplans: Es construeix cada regió de Voronoi per separat mitjançant la intersecció de $n-1$ semiplans. La construcció d'aquests semiplans es realitza mitjançant un algorisme de "divideix i venceràs". Posteriorment, es computarà la intersecció d'aquests $n-1$ semiplans per donar origen a $V(p_i)$. El seu cost computacional és molt alt.

- Algorisme Incremental: Tot i la seva complexitat quadràtica, ha estat el més popular per construir el diagrama. El seu funcionament és el següent:
 1. Construït el diagrama $S = \{p_1, \dots, p_n\}$, punt a punt. S'afegeix un nou punt p_{n+1} .
 2. Es busca la regió $V(p_k)$ on es trobi el punt p_{n+1} .
 3. Es traça la mediatriu del segment $p_{n+1}p_k$ i la mediatriu del segment $p_{n+1} p_i$ per cada p_i veí del vèrtex p_k .
 4. S'elimina les porcions d'aresta i els vèrtexs que queden dins de la nova regió.

- Divideix i Venceràs: La seva complexitat és asimptòticament òptima. Els passos fonamentals de l'algorisme són:
 1. Es divideix S en dos subconjunts S_1 i S_2 , aproximadament de la mateixa grandària..
 2. Es calcula recursivament els diagrames $V(S_1)$ i $V(S_2)$.
 3. A partir dels diagrames obtinguts, calcular $V(S)$. Calculant la poligonal, i eliminant les línies de $V(S_1)$ i $V(S_2)$ a la dreta (esquerra) de la poligonal.

- Algorisme de Fortune: Un algorisme intel·ligent que sorgeix l'any 1985, el qual resulta tan simple com l'incremental però més eficient. És un algorisme "d'escombrat" (*sweep line*), d'esquerra a dreta construeix el diagrama de Voronoi a mesura que van apareixent els punts en aquest escombrat del pla.

El càlcul del diagrama de Voronoi està implementat en diferents llibreries de Python. En el nostre cas, farem servir la llibreria Scipy.

Models estadístics

Regressió Lineal

La regressió lineal és un enfoc per tal de modelar la relació entre la variable resposta i una o més variables explicatives. Relaciona la variable resposta Y amb p variables explicatives, X_j amb $j = 1 \dots p$, o qualsevol transformació d'aquestes que generin un hiperplà de paràmetres B_j . La formulació matemàtica que segueix el model és la següent:

$$Y = B_0 + B_1X_1 + \dots + B_pX_p + e$$

on e és una variable aleatòria que recull els factors no controlables, en altres paraules és el terme que representa l'atzar. Mentre que B_0 és el terme constant.

Si $j = 1$ parlem de regressió lineal simple, altrament parlem de regressió lineal múltiple.

La regressió lineal té diferents usos pràctics, però els dos més comuns són els de predir i explicar les variacions en la variable resposta.

Els paràmetres B_j estimats representen els coeficients de la regressió. Aquests coeficients medeixen la influència que té cada variable explicativa sobre la variable resposta. Per estimar els coeficients s'utilitza el mètode de mínims quadrats. Aquest mètode busca la corba que passi tan a prop com sigui possible dels punts de les dades, de tal manera que sigui mínima la suma dels residus elevats al quadrat entre els valors de la corba i dels punts observats

La fórmula matemàtica que s'utilitza per estimar els paràmetres és la següent (representada en forma matricial):

$$\hat{B} = (X^t X)^{-1} X^t Y$$

Gradient Tree Boosting

El gradient boosting és una tècnica d'aprenentatge automàtic utilitzada per l'anàlisi de regressió i per problemes de classificació, aquesta tècnica produeix un model en forma d'un conjunt de models de predicció, en aquest cas un conjunt d'arbres de decisió. Aquest mètode pertany a la família dels *ensembles* (tècnica que crea múltiples models i els combina).

El model es construeix de forma escalonada a partir de la tècnica boosting, combinant resultats dèbils per tal d'obtenir resultats robusts. Generalitza el model aplicant una funció de pèrdua.

Un avantatge d'utilitzar la tècnica de combinar arbres de decisió com el gradient boosting és que poden proporcionar estimacions de la importància de les variables que han estat introduïdes al model de forma explicativa, mostrant una puntuació relativa de la importància per cada atribut.

A partir d'aquesta puntuació es pot ordenar els atributs i comparar la importància entre uns i altres.

Capítol 4: ALGORISME PEL CÀLCUL D'ESPAI

Introduït el marc teòric del diagrama de Voronoi, en aquest apartat es detalla el procediment que s'ha realitzat manualment a través del Python (veure ANNEX 1), amb la finalitat de construir l'algorisme de càlcul d'espai.

L'explicació dels passos que s'han seguit per a la construcció de l'estadístic se separen en dos passos principals: el primer, l'acotació del diagrama de Voronoi on es retalla el diagrama de Voronoi a la mida del camp per tal que totes les tessel·les dels jugadors siguin finites, el segon pas és el del càlcul de l'àrea que ocupa cada tessel·la en el terreny de joc.

Pas 1: Acotació del diagrama de Voronoi

Un cop construït el diagrama de Voronoi, una de les limitacions que ens sorgeixen en la seva aplicació en un partit de futbol és que obtenim tant tessel·les de Voronoi infinites com tessel·les finites amb una part d'aquesta fora del camp.

Les tessel·les de Voronoi infinites són les que corresponen a jugadors que es troben en l'envolupant convexa del conjunt format per tots els jugadors.

En les figures dels diagrames de Voronoi, les arestes infinites es mostren amb línia discontinua, mentre que les arestes finites es dibuixen mitjançant una línia contínua (veure Figura 2).

De tal manera que per calcular l'àrea dominada per cada jugador, hem de retallar el diagrama de Voronoi a la mida del camp, en aquest cas de 105x68m.

No és trivial l'acotació de les regions de Voronoi, de manera que només es consideri el tros que cau dins del camp, degut a la casuística amb que ens podem trobar. El mètode més robust per realitzar aquest retallat es basa en la idea de Bobach et al. (2009) d'afegir punts fantasma que forcin que els límits del camp siguin arestes de Voronoi.

S'afegeixen punts fantasma per cada jugador que defineix una tessel·la que talla algun costat del camp. Per cada un d'aquests punts afegirem entre un i tres punts fantasma. El fet d'afegir aquest punt, força que el costat del camp sigui una aresta del Voronoi, ja que és la mediatriu entre els dos punts, l'original i l'afegit (punt fantasma).

Les diferents llibreries que calculen diagrames de Voronoi no tenen funcions per tallar ni per incorporar punts fantasma, com a conseqüència s'ha hagut d'implementar manualment.

Les diferents tessel·les que es troben a l'aplicar el diagrama de Voronoi, abans de retallar-lo, es classifiquen entre finites i infinites, dins d'aquesta classificació es troben diferents tipus:

• **Tessel·la finita:**

- *Tots els vèrtexs de Voronoi es troben dins del camp, aquest tipus de tessel·la és la ideal, a la qual no li hem de realitzar cap transformació. Per exemple, la tessel·la pintada en vermell, en la Figura 2, satisfà aquestes condicions.*

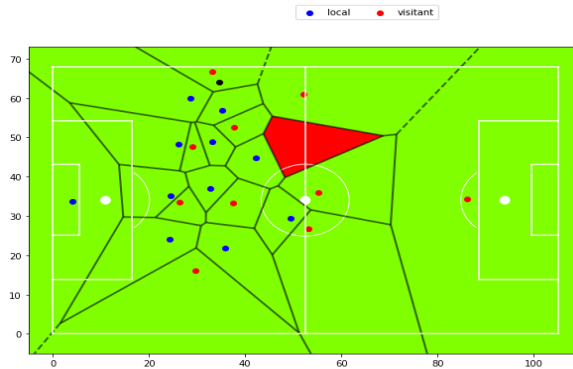


FIGURA 2: Exemple de tessel·la finita on tots els vèrtexs es troben dins del camp

- *Algun vèrtex de Voronoi es troba fora del camp, en aquest cas tenim dues arestes de Voronoi que tallaran la frontera del camp. Dins d'aquest subtipus existeixen dos altres tipus, segons si les arestes tallen o no pel mateix costat del camp:*

→ *Tallen pel mateix costat:* Afegim un punt fantasma, simètric del nostre jugador, respecte a la frontera (costat del camp) que estem tallant. Veure l'exemple de la tessel·la pintada en vermell en la Figura 3.

→ *No tallen pel mateix costat:* Afegim dos punts fantasmes, per cada costat de la frontera, on cada punt fantasma és simètric al nostre jugador. Veure l'exemple de la tessel·la pintada en blau en la Figura 3.

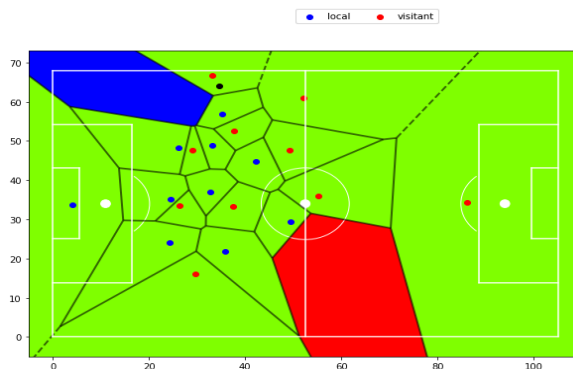


FIGURA 3: Exemple de tessel·la finita on no tots els vèrtexs es troben dins del camp

- **Tessel·la infinita:**

En aquest cas hem d'extremar les precaucions, ja que la frontera del camp la poden tallar tant arestes finites com infinites. Per mirar quins punts fantasmes s'han de posar, igual que en el cas anterior, es consideren les arestes de la tessel·la de Voronoi que tallen la frontera.

Ara fins i tot ens podem trobar que una o dues de les arestes de Voronoi tallin costats diferents del camp de manera que deixin dos còrnors dins de la tessel·la considerada. En aquest cas haurem d'afegir tres punts fantasmes associats al jugador corresponent a la tessel·la en qüestió, tal com s'observa en la tessel·la de color blau pintada en la Figura 4

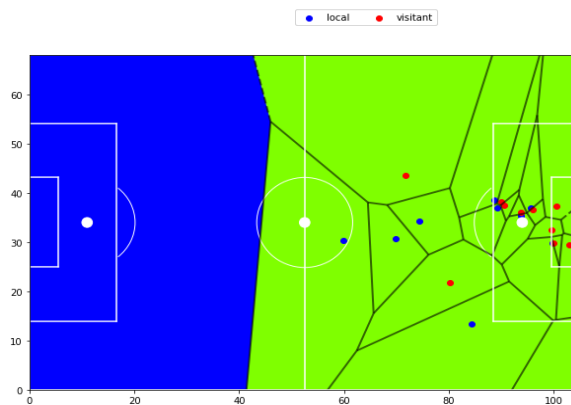


FIGURA 4: Exemple de tessel·la infinita

Les figures 5,6 i 7 mostren tot el procés de retallar el Voronoi mitjançant el mètode descrit anteriorment.

- 1) La Figura 5 presenta el diagrama de Voronoi corresponent als jugadors en un moment de joc:

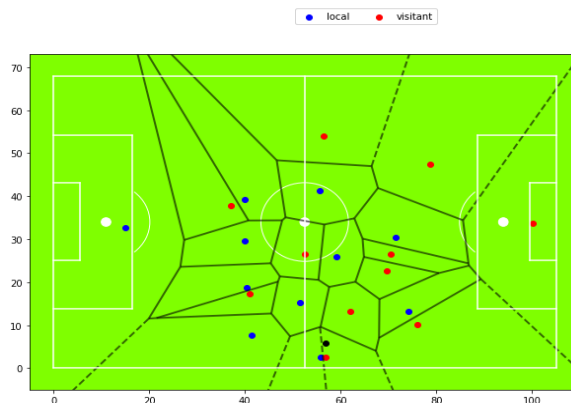


FIGURA 5: Diagrama de Voronoi sense cap transformació en un moment del partit

- 2) La Figura 6 mostra la incorporació dels punts fantasmes necessaris per retallar les tessel·les que ho requereixen a partir de la mida del camp. Dins del camp es visualitzen els punts originals, mentre que fora d'aquest entorn veiem els punts fantasmes.

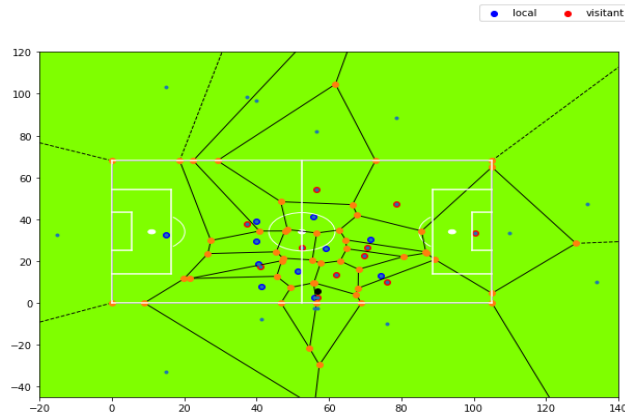


FIGURA 6: Diagrama de Voronoi junt amb els punts fantasmes

- 3) Finalment, ens quedem només amb les tessel·les del nou Voronoi que corresponen als jugadors, la Figura 7 indica com:

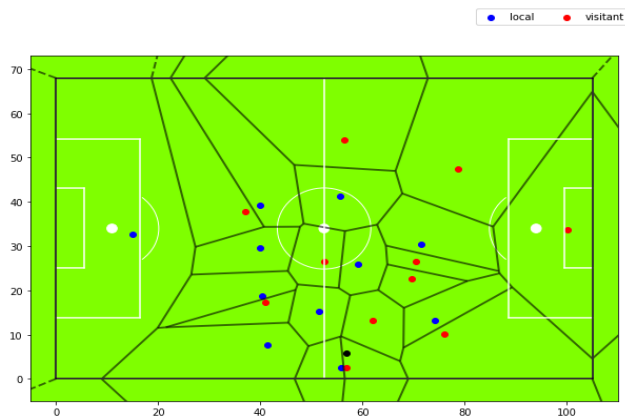


FIGURA 7: Diagrama de Voronoi retallat

Pas 2: Càlcul de l'àrea

Un cop retallat el diagrama de Voronoi es calcula l'àrea de la tessel·la definida per cada jugador, el càlcul es realitza descomponent cada tessel·la en triangles i calculant l'àrea de cada triangle, un cop calculades se sumen, obtenint així l'àrea que domina cada jugador al camp. En altres paraules, l'espai de què disposa cada jugador. La descomposició en triangles és possible perquè cada tessel·la és convexa i finita.

Per donar un context ofensiu del perill que representa l'àrea que domina un jugador, no es considera tota la regió definida pel camp de joc, sinó que s'agafa només la part de l'àrea en zona d'atac. S'estableix una divisió sobre l'àrea total del camp, s'agafa l'àrea que queda en l'últim terç del camp, agafant així gran part de l'espai que hi ha en zona ofensiva i una part de l'espai que es troba en la darrera part de la zona de creació. Una altra alternativa interessant seria la d'agafar únicament la de l'últim quart del camp, és a dir l'espai de la part de finalització del joc.

La diferència entre el primer indicador (considerant el darrer terç del camp de joc) i l'alternatiu (considerant únicament l'últim quart del camp) és que el primer pot ser més adient per un rival defensiu que acostuma a "posar l'autobús", és a dir una formació molt defensiva. Mentre que per altra banda, l'indicador alternatiu pot tenir més sentit quan el rival és un equip menys especulador i més ambiciós, també té més de sentit per jugades de contraatac on els espais s'engrandeixen.

Tot i això, aquesta ponderació és subjectiva a l'objectiu a estudiar, concretament al rival a estudiar i al pla de joc previst a executar. En aquest cas es defineix la divisió d'un terç.

D'aquesta forma calculem un indicador sobre l'espai que disposa un equip en zona ofensiva, aquest és un índex del perill que genera un equip en cada moment. L'indicador de l'espai es presenta essent una agrupació de l'espai dels jugadors d'un mateix equip, tot i això també resulta interessant diferenciar l'espai disposat per cada un dels jugadors.

En la nostra base de dades i les que s'utilitzen professionalment, cada moment equival a 0,04 segons, de tal forma que per obtenir un indicador sobre l'espai d'una jugada es calcula l'espai creat en cada moment i es realitza la mitjana pels moments que componen la jugada, és a dir es calcula la mitjana dels espais de tots els moments que defineixen la jugada.

Eficiència computacional de l'algorisme:

Detallat el procediment per tal de crear l'algorisme per calcular l'espai, el darrer pas consisteix a avaluar la seva capacitat computacional. Un punt molt important de tot algorisme, és el temps que tarda la seva execució per tal de saber si és aplicable o no en la pràctica, sobretot si cal aplicar-lo en temps real.

En aquest cas, s'exposa l'eficiència computacional per moment, en altres paraules: el temps que tarda l'algorisme a calcular l'espai que disposa cada jugador en un determinat moment. També es calcula el temps computacional en calcular l'espai total que s'ha acumulat durant la primera part, ja que potser no es disposen de les dades en temps reals i s'obtenen en finalitzar la primera meitat. Els resultats són els següents:

- Un moment determinat: 0,008 s
- Primera meitat del partit: 46 s

Per tant, es pot dir que l'algorisme presentat és eficient computacionalment, essent possible calcular l'espai en temps real. A més, si no es disposessin les dades posicionals fins un cop finalitzada la primera meitat, també es podria executar el mètode presentat en la mitja part de joc.

Capítol 5: SOFTWARE UTILITZAT

Per què Python i no R?

Durant el grau d'estadística el programari que s'ha utilitzat majoritàriament ha estat l'R, tot i això en aquest TFG s'ha utilitzat Python a l'hora d'aplicar la metodologia comentada.

Aquesta elecció ha estat causada principalment per dos motius. El motiu principal ha estat aprendre Python, ja que és un llenguatge molt utilitzat, tant en l'àmbit empresarial com en l'àmbit acadèmic, el qual no l'havia utilitzat fins al realitzar les pràctiques curriculars. Així que per continuar agafant experiència amb aquest llenguatge vaig decidir realitzar tot el TFG íntegrament amb Python.

L'altre motiu, tot i ser secundari, va estar que un cop dissenyada la idea a aplicar vaig trobar el paquet *Voronoi* de la llibreria *Scipy* i el seu disseny em va acabar de convèncer amb l'elecció del programari.

→ Llibreries de Python utilitzades

Per desenvolupar el treball i presentar els resultats s'han utilitzat diferents llibreries de Python, les quals es detallen a continuació:

- ✓ **Pandas:** Per la manipulació i anàlisi de les dades. Ofereix estructura de dades i operacions per manipular taules numèriques i sèries temporals.
- ✓ **Csv:** Per llegir fitxers CSV.
- ✓ **Matplotlib:** Per generar gràfiques a partir de dades contingudes en llistes o en vectors.
- ✓ **Numpy:** Per crear vectors i matrius de grans dimensions. Ofereix un gran conjunt de funcions matemàtiques per utilitzar.
- ✓ **Scipy:** Biblioteca on es troben algorismes matemàtics per optimitzar, interpolar, integrar, etc.
- ✓ **Math:** Mòdul estàndard de Python on es troben moltes funcions matemàtiques.
- ✓ **Sklearn:** Llibreria més famosa de Machine Learning per Python, conté moltes eines eficients de ML i de modelització estadística.
- ✓ **Seabron:** Per visualitzar les dades de forma atractiva, la llibreria està basada en matplotlib.
- ✓ **Plotly:** Per gràfics de qualitat i interactius. També s'utilitza per visualitzar les dades mitjançant taules.

Capítol 6: DADES

Origen

Un dels problemes més habituals en l'anàlisi no-professional és la manca de dades de qualitat per a la realització de l'estudi. En l'àmbit futbolístic aquesta problemàtica augmenta, ja que la recollida d'aquestes dades suposa un cost elevat, com a conseqüència la seva propietat té un cost massa car per a un particular.

Tot i això, gràcies a les dades de mostra que proporciona l'empresa *Metrica-Sports*, s'ha pogut procedir a la realització del treball.

Metrica-Sports és una de les empreses més importants en l'àmbit de Football Analytics, el seu projecte sorgeix l'any 2013 i es basa en proporcionar dades i en la seva visualització, a més d'aportar solucions millorant així el rendiment.

Dins del seu projecte de dades obertes per fomentar l'estadística esportiva, l'empresa va decidir publicar diferents arxius, entre els quals hi ha un partit on es proporciona la informació sobre totes les accions realitzades pels jugadors durant tot el partit i la posició que ocupen els jugadors en tot moment.

Descripció

En referència a les dades per un tema de confidencialitat, el partit està anonimitzat i per tant es desconeixen els equips i els jugadors. El partit està dividit en tres bases:

- *Events*: Descriu l'acció que es realitza amb la pilota (tipus i subtipus), en quin moment es realitza, on comença i on acaba, la seva duració i per quin jugador es realitza. (Veure Base de Dades 1)
- *Tracking_Home*: Proporciona la posició que ocupa cada jugador de l'equip local en cada moment. (Veure Base de Dades 2)
- *Tracking_Away*: Proporciona la posició que ocupa cada jugador de l'equip visitant en cada moment. (Veure Base de Dades 3)

En la base *Events*, cada fila representa una acció (conjunt de jugades fins que la pilota canvia d'equip). Les variables que es troben són:

- Team: Variable binària que representa quin equip realitza l'acció
- Type: Tipus de l'acció, descriu l'acció
- Subtype: Algunes accions tenen un subtipus
- Period: Variable binària. 1: primera part, 2: segona part
- Start Frame: Identificador d'on comença l'acció
- Start Time: Variable numèrica, en segons, que indica en quin moment comença l'acció.
- End Frame: Identificador d'on finalitza l'acció
- End Time: Variable numèrica, en segons, que indica en quin moment acaba l'acció.
- From: Indica el jugador que comença l'acció
- To: Indica el jugador que s'afavoreix de l'acció, el receptor
- Start X: Coordenada X on comença l'acció
- Start Y: Coordenada Y on comença l'acció
- End X: Coordenada X on finalitza l'acció
- End Y: Coordenada Y on finalitza l'acció

Per altra banda en la base de *Tracking*, el format és idèntic tant per l'equip local com per l'equip visitant. Cada fila representa un moment diferent, cada 0,04 s. El significat de les seves variables, és el següent:

- Period: Variable binària. 1: primera part, 2: segona part
- Frame: Identificador per cada extracció de la posició dels jugadors
- Time: Variable numèrica que representa el moment en què té lloc cada moment, s'expressa en segons.
- Player'n'_X: Coordenada X que ocupa el jugador n
- Player'n'_Y: Coordenada Y que ocupa el jugador n
- Ball_X: Coordenada X que ocupa la pilota
- Ball_Y: Coordenada Y que ocupa la pilota

Coordenades (X,Y):

Malgrat que les dimensions del camp de futbol són de 105x68 metres. En aquest cas, les coordenades representades van de 0 a 1 en cada eix, sent el (0,0) la part superior esquerra; (1,1) la part inferior dreta, i el (0.5,0.5) el punt mitjà del camp.

Team	Type	Subtype	Period	Start Frame	Start Time [s]	End Frame	End Time [s]	From	To	Start X	Start Y	End X	End Y	
0	Away	SET PIECE	KICK OFF	1	1	0.04	0	0.00	Player19	NaN	NaN	NaN	NaN	
1	Away	PASS	NaN	1	1	0.04	3	0.12	Player19	Player21	0.45	0.39	0.55	0.43
2	Away	PASS	NaN	1	3	0.12	17	0.68	Player21	Player15	0.55	0.43	0.58	0.21
3	Away	PASS	NaN	1	45	1.80	61	2.44	Player15	Player19	0.55	0.19	0.45	0.31
4	Away	PASS	NaN	1	77	3.08	96	3.84	Player19	Player21	0.45	0.32	0.49	0.47
...	
1740	Home	PASS	NaN	2	143361	5734.44	143483	5739.32	Player12	Player13	0.60	0.33	0.19	0.95
1741	Home	PASS	NaN	2	143578	5743.12	143593	5743.72	Player13	Player4	0.09	0.88	0.14	0.69
1742	Home	BALL LOST	INTERCEPTION	2	143598	5743.92	143618	5744.72	Player4	NaN	0.13	0.69	0.07	0.61
1743	Away	RECOVERY	BLOCKED	2	143617	5744.68	143617	5744.68	Player16	NaN	0.05	0.62	NaN	NaN
1744	Away	BALL OUT	NaN	2	143622	5744.88	143630	5745.20	Player16	NaN	0.05	0.63	0.03	1.01

1745 rows x 14 columns

BASE DE DADES 1: Base Events

Period	Frame	Time [s]	Player11_X	Player11_Y	Player1_X	Player1_Y	Player2_X	Player2_Y	Player3_X	...	Player10_X	Player10_Y	
2	1	1	0.04	0.00082	0.48238	0.32648	0.65322	0.33701	0.48863	0.30927	...	0.55243	0.43269
3	1	2	0.08	0.00096	0.48238	0.32648	0.65322	0.33701	0.48863	0.30927	...	0.55243	0.43269
4	1	3	0.12	0.00114	0.48238	0.32648	0.65322	0.33701	0.48863	0.30927	...	0.55243	0.43269
5	1	4	0.16	0.00121	0.48238	0.32622	0.65317	0.33687	0.48988	0.30944	...	0.55236	0.43313
6	1	5	0.2	0.00129	0.48238	0.32597	0.65269	0.33664	0.49018	0.30948	...	0.55202	0.43311

5 rows x 33 columns

BASE DE DADES 2: Base Tracking Home

Period	Frame	Time [s]	Player25_X	Player25_Y	Player15_X	Player15_Y	Player16_X	Player16_Y	Player17_X	...	Player24_X	Player24_Y	
2	1	1	0.04	0.90509	0.47462	0.58393	0.20794	0.67658	0.46710	0.6731	...	0.37833	0.27383
3	1	2	0.08	0.90494	0.47462	0.58393	0.20794	0.67658	0.46710	0.6731	...	0.37833	0.27383
4	1	3	0.12	0.90434	0.47463	0.58393	0.20794	0.67658	0.46710	0.6731	...	0.37833	0.27383
5	1	4	0.16	0.90377	0.47463	0.58351	0.20868	0.6764	0.46762	0.67279	...	0.37756	0.27473
6	1	5	0.2	0.90324	0.47464	0.58291	0.21039	0.67599	0.46769	0.67253	...	0.37663	0.27543

5 rows x 33 columns

BASE DE DADES 3: Base Tracking Away

Capítol 7: RESULTATS

Es divideix l'apart de resultats en quatre subapartats: descriptiu de les dades crues, quantificació de l'espai, utilització de l'espai i modelització de l'espai.

Descriptiva general

Abans d'aplicar l'algorisme presentat anteriorment per quantificar l'espai, es realitza un anàlisi descriptiu de les dades que disposem amb l'objectiu de fer-nos una idea general del què ha succeït al partit. Fent una taula de freqüències de les accions més importants d'un partit ens fem una idea del què ha succeït. Aquesta informació es presenta en la taula 1.

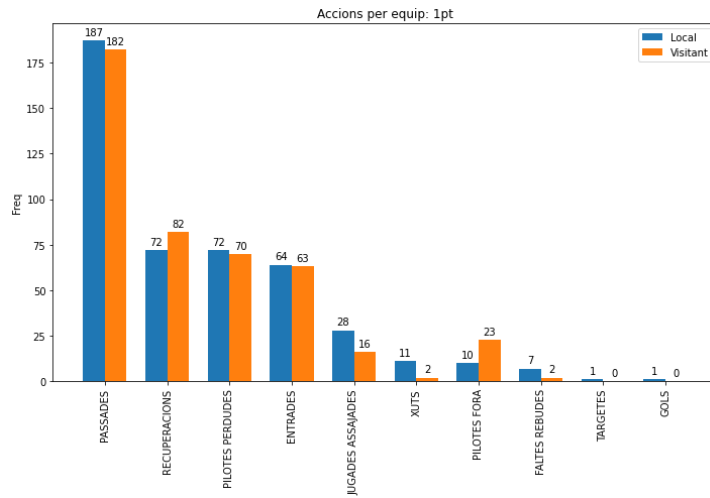
Acció	Local	Visitant
Gols	3	0
Xuts	18	6
Xuts a porteria	7	3
Possessió (%)	53	47
Passades	437	362
Faltes	7	15
Targetes grogues	2	2
Targetes vermelles	0	0
Fores de joc	0	1
Córners	8	3
Nº Jugades	917	828

TAULA 1: Freqüència d'accions per equip

A partir de la Taula 1 ja veiem que l'equip local ha guanyat i a més ha dominat el partit. L'equip de casa ha dominat en possessió, en passades i en xuts, aquest domini s'ha traduït en un resultat de tres gols a zero.

Un cop tenim una idea general del què ha passat durant els 90 minuts en el terreny de joc, realitzem dos gràfics de freqüències, el Gràfic 1 i Gràfic 2, sobre les accions més importants i més nombroses en cada part, l'objectiu d'aquests dos gràfics és visualitzar si existeixen diferències entre la primera i la segona part, identificar si una de les dues parts és el causant del rendiment global de l'equip visitant, o aquest ha estat constant les dues parts.

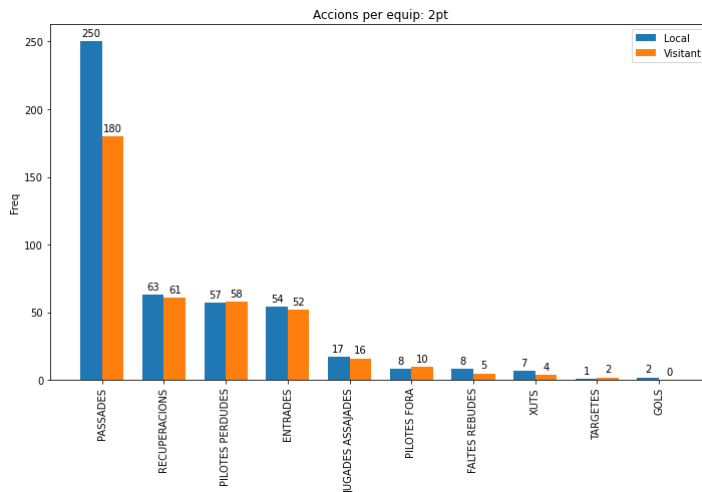
→ Primera part:



GRÀFIC 1: Número d'accions per equip en la 1ªpart

Durant els primers 45 minuts de joc, observem (veure Gràfic 1) una clara diferència en el nombre de xuts realitzats i en el nombre d'accions a pilota parada realitzades entre l'equip local i el visitant, aquesta diferència s'ha traduït en un gol a favor per l'equip de casa.

→ Segona part:



GRÀFIC 2: Número d'accions per equip en la 2ªpart

Malgrat que la diferència entre els xuts realitzats pels dos equips es redueix en la segona part, la diferència de gols marcats entre els dos equips puja, essent dos gols per l'equip local i zero pel visitant. A banda de la diferència entre els gols, el rendiment dels equips s'observa

en el nombre de passades realitzades per cada un, exactament en la segona part l'equip local ha realitzat 70 passes més que el visitant.

Un cop obtinguda una idea de com s'ha desenvolupat el partit, visualitzem la mediana i la variància dels moviments de cada jugador durant el partit amb l'objectiu de fer-nos així una idea de quina ha estat la zona per on s'ha mogut. A causa del canvi de camp en la mitja part, dividim l'anàlisi tant per equip com per primera i segona part.

És important recordar que a partir de la coordenada X ens fem una idea de la zona horitzontal per on s'ha mogut el jugador, mentre que la coordenada Y representa l'altura. Valors pròxims a 1 en l'eix X representen la porteria de la part dreta (105m), mentre que valors pròxims a 1 en l'eix Y representen l'alçada superior del camp (68m).

Per facilitar la interpretació, convertim l'eix Y en el valor absolut de $(Y-0.5)$, d'aquesta forma podrem saber si el jugador s'ha mogut pròxim a la banda o a la part central del camp, a més amb aquesta conversió no ens hem de preocupar de si el jugador es troba en la banda esquerra o en la dreta.

En primer lloc s'estudien els moviments associats a la coordenada X, per fer-nos una idea de la distància que separa la porteria del rival als jugadors que l'ataquen (veure Taula 2).

→ Primera part:

Primer comencem amb l'equip local, el primer que veiem és que a causa de que s'ha realitzat una substitució durant la primera part han participat 12 jugadors diferents durant la primera part.

	Player11	Player1	Player2	Player3	Player4	Player5
Variància	0,0068	0,0427	0,0382	0,0430	0,0290	0,0377
Mediana	0,1195	0,4313	0,3838	0,3816	0,4159	0,5507
	Player6	Player7	Player8	Player9	Player10	Player12
Variància	0,0394	0,0338	0,0418	0,0334	0,0273	0,0269
Mediana	0,5018	0,4516	0,5777	0,6305	0,6073	0,5078

TAULA 2: Variància i mediana dels moviments horitzontals dels jugadors locals 1ªpt

A partir de la mediana ens fem una idea de quina ha estat la posició que ha acostumat a ocupar un jugador, mentre que la variància ens guia si el jugador s'ha mogut molt o no en l'eix horitzontal del camp.

En la primera part, els jugadors que s'han situat més en la zona atacant han estat els jugadors Player9 i Player10, mentre que tots els jugadors mostren un grau de moviment semblant pel camp, a destacar el Player3 com el que més s'ha mogut.

En la Taula 3 visualitzem les mateixes dades per l'equip visitant.

	Player25	Player15	Player16	Player17	Player18	Player19
Variància	0,0063	0,0294	0,0233	0,0304	0,0412	0,0353
Mediana	0,8982	0,6588	0,6706	0,6568	0,5446	0,5342
	Player20	Player21	Player22	Player23	Player24	
Variància	0,0334	0,0425	0,0413	0,0330	0,0404	
Mediana	0,5929	0,5189	0,5890	0,4289	0,4227	

TAULA 3: Variància i mediana dels moviments horitzontals dels jugadors visitants 1ªpt

En aquest cas, veiem que els jugadors que han representat més perill han estat clarament el Player24 i Player23. Per altra banda els que han ofert més moviments, han estat el Player21 i Player22 de forma gairebé idèntica.

→ Segona part:

Visualitzant les dades de la segona part de l'equip local en la Taula 4, trobem moltes similituds en els moviments realitzats en la primera part.

	Player11	Player2	Player3	Player4	Player5
Variància	0,0044	0,0271	0,0278	0,0367	0,0414
Mediana	0,9159	0,6737	0,6821	0,6397	0,5308
	Player6	Player7	Player8	Player9	Player10
Variància	0,0456	0,0354	0,0437	0,0361	0,0384
Mediana	0,5849	0,6213	0,4805	0,4287	0,4832
	Player12	Player13	Player14		
Variància	0,0309	0,0247	0,0326		
Mediana	0,5635	0,4307	0,4617		

TAULA 4: Variància i mediana dels moviments horitzontals dels jugadors locals 2ªpt.

En el segon període el Player13 i Player14 han substituït al Player9 i Player10, i en atac s'han mogut a una altura de la porteria semblant a la que s'havien mogut els jugadors que han substituït.

I ja per finalitzar, en la Taula 5 veiem el comportament dels jugadors de l'equip visitant durant aquest segon període.

	Player25	Player15	Player16	Player17	Player18
Variància	0,0054	0,0370	0,0269	0,0470	0,0354
Mediana	0,1178	0,3907	0,3888	0,4513	0,5477
	Player19	Player20	Player21	Player22	Player23
Variància	0,0337	0,0323	0,0311	0,0420	0,0286
Mediana	0,5395	0,4102	0,4957	0,4699	0,6345
	Player24	Player26	Player27	Player28	
Variància	0,0376	0,0228	0,0350	0,0308	
Mediana	0,6286	0,6197	0,5247	0,5028	

TAULA 5: Variància i mediana dels moviments horitzontals dels jugadors visitants 2^opt.

Al Player23 i Player24, s'ha afegit el Player26 com a jugadors més perillosos per l'equip rival. En el segon període el Player17 s'ha anat movent més pel camp que durant la primera part del joc.

Un cop s'ha estudiat el moviment en l'eix horitzontal, s'estudia l'eix vertical. Per obtenir una idea de per quina alçada es mou cada jugador, en altres paraules quins són els jugadors que es mouen més per les bandes i quins ho fan més per la zona central, valors pròxims a 0 signifiquen la part central del camp, mentre que valors grans mostren que el jugador s'ha mogut més per la banda.

En referència a l'equip local, en la Taula 6 observem la distribució de la coordenada Y.

	Player11	Player1	Player2	Player3	Player4
Variància	0,0025	0,0243	0,0362	0,0193	0,0678
Mediana	0,0231	0,2568	0,1147	0,0842	0,2311
	Player5	Player6	Player7	Player8	Player9
Variància	0,0595	0,0395	0,0738	0,0878	0,0308
Mediana	0,2041	0,1306	0,2330	0,2734	0,1281

	Player10	Player12	Player13	Player14
Variància	0,0582	0,0282	0,0308	0,0258
Mediana	0,1732	0,1134	0,1091	0,2279

TAULA 6: Variància i mediana dels moviments verticals dels jugadors locals

Observem que els jugadors que es mouen més en posicions d'extrem són el Player8 i Player1, per altra banda el Player8 i el Player7 són els que mostren més variabilitat respecte aquest eix.

Per altra banda, per la part de l'equip visitant la distribució del moviment dels seus jugadors en l'eix Y queda de la següent manera:

	Player25	Player15	Player16	Player17	Player18
Variància	0,0048	0,0380	0,0153	0,0386	0,0616
Mediana	0,0268	0,1508	0,0731	0,2710	0,2115
	Player19	Player20	Player21	Player22	Player23
Variància	0,0380	0,0343	0,0447	0,0931	0,0399
Mediana	0,1262	0,1340	0,1555	0,3116	0,1431
	Player24	Player26	Player27	Player28	
Variància	0,0323	0,0201	0,0169	0,0320	
Mediana	0,1283	0,0666	0,2691	0,1485	

TAULA 7: Variància i mediana dels moviments verticals dels jugadors visitants

En aquest cas el Player22 és el que ha tingut una presència més alta en el joc de banda, a més és el que s'ha mogut més en l'eix Y.

Quantificació de l'espai

Aquest subapartat es divideix en tres subapartats segons el conjunt de l'espai estudiat: pel conjunt de l'equip, pel porter i pels jugadors.

Espai pel conjunt de l'equip

Un cop introduïdes les dades i realitzada la descriptiva per obtenir una idea general del que ha succeït en el terreny de joc, s'aplica el mètode descrit anteriorment per obtenir l'espai que ha generat cada equip durant el partit. Aquesta mesura està representada en m².

Seguint amb l'objectiu de la mesura creada, l'aproximació sobre l'espai es calcularà en les accions on l'equip hagi disposat de la pilota, ja que l'objectiu és utilitzar aquesta mètrica com una referència sobre el perill ofensiu generat per un equip.

Aplicat en totes les accions on l'equip ha estat el protagonista, on ha tingut la pilota en zona ofensiva (darrer terç del camp), es realitza la corresponent suma acumulada de l'espai per obtenir l'espai que ha generat cada equip durant el partit. Els resultats obtinguts han estat els que es presenten en la Taula 8:

Equip Local	Equip Visitant
410.786,3 m ²	347.903,9 m ²

TAULA 8: Espai total generat durant el partit

S'obté que l'equip local, el vencedor del partit per 3-0, ha estat clarament superior durant els noranta minuts també a l'hora de generar espais en zona ofensiva. En altres paraules, l'equip local ha disposat de més metres quadrats en zona d'atac respecte a l'equip visitant, aquesta diferència concorda amb el resultat del partit i amb les ocasions que ha generat cada equip. Per tal d'identificar alguna possible diferència entre el primer i el segon període, es descompon l'espai entre les dues parts per veure així si la diferència entre els equips ha estat constant (veure Taula 9).

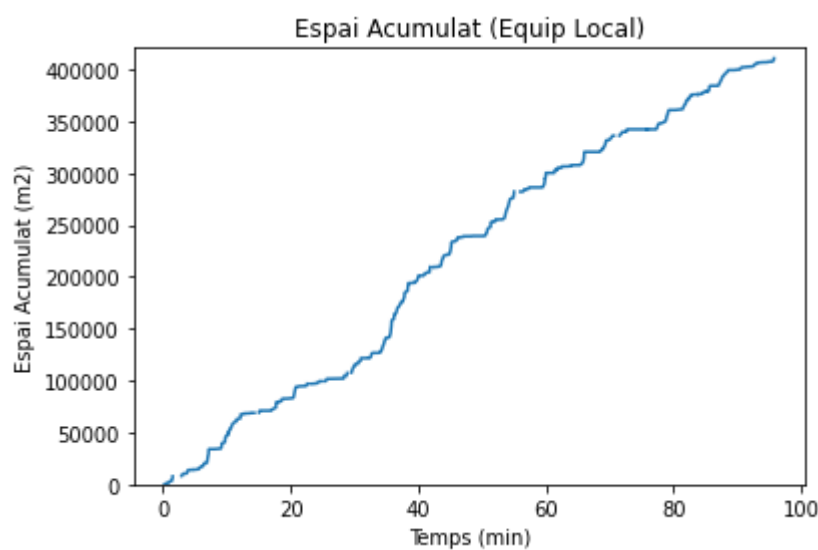
	Equip Local	Equip Visitant
1ª Part	239.074,3 m ²	157.527,4 m ²
2ª Part	171.712 m ²	190.376,5 m ²

TAULA 9: Espai generat en cada part

Tot i que en referència als gols marcats la diferència més gran ha estat en la segona part, ja que se n'ha marcat un més que en la primera part, tant en el nombre de xuts realitzats com en l'espai generat la diferència més important ha estat en el primer temps. La qual cosa ens porta a pensar que la primera part ha estat el moment en què l'equip local ha generat més perill tot i que no ha estat tan encertat de cara a porteria com en la segona part.

Amb l'objectiu de poder identificar els bons i mals moments de l'equip a l'hora de generar espai durant el partit, es realitza el gràfic de l'espai acumulat durant el transcurs dels noranta minuts.

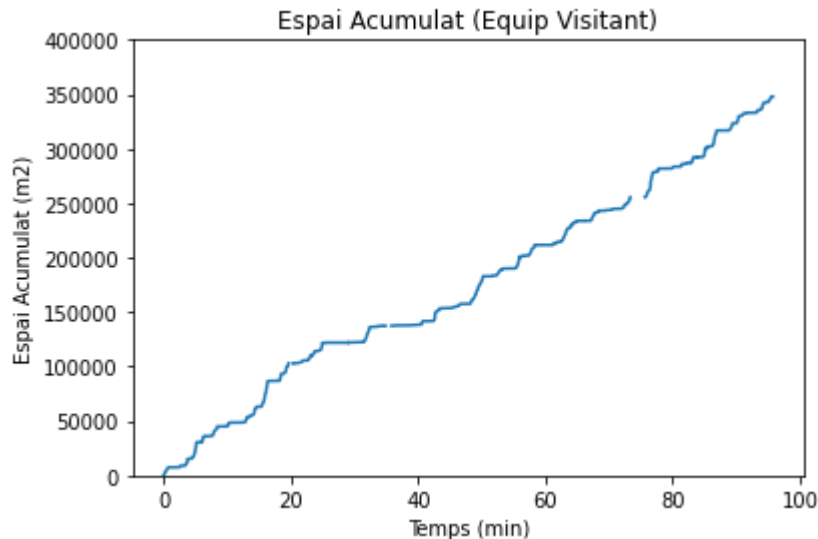
En primer lloc es realitza el gràfic de l'equip local (veure Gràfic 3).



GRÀFIC 3: Evolució de l'espai acumulat per l'equip local

Visualment veiem que el moment en que l'equip local ha vist incrementat de manera molt considerable l'espai generat ha estat en l'interval entre els minuts trenta-cinc i quaranta de la primera part. Per altra banda en l'interval entre el minut seixanta (després de marcar el segon gol) i el vuitanta, l'espai generat per l'equip local ha estat força reduït, la qual cosa podria ser deguda a un relaxament causat per l'avantatge en el marcador.

Estudiat el comportament de l'espai local, es realitza el mateix gràfic (veure Gràfic 4), però per l'equip visitant per veure on se situen els bons i mals moments.



GRÀFIC 4: Evolució de l'espai acumulat per l'equip visitant

Visualment veiem que el moment més crític per l'equip visitant ha estat entre el minut vint i quaranta de la primera part, en aquest interval l'espai generat ha estat molt poc. Per altra banda, es visualitza que la reacció de l'equip visitant ha estat en el darrer tram de partit (els darrers trenta minuts), observem que es visualitza un creixement de l'acumulació de l'espai total. Tal com ens indica el resultat aquesta reacció ha arribat massa tard.

Calculat l'espai generat per cada equip i visualitzats els moments on els equips han estat més i menys encertats, s'intenta cerca alguna relació entre l'espai i el nombre de passes realitzats i també entre l'espai i la distància recorreguda.

Com a aficionat al futbol, crec fermament en la pràctica del "joc de posició" com a estil de joc (segons Albert Capellas: "el joc de posició és un concepte global, ofensiu i defensiu a la vegada, basat en controlar el joc a través de la possessió de la pilota, la posició dels jugadors i la pressió a la pilota."). Aquest estil de joc considera que els avantatges en una jugada es creen mitjançant passes i treballant la jugada, com més passes es realitzin més espai es generarà a causa que la defensa rival es desorganitzarà. A més a més, crec que cada cop es valora més als equips que corren quan hauria de ser una mala virtut, ja que si et posicones correctament minimitzes la distància a recórrer.

Seguint aquesta idea el primer a mirar és el grau de correlació entre l'espai acumulat i el nombre de passades acumulades, separant l'anàlisi per equip.

En els dos casos, aplicant la correlació de Spearman, obtenim un grau de correlació molt gran entre les dues variables.

En ambos casos, tant referent a l'equip local com a l'equip visitant, s'obté una correlació de 0,99. La qual cosa ens indica un grau d'associació molt elevat entre la variable referent a les passades i la referent a l'espai calculat.

Seguint amb la idea introduïda sobre el "joc de posició" s'introdueix la variable distància recorreguda, que representa la distància que recorre en cada moment l'equip.

Primer de tot es calcula la distància total que recorre un equip independentment de si controla o no la pilota. La distància per equip obtinguda és la que presenta en la Taula 10.

Equip Local	Equip Visitant
108.945,4 m	112.460,2 m

TAULA 10: Distància recorreguda total durant el partit

S'obté que l'equip visitant, el que ha produït menys espai, ha estat el que més distància ha recorregut durant el partit.

Continuant amb la idea que l'espai generat (i el perill ofensiu) no depenen de la distància recorreguda per l'equip, es realitza la correlació de Spearman entre el variable espai i la variable distància recorreguda per acció.

Altre cop s'avalua l'estadístic de correlació en les dades referents als dos equips. En els dos casos s'obté un resultat semblant, una correlació de 0,11 en les dades que fan referència a l'equip local i de 0,12 en les dades de l'equip visitant, la qual cosa indica una associació feble entre les dues variables.

Espai porters

Tot i que el mètode exposat per calcular l'espai fa referència a la zona ofensiva, també es pot calcular per qualsevol jugador del camp independentment de la posició que ocupa. En particular es decideix calcular l'àrea (espai) que ocupa cada porter en cada moment independentment de si el seu equip disposa o no de la pilota, després se sumen tots els moments obtenint així l'espai que ha anat acumulant cada porter durant tot el partit. Notem que això ve donat per les tesselles de Voronoi definides per cada porter.

La idea d'aquesta mètrica recau en el fet que un major espai per un dels dos porters significa que l'equip del porter que té més espai és el que està jugant a camp contrari, en conseqüència està generant més perill que el seu rival. També és una bona mesura per veure ràpidament en quina línia planta la defensa cada equip, com més espai

disposa el porter, més lluny es troba la línia defensiva i també serveix per identificar jugades de perill, ja que en accions de perill el porter que defensa l'acció disposarà d'un espai molt petit.

En el partit es calcula la mitjana de l'àrea controlada pels porters durant les accions de joc i també la controlada en tot el partit, la qual representa la suma de l'àrea que es controla en cada moment. Els resultats es presenten en la Taula 11.

	Equip Local	Equip Visitant
Mitjana	1.322,8 m ²	1.209,2 m ²
Acumulació	54.274.253,1 m ²	49.614.941,9 m ²

TAULA 11: Espai mitjà i acumulat pels dos porters durant el partit

Amb l'espai creat per cada porter arribem a la mateixa conclusió, l'equip local ha creat més espai respecte a l'equip visitant. En altres paraules, els jugadors de l'equip local han estat més estona a camp contrari que els jugadors rivals, d'aquesta forma els defenses locals s'han situat en posició més avançada i la distància entre els defenses i el porter ha estat més elevada en l'equip local que en l'equip visitant.

Espai jugadors

Malgrat que l'objectiu principal en quantificar l'espai és que es mostri com una mesura col·lectiva, ja que un bon posicionament ha de ser col·lectiu i no individual. Aquesta mesura també es pot descompondre de forma individual analitzant l'àrea de les tesselles de Voronoi que van definint en cada moment, amb l'objectiu d'identificar si algun jugador de zona ofensiva (sols es considera l'espai que es genera en el darrer terç del camp) està oferint un rendiment per sota de les seves possibilitats i seria possible una millora en la seva col·locació dins del terreny de joc, per tant generalment l'espai del porter serà nul.

En primer lloc es realitza la descomposició de l'espai generat per l'equip local entre els jugadors que han pres part en el partit. Com que ja s'ha vist bé l'espai generat en la primera part no ha estat proporcional al de la segona part, es descompon altre cop entre la primera i la segona part. Els resultats obtinguts han estat els que es presenten en la Taula 12.

→ Primera part:

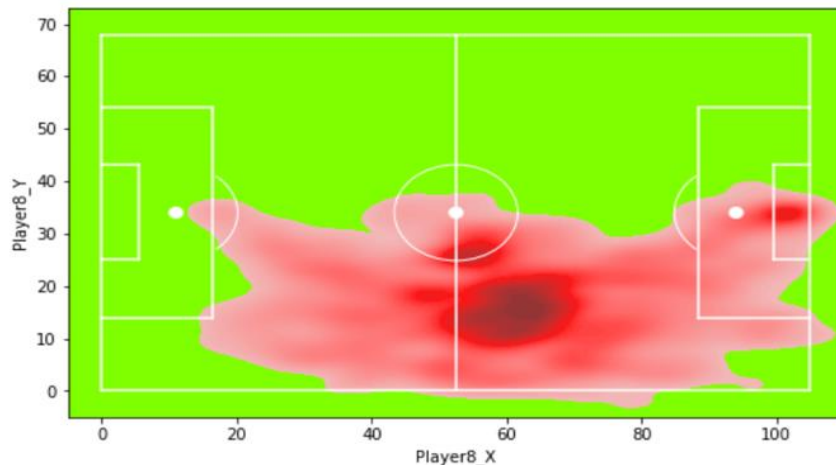
(*Out*: Jugador ha estat substituït, *In*: Jugador ha entrat al camp)

Player11 (Porter)	0 m ²
Player1 (Out)	3.784,8 m ²
Player2	8.699,6 m ²
Player3	2.312,9 m ²
Player4	8.260,8 m ²
Player5	39.473,8 m ²
Player6	12.379,4 m ²
Player7	21.787,1 m ²
Player8	77.350,1 m ²
Player9	27.256 m ²
Player10	31.449 m ²
Player12 (In)	6.320,9 m ²
Player13	0 m ²
Player14	0 m ²

TAULA 12: Espai disposat per cada jugador local durant la 1ª pt

Descompost l'espai total creat en la primera part es veu com que el Player8 ha estat el generador principal de perill, generant quasi el doble d'espai respecte del Player5, el qual ha estat el segon jugador que n'ha produït més.

Per visualitzar els moviments del Player8, es visualitza el seu mapa de calor durant els quaranta-cinc primers minuts (veure Mapa de Calor 1).



MAPA DE CALOR 1: Moviments del Player8 durant la 1ª pt

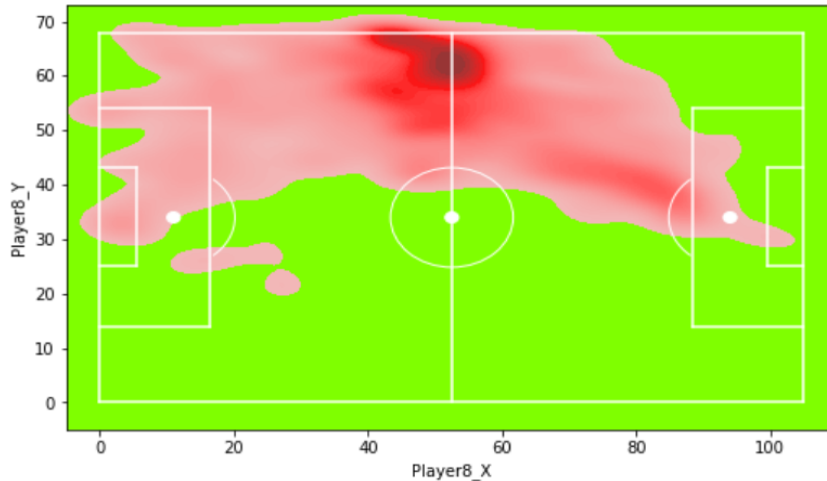
A través del mapa de calor es pot visualitzar per quina zona s'ha mogut el Player8, es visualitza que el seu atac s'ha focalitzat en la banda dreta. Per altra banda, a partir d'aquest gràfic i aquests resultats l'equip rival hauria de realitzar alguna modificació tàctica per poder aturar aquesta allau de perill del Player8.

→ Segona part:

Player11 (Porter)	0 m ²
Player1	0 m ²
Player2	1.145,3 m ²
Player3	1.132,1 m ²
Player4	12.476,3 m ²
Player5	22.735,3 m ²
Player6 (Out)	4.283,5 m ²
Player7	11.649,8 m ²
Player8	45.814 m ²
Player9	19.858 m ²
Player10 (Out)	20.813,4 m ²
Player12	5.071,4 m ²
Player13 (In)	13.006,2 m ²
Player14 (In)	13.726,7 m ²

TAULA 13: Espai disposat per cada jugador local durant la 2^o pt

A partir de la Taula 13, es veu com altre cop el Player8 ha estat el jugador que més espai ha disposat i altre cop el Player5 ha estat el segon que més n'ha realitzat, tot i això aquest espai s'ha vist reduït en comparació al disposat en la primera meitat. Aquesta reducció pot ser deguda a una bona modificació tàctica de l'entrenador visitant, o bé pot estar generada per una relaxació dels jugadors local al disposar d'un marcador ampli. Visualitzem altre cop el mapa de calor del Player8 per veure com han estat els seus moviments en aquesta segona part (veure Mapa de Calor 2).



MAPA DE CALOR 2: Moviments del Player8 durant la 2^opt

La primera idea a extreure és que el Player8 no ha produït tant de perill com en la primera part, aquest cop no ha tingut tanta presència en el darrer terç d'atac i la seva posició ha estat reduïda uns metres, movent-se més lluny de la porteria rival.

En referència a l'equip visitant, les dades es mostren en les Taules 14 i 15.

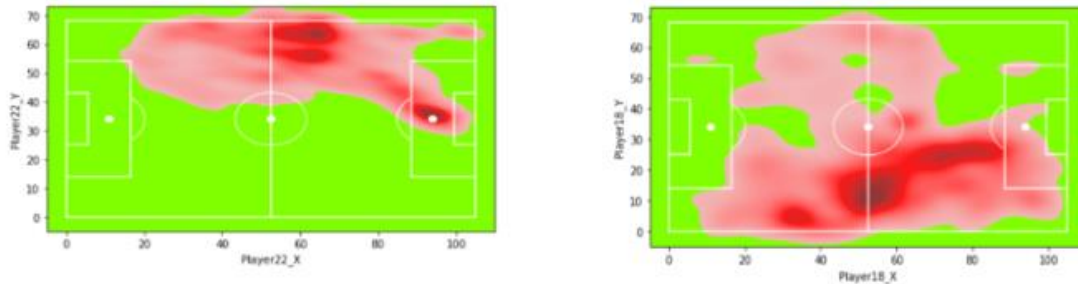
→ Primera part:

Player25 (Porter)	0 m ²
Player15	614,3 m ²
Player16	74,4 m ²
Player17	2.744,66 m ²
Player18	37.361,1 m ²
Player19	8.600,7 m ²
Player20	1.636,1 m ²
Player21	9.754,4 m ²
Player22	44.794,3 m ²
Player23	31.918,1 m ²
Player24	20.029,3 m ²
Player26	0 m ²
Player27	0 m ²
Player28	0 m ²

TAULA 14: Espai disposat per cada jugador visitant durant la 1^a pt

El jugador visitant més perillós durant la primera part ha estat el Player22, seguit de molt a prop pel Player18.

Per fer-nos una idea dels moviments dels dos jugadors, visualitzem el mapa de calor dels dos jugadors.



MAPA DE CALOR 3: Moviments del Player 22 (esquerra) i Player 18(dreta) durant la 1ª pt

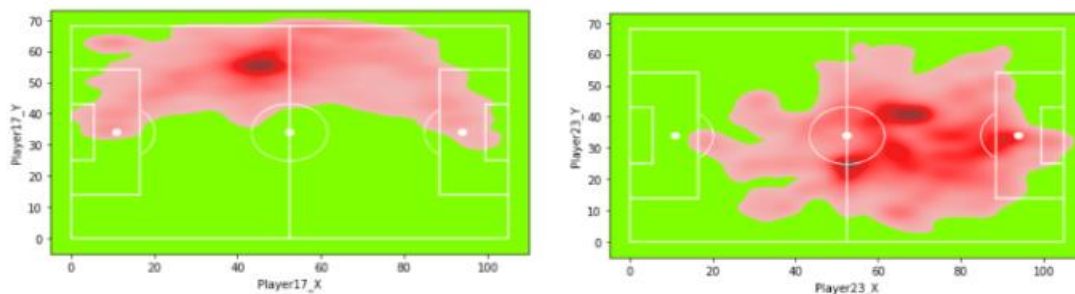
Mitjançant els Mapa de Calor 3, es veu com els dos jugadors no han tingut tanta presència en àrea atacant com n'ha tingut el Player8 de l'equip local, recordar que l'equip visitant en la primera part ha atacat a la porteria esquerra, on la presència de calor ha estat gairebé nul·la. La principal diferència entre aquests dos jugadors (Player22 i Player18) ha estat que tot i que el primer ha disposat més espai, els moviments del segon han estat més amplis i ha tingut presència en més llocs del camp, movent-se amb més llibertat. També es pot identificar com el Player24, jugador d'atac, ha disposat una quantitat d'espai força reduïda en comparació a la dels seus companys d'atac.

→ Segona part:

Player25 (Porter)	0 m ²
Player15	1.798,2 m ²
Player16	608 m ²
Player17	3.6334 m ²
Player18	26.323,6 m ²
Player19 (Out)	8.391,2 m ²
Player20	1.870,7 m ²
Player21	9.870,8 m ²
Player22 (Out)	15.887,6 m ²
Player23	36.326,8 m ²
Player24 (Out)	11.497,8 m ²
Player26 (In)	11.419,5 m ²
Player27 (In)	24.983,3 m ²
Player28 (In)	5.065,1 m ²

TAULA 15: Espai disposat per cada jugador visitant durant la 2ª pt

En l'equip local, els jugadors que més espai havien disposat durant la primera part també han estat els protagonistes en la segona, cosa que no ha succeït en l'equip visitant. El Player 22 ha estat substituït mentre que el Player18 ha reduït la seva presència en zona d'atac. També ha estat substituït el Player24, el qual havia tingut un impacte mínim en la primera part. Durant aquests darrers quaranta-cinc minuts els jugadors que més espai han disposat han estat el Player23 i el Player17, els dos han fet un pas endavant en aquesta segona part, ja que han augmentat de forma considerable l'espai ofensiu, a més han disposat més perill en porteria rival. Aquest augment junt amb la disminució de l'espai del Player18 ens fan pensar que l'entrenador visitant ha realitzat alguna modificació en la posició dels seus jugadors. Visualitzem els mapes de calor per fer-nos una idea dels moviments del Player23 i el Player17.



MAPA DE CALOR 4: Moviments del Player 17 (esquerra) i Player 23(dreta) durant la 2º pt

En aquest segon període els jugadors visitants han tingut més presència en zona d'atac, sobretot el Player23 el qual ha ofert una gran varietat de moviments per tot el centre del camp. El Player17 ha tingut un impacte en l'extrem esquerre on ha tingut un recorregut molt ampli per tota la banda.

Optimització (Utilització de l'espai)

Un cop l'espai ofensiu ha estat creat mitjançant un bon posicionament dels jugadors, l'equip ha de moure la pilota amb sentit per aprofitar l'espai, amb l'objectiu final de convertir la jugada en gol.

Seguint amb aquesta línia, es descompon l'espai segons la posició de l'eix vertical, exactament en tres zones: zona superior, zona central i zona inferior, aquestes tres zones del camp es divideixen en parts iguals. Mitjançant aquesta descomposició, a part de saber si s'ha aprofitat l'espai, també es pot estudiar si existeixen diferències entre zones i valorar més a fons l'actuació dels jugadors encarregats d'atacar o cobrir cada zona.

A banda de diferència entre equip local i equip visitant, també és pertinent diferenciar entre primera i segona part.

→ Equip local:

En referència a l'equip local, durant la primera part s'ha vist que havia generat un espai total de 239.074,3 m² en l'últim terç ofensiu. Per contextualitzar aquest espai, es descompon aquest número entre les tres zones comentades anteriorment. Un cop diferenciat l'espai entre zona superior, zona central i zona inferior, la distribució de l'espai total és la de la Taula 16.

Espai Superior	84.482,7 m ²
Espai Central	46.278,4 m ²
Espai Inferior	108.313,1 m ²

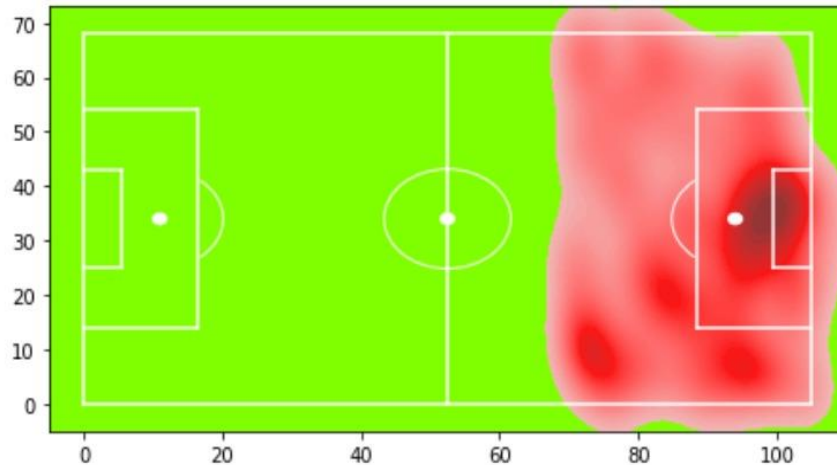
TAULA 16: Distribució per zones de l'espai generat per l'equip local durant la 1ª pt

A partir d'aquesta distribució de l'espai s'extreu la primera idea, l'equip local durant la primera part s'ha mogut més per la banda dreta, la qual representa l'espai inferior en la primera part. Respecte a l'àrea situada en la zona central és important contextualitzar aquest nombre, ja que l'àrea central sempre serà inferior respecte a la de les bandes ja que en l'àrea del porter acostuma a haver-hi una gran acumulació de defenses. Per tant, no té molt de sentit comparar l'espai d'aquesta zona amb el creat en els extrems del camp.

Estudiada la distribució de l'espai creat per l'equip local en la primera part, es visualitza el mapa de calor (veure ANNEX 2) de la pilota en les accions ofensives locals durant el primer període, amb l'objectiu de veure si la pilota ha utilitzat la mateixa distribució que l'espai,

concretament si l'equip local ha aprofitat la superioritat proporcionada pel posicionament dels seus jugadors.

Visualitzem el mapa de calor (veure Mapa de Calor 5) que ha seguit la pilota en les accions locals de la primera part en l'últim terç de camp (part ofensiva):



MAPA DE CALOR 5: Moviments de la pilota en possessió local durant la 1ª pt

Mitjançant el gràfic s'extreuen dues idees clau, la primera és que l'equip local tal com ja s'havia vist abans ha realitzat molt de perill durant la primera part, aquesta hipòtesi es corrobora amb el mapa de calor, en la zona on es troba la porteria rival es visualitza una intensitat molt alta, en altres paraules que la pilota hi ha arribat molt. La segona idea extreta mitjançant el mapa de calor és que aquest perill s'ha generat majoritàriament per la banda dreta, on l'equip de casa ha tingut més presència respecte a la banda esquerra, per tant els jugadors locals han aprofitat l'espai que s'ha generat en aquesta part del camp.

Estudiada la distribució que segueix l'espai en l'eix Y durant la primera part, es fa el mateix estudi, però en la segona part (veure Taula 17).

Anteriorment ja s'ha vist com en la segona part s'ha reduït l'espai generat per l'equip local en comparació al realitzat en la primera part. Tot i això, la distribució total de l'espai generat per l'equip local durant la segona meitat ha estat la següent:

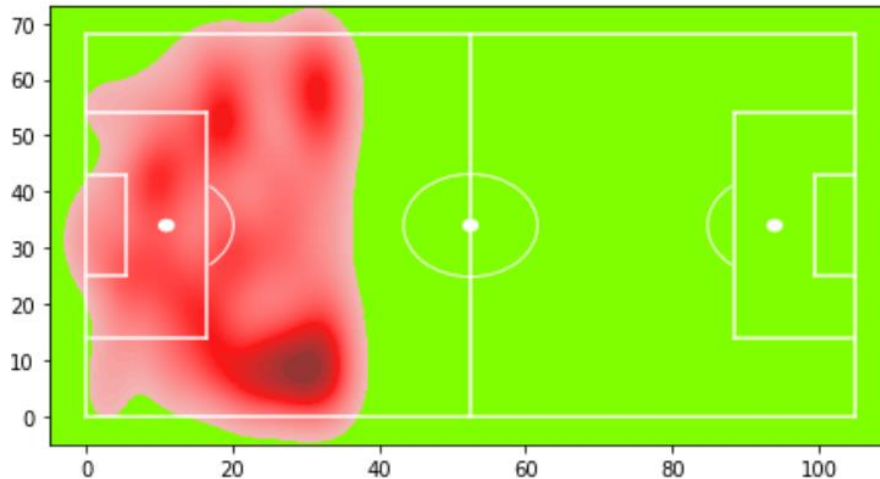
Espai Superior	75.030,7 m ²
Espai Central	32.108,7 m ²
Espai Inferior	64.572,6 m ²

TAULA 17: Distribució per zones de l'espai generat per l'equip local durant la 2ª pt

És important recordar que els jugadors que en la primera part han generat l'espai inferior, pel canvi de camp entre meitat, ara són els protagonistes de l'espai superior. En relació amb la

distribució de l'espai generat durant la primera part veiem que és força semblant en escala relativa al d'aquesta segona part.

Per saber si l'espai superior ha estat aprofitat al màxim, es visualitza el mapa de calor que segueix la pilota en aquesta segona part (veure Mapa de Calor 6).



MAPA DE CALOR 6: Moviments de la pilota en possessió local durant la 2ªpt

Visualitzant el mapa de calor de la trajectòria que segueix la pilota en la zona d'atac i l'espai generat en aquesta segona part, les dues idees més rellevants a extreure són: l'equip local ha generat menys perill, la intensitat de la zona vermella no és tan intensa dins l'àrea rival com ho era en la primera part. L'altra idea important que s'extreu és que en aquesta segona part no ha aprofitat la distribució de l'espai, distribuint la pilota majoritàriament per la part inferior del camp, tot i tenir més espais en la zona superior.

→ Equip visitant:

Un cop analitzat com s'ha distribuït l'espai local durant el partit i amb quin nivell s'ha aprofitat, s'elabora la mateixa anàlisi per a l'equip visitant, el qual com ja s'ha vist en els primers apartats ha generat menys espai en comparació a l'equip local.

Primer de tot s'estudia la distribució de l'espai generat durant la primera meitat, recordar que en aquesta part l'equip visitant ataca a la porteria situada a l'esquerra del camp.

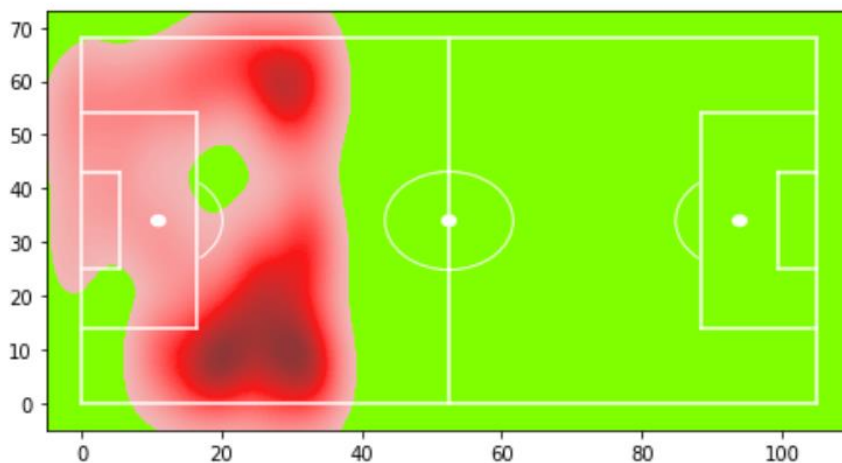
L'espai total generat per l'equip visitant en aquesta primera part es distribueix segons es mostra en la Taula 18.

Espai Superior	74.396,5 m ²
Espai Central	25.352 m ²
Espai Inferior	57.779,8 m ²

TAULA 18: Distribució per zones de l'espai generat per l'equip visitant durant la 1^a pt

Mitjançant la distribució de l'espai generat s'observa com l'equip visitant durant la primera part gairebé ha generat el mateix espai en la zona superior que la suma de l'espai generat en la part central més l'inferior. Per tant, es pot parlar d'una mala actuació en termes de col·locació i moviment tant en la zona central com en l'inferior del camp durant aquesta primera part, els jugadors situats en la banda esquerra de l'equip visitant no han tingut una actuació notable en aquests quaranta-cinc primers minuts. En referència a l'espai central tal com s'ha comentat anteriorment no és comparable amb el generat en les bandes tot i això el rendiment en aquesta zona també és millorable.

Malgrat aquesta possible millora, es visualitza el moviment de la pilota en l'últim terç ofensiu quan ataca l'equip visitant, per visualitzar si l'equip ha aprofitat o no l'espai situat en l'àrea superior.



MAPA DE CALOR 7: Moviments de la pilota en possessió visitant durant la 1ª pt

Mirant el mapa de calor, es veu com el perill generat per l'equip visitant de cara a porteria durant aquest primer temps ha estat molt baix, la intensitat (calor) en l'àrea rival és molt feble, aquest fet ja s'havia vist anteriorment (Taula 9). La novetat observada amb el mapa de calor és que tot i que gairebé la meitat de l'espai generat s'ha localitzat en la part superior, generalment l'equip visitant ha distribuït la pilota per la part inferior, desaprofitant així l'àrea disponible en la part superior.

Comparant el present mapa de calor amb el de l'equip local en la primera part, s'observa com en aquest cas a diferència de l'equip local, l'equip visitant ha mogut molt poc la pilota en algunes parts del camp.

Un cop vist com s'ha distribuït i com s'ha desaprofitat l'espai generat per l'equip visitant durant la primera part, s'elabora la mateixa anàlisi sobre l'espai generat durant la segona part.

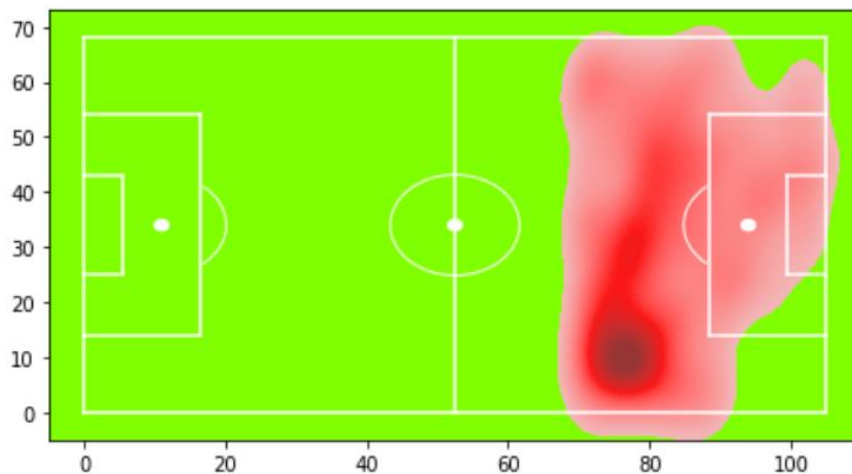
En aquesta segona part l'equip visitant ha augmentat l'espai generat respecte al generat durant el primer temps. Per intentar explicar aquest augment d'espai descomponem l'espai total en les tres zones de l'eix vertical (Taula 19).

Espai Superior	70.650,2 m ²
Espai Central	38.094,8 m ²
Espai Inferior	81.631,5 m ²

TAULA 19: Distribució per zones de l'espai generat per l'equip visitant durant la 2^o pt

A partir dels resultats obtinguts es veu com els jugadors que en la primera part se situen en la part superior continuen essent els que més espai generen, aquest cas en la zona inferior. Per altra banda, es veu com els jugadors que estaven situats en la part inferior i en la part central durant la primera part, els quals generaven poc espai, han millorat la seva actuació generant més espai en la part central i en l'espai superior. Sembla que aquests jugadors han estat els causants de l'augment d'espai entre la primera i la segona part.

Un cop analitzada la distribució de l'espai visitant en aquesta segona part, es realitza el corresponent mapa de calor per veure si aquest cop s'ha aprofitat.

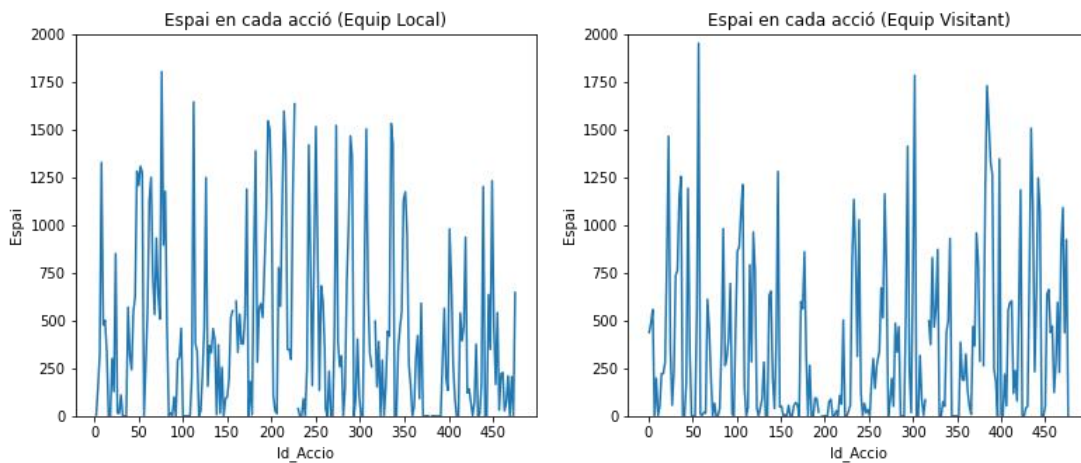


MAPA DE CALOR 8: Moviments de la pilota en possessió visitant durant la 2^o pt

A partir del Mapa de Calor 8 es veu com aquest cop sí que s'ha aprofitat de forma més considerable l'espai generat, tot i això també sembla que s'ha sobreexplotat la zona inferior i s'hauria d'haver intentat jugar també per la zona superior on s'ha generat bona quantitat

d'espai. Un altre punt important que es visualitza amb el mapa és que els extrems drets de l'equip visitant en comptes de jugar tant pel centre, havien d'haver buscat més l'u contra u en aquesta banda dreta encarant els defenses de l'equip local, amb la finalitat de buscar més el joc a la banda dreta de la porteria rival. Tot i que s'ha generat poc perill, s'observa que ha estat més gran en la segona part que a la primera part.

A vegades ens pot interessar fer una anàlisi menys general, d'un període de temps menys curt, amb la idea d'obtenir resultats durant el partit i no haver d'esperar a la mitja part per veure si s'ha de realitzar algun canvi en el sistema de joc. Un altre enfocament pot ser el d'elaborar la mateixa anàlisi anterior però en comptes per cada part, per cada 10/15 minuts. Amb aquest objectiu es visualitza la sèrie de l'espai obtingut en cada acció. En el Gràfic 5 es visualitzen els resultats pels dos equips.



GRÀFIC 5: Espai generat per equip en cada acció (ordenades cronològicament)

Aquest tipus de gràfic ens permet visualitzar les accions més crítiques, on l'espai era molt elevat i les accions poc efectives on s'ha generat poc espai.

Tenint en compte que l'objectiu és estudiar si l'espai generat ha estat o no aprofitat, ens quedem amb les accions en què l'espai ha estat elevat per veure si en aquests moments s'ha jugat amb criteri, aprofitant els espais lliures.

La metodologia utilitzada és la mateixa que l'aplicada anteriorment, però ara en comptes d'agafar tota la part s'agafa un interval d'accions en què l'espai ha estat elevat.

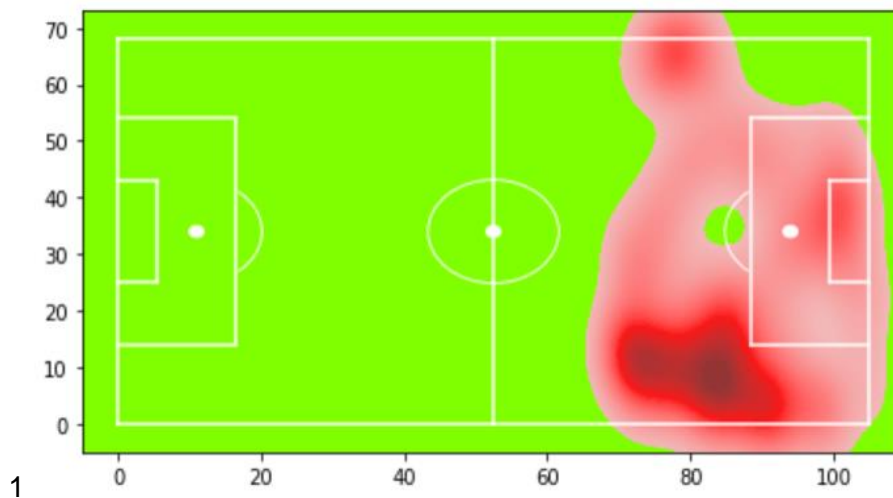
Com a exemple, s'utilitza per a cada equip el conjunt d'accions en què l'equip ha generat més espai durant tot el partit. Es visualitza el resum en forma de taula de l'interval de moviments i el corresponent mapa de calor de la pilota durant l'acció.

→ Equip Local:

Per l'equip local s'escull l'interval de joc entre els minuts 32 i 39 (veure Taula 20).

Nombre d'accions	19
Espai	15.314,9 m ²
Espai Superior	5.553,1 m ²
Espai Central	3.215,4 m ²
Espai Inferior	6.546,4 m ²

TAULA 20: Resum del joc de l'equip local entre el min. 32 i 39



MAPA DE CALOR 9: Moviments de la pilota en possessió local entre el min. 32 i 39

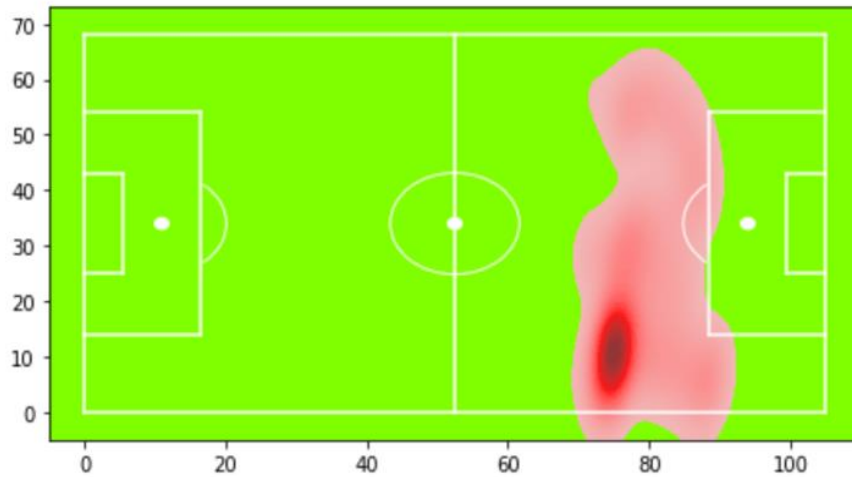
A partir de la Taula 20 es visualitza la distribució de l'espai generat en aquest interval de set minuts. A través del Mapa de Calor 9, sembla que l'equip local ha jugat amb força criteri en aquest interval de temps i el perill ha arribat dins l'àrea rival. Tot i això hagués estat més adient jugar una mica més en la banda esquerra (eix superior) on també s'ha generat molt espai.

→ Equip Visitant:

Per part de l'equip visitant com a exemple se selecciona un interval dels darrers moments del partit. Exactament entre els minuts 83 i 86 que es mostra en la Taula 21.

Nombre d'accions	11
Espai	6725 m ²
Espai Superior	2449,7 m ²
Espai Central	1147,3 m ²
Espai Inferior	3128 m ²

TAULA 21: Resum del joc de l'equip visitant entre el min. 83 i 86



MAPA DE CALOR 10: Moviments de la pilota en possessió visitant entre el min. 83 i 86

La idea més important a extreure a partir del mapa de calor és que el perill generat per l'equip visitant en aquest interval no li ha permès arribar fins a l'àrea rival. Això pot ser degut a no haver aprofitat de forma correcta l'espai generat, tot i ser on més espai s'ha creat s'ha focalitzat massa el joc en la banda dreta, hagués estat més oportú jugar també per l'extrem esquerre, on els jugadors també disposaven de bona part de l'espai total.

Modelització

Per intentar buscar alguna explicació sobre l'espai creat en el partit, a banda de la posició que ocupen els jugadors en el camp, es modelitza l'espai generat. L'objectiu és poder analitzar l'espai que ha generat cada equip, identificar possibles errors de posició i alhora enfortir les accions positives dutes a terme.

El primer punt a comentar és que s'agrupa l'espai de cada jugada dins la seva acció, es crea un identificador per acció per poder agrupar les jugades que la formen. Un cop agrupat l'espai per acció, es realitza la seva mitjana sobre aquesta agrupació.

Amb l'objectiu de cercar alguna causalitat de l'espai, es creen variables que el poden explicar i s'estudia la seva inclusió al model com a variables explicatives. Les noves variables són:

- **N_Jugadors:** Representa la mitjana del nombre de jugadors que actuen en una acció. La funció de la variable és estudiar si com més jugadors hi ha a la zona ofensiva, més perill es crearà.
- **Pilota_Pos:** Indica per quina alçada del camp es mou la pilota (coordenada y de la posició de la pilota). Aquesta variable ens pot indicar si el joc per alguna banda o pel centre té alguna influència sobre l'espai generat durant el partit. El domini de la variable és de 0 a 68, essent 68 la part superior del camp i 0 la part inferior del camp.
- **Temps_Accio:** El temps en segons que dura l'acció. L'objectiu d'incloure aquesta variable és veure si la duració d'una acció es reflecteix en l'espai, si existeix relació entre espai i temps.

A banda d'aquestes variables, també s'inclouen com a variables explicatives les que expliquen el tipus d'acció que s'ha dut a terme en cada jugada i el nombre de vegades que ha succeït una determinada acció. L'objectiu d'aquestes variables és poder traçar la seqüència de cada acció.

Malgrat que amb la metodologia presentada anteriorment ja es poden extreure resultats consistents, els quals donen lloc a conclusions sobre el partit. Pot ser d'interès modelitzar l'espai amb la finalitat de buscar alguna explicació més concreta sobre l'espai generat.

A l'hora de modelitzar es poden aplicar diferents mètodes, tot i això en aquest cas com que l'objectiu principal és el d'intentar explicar la variable resposta i no el de realitzar prediccions sobre ella, s'utilitzen dos mètodes per modelitzar. S'utilitza el clàssic model lineal (ML) i

l'algorisme *Gradient Tree Boosting (GBDT)*, el qual pertany a la família dels mètodes *ensemble*.

A partir del ML s'estudia el grau d'ajustament que existeix entre la variable resposta i les variables explicatives, mitjançant el contrast individual es comprova el nivell de significació dels paràmetres associats a les variables.

Per altra banda, mitjançant el *GBDT* es visualitza la importància de les diferents variables explicatives amb el mateix objectiu, el d'explicar l'espai de cada equip.

En els dos casos es realitzen dos models, un per cada equip. Cada observació representa l'espai obtingut durant cada acció.

Es realitza un estudi d'inclusió de les variables en el model, un cop finalitzat es decideix que les que actuen com a explicatives són:

- **N_Jugadors:** La mitjana de jugadors diferents que actuen durant l'acció.
- **Pilota_pos:** Posició mitjana de la pilota en referència a l'eix Y. S'escala aquesta variable calculant-la com el valor absolut de la diferència l'altura mitjana i 34 (punt central de l'eix Y). D'aquesta forma representa si la pilota durant l'acció s'ha mogut per la banda o pel centre
- **Temps_Accio:** Temps total en segons que dura l'acció
- **Passades:** Nombre de passes que es realitzen durant l'acció
- **Recuperacio:** Variable binària representada en 1 si l'acció comença amb una recuperació, altrament 0.
- **Xut:** Variable binària representada en 1 si l'acció finalitza amb xut, altrament 0.
- **Falta rebuda:** Variable binària representada en 1 si l'acció és interrompuda per una falta, altrament 0.

En referència al model utilitzat per explicar l'espai generat per cada equip, els paràmetres definits per ajustar el model han estat els que es presenten en la Taula 22.

<i>N_estimators</i> (Nombre d'arbres a entrenar)	50
<i>Max_depth</i> (Profunditat màxima dels paràmetres individuals de regressió, limita el nombre de nodes en l'arbre)	4
<i>Min_samples_split</i> (Nombre mínim d'observacions per dividir un node)	5
<i>Learning_rate</i> (Redueix la contribució de cada arbre, amb l'objectiu d'evitar el sobre ajustament)	0.1
<i>Loss</i> (Funció de pèrdua a optimitzar)	'ls': least squares

TAULA 22: Paràmetres definits en el GBDT

Referent a l'espai generat per l'equip local, visualitzem tant els resultats obtinguts mitjançant el ML com els obtinguts amb el GBDT.

Primer de tot, es visualitzen els resultats obtinguts amb el ML en la Taula 23.

	R^2	Estadístic F	Prob (Est. F)	AIC
	Coefficient	Error Estàndard	t	P > t
Constant	303,71	77,44	3,92	< 0,01
Falta rebuda	-127,49	64,50	-2,00	0,05
Passades	-40,15	16,51	-2,43	0,02
Recuperacio	-9,12	41,33	-0,22	0,82
Xut	67,87	47,41	1,43	0,15
Pilota_pos	-0,15	2,23	-0,07	0,95
N_Jugadors	113,92	7,46	15,26	< 0,01
Temps_Accio	14,56	11,21	1,30	0,19

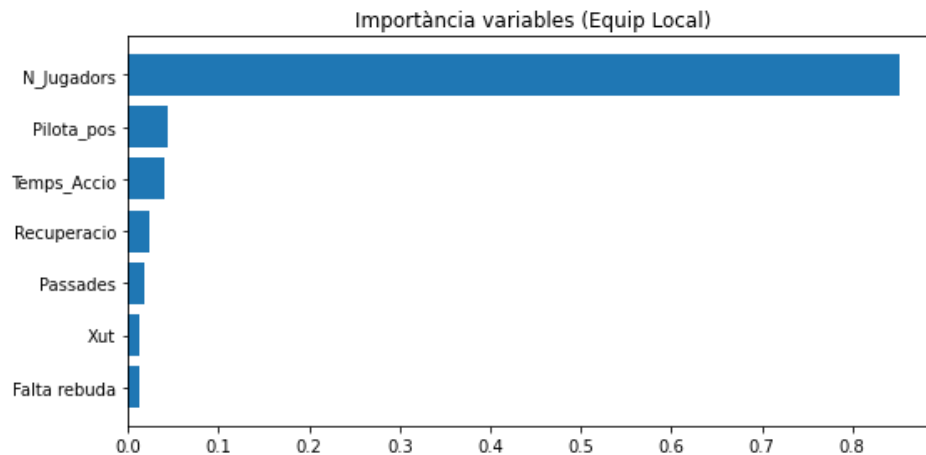
TAULA 23: Resultats del model ML referent a l'espai generat per l'equip local

A partir del ML ajustat, el primer que es veu és que el model ajustat és significatiu globalment, en altres paraules almenys una de les variables explicatives aporten informació per explicar l'espai generat. Continuant amb aquest punt veiem que obtenim un R^2 igual a 0,844, la qual cosa significa que la variable resposta està explicada un 84% per les variables explicatives incloses al model.

Per altra banda, ens fixem en la significació individual dels paràmetres que componen el model. Tres paràmetres associats a les variables incrustades en el model són significatius

amb un nivell de significació del 5%, concretament els paràmetres associats a les variables: *falta rebuda*, *passades* i *n_jugadors*.

Per complementar els resultats obtinguts amb el ML, visualitzem la importància de cada variable dins del model, segons el GBDT (veure Gràfic 6).



GRÀFIC 6: Importància de les variables en el GBDT (Equip local)

Modelitzat l'espai local generat en els noranta minuts, mitjançant els dos models s'observa clarament que la variable més important és la que representa el nombre de jugadors que prenen part en l'acció.

Per la resta de variables hi ha diferències del nivell d'importància entre el model ML i el GBDT. A partir del ML, a banda del variable nombre de jugadors, s'obté que el paràmetre associat a la variable *passades* i a la variable *falta rebuda* són significatius al nivell de significació establert (5%). El coeficient associat al paràmetre de la variable *falta rebuda* dona a entendre que si en una acció amb molt d'espai l'equip defensor comet falta, això li permet reduir aquest espai, ja que al aturar el joc es pot organitzar correctament.

Per altra banda, en el GBDT es veu com la variable referent a la posició de la pilota també té pes, aconseguint més espai mitjançant el joc per banda respecte al joc central. La duració total de l'acció també pren part en la importància de l'espai, com més llarga (elaborada) la jugada més espai podem esperar a generar.

Si ens fixem en les accions, sembla que després d'una recuperació podem esperar més espai la qual cosa és degut al fet que quan es recupera la pilota molts cops l'equip rival es troba desubicat. El nombre de passades i si la jugada ha finalitzat o no en xut també prenen importància en l'espai generat per l'equip local. Que les jugades que finalitzin en xut tenen importància dins de l'espai generat es veu a partir de què el coeficient associat a la variable *xut* és positiu, en altres paraules si la jugada ha finalitzat en xut, l'espai generat serà major.

Estudiat el comportament de l'espai local es realitza el mateix procediment però per l'espai generat per l'espai visitant.

El primer punt, altre cop consisteix en l'ajustament del ML, el qual es mostra en la Taula 24.

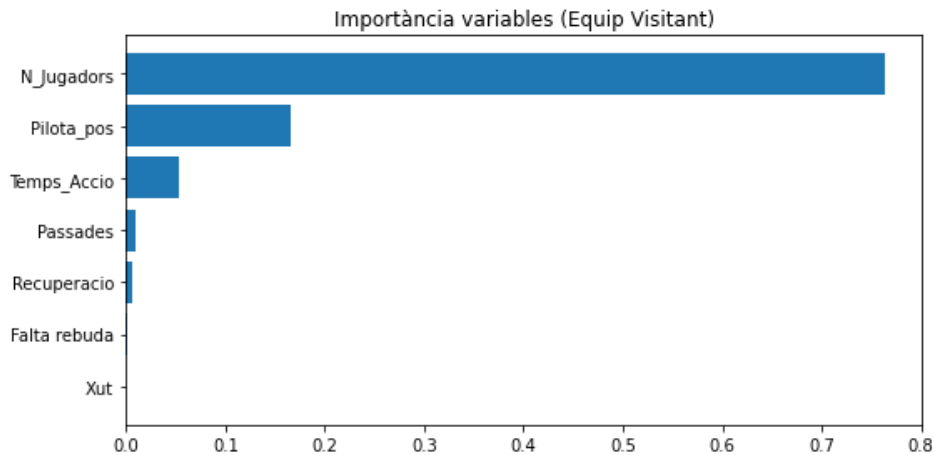
	R^2	Estadístic F	Prob (Est. F)	AIC
	0,68	18,24	9,36e-13	922,30
	Coefficient	Error Estàndard	t	P > t
Constant	389,86	94	4,14	< 0,01
Falta rebuda	-226,96	122,53	-1,85	0,07
Passades	-20,29	22,55	-0,90	0,37
Recuperacio	-35,1	50,60	-0,69	0,49
Xut	35,67	112,26	0,31	0,75
Pilota_pos	-1,75	2,86	-0,60	0,55
N_Jugadors	109,87	12,43	8,83	< 0,01
Temps_Accio	-7,25	14,98	-0,48	0,63

TAULA 24: Resultats del model ML referent a l'espai generat per l'equip visitant

Aquest cop obtenim uns resultats que indiquen menys associació entre la variable resposta i les variables explicatives. El coeficient de determinació ha passat de ser 0,84 en el cas de l'equip local a 0,68 en l'equip visitant.

L'altre punt important a comentar és que en aquest sols el paràmetre associat a la variable *N_Jugadors* és significatiu, a banda de la constant del model. El paràmetre associat a la variable *falta rebuda* es troba vora el llindar del nivell de significació.

Continuant amb l'anàlisi de l'espai generat per l'equip visitant, es realitza la mateixa modelització, però aquest cop utilitzant el GBDT i visualitzem en el gràfic 7 la importància de cada variable dins del model:



GRÀFIC 7: Importància de les variables en el GBDT (Equip visitant)

Veiem altre cop que la variable *N_Jugadors* és la que té més importància dins del model, la qual cosa té molt sentit, ja que com més jugadors en camp ofensiu l'espai serà més gran.

En aquest cas també apareixen diferències en el nivell d'importància entre el ML i el GBDT. Mentre que en el ML el paràmetre associat a la variable *falta rebuda* era el segon més significatiu, mitjançant el GBDT s'obté un nivell d'importància gairebé nul·la.

Tornant al GBDT, en referència a la variable que representa la posició de la pilota en l'eix vertical, altre cop en el GBDT és la segona variable amb més importància dins del model, a més en aquest segon model aquesta importància és més del doble obtingut en el primer model, el de l'espai local.

La duració de l'acció té la mateixa importància que en l'anterior model.

Respecte a les variables que expliquen el transcurs de l'acció veiem que en aquest model sols la variable que referencia a les passades realitzades durant l'acció té pes. Que la variable recuperació no tingui gairebé cap pes sembla que és símptoma de la bona col·locació de l'equip local, ja que quan han perdut una pilota han estat ben posicionats i en conseqüència l'equip visitant no ha pogut aprofitar la recuperació de forma efectiva. La variable del nombre de xuts, a diferència del model GBDT de l'espai local, ara té un pes nul a causa del poc perill de cara la porteria que ha realitzat l'equip visitant.

Capítol 8: DISCUSSIÓ DELS RESULTATS

En aquest apartat farem una discussió dels resultats obtinguts pel mètode proposat i també farem notar algunes limitacions que s'han observat en aquest treball.

Primer de tot, a través dels resultats obtinguts en l'apartat descriptiu de les dades, s'observa el punt més important de tot partit de futbol, el resultat, el qual ha estat de tres gols a zero, favorable per l'equip local. A més a més, en estudiar la freqüència de les accions separades per la primera i segona meitat s'ha vist com la primera part ha estat el període quan l'equip local ha realitzat més xuts.

A l'executar l'algorisme al partit, s'observen diferents resultats interessants a comentar. En calcular l'espai total de cada equip, es veu com l'equip local és el que genera més espai durant el partit, aquesta diferència recau principalment en la primera part en què en termes d'espai, l'equip local ha estat molt superior, concretament la diferència es troba entre els minuts 35 i minuts 60. A més es veu com la reacció de l'equip visitant arriba massa tard, ja a les acaballes del partit.

A través de l'espai que disposa cada porter durant el partit es veu com el porter local és el que disposa de més espai en la seva àrea, la seva mitjana durant el partit és lleugerament superior a 100 m² respecte a l'àrea que disposa el porter visitant, aquest resultat indica que la defensa de l'equip local ha estat més allunyada de la seva porteria, per tant el perill s'ha focalitzat més en l'àrea visitant que en l'àrea local.

Un punt molt interessant del mètode proposat és què es pot descomposar l'espai entre els jugadors d'un equip. En referència a l'equip local, tant en la primera part com en la segona part el Player8 ha estat el jugador que ha disposat de més espai. En l'equip visitant s'ha vist com en la primera part el Player18 i Player22 han estat els que més espai han disposat, a més s'ha vist com el seu altre davanter, Player24 ha disposat de molt poc espai en aquesta primera part, en la segona part l'equip visitant realitza canvis que es noten en l'espai generat, concretament en el Player23 i Player17 els quals guanyen protagonisme essent els dos que generen més espai de l'equip.

Seguint amb la idea de descomposar l'espai entre els jugadors, es divideix l'espai per la zona vertical, diferenciant entre zona superior, zona central i zona inferior, aquesta descomposició permet detectar si alguna zona vertical ha estat la causant de l'espai.

A partir de la descomposició de l'espai per zones junt amb l'aplicació de mapes de calor del moviment de la pilota s'obtenen resultats del nivell d'utilització de l'espai que s'ha generat.

Els mapes de calor ens permet visualitzar per quines zones s'ha mogut més la pilota i detectar si s'ha mogut massa per alguna zona on l'espai era mínim, o en canvi la distribució de la pilota

ha estat semblant a la de l'espai. També, a través del grau de calor en l'àrea rival es pot visualitzar el grau de perill que s'ha generat.

L'equip local en la primera part crea molt espai en la zona inferior (banda dreta), a més l'aprofita molt correctament i crea molt de perill en la porteria rival. Per altra banda, en la segona part l'espai està més repartit en les dues zones, com a punt negatiu enfoca massa el joc en la part inferior (banda esquerra).

L'equip visitant en la primera part genera molt poc espai en la zona central i zona inferior (banda esquerra), a més no aprofita l'espai generat en la zona superior i el risc que genera en la porteria rival és gairebé nul. En la segona part genera més espai i més repartit, millora el rendiment en la zona central i zona inferior. També genera més risc a la porteria rival, però no acaba d'aprofitar l'espai de forma correcta, jugant massa en la zona inferior (banda dreta). En l'apartat de la modelització de l'espai, s'ha vist com la variable més rellevant és el nombre de jugadors que prenen part en l'acció. Seguint de la variable "Pilota_pos", on en l'equip visitant prenia més importància el fet de la posició vertical en què es jugava. A més a més, s'ha vist com l'equip que més espai ha generat ha obtingut com a rellevant la variable recuperació, en altres paraules l'equip visitant no s'ha ubicat de la forma més correcta per evitar la sortida post-pèrdua.

Un cop vist i comentat els resultats sorgeixen avantatges i limitacions de l'algorisme. Els dos principals avantatges són l'eficiència computacional de l'algorisme proposat i la quantitat de resultats que es poden obtenir (espai per equip, per porter, per jugadors, utilització de l'espai...) tal com s'ha vist en l'apartat de RESULTATS.

Per altra banda, s'han observat dues limitacions del mètode proposat en aquest treball. La primera és que no es té en consideració la velocitat dels jugadors (a l'estar anonimitzat el partit tampoc s'hauria pogut provar) ni el seu cansament, aquests dos factors influeixen a l'hora d'ocupar espais i seria interessant que aquestes dues variables entressin en l'algorisme. L'altre inconvenient que s'ha observat ha estat dins l'espai descompost pels jugadors, on no es pot quantificar l'espai que genera cada jugador de forma indirecta, en atraure defenses.

Capítol 9: CONCLUSIONS

Primer de tot i la més important, el mètode desenvolupat per al càlcul d'espai té sentit per tal d'avaluar la ubicació dels jugadors d'un equip quan estan atacant, a banda de quantificar el posicionament de l'equip, també és un indicador del perill generat.

El segon punt important a extreure del mètode aplicat és que serveix per avaluar el que estan fent tots els jugadors en tot moment independentment de si tenen o no la pilota. Seguint amb aquest punt, es poden identificar els jugadors ofensius que no estiguin realitzant una labor correcta i intentar revertir la situació, per altra banda també es pot identificar els jugadors rivals que estan fent una feina molt bona i evitar que ho segueixin fent.

A més a més, s'ha vist que comparar l'espai que disposa cada porter també és una mesura molt útil per tal d'identificar clarament quin equip està generant perill i identificar quin equip planta l'altura defensiva més alta.

Una altra conclusió que s'extreu amb l'estadístic definit en el treball és que la distribució de l'espai generat en les diferents zones del camp es complementa molt bé amb els mapes de calor, a partir d'aquesta combinació es poden extreure resultats més interessants en analitzar un partit i també conclusions sobre el nivell d'utilitat que un equip ha tret de l'espai generat.

També s'ha vist que la modelització de l'espai permet extreure conclusions sobre l'origen d'aquest i del seu generador, a més de la importància de cada variable.

El mètode exposat a banda de ser útil per mesurar l'espai i el perill generat, resulta ser un mètode amb diverses aplicacions, molt eficient computacionalment, la qual cosa permet mostrar resultats a temps real.

Un punt negatiu que s'ha observat en aquest treball és que el fet de no tenir una bossa de partits per tal d'analitzar l'algorisme creat, ha comportat no poder extreure conclusions sobre els factors influents a l'hora de generar l'espai.

Com a punt de futur, per tal d'ampliar el mètode descrit en el treball, una possible millora, tenint en compte les limitacions que s'han trobat en l'algorisme, seria la de ponderar l'espai a través de mètodes bayesians tenint en compte la velocitat del jugador, la seva distància recorreguda i el temps que porta jugant en el partit. D'aquesta forma s'aconseguirien nous indicadors.

BIBLIOGRAFIA

- Fernandez, J. i Bornn, L. (2018). Wide Open Spaces: A statistical technique for measuring space creation in professional soccer. In *MIT Sloan Sports Analytics Conference, 2018*.
- Rein, R., Raabe, D., Memmert, D. i Perl, J. (2016). Evaluation of changes in space control due to passing behavior in elite soccer using Voronoi-cells. *Advances in Intelligent Systems and Computing*, 179-183.
- Horton, M., Gudmundsson, J., Chawla, S. i Estephan, J. (2014). Classification of Passes in Football Matches using Spatiotemporal Data. *ACM Transactions on Spatial Algorithms and Systems*, 3 (2), 1-6:30.
- Fonesca, S., Milho, J., Travassos, B. i Araújo, D. (2012). Spatial dynamics of team sports exposed by Voronoi diagrams. *Human Movement Science*, 31 (6), 1652-1659.
- Kise, K., Sato, A. i Iwata, M. (1997). Segmentation of Page Images Using the Area Voronoi Diagram. *Computer Vision and Image Understanding*, 70 (3), 370-382.
- Mahdavi, S., Mohades, A., Salehzadeh, A. Jahandieh, S. i Masoudi-Nejad, A. (2012). Computational analysis of RNA-protein interaction interface via the Voronoi diagram. *Computational Biology and Chemistry*, 47, 16-23.
- Nowak, A. (2015). Application of Voronoi diagrams in contemporary architecture and town planning. *Challenges of Modern Technology*, 6 (2), 30-34.
- Stewart, C. i Van der Ree, R. (2010). A Voronoi diagram based population model for social species of wildlife. *Ecological Modelling*, 221 (12), 1554-1568.
- Bobach, T., Farin, G., Hansford, D. i Umlauf, G. (2008). Natural neighbor extrapolation using ghost points. *Computer-Aided Design*, 41, 350-365.
- [Pàgina web de l'empresa Metrica-Sports]. <https://metrica-sports.com/>
- Sanchez, A. (s.d.). *¿Cuales son las medidas de un campo de futbol?*. Recuperat el 7 de Març de 2021, de Entrenador Futbol: <https://entrenadorfutbol.es/medidas-campo-futbol/>
- Diagramas de Voronoi* (s.d.). Recuperat el 17 de Febrer de 2021, de Geometría Computacional: <http://asignatura.us.es/fgcitig/contenidos/gctem3ma.htm>
- Polígonos de Thiessen* (2021). Recuperat el 27 de Febrer de 2021, de Wikipedia: https://es.wikipedia.org/wiki/Pol%C3%ADgonos_de_Thiessen
- Expósito, C. (s.d.). *Los Diagramas de Voronoi: La forma matematica de dividir el mundo*. Recuperat de <https://mediateca.educa.madrid.org/>

Brownlee, J. (2016). *A Gentle Introduction to the Gradient Boosting Algorithm for Machine Learning*. Recuperat el 8 de Maig de 2021, de Machine Learning Master: <https://machinelearningmastery.com/gentle-introduction-gradient-boosting-algorithm-machine-learning/>

Linear regression (2021). Recuperat el 8 de Maig de 2021, de Wikipedia: https://en.wikipedia.org/wiki/Linear_regression

Least squares (2021). Recuperat el 8 de Maig de 2021, de Wikipedia: https://en.wikipedia.org/wiki/Least_squares

ANNEX

Annex 1

La part de programació en aquest treball ha estat important (per consultar tot el codi https://github.com/xamapl99/TFG_Codi, demanar permís a xaviermartinezpla@gmail.com per tenir-hi accés) . En aquest apartat es mostren les dues funcions principals (amb les seves funcions auxiliars), la funció que retalla el Voronoi i la funció que calcula l'espai que disposa un equip tenint en compte la constant per la qual volem tallar la zona ofensiva (en el nostre cas ha estat el darrer terç).

Retallar el Voronoi a les mides del camp

```
def vertexs_fora_domini(vertexs, domini):
    #vertexs és una llista de punts i domini és una llista de dos subllistes de dos punts[[minx, miny], [maxx, maxy]]
    #Dóna els índexs dels vèrtexs que cauen fora del domini
    index_llista=[]
    for i in range(len(vertexs)):
        if domini[0][0] <= vertexs[i][0] <= domini[1][0] and domini[0][1] <= vertexs[i][1] <= domini[1][1]:
            pass
        else:
            index_llista.append(i)
    return index_llista

def tot_Dins(vertexs, l_index, domini):
    #l_index és una llista d'índexos de vèrtexs.
    #Mirarà si tots els vèrtexs que tenen l'índex a l_index estan dins del domini o no
    l_v_f= vertexs_fora_domini(vertexs, domini)
    for i in l_index:
        if i in l_v_f:
            return False
    return True

def algun_vertex_fora(l_vertex, l_vertex_fora):
    #això treballa amb índexos
    for v in l_vertex:
        if v in l_vertex_fora:
            return True
    return False

def direccioMediatriu(p1, p2):
    #Vector que dona la direcció de la mediatriu entre els dos punts.
    x=p2[0]-p1[0]
    y=p2[1]-p1[1]
    modul= math.sqrt(x**2+y**2)
    return [y/modul, -x/modul]

def distancia(p1, p2):
    return ((p1[0]-p2[0])**2 + (p1[1]-p2[1])**2)**0.5
```

```

def talla_domini(p1, p2, domini):
    #El primer punt estarà fora del domini i el segon dins.
    #Passem punts i no índexs. domini ens ve donat com una llista de dos subllistes de dos floats
    #Retorna el punt intersecció i el costat en el qual talla.
    if p1[0] < domini[0][0]:
        lamb= (domini[0][0] - p2[0])/(p1[0]-p2[0])
        py = lamb*p1[1] + (1-lamb)*p2[1]
        if domini[0][1] <=py <= domini[1][1]:
            px = domini[0][0]
            return [px, py], 1
        else:
            if p1[1] < domini[0][1]:
                lamb = (domini[0][1] - p2[1])/(p1[1]-p2[1])
                px = lamb*p1[0] + (1-lamb)*p2[0]
                return [px, domini[0][1]], 0
            else:
                lamb = (domini[1][1] - p2[1])/(p1[1]-p2[1])
                px = lamb*p1[0] + (1-lamb)*p2[0]
                return [px, domini[1][1]], 2
    elif p1[0] > domini[1][0]:
        lamb= (domini[1][0] - p2[0])/(p1[0]-p2[0])
        py = lamb*p1[1] + (1-lamb)*p2[1]
        if domini[0][1] <=py <= domini[1][1]:
            px = domini[1][0]
            return [px, py], 3
        else:
            if p1[1] < domini[0][1]:
                lamb = (domini[0][1] - p2[1])/(p1[1]-p2[1])
                px = lamb*p1[0] + (1-lamb)*p2[0]
                return [px, domini[0][1]], 0
            else:
                lamb = (domini[1][1] - p2[1])/(p1[1]-p2[1])
                px = lamb*p1[0] + (1-lamb)*p2[0]
                return [px, domini[1][1]], 2
    elif p1[1] < domini[0][1]:
        lamb= (domini[0][1]-p2[1])/(p1[1] -p2[1])
        px = lamb*p1[0] + (1-lamb)*p2[0]
        return [px, domini[0][1]], 0
    else:
        lamb =(domini[1][1]-p2[1])/(p1[1]-p2[1])
        px = lamb*p1[0] + (1-lamb)*p2[0]
        return [px, domini[1][1]], 2

def talla_domini_vector(p, v, domini):
    #p és interior al domini i v és un vector de mòdul 1. Busquem el punt intersecció p+lambdav amb el domini
    dist= max(domini[1][0]-domini[0][0],domini[1][1]-domini[0][1] )
    p_nou = [p[0] + v[0]*dist*2, p[1]+v[1]*dist*2] #aquest punt està fora, per tant cridarem la funció anterior
    return talla_domini(p_nou, p, domini)

def intersecta_segments(p1, p2, q1, q2):
    #el segment p1p2 talla el segment q1q2 i volem saber quin és el punt intersecció
    b=np.array([q2[0]-p2[0], q2[1]-p2[1]])
    a=np.array([[p1[0]-p2[0], q2[0]-q1[0]], [p1[1]-p2[1], q2[1]-q1[1]] ])
    beta=np.linalg.solve(a, b)[1]
    alfa = np.linalg.solve(a, b)[0]
    return alfa, beta
    #[q1[0]*beta+q2[0]*(1-beta),q1[1]*beta+q2[1]*(1-beta) ]

```

```

def p_vertex_domini(index1, index2):
    #index1, index2 són dos valors diferents entre 0, 1, 2, 3
    if (index1==0 and index2==1) or (index2==0 and index1==1):
        return 0
    elif (index1==1 and index2==2) or (index2==1 and index1==2):
        return 1
    elif (index1==2 and index2==3) or (index2==2 and index1==3):
        return 2
    elif (index1==0 and index2==3) or (index2==0 and index1==3):
        return 3
    else:
        return -1

```

#També pot passar que haguem d'afegir més d'un punt. Els punts a afegir no depenen només de l'índex
def vertexs_domini_2(i1, i2, domini, punt, totspunts, nvert):

```

    if (i1==0 and i2==2) or (i1==2 and i2==0):
        #serà 0 i 1 o 2 i 3
        d1= distancia(domini[0], punt)
        d2= distancia(domini[0], totspunts[0])
        for x in totspunts:
            d2= min(d2, distancia(domini[0], x))
        if d1 < d2:
            if i1==0 and i2==2:
                return [nvert, nvert+1]
            else:
                return [nvert+1, nvert]
        else:
            if i1==0 and i2==2:
                return [nvert+3, nvert+2]
            else:
                return [nvert+2, nvert+3]
    else:
        #serà 0,3 o 1, 2
        d1= distancia(domini[0], punt)
        d2= distancia(domini[0], totspunts[0])
        for x in totspunts:
            d2= min(d2, distancia(domini[0], x))
        if d1 < d2:
            if i1==1 and i2==3:
                return [nvert, nvert+3]
            else:
                return [nvert+3, nvert]
        else:
            if i1==3 and i2==1:
                return [nvert+3, nvert+1]
            else:
                return [nvert+1, nvert+3]

```

#Una aresta pot tallar dos costats del domini

```

def inicialitza_infor_arestes(DV, domini):
    vertexs_fora= vertexs_fora_domini(DV.vertices, domini)
    linfor_arestes=[]
    for i in range(len(DV.ridge_vertices)):
        p1= DV.ridge_vertices[i][0]
        p2=DV.ridge_vertices[i][1]
        if -1 not in [p1, p2] and p1 not in vertexs_fora and p2 not in vertexs_fora:
            #Aresta inclosa en el domini. No haurem de fer-li res

```

```

linfor_arestes.append([1, True, 0, None, None, None, None])
elif -1 not in [p1, p2]:
    #Aresta finita però no inclosa en el domini. S'haurà de tallar
    linfor_arestes.append([0, False, None, None, None, None, None])
else:
    linfor_arestes.append([-1, False, None, None, None, None, None]) #Si el booleà és fals vol dir que no
s'ha tallat l'aresta infinita.
    #Si el booleà és cert, l'aresta ja ha estat tallada. Guardem l'índex en la llista de l_punts_frontera
    return linfor_arestes

```

#Funcio que talla una aresta finita tal que té un vèrtex dins i l'altre fora del domini

```

def tracta_aresta_finita(pfora, pdins, DV, domini, linfor, l_punts_frontera):
    #pfora i pdins són índexos de vèrtex. El primer està fora del domini i el segon dins
    #Primer buscarem [pfora, pdins] dins de DV.ridge_vertices
    for i in range(len(DV.ridge_vertices)):
        x= DV.ridge_vertices[i]
        if pfora in x and pdins in x:
            #Vol dir que x fa referència a l'aresta que volem tallar
            if linfor[i][1]:
                #En aquest cas vol dir que aquesta aresta ja ha estat tallada, per tant no hem de tornar-ho a fer
                return linfor[i][3], linfor[i][4]
            else:
                p1, n1= talla_domini(DV.vertices[pfora], DV.vertices[pdins], domini)
                linfor[i][1]=True
                l_punts_frontera.append(p1)
                linfor[i][2] = 1
                linfor[i][3] = len(l_punts_frontera) -1 #index en el que hem posat el punt
                linfor[i][4] = n1
                return linfor[i][3], linfor[i][4]

```

def tracta_aresta_finita_2_fora(p1, p2, DV, domini, linfor, l_punts_frontera):

```

    #Ara p1 i p2 ens donen vèrtexs que estan fora del domini els dos
    #Primer buscarem l'aresta dins de DV.ridge_vertices per si ja ha estat tractada
    for i in range(len(DV.ridge_vertices)):
        x= DV.ridge_vertices[i]
        if p1 in x and p2 in x: #hem trobat l'aresta
            if linfor[i][1]:
                return linfor[i][2:]
            else:
                punt1=DV.points[DV.ridge_points[i][0]]
                punt2=DV.points[DV.ridge_points[i][1]]
                p_v1=DV.vertices[p1]
                p_v2=DV.vertices[p2]
                #Aquests són els dos punts donats que separa l'aresta que ens ocupa
                alfa, beta=intersecta_segments(p_v1, p_v2, punt1, punt2) #aquest punt està dins
                if 0<= alfa<= 1 and 0<= beta<= 1:
                    p_aux = [punt1[0]*beta+punt2[0]*(1-beta),punt1[1]*beta+punt2[1]*(1-beta) ]

                p_front1, n1= talla_domini(p_v1, p_aux, domini)
                p_front2, n2= talla_domini(p_v2, p_aux, domini)
                l_punts_frontera.append(p_front1)
                l_punts_frontera.append(p_front2)
                linfor[i][1] = True
                linfor[i][2]=2
                linfor[i][3] = p_front1
                linfor[i][4]= n1
                linfor[i][5] = p_front2
                linfor[i][6]= n2

```

```

        return linfor[i][2:]
    else:
        linfor[i][1] = True
        linfor[i][2]=0
        return linfor[i][2:]

def tracta_aresta_infinita_extrem_dins(ind_aresta, DV, domini, linfor, l_punts_frontera):

    #Les arestes infinites no venen univocament determinades per l'extrem, passarem l'índex que la determinen
    if linfor[ind_aresta][1]:
        #l'aresta ja havia estat tractada, potser en el cas de tractar una altra tessel·la
        return linfor[ind_aresta][3:]
    else:
        #l'aresta que considerem és la que està a la posició i. Mirem de quins punts és la mediatriu
        punt1 = DV.points[DV.ridge_points[ind_aresta][0]]
        punt2 = DV.points[DV.ridge_points[ind_aresta][1]]

        #l'aresta que considerem és la mediatriu entre punt1 i punt2. Agafem el vector normalitzat.
        vector= direccioMediatriu(punt1, punt2)

        if DV.ridge_vertices[ind_aresta][0] == -1:
            extrem = DV.vertices[DV.ridge_vertices[ind_aresta][1]]
            id_extrem = DV.ridge_vertices[ind_aresta][1]
        else:
            extrem = DV.vertices[DV.ridge_vertices[ind_aresta][0]]
            id_extrem = DV.ridge_vertices[ind_aresta][0]

        for j in range(len(DV.ridge_vertices)):
            if id_extrem in DV.ridge_vertices[j] and ind_aresta != j:
                if DV.ridge_points[j][0] not in DV.ridge_points[ind_aresta]:
                    punt3 =DV.points[DV.ridge_points[j][0]]
                else:
                    punt3 = DV.points[DV.ridge_points[j][1]]
                break
        p_aux=[extrem[0]+ vector[0]*0.1, extrem[1]+ vector[1]*0.1]

        if distancia(p_aux, punt3) < distancia(p_aux, punt1):
            v=[-vector[0], -vector[1]]
            p, ind=talla_domini_vector(DV.vertices[id_extrem], v, domini)
        else:
            p, ind=talla_domini_vector(DV.vertices[id_extrem], vector, domini)

        l_punts_frontera.append(p)
        linfor[ind_aresta][1]=True
        linfor[ind_aresta][2] =1
        linfor[ind_aresta][3] = len(l_punts_frontera)-1
        linfor[ind_aresta][4] = ind
        return linfor[ind_aresta][3], linfor[ind_aresta][4]

def punt_fantasma(punt, x, domini):
    px= punt[0]
    py = punt[1]
    if x == 0:
        return [px, py - 2*(py-domini[0][1])]
    elif x == 2:
        return [px, py + 2*(domini[1][1]-py)]
    elif x == 1:

```

```

    return [px - 2*(px-domini[0][0]), py]
elif x == 3:
    return [px+ 2*(domini[1][0]-px), py]

def falta(x, p, domini, points):
    p_aux=punt_fantasma(p, x, domini)
    d= distancia(p_aux, p)
    for punt in points:
        if distancia(p_aux, punt) < d:
            return False
    return True

def punts_fantasma_Voronoi(punts, domini):

    # domini és el rectangle pel qual retallarem

    DV=scipy.spatial.Voronoi(punts)

    vertextsVoronoi = DV.vertices
    nVertexs= len(DV.vertices)

    #Calculem els vèrtexs que estan fora del domini
    v_fora= vertexts_fora_domini(DV.vertices, domini)
    v_foraAM= v_fora+[-1]

    punts_fantasma=[] #Guardarem les tessell·les un cop han estat tallades, si convé

    l_p_frontera=[[domini[0][0],domini[0][1]], [domini[0][0],domini[1][1]], [domini[1][0],domini[1][1]], [domini[1][0],
domini[0][1]]]
    #Guardarem els punts en la frontera del quadrat degut a retallar pel domini. Inicialitzem amb els vèrtexs del
domini
    #Per no calcular dues vegades el mateix punt guardarem informació de les arestes que s'han de tallar

    linfor=inicialitza_infor_arestes(DV, domini) #llista on guardarem informació sobre les arestes de Voronoi

    for i in range(len(DV.points)):

        l_regio=DV.regions[DV.point_region[i]] #tessel·la associada al punt i. Hi ha els índexs dels vèrtexs

        if -1 not in l_regio and tot_Dins(DV.vertices, l_regio, domini):
            pass
        else:
            costats =set()
            for k in range(len(l_regio)):
                seg= (k+1)%len(l_regio)
                if l_regio[k] not in v_foraAM and l_regio[seg] not in v_foraAM:
                    #aresta finita inclosa en el domini. No cal fer res
                    pass
                elif l_regio[k] in v_fora and l_regio[seg] not in v_foraAM:
                    idpunt, cost=tracta_aresta_finita(l_regio[k], l_regio[seg], DV, domini, linfor, l_p_frontera)
                    costats.add(cost)

                elif l_regio[seg] in v_fora and l_regio[k] not in v_foraAM:
                    idpunt, cost=tracta_aresta_finita(l_regio[seg], l_regio[k], DV, domini, linfor, l_p_frontera)
                    costats.add(cost)

                elif l_regio[k] in v_fora and l_regio[seg] in v_fora:
                    resultat=tracta_aresta_finita_2_fora(l_regio[k], l_regio[seg], DV, domini, linfor, l_p_frontera)

```

```

if resultat[0]==0:
    pass
elif resultat[0]==2:
    costats.add(resultat[2])
    costats.add(resultat[4])

elif (l_regio[k]==-1 and l_regio[seg] not in v_foraAM):
    l_identi=[]
    for j in range(len(DV.ridge_vertices)):
        if l_regio[k] in DV.ridge_vertices[j] and l_regio[seg] in DV.ridge_vertices[j]:
            if i in DV.ridge_points[j]:
                l_identi.append(j)
    for x in l_identi:
        resul=tracta_aresta_infinita_extrem_dins(x, DV, domini, linfor, l_p_frontera)
        costats.add(resul[1])

elif (l_regio[seg]==-1 and l_regio[k] not in v_foraAM):
    l_identi=[]
    for j in range(len(DV.ridge_vertices)):
        if l_regio[k] in DV.ridge_vertices[j] and l_regio[seg] in DV.ridge_vertices[j]:
            if i in DV.ridge_points[j]:
                l_identi.append(j)
    for x in l_identi:
        resul=tracta_aresta_infinita_extrem_dins(x, DV, domini, linfor, l_p_frontera)
        #l_regio_retallada.append(l_regio[k])
        #l_regio_retallada.append(resul)
        costats.add(resul[1])

for x in costats:
    p_fantasma= punt_fantasma(DV.points[i], x, domini)
    punts_fantasma.append(p_fantasma)
if len(costats)== 2:
    if costats==set({0, 1}) or costats== set({1, 2}) or costats==set({2, 3}) or costats==set({3, 0}):
        pass
    elif costats==set({0, 2}):
        if falta(1, DV.points[i], domini, DV.points):
            p_fantasma= punt_fantasma(DV.points[i], 1, domini)
            punts_fantasma.append(p_fantasma)
            costats.add(1)
        elif falta(3, DV.points[i], domini, DV.points):
            p_fantasma= punt_fantasma(DV.points[i], 3, domini)
            punts_fantasma.append(p_fantasma)
            costats.add(3)
    elif costats==set({1, 3}):
        if falta(0, DV.points[i], domini, DV.points):
            p_fantasma= punt_fantasma(DV.points[i], 0, domini)
            punts_fantasma.append(p_fantasma)
            costats.add(0)

        elif falta(2, DV.points[i], domini, DV.points):
            p_fantasma= punt_fantasma(DV.points[i], 2, domini)
            punts_fantasma.append(p_fantasma)
            costats.add(2)
#print ('costats que talla', i, costats, DV.points[i])
return punts_fantasma

```

Càlculer l'àrea

def tallar_pt_espai(p1,p2,constant):##constant és el punt x per on volem tallar en el nostre cas sols volem el darrer terç ofensiu

```
    lbmd=(constant-p2[0])/(p1[0]-p2[0])
    pt_y=lbmd*p1[1]+(1-lbmd)*p2[1]
    pt_nou=np.array([constant,pt_y])
    return(pt_nou)
```

def area_Triangle(p1, p2, p3):

```
    v1x= p3[0] -p1[0]
    v1y= p3[1] -p1[1]
    v2x= p2[0] -p1[0]
    v2y= p2[1] -p1[1]
    return abs((v1x*v2y-v1y*v2x)/2)
```

def area_poligon_porter(poligon, centre):

```
    area =0
    for i in range(len(poligon)-1):
        area = area + area_Triangle(poligon[i], poligon[i+1], centre)
    area = area + area_Triangle(poligon[-1], poligon[0], centre)
    return area
```

def nova_area_poligon (poligon):

```
    suma_total=0
    for i in range (1,len(poligon)-1):
        suma_total=suma_total+area_Triangle(poligon[0],poligon[i],poligon[i+1])

    return(suma_total)
```

def retocar_punts (punts):

```
    for i in range(0,len(punts)):
        if punts[i][0]>105:
            punts[i][0]=104.99
        if punts[i][0]<0:
            punts[i][0]=0.01
        if punts[i][1]>68:
            punts[i][1]=67.99
        if punts[i][1]<0:
            punts[i][1]=0.01
```

def poligon_nou_retallat_part_dreta (pts_poligon,constant):

```
    situacio_pt=[]
    for i in range(len(pts_poligon)):
        if pts_poligon[i][0]<constant:
            situacio_pt.append(False)
        else:
            situacio_pt.append(True)
    nou_poligon=[]
    for i in range(len(situacio_pt)):
        k=(i+1)%len(situacio_pt)
        if situacio_pt[i]==True and situacio_pt[k]==True:
            nou_poligon.append(pts_poligon[i])
        elif situacio_pt[i]==True and situacio_pt[k]==False:
            nou_poligon.append(pts_poligon[i])
            nov_pt=tallar_pt_espai(pts_poligon[i],pts_poligon[k],constant)
            nou_poligon.append(nov_pt)
```



```

elif situacio_pt[i]==False and situacio_pt[k]==True:
    nov_pt=tallar_pt_espai(pts_poligon[i],pts_poligon[k],constant)
    nou_poligon.append(nov_pt)
return(nou_poligon)

```

```
def poligon_nou_retallat_part_esquerra (pts_poligon,constant):
```

```

situacio_pt=[]
for i in range(len(pts_poligon)):
    if pts_poligon[i][0]>constant:
        situacio_pt.append(False)
    else:
        situacio_pt.append(True)
nou_poligon=[]
for i in range(len(situacio_pt)):
    k=(i+1)%len(situacio_pt)
    if situacio_pt[i]==True and situacio_pt[k]==True:
        nou_poligon.append(pts_poligon[i])
    elif situacio_pt[i]==True and situacio_pt[k]==False:
        nou_poligon.append(pts_poligon[i])
        nov_pt=tallar_pt_espai(pts_poligon[i],pts_poligon[k],constant)
        nou_poligon.append(nov_pt)
    elif situacio_pt[i]==False and situacio_pt[k]==True:
        nov_pt=tallar_pt_espai(pts_poligon[i],pts_poligon[k],constant)
        nou_poligon.append(nov_pt)
return(nou_poligon)

```

```
def area_total_dreta(points, domini,constant): ##Calcular l'area per l'equip que marca a la dreta
```

```

lf=punts_fantasma_Voronoi(points, domini)
l=np.concatenate((points, np.array(lf)))
DV2= scipy.spatial.Voronoi(l)
suma = []
for i in range(len(points)):
    m=[]
    poligon=DV2.vertices[(DV2.regions[DV2.point_region[i]])]
    for j in range(len(poligon)):
        if poligon[j][0]<constant:
            m.append(True)
    if len(m)==len(poligon):##tots fora
        suma.append(0)
    elif len(m)==0:      ##tots dins
        res_area_tot=nova_area_poligon(poligon)
        suma.append(res_area_tot)
    elif len(m)>0 and len(m)<len(poligon):##alguns fora,alguns dins
        poligon_v2=poligon_nou_retallat_part_dreta(poligon,constant)
        res_area_v2=nova_area_poligon(poligon_v2)
        suma.append(res_area_v2)
return(suma)

```

```
def area_total_esquerra(points, domini,constant): ##Calcular l'area per l'equip que marca a la esquerra
```

```

lf=punts_fantasma_Voronoi(points, domini)
l=np.concatenate((points, np.array(lf)))
DV2= scipy.spatial.Voronoi(l)
suma = []
for i in range(len(points)):
    m=[]
    poligon=DV2.vertices[(DV2.regions[DV2.point_region[i]])]
    for j in range(len(poligon)):

```

```
if poligon[j][0]>constant:
    m.append(True)
if len(m)==len(poligon):##tots fora
    suma.append(0)
elif len(m)==0:    ##tots dins
    res_area_tot=nova_area_poligon(poligon)
    suma.append(res_area_tot)
elif len(m)>0 and len(m)<len(poligon):##alguns fora,alguns dins
    poligon_v2=poligon_nou_retallat_part_esquerra(poligon,constant)
    res_area_v2=nova_area_poligon(poligon_v2)
    suma.append(res_area_v2)
return(suma)
```

Annex 2

Tal i com s'ha vist en el subapartat Optimització (Utilització de l'Espai), l'estadístic generat en aquest treball descompost en les tres zones verticals (zona superior, central i inferior) es complementa molt bé amb el mapa de calor del moviment de la pilota.

Per crear el mapa de calor per un moment determinat, per tal d'analitzar si l'espai ha estat utilitzat o no, ens ajudem de la llibreria *seaborn* seguint el següent procés:

- 1) Emmagatzem les coordenades de la pilota, tant la coordenada X com la Y, en els moments en què volem dibuixar el mapa de calor i en les accions on l'equip que estem analitzant té la pilota. Cada coordenada es guarda en un vector diferent, la coordenada X en la variable "pos_pilota_X", mentre que la coordenada Y en la variable "pos_pilota_Y".
- 2) Amb la funció *kdeplot* que trobem en la llibreria *seaborn* li introduïm les variables "pos_pilota_X" i "pos_pilota_Y" per realitzar el mapa de calor corresponent. Per tal de visualitzar el mapa de calor dins del camp de futbol, es pinta el camp amb les seves mides corresponents i detalls.

El codi de Python per realitzar aquest procés és el següent:

```
fig=plt.figure()
fig.set_size_inches(7.5, 4)##Per fer-ho gran
ax = plt.gca()
ax.set_facecolor('chartreuse')
ax1=fig.add_subplot(1,1,1)
plt.xlim(-5,110)
plt.ylim(-5,73)
##Dimensions del camp 105x68m
plt.plot([0,0],[0,68], color="white")
plt.plot([0,105],[68,68], color="white")
plt.plot([105,105],[68,0], color="white")
plt.plot([105,0],[0,0], color="white")
plt.plot([52.5,52.5],[0,68], color="white")
##Àrea esquerra
#Gran
plt.plot([16.5,16.5],[54.16,13.84],color="white")
plt.plot([0,16.5],[54.16,54.16],color="white")
plt.plot([16.5,0],[13.84,13.84],color="white")
#Petita
plt.plot([0,5.5],[43.16,43.16],color="white")
```

```
plt.plot([5.5,5.5],[43.16,25],color="white")
plt.plot([5.5,0],[25,25],color="white")
##Àrea dreta
#Gran
plt.plot([105,88.5],[54.16,54.16],color="white")
plt.plot([88.5,88.5],[54.16,13.84],color="white")
plt.plot([88.5,105],[13.84,13.84],color="white")
#Petita
plt.plot([105,99.5],[43.16,43.16],color="white")
plt.plot([99.5,99.5],[43.16,25],color="white")
plt.plot([99.5,105],[25,25],color="white")
centreCircle = plt.Circle((52.5,34),9.15,color="white",fill=False)
centreSpot = plt.Circle((52.5,34),1,color="white")
leftArc = Arc((11,34),height=18.3,width=18.3,angle=0,theta1=310,theta2=50,color="white")
rightArc= Arc((94,34),height=18.3,width=18.3,angle=0,theta1=130,theta2=230,color="white")
leftPenSpot = plt.Circle((11,34),1,color="white")
rightPenSpot = plt.Circle((94,34),1,color="white")
#Dibuixar els cercles del camp
ax.add_patch(centreCircle)
ax.add_patch(centreSpot)
ax.add_patch(leftPenSpot)
ax.add_patch(rightPenSpot)
ax.add_patch(rightArc)
ax.add_patch(leftArc)
#Per fer el mapa de calor
sns.kdeplot(pos_pilota_X,pos_pilota_Y, shade=True,n_levels=50,color="red")
```

