

MÁSTER EN BUSINESS ANALYTICS

**“FutHub: solución tecnológica para el análisis
táctico y la mejora del rendimiento en el fútbol
semiprofesional”**

Presentado por:

GUILLERMO RICO SALANOVA

IÑIGO GANDARIAS GARCÍA

LUIS JAVIER MONCADA RODRÍGUEZ

Dirigido por:

VICTOR YESTE

VICENTE CASTILLO

CURSO ACADÉMICO 2024-2025

Resumen

El presente Trabajo de Fin de Máster aborda el análisis de datos aplicado al fútbol semiprofesional, un campo donde la utilización de herramientas tecnológicas especializadas aún es restringida. En contraste con el ambiente laboral, donde el análisis táctico basado en datos está totalmente incorporado, numerosos equipos semiprofesionales no cuentan con soluciones ajustadas a sus recursos y requerimientos, lo que provoca una brecha en la habilidad para incrementar el desempeño y tomar decisiones basadas en evidencia.

Con el objetivo de reducir esta brecha, nace FutHub, una solución tecnológica creada especialmente para el análisis táctico y la optimización del desempeño en clubes de fútbol semiprofesional. La plataforma facilita la recolección, tratamiento y presentación de información esencial vinculada con el juego y los entrenamientos. Su diseño se enfoca en proporcionar una interfaz fácil de usar y características útiles que simplifiquen su utilización por equipos técnicos con habilidades restringidas en análisis de datos o sin acceso a departamentos especializados.

Para la creación de FutHub, se han utilizado programas como Lovable y el lenguaje de programación Python, implementándose con una metodología fundamentada en el análisis de necesidades reales adquiridas a través de entrevistas con entrenadores y analistas, un diseño enfocado en el usuario y la validación mediante pruebas piloto con equipos reales pertenecientes al sector semiprofesional. Los hallazgos de estas pruebas señalan que la herramienta aporta de manera significativa a un entendimiento más profundo del juego, optimiza la elaboración de entrenamientos más particulares y facilita una valoración más imparcial del desempeño tanto en grupo como individual.

Palabras clave

Análisis de datos, fútbol semiprofesional, rendimiento deportivo, tecnología deportiva, táctica, visualización de datos, solución tecnológica, FutHub.

Abstract

This master's thesis addresses data analysis applied to semi-professional football, a field where the use of specialized technological tools is still limited. In contrast to the workplace, where data-driven tactical analysis is fully incorporated, many semi-professional teams lack solutions tailored to their resources and requirements, resulting in a gap in the ability to improve performance and make evidence-based decisions.

To reduce this gap, we suggest FutHub, a technological solution specifically designed for tactical analysis and performance optimization in semi-professional football clubs. The platform facilitates the collection, processing, and presentation of essential game-related and training data. Its design focuses on providing an easy-to-use interface and useful features that simplify its use by technical teams with limited data analysis skills or without access to specialized departments.

FutHub was created using programs such as Lovable and the Python programming language. It was implemented using a methodology based on the analysis of real needs acquired through interviews with coaches and analysts, a user-centered design, and validation through pilot testing with real teams from the semi-professional sector. The findings from these tests indicate that the tool significantly contributes to a deeper understanding of the game, optimizes the development of more targeted training sessions, and facilitates a more impartial assessment of both group and individual performance.

Keywords

Data analysis, semi-professional football, sports performance, sports technology, tactics, data visualization, technological solutions, FutHub.

<i>Resumen</i>	2
<i>Palabras clave</i>	2
<i>Abstract</i>	3
<i>Keywords</i>	3
1. Introducción	9
2. Justificación	10
2.1 Objetivos de Desarrollo Sostenible (ODS)	11
2.2 Problema y finalidad	12
3. Objetivos	13
3.1 Objetivo general	13
3.2 Objetivos específicos	13
4. Marco teórico	14
4.1 Introducción al Análisis de Datos	14
4.2 Análisis de Datos en el Deporte	16
4.2.2. Impacto del análisis de datos en el rendimiento y la estrategia	19
4.2.3. Diferencias entre el deporte profesional y semiprofesional	21
4.3 Fuentes y Tipos de Datos en el Deporte	23
4.3.1. Datos de rendimiento deportivo	23

4.3.2. Datos físicos y fisiológicos	24
4.3.3. Datos tácticos y posicionales	24
4.4 Tecnologías para la Captura y Procesamiento de Datos Deportivos	25
4.4.1. Sensores y wearables	25
4.4.2. Sistemas de tracking y geoposicionamiento.....	25
4.4.3. Herramientas de análisis y visualización de datos	26
4.5 Visión por Computador (Computer Vision) Aplicada al Deporte.....	26
4.5.1. Fundamentos de la visión por computador.....	26
4.5.2. Aplicaciones en la detección y seguimiento de jugadores	27
4.5.3. Análisis automático de eventos: goles, faltas, posesión, etc.	27
4.5.4. Limitaciones y desafíos actuales	28
4.6 Herramientas de Análisis de Eventos Deportivos.....	29
4.6.1. Plataformas comerciales	29
4.6.2. Herramientas open source y bibliotecas técnicas	30
4.6.3. Modelos y frameworks de IA utilizados en deportes	31
4.6.4. Integración de flujos de trabajo analíticos en clubes semiprofesionales	32
4.7 Retos y Oportunidades del Análisis de Datos en el Deporte Semiprofesional.....	33
4.7.1. Limitaciones técnicas, económicas y humanas	33
4.7.2. Potencial transformador de un enfoque analítico	33

4.7.3. Casos emergentes y tendencias en clubes semiprofesionales.....	34
4.8 Estudio de Caso: Aplicación Real del Análisis de Datos en el Fútbol	34
4.8.1. Contextualización del caso seleccionado (club, categoría, entorno)	34
4.8.2. Herramientas utilizadas y metodología aplicada.....	35
4.8.3. Resultados obtenidos y análisis del impacto	35
4.8.4. Lecciones aprendidas y aplicabilidad a otros entornos semiprofesionales.....	36
5. Metodología.....	38
5.1 Diagrama de Gantt.....	38
5.2 Extracción de datos: Computer Vision aplicada al fútbol	41
5.2.1 Selección del material audiovisual.....	41
5.2.2 Preprocesamiento del vídeo	42
5.2.3 Detección de jugadores y balón	43
5.2.4. Algoritmo de tracking y generación de trayectorias.....	47
5.2.5. Clusterización de jugadores	51
5.2.6. Detección del terreno de juego y proyección de jugadores y balón sobre él.....	52
5.2.7. Generación de data sets con la geolocalización de jugadores y balón.....	57
5.3 Procesamiento de datos: Cálculo de eventos futbolísticos.....	58
5.3.1. Diseño de reglas y lógica para detección de eventos	58
5.3.2 Asociación temporal y espacial de eventos.....	65

5.3.3 Creación del dataset de eventos	66
5.3.4 Validación interna y consistencia de los eventos generados:	67
5.4 Trabajo con cuerpo técnico	68
5.4.1 Primer contacto con el cuerpo técnico	68
5.4.2 Cálculo de estadísticas individuales de rendimiento físico y deportivo:	69
5.4.3 Calculo de la evaluación del comportamiento del equipo en situaciones tácticas específicas	70
5.5 Visualización de datos: Análisis táctico y de rendimiento	73
5.5.1. Librerías Python utilizadas para visualización	74
5.5.2. Creación del Dashboard con Lovable	76
6. Modelo de Negocio.....	87
6.1 Estudio de Mercado	88
6.2 validación y prueba del producto (MVP).....	90
6.3 Funcionamiento general.....	91
6.4 Modelo de ingresos.....	94
6.5 Infraestructura técnica	95
7. Conclusiones	98
8. Futuras líneas de investigación.....	100
9. Referencias.....	104

10.	<i>Índice de figuras</i>	109
	10.1 Índice de tablas	109
	10.2 Índice de ilustraciones	109
	10.3 Índice de ecuaciones	110
11.	<i>Utilización de herramientas de inteligencia artificial en este trabajo</i>	111
12.	<i>Anexos</i>	112

1. Introducción

El análisis de datos en el ámbito deportivo se está consolidando como un componente esencial para las instituciones que forman parte de este sector. La continua innovación tecnológica y la creciente disponibilidad de datos han transformado la manera en que se toman decisiones, convirtiendo la información en un activo altamente cotizado. Sin embargo, esta evolución se ha concentrado principalmente en el deporte profesional, quedando el ámbito semiprofesional al margen por distintas razones, entre ellas la falta de recursos o conocimiento especializado.

Este trabajo de fin de máster tiene como objetivo explorar el potencial del análisis de datos y la toma de decisiones “data-driven” dentro del contexto del deporte semiprofesional. Para ello, se aplicarán los conocimientos, técnicas analíticas y herramientas adquiridas a lo largo del Máster en Business Analytics, con el fin de desarrollar un enfoque estratégico basado en datos.

Dado nuestro contexto geográfico y la cercanía cultural con este deporte, el estudio se centrará específicamente en el fútbol semiprofesional en España. Esta elección no solo se debe a la accesibilidad y conocimiento del entorno, sino también a las distintas oportunidades de negocios que se pueden detectar

A lo largo del proyecto se realizará un análisis del sector, incluyendo una revisión de antecedentes y casos de uso actuales, así como la exploración de herramientas tecnológicas que puedan facilitar la integración de estrategias analíticas en entidades deportivas semiprofesionales.

De manera paralela, se desarrollará una herramienta que permita analizar eventos deportivos de manera sistemática. Esta herramienta tendrá como finalidad transformar observaciones cualitativas en datos estructurados que puedan ser utilizados posteriormente para generar análisis, identificar patrones de rendimiento y apoyar la toma de decisiones dentro de las organizaciones deportivas semiprofesionales.

2. Justificación

El análisis de datos en el deporte profesional actual se ha consolidado como una herramienta fundamental para optimizar el rendimiento tanto de los jugadores como de los equipos. Lejos de ser una simple tendencia, se ha convertido en un pilar estratégico en la gestión deportiva moderna. Hoy en día, las instituciones de élite del deporte obtienen información detallada mediante tecnologías avanzadas como el análisis de vídeo, sensores biométricos y software especializado en big data. Estas herramientas permiten recopilar y procesar grandes volúmenes de datos en tiempo real, generando una visión profunda y precisa del rendimiento individual y colectivo.

La creciente relevancia del análisis de datos en el fútbol no es una moda, sino una transformación en la forma en que se entiende gestiona y compite en este deporte. Figuras clave del panorama futbolístico internacional han reconocido públicamente el impacto de los datos en este deporte. Ramón Rodríguez Verdejo (“Monchi”), director deportivo del Sevilla FC, lo resume con claridad: *“Hoy no se puede fichar sin datos. Lo que antes era intuición ahora se complementa con la información objetiva.”* Esta afirmación refleja cómo la intuición, aunque sigue siendo importante, ya no es suficiente por sí sola en la toma de decisiones deportivas. (Contreras, 2020)

Los datos permiten validar hipótesis, reducir el margen de error y optimizar cada movimiento, desde el fichaje de un jugador hasta la preparación de un partido. En la misma línea, Ian Graham, ex director de investigación del Liverpool FC, subraya que *“el trabajo del analista de datos no es reemplazar al entrenador, sino darle las mejores herramientas para tomar decisiones.”* Esta visión resalta el carácter complementario del análisis: no se trata de sustituir la experiencia y el conocimiento del cuerpo técnico, sino de potenciarlo con información precisa, contextualizada y útil. (Schoenfeld, 2019)

2.1 Objetivos de Desarrollo Sostenible (ODS)

Este Trabajo Final de Máster está en consonancia con dos Objetivos de Desarrollo Sostenible definidos por la Agenda 2030 de las Naciones Unidas: el ODS 3 (Salud y bienestar) y el ODS 4 (Educación de calidad), ambos con una relación directa con el efecto que las tecnologías de análisis de datos pueden generar en el sector deportivo, incluso en entornos no deportivos (*Naciones Unidas, s.f.-a; Naciones Unidas, s.f.-b*).

Respecto al ODS 3, este proyecto sugiere una solución tecnológica que ayuda a potenciar la salud y el bienestar de los futbolistas semiprofesionales. Mediante la utilización de dashboards personalizados que examinan el desempeño físico e individual, los equipos pueden adquirir datos exhaustivos sobre la carga física, el progreso del estado físico y posibles indicadores de sobre entrenamiento o peligro de lesión. Esta habilidad de seguimiento continuo facilita a entrenadores y preparadores físicos la toma de decisiones más fundamentadas que prevengan problemas de salud y promuevan una organización más balanceada de las sesiones de entrenamiento. En este contexto, la herramienta no solo mejora el desempeño en el deporte, sino que también fomenta una práctica deportiva más segura, saludable y sostenible en el largo plazo.

En contraposición, el proyecto cumple con el ODS 4, ya que fomenta el acceso a una educación de alta calidad mediante el aprendizaje constante en el sector deportivo. Los dashboards funcionan como herramientas educativas para jugadores y equipos técnicos, simplificando la comprensión de información esencial en términos técnicos, tácticos y físicos. Esto promueve una cultura de perfeccionamiento continuo fundamentada en pruebas, en la que los jugadores pueden entender de manera más efectiva su propio desempeño y los entrenadores pueden mejorar sus tácticas de formación y competencia. La inclusión del análisis de datos en el proceso de formación transforma el ambiente del fútbol semiprofesional en un ámbito de entrenamiento donde la tecnología funciona como un impulsor del aprendizaje y la profesionalización.

Esta propuesta evidencia que el uso de tecnologías de análisis de datos accesibles en el fútbol no solo influye de manera positiva en el desempeño deportivo, sino que también favorece el crecimiento personal, la salud física y la educación constante de todos los participantes, en consonancia con los principios esenciales de los ODS mencionados.

2.2 Problema y finalidad

La realidad se diferencia claramente de la situación del fútbol amateur y juvenil, en la que el acceso a estos recursos continúa siendo bastante restringido debido a obstáculos económicos, logísticos y técnicos. En la actualidad, los equipos de fútbol amateur no cuentan con instrumentos accesibles que les faciliten evaluar su desempeño basándose en datos objetivos. La mayor parte de los estudios se llevan a cabo manualmente o mediante observaciones subjetivas, lo que restringe la mejora constante. El objetivo de este proyecto es desarrollar una aplicación que pueda obtener, manejar y mostrar datos pertinentes (tales como posesión, áreas de juego, acciones sobresalientes o estadísticas individuales) mediante el análisis automático de vídeos de partidos. *(Sia Academy, n.d.)*

En años recientes, el aumento en la disponibilidad de instrumentos de grabación asequibles (como smartphones con cámaras de alta definición o cámaras deportivas asequibles) ha creado una nueva oportunidad para disminuir esta desigualdad. Cada vez más equipos registran sus partidos, aunque pocas veces poseen los medios o la habilidad para obtener información relevante de dicho material. Esto sugiere un panorama donde hay un enorme volumen de datos potenciales sin explotar. *(Catapult, 2025)*

El proyecto busca democratizar el acceso al análisis táctico y técnico mediante herramientas automatizadas de análisis de vídeo, aprovechando técnicas de visión por computador, aprendizaje automático y procesamiento de datos. Esta iniciativa no solo pretende mejorar el rendimiento deportivo en contextos de menor presupuesto, sino también fomentar una cultura del dato más inclusiva, que acerque los beneficios del análisis profesional a todos los niveles del fútbol.

3. Objetivos

3.1 Objetivo general

En este contexto, la presente propuesta de Trabajo de Fin de Máster tiene como objetivo el desarrollo de una aplicación funcional que permita analizar y visualizar datos relevantes de partidos de fútbol amateur y juvenil a partir del procesamiento automático de vídeos.

3.2 Objetivos específicos

Para alcanzar este objetivo general, se han definido una serie de objetivos específicos que guían el desarrollo técnico y metodológico de la propuesta.

- Investigar técnicas de visión por computador y aprendizaje automático aplicadas al análisis de partidos de fútbol.
- Diseñar una arquitectura de sistema que permita el procesamiento eficiente de vídeos.
- Implementar algoritmos para el reconocimiento de jugadores, detección de eventos y extracción de estadísticas.
- Desarrollar una interfaz accesible para usuarios no expertos que permita consultar informes y visualizar jugadas clave.
- Validar la utilidad de la aplicación mediante pruebas con equipos amateur reales.

4. Marco teórico

4.1 Introducción al Análisis de Datos

4.1.1 Business Analytics

El análisis de datos se ha convertido en una de las áreas más innovadoras en el sector empresarial. Cada vez es más visible la implementación de departamentos o empleados especializados en análisis de negocios. ¿Entonces, a que se refiere el business analytics?

En términos generales el término business analytics se refiere al análisis de datos aplicados a los negocios. Un business analyst es capaz de recopilar, procesar analizar y presentar datos con el objetivo de tomar decisiones informadas. A través de este proceso es que las empresas son capaces de descubrir patrones o tendencias que les permitirá entender resultados internos, analizar la situación del mercado actual o predecir situaciones futuras y adelantarse ante su competencia. (*Esdan Business School, 2024*)

Encontramos distintos tipos de análisis dentro de la rama de business analytics:

Análisis descriptivo: El análisis descriptivo tiene como objetivo principal proporcionar una visión comprensible de los eventos pasados y haciendo uso de estadísticas y visualizaciones. Esta aproximación facilita la interpretación de grandes volúmenes de datos, contribuyendo a una comprensión compartida por parte de los distintos actores organizativos.

Análisis de diagnóstico: El análisis de diagnóstico busca profundizar en las causas que explican los comportamientos observados en el pasado. Se centra en identificar patrones, anomalías o relaciones causales mediante técnicas como la minería de datos, análisis de correlación y segmentación de información.

Análisis predictivo: El análisis predictivo se basa en la utilización de modelos estadísticos y algoritmos de aprendizaje automático para anticipar comportamientos o eventos futuros. Este tipo de análisis se fundamenta en los datos históricos recopilados a través de análisis

descriptivos y diagnósticos, con el fin de construir modelos capaces de estimar probabilidades y tendencias.

Análisis Prescriptivo: Finalmente, el análisis prescriptivo representa la evolución natural del análisis predictivo, ya que no solo estima qué podría suceder, sino que además propone acciones concretas para alcanzar un resultado deseado. Este tipo de análisis integra técnicas de optimización, simulación y toma de decisiones automatizada.

4.1.2 Principios de una cultura “data-driven”

La cultura *data-driven* hace referencia a una filosofía organizacional en la que las decisiones estratégicas, operativas y tácticas se fundamentan sistemáticamente en el análisis de datos. Según McKinsey, las empresas orientadas por los datos en 2025 no solo utilizarán la información como apoyo, sino que la integrarán profundamente en todos los niveles de toma de decisiones.

De acuerdo con *McKinsey & QuantumBlack*, para 2025 las organizaciones que realmente sean *data-driven* deberán haber avanzado en estas siete dimensiones clave:

Integración de los datos en cada decisión, interacción y proceso

El uso de datos debe estar presente “de forma natural y regular” en el trabajo diario de casi todos los empleados. En lugar de resolver problemas mediante largos planes, se emplean técnicas analíticas para tomar decisiones rápidas y fundamentadas.

Procesamiento y entrega de datos en tiempo real

Las empresas deben capturar, procesar y visualizar datos casi instantáneamente para que las decisiones se basen en información actual. Esto requiere infraestructuras de análisis en tiempo real y tecnologías edge computing.

Almacenamiento flexible de datos integrados y listos para usar

Se adoptan bases de datos modernas (por ejemplo, series temporales, grafos, NoSQL) que permiten integrar fuentes estructuradas y no estructuradas de manera ágil, facilitando el consumo instantáneo de datos.

Modelo operativo que trata los datos como productos

Los datos se gestionan como productos: cada activo tiene un equipo responsable, asegurando calidad, acceso, seguridad y capacidad de evolución continua.

Rol ampliado del Chief Data Officer (CDO)

El CDO pasa de gestionar gobernanza y normativas a generar valor corporativo, liderando la estrategia de datos, creando ingresos vía monetización y colaborando estrechamente con unidades de negocio.

Pertenencia a ecosistemas de datos

Las organizaciones colaboran activamente con socios externos, compartiendo datos en plataformas comunes o mercados de datos. Esto permite generar conocimientos más ricos a partir de la combinación de fuentes diversas.

Gestión de datos priorizada y automatizada para privacidad, seguridad y resiliencia

Se implementan procesos automatizados de gestión de datos, garantizando protección, calidad constante y recuperación ágil ante incidentes, con un enfoque ético y normativo robusto.

4.2 Análisis de Datos en el Deporte

La innovación ha llegado con mucha inercia en el deporte, iniciando una transformación impulsada por el avance de las tecnologías de la información y el análisis de datos. El análisis deportivo, entendido como la recopilación, procesamiento e interpretación sistemática de datos relacionados con el rendimiento, las tácticas y otros aspectos del juego, se ha convertido en una herramienta fundamental para la toma de decisiones estratégicas en clubes, entrenadores, cuerpos técnicos y todas las partes implicadas en las instituciones deportivas.

Gracias al desarrollo de sensores, cámaras, software de seguimiento y técnicas de inteligencia artificial, es posible capturar enormes volúmenes de información durante eventos deportivos. Estos datos pueden abarcar desde métricas físicas de los atletas, hasta eventos tácticos o individuales que afectan el desarrollo del juego o el rendimiento de un equipo. En disciplinas como el fútbol, el baloncesto, el béisbol o el tenis, el uso de estas herramientas permite obtener ventajas competitivas, optimizar el rendimiento y reducir riesgos de lesiones.

El análisis deportivo ha dejado de ser exclusivo de contextos científicos o académicos para convertirse en un componente clave de la gestión profesional en el ámbito deportivo. A través de técnicas estadísticas y algoritmos de machine learning, se pueden identificar patrones, predecir resultados, evaluar el rendimiento individual y colectivo, e incluso mejorar la experiencia del aficionado. Esta evolución ha dado lugar a un enfoque más objetivo y basado en evidencia para comprender y gestionar el deporte moderno. *(Análisis Deportivo: Cómo Utilizan El análisis De Datos Los Distintos Deportes, 2024)*

4.2.1. El análisis de datos en el sector futbolístico

El análisis de datos ha revolucionado profundamente el funcionamiento del sector deportivo en los últimos tiempos. El desarrollo de tecnologías que ayudan en la recogida de datos (como pueden ser GPS, videocámaras o dispositivos de monitoreo fisiológico) y la creación de herramientas analíticas ha permitido a clubes, entrenadores, analistas a tomar decisiones más basadas y respaldadas en evidencias cuantificables.

En términos generales, el análisis de datos en el ámbito deportivo se aplica en 3 áreas fundamentales: rendimiento de jugador, estrategia de equipo y prevención de lesiones. Por ejemplo, es posible supervisar la carga física de un futbolista para evitar sobreesfuerzos y de esta manera prevenir lesiones. Por otro lado, el análisis de datos permite analizar los patrones de juego del equipo y del rival para diseñar estrategias que permitan contrarrestar aquellas debilidades internas o fortalezas del rival. El uso de los análisis también ha permitido que los

cuerpos técnicos puedan analizar indicadores técnicos como precisión de pases, eficacia de disparos o posicionamiento defensivo.

A su vez, la expansión de plataformas comerciales de análisis deportivo ha facilitado el acceso a grandes volúmenes de datos, incluso para equipos con menor presupuesto. Este fenómeno ha promovido la profesionalización de nuevas profesiones como los data analysts deportivos o los performances analysts, cuya presencia es cada vez mayor en cuerpos técnicos.

El fútbol, ha pasado de ser un deporte tradicional sujeto a la intuición, la experiencia y la percepción subjetiva, hacia la incorporación del Big Data en los procesos de análisis y toma de decisiones. Hoy en día, el uso sistemático de grandes volúmenes de datos ha revolucionado la forma en que se interpreta el juego, permitiendo una comprensión más profunda, objetiva y precisa de cada componente que lo compone.

Gracias a la implementación de tecnologías mencionadas anteriormente, es posible registrar y analizar millones de datos en tiempo real sobre variables físicas, técnicas y tácticas.

El Big Data ha introducido un enfoque basado en evidencia para el diseño de estrategias, la preparación de partidos y la evaluación del rendimiento. Los cuerpos técnicos pueden ajustar esquemas tácticos basándose en datos históricos y en simulaciones predictivas, mientras que los clubes pueden tomar decisiones más acertadas en el mercado de fichajes mediante modelos de valoración de jugadores basados en datos objetivos.

Además, el uso de Big Data también ha potenciado el ámbito de la prevención de lesiones, mediante el monitoreo constante de la carga física y la fatiga acumulada, lo cual permite personalizar entrenamientos y reducir riesgos. Todo esto ha contribuido a la profesionalización del fútbol en todos sus niveles, generando una ventaja competitiva para aquellos clubes que han sabido incorporar la analítica de datos de manera estratégica. (*Universidad Alfonso X el Sabio, 2025*)

4.2.2. Impacto del análisis de datos en el rendimiento y la estrategia

El impacto del análisis de datos en el rendimiento y la estrategia es evidente tanto a nivel individual como colectivo. A nivel individual, el uso de datos permite analizar el progreso de los deportistas, tanto en física, como técnica o táctica, esto ayuda a la personalización de entrenamientos y planes de recuperación. Mientras que, a nivel colectivo, los datos facilitan la modificación de esquemas tácticos, la optimización de alineaciones y la toma de decisiones. Es claro el Big Data, ha transformado profundamente la estrategia y el rendimiento en el fútbol moderno. Esta herramienta se ha convertido en un elemento esencial en la planificación de entrenamientos, el diseño táctico, la prevención de lesiones y la toma de decisiones durante la competición. Gracias a la capacidad para recopilar, procesar e interpretar grandes volúmenes de datos en tiempo real, los cuerpos técnicos ahora disponen de una base empírica sólida sobre la cual optimizar cada detalle del juego.

En cuanto a la estrategia colectiva, el Big Data permite analizar patrones de comportamiento propios y del rival: zonas de mayor vulnerabilidad, fases del juego en las que se pierde posesión, o tendencias de presión y recuperación. Un caso emblemático es el del Getafe CF dirigido por Pepe Bordalás, cuya propuesta táctica basada en la agresividad en balones divididos y segundas jugadas fue reforzada por análisis detallados que identificaban qué jugadores ganaban más duelos aéreos y contra qué rivales resultaba más ventajoso aplicar esa estrategia. (*Julián Burgos, 2022*)

Respecto al rendimiento individual, el análisis de datos permite identificar fortalezas y debilidades técnicas de cada jugador. Por ejemplo, si se detecta que un delantero como Ángel Rodríguez sólo convierte un bajo porcentaje de jugadas peligrosas tras recibir el balón en tres cuartos de campo, el cuerpo técnico puede tomar decisiones de alineación o enfocar el entrenamiento a mejorar la toma de decisiones en esas situaciones.

Además, en el ámbito físico, los sensores y plataformas de monitorización permiten optimizar la carga de trabajo y la recuperación de los jugadores, reduciendo el riesgo de lesiones. Se monitorizan variables como distancia recorrida, esfuerzo explosivo, calidad del sueño o

recuperación postpartido. Esta integración permite personalizar el entrenamiento y establecer protocolos de prevención. Así, se maximiza el rendimiento sin comprometer la salud del futbolista.

También es posible destacar el uso del Big Data en el ámbito de los fichajes, permitiendo evaluar objetivamente el encaje de un jugador en un sistema de juego determinado. El caso de Monchi en el Sevilla FC es paradigmático: usando plataformas de análisis masivo, seleccionó al central Diego Carlos aplicando filtros sobre balones aéreos ganados y precisión en pases largos, lo cual condujo a una contratación eficiente y alineada con los requerimientos tácticos de Julen Lopetegui. (*CONTRERAS, 2020*)

Un ejemplo internacional de éxito estratégico impulsado por el análisis de datos fue el del Leicester City en la temporada 2015/2016, cuando se proclamó campeón de la Premier League contra todo pronóstico. Utilizando herramientas como Prozone Sports, el club identificó perfiles de jugadores subvalorados, pero estadísticamente eficientes, como N'Golo Kanté, y diseñó un sistema de juego optimizado que maximizó las virtudes del plantel. La estrategia defensiva compacta, el juego directo y el contraataque veloz no fueron casuales, sino fruto de decisiones respaldadas por datos rigurosamente analizados. (*López-Zafra, 2019*)

Un claro ejemplo puede ser Stats Perform, una plataforma que fusiona inteligencia artificial con datos precisos suministrados por Opta, facilitando a los entrenadores la observación de patrones en los partidos, así como la detección de fallos, y la predicción de resultados futuros. El estudio de eventos (como duelos, tiros o posesión) junto al análisis posicional se han establecido como componentes esenciales para la toma de decisiones.

De la misma manera, el uso de machine learning y algoritmos predictivos en Stats Perform empieza a cobrar importancia en la predicción de lesiones, la identificación de talento o la proyección de beneficios futuros. El análisis de datos ha dejado de ser una herramienta auxiliar para ser un elemento esencial en la planificación deportiva.

El uso de datos en el fútbol ha generado una transformación en la forma en que los equipos entienden y gestionan el juego, impactando directamente en el rendimiento y la estrategia en las distintas áreas. Gracias a la innovación del análisis de datos, los clubes cuentan hoy con herramientas más precisas para la toma de decisiones, desde la planificación táctica hasta el scouting de jugadores. Este enfoque permite identificar patrones de juego, anticipar comportamientos del rival y ajustar las estrategias de forma más eficiente. A nivel individual, los datos permiten evaluar de manera detallada el rendimiento de cada jugador, lo que facilita intervenciones personalizadas para trabajar en fortalezas y detectar debilidades. Además, el monitoreo de cargas físicas a través de métricas cuantificables ha contribuido significativamente a la prevención de lesiones, optimizando el estado físico de los jugadores a lo largo de la temporada.

El Barrow AFC en Inglaterra, es un ejemplo claro que demuestra cómo incluso clubes con recursos limitados han podido mejorar su rendimiento competitivo mediante el uso de datos, lo que valida su utilidad más allá de los grandes presupuestos. En definitiva, el análisis de datos se ha consolidado como una herramienta clave en la evolución moderna del fútbol, no solo como complemento, sino como herramienta fundamental en el día a día de los clubes y los entrenadores. (*Hudl Statsbomb Customer Case Study: Barrow AFC - Hudl Statsbomb | Data Champions, 2023*)

4.2.3. Diferencias entre el deporte profesional y semiprofesional

A pesar del enorme avance que ha supuesto el análisis de datos en el fútbol profesional, su implementación en el contexto semiprofesional sigue enfrentando diversas barreras que condicionan tanto su adopción como su efectividad. Estas limitaciones pueden clasificarse en tres grandes áreas: limitaciones tecnológicas, económicas y humanas.

1. Limitaciones tecnológicas

Una de las principales diferencias entre el fútbol profesional y el semiprofesional es el acceso desigual a las herramientas tecnológicas. Mientras que los grandes clubes disponen de sistemas avanzados de tracking, plataformas de big data y software de video análisis, los equipos semiprofesionales deben conformarse con soluciones más limitadas o incluso recurrir al análisis manual. El uso de tecnologías como cámaras multicámara, sensores GPS o software como Wyscout, StatsBomb o Instat resulta inalcanzable para muchos de estos equipos, que carecen de la infraestructura necesaria para su implementación. (*Sports Data Campus, 2025*)

2. Limitaciones económicas

El análisis de datos, especialmente a gran escala, requiere una inversión considerable. No solo en licencias de software especializado, sino también en hardware, almacenamiento de datos, suscripciones a bases de datos y personal cualificado. Para muchos clubes semiprofesionales, cuyo presupuesto se destina casi en su totalidad a cuestiones operativas básicas (viajes, equipamiento, salarios mínimos), este tipo de inversión es inviable. Incluso en casos donde existe voluntad de incorporar el análisis, las prioridades presupuestarias no permiten desarrollar departamentos de analítica de forma sostenible.

3. Limitaciones humanas y formativas

Otra gran barrera se encuentra en la falta de personal cualificado. Los clubes profesionales cuentan con analistas de datos, técnicos deportivos con formación específica y departamentos multidisciplinares que trabajan de manera coordinada. En el ámbito semiprofesional, por el contrario, muchas funciones son asumidas por una misma persona (por ejemplo, el entrenador que también hace de preparador físico y analista). La falta de tiempo y formación específica impide que se pueda extraer valor real de los datos, incluso cuando se cuenta con herramientas básicas.

A su vez, se nota un aumento en la adopción de tecnologías asequibles en el ámbito del deporte semiprofesional. El avance de soluciones asequibles, tales como cámaras de grabación automáticas, plataformas de análisis simplificadas y la disponibilidad de contenidos educativos

en línea, está comenzando a disminuir la diferencia entre ambos niveles. En esta línea, el futuro se proyecta hacia una gradual profesionalización del análisis de datos en todas las esferas del deporte. *(Díaz & Ritchie, 2021)*

4.3 Fuentes y Tipos de Datos en el Deporte

El análisis del rendimiento en el fútbol se ha sofisticado considerablemente con la incorporación de tecnologías que permiten recopilar datos de forma precisa, constante y automatizada. FIFPRO identifica cuatro grandes categorías de datos fundamentales en este proceso: físicos, técnicos, tácticos y biométricos.

Cada una de estas categorías aporta una perspectiva específica. Los datos físicos permiten comprender el esfuerzo y la exigencia del juego desde el punto de vista del rendimiento fisiológico. Los datos técnicos recogen las acciones realizadas por el jugador con el balón y su eficacia en la ejecución. Los datos tácticos se centran en el posicionamiento, la interacción con los compañeros y la integración en los esquemas del equipo. Por último, los datos biométricos se vinculan con aspectos relacionados con la salud, la prevención de lesiones y el bienestar general del deportista.

Este enfoque integral permite no solo analizar lo que sucede durante un partido, sino también planificar entrenamientos, diseñar estrategias, prevenir sobrecargas y adaptar decisiones a las necesidades individuales de cada futbolista. No obstante, como también subraya FIFPRO, el uso de estos datos debe realizarse dentro de un marco ético que respete la privacidad, la transparencia y los derechos de los jugadores sobre su propia información. *(Datos De Rendimiento De Los Futbolistas - FIFPRO World Players' Union, n.d.)*

4.3.1. Datos de rendimiento deportivo

Los datos de rendimiento son los que detallan el desempeño de un atleta o de un equipo durante el torneo. Para deportes colectivos como el fútbol, estos datos abarcan indicadores como el

número de pases finalizados, disparos a puerta, intercepciones, recuperaciones, duelos ganados, entre otros. Su estudio posibilita valorar la efectividad técnica y táctica de los jugadores, además de detectar patrones de éxito o debilidad.

En el ámbito individual, se puede evaluar el rendimiento a través de indicadores como la eficacia en acciones fundamentales (como regates, disparos o entradas), mientras que en el ámbito colectivo se examinan elementos como la posesión, la organización defensiva u ofensiva, y la conducta del equipo en diferentes etapas del juego. Plataformas como Instat, Wyscout o StatsBomb ofrecen este tipo de datos de manera organizada mediante eventos clasificados y estadísticas sofisticadas.

4.3.2. Datos físicos y fisiológicos

Estos datos proporcionan información sobre la condición física del jugador, su carga física y sus habilidades de resistencia, velocidad o potencia. Se obtienen mediante dispositivos como pulsómetros, acelerómetros, sistemas GPS o bandas de frecuencia del corazón. Dentro de los indicadores más empleados se incluyen la distancia total cubierta, el número de sprints, la velocidad máxima, la frecuencia cardíaca media y el esfuerzo detectado.

El estudio físico y fisiológico no solo facilita la mejora del rendimiento a través de una planificación más adecuada del entrenamiento, sino que también resulta esencial para prevenir lesiones. La regulación de la carga tanto aguda como crónica, o la utilización de indicadores como el Player Load (métrica que mide el esfuerzo físico), contribuyen a identificar instantes de sobrecarga y a elaborar tácticas de recuperación más eficaces.

4.3.3. Datos tácticos y posicionales

Los datos tácticos y posicionales proporcionan una visión más estructural del juego, facilitando el análisis de cómo los jugadores se ubican y se mueven durante el encuentro. Su principal

origen son los sistemas de monitorización por vídeo o sensores, los cuales registran coordenadas geográficas en tiempo real. Estas coordenadas facilitan la determinación de las distancias entre los jugadores, las líneas de pase, la ocupación de espacios y las formaciones dinámicas.

En años recientes, el análisis posicional ha ganado un significativo papel debido a tecnologías como el seguimiento de datos y el evento de datos, que facilitan la reconstrucción de juegos, la evaluación de la compacidad del equipo o el análisis de la sincronización entre líneas. Este tipo de estudio es esencial para entender el comportamiento táctico y su influencia en el desempeño grupal.

4.4 Tecnologías para la Captura y Procesamiento de Datos Deportivos

4.4.1. Sensores y wearables

Los sensores portátiles (“wearables”), han revolucionado la manera en que se supervisa el rendimiento físico de los atletas. Estos dispositivos, que pueden incorporarse en chalecos, bandas, camisetas o incluso calzado, recopilan información fisiológica y cinematográfica en tiempo real. Algunos de los sensores más habituales son los de aceleración, giroscopios, magnetómetros, pulsómetros y GPS.

Los wearables facilitan el registro de variables en situaciones reales de actividad física o competencia, tales como la frecuencia cardíaca, la aceleración, la carga de impacto, la distancia cubierta o la rapidez. Este tipo de información es esencial para regular el volumen de entrenamiento, evitar lesiones y valorar el desempeño físico. Además, cada vez más atletas semiprofesionales y amateurs emplean relojes inteligentes o bandas deportivas

4.4.2. Sistemas de tracking y geoposicionamiento

Los sistemas de seguimiento o tracking posibilitan registrar la ubicación y el desplazamiento de los jugadores (y, en ciertas situaciones, del balón) a través del tiempo. Hay diversas técnicas

para recolectar estos datos, entre las que sobresalen los sistemas GPS, los sistemas ópticos basados en cámaras (como TRACAB o Second Spectrum) y los sistemas de radiofrecuencia.

En el fútbol, estos sistemas facilitan la reconstrucción de trayectorias de los jugadores durante el encuentro, el análisis de su posición, la evaluación de la compacidad del equipo o la identificación de desalineaciones tácticas. Los datos recolectados se guardan en coordenadas geográficas (x, y) sincronizadas con el tiempo (t), lo que permite investigaciones dinámicas y espaciales del desempeño. (*Tecnología Vestible En El Deporte, 2024*)

4.4.3. Herramientas de análisis y visualización de datos

Una vez recolectados, los datos deben ser procesados, examinados e interpretados para crear valor. En este contexto, los programas de análisis y visualización desempeñan un rol esencial. Herramientas como Hudl, LongoMatch, Sportscode, Tableau, Python o Power BI se emplean extensamente para ordenar, examinar y mostrar los datos de manera óptima y entendible.

Estas plataformas facilitan la identificación de sucesos (como pasos, tiros, faltas) hasta la visualización sofisticada de mapas de calor, redes de pases, áreas de acción y diagramas de tiempo. Algunas incluyen algoritmos de inteligencia artificial para identificar tendencias y realizar pronósticos, mientras que otras se enfocan en simplificar la toma de decisiones tácticas a través de dashboards interactivos.

4.5 Visión por Computador (Computer Vision) Aplicada al Deporte

4.5.1. Fundamentos de la visión por computador

El *computer vision* fusiona métodos de procesamiento de imágenes, deep learning, identificación de patrones y análisis espacial para obtener datos relevantes de fuentes visuales. Zheng et al. (2025) insinúan que dentro de los fundamentos técnicos más relevantes en este ámbito tenemos:

- Detección de objetos: se emplean redes neuronales convolucionales (CNN) como YOLO (You Only Look Once), Faster R-CNN o SSD para identificar y categorizar elementos en una imagen (jugadores, balón, árbitros, líneas de campo, entre otros)
- Segmentación semántica: facilita el reconocimiento y la separación de zonas concretas de una imagen (como el campo, las áreas, las porterías).
- Seguimiento (tracking): Métodos como SORT, Deep SORT o algoritmos basados en filtros Kalman posibilitan el seguimiento del desplazamiento de un objeto a través del tiempo en una secuencia de vídeo.
- Reconstrucción 2D/3D: a través de modelos geométricos o visión estereoscópica, es posible calcular la ubicación y dirección de los jugadores en el espacio.

4.5.2. Aplicaciones en la detección y seguimiento de jugadores

Una de las aplicaciones más relevantes de la visión computacional en el ámbito deportivo es la identificación y monitorización automática de los jugadores y del balón. Mediante la formación de modelos de detección de objetos en grandes cantidades de datos etiquetados, se puede identificar con exactitud a cada participante en cada fotograma de un vídeo.

El monitoreo constante facilita la reconstrucción de rutas, la medición de velocidades y el análisis de patrones de movimiento sin la necesidad de dispositivos (como GPS o sensores). Esto resulta particularmente útil en competencias oficiales donde no es viable el uso de wearables, o en situaciones de presupuesto reducido donde no se cuenta con estos dispositivos.

4.5.3. Análisis automático de eventos: goles, faltas, posesión, etc.

La visión computacional también se emplea para el estudio automático de sucesos fundamentales en los juegos, tales como goles, faltas y tenencia del balón. Al fusionar métodos de detección de objetos y seguimiento, estos sistemas tienen la capacidad de detectar y documentar automáticamente sucesos relevantes, produciendo estadísticas exhaustivas y

resumen del juego. Este tipo de evaluaciones brinda a entrenadores y analistas una perspectiva más exacta e imparcial del progreso del partido, lo que facilita la toma de decisiones fundamentadas y la creación de tácticas más eficaces.

Algunas soluciones avanzadas, como las de Second Spectrum, SkillCorner o sistemas creados por clubes con departamentos de Investigación y Desarrollo, emplean esta tecnología para producir estadísticas avanzadas de forma automática y en directo. Además, es un pilar fundamental para sistemas de retransmisión mejorada. (*Vidal-Codina, F., Evans, N., El Fakir, B. et al, 2022*)

4.5.4. Limitaciones y desafíos actuales

A pesar de los progresos, el uso del computer vision en el deporte todavía se topa con restricciones y retos significativos, como la calidad del video. Los videos de baja resolución, mala iluminación o ángulos restringidos disminuyen considerablemente la exactitud de los algoritmos. Otras restricciones son las oclusiones y solapamientos, que, en deportes como el fútbol, los jugadores a menudo se encuentran en contacto directo, lo que complica la segmentación y/o la monitorización. Respecto a la capacidad informática, el análisis en tiempo real continúa siendo costoso, aunque cada vez se vuelve más accesible gracias a las GPUs y los servicios de nube. Finalmente, los modelos formados en un torneo o estadio no siempre logran ajustarse adecuadamente a otros entornos, cámaras o niveles.

La superación de las restricciones actuales de la visión computacional permitirá ampliar su uso significativamente más allá del ámbito deportivo de élite, generando nuevas posibilidades en entornos contemporáneos como el deporte semiprofesional, amateur o educativo. En estas etapas, donde usualmente no se cuenta con recursos de tecnología de vanguardia ni con equipos especializados en análisis de datos, la capacidad de automatizar el seguimiento, evaluación y análisis táctico a partir de sencillas grabaciones en vídeo constituye un avance de calidad.

4.6 Herramientas de Análisis de Eventos Deportivos

4.6.1. Plataformas comerciales

En la actualidad, encontramos distintas plataformas comerciales dedicadas al análisis de eventos deportivos a nivel global, centradas principalmente en el análisis de competiciones a nivel profesional, de campeonatos tanto nacionales como internacionales. A lo largo de este apartado, vamos a realizar una revisión de las más utilizadas a nivel global por distintos profesionales dentro del sector.

En el mercado existen varias plataformas que realizan este tipo de análisis de eventos deportivos, centrándose en el fútbol. Se podrían separar por un lado aquellas que utilizan un análisis dual, es decir, que además de aplicar técnicas de análisis de vídeo como sería el *computer vision*, se apoyan también en operadores humanos que bien pueden estar en el propio campo, o visualizar el evento posteriormente. Estos operadores humanos se encargarían de anotar eventos sucedidos dentro del partido, como pases, faltas...etc. Además, anotan quién realiza ese evento y en qué minuto. El hecho de utilizar operadores humanos les permite hacer una comparación de lo detectado por sus sistemas automatizados y darle robustez a sus análisis. En este sentido, las plataformas más utilizadas serían *WyScout* y *Opta* (ahora perteneciente a *Stats Perform*). Ambas plataformas preparan análisis detallados e informes para tanto equipos como para jugadores concretos.

Por otro lado, habría plataformas que se centran en intentar automatizar el proceso de análisis deportivo, utilizando únicamente inteligencia artificial y técnicas de machine learning. La más utilizada y con más renombre a nivel profesional dentro de este tipo de soluciones tecnológicas sería *Second Spectrum* que es el proveedor oficial de datos de *tracking* para grandes ligas profesionales como sería el caso de la Premier League o la MLS (Major League Soccer). En este caso se utilizan cámaras especializadas para realizar tracking 3D, detección automática de jugadas, o visualizaciones interactivas.

También cabe destacar en este sentido, la plataforma *Metrica Sports*, que permite que sea el propio usuario quien suba el vídeo a la plataforma y le permite hacer un tracking de jugadores y del campo de forma manual, siendo el usuario quien anota en el vídeo los elementos necesarios para que el modelo aprenda y posteriormente se realice de forma autónoma. Esta plataforma es utilizada por equipos de un nivel semiprofesional, que cuentan con un personal especializado en el análisis mucho más limitado. (Carpio, 2023)

4.6.2. Herramientas open source y bibliotecas técnicas

El análisis de datos aplicado en el deporte no solo se limita a plataformas que ofrecen análisis de distintos partidos o eventos deportivos. Desde hace unos años y gracias al auge de las herramientas de software libre, los equipos pueden ahora elaborar sus propios sistemas tanto de detección como de tracking y procesamiento de los datos obtenidos en los procesos anteriormente mencionados.

Empezaremos hablando de distintas librerías de Python utilizadas principalmente para el procesamiento de vídeo. Una de las librerías más utilizadas en este sentido sería OpenCV. Esta librería nos permite realizar tareas relacionadas con el procesamiento de vídeo, como sería leer y procesar vídeo frame a frame, realizar tareas de preprocesamiento, como el redimensionado de imágenes, o realizar anotaciones sobre el propio vídeo.

Otra de las librerías utilizadas sería *supervision*, la cuál es una librería más moderna, desarrollada por *Roboflow* que permite la simplificación de procesos a la hora de manipular los datos obtenidos a través del *computer vision*. Algunas de sus funciones serían la de utilizar los datos obtenidos para anotar en el vídeo las líneas del campo, las *bounding boxes*...etc.

Otras librerías utilizadas no solo en un contexto de análisis de datos aplicado a eventos deportivos, sino en multitud de escenarios serían Pandas, Numpy o Scipy. Estas librerías en este contexto se aplicarán principalmente en la etapa de procesado de los resultados obtenidos

tras el *computer vision*, y sobre todo para el posterior cálculo de las distintas estadísticas que se mostrarán en los *dashboards*

4.6.3. Modelos y frameworks de IA utilizados en deportes

Para realizar todo el procesamiento del vídeo, y extraer la información necesaria para la detección de eventos, sería necesaria la utilización de distintos modelos de detección. El modelo utilizado para este trabajo y uno de los más reconocidos a nivel mundial es YOLO, en este caso la versión 8. Las ventajas de utilizar este modelo son tanto su precisión como su velocidad. Además, se puede utilizar para distintos análisis dentro del mismo proyecto, como sería por un lado la detección de jugadores, y por otro lado la detección de los distintos puntos distintivos de un campo de fútbol. Otro modelo también utilizado en *computer vision* pero para otros propósitos como sería el caso de MediaPipe el cual se utiliza principalmente para la detección mediante vídeo de poses y posiciones del cuerpo humano.

Por otro lado, tras las primeras detecciones dentro del vídeo, los siguientes modelos que se utilizan serían todos aquellos relacionados con el tracking de los objetos detectados y que permite asignar a cada objeto detectado una identidad, y mantenerla a lo largo de los frames en los que se detecta dicho objeto. En este sentido, se pueden utilizar modelos como ByteTrack o DeepSORT, los cuales no solo integran el tracking de los objetos, sino que también realizan predicciones para mantener un seguimiento robusto y salvando posibles oclusiones del objeto que pueden producirse a lo largo del vídeo.

Si hablamos de los distintos frameworks utilizados en este tipo de proyectos, debemos hablar principalmente de PyTorch, el cuál es ampliamente utilizado por su gran flexibilidad y eficiencia. Es por ello, que muchos de los modelos de *computer vision* están implementados en este framework. Otro framework utilizado tanto en los deportes como en otros ámbitos sería Tensor Flow, el cual se utiliza más en aplicaciones ya que ofrece más facilidades a la hora de desplegar modelos de IA y tiene una interfaz más accesible a nivel de entrenamiento. (*La Inteligencia Artificial En El Deporte, 2024*)

4.6.4. Integración de flujos de trabajo analíticos en clubes semiprofesionales

El progreso de las herramientas de análisis deportivo, tanto comerciales como de código abierto, ha permitido que los clubes semiprofesionales comiencen a usar métodos basados en datos sin necesidad de grandes presupuestos. A pesar de tener menos recursos que el fútbol profesional, las tecnologías actuales y los procesos escalables ahora permiten el uso práctico del análisis táctico y de rendimiento en estas situaciones.

El flujo de trabajo que se lleva a cabo para la realización de diferentes análisis es muy similar tanto para clubes de un nivel semiprofesional como para un club totalmente profesional.

El flujo de trabajo comenzaría con la recopilación de datos, que en primera instancia sería aplicar computer vision sobre vídeos captados con diferentes cámaras para realizar todo el proceso de detección y tracking de los jugadores y balón. El objetivo de esto sería la obtención de datos relacionados con la geolocalización de los jugadores y el balón en el campo para pasar al siguiente paso.

El paso intermedio sería el procesamiento de dichos datos, que se utilizarán para la detección de eventos, como sería el caso de pases, tiros, intercepciones, conducciones...etc. Este proceso se realizaría mediante algoritmos que cotejan la posición del balón con la de los jugadores en un momento concreto (como podría ser un frame) obteniendo conclusiones de qué acción se realizó, qué jugador la realizó y a qué equipo pertenecía. Con esto se construiría el data set utilizado en el último paso.

El último paso sería el de la visualización de las distintas estadísticas y resultados en el formato que el equipo en cuestión estuviera solicitando, es para ello para lo que se realizan reuniones periódicas para determinar qué datos quiere el cuerpo técnico ver, y de qué forma.

4.7 Retos y Oportunidades del Análisis de Datos en el Deporte Semiprofesional

4.7.1. Limitaciones técnicas, económicas y humanas

Como hemos mencionado anteriormente, la principal limitación del deporte semiprofesional en este ámbito es la falta de recursos técnicos y económicos. En contraste con el ámbito deportivo profesional, que cuenta con departamentos específicos de análisis y tecnología de alta calidad, en los clubes semiprofesionales la infraestructura suele ser escasa o nula. A menudo no contamos con cámaras automatizadas, sensores GPS, programas especializados ni bases de datos estandarizadas.

Asimismo, los recursos humanos son limitados. En la mayoría de las situaciones, los entrenadores tienen que desempeñar varias responsabilidades (preparación física, dirección técnica, análisis, etc.), lo que complica una dedicación específica al manejo de datos. Finalmente, las restricciones temporales y logísticas (dado que muchos técnicos y jugadores no se enfocan únicamente en el deporte) obstaculizan la aplicación sistemática de metodologías de análisis.

4.7.2. Potencial transformador de un enfoque analítico

Pese a las restricciones presentes, el análisis de datos constituye una oportunidad estratégica para los clubes semiprofesionales. Incluso con recursos limitados, el uso de un enfoque analítico facilita la optimización de entrenamientos, la toma de decisiones más basadas en evidencia y la mejora del rendimiento tanto individual como grupal.

La disponibilidad de herramientas asequibles, como programas de análisis elemental, aplicaciones para móviles o cámaras automáticas, ha disminuido las barreras de entrada. Estas soluciones posibilitan la puesta en marcha de procesos de análisis sin la necesidad de inversiones significativas, promoviendo una profesionalización gradual. La incorporación de datos promueve una cultura más consistente e imparcial en el equipo técnico, lo que podría resultar en mejores rendimientos deportivos y una mayor habilidad para captar talento.

4.7.3. Casos emergentes y tendencias en clubes semiprofesionales

Recientemente, se han registrado numerosos ejemplos de clubes semiprofesionales que han empezado a incluir el análisis de datos en su estructura de operaciones. Algunos lo realizan mediante alianzas con universidades o centros de investigación, incorporando a estudiantes o profesionales jóvenes en actividades que llevan a cabo tareas de análisis táctico o de desempeño. Otros clubes han decidido implementar plataformas tecnológicas automatizadas que facilitan la recopilación y representación de datos.

Por ejemplo, muchos clubes han implementado el uso de cámaras como Veo, que registran partidos de manera independiente y producen resúmenes automáticos. Plataformas como Hudl Assist o Instat también han iniciado la provisión de planes personalizados para dispositivos de presupuesto reducido. Además, la expansión del conocimiento libre en deportes como el fútbol ha facilitado el surgimiento de comunidades de análisis autónomas que comparten recursos y métodos de fácil acceso.

4.8 Estudio de Caso: Aplicación Real del Análisis de Datos en el Fútbol

4.8.1. Contextualización del caso seleccionado (club, categoría, entorno)

El Leicester City Football Club, un equipo tradicionalmente considerado de segunda línea en el fútbol inglés, sorprendió al mundo del deporte al proclamarse campeón de la Premier League en la temporada 2015–2016. Esta hazaña fue calificada por muchos como uno de los mayores logros deportivos en la historia moderna, teniendo en cuenta que el club partía con una probabilidad de 5000 a 1 según las casas de apuestas para ganar el campeonato. (*Creasey, S. 2016*)

Previo a esa temporada, el Leicester luchaba por mantenerse en la categoría y disponía de un presupuesto significativamente inferior al de los clubes tradicionales como Manchester City, Arsenal o Chelsea. En ese contexto de desventaja estructural, el club decidió apostar por una

estrategia innovadora: aplicar el análisis avanzado de datos como herramienta clave para la toma de decisiones tanto deportivas como operativas. A través de un enfoque científico del rendimiento, el club logró competir al más alto nivel con una plantilla modesta pero extraordinariamente optimizada.

4.8.2. Herramientas utilizadas y metodología aplicada

Para implementar esta visión, el Leicester colaboró con Prozone Sports, una empresa especializada en sistemas de análisis de rendimiento basados en visión computacional, estadística avanzada y seguimiento en tiempo real. Prozone instaló un sistema de múltiples cámaras en el estadio y centros de entrenamiento, lo que permitió capturar datos detallados sobre la ubicación de cada jugador, sus trayectorias, zonas de acción, esfuerzos físicos y relaciones espaciales con los compañeros y rivales.

Además, se utilizaron tecnologías de GPS y sensores portátiles, que permitían medir variables como velocidad, aceleración, carga de trabajo y distancia recorrida durante los entrenamientos y partidos. Estos datos eran luego procesados y visualizados mediante dashboards personalizados que permitían al cuerpo técnico tomar decisiones tácticas y de planificación física basadas en evidencia.

Metodológicamente, el equipo utilizó modelos estadísticos como los Expected Goals (xG) para evaluar la calidad de las oportunidades creadas y concedidas. También se recurrió a análisis de patrones recurrentes del rival, detección de vulnerabilidades en determinados momentos del partido, y segmentación de los datos según las fases del juego. Esta metodología permitió identificar los momentos ideales para aplicar presión, recuperar el balón y lanzar ataques rápidos, aprovechando al máximo las características físicas y tácticas del equipo. *(Otto-Friedrich-Universität Bamberg, 2022).*

4.8.3. Resultados obtenidos y análisis del impacto

El impacto del análisis de datos en el Leicester City fue directo y contundente. En términos tácticos, el equipo se convirtió en una de las formaciones más eficientes de la liga, destacando por su solidez defensiva, su capacidad de contraataque y su aprovechamiento de errores del rival. Todo esto fue producto de una planificación táctica basada en los datos analizados, que permitió identificar y explotar ventajas en tiempo real.

En cuanto al rendimiento físico, el uso de sensores permitió individualizar las cargas de entrenamiento, prevenir lesiones y maximizar la disponibilidad de los jugadores clave. La plantilla se mantuvo notablemente saludable durante toda la temporada, lo que permitió una continuidad táctica muy poco frecuente en equipos con menor profundidad de banquillo.

Una de las decisiones estratégicas más exitosas fue la incorporación de jugadores mediante análisis de datos, como Riyad Mahrez y N’Golo Kanté, quienes fueron identificados por tener un rendimiento subestimado en sus anteriores equipos, pero que estadísticamente encajaban con el estilo de juego deseado por el cuerpo técnico. Ambos futbolistas resultaron determinantes en la conquista del título.

Finalmente, el club obtuvo no solo el campeonato, sino también un incremento extraordinario de valor de mercado, visibilidad global y retorno económico, validando el análisis de datos como un camino viable hacia el éxito deportivo sostenible. (*B EYE | Data, AI & EPM Consulting, 2023*)

4.8.4. Lecciones aprendidas y aplicabilidad a otros entornos semiprofesionales

El caso del Leicester City proporciona aprendizajes muy valiosos que pueden ser adaptados a contextos semiprofesionales con los recursos adecuados. En primer lugar, demuestra que el análisis de datos no es exclusivo de clubes con presupuestos millonarios, sino que puede ser implementado de forma inteligente y estratégica para maximizar el rendimiento incluso en equipos con recursos limitados.

Una de las lecciones más importantes es que, incluso con herramientas accesibles (como grabación en vídeo desde un punto elevado y software de análisis básico), se pueden generar patrones de comportamiento colectivo e individual útiles para la toma de decisiones tácticas. Asimismo, el uso de datos para identificar jugadores con alto rendimiento potencial a bajo coste es una estrategia que puede ser extremadamente útil en categorías donde el scouting tradicional es limitado.

Por otra parte, la planificación física basada en seguimiento de cargas, aunque en Leicester se hizo con tecnología avanzada, puede replicarse a menor escala mediante registros manuales, aplicaciones móviles de seguimiento y reportes subjetivos bien estructurados.

5. Metodología

5.1 Diagrama de Gantt

Para coordinar y administrar de manera eficaz el desarrollo del Trabajo de Fin de Máster, hemos elaborado un cronograma detallado que se extiende desde el 1 de marzo hasta el 11 de julio de 2025, fecha en que se defenderá el proyecto. Este cronograma se muestra a través de un diagrama de Gantt, una herramienta que nos simplifica la determinación del tiempo estimado de cada tarea, sus fechas de inicio y finalización, y las conexiones entre las tareas. Además de detectar posibles desvíos respecto a los plazos establecidos y a una toma de decisiones más eficaz para alterar el plan de trabajo si fuera necesario.

La planificación se ha estructurado en 8 fases principales: planificación y definición de proyecto, investigación, diseño del sistema y enfoque técnico, desarrollo y pruebas, evaluación y validación, redacción del documento, revisión final y preparación de la defensa. Cada una de estas etapas se ha dividido en subtareas específicas, asignándoles una duración estimada en función de su complejidad y dependencia con otras actividades.

Este enfoque no solo nos permite llevar un seguimiento sistemático del progreso del proyecto, sino que también facilita la detección temprana de posibles desviaciones respecto al plan previsto. El cronograma ha sido diseñado con suficiente flexibilidad para adaptarse a posibles imprevistos y se irá actualizando conforme avance el desarrollo del TFM.

Tabla 1: Fases del proyecto

FASE 01 – Planificación y definición del proyecto
Duración: Semana 1-2 (1–14 mar)
<ul style="list-style-type: none"> ○ Confirmar tema del TFM ○ Establecer objetivos generales y específicos ○ Definir requisitos de las entregas intermedias ○ Definir estructura general del TFM ○ Planificar un cronograma general
FASE 02 – Investigación

Duración: Semanas 3-4 (15–28 mar)

- Recopilación de artículos académicos y recursos técnicos
- Análisis de apps o sistemas similares (scouting de benchmarks)
- Elaboración del formulario para entrenadores

FASE 03 – Diseño del sistema y enfoque técnico

Duración: Semanas 5-7 (29 mar – 18 abr)

- Pruebas piloto, con fragmentos de videos cortos
- Elección de tecnologías (librerías, frameworks, etc.)
- Definición de arquitectura general (backend, frontend, análisis)
- Planificación de desarrollo por módulos

FASE 04 – Desarrollo y pruebas

Duración: Semanas 8–16 (19 abr – 10 jun)

- 4.1 – Relación con el cuerpo técnico
 - Contactar con un cuerpo técnico
 - Identificar las necesidades/requisitos del cuerpo técnico
 - Presentar una propuesta inicial de visualización para su validación y feedback
- 4.2 – Módulo de análisis de vídeo
 - Carga y segmentación de vídeos
 - Extracción de eventos o frames clave
 - Conversión de datos a métricas útiles
- 4.3 – Backend y lógica de datos
 - Modelo de base de datos
 - API para comunicar análisis y visualización
 - Validación con ejemplos
- 4.4 – Interfaz gráfica
 - Maquetación de la vista de resultados
 - Pruebas funcionales
 - Ajustes de diseño y usabilidad

FASE 05 – Evaluación y validación

Duración: Semanas 15-16 (6 – 13 jun)

- Analizar resultados y rendimiento
- Recoger feedback del cuerpo técnico
- Documentar mejoras/líneas futuras

FASE 06 – Redacción del documento TFM

Duración: Semanas: 3 (15 - 18 mar), 6-7 (11–18 abril), 10-16 (3 may - 15 jun)

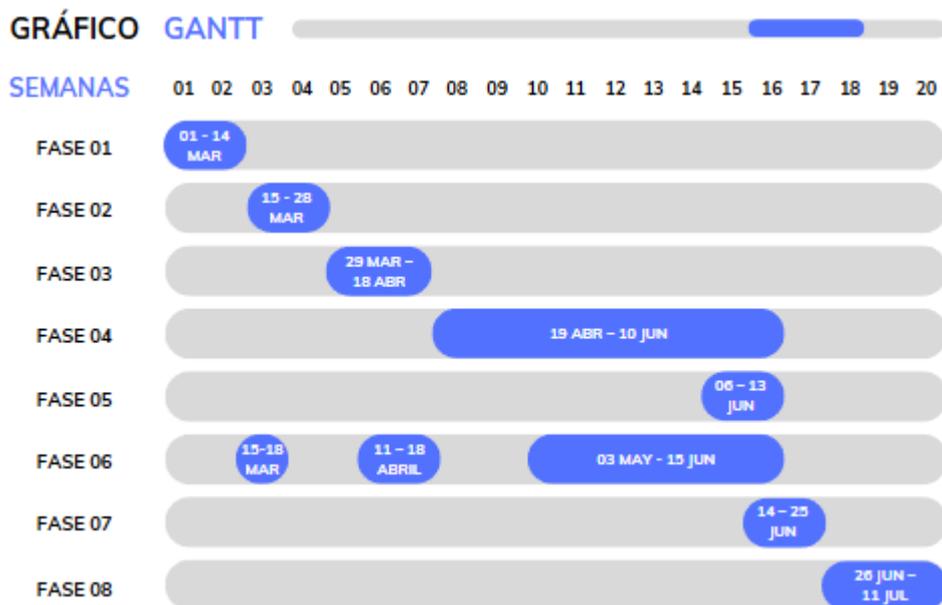
- Redactar marco teórico, metodología, desarrollo, resultados...

<ul style="list-style-type: none"> ○ Incluir capturas de pantalla y esquemas ○ Incluir citas bibliográficas
<p>FASE 07 – Revisión final</p> <p>Duración: Semanas 16–17 (14–25 jun)</p>
<ul style="list-style-type: none"> ○ Revisión del documento final (citas, menciones, índice, gramática...) ○ Revisión con el tutor y correcciones ○ Entrega final el 25 de junio
<p>FASE 08 – Presentación/defensa final</p> <p>Duración: Semanas 18-20 (26 jun – 11 jul)</p>
<ul style="list-style-type: none"> ○ Preparación del resumen y presentación de defensa ○ Ensayo de la defensa

Fuente: Elaboración propia

Teniendo en cuenta las 8 fases del proyecto y sus tiempos, así quedaría el Gantt.

Ilustración 1: Diagrama de Gantt



Fuente: Elaboración propia

5.2 Extracción de datos: Computer Vision aplicada al fútbol

En este apartado se va a explicar el procedimiento utilizado para la obtención automática de datos a partir de contenido audiovisual de partidos de fútbol, empleando técnicas de computer vision. Este procedimiento proporciona la base para la futura generación de métricas posicionales, tácticas y rendimiento colectivo.

Este proceso, se ha realizado utilizando como base una guía pública, publicada por Roboflow en su canal de YouTube y en el que se detallan los pasos necesarios para realizar todo lo relativo al análisis de un vídeo deportivo mediante computer vision.

Este es un pequeño esquema de los pasos que se han seguido a la hora de realizar todo el proceso:

5.2.1 Selección del material audiovisual

El material utilizado para el análisis consiste en grabaciones de partidos completos, adquiridas a través de cámaras fijas ubicadas en distintas posiciones elevadas de las instalaciones. Se le ha otorgado prioridad a la calidad de los videos, fijando una calidad mínima de 720p, al mismo tiempo también hemos querido garantizar un mínimo de fotogramas por segundo, tratando videos con un mínimo de 25 fps, de esta forma garantizamos una captura fluida de movimientos de los elementos del terreno de juego, dado que una frecuencia reducida puede obstaculizar la interpolación de trayectorias y provocar pérdidas importantes en escenas de movimientos acelerados.

Las grabaciones se han llevado a cabo desde un punto de vista alto y enfocado longitudinalmente con gran parte del terreno de juego. Con esta visión queremos potenciar la visibilidad conjunta de una gran parte de los elementos, además de simplificar la identificación sin grandes oclusiones entre elementos. Por lo que, hemos descartado videos demasiado próximos o con inestabilidad de la cámara, debido a que dificultan la homografía del campo.

En situaciones en las que se empleó una cámara panorámica automática (como Veo), se ha seleccionado el modo de grabación integral del terreno, evitando versiones recortadas que siguen la acción con zoom. Siendo esta visión beneficiosa desde el punto de vista de la retransmisión, son insuficientes para un análisis estructurado del dato.

Asimismo, se ha tenido en cuenta el entorno ambiental de las grabaciones, optando por los encuentros disputados con una luz natural notable. Estos elementos influyen directamente en la calidad del análisis de imagen, principalmente en labores de identificación de bordes, colores y segmentación. Respecto al ambiente del terreno, se ha hecho énfasis en que el césped posee una tonalidad homogénea y claramente delimitada, facilitando las tareas de segmentación y aplicación de la transformación homográfica. Del mismo modo, se han descartado grabaciones con cámaras con lentes gran angular, debido a que generan deformaciones geométricas en los extremos del terreno de juego.

Para poder garantizar la solidez del sistema y la utilidad en diversas situaciones, se escogieron diversos partidos que se ajustaban a diferentes jornadas, equipos y condiciones de juego, buscando que esta diversidad posibilite valorar la coherencia del sistema de detección ante modificaciones en los equipos, disposiciones tácticas y condiciones del entorno.

5.2.2 Preprocesamiento del vídeo

La fase de preprocesamiento del video es crucial para asegurar la calidad y consistencia de los datos de entrada antes de implementar algoritmos de identificación y tracking. Se prepara el contenido visual para que sea compatible con los modelos de computer vision, minimizando errores causados por condiciones no ideales del video original.

Cada video utilizado se fragmentó en una serie de fotogramas (imágenes estáticas), lo que ha facilitado el uso de técnicas de detección de imagen a imagen. Para realizar este proceso se ha utilizado la librería OpenCV en Python, utilizando el objeto `cv2.VideoCapture()`, para producir un fotograma por cada momento del video según el framerate original. Esta extracción también

ha posibilitado la supervisión granular del tiempo de análisis, permitiendo la sincronización posterior con los eventos relevantes del partido.

Los fotogramas obtenidos han sido ajustados a una resolución estándar acorde al modelo YOLOv8, siendo el objetivo del redimensionamiento el uniformar el tamaño de la entrada de imágenes. Igualmente, todos los fotogramas son convertidos al mismo formato de color (RGB) y a un formato de archivo (png), asegurando la integridad visual durante el proceso de procesamiento.

A continuación, se normalizan y estructuran las detecciones válidas en un formato de datos que simplifique su utilización futura en etapas de monitoreo y análisis de posición. Por esta razón, se registran tanto las coordenadas como las dimensiones de cada objeto detectado, sumado al número de fotograma correspondiente, generando una base de datos temporal organizada en estructuras como listas de Python o arrays de NumPy.

5.2.3 Detección de jugadores y balón

El primero de los pasos dentro del flujo de trabajo a llevar a cabo para la realización de análisis dentro de un entorno deportivo sería la detección de los distintos intervinientes dentro de un evento deportivo, en nuestro caso dentro de un partido de fútbol. Esta etapa del flujo de trabajo constituye una de las más importantes dentro de todo el proceso, ya que todos los datos relativos a geolocalización de los jugadores que podamos recolectar vendrán determinados por la precisión que obtengamos a la hora de inferir con el modelo. Estos datos serán los que nos permitirán realizar el resto de las etapas.

En nuestro caso, los elementos que buscaremos detectar serán por un lado los jugadores, por otro lado, el balón y por último los árbitros. El motivo de la detección de los árbitros puede parecer a priori no tener sentido para el propósito que se persigue con este proyecto, pero es importante destacar que, los árbitros pertenecen a uno de los grupos intervinientes dentro de un partido de fútbol, y si no se buscara detectarlos activamente y separarlos del resto de

elementos (jugadores y balón) podría causar errores y distorsiones debido a que podrían ser detectados como jugadores de uno de los equipos.

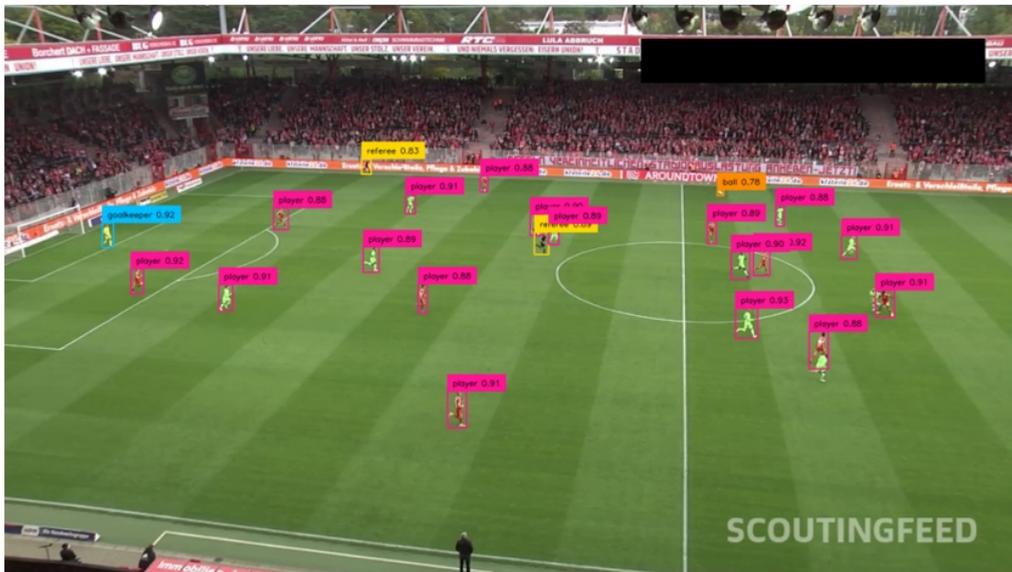
5.2.3.1. Modelo de detección utilizado (e.g. YOLOv8)

Para empezar a hablar de la detección de elementos claves dentro de nuestro entorno de trabajo, debemos de hablar en primer lugar de la herramienta que utilizaremos para ello. En este caso, el modelo de computer vision que utilizaremos para realizar la tarea de detectar a los jugadores, balón y árbitros será YOLO, concretamente la versión 8.

Han sido varias las características que nos han hecho decantarnos por la utilización de este modelo, como sería por un lado la gran cantidad de información que hemos podido encontrar para la correcta utilización e inferencia de este modelo, así como la precisión y velocidad de procesamiento. Pero si hay una característica importante en un entorno deportivo, concretamente de deportes colectivos, es la capacidad del modelo de poder realizar detecciones simultáneas dentro de un mismo frame, ya que, en nuestro caso, el modelo debe de poder manejar hasta 22 detecciones de jugadores a la vez, así como una para el balón y al menos 3 para los árbitros.

Antes de hablar de la propia detección en sí, debemos de hablar del entrenamiento del modelo, ya que, aunque el modelo utilizado ha sido YOLO v8, nos hemos apoyado de la plataforma Roboflow para utilizar un modelo que ellos ya tienen pre entrenado para la detección de los distintos elementos claves dentro de un entorno futbolístico. Este modelo pre entrenado se llama *football_players_detection* y es de uso público, únicamente se necesitaría una clave de API para poder utilizarlo, que es algo que se te proporciona desde la propia plataforma cuando te registras.

Ilustración 2: Detección de objetos



Fuente: Modificado de ScoutingFeed

Por otro lado, aunque finalmente nos decantamos por utilizar este modelo pre entrenado, también se planteó la idea de entrenar directamente un modelo, con una galería de imágenes ya anotadas, que se podría encontrar en la propia plataforma de Roboflow de forma pública. Finalmente se decidió utilizar el modelo pre entrenado debido principalmente a los buenos resultados y la robustez en las detecciones que este obtenía.

Una vez le proporcionamos al modelo un vídeo y lo ejecutamos, se comienza a detectar frame por frame cada una de las clases reconocidas en el modelo, que en nuestro caso son 4 (jugador, portero, árbitro y balón). En el caso de la diferenciación de porteros, árbitros, y jugadores, esta se realiza de una forma muy simple, y es que el modelo ya estaba entrenado previamente separando por clases, y esto se hacía principalmente por el color del uniforme respecto al resto de jugadores. En el caso de los porteros, más adelante definiremos una función para detectarlos siempre, pero, además, los porteros además de llevar un uniforme se sitúan en una zona determinada de la imagen.

El resultado de esta detección se visualizará en forma de *bounding box* que nos indicaría dónde está detectando al jugador y en qué clase se le ha clasificado.

Finalmente, por cuestiones estéticas, podemos eliminar la etiqueta de la clase de detección de la que se trata, y sustituir la *bounding box* por una elipse, de forma que la imagen queda mucho más limpia cuando la visualizamos.

Cabe destacar, que a pesar de que el modelo es bastante preciso, no siempre detecta el 100% de los objetos dentro del vídeo, y siempre hay frames en los que puede ser que alguno de los jugadores o incluso el balón no se detecte. Este es un proceso que más adelante del proceso solucionaremos mediante la interpolación de las posiciones faltantes.

Por el contrario, puede suceder que se registren detecciones que no correspondan con la realidad o con alguno de los elementos que se pretende detectar, es por ello que se establece un umbral de confianza del 30%, lo que permite que no seamos excesivamente restrictivos y que posibles detecciones no se utilicen, pero evitaríamos recibir falsos positivos.

Por último, los datos obtenidos de las detecciones en este punto se guardaron de dos maneras distintas; por un lado, se utilizó una lista vacía que se fue rellenando con la información de las distintas detecciones. La otra forma de guardar el progreso, de una forma más visual y que nos permitía la evaluación de errores era guardar todo el vídeo con las anotaciones realizadas, para ello se utilizó la librería de *supervision*, de la que ya hablamos en el marco teórico.

5.2.3.2. Técnicas de filtrado y asignación de identificadores (*player_id*, *team_id*)

Después de efectuar las detecciones en cada frame, se realiza un proceso de filtrado con la finalidad de eliminar posibles fallos originados en el modelo de detección. Primero, se suprimen las detecciones duplicadas utilizando métodos de *non-maximum suppression* (NMS), garantizando que para cada objeto solo se mantenga la detección con un nivel de confianza superior. Además, se establecen límites de confianza establecidos con antelación que facilitan la eliminación de detecciones con alta posibilidad de ser falsos positivos.

Después de este primer filtrado, se lleva a cabo la atribución de identificadores únicos para los jugadores (`player_id`) y los equipos (`team_id`). Esta tarea se lleva a cabo teniendo en cuenta tanto los datos recabados durante la identificación (como la categorización de las clases: jugador, guardameta, árbitro, balón) como los datos de ubicación y seguimiento suministrados por el algoritmo de seguimiento. Así, se logra establecer una relación consistente y estable entre los objetos identificados a lo largo del tiempo, lo que facilita la creación de rutas individuales para cada jugador y el balón, además de vincular cada acción de juego con los participantes implicados.

5.2.4. Algoritmo de tracking y generación de trayectorias

Una vez identificados los elementos significativos en cada fotograma (jugadores, árbitros y balón), es imprescindible conectar estas detecciones a través del tiempo para generar trayectorias. A este procedimiento se le llama tracking y facilita el análisis del movimiento de cada uno de los elementos en el transcurso del partido. El propósito del tracking es asignar un identificador constante a cada objeto identificado (`player_id` o `ball_id`), de tal forma que se pueda reconstruir su trayectoria frame a frame, incluso en los casos de solapamientos visuales o movimientos rápidos. Para conseguirlo se utilizan algoritmos de seguimiento multi-objeto (MOT) que fusionan datos espaciales y temporales.

Para este proyecto, el seguimiento se ha llevado a cabo a través de un algoritmo llamado *ByteTrack* que es un modelo creado específicamente para usarse con YOLO, aunque también se valoró la utilización del modelo DeepSort. Las principales diferencias entre ambos modelos, es que DeepSort utiliza redes neuronales adicionales (En su caso CNN) lo que lo hace un modelo más pesado computacionalmente. Además, DeepSort utiliza dicho modelo adicional para realizar asociaciones por apariencia, lo que lo puede hacer interesante si se detecta la pérdida del tracking de forma usual dentro de los distintos jugadores.

Finalmente, nos decantamos por utilizar ByteTrack debido por un lado a su alta eficiencia y rapidez a la hora de realizar el tracking. Por otro lado, nos ofrecía la posibilidad de utilizarlo sin necesidad de entrenar a otro modelo para que este funcionara correctamente.

Esta segunda etapa dentro del proceso es clave para la posterior detección de eventos, ya que son las posiciones que se van guardando, tanto para el balón como para los jugadores, las que nos permitirán detectar eventos en los que hay disputa por el balón entre varios jugadores, o poder identificar un saque de banda o de córner por la posición desde donde sale el balón.

5.2.4.1. Seguimiento de jugadores

Para trackear a los jugadores a través del vídeo, se aplica un algoritmo de seguimiento multiobjetivo fundamentado en la librería *supervision* y el seguimiento *ByteTrack*, reconocido por su exactitud en situaciones con oclusiones parciales y movimientos acelerados.

El proceso inicia dividiendo las detecciones en dos grupos: el balón y los jugadores. Las detecciones de los jugadores son sometidas a un proceso de supresión de *non-maximum suppression* (no-máxima) con un umbral de 0.5, lo que facilita la eliminación de duplicados y la conservación únicamente de las detecciones más confiables por objeto.

A continuación, estas detecciones mejoradas se transfieren al algoritmo de tracking, en este caso *ByteTrack*. El algoritmo realiza un emparejamiento entre las detecciones detectadas a lo largo de los distintos frames. Para realizar este emparejamiento se utiliza únicamente la información geométrica obtenida en el proceso de detección (al contrario que *DeepSort*, que como hemos comentado antes, en este punto utilizaría referencias visuales, no geométricas). En este momento, *ByteTrack* utilizaría el solapamiento espacial para determinar si para un jugador detectado sus bounding boxes se van solapando a lo largo de los frames (esto es lo correspondiente a un movimiento natural de un objeto moviéndose). Para determinar este solapamiento espacial de las bounding boxes, se calcula un indicador llamado *Intersection Over Union (IoU)*, el cual determina el grado de solapamiento que hay entre la bounding box de una detección en un frame y en el siguiente. De esta forma evalúa si se trata del mismo objeto o de un objeto distinto.

La fórmula utilizada para realizar este cálculo sería la siguiente:

Ecuación 1: Calculo de Intersection Over Union (IoU)

$$\text{IoU} = \frac{|B_{\text{pred}} \cap B_{\text{gt}}|}{|B_{\text{pred}} \cup B_{\text{gt}}|}$$

Fuente: Elaboración propia

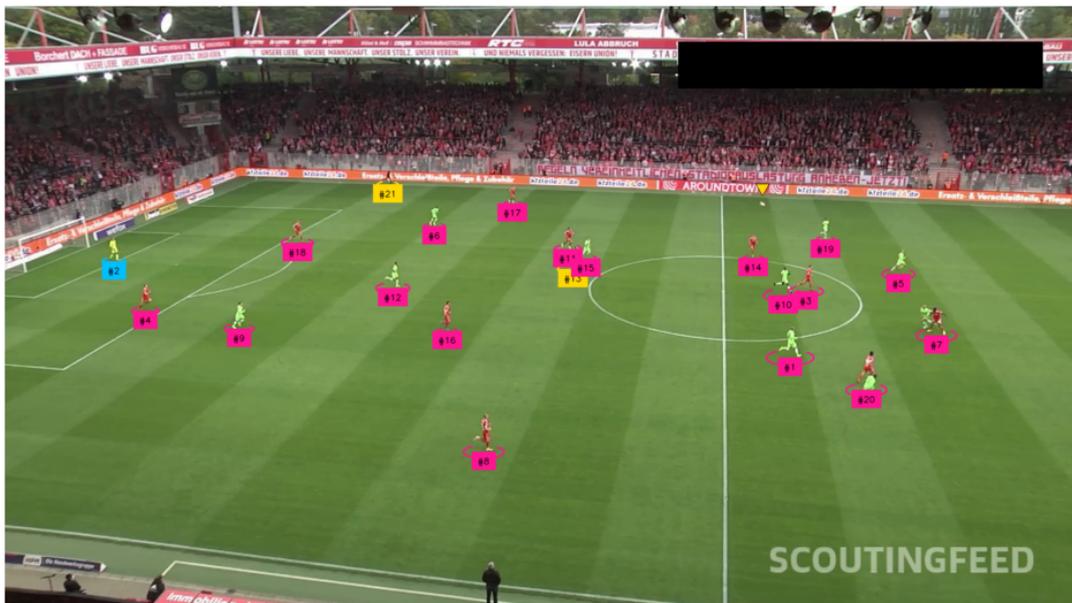
Donde B_{pred} se corresponderá con la detección predicha (la nueva detección), B_{gt} sería la detección de referencia (el frame anterior), donde \cap correspondería a la intersección entre ambas, y \cup correspondería a la unión de entre ambas detecciones. Todo ello bajo $|\cdot|$ que correspondería el área de región.

El resultado de esta ecuación será un número entre 0 y 1 siendo 0 objetos distintos y 1 el mismo objeto idéntico. Usualmente, se consideran que resultados del entorno de entre 0.4 y 0.5 ya se consideran el mismo elemento. Además, la innovación de este modelo radica en que combina detecciones de alta confianza con detecciones de baja confianza para utilizar detecciones de jugadores parcialmente ocultos que otros modelos descartarían. Estas detecciones de baja confianza se utilizarían en una segunda fase para no entren en conflicto con las de alta confianza.

Los resultados se guardan en una estructura llamada *tracked_detections_by_frame*, en la que cada entrada se relaciona con las detecciones con seguimiento vinculadas a un frame específico. En la representación gráfica sobre el video, añadimos elipses para representar a cada jugador e insertamos etiquetas numéricas (#ID) para poder facilitar su seguimiento visual. De esta manera, podremos en etapas posteriores construir un data set con dichos datos almacenados para poder proyectar sobre un campo virtual los movimientos de los jugadores.

El resultado del tracking sería el siguiente:

Ilustración 3: Asignar #ID a las detecciones



Fuente: Modificado de ScoutingFeed

5.2.4.2. Seguimiento del balón

El seguimiento del balón se trata de manera distinta dado a las características que posee (tamaño pequeño y trayectorias irregulares). Esto se ve reflejado en jugadas como serían tiros que resultan bloqueados por un jugador, despejes o rebotes. Esto hace que los modelos tradicionales no resulten siempre eficaces a la hora de realizar un seguimiento del balón.

La identificación de las detecciones del balón se realiza mediante la clase `ball_id`, que se definió previamente cuando se realizó la detección de objetos. Además, se le aplica una ampliación artificial de su bounding box a través de la función `sv.pad_boxes()`, con el objetivo de simplificar su visualización y posterior procesamiento algorítmico para la detección y cálculo de eventos .

A diferencia de los jugadores, el balón no se encuentra bajo el seguimiento extendido por medio de *ByteTrack*, sino que mantiene su identificación directa de un frame a otro, ya que utilizando *ByteTrack* sería difícil asignarle un identificador persistente a lo largo de los frames. Sin embargo, su información se guarda en la estructura *ball_detections_by_frame*, lo que facilita la reconstrucción posterior de su trayectoria global.

5.2.5. Clusterización de jugadores

Para poder finalizar el proceso de procesamiento de vídeo respecto a los jugadores, debemos de finalizar separando a los jugadores por equipos, ya que, para nuestros análisis posteriores, es imprescindible identificar qué eventos han sido producidos por jugadores pertenecientes a un equipo concreto (Si lo que pretendemos es realizar un informe para uno de los equipos participantes en el evento).

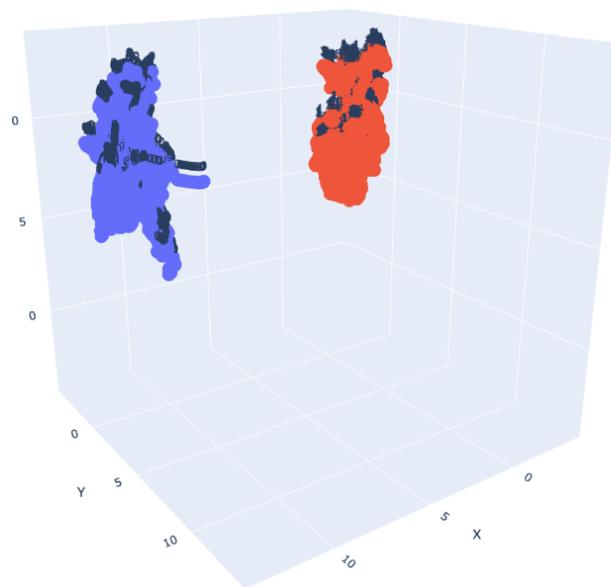
Para ello, primer paso sería obtener las características visuales por las cuáles vamos a dividir las distintas detecciones asociadas a los jugadores. En nuestro caso, utilizamos las bounding boxes para recortar el área del frame que queda dentro de cada una de dichas bounding boxes, dando como resultado que tenemos los recortes de cada uno de los jugadores para cada uno de los frames.

Una vez ya hemos obtenido dichas características, el siguiente paso sería vectorizar dichas características. Esto quiere decir transformar las características de la imagen como sería el color de la camiseta en un valor numérico. Esto lo realizaremos utilizando *SigLip* (Sigmoid Language-Image Pretraining). Se trata de un modelo auto supervisado, cuyo componente visual resulta muy útil para tareas relacionadas con la identificación de elementos visualmente parecidos.

Una vez ya tenemos vectorizadas las características (principalmente el color de la camiseta), utilizaremos un algoritmo de clasificación para clasificar las distintas detecciones en dos grandes grupos diferenciados; equipo 1 y equipo 0. Para realizar esto, utilizaremos KMeans, el

cuál es un algoritmo ampliamente utilizado y que al haber usado SigLip (No sensible a iluminación o a tonos de camiseta similares) nos permite tener una precisión bastante buena. Para indicar que solo habrá dos equipos, indicamos el argumento dentro de la llamada al modelo indicando $n_clusters = 2$.

Ilustración 4: Algoritmo de clasificación



Fuente: Elaboración propia

Posteriormente se realizó un gráfico de puntos para visualizar de forma gráfica la separación de ambos clusters y comprobar que no había fallos en los que se detectaran mal los distintos equipos. La visualización sería la siguiente:

5.2.6. Detección del terreno de juego y proyección de jugadores y balón sobre él.

Una vez que obtuvimos las coordenadas de los jugadores y el balón en cada frame, el siguiente paso es proyectar estas posiciones sobre una representación estandarizada del campo de juego, o lo que sería lo mismo, un campo virtual que represente las mismas medidas que el campo que se encuentra en el vídeo. Este procedimiento de normalización espacial es esencial para asegurar que los datos obtenidos sean equivalentes a los captados sobre la realidad a través de

la cámara, sin importar aspectos como la resolución del vídeo, el ángulo de la cámara o el tamaño del campo durante la grabación.

5.2.6.1 Detección del terreno de juego:

En el caso de la detección del terreno de juego, debemos empezar por detectar los puntos más relevantes y que necesitaremos para el cálculo de las dimensiones reales del campo captado en cámara. Algunos de estos puntos claves a destacar serían las distintas referencias del círculo central del campo, así como la unión entre la línea central del campo y las líneas correspondientes a las bandas. Otros puntos clave serían las esquinas del campo, y las esquinas de las distintas áreas (Tanto el área grande como el área pequeña)

Es por ello, que para hacer esta labor tendremos que un modelo específico para este propósito. En nuestro caso, nos apoyaremos de nuevo en los modelos que nos proporciona un modelo pre entrenado usando YOLO v8 pose, que está especializado para detectar puntos concretos en contextos de posturas del cuerpo humano o de la cara, pero que entrenado de la forma correcta se puede reconvertir para utilizarlo en la detección de puntos dentro de un campo de fútbol. En este caso, desde Roboflow entrenaron a este modelo definiendo 32 puntos de interés dentro de un campo de fútbol.

Una vez se hizo la inferencia, las detecciones de puntos se vieron de la siguiente manera dentro del vídeo (Ya que se anotaron dichos puntos dentro del vídeo):

Ilustración 5: Detección del terreno de juego



Fuente: Modificado de ScoutingFeed

Como se puede apreciar en la imagen, hay puntos que no están situados exactamente donde se debería de encontrar una línea o uno de los vértices de unión entre distintas líneas.

Para solventar este problema, lo que se hizo fue filtrar los *keypoint* detectados, haciendo uso del coeficiente de confianza que nos ofrece el modelo a la hora de realizar las detecciones. Se estableció un umbral de confianza del 30%, con ello, el resultado obtenido fue el siguiente:

Ilustración 6: Detección del campo después de filtrar los keypoints



Fuente: Modificado de ScoutingFeed

Observando de nuevo la imagen, vemos que ya no se encuentran puntos fuera de lugar y por tanto, podemos confirmar que el modelo está haciendo las detecciones correctamente.

5.2.6.2 Proyección del campo y jugadores en un campo virtual.

Para poder hacer la proyección del campo captado en cámara y pasarlo a un plano 2D, tendremos que realizar el cálculo de una homografía, que lo que permite es relacionar dos planos mediante una matriz 3x3.

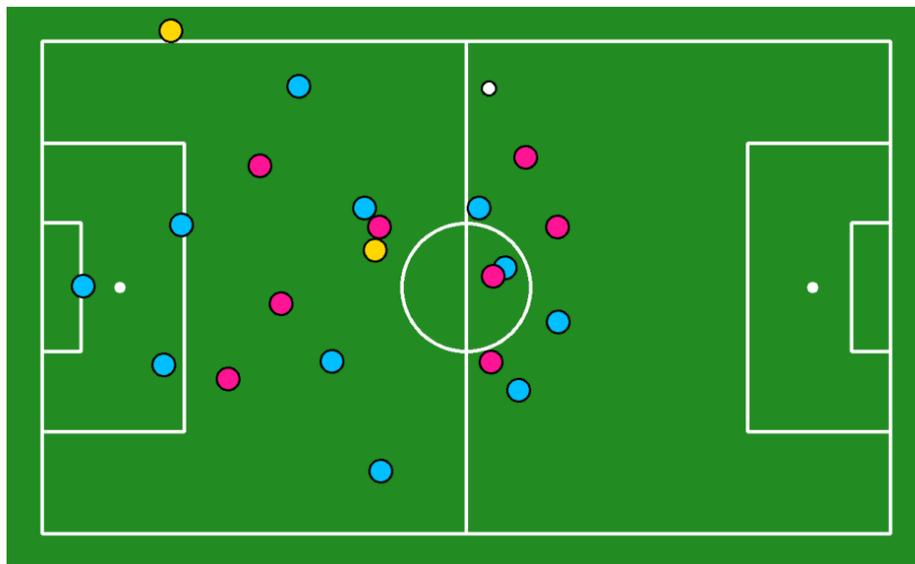
En nuestro caso, nos apoyaremos de la función ViewTransformer, la cual es una función recogida en el repositorio de github público de Roboflow y que nos ha permitido realizar este cálculo. Para ello, se utilizan dos grupos de inputs en forma de puntos, el primero de los grupos se refiere a los puntos que hemos identificado en la imagen original mediante la detección de los keypoints con el modelo anteriormente mencionado. El segundo grupo de puntos, corresponderían a los mismos puntos de referencia, pero esta vez dentro de un campo virtual.

Dichos puntos se establecen con una lista ya predefinida que se proporciona en el mismo repositorio del cuál se extra la función de ViewTransformer.

Finalmente, una vez ya tenemos configurado y proyectado el campo, solo faltaría incluir en ese campo virtual la posición tanto de los jugadores como del balón. Para ello, se convierten las coordenadas detectadas en el vídeo original a un sistema de referencia común, haciendo que podamos dibujar de forma visual con círculos, los distintos jugadores del mismo color que se le asignó a su equipo dentro de nuestro campo virtual. Para realizar esta tarea, se utilizaron funciones que de nuevo se encontraban en el repositorio público de Roboflow. Algunas de estas funciones serían `draw_pitch` (para dibujar el campo) o `draw_point_on_pitch` (para poder dibujar los puntos que representan a los jugadores dentro del campo).

El resultado de estas visualizaciones sería algo similar a lo siguiente:

Ilustración 7: Visualización 2D



Fuente: Elaboración propia

Como vemos en la imagen, los jugadores se separan en dos equipos, y vemos dos puntos amarillos que representarían los árbitros. Por un lado, tendríamos uno dentro del campo (árbitro

de campo), y por otro lado vemos que se detecta y se dibuja el asistente o juez de línea, justo en el borde de la línea de banda (donde se colocan dichos árbitros)

5.2.7. Generación de data sets con la geolocalización de jugadores y balón

Finalmente, para terminar con esta etapa dentro de la extracción de los datos, lo que se propone es almacenar dichos datos dentro de un archivo CSV para su posterior análisis. Estos datos son los que se utilizarán la etapa de procesamiento de datos para hacer la detección de eventos dentro de un partido de fútbol.

Proponemos dividir los datos relativos a la geolocalización de jugadores y balón en dos data sets diferenciados, uno para los jugadores y otro para el balón. El motivo de esto es que, si no los separamos, tendríamos muchas filas en las que habría nulos, ya que las columnas que aplican para el guardado de datos de jugadores no son las mismas que las del balón.

Las variables dentro de nuestro data set de jugadores serían:

Tabla 2: Variables del data set (*players_tracking*)

Frame	Tracker_id	Team_id	X	Y
Número de frame de la detección	ID del jugador	ID del equipo al que pertenece el jugador (0,1)	Coordenadas X dentro del campo	Coordenadas Y dentro del campo

Fuente: *Elaboración propia*

Por otro lado, el data set correspondiente a las posiciones del balón se compondría de las siguientes variables:

Tabla 3: Variables del dataset (*ball_tracking*)

Frame	X	Y
Número de frame de la detección	Coordenadas X dentro del campo	Coordenadas Y dentro del campo

Fuente: Elaboración propia

5.3 Procesamiento de datos: Cálculo de eventos futbolísticos

Tras obtener y normalizar los datos posicionales de jugadores y balón, el siguiente paso es utilizar los datos de posición para convertir dichos datos en eventos dentro del partido. Estos eventos simbolizan acciones significativas del partido (pases, regates, disparos, posesiones...) y nos permiten reconstruir la dinámica del partido desde una perspectiva analítica, facilitando el análisis de comportamientos individuales y colectivos.

Estos eventos son los que luego utilizaremos para realizar los distintos dashboards y paneles que pueden ayudar al equipo en cuestión a entender las fortalezas y debilidades en su juego.

5.3.1. Diseño de reglas y lógica para detección de eventos

La identificación de sucesos futbolísticos basándose en datos posicionales no etiquetados requiere la creación de un conjunto de normas lógicas que faciliten la interpretación del comportamiento del balón y de los jugadores basándose en su desplazamiento espacial y temporal.

Este enfoque se fundamenta en lógica condicional y relaciones espaciales y temporales derivados directamente de las trayectorias previamente adquiridas a través de técnicas de *computer vision*. Esto quiere decir, utilizar la información de posición obtenida en el proceso de *computer vision* para poder detectar cuándo, quién y donde se produce un evento.

Estas reglas persiguen reconocer eventos esenciales como pases, desplazamientos, tiros o recuperaciones. Cada clase de evento se configura como un modelo particular en la dinámica

del balón y su vínculo con los jugadores, considerando factores como la velocidad de movimiento, la dirección y la proximidad relativa respecto al balón y otros jugadores. El procedimiento se estructura en tres pasos fundamentales.

El primero de esos pasos, es el análisis de la trayectoria del balón. Se analizan las modificaciones en la dirección, la aceleración o la interrupción de su movimiento. Estos indicios facilitan la deducción del comienzo o fin de un evento.

El siguiente paso es el de la proximidad con los jugadores. Se determina si un jugador tiene el balón basándose en su proximidad y coordinación de movimiento. Esta condición es fundamental para asignarle acciones personales (como la conducción o el pase).

El último de los pasos es el de las transiciones y cambios de posesión. Se registran cambios en el control entre jugadores o equipos, lo que facilita el registro de eventos como pases, robos, asistencias o posesiones perdidas.

En este sentido, podríamos encontrar tres niveles de detección de eventos. Por un lado, los eventos más básicos se detectarían en el primer paso, es decir, viendo únicamente las trayectorias del balón y su posición respecto a las líneas del campo. Eventos de este tipo serían saques de esquina, saques de puerta, córneres...etc.

El segundo nivel de detección vendría dado ya por la intervención de la información de posición tanto de los jugadores como del balón, este nivel de detección sería únicamente de toques, que se considerarían cuando el balón se aproxima a un jugador y posteriormente sale del mismo jugador en otra dirección.

El tercer nivel de detección vendría dado por la intervención de más de un jugador y más de un toque, es aquí donde tendríamos el grueso de eventos relevantes como serían los pases, robos, intercepciones, tiros...etc.

El último nivel sería la detección de eventos de difícil automatización, principalmente de eventos que implican una detención del juego, como sería el caso de las faltas, cambios o fueras

de juego. En el caso de estos eventos, automatizarlos podría suponer el riesgo de perder precisión.

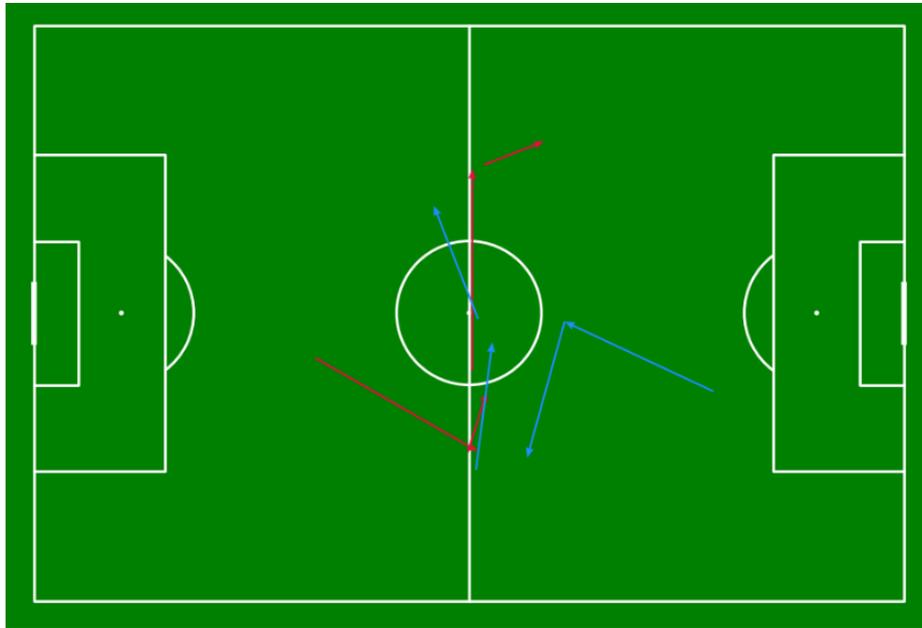
5.3.1.1. Pases

Un pase se entiende como un traspaso libre del balón entre dos jugadores del mismo equipo. Se definirá como pase cuando el balón se desplace del jugador en posesión (emisor) y posteriormente se acerque a otro jugador del mismo equipo (receptor), siempre y cuando no haya ninguna interrupción en la posesión por un jugador del equipo rival. En este sentido, puede haber rivales que entren dentro de la trayectoria, pero si observamos que la trayectoria y movimiento del balón no cambia a pesar de haber un rival entre medio se considerara pase completado.

Por otro lado, sería interesante realizar la división entre pases completados y pases intentados. Un pase intentado se producirá cuando un jugador tenga el balón y le dé una trayectoria con un sentido directo o indirecto a uno de sus compañeros de equipo.

Se anotan las coordenadas del comienzo y finalización del pase, así como su trayectoria. Los identificadores del jugador (*player_id*) emisor y receptor, además del intervalo temporal del evento. Todo ello nos permitirá hacer análisis sobre qué jugadores realizan qué tipo de jugadas, en que zonas del campo...etc.

Ilustración 8: Visualización de los pases



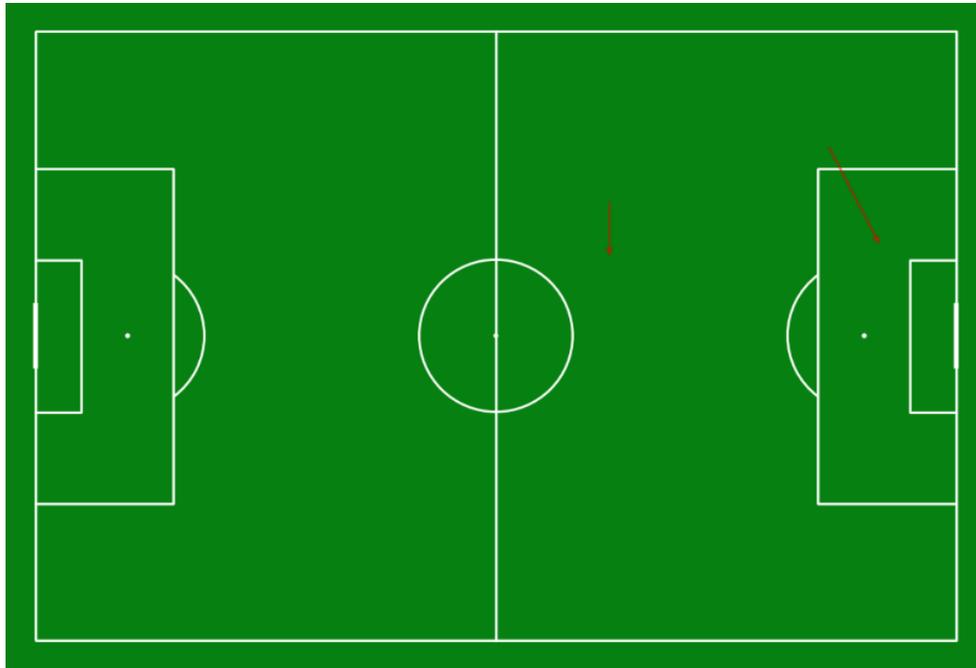
Fuente: Elaboración propia

5.3.1.2 Conducciones

Se produce una conducción cuando un jugador conserva el dominio del balón mientras se mueve por el campo. Esta conducción se detectará cuando el balón se mantenga a una distancia reducida del jugador (por debajo del umbral establecido), el jugador y el balón se desplacen en el mismo sentido durante un número mínimo de fotogramas consecutivos, y no exista otro jugador cercano que tenga la capacidad de interferir con la posesión.

Este evento se representará como un segmento con comienzo, final, duración y distancia recorrida.

Ilustración 9: Visualización de las conducciones



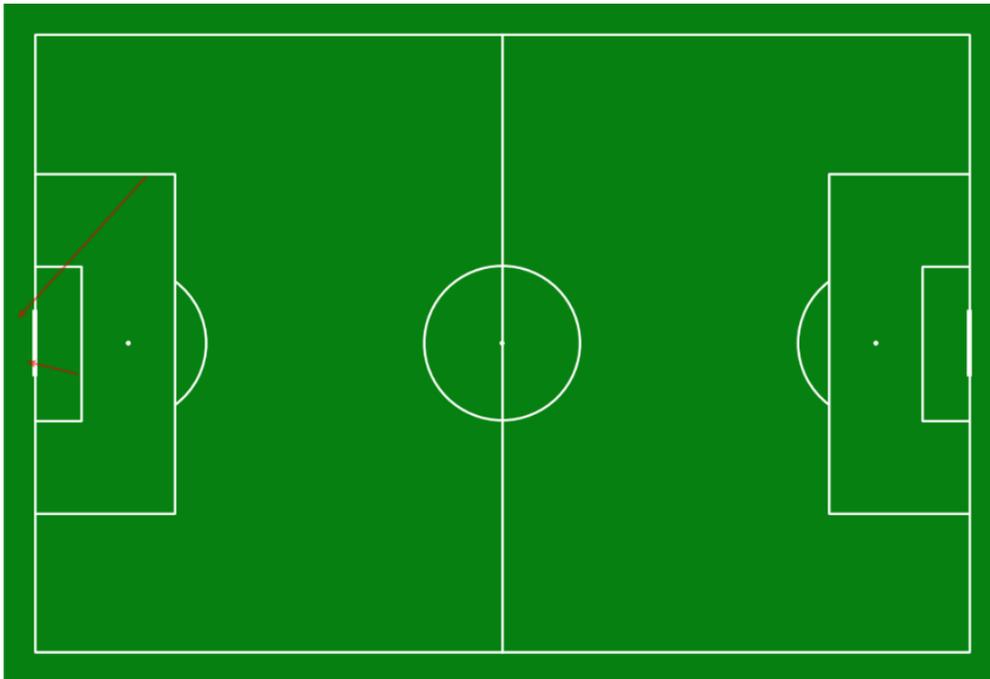
Fuente: Elaboración propia

5.3.1.3 Tiros

Un tiro se define como un movimiento veloz del balón hacia la portería rival, con el propósito de finalizar la jugada. Este evento se detectará cuando el balón pase de la posesión de un jugador a sobrepasar la línea del campo donde se encuentra la portería rival (*eje_y*), no haya presencia de ningún jugador en su recorrido inmediato, el evento ocurra en la parte del campo rival.

Este evento se representará con las coordenadas de la posición de inicio y final, id del jugador y distancia recorrida. Es posible añadir una inferencia acerca del desenlace del tiro (gol, parada, fuera...), basándose en su trayectoria final.

Ilustración 10: Visualización de los tiros



Fuente: Elaboración propia

5.3.1.4 Robos, recuperaciones e intercepciones

En este punto, es importante definir en cada momento quién es tanto el equipo como el jugador que tiene el control del balón, ya que en el momento que el control del balón cambie de equipo, sabremos que ha ocurrido uno de los escenarios de cambio de posesión.

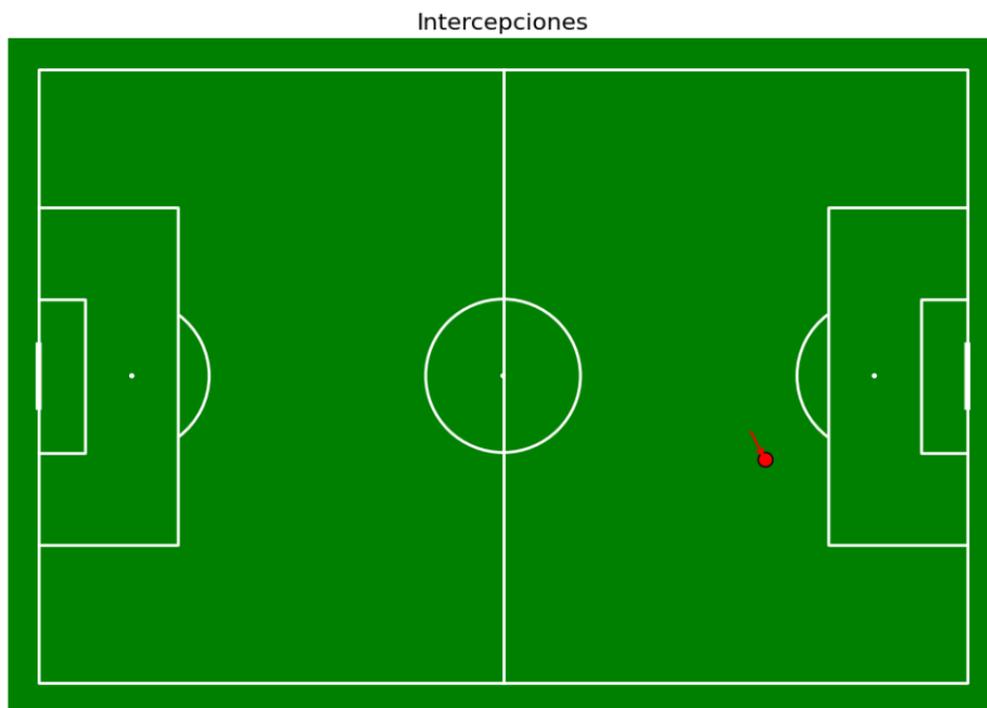
En este caso, los robos y recuperaciones del control de posesión pueden deberse a diferentes eventos, ya que puede ser que se le arrebatase el balón a un rival cuando lo tiene, o que se intercepte un pase que el jugador rival intentó realizar. Es por ello, que lo primero que deberemos de hacer es trazar un círculo de una distancia determinada alrededor de cada uno de los jugadores dentro del campo. Este espacio será su zona de control, cuando el balón este dentro de su zona y no haya otro jugador del equipo contrario dentro de su zona de control.

Cuando haya dos jugadores en la misma zona de control y esté el balón también dentro de la zona, se considerará que estará en disputa.

Se considerará robo o recuperación cuando un jugador tenga en su posesión el balón y otro jugador entre en su zona y acabe consiguiendo pasarlo a otro compañero de su equipo o produzca un evento de conducción.

Por otro lado, una intercepción se producirá cuando un jugador consiga tomar el control de un balón que venía de producir un evento de pase intentado.

Ilustración 11: Visualización de las intercepciones



Fuente: Elaboración propia

5.3.2 Asociación temporal y espacial de eventos

Después de identificar y definir los eventos del partido a través de reglar lógicas, es crucial vincular estos eventos de forma precisa a su contexto temporal y espacial. Esta fase asegura que cada acción detectada no solo sea entendida de manera abstracta, sino que también pueda situarse de manera precisa en el transcurso del partido y en el campo de juego, lo que facilita la realización de análisis tácticos y visualizaciones posteriores.

Cada evento identificado se relaciona con una ventana de tiempo concreta dentro del video, evidenciada en cantidad de fotogramas o marcas temporales (*timestamp*). Por ello, se guardan para cada acción, el fotograma inicia, que se refiere al momento donde se identifica el inicio de la acción, el fotograma final, momento en el que se percibe como ha finalizado el evento, y la duración de la acción en cantidad de fotogramas o segundos transcurridos. Esta información es fundamental para coordinar las acciones con las secuencias del partido, facilitando su revisión manual o visualización automática.

A Al mismo tiempo, cada evento se vincula a un grupo de coordenadas normalizadas, vinculadas al terreno de juego virtual (2D). Dependiendo del tipo de acción, se registrará una coordenada única (en caso de robo de posesión, falta...), dos coordenadas de comienzo y final de la acción (en caso de pases, conducciones o disparos), trayectorias completas si se pretende simular el movimiento del jugador o del balón durante el partido. La precisión a la hora de la localización facilita el trabajo a la hora de crear mapas de calor, representaciones vectoriales o análisis de áreas detalladas del terreno de juego (áreas de ataque, duelos...).

Además, la relación temporal y espacial se enriquece con otros datos relevantes ya obtenidos anteriormente, como los identificadores de jugadores (*player_id*) y de equipo (*team_id*), situación del encuentro en ese instante (minuto de juego, resultado, fase ofensiva/defensiva...), lugar relativo de otros jugadores en el momento del evento (útil para analizar la presión o los espacios libres en el terreno de juego...). Esta integración transforma a cada evento en una unidad de análisis llena de contexto, que puede emplearse tanto para análisis personalizados como para modelado grupal.

5.3.3 Creación del dataset de eventos

Una vez hemos realizado la detección de eventos, debemos de recopilar todos estos datos en un solo dataset que recoja todos los eventos que se han producido a lo largo del partido. Es crucial que tales eventos, puedan ser asociados a tres elementos fundamentales; Por un lado, deben poder asociarse a un jugador, quien realiza la acción, opcionalmente, habrá un segundo jugador que interviene, como podría ser el que recibe el pase. Otro de los elementos asociativos, sería el frame en el que se produjo la acción y el frame en el que terminó la acción. Con esta información, podremos calcular el espacio temporal en el que se produjo, pudiendo dividir así los eventos por fases del partido.

El último elemento y más importante, sería la ubicación en la que se produce el evento, identificando las coordenadas X e Y en la que empezó el evento, y opcionalmente, las coordenadas X e Y en las que acaba, esto nos serviría no solo para identificar dónde se produjo el evento, sino que también lo utilizaremos para pintar las trayectorias de los eventos dentro del campo.

Es por ello por lo que las variables que tendremos dentro de nuestro dataset serán:

- **Equipo:** Equipo que realizó la acción
- **Tipo de evento:** Tipo de evento concreto del que se trata (Pase, tiro, intercepción...)
- **Frame_inicio:** Frame en el que se inició la acción
- **Frame_final:** Frame en el que acabó el evento
- **Jugador_inicio:** ID del jugador que realizó el evento
- **Jugador_fin:** ID del jugador que acabó el evento (Principalmente utilizado para los pases entre jugadores)
- **X_inicio:** Coordenadas X donde se inició el evento
- **Y_inicio:** Coordenadas Y donde inició la acción
- **X_final:** Coordenadas X donde finalizó la acción (Para acciones que tienen una trayectoria como serían pases o tiros)
- **Y_final:** Coordenadas Y donde finalizó la acción

Finalmente, con propósito de poder realizar un análisis completo de un partido completo, con diferentes tipos de eventos, hemos procedido a utilizar una base de datos ya con la estructura anteriormente mencionada. Esta base de datos se encuentra en la plataforma Kaggle a través del siguiente enlace:

<https://www.kaggle.com/datasets/ekrembayar/metrica-sports-football-tracking-data>

Esta base de datos pertenece a la plataforma Metrica Sports y es de uso público. Además de los datos de los distintos eventos para un partido de fútbol, también cuenta con los datos de tracking, por lo que podremos juntarlo todo en visualizaciones en apartados futuros.

5.3.4 Validación interna y consistencia de los eventos generados:

Para asegurar que los eventos que se están generando se corresponden con la realidad, se procedió a realizar una validación interna de los mismos.

Por un lado, se hizo una verificación temporal de los distintos eventos, es decir, que se trató de verificar que los eventos se encuentran secuenciados de forma correcta, esto haría que no pudiera haber la posibilidad de que para el mismo frame hubiera más de un evento, evitando así duplicidades erróneas. Esto lo podemos visualizar directamente en el dataset de eventos que extraemos del procesamiento del vídeo, ya que el número de frame debe ser único.

Otra de las validaciones que se han realizado ha sido la consistencia espacial de los eventos detectados. Para ello hemos utilizado tanto los datos relativos a los eventos detectados, así como los datos relativos tanto al tracking de los jugadores como del balón. Esto nos permite asegurarnos que los eventos detectados y sobre todo los jugadores involucrados en ellos, tienen unas coordenadas espaciales iguales a las que se detectaron en el tracking. Esto es lógico, ya que como hemos dicho en apartados anteriores, son las propias coordenadas obtenidas en el tracking las que se han utilizado para posteriormente hacer la detección de los eventos.

Por último, una de las prácticas que nos ha ayudado mucho a poder verificar los eventos que estábamos detectando, ha sido la selección de vídeos. Mediante la selección de vídeos de una extensión muy limitada (de alrededor de 30 segundos) en los que teníamos la constancia que se producían eventos de la tipología que queríamos detectar, hemos podido comprobar en el mismo momento de haber hecho la generación de eventos, si la cantidad y la tipología de eventos detectados se correspondían con lo que habíamos visto en el vídeo y que por tanto esperábamos.

Esta combinación de diferentes metodologías de comprobación y verificación de la exactitud de los eventos que estábamos detectando, hemos podido ir detectando los distintos errores que íbamos detectando, y así poder ir ajustando la lógica de detección en función de los mismos.

5.4 Trabajo con cuerpo técnico

5.4.1 Primer contacto con el cuerpo técnico

Como parte del proceso de desarrollo de este proyecto, realizamos una primera reunión con el equipo técnico de un equipo de fútbol amateur, con la finalidad de conocer sus necesidades y gustos respecto al dashboard de análisis. La implicación directa de los entrenadores y analistas nos parece crucial para garantizar que las métricas analizadas sean las correctas y que al mismo tiempo la forma de mostrarlas lo sean también, ajustándonos de manera exacta a las necesidades reales.

En esta reunión, el equipo técnico ha especificado los elementos clave que buscan observar después de concluir un partido, incluyendo métricas tanto individuales como grupales, estadísticas de posesión, mapas de calor, rendimiento físico de los jugadores, áreas de recuperación y pérdida de balón, y sucesos significativos durante el partido. Toda esta información facilitará la creación de un panel de control intuitivo, que simplifique la toma de decisiones técnicas y el estudio subsiguiente de los partidos.

Las especificaciones que hemos obtenido en esta reunión se utilizarán como base para el diseño funcional de la interfaz de usuario y la creación de los módulos de análisis que conformarán el dashboard final.

5.4.2 Cálculo de estadísticas individuales de rendimiento físico y deportivo:

Una vez que el cuerpo técnico nos indicó las necesidades que tenían en cuanto a qué métricas querían observar, se determinó la diferencia entre las estadísticas individuales relativas al aspecto del rendimiento físico dentro del terreno de juego y las que son relativas al desempeño en el ámbito deportivo y de acciones individuales.

Si hablamos del aspecto físico, las métricas que se nos solicitó calcular son relativas principalmente a las distancias recorridas, y a las tipologías de esfuerzo que se estaban aplicando y en qué medida. Esto se hizo, ya que para el cuerpo técnico era importante conocer con qué intensidad trabajaban sus jugadores y cuánto tiempo podían sostener esta intensidad. Las métricas calculadas fueron:

- **Distancia recorrida:** Para ello nos servimos de las posiciones del tracking, y sumamos la distancia en línea recta entre los distintos puntos en los que se detectó ese jugador
- **Distancia con alta intensidad:** Para ello, previamente debemos definir la velocidad entre distintos puntos, y si supera un umbral de 20 km/h entraría dentro de esta categoría. Finalmente, sumaríamos la distancia total de los segmentos.
- **Distancia a sprint:** En este caso, haríamos lo mismo que para la métrica anterior, pero esta vez estableciendo un umbral de velocidad de 25 km/h
- **Sprint válidos:** En este caso hemos seleccionado el número de tramos, en los que el jugador mantuvo el estado de sprint por un umbral mínimo de 25 frames (1 segundo)
- **Velocidad máxima alcanzada:** Una vez calculada la velocidad en los distintos tramos, seleccionamos la mayor.
- **Número de desaceleraciones:** Con las velocidades calculadas, podemos calcular los cambios bruscos de velocidad, en este caso, bajadas bruscas de velocidad y que se producen en un lapso de al menos 25 frames (Evitando posibles fallos del tracking.)

- **Ratio descanso/trabajo:** Es una ratio en el que calculamos la relación entre velocidades que se podrían considerar de trabajo, y los tramos con velocidades inferiores, y por tanto consideradas de descanso

Por otro lado, si hablamos del aspecto del rendimiento deportivo, a partir del dataset de eventos, pudimos construir tres diferentes datasets que contienen información sobre los tres aspectos básicos:

El primero de ellos sería el aspecto ofensivo, en el que encontraríamos métricas principalmente de ataque, como serían los tiros, regates completados, goles, asistencias...etc.

El segundo aspecto sería el defensivo, que incluye todo lo relativo tanto a las jugadas positivas para el equipo, como sería el caso de las recuperaciones, intercepciones, despejes... y lo relativo a jugadas que perjudican al equipo, como serían las faltas o las tarjetas.

Por último, tendríamos el dataset con datos de juego, es decir, cómo el jugador ha contribuido al correcto funcionamiento del equipo en cuanto a creación del juego se refiere. Esto incluiría la cantidad de pases completados, pérdidas, centros...etc.

Con estas métricas de rendimiento deportivo, sería posible para el cuerpo técnico, analizar cómo se comporta un jugador en los distintos aspectos relativos al juego, permitiéndole ajustar sus tácticas, al perfil de jugadores con los que cuentan en plantilla, maximizando el potencial de estos.

5.4.3 Calculo de la evaluación del comportamiento del equipo en situaciones tácticas específicas

El último apartado que nos pidió el cuerpo técnico es el de la Táctica Analítica, este apartado trata de un estudio minucioso del comportamiento del equipo, considerando cuatro etapas principales del juego; la primera de ellas hace referencia al momento con balón (MCB), otra al momento sin balón (MSB), otra a la transición defensa-ataque y la última en cambio transición ataque-defensa. Para cada etapa es necesario considerar y documentar los siguientes elementos:

Comenzamos con la fase del **momento con balón (MCB)**, donde el cuerpo técnico necesita que examinemos la conducta del equipo cuando posee el balón, para ello nos dieron unos elementos a considerar:

- **Zonas de pases:** Identificación de las áreas del terreno donde se llevan a cabo los pases, tanto durante la etapa de construcción como en las áreas de finalización, de esta forma el cuerpo técnico quiere buscar las áreas donde el equipo se siente más cómodo con el balón.
- **Perdidas:** Registros y un mapa de calor de las pérdidas de balón, estudiando en qué áreas del campo ocurren y en qué situaciones, lo que ayudaría al cuerpo técnico a identificar en que áreas del campo son menos eficaces con la posesión.
- **Pases para llegar al campo rival:** Evaluación de los pases que posibilitan al equipo avanzar y traspasar la línea divisoria hacia el campo adversario, pudiendo analizar la forma más eficaz de comenzar con el ataque.
- **Pases en campo contrario:** Análisis de las acciones efectuadas en el campo del equipo adversario, valorando la circulación, las combinaciones y la ocupación de espacios.
- **Centros al área:** Evaluación y ubicación de los centros enviados al área contraparte, evaluando su origen y eficacia.

Continuamos con la etapa del **momento sin balón (MSB)**, en la que el equipo técnico solicita que examinemos el comportamiento del equipo cuando no posee la posesión, enfocándonos en los siguientes factores:

- **Alturas de las líneas:** Vigilar la posición de las distintas líneas del equipo (defensa, mediocampo y ataque) durante la etapa defensiva. Esto facilitará al equipo técnico la identificación del nivel de presión, el bloque defensivo y el grado de riesgo que el equipo asume.
- **Distancia entre los jugadores:** Evaluación de las separaciones que los jugadores establecen entre ellos, con el fin de evaluar el grado de compasión, la ocupación de espacios y la habilidad para limitar las alternativas al oponente.

- **Distancia entre líneas:** Analizar los metros que separan las distintas líneas del equipo, permite detectar posibles desequilibrios o aciertos en la estructura defensiva.
- **Áreas de presión:** Documentación de las zonas del campo en las que el equipo aplica la presión sobre el oponente, lo que permite al equipo técnico determinar si la presión se aplica en el campo propio, en el campo adversario o en áreas intermedias, y de qué manera.

Seguimos con la fase de la **transición defensa-ataque (TDA)**, donde el equipo técnico requiere que examinemos el comportamiento del equipo al instante siguiente de recuperar el balón, enfocándonos en los siguientes aspectos:

- **Recuperación y zona:** Determinar las áreas del campo en las que el equipo consigue recuperar el balón, lo que permite evaluar en qué zonas se está mostrando más eficaz en la defensa y dónde se inician las transiciones ofensivas.
- **Pases hacia adelante:** Registro de los pases verticales o de avance que se llevan a cabo justo después de la recuperación, con el propósito de valorar la habilidad del equipo para emprender ataques veloces.
- **Pases hacia atrás:** Análisis de los pases en dirección opuesta a la portería adversaria después de la recuperación, con el objetivo de consolidar la posesión, reestructurar al equipo y organizar el ataque de forma más detallada.
- **Pases para llegar al área rival:** Evaluación de los pases realizados después de la recuperación que buscan llegar a posiciones de finalización o áreas cercanas al área contraria, con el fin de evaluar la eficacia de las transiciones rápidas.

Por último, tenemos la **transición ataque-defensa (TAD)**, donde el equipo técnico pide analizar la conducta del equipo después de perder el balón, teniendo en cuenta los siguientes factores:

- **Zona de la pérdida:** Documentación de las zonas del campo donde el equipo pierde la propiedad, lo que permitirá reconocer áreas de riesgo y modificar la estrategia de presión después de la pérdida.

- **Metros corridos hasta finalizar defensa:** Evaluación de la distancia que el equipo cubre desde la pérdida de la pelota hasta que consigue frenar o neutralizar el ataque adversario, lo que facilitará la valoración de la rapidez y efectividad en la reestructuración defensiva.
- **Jugadores que saltan a la presión tras pérdida:** Evaluación de la cantidad de jugadores que reaccionan de inmediato a la pérdida e intentan aplicar presión para recuperar el balón de manera rápida, proporcionando al equipo técnico un entendimiento de la intensidad y coordinación en la presión después de la pérdida.

Para realizar este estudio de este apartado vamos a emplear tres bases de datos fundamentales: el data set de eventos, el data set de seguimiento (tracking) de jugadores y el data set de seguimiento de balones. La mezcla de estos tres conjuntos de datos nos ofrece una perspectiva integral del comportamiento táctico del equipo en las distintas etapas del juego.

5.5 Visualización de datos: Análisis táctico y de rendimiento

Para facilitar la interpretación y el análisis de los datos obtenidos, hemos creado una capa de visualización que facilita la representación de la información más significativa de manera clara y comprensible. Para llegar a este punto, hemos utilizado tanto herramientas de programación como soluciones dirigidas a los usuarios finales.

En un principio, hemos acudido a varias librerías de visualización de datos en Python, tales como *Matplotlib*, *Mplsoccer* y *Seaborn*, las cuales facilitaron la creación de gráficos estáticos e interactivos valiosos durante el desarrollo y análisis exploratorio de la información. Estas herramientas fueron esenciales para confirmar las hipótesis iniciales, identificar patrones y orientar las decisiones en el diseño de la plataforma.

Posteriormente, con el propósito de desarrollar una interfaz intuitiva y práctica para usuarios, hemos optado por Lovable, una herramienta No-code que te permite crear aplicaciones web y productos digitales usando inteligencia artificial generativa, la cual nos ha facilitado la creación de un dashboard interactivo. (*NoCodeHackers, 2025*)

5.5.1. Librerías Python utilizadas para visualización

Como hemos comentado previamente, durante el desarrollo de la plataforma se han utilizado varias librerías de Python con la finalidad de producir representaciones visuales precisas, claras y ajustadas a los distintos tipos de datos futbolísticos estudiados. Estas herramientas nos han facilitado tanto examinar la información en etapas iniciales del análisis como generar visualizaciones finales para su exposición y utilización en el tablero de control.

Las principales librerías utilizadas fueron Matplotlib, MplSoccer y Seaborn, cada una con un propósito específico dentro del proceso de visualización.

Matplotlib

Matplotlib es una de las librerías de visualización más establecidas en el entorno de Python. Su versatilidad y extenso respaldo la hacen una herramienta imprescindible para la elaboración de gráficos estáticos y personalizados. En nuestro caso, hemos utilizado Matplotlib como fundamento para elaborar visualizaciones genéricas, aunque principalmente la hemos utilizado para la composición de estilos gráficos y elementos visuales (colores, tipos de letra, leyendas...). (*Matplotlib Developers, 2025*).

Además, su vinculación con otras librerías, como MplSoccer y Seaborn, también permitió mantener una uniformidad visual en todo el ambiente gráfico.

Mplsoccer

MplSoccer es una librería dedicada a la visualización de datos de fútbol, ampliando las capacidades de Matplotlib con instrumentos creados específicamente para este deporte (*mplsoccer, 2025*). Entre sus características más destacadas se encuentran la capacidad de:

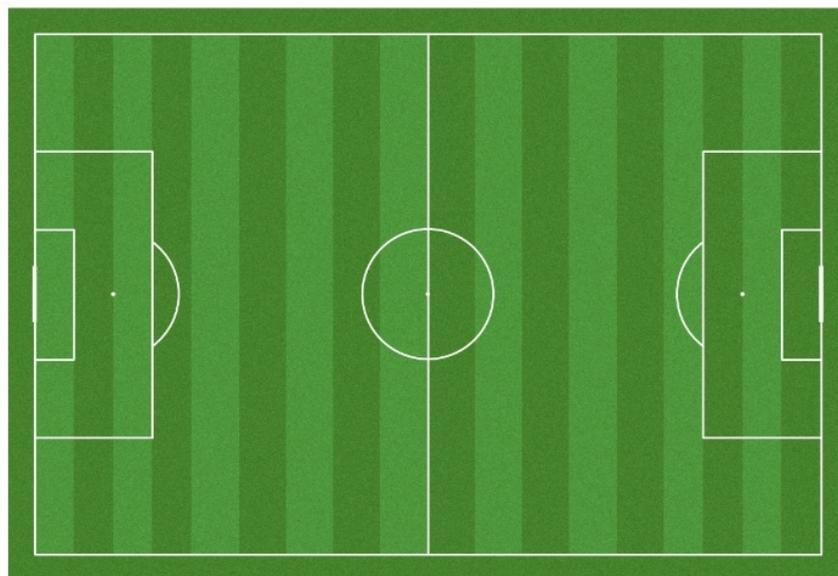
- Dibujar campos de fútbol bajo diferentes ángulos (horizontal, vertical, radiales, etc.)
- Reflejar sucesos en el terreno de juego, tales como pasos, disparos, posiciones intermedias, áreas de calor, entre otros.

- Importar y manejar de forma directa datos públicos, como los suministrados por StatsBomb.

En este proyecto, MplSoccer ha desempeñado un papel crucial en la representación espacial de los partidos y acontecimientos esenciales. Principalmente aplicándose para visualizar mapas de tiros/pases por jugador o equipo, representar trayectorias de pases en situaciones concretas o elaborar mapas de calor a medida para el posicionamiento medio y las áreas de acción.

Estas visualizaciones han aportado un plus al análisis táctico y técnico de los partidos, algo particularmente beneficioso en el ámbito amateur, donde estas herramientas a menudo no se encuentran accesibles.

Ilustración 12: Visualización terreno de juego (base)



Fuente: Elaboración propia

Seaborn

Seaborn es una destacada biblioteca basada en Matplotlib, enfocada en la creación de gráficos estadísticos con un diseño elegante y coherente. Proporciona una interfaz fácil para diseñar

visualizaciones complejas con escasas líneas de código, lo cual simplifica el análisis exploratorio de la información. (*Waskom, 2025*)

Durante la creación de esta plataforma, se ha empleado principalmente para observar vínculos entre diferentes métricas a través de mapas de calor (heat maps), y debido a su habilidad para fusionar de manera sencilla datos numéricos y categóricos, Seaborn ha resultado útil en la etapa de validación del modelo de datos y en la identificación de patrones.

5.5.2. Creación del Dashboard con Lovable

Tras finalizar el análisis y la representación de los datos a través de herramientas de desarrollo, se llevó a cabo la creación de un dashboard interactivo como producto final diseñado para el usuario. Para lograrlo, hemos optado por emplear Lovable, una plataforma sin codificación que facilita la creación de visualizaciones dinámicas de manera fácil, intuitiva y con una estética profesional.

La implementación de Lovable satisface la demanda de proporcionar una solución práctica y asequible para perfiles no expertos, tales como entrenadores, jugadores o analistas de ambiente aficionado, que necesitan examinar datos y métricas de manera rápida y entendible, sin requerir habilidades en programación o análisis de datos.

Hemos dividido el dashboard en 6 módulos distintos, de esta manera buscamos organizar la información por grupos para así poder mejorar la experiencia usuario, y al mismo tiempo darle ese toque organizado y profesional. Estos son los módulos del dashboard: Dashboard (página principal), Analytics, Management, Players, Calendar, Videos y por último, Tactical Board.

5.5.2.1 Dashboard (página principal)

En la página principal del dashboard, la atención se enfoca principalmente en información del equipo, con la finalidad de proporcionar una perspectiva global de la condición presente del

equipo. Primero, se muestra un resumen de la temporada, en el que se señala la cantidad total de encuentros jugados, los resultados alcanzados (victorias, empates y derrotas) y el saldo de goles a favor y en contra. Este bloque ofrece una representación visual y rápida del desempeño total del equipo durante la temporada.

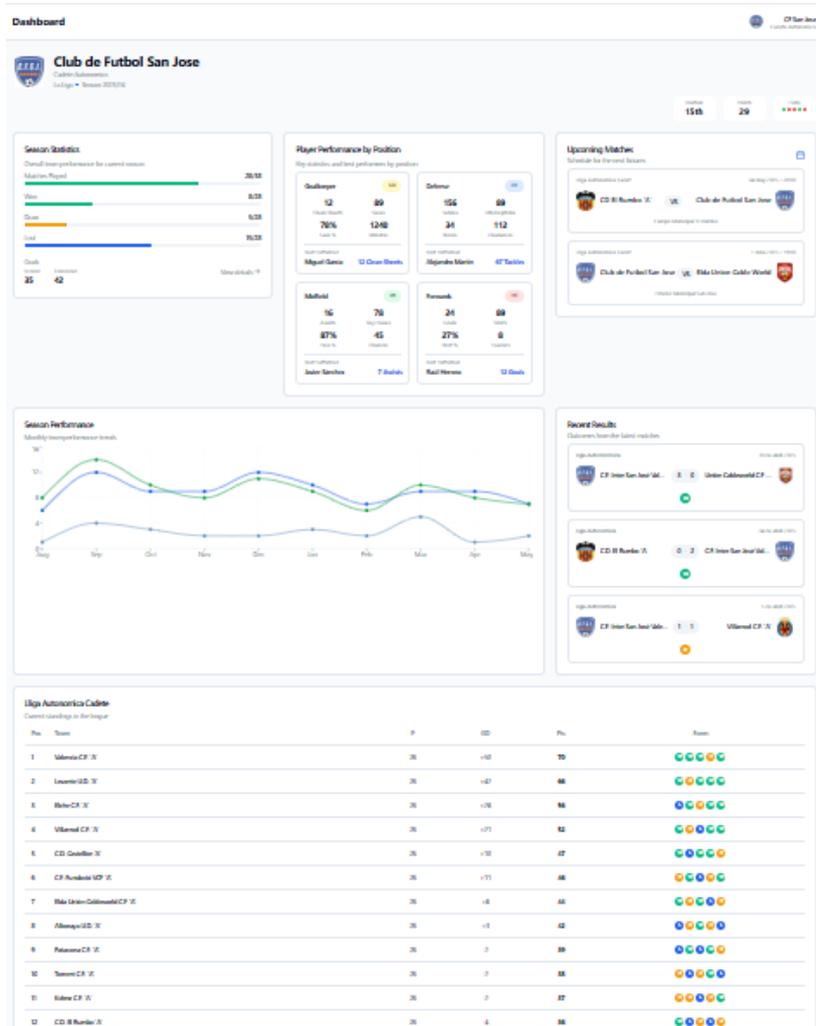
A continuación, se incluye un bloque de desempeño de futbolistas por posición, donde los jugadores se organizan de acuerdo con su posición (porteros, defensas, centrocampistas y delanteros).

En el caso de los porteros, se representa las redes vacías, las paradas, el porcentaje de paradas y los minutos totales, y justo al lado el portero que mejor rendimiento ha tenido, con la estadística que mejor le represente. En los defensas, las entradas, los bloques, las intercepciones y barridas, y al igual que antes, el jugador que más aporte a estas estadísticas. En los centrocampistas, asistencias, % de pases, oportunidades y pases claves, y por último los delanteros, goles, tiros, % de tiros y remates de cabeza.

Seguido, se incorporan dos secciones informativas relacionadas con la programación de los partidos: en un lado, se presentan los próximos encuentros, con información como fecha, hora, rival y lugar de juego; en otro, se presenta un resumen de los últimos partidos jugados, lo que permite apreciar el progreso reciente del equipo.

Finalmente, se incorporan dos bloques extra destinados al análisis del desempeño grupal. El primer gráfico mensual muestra la progresión del equipo a través del tiempo, exhibiendo indicadores como puntos logrados, goles anotados y goles recibidos por mes. El segundo bloque presenta la tabla actualizada de la liga, incluyendo datos fundamentales como la posición de cada equipo, puntos totales, diferencia de goles y racha en los cinco últimos encuentros.

Ilustración 13: Dashboard (inicio)



Fuente: Elaboración propia

5.5.2.2 Analytics

Esta sección es posiblemente la más relevante del dashboard, puesto que contiene todos los datos obtenidos de los partidos. Para simplificar su entendimiento y conforme a las directrices del cuerpo técnico, la hemos segmentado en tres subpestañas, cada una enfocada en un tipo de información particular.

A pesar de que el equipo técnico estableció inicialmente la organización, durante el desarrollo hemos sugerido mejoras y modificaciones, que en su mayoría fueron aceptadas e incorporadas, potenciando de esta manera la utilidad final del tablero de control.

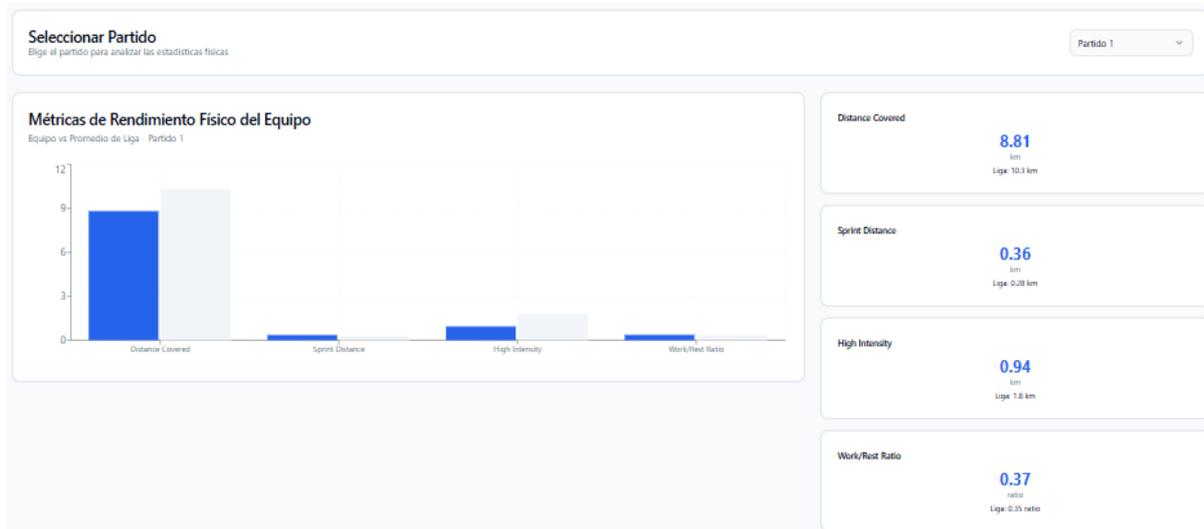
5.5.2.2.1 Physical stats

La primera de las subpestañas se enfoca en el estudio de los datos físicos, un sector crucial para valorar el desempeño tanto individual como grupal en cada encuentro. Ya que las circunstancias particulares de cada partido (como el oponente, el campo de juego o el clima) pueden tener un impacto significativo en los resultados físicos, hemos implementado un filtro interactivo que facilita la elección del partido que se quiere examinar, permitiendo una valoración más exacta.

En la sección superior de la subpestaña, se muestra un gráfico comparativo donde se comparan varias medidas físicas —tales como la distancia total recorrida, la distancia en sprint, los metros a alta intensidad y la relación entre trabajo y descanso— entre el equipo completo y el promedio liga. Justo debajo, estas mismas estadísticas se presentan de manera específica, lo que facilita la observación de los valores.

En la sección inferior se ha incorporado una tabla exhaustiva con los datos físicos de cada jugador, lo que facilita un estudio más detallado y particular del desempeño individual. Esta tabla contiene campos como el nombre y posición del jugador (para simplificar su reconocimiento), además de las métricas físicas más relevantes: distancia total recorrida (km), intensidad alta (km), distancia en sprint (km), relación trabajo/descanso, cantidad de sprints, desaceleraciones y la velocidad máxima lograda (km/h). Esta información es particularmente valiosa para el cuerpo técnico al estimar la carga física de cada jugador y organizar entrenamientos o rotaciones.

Ilustración 14: Estadísticas físicas



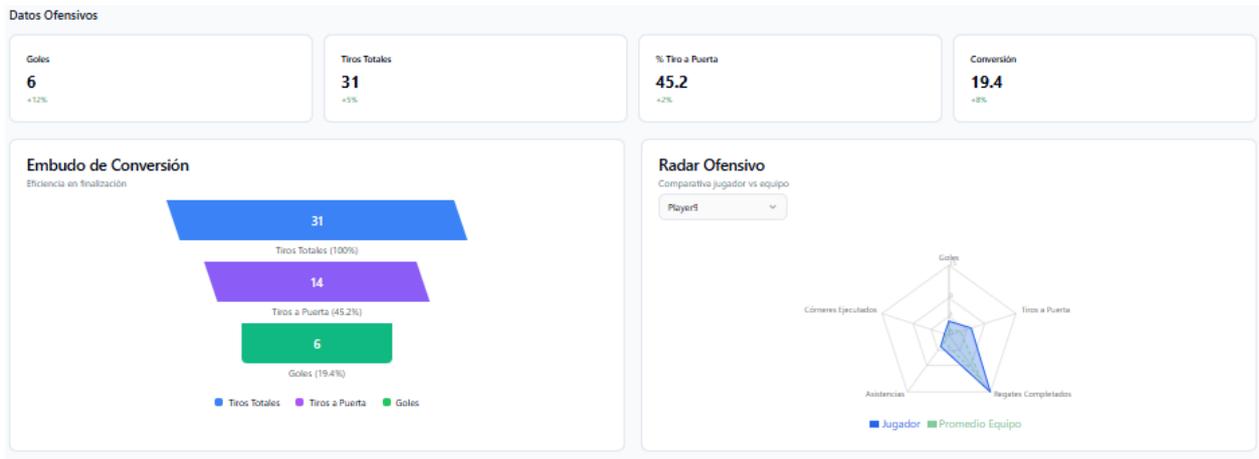
Fuente: Elaboración propia

5.5.2.2.2 Game analytics

La segunda subpestaña se enfoca en el estudio de los datos vinculados al juego, que se han repartido nuevamente en tres áreas: ofensiva, defensiva y juego general. Esta separación se debe a la exigencia de examinar de forma particular el comportamiento del equipo en cada una de estas etapas del juego.

En la parte ofensiva, la atención primordial se centra en los disparos a puerta y su efectividad. Inicialmente, se muestran varios segmentos de información esenciales: goles anotados, disparos totales, porcentaje de disparos a puerta y la relación de conversión a gol, indicadores que facilitan la valoración de la habilidad del equipo para marcar goles. A continuación, se presenta un embudo de conversión de tres niveles (tiros, disparos a puerta y goles), que facilita la visualización del desempeño grupal al concluir los juegos. Además, se muestra un diagrama radial (radar) de cinco ejes, donde se contrastan algunas características ofensivas del jugador escogido con la media del equipo, lo que permite un análisis individual dentro del marco del equipo.

Ilustración 15: Estadísticas de juego (ofensivo)



Fuente: elaboración propia

En la zona defensiva, se preserva una estructura parecida. Primero, se presentan indicadores esenciales como recuperaciones totales, despejes efectuados, infracciones cometidas y tarjetas obtenidas, junto con su fluctuación porcentual en relación con los partidos previos, lo que facilita la observación de tendencias actuales. Luego, se incorpora un histograma de recuperaciones por juego, junto con un diagrama de líneas que muestra la progresión del porcentaje de victorias obtenidas, un indicador esencial para evaluar el desempeño defensivo del equipo.

Finalmente, en la sección dedicada al juego general, se presenta un diagrama de líneas que ilustra la evolución de pases totales, completados y fallados por juego, junto con un histograma de posesiones perdidas, lo que facilita la evaluación de la habilidad del equipo para conservar la posesión del balón. Esta parte concluye con un resumen de cuatro métricas fundamentales vinculadas al juego: exactitud en los pases (%), pérdidas totales de balón, centros ejecutados y pases largos, ofreciendo una perspectiva global del estilo de juego y su eficacia.

5.5.2.2.3 Tactical analytics

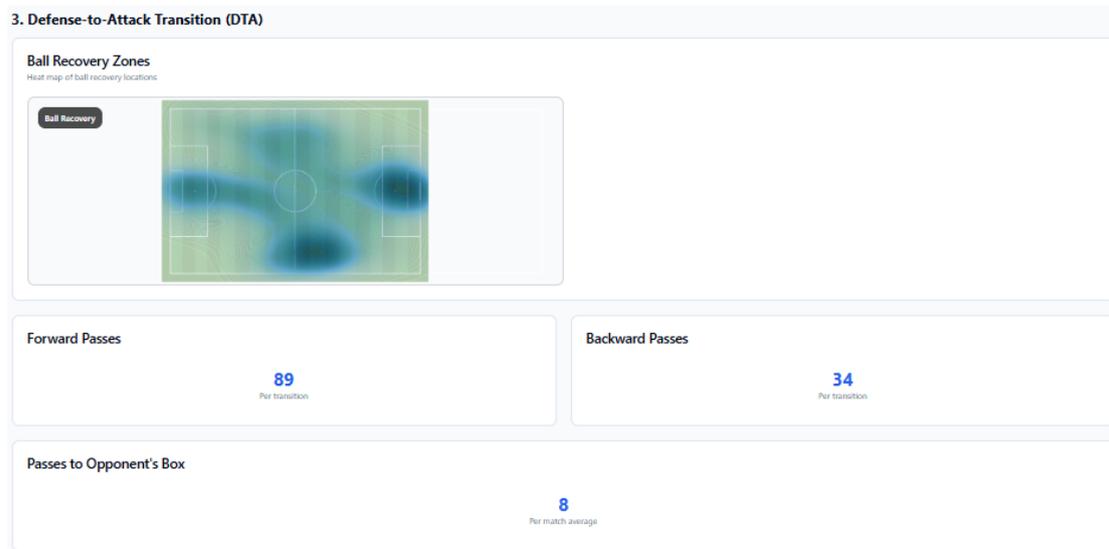
La última sección del módulo Analytics se enfoca en el estudio de elementos tácticos, esenciales para entender el comportamiento grupal del equipo tanto en etapas con balón como sin él, además de en los instantes de transición entre ataque y defensa.

La primera parte se centra en los instantes de posesión, donde se examinan las áreas de pase y las pérdidas de balón a través de mapas de calor, los cuales facilitan la identificación de patrones y zonas de riesgo o productividad. Además, se incorporan métricas fundamentales como la cantidad de centros dirigidos al área adversaria, el número de pases efectuados en el campo contrario o la media de pases requeridos para alcanzar el área adversaria, indicadores que contribuyen a valorar la eficacia y el estilo de juego ofensivo del equipo.

A continuación, se muestran los datos vinculados a los instantes sin juego. En esta parte, se emplea un mapa en dos dimensiones para ilustrar la altura media de las líneas de juego (defensa, mediocampo y delantera), lo que facilita la visualización del posicionamiento táctico del equipo. Además, se presenta un mapa de calor de las áreas de presión, junto con las separaciones entre líneas y entre jugadores, lo que proporciona datos acerca del nivel de compacidad, organización y sincronización defensiva.

A continuación, se examinan las transiciones entre defensa y ataque. Para ello, se emplea un mapa de calor de las áreas de recuperación de balón, complementado con estadísticas como la media de pases por partido hasta llegar al área del oponente, además del número de pases hacia adelante y pases hacia atrás, que facilitan la comprensión de la rapidez y verticalidad de las transiciones ofensivas.

Ilustración 16: Estadísticas tácticas (DTA)



Fuente: Elaboración propia

Finalmente, la sección finaliza con el estudio de la transición entre el ataque y la defensa, en la que se puede apreciar la conducta del equipo después de perder el balón. Se proporciona un mapa de calor de las áreas donde ocurren las pérdidas de posesión, una tabla con los jugadores que ejercen mayor presión después de una pérdida, y una métrica que indica el promedio de metros avanzados por cada línea hasta finalizar la etapa defensiva. Todo esto posibilita valorar la habilidad de respuesta, organización y presión después de la pérdida, elementos cruciales en la rápida recuperación y el equilibrio defensivo.

5.5.2.3 Management

La sección de administración inicia mostrando datos generales acerca de la plantilla actual, que incluyen el número de jugadores registrados, las competencias en las que el equipo participa, los encuentros pendientes y los documentos a disposición. A continuación, se detalla cada uno de estos componentes, proporcionando características como la alteración de la plantilla o la inclusión de nuevos jugadores, la modificación del calendario (diseñada para cuando se confirmen fechas y lugares de los encuentros), y la configuración del club (datos generales, permisos de acceso, etc.). Además, facilita la administración de documentos como contratos o actas, la actualización de datos vinculados a las competencias, la consulta de elementos financieros (presupuesto, ingresos y costos) y, por último, la revisión de la actividad más reciente en la plataforma.

5.5.2.4 Players

La siguiente pestaña recopila información sobre los jugadores de la plantilla. En este caso creemos que lo más óptimo es poner en la parte superior información general, como son el número total de jugadores, edad promedia, goles totales, asistencias... Y después proyectar una lista interactiva con todos los jugadores que completan la plantilla, en la que saldrá la siguiente información; dorsal, nombre, posición, edad, apariciones, minutos, goles, asistencias y tarjetas recibidas (tanto amarillas como rojas). Para facilitar al usuario la búsqueda de jugadores, también hemos añadido un filtro de búsqueda por nombre/dorsal y otro filtro por posición.

Ilustración 17: Estadísticas jugadores

Players CF San José
Categoría Sub-18

Total Players
10

Average Age
16.5

Total Goals
38

Total Assists
25

Squad Overview
Manage and view player information Add Player

🔍 Search by name or shirt number... All Positions: ▼

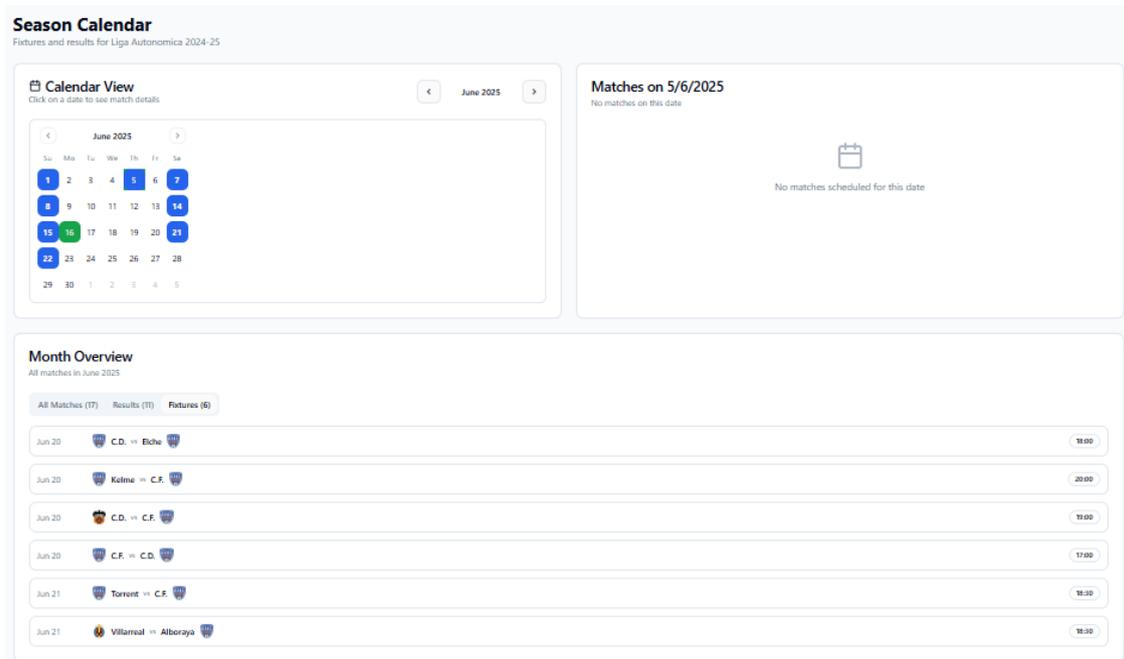
#	Player	Position	Age	Apps	Minutes	Goals	Assists	Cards
1	Miguel García Rodríguez <small>Spain</small>	Goalkeeper	17	14	1240	0	0	🟡
2	Alejandro Martín López <small>Spain</small>	Defender	16	13	1180	1	2	🟡
4	Pablo Fernández Silva <small>Spain</small>	Defender	17	15	1320	0	1	🟡🔴
5	Carlos Ruiz Moreno <small>Spain</small>	Defender	16	11	960	2	0	🟡
6	Javier Sánchez Díez <small>Spain</small>	Midfielder	17	16	1450	3	5	🟡
8	Sergio González Vega <small>Spain</small>	Midfielder	16	14	1220	2	7	🟡
10	David Jiménez Castro <small>Spain</small>	Midfielder	17	15	1380	6	4	🟡
11	Adrián Torres Blanco <small>Spain</small>	Forward	16	13	1150	8	2	🟡
9	Raúl Herrera Peña <small>Spain</small>	Forward	17	14	1280	12	3	🟡
7	Rubén Morales Ortega <small>Spain</small>	Forward	16	10	890	4	1	-

Fuente: Elaboración propia

5.5.2.5 Calendar

En esta pestaña, como bien se puede observar en el nombre, hemos decidido hacer referencia al calendario de la temporada del equipo. En este caso hemos decidido utilizar 3 widgets para este apartado, el primero sería una visualización mensual de un calendario, en el que estarían marcadas las fechas en las que el equipo tiene, al lado tendríamos otro widget que sería para el calendario del día, el cual estaría enfocado a mostrar las horas de los distintos eventos, como llegada al estadio, inicio del partido... Por último, hemos optado también por visualizar el calendario de todos los partidos de la temporada, agrupándolos por 3 grupos (todos, completados y futuros).

Ilustración 18: Calendario



Fuente: Elaboración propia

5.5.2.6 Videos

Además, se ha introducido una nueva pestaña llamada "Vídeos", que facilita el almacenamiento y manejo de grabaciones de partidos. En esta sección, los usuarios pueden observar la cantidad total de vídeos guardados, la cantidad que ya ha sido analizada y la cantidad que aún necesita ser revisada. Además, se proporciona acceso inmediato a los vídeos más recientes, lo que es particularmente beneficioso para que los jugadores puedan evaluar sus actuaciones y revisar grabaciones de encuentros previos de manera rápida y fácil.

5.5.2.7 Tactical Board

Por último, hemos decidido añadir la pestaña de "Tactical Board", esta tiene como función principal ofrecer al entrenador una pizarra virtual en la que puede preparar tácticas o jugadas ensayadas para después explicársela a sus jugadores.

Dentro de esta pestaña, hemos añadido 2 subpestañas, “Main tactical board” y “Corner kicks”, en la primera se puede visualizar todo el terreno de juego, los jugadores de ambos equipos ilustrados por círculos rojos y azules, y un balón representado por un círculo negro, mientras que en la segunda solo se visualiza una portería, esto pensando en que facilite el uso en jugadas de córner, faltas cercanas o cualquier jugada más precisa.

Ilustración 19: Pizarra interactiva



Fuente: Elaboración propia

6. Modelo de Negocio

FutHub nace como una solución integral que combina el análisis automatizado de partidos, la visualización de datos relevantes y la gestión operativa del club, todo dentro de un mismo entorno digital. A diferencia de las soluciones existentes en el mercado, orientadas mayoritariamente a clubes profesionales con grandes recursos, nuestra plataforma propone un modelo de negocio basado en la suscripción, permitiendo a los entrenadores subir grabaciones

de sus partidos —cumpliendo ciertos estándares técnicos— para recibir informes analíticos procesados mediante herramientas de inteligencia artificial y visión por computadora.

Durante el presente año, el proyecto validará su viabilidad a través del desarrollo de un Producto Mínimo Viable (MVP) que será implementado en colaboración con un club semiprofesional. Esta primera fase de prueba tiene como objetivo no solo comprobar la funcionalidad técnica de la solución, sino también recoger feedback directo de los usuarios, evaluar la aceptación en condiciones reales y optimizar el producto antes de su lanzamiento comercial.

A lo largo de este apartado se detallarán el estudio de mercado realizado, los principales competidores del sector, el funcionamiento general de FutHub, el enfoque técnico del sistema y los pilares estratégicos que sostienen este modelo de negocio innovador.

Ilustración 20: Logotipo FutHub



Fuente: Elaboración propia

6.1 Estudio de Mercado

FutHub se orienta principalmente al segmento de clubes semiprofesionales, amateur y de categorías formativas, un ecosistema caracterizado por su alto nivel de competitividad deportiva pero limitado acceso a recursos tecnológicos avanzados. Estos clubes, ubicados mayoritariamente en ligas regionales o divisiones inferiores, enfrentan la necesidad constante de conocer su rendimiento en contextos de recursos escasos y estructuras organizativas reducidas. Sin embargo, en los últimos años ha surgido un creciente interés por adoptar herramientas digitales que permitan profesionalizar la gestión interna y mejorar el análisis

técnico-táctico, especialmente por parte de entrenadores jóvenes, formados en entornos académicos donde la tecnología aplicada al deporte ya es parte del proceso de formación.

La propuesta de FutHub surge como respuesta a una necesidad concreta y latente: permitir a estos clubes el acceso a una plataforma accesible que combine análisis automatizado de partidos y herramientas de gestión. Frente a la ausencia de soluciones adaptadas al nivel semiprofesional que cumplan con criterios de simplicidad, eficacia y sostenibilidad económica, se presenta una oportunidad de mercado. Así, el contexto del fútbol modesto se presenta como un espacio factible para la introducción de nuestra solución capaces de integrarse en la operativa diaria del cuerpo técnico y mejorar significativamente su capacidad de análisis y organización.

Actualmente, muchos clubes semiprofesionales no tienen acceso a sistemas de análisis estructurados y dependen exclusivamente del juicio subjetivo del entrenador y el análisis manual. Esta limitación les impide identificar patrones de juego, áreas de mejora y rendimiento real del equipo y sus jugadores. Al mismo tiempo, no existen plataformas asequibles y específicamente pensadas para su contexto y dinámica operativa.

La necesidad del mercado no es solo tecnológica, sino también organizativa: muchos clubes carecen de herramientas básicas para llevar un registro sistematizado de jugadores, estadísticas, entrenamientos y partidos.

FutHub se posiciona entre soluciones altamente profesionales y costosas, y plataformas genéricas poco adaptadas al fútbol base o amateur. A continuación, se presenta un análisis comparativo de los principales competidores:

Tabla 4: *Comparación de competidores*

PLATAFORMA	DESCRIPCIÓN	LIMITACIONES
WYSCOUT	Plataforma profesional de scouting y análisis global.	Coste elevado, diseñado para clubes profesionales.

INSTAT	Análisis avanzado de video y datos.	Enfoque elitista, tarifas inaccesibles para clubes modestos.
NACSPORT	Software de análisis manual por etiquetas.	Licencias de software costosas, curva de aprendizaje alta.
HUDL	Muy utilizada en deporte base, enfocada en video.	Menor énfasis en análisis táctico automatizado.
FOOTTERS	Plataforma de streaming para fútbol modesto.	No ofrece análisis; centrado en la retransmisión.
SOFASCORE	Apps de estadísticas y predicción para usuarios.	No orientadas a entrenadores ni análisis técnico interno.

Fuente: Elaboración propia

6.2 validación y prueba del producto (MVP)

Como parte del proceso de validación del modelo de negocio de FutHub, durante la próxima temporada se pondrá en marcha un Producto Mínimo Viable (MVP) que será probado en un contexto real y competitivo, trabajando de la mano con un cuerpo técnico profesional. Esta fase no solo permitirá testar la funcionalidad técnica de la plataforma, sino también evaluar su adopción, utilidad práctica y potencial de escalabilidad dentro del ecosistema del fútbol semiprofesional.

El MVP se desarrollará en colaboración con el club en el que estará involucrado el cuerpo técnico compuesto por Adrián Dioume (entrenador principal), Vicent (asistente técnico), Xavi Aguilar (preparador físico) y Salva Fenoll (delegado). Este equipo multidisciplinar aporta experiencia directa en el fútbol semiprofesional y una comprensión profunda de las dinámicas tácticas, físicas y organizativas que caracterizan a este tipo de equipos. Su participación

garantiza que el proceso de validación se lleve a cabo bajo criterios profesionales, con foco tanto en el rendimiento deportivo como en la viabilidad operativa del producto.

Durante esta fase piloto, se trabajará con el equipo real que el cuerpo técnico dirigirá en la temporada, aplicando FutHub como herramienta central para el análisis de los partidos, la gestión de la plantilla y el seguimiento del rendimiento físico y técnico de los jugadores. Esta implementación en condiciones reales permitirá no solo observar cómo la plataforma se integra en la rutina de trabajo semanal, sino también recoger datos cuantitativos y cualitativos sobre la experiencia del usuario, el ahorro de tiempo en el análisis táctico, la utilidad de las visualizaciones generadas y las posibilidades de mejora del producto.

Los aprendizajes derivados de esta primera experiencia servirán de base para realizar ajustes técnicos, refinar la propuesta de valor y preparar el lanzamiento comercial de FutHub a una escala mayor, adaptado a las verdaderas necesidades del mercado objetivo.

6.3 Funcionamiento general

A nivel general, podríamos dividir el funcionamiento de la plataforma en 3 aspectos, desde que el usuario ingresa en la plataforma hasta que obtiene el resultado final. Aun así, durante todo el proceso, el usuario podría ver distintos resultados intermedios antes de llegar al dashboard, que sería el resultado final.

La primera de las etapas sería la carga del vídeo y el procesamiento del mismo. En este sentido, el usuario subiría a la propia plataforma el vídeo, siempre y cuando cumpla unos requisitos mínimos, tanto de calidad como de configuración de cámara y ángulo. Sería la propia plataforma la que proporcionaría al usuario las características básicas del vídeo para que este fuera procesado. Toda la información se extraería de forma automática, de forma que se podrían sacar diferentes frames de la serie temporal, y hacer la comparativa para ver si se cumplen los requisitos de visibilidad, resolución, consistencia en el ángulo...etc. Esto haría que cuando el vídeo pasara al segundo paso, la cantidad de errores en la detección y tracking se minimizara mucho. Toda esta información se guardará en una base de datos relacional en la que se le asigna

un id a cada uno de los partidos. Esto nos permite procesar varios vídeos simultáneamente sin que se pierdan datos importantes en el proceso.

Dentro de este aspecto, también entraría todo lo relacionado con el procesado del vídeo, tanto la detección de los distintos jugadores sobre el campo, así como el tracking de estos a lo largo de todo el vídeo, y la proyección del campo para poder geolocalizar en todo momento a los jugadores. Toda esta información, se iría guardando en una base de datos, usando como identificadores tanto un ID concreto para ese equipo, como un ID para cada uno de los jugadores. Esto se hace ya que esta base de datos es crucial para el cálculo de las distintas estadísticas y métricas, así como para la realización de distintas visualizaciones.

La segunda etapa, está relacionada con el cálculo de las distintas métricas tanto físicas como deportivas para los distintos jugadores y a nivel colectivo. Para el cálculo de estas métricas haríamos uso de la base de datos que ya construimos en el paso anterior. Además, para poder ver el progreso y la comparativa con otros partidos de la temporada, se utilizará un ID para cada partido, así como comentamos al principio de la primera etapa. Esto nos permite la trazabilidad de las distintas métricas a lo largo de toda la temporada.

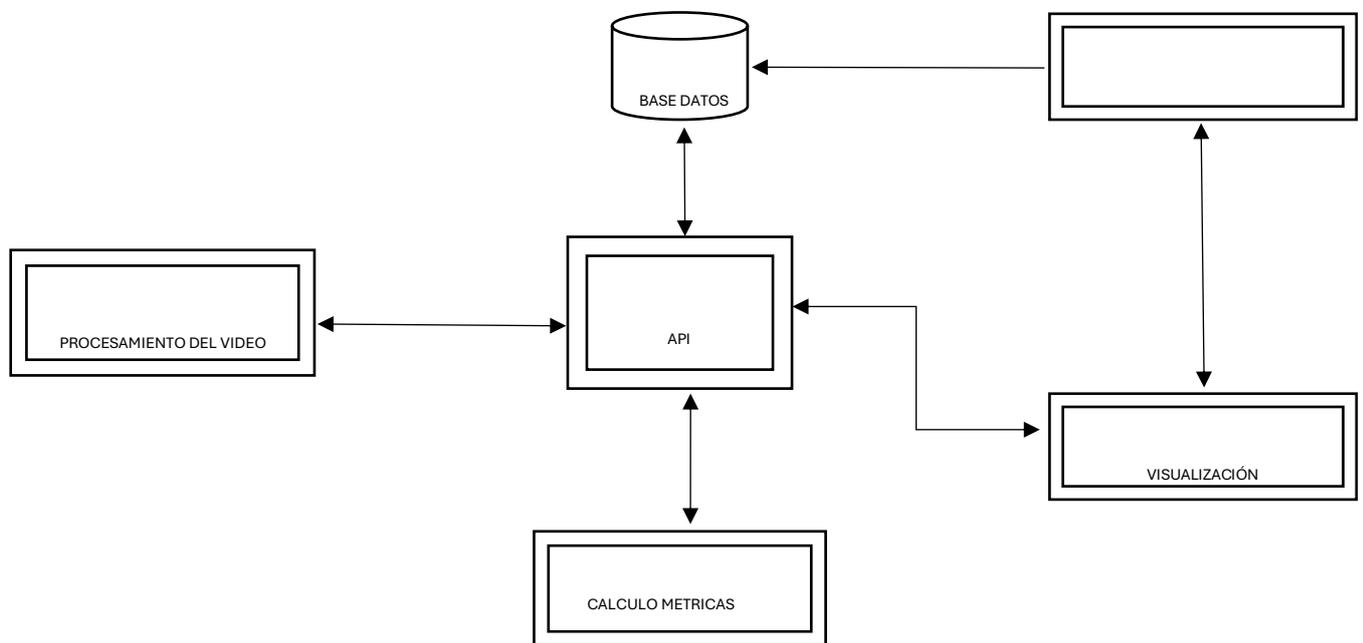
Puede ser en este punto, en el que el usuario pueda configurar su perfil en función de las métricas que le gustaría visualizar, en función del plan que hubiera seleccionado.

Por último, tendríamos lo que sería el output que el cliente recibiría, en este caso en forma de dashboard interactivo, que se nutriría de las distintas bases de datos que se han ido creando a lo largo del proceso, y que permitirían al usuario navegar a través de los distintos tipos de análisis y tener una visión completa del rendimiento de sus jugadores en los distintos ámbitos del fútbol.

Todas estas etapas, estarían conectadas a través de una API que es lo que permitiría el flujo de información y datos entre las diferentes etapas del proceso

Para tener una imagen mucho más explicativa de todo lo relacionado con la funcionalidad. Se podría visualizar en forma de esquema, en donde se pueden ver las diferentes interconexiones entre las etapas y distintos análisis que se realizan para llegar hasta la visualización final, que sería lo que el usuario vería a través de la plataforma.

Ilustración 21: Esquema de la funcionalidad



Fuente: Elaboración propia

Como mencionamos anteriormente, vemos en el centro del diagrama una API, encargada del flujo de datos e información entre los diferentes procesos. El usuario subiría a través de la plataforma directamente a la base de datos el vídeo que se quisiera procesar. Dicho vídeo pasaría por las distintas fases del análisis, siendo la API la que movería toda la información de un proceso a otro, así como hacia la base de datos, donde se almacenarán tanto los vídeos originales, como todos los datos obtenidos.

Dichos datos, son los que posteriormente se mostrarán en el dashboard que el usuario final verá, y que podría llegar a personalizar, de ahí que se muestre una flecha que va en ambas direcciones entre la visualización de datos y el usuario.

6.4 Modelo de ingresos

Tabla 5: Modelo de ingresos

Plan	Precio estimado (€ / mes)	Público objetivo	Características
Basic	79,99	Escuelas de fútbol, equipos amateur	1 análisis de partido por semana, gestión básica del equipo, estadísticas individuales 1 plantilla activa.
Pro	129,99	Clubes semiprofesionales con cuerpo técnico reducido	Hasta 3 análisis por semana, gestor de entrenamientos, reportes técnicos, estadísticas por jugador, hasta 2 plantillas.
Premium	169,99	Clubes organizados con estructura técnica completa	Hasta 5 análisis por semana, visualización avanzada, comparativas históricas, múltiples usuarios, integración de informes físicos y tácticos, hasta 4 equipos.
Elite	1000	Clubes de alto rendimiento, academias profesionales, federaciones	Análisis profundo y personalizado, informes adaptados al modelo de juego, consultoría mensual con analistas de FutHub, visualizaciones interactivas, evaluación táctica detallada, soporte dedicado.

Fuente: Elaboración propia

Este modelo se basa en un principio clave: permitir que cualquier club, sin importar su tamaño o presupuesto, pueda acceder a herramientas profesionales de análisis y gestión deportiva. La tarifa básica permite a pequeños equipos comenzar con una versión funcional del producto, mientras que los planes superiores ofrecen más capacidad de análisis, automatización y soporte técnico. Esta estructura también incentiva la progresión entre planes conforme el club crece o profesionaliza su estructura.

A diferencia de modelos de pago por uso o licencias anuales cerradas, la suscripción mensual ofrece flexibilidad, facilidad de cancelación y una barrera de entrada más baja. Esto es especialmente importante en el fútbol semiprofesional, donde los presupuestos suelen ser ajustados y volátiles.

Además de las suscripciones individuales por club, FutHub puede escalar el modelo ofreciendo planes multiusuario para escuelas de fútbol, academias privadas o federaciones regionales. Estos acuerdos institucionales permitirían gestionar múltiples equipos bajo una sola cuenta administradora, facilitando la adopción masiva del sistema.

6.5 Infraestructura técnica

Con el objetivo de que el proyecto sea escalable, se ha decidido utilizar un modelo cloud-native, es decir, aprovechar todas las características y posibilidades que ofrece la computación en la nube, tanto a nivel de almacenamiento de datos, como a nivel de computación y procesamiento de vídeos.

En el caso del almacenamiento de datos, no será igual para todos los tipos de datos que guardemos, ya que deberemos de utilizar métodos diferentes para guardar los vídeos originales o las estadísticas de los distintos jugadores.

En el caso de los vídeos que se utilizan para la extracción de las detecciones y el tracking, se almacenarían utilizando Amazon Web Services (AWS), concretamente Amazon S3, que nos

permite no tener un límite de capacidad de almacenamiento (ideal para vídeos debido a su tamaño pesado), así como poder establecer el ciclo de vida de los vídeos dentro de la plataforma, el cuál sería alrededor de 30 días. Después de ese tiempo el vídeo se eliminaría, para liberar espacio y debido a su baja utilidad en el futuro.

En el caso de los datos relacionados con el tracking de jugadores y balón, se usarían archivos estructurados con un formato Parquet, que posteriormente se guardarían un data lake para su posterior tratamiento y análisis. Esto se podría hacer de nuevo en Amazon Web Services para tener la información un poco más centralizada.

Por último, los datos relacionados con las métricas de los distintos jugadores, así como los resultados de los partidos, identificadores de equipos y jugadores se almacenarían en tablas relacionadas, en una base de datos como PostgreSQL, ya que además de ser gratuito y robusto, nos permite acceder a esa información, que será la que principalmente vea el usuario final de una forma rápida y eficiente, mejorando la optimización y la velocidad de acceso al servicio.

En el caso de las distintas inferencias a realizar por el modelo de computer vision (YOLO v8) y los modelos de keypoint detection, se utilizarían también servicios en la nube, que nos permitan realizar todos estos procesos de una forma mucho más eficiente y rápida, cosa que en un ordenador local no se podría hacer. El servicio a utilizar sería preferiblemente Amazon Web Services, ya que además de que los vídeos que se van a procesar ya se encuentran en esta plataforma, tienen soluciones como el caso de AWS Lambda que nos permiten realizar la inferencia con modelos de GPU de Nvidia y sobre todo realizar este proceso para varios vídeos a la vez, lo que ahorraría tiempos de espera clave.

Para la detección de eventos y el posterior cálculo de métricas, se usaría AWS Step Functions junto con AWS Lambda, de esta forma, podemos hacer la detección de eventos con AWS Lambda, y mediante Step Functions desencadenar una serie de acciones de forma secuencial para el cálculo de las distintas estadísticas de los distintos jugadores. Todo ello nos permite introducir cambios sin tener que modificar todo el pipeline de acciones.

La comunicación con el usuario final que recibe el análisis sería mediante una interfaz web que le permita al usuario ver los distintos informes en forma de dashboard o panel de control desde cualquier dispositivo. Además, a través de esa interfaz web, el usuario sería capaz de subir los vídeos, para que posteriormente fueran procesados y se realizara todo el proceso asociado al procesamiento del vídeo. Otra de las dificultades que hemos afrontado a la hora de realizar el trabajo han sido el tratamiento de los vídeos y la integración de las diferentes metodologías necesarias para llevar a cabo todo el flujo de trabajo, como sería el caso del object detection, el tracking, y la detección de los diferentes eventos. Ante esto, nuestra solución se centró en la constante validación interna de los resultados obtenidos a través de la verificación de que los resultados con el vídeo original. Esto se tradujo, por ejemplo, en que la precisión de

7. Conclusiones

A través de este proyecto hemos tratado de demostrar la viabilidad de aplicar una metodología basada en datos orientada a un contexto futbolístico en el que en muchas ocasiones no es posible aplicar este tipo de enfoque debido a las limitaciones económicas y de personal. Queda demostrado que, mediante el enfoque planteado, puede transformarse una grabación (que ciertamente debe tener unas características básicas ya definidas) en métricas completamente estructuradas y que se pueden visualizar en distintos paneles de control para ayudar al equipo técnico a tomar decisiones basadas en datos y maximizar el rendimiento de los jugadores y por consiguiente del equipo en el ámbito deportivo.

En el transcurso de la realización del trabajo se han identificado múltiples retos, tanto relacionados con la calidad y la cantidad de vídeos disponibles y óptimos para la realización de toda la parte de “computer vision”. Además, la variabilidad espacial y temporal que presentan los datos asociados al vídeo hacen que la interpretación y clasificación de las acciones sea complicada. Otra de las dificultades encontradas en la realización del proyecto ha sido la integración de las distintas fases del flujo de trabajo del procesamiento del vídeo, como sería el caso de object detection, tracking y detección de eventos para poder replicarlo de forma sucesiva y automática a los distintos vídeos a los que nos podamos enfrentar en el futuro. En este sentido, la validación interna ha sido clave para poder asegurar la consistencia de los resultados obtenidos, así como su fiabilidad. Esto se ha llevado a cabo, aplicando el procesamiento de vídeo sobre clips de una duración determinada, que nos permitían saber que acciones habían sucedido en dicho vídeo.

Por otro lado, los resultados obtenidos los podríamos dividir en la realización de todo el procesamiento del vídeo, así como haber sido capaces de construir datasets en los que se recojan frame por frame todas las posiciones de los jugadores dentro del campo, así como del balón. Por otro lado, hemos obtenido distintas métricas cuantitativas como sería tanto las métricas de rendimiento físico, así como métricas tácticas tanto a nivel individual como a nivel colectivo.

Una vez realizado tanto el proceso de extracción del dato como su procesamiento posterior para construir una base de datos que nos permita extraer conclusiones, se procedió a la creación tanto de visualizaciones específicas del ámbito futbolístico como un panel de control que incluye todos los ámbitos de este deporte.

Además del trabajo realizado, también se ha validado todo ello junto a un cuerpo técnico real, con los que se ha tenido constante contacto, y quienes nos han dado feedback de cómo querían tener los datos presentados, qué datos querían visualizar y qué implementaciones y mejoras se debían de realizar para acercarse lo máximo posible a un entorno profesional. De este intercambio de información se destaca que la herramienta que les presentábamos facilita la objetividad en los análisis postpartido, así como supone un gran ahorro de tiempo para el cuerpo técnico.

El desarrollo de este trabajo ha permitido explorar las posibilidades del análisis de datos aplicado al fútbol semiprofesional, así como las condiciones técnicas, humanas y económicas necesarias para su implementación real. A lo largo de la investigación, ha quedado claro que existen limitaciones tanto en el acceso a tecnologías adecuadas como en la disponibilidad de personal especializado, lo cual representa una barrera clara para que muchos clubes amateurs puedan adoptar herramientas de análisis.

Estas barreras han sido experimentadas directamente en el proceso de desarrollo de FutHub. La necesidad de contar con sistemas de grabación estables, plataformas de procesamiento de vídeo, almacenamiento en la nube, y algoritmos fiables, implica una inversión inicial que muchos clubes no pueden asumir. A esto se suma la falta de tiempo y formación por parte de los cuerpos técnicos, quienes en muchas ocasiones deben asumir múltiples funciones sin el apoyo de estructuras organizativas amplias.

Teniendo esto en cuenta, se entiende por qué el análisis de datos aún no ha llegado al deporte aficionado, pese a su valor comprobado en el alto rendimiento. La analítica deportiva sigue siendo percibida como una herramienta reservada a clubes profesionales, donde los recursos permiten construir departamentos dedicados exclusivamente a esto.

Sin embargo, este trabajo no pretende limitarse al diagnóstico de un problema, sino proponer una solución viable y escalable. FutHub nace como un primer paso hacia la democratización del análisis de rendimiento deportivo, con el objetivo de que se convierta en una parte esencial del entrenamiento, la planificación y el desarrollo táctico en el deporte. A través de un modelo accesible, adaptable y progresivamente automatizado, este proyecto busca crear un puente entre el mundo profesional y las realidades del fútbol formativo y amateur.

8. Futuras líneas de investigación

Este trabajo de fin de máster es una propuesta que nace desde la pasión del grupo tanto en el área deportiva como la integración del análisis de datos en este sector. Por ende, FutHub no solo representa una propuesta técnica y conceptual para la aplicación del análisis de datos en el fútbol semiprofesional, sino que también constituye el primer paso hacia la construcción de un producto con proyección real en el mercado deportivo. Como se menciona anteriormente, durante los próximos meses desarrollaremos un mvp el cual tendrá como meta la estructuración de un modelo de negocio viable a través de la prueba y error con estructuras deportivas del mundo real. En adelante, se plantean diversas líneas de investigación y desarrollo que permitirán consolidar y escalar el proyecto.

Automatización mediante inteligencia artificial

Uno de los principales puntos a desarrollar en futuras etapas es el desarrollo del actual modelo para permitir que el ciclo completo sea automatizado. La principal propuesta de valor es crear una herramienta accesible para los equipos semiprofesionales para que sin necesidad de intervención manual se puedan generar análisis de los partidos. Esto implicaría investigar e implementar modelos avanzados de visión por computadora y aprendizaje automático capaces de:

- Detectar automáticamente eventos del juego completo.
- Identificar formaciones y posiciones de los jugadores a lo largo del partido.
- Generar visualizaciones tácticas personalizadas
- Adaptar los informes al estilo de juego y a los objetivos específicos de cada equipo.

Sin embargo, el desarrollo de esta propuesta se enfrenta a desafíos importantes a nivel tecnológico y operativo. Por un lado, el entrenamiento de modelos de aprendizaje automático requiere acceso a hardware especializado, como ordenadores de alto rendimiento o unidades de procesamiento gráfico (GPUs), recursos que no siempre están al alcance de startups en etapas tempranas. Además, la construcción de modelos fiables implica una considerable inversión de tiempo, ya que requiere la recopilación y etiquetado manual de grandes cantidades de datos, el diseño de arquitecturas algorítmicas, la iteración sobre resultados, y la validación en contextos reales de competición.

Por tanto, esta línea de investigación debe abordarse como un proceso progresivo, estructurado en fases. En una etapa inicial, se priorizarán modelos más simples que permitan automatizar tareas específicas, mientras que la automatización se proyectará como un objetivo a medio-largo plazo, a medida que se consoliden los recursos técnicos y humanos necesarios.

Expansión a otras disciplinas

Una de las direcciones más prometedoras consiste en adaptar la tecnología desarrollada para otros deportes, en los cuales también podríamos encontrar una oportunidad de negocio. Hemos encontrado que como factor común los deportes en general requieren análisis táctico, seguimiento de rendimiento y gestión técnica. Es por eso que en un futuro se ha analizado el desarrollo de una arquitectura modular del sistema desarrollado permitiría personalizar algoritmos y visualizaciones para contextos deportivos diversos.

De hecho, ya se ha establecido contacto preliminar con los responsables de una aplicación vinculada al mundo del pádel, quienes han mostrado interés en explorar la posibilidad de integrar un sistema automatizado de análisis de vídeo para este deporte. Esta colaboración futura no solo valida el interés comercial más allá del fútbol, sino que también abre una línea de investigación aplicada en deportes de raqueta, donde el análisis técnico es clave para el desarrollo de jugadores.

Anexo I - Notas Dashboard Analysis

El presente documento tiene como objetivo servir de base para la construcción del dashboard, y agradeceríamos contar con su experiencia y conocimientos durante este proceso. Tal como se mencionó en nuestra reunión, el dashboard estará estructurado en dos secciones principales:

1. Analítica General: proporcionará una visión global de la situación actual del club

dentro de la liga, junto con datos generales relevantes.

2. Analítica Específica: incluirá información detallada organizada en tres áreas clave:

- Preparación física
- Juego analítico
- Táctica analítica

Este documento es solo un punto de partida. Les solicitamos su apoyo para completar o sugerir los indicadores y datos que consideren más relevantes para cada sección. Su aporte será clave para que el dashboard sea realmente útil y representativo del trabajo del cuerpo técnico.

Preparación física: Indicadores relacionados con el estado físico de los jugadores

- Distancia total recorrida
- Distancia recorrida alta intensidad (>20 km/h aprox.)
- Distancia recorrida sprint (>25km/h aprox.)
- Ratio trabajo:descanso (Tomar referencia 15km/h aprox)
- Número sprints
- Número desaceleraciones (pérdida velocidad desde 20km/h)
- Velocidad Máxima alcanzada

Juego analítica: Datos objetivos y métricas que permiten evaluar el rendimiento del equipo y los jugadores:

- Incluiría la fase ABP (Acciones a balón parado)

Táctica analítica: Evaluación del comportamiento del equipo en situaciones tácticas específicas, como (4 fases):

- Momento Con Balón (MCB)
 1. Zonas de pases
 2. Pérdidas
 3. Pases para llegar a campo contrario
 4. Pases en campo contrario
 5. Centros al área
- Momento Sin Balón (MSB)
 1. Alturas líneas
 2. Distancia jugadores
 3. Distancia líneas
 4. Zonas de presión
- Transición Defensa- Ataque (TDA)
 1. Recuperación y zona
 2. Pases hacia adelante
 3. Pases hacia atrás
 4. Pases para llegar a área rival
- Transición Ataque- Defensa (TAD)
 1. Zona de la pérdida
 2. Metros corridos hasta finalizar defensa
 3. Jugadores que saltan a la presión tras perdida

9. Referencias

Análisis deportivo: Cómo utilizan el análisis de datos los distintos deportes. (2024, February 22). DataCamp. Retrieved June 18, 2025, from

<https://www.datacamp.com/es/blog/sports-analytics-how-different-sports-use-data-analysis>

B EYE | Data, AI & EPM Consulting. (2023, October 31). Sports Analytics: A Complete Handbook for Organizations. LinkedIn. Retrieved June 18, 2025, from

<https://www.linkedin.com/pulse/sports-analytics-complete-handbook-organizations-b-eye-ltd-s0lgf?>

Burgos, J. (2022, Febrero 19). Jordi Cruyff introdujo el método Zone7 a través del Getafe de Bordalás. Diario AS. Retrieved June 18, 2025, from

https://as.com/futbol/2022/02/19/primera/1645292194_314179.html#:~:text=EI%20Getafe%20de%20Bordal%C3%A1s%20a%20trav%C3%A9s%20de,inteligencia%20artificial%20aplicado%20al%20f%C3%BAtbol%20llamada%20Zone7

Carpio, G. (2023, marzo 29). Las principales plataformas proveedoras de datos de fútbol: ¿cuál es la mejor para ti? Futbol con datos. Retrieved junio 18, 2025, from

<https://futbolcondatos.com/las-principales-plataformas-proveedoras-de-datos-de-futbol-cual-es-la-mejor-para-ti/>

Catapult. (2025, January 8). Análisis de vídeo en el fútbol: Mejorar las Tácticas de Juego. Catapult. Retrieved June 18, 2025, from [https://www.catapult.com/es/blog/video-](https://www.catapult.com/es/blog/video-analysis-in-football)

[analysis-in-football](https://www.catapult.com/es/blog/video-analysis-in-football)

CONTRERAS, E. (2020, February 9). Así ficha Monchi: departamento I+D, data y algoritmo entre 18.000 futbolistas. MARCA. Retrieved June 18, 2025, from <https://www.marca.com/primer-plana/2020/02/08/5e3709e822601d8a5b8b458e.html>

Creasey, S. (2016, 16 de mayo). Foxy Leicester City FC won Premiership with data analytics. Computer Weekly. Recuperado el 18 de junio de 2025 de <https://www.computerweekly.com/news/450296302/Foxy-Leicester-City-FC-won-Premiership-with-data-analytics>

Datos de rendimiento de los futbolistas - FIFPRO World Players' Union. (n.d.). FIFPro. Retrieved June 18, 2025, from <https://fifpro.org/es/apoyar-a-los-y-las-futbolistas/competencias-innovacion-y-crecimiento/datos-de-rendimiento-de-los-futbolistas>

Díaz, M., & Ritchie, D. (2021, August 1). Big Data en el fútbol sin recursos: ¿cómo aplicarlo? Nacsport. Retrieved June 18, 2025, from <https://www.nacsport.com/blog/es-es/Tips/big-data-futbol>

Esden Business School. (2024, 01 08). ¿Qué es el Business Analytics y para qué sirve? Esden Business School. Retrieved 06 18, 2025, from <https://esden.es/que-es-business-analytics-y-para-que-sirve/>

Hudl Statsbomb Customer Case Study: Barrow AFC - Hudl Statsbomb | Data Champions. (2023, February 28). StatsBomb. Retrieved June 18, 2025, from <https://statsbomb.com/news/statsbomb-customer-case-study-barrow-afc/>

International Journal of Innovative Research in Technology. (2024). Strategic Edge [enfoque en modelos predictivos y Big Data en deporte]. IJIRT, 612–617. Recuperado de https://ijirt.org/publishedpaper/IJIRT168344_PAPER.pdf

La inteligencia artificial en el deporte. (2024, May 8). Tokio School. Retrieved June 18, 2025, from <https://www.tokioschool.com/noticias/inteligencia-artificial-deporte/>

López-Zafra, J. M. (2019, 05 07). Alcanzar un sueño. El Confidencial. Retrieved junio 18, 2025, from https://blogs.elconfidencial.com/economia/big-data/2019-05-07/leicester-city-metodologia-cientifica-lesiones_1984706/

Matplotlib Developers. (2025). Matplotlib: Visualization with Python. Recuperado el 18 de junio de 2025, de <https://matplotlib.org/>

mplsoccer. (2025). mplsoccer documentation. Recuperado el 18 de junio de 2025, de <https://mplsoccer.readthedocs.io/en/latest/>

Naciones Unidas. (s.f.-b). Educación de calidad. <https://www.un.org/sustainabledevelopment/es/education/>

Naciones Unidas. (s.f.-a). Salud y bienestar. <https://www.un.org/sustainabledevelopment/es/health/>

NoCodeHackers. (2025, junio 11). Lovable: Qué es, cómo funciona y por qué usarlo en 2025. NocodeHackers. Retrieved June 18, 2025, from <https://www.nocodehackers.es/herramientas-no-code/lovable>

Otto-Friedrich-Universität Bamberg. (2022). Case 01 – Premier-League [Documento de estudio de caso sobre la aplicación de TI en la PL]. Studocu. Recuperado el 18 de junio de 2025 de <https://www.studocu.com/de/document/otto-friedrich-universitat-bamberg/einfuehrung-in-die-wirtschaftsinformatik/case-01-premier-league/37806603>

Roboflow. (2023, septiembre 7). Computer Vision Models: What They Are and How They Work [Video]. YouTube. <https://www.youtube.com/watch?v=aBVGKoNZQUw>

Schoenfeld, B. (2019, May 29). El arma secreta del Liverpool: el análisis de datos. The New York Times. Retrieved June 18, 2025, from <https://www.nytimes.com/es/2019/05/29/espanol/liverpool-champions.html>

Seaborn: Statistical data visualization. (2025). Waskom, M. L. Recuperado el 18 de junio de 2025, de <https://seaborn.pydata.org/>

Sia Academy. (n.d.). Las diferencias entre el fútbol amateur y el profesional: una mirada a través de SIA Academy. International football academy in Spain. Retrieved June 18, 2025, from <https://soccerinteraction.com/es/diferencias-futbol-amateur-profesional>

Sports Data Campus. (2025, 25 de marzo). 6 software de videoanálisis de fútbol que no te puedes perder. Recuperado el 18 de junio de 2025, de <https://sportsdatacampus.com/los-6-software-de-videoanalisis-de-futbol/>

Tecnología vestible en el deporte. (2024, June 27). Catapult. Retrieved June 18, 2025, from <https://www.catapult.com/es/blog/tecnologia-portatil-en-los-deportes>

The data-driven enterprise of 2025. (2022, January 28). McKinsey. Retrieved June 18, 2025, from <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-data-driven-enterprise-of-2025>

Universidad Alfonso X el Sabio. (2025). Big Data en el fútbol: aplicaciones y beneficios. Blog de Deporte. Recuperado el 18 de junio de 2025, de UAX. <https://www.uax.com/blog/deporte/impacto-big-data-en-el-futbol>

Vidal-Codina, F., Evans, N., El Fakir, B. et al. Automatic event detection in football using tracking data. Sports Eng 25, 18 (2022). <https://doi.org/10.1007/s12283-022-00381-6>

Zheng, F., Al-Hamid, D. Z., Chong, P. H. J., Yang, C., & Li, X. J. (2025). A Review of Computer Vision Technology for Football Videos. Information, 16(5), 355. <https://doi.org/10.3390/info16050355>

10. Índice de figuras

10.1 Índice de tablas

Tabla 1: Fases del proyecto.....	38
Tabla 2: Variables del dataset (players_tracking).....	57
Tabla 3: Variables del dataset (ball_tracking)	57
Tabla 4: Comparación de competidores	89
Tabla 5: Modelo de ingresos.....	94

10.2 Índice de ilustraciones

Ilustración 1: Diagrama de Gantt	40
Ilustración 2: Detección de objetos	45
Ilustración 3: Asignar #ID a las detecciones	50
Ilustración 4: Algoritmo de clasificación.....	52
Ilustración 5: Detección del terreno de juego.....	54
Ilustración 6: Detección del campo después de filtrar los keypoints	55
Ilustración 7: Visualización 2D.....	56
Ilustración 8: Visualización de los pases.....	61
Ilustración 9: Visualización de las conducciones.....	62
Ilustración 10: Visualización de los tiros	63
Ilustración 11: Visualización de las intercepciones	64

Ilustración 12: Visualización terreno de juego (base).....	75
Ilustración 13: Dashboard (inicio).....	78
Ilustración 14: Estadísticas físicas.....	80
Ilustración 15: Estadísticas de juego (ofensivo).....	81
Ilustración 16: Estadísticas tácticas (DTA).....	83
Ilustración 17: Estadísticas jugadores	85
Ilustración 18: Calendario	86
Ilustración 19: Pizarra interactiva	87
Ilustración 21: Logotipo FutHub.....	88
Ilustración 22: Esquema de la funcionalidad	93
 10.3 Índice de ecuaciones	
Ecuación 1: Calculo de Intersection Over Union (IoU).....	49

11. Utilización de herramientas de inteligencia artificial en este trabajo

En la realización del trabajo, se han utilizado diferentes herramientas de inteligencia artificial con el fin tanto de asegurar la calidad del proyecto, como para la realización de diferentes validaciones. En este apartado vamos a detallar las diferentes herramientas que se han utilizado, así como la finalidad de las mismas y algunos de los prompts utilizados.

La principal herramienta utilizada ha sido ChatGpt, que se ha utilizado principalmente para entender algunos de los conceptos relacionados con computer vision, que requerían de alguna explicación extra y que no hemos encontrado en las fuentes consultadas. Un ejemplo de prompt utilizado ha sido “En un entorno de computer vision, qué es la homografía y que aplicabilidad tiene a la hora de proyectar un terreno de juego captado en vídeo”. Además, se ha utilizado para realizar validaciones en cuanto al código utilizado, con prompts como “Este es el código utilizado para la detección de intercepciones, revisa si la lógica para detectar el cambio de posesión se está realizando correctamente”

Por otro lado, se ha utilizado para el dashboard, se ha utilizado la herramienta Lovable, que nos ha permitido a través de diferentes prompts, construir una interfaz web, que alberga tanto los gráficos que queríamos representar así como un prototipo de plataforma del que hemos hablado en el apartado de Modelo de negocio. Algunos de los prompts utilizados serían del estilo “Este es el script json que contiene mis estadísticas ofensivas, incorporalos para que se utilicen en las diferentes visualizaciones de la pestaña Game Analytics” o “Cambia las variables del gráfico de radar por Tiros, Regates completados, Faltas cometidas, Asistencias y centros”

12. Anexos

Anexo I: Documento elaborado por el cuerpo técnico en el que se detallan los parámetros y métricas considerados relevantes para el estudio.

Anexo II: Conjunto de datos elaborado por el autor a partir de información obtenida de la plataforma Kaggle, utilizado para la creación del panel de visualización.

Anexo III: Segundo conjunto de datos de elaboración propia, generado a partir de datos extraídos de Kaggle, empleado en el desarrollo del dashboard.

Anexo IV: Tercer conjunto de datos de elaboración propia basado en información proveniente de Kaggle, utilizado en la construcción del panel de visualización.

Anexo V: Resultados generados tras el proceso de entrenamiento del modelo propuesto.

Anexo VI: Scripts desarrollados para el entrenamiento del modelo implementado.