

Trabajo Fin de Máster

Máster en Ingeniería Industrial

Desarrollo de una aplicación interactiva para el apoyo a la toma de decisiones para Direcciones Deportivas Data Driven en fútbol

Autor: Carlos Soria Polo

Tutor: Alicia Robles Velasco

Dpto. Organización Industrial y Gestión de Empresas II
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2023



Trabajo Fin de Máster
Máster en Ingeniería Industrial

Desarrollo de una aplicación interactiva para el apoyo a la toma de decisiones para Direcciones Deportivas Data Driven en fútbol

Autor:
Carlos Soria Polo

Tutor:
Alicia Robles Velasco
Profesor Ayudante Doctor

Dpto. Organización Industrial y Gestión de Empresas II
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla
Sevilla, 2023

Trabajo Fin de Máster: Desarrollo de una aplicación interactiva para el apoyo a la toma de decisiones para
Direcciones Deportivas Data Driven en fútbol

Autor: Carlos Soria Polo

Tutor: Alicia Robles Velasco

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

El Secretario del Tribunal

Sevilla, 2023

Agradecimientos

Aprovecho estas líneas para agradecer a todas las personas que han contribuido de alguna manera en la realización de este Trabajo Fin de Máster. Su apoyo moral y afectivo ha sido fundamental para llevar a cabo este proyecto con éxito.

Agradezco a mi familia por su apoyo incondicional y comprensión a lo largo de mi trayecto académico, así como por su constante estímulo y motivación para alcanzar mis metas.

A mis profesores y, en especial a mi tutora de trabajo, Alicia Robles Velasco, que merecen un reconocimiento especial por su orientación experta a lo largo de mi vida académica. Su experiencia y conocimientos han sido esenciales para mi desarrollo como estudiante y, sobre todo, como persona.

Además, quiero agradecer a mis amigos y compañeros de clase por su colaboración, ayuda y apoyo para afrontar todo lo que nos hemos encontrado por el camino.

Poder hacer este trabajo y seguir la línea de investigación empezada en el Trabajo Fin de Grado ha sido muy importante para mí, el fútbol me apasiona y poder vincularlo al ámbito más ingenieril y científico ha sido todo un reto que, sin duda, ha merecido la pena.

Carlos Soria Polo

Sevilla, 2023

Resumen

Este Trabajo Fin de Máster tiene como objetivo principal concretar la creación de una dirección deportiva en un club de fútbol que utilice los datos en su toma de decisiones con el fin de facilitar las actividades de organización y gestión del departamento, así como facilitar la labor de scouting en la Dirección Deportiva de clubes de fútbol, mediante la creación de una "Dirección Deportiva Data Driven".

El proyecto define el personal y los recursos necesarios, así como el cronograma de trabajo para la realización de planificaciones de plantilla en tiempo y forma.

Este enfoque busca la integración directa, gracias a una aplicación interactiva y programada en Python con la ayuda de la librería Streamlit de la interpretación de datos en las operaciones de planificación de plantillas, fusionando el conocimiento de los scouts con el análisis de datos.

Para dar un enfoque realista y veraz a la aplicación, se propone adicionalmente un caso de uso real de una alternativa de planificación de plantilla para una ventana de fichajes, en la que, teniendo en cuenta las restricciones pertinentes, se pondrá a prueba las habilidades de la aplicación para hallar propuestas de valor para el equipo.

Palabras Clave: Dirección Deportiva Data Driven, Machine Learning, Scouting, Análisis de Datos, Planificación de Plantillas, Aplicación Interactiva.

Abstract

The main objective of this Master's Thesis is to establish a sports management approach within a football club that leverages data for decision-making, streamlining departmental organization, management activities, and scouting tasks in sports management through the creation of a "Data-Driven Sports Management." The project outlines the required personnel, resources, and a work schedule for timely squad planning.

This approach seeks the direct integration of data interpretation into squad planning operations, facilitated by an interactive Python application developed using the Streamlit library. It bridges the gap between traditional scouting knowledge and data analysis.

To provide a realistic and practical dimension to the application, a real-use case is additionally proposed for squad planning within a transfer window, testing the application's ability to generate valuable proposals for the team while considering relevant constraints.

Keywords: Data-Driven Sports Management, Machine Learning, Scouting, Data Analysis, Squad Planning, Interactive Application.

Índice

Agradecimientos	vii
Resumen	ix
Abstract	xi
Índice	xiii
Índice de Tablas	xv
Índice de Figuras	xvii
1 Introducción	1
1.1. <i>Objetivo</i>	2
1.2. <i>Justificación de la continuación de la línea de desarrollo respecto a mi Trabajo Fin de Grado</i>	3
1.3. <i>Machine Learning</i>	4
1.2.1. Tipos de aprendizaje	6
2 Estado del arte sobre Direcciones Deportivas Data Driven	9
2.1 <i>Caso de éxito: Toulouse F.C.</i>	9
2.2 <i>Caso de éxito: Sevilla F.C.</i>	10
3 Confección de una Dirección Deportiva Data Driven	17
3.1. <i>Organización y costes asociados</i>	19
3.1.1 Personal clave	19
3.1.1.1 Determinación de número de profesionales necesarios por cada rol	21
3.1.1.2 Inversión salarial	23
3.1.2 Recursos necesarios	25
3.2. <i>Cronograma de planificación</i>	29
3.2.1. Fase I – Análisis de jugadores	30
3.2.1.1. Primer paso: Mapeo	31
3.2.1.2. Segundo paso: Cribado.....	32
3.2.2. Fase II – Definición de perfiles	34
3.2.2.1. Primer paso: Reunión con el entrenador.....	35

3.2.2.2.	Segundo paso: Selección.....	36
3.2.3.	Fase III – Confección de la plantilla	37
3.2.3.1.	Primer paso: Priorización	37
3.2.3.2.	Segundo paso: Negociación y compraventa de jugadores	37
4	Planificación de la plantilla.....	41
4.1.	<i>Lenguajes de programación y software utilizado</i>	<i>42</i>
4.2.	<i>Proceso ETL</i>	<i>48</i>
4.2.1.	Extracción de los datos	48
4.2.1.1.	Métricas relevantes.....	50
4.2.2.	Transformación de los datos	58
4.2.2.1.	Estandarización de variables	58
4.2.3.	Carga de los datos	60
4.3.	<i>Desarrollo de la aplicación con Streamlit</i>	<i>60</i>
4.3.1.	Pestaña 1: Inicio.....	61
4.3.2.	Pestaña 2: Asignación de Ligas.....	62
4.3.3.	Pestaña 3: Evaluación de los scouts.....	66
4.3.4.	Pestaña 4: Rating.....	75
4.3.5.	Pestaña 5: Clustering.....	86
4.4.	<i>Resultado</i>	<i>93</i>
4.4.1.	Desarrollo de la Fase I	93
4.4.2.	Desarrollo de la Fase II	95
4.4.2.1.	Generación de informe de necesidades	95
4.4.2.2.	Análisis del informe de necesidades en la aplicación	101
4.4.2.3.	Simulación de la planificación	103
4.4.2.4.	Generación de informes finales	110
4.4.2.5.	Planificación final.....	116
5	Conclusiones.....	119
	Referencias.....	121

ÍNDICE DE TABLAS

Tabla 1: Presupuesto I – Personal	25
Tabla 2: Presupuesto II - Recursos	29
Tabla 3: Relación entrada-salida de información y fechas de la Fase I	31
Tabla 4: Relación entrada-salida de información y fechas de la Fase II	34
Tabla 5: Relación entrada-salida de información y fechas de la Fase III	37
Tabla 6: Compra de jugadores Sevilla F.C. (Mercado de verano 2023/2024)	97
Tabla 7: Venta de jugadores Sevilla F.C. (Mercado de verano 2023/2024)	98
Tabla 8: Lista final de jugadores para la posición mediapunta	105
Tabla 9: Lista final de jugadores para la posición portero	106
Tabla 10: Lista final de jugadores para la posición central	106
Tabla 11: Lista final de jugadores para la posición lateral izquierdo	107
Tabla 12: Lista final de jugadores para la posición centrocampista (pivote)	107
Tabla 13: Lista final de jugadores para la posición centrocampista (box-to-box)	108
Tabla 14: Lista final de jugadores para la posición centrocampista (toque)	108
Tabla 15: Lista final de jugadores para la posición extremo izquierdo	109
Tabla 16: Lista final de jugadores para la posición delantero	109
Tabla 17: Planificación final en el Sevilla F.C. para el mercado de verano 2023/2024	117

ÍNDICE DE FIGURAS

Figura 1: Diferencia gráfica entre los tres tipos de aprendizaje en ML (Mathworks, s.f.)	8
Figura 2: Crecimiento de la imagen de marca del Sevilla F.C. (Sevilla F.C., 2021)	12
Figura 3: Máster en Big Data aplicado al Scouting en Fútbol (Sevilla FC, 2021)	13
Figura 4: Organigrama del Sevilla F.C. (Sevilla F.C., s.f.)	19
Figura 5: Visualizaciones de la aplicación Wyscout (Wyscout, s.f.)	26
Figura 6: Comparativa de rendimiento entre Karim Benzema y Harry Kane (OptaAnalyst, 2023)	27
Figura 7: Cronograma de las principales fases para la planificación de una plantilla (Fuente: Elaboración propia)	39
Figura 8: Esquema de procesos para realizar una planificación de plantilla (Fuente: Elaboración propia)	42
Figura 9: Logo oficial de Python	42
Figura 10: Logo oficial de Spyder	43
Figura 11: Logo oficial de Anaconda Navigator	43
Figura 12: Logo oficial de Streamlit	44
Figura 13: Logo oficial de Pandas	44
Figura 14: Logo oficial de Scikit-Learn	45
Figura 15: Logo de Matplotlib	45
Figura 16: Logo de NumPy	45
Figura 17: Logo de SQLite	46
Figura 18: Logo de BeautifulSoup	46
Figura 19: Logo de Jupyter Lab	47
Figura 20: Logo de Power BI	48
Figura 21: Logo de FBref	49
Figura 22: Esquema táctico 1-4-2-3-1 (Luque, 2013)	50

Figura 23: Cálculo de xG de Gabriel Jesús en la temporada 2019-2020 (Whitmore, 2021)	57
Figura 24: Distribución normal estándar (Disfrutalasmaticas, s.f.)	59
Figura 25: Visión general (Pestaña 1: Inicio)	61
Figura 26: Información Expanders (Pestaña 1: Inicio)	62
Figura 27: Visión general (Pestaña 2: Asignación de Ligas)	63
Figura 28: Selector de rol (Pestaña 2: Asignación de Ligas)	63
Figura 29: Contraseña incorrecta (Pestaña 2: Asignación de Ligas)	64
Figura 30: Contraseña correcta (Pestaña 2: Asignación de Ligas)	64
Figura 31: Ejemplo de reparto de ligas para scouts (Pestaña 2: Asignación de ligas)	65
Figura 32: Ejemplo de reparto de ligas para analistas de datos (Pestaña 2: Asignación de ligas)	65
Figura 33: Visión general (Pestaña 3: Evaluación Scout)	66
Figura 34: Opción 'Generar Nuevo Informe' (Pestaña 3: Evaluación Scout)	67
Figura 35: Ejemplo de sliders 'Perfil Físico' (Pestaña 3: Evaluación Scout)	72
Figura 36: Cuadro de texto para observaciones generales (Pestaña 3: Evaluación Scout)	73
Figura 37: Selectores de consulta de informes (Pestaña 3: Evaluación Scout)	73
Figura 38: Error de informe (Pestaña 3: Evaluación Scout)	73
Figura 39: Informe de Jon Moncayola por Scout 5 (Pestaña 3: Evaluación Scout)	74
Figura 40: Visión general (Pestaña 4: Rating)	75
Figura 41: Expanders de configuración (Pestaña 4: Rating)	75
Figura 42: Filtro de ligas (Pestaña 4: Rating)	76
Figura 43: Filtro de posición (Pestaña 4: Rating)	77
Figura 44: Filtro de edad (Pestaña 4: Rating)	77
Figura 45: Filtro de minutos (Pestaña 4: Rating)	78
Figura 46: Visualización de características relevantes (Pestaña 4: Rating)	84
Figura 47: Tabla final de Ratings (Pestaña 4: Rating)	85
Figura 48: Visión general (Pestaña 5: Clustering)	86
Figura 49: Opciones de personalización Machine Learning (Pestaña 5: Clustering)	87
Figura 50: Personalización nº1 Machine Learning (Pestaña 5: Clustering)	88

Figura 51: Personalización nº2 Machine Learning (Pestaña 5: Clustering)	88
Figura 52: Personalización nº3 Machine Learning (Pestaña 5: Clustering)	89
Figura 53: Opción nº4 Machine Learning (Pestaña 5: Clustering)	90
Figura 54: Opción nº5 Machine Learning (Pestaña 5: Clustering)	91
Figura 55: Resultado final en tabla (Pestaña 5: Clustering)	92
Figura 56: Resultado final en gráfico (Pestaña 5: Clustering)	92
Figura 57: Reparto de Ligas para Scouts (Resultado)	94
Figura 58: Reparto de Ligas para Analistas de Datos (Resultado)	94
Figura 59: Plantilla Sevilla F.C. (Temporada 22/23)	96
Figura 60: Priorización de características en la pestaña 'Rating'	102
Figura 61: Tabla final para el puesto mediapunta	103
Figura 62: Tabla final de jugadores para la posición mediapunta obtenido de la pestaña Clustering	104
Figura 63: Gráfico final de jugadores para la posición mediapunta obtenido de la pestaña Clustering	105
Figura 64: Información general de Folarin Balogun (Pestaña 1 del informe final en Power BI)	110
Figura 65: Estadísticas generales de Folarin Balogun (Pestaña 2 del informe final en Power BI)	111
Figura 66: Información (I) 'Evaluación Scout' (Pestaña 3 del informe final en Power BI)	112
Figura 67: Información (II) 'Evaluación Scout' (Pestaña 4 del informe final en Power BI)	112
Figura 68: Información (III) 'Evaluación Scout' (Pestaña 5 del informe final en Power BI)	113
Figura 69: Información 'Evaluación Scout' comparativa delanteros (Pestaña 5 del informe final en Power BI)	113
Figura 70: Representación de percentiles de Folarin Balogun (Pestaña 6 del informe final en Power BI)	114
Figura 71: Comparativa de rendimiento a través de la representación de los percentiles (Pestaña 6 del informe final en Power BI)	115

1 INTRODUCCIÓN

En los últimos tiempos, el análisis de datos se ha convertido en una herramienta fundamental en la modernización y digitalización de la industria. Este cambio de enfoque en la gestión empresarial está presente en una amplia gama de sectores, como finanzas, seguros, informática y videojuegos, entre otros. El hecho de que el mundo se encuentre inmerso en la era digital ha llevado a las empresas a generar una gran cantidad de datos a través de diversas fuentes, como aplicaciones móviles, dispositivos IoT y redes sociales.

El desarrollo técnico en el tratamiento de datos proporciona información valiosa en cuestiones que afectan a la operatividad de la empresa, como pueden ser el comportamiento y preferencias de los clientes, oportunidades de crecimiento o tendencias del mercado. Además de repercutir en una mejora en la eficiencia de los procesos, la implementación de estas técnicas puede incurrir en una reducción y, por ende, una mejor gestión de costes, gracias a la detección de ineficiencias o gastos superfluos.

La analítica de datos ha adquirido tanta relevancia en el desarrollo tecnológico actual que se ha convertido en uno de los campos con mayor demanda laboral. En este campo, existe una gran variedad de perfiles que, de forma combinada, ayudan a hacer más eficiente la extracción, transformación y posterior interpretación de los datos. Estos perfiles van desde los ingenieros de datos, encargados de diseñar el ecosistema de recolección de datos para las empresas; los científicos de datos, personas con un profundo conocimiento en matemática y estadística que codifican técnicas de inteligencia artificial y aprendizaje automático aplicadas a cada caso según convenga; hasta los analistas de datos, que interpretan los datos en el contexto empresarial y de negocios para llegar a conclusiones a partir de esta interpretación.

Los deportes y, más concretamente, el fútbol, que es la temática principal en este proyecto, no son ajenos al desarrollo tecnológico incesante que el mundo está viviendo. La digitalización y el uso de los datos está muy presente en el mundo del fútbol, desde el uso de redes sociales para generar ‘engagement’ (término en inglés que se refiere a crear vínculo emocional en redes sociales entre el usuario y la cuenta) hasta la analítica de datos aplicada a prevención de lesiones, previsión de demanda en las tiendas o, como es el caso de este proyecto, el estudio del mercado de compraventa de jugadores (en adelante, scouting, por su terminología en inglés).

En los clubes de fútbol, el departamento encargado de confeccionar la plantilla de jugadores, o lo que es lo mismo, de realizar el scouting, es la dirección deportiva. El scouting es el proceso por el cual los equipos, gracias al visionado de partidos o al análisis de datos, son capaces de encontrar jugadores que, potencialmente pueden ser objeto de adquisición por parte del equipo. Esta labor se debe hacer teniendo en cuenta una organización concreta del trabajo y adecuándose a las necesidades y capacidades económicas del club.

Esta idea es en la que incide este trabajo, conseguir implementar el uso de los datos en la operativa de la dirección deportiva, con la finalidad de facilitar y automatizar las labores de scouting. Para ello, el principal objetivo de este trabajo es confeccionar lo que se conoce como ‘Dirección Deportiva Data Driven’, que es un

tipo de dirección deportiva en fútbol que incorpora recursos necesarios en la analítica de datos para poder implementarla en la toma de decisión en la compraventa de jugadores.

1.1. Objetivo

Como se ha comentado en el epígrafe anterior, el objetivo general de este Trabajo Fin de Máster es facilitar las tareas de scouting en la Dirección Deportiva de clubes de fútbol, a partir de la confección de lo que se conoce en el entorno de los datos en fútbol como ‘Dirección Deportiva Data Driven’. Este tipo de dirección deportiva busca integrar, de manera directa, la interpretación de los datos en la operativa de las planificaciones de la plantilla del club, combinando el conocimiento de los scouts (visión más subjetiva y tradicional) con la analítica de datos (visión más objetiva y novedosa). Para poder integrar estos dos puntos de vista de manera orgánica, se necesita personal especializado en el mercado y en fútbol (scouts) y personal que tenga un punto de vista más técnico y tecnológico (analistas, científicos y ingenieros de datos).

Para su confección, los objetivos particulares a cumplir son los siguientes:

- 1) **Propuesta de organización de la ‘Dirección Deportiva Data Driven’:** En este punto, se buscará definir los miembros que deberían componerla. Para su buen funcionamiento, es necesario organizar la fuerza de trabajo del personal en un orden cronológico, para poder establecer un calendario de trabajo ordenado, organizado y con cierto grado de automatización, teniendo en cuenta las ventanas de mercado (invierno y, sobre todo, verano), y cualquier restricción temporal adicional. De este modo, quedaría definida la estructuración de la dirección deportiva.
- 2) **Extracción de datos:** Una vez definida de forma teórica la organización de la Dirección Deportiva Data Driven, lo primero que hay que definir es el conjunto de datos a utilizar para realizar la aplicación. En este conjunto de datos se recoge el rendimiento de una gran cantidad de jugadores de fútbol que, con el correspondiente tratamiento, van a ser utilizados para poder aplicar técnicas de Machine Learning para su posterior interpretación y análisis.

Es necesario contar con datos de alta calidad, lo cual puede resultar complicado de obtener, especialmente para clubes con recursos limitados que no pueden hacer frente a los altos costos de los proveedores de datos en el fútbol. Por lo tanto, es importante establecer pautas de trabajo y recolección de datos que tengan en cuenta la limitación económica para pagar las tarifas de los proveedores. En este TFM, se propone utilizar fuentes de datos gratuitas que, con el tratamiento adecuado, también pueden ser de gran utilidad.

- 3) **Creación de una aplicación para la planificación de la plantilla:** Se propone una herramienta para poder automatizar la organización y gestión de la Dirección Deportiva Data Driven y el análisis y búsqueda de jugadores objetivo en el mercado a partir de la analítica de datos. Este análisis de datos en las direcciones deportivas se realiza a partir de algoritmos basados en modelos de Machine Learning. Esta aplicación está diseñada para que todos los miembros de la dirección deportiva y del

cuerpo técnico que no tengan conocimientos matemáticos y estadísticos, sean capaces también de poder utilizarla e interpretarla de una manera cómoda y fácil. Con el fin de facilitar este análisis y gestión de los datos, se propone desarrollar una aplicación respaldada por algoritmos de Machine Learning implementados en Python.

- 4) **Caso práctico:** A partir de la organización definida de la ‘Dirección Deportiva Data Driven’ propuesta y de la aplicación desarrollada, se plantea, a partir de unas restricciones concretas de presupuesto, una propuesta de planificación de la plantilla del equipo acorde a los plazos definidos en el primer punto.
- 5) **Generación de informes:** Los jugadores objeto de adquisición por parte del club una vez finalizado el caso práctico, se generarán unos informes en los que se visualice información general de los jugadores, gráficos y visualizaciones obtenidos del uso de la aplicación y de la implementación de herramientas de visualización de datos y, por último, algunas interpretaciones relevantes a añadir a partir del análisis de datos.

De este modo, quedaría definido, de forma completa, tanto la gestión y organización de la ‘Dirección Deportiva Data Driven’, gracias al uso de una aplicación programada exclusivamente para automatizar el proceso de scouting en la dirección deportiva y, además, aportando un caso práctico real para testear la aplicabilidad de esta propuesta de dirección deportiva. El siguiente paso es entender cómo se articula la organización de la ‘Dirección Deportiva Data Driven’ teóricamente. Para ello, es necesario entender qué es el aprendizaje automático y, por tanto, qué es el Machine Learning.

1.2. Justificación de la continuación de la línea de desarrollo respecto a mi Trabajo Fin de Grado

El presente Trabajo Fin de Máster (en adelante, TFM) irá en la línea de desarrollo del Trabajo Fin de Grado (en adelante, TFG) que realicé para la finalización de los estudios del Grado en Ingeniería de las Tecnologías Industriales, titulado “Diseño y aplicación de técnicas de machine learning para optimizar el Scouting en clubes de fútbol”.

Dicho TFG incidía, en su mayor parte, en la recolección de datos a partir de una aplicación de estadísticas básicas de fútbol (Sofascore) y utilizaba algoritmos de clasificación basados en aprendizaje automático para segmentar el conjunto de futbolistas en función de las métricas de rendimiento de cada uno. Sin embargo, en dicho TFG no se hizo hincapié en cómo se debería estructurar un departamento específico encargado de este proceso de implementación de técnicas basadas en analíticas de datos, sino que simplemente se limitó a definir la herramienta de estudio de estos.

Es por ello por lo que se ha considerado de relevancia realizar un estudio de cómo, qué y cuánta inversión se necesitaría para constituir una Dirección Deportiva Data Driven para poder combinar la visualización y análisis de forma tradicional mediante la visualización de partidos, con el análisis de datos utilizando

herramientas de inteligencia artificial.

Además de definir la estructura y la cronología temporal que debe seguir una Dirección Deportiva Data Driven durante una planificación, se estudiará el proceso de recogida de datos. Debido al creciente uso de datos en el mundo del fútbol y a la democratización de estos, en este lapso de tiempo entre la realización del TFG y el TFM, se han introducido nuevas bases de datos, tanto de pago como gratuitas, más profesionales y con unas métricas de rendimiento más complejas, lo que permite hacer un estudio más intensivo en la actualidad de los jugadores.

Otro punto importante a considerar es que, para la construcción de una Dirección Deportiva Data Driven, sería conveniente disponer de una herramienta interactiva que lleve implementados los algoritmos de Machine Learning, además de otras funcionalidades. Esto es de gran importancia ya que las direcciones deportivas que basan su operativa, en parte, en la analítica de datos, disponen de herramientas informáticas para facilitar y automatizar el desarrollo, por lo que, realizar un estudio y desarrollo en este aspecto no realizado anteriormente, se acerca a lo que puede ser una gestión real de una dirección deportiva.

Por último, otra línea de desarrollo tiene que ver con la generación de informes de rendimiento, comúnmente usados por las direcciones deportivas para realizar las labores de scouting. Estos informes servirán, sobre todo, al Director Deportivo para poder decidir la planificación final a acometer, gracias a la información proporcionada por los miembros de la Dirección Deportiva Data Driven.

Por todas estas cuestiones, se ha visto conveniente realizar una continuación de la línea de desarrollo comenzada en el Trabajo Fin de Grado realizado, lo que proporcionará al lector una visión más amplia y realista de cómo se debe integrar el uso de los datos en un club de fútbol.

1.3. Machine Learning

Antes de entrar en definir los procesos necesarios para realizar scouting en un club de fútbol, es necesario entender cómo se sustentan teóricamente las técnicas de aprendizaje automático que dan base al estudio y análisis de los algoritmos utilizados, conociendo qué tipos de algoritmos se utilizan hoy en día, qué factores determinan la utilización de un algoritmo u otro y cuáles son los resultados que arroja cada tipo de algoritmo.

Dentro de la analítica de datos, se ha desarrollado en mayor medida el uso de técnicas de inteligencia artificial o el aprendizaje automático (en adelante, Machine Learning, por su denominación en inglés). Las técnicas de Machine Learning se han extendido a prácticamente todos los sectores empresariales que tengan una mínima perspectiva de expansión y crecimiento, transformando la forma en que abordan sus desafíos y optimizan sus operaciones.

Según (Amazon, s.f.), el Machine Learning (o aprendizaje automático en español) es la ciencia de desarrollo e implementación de algoritmos y modelos estadísticos que utilizan los sistemas de computación basándose en patrones observados en los conjuntos de datos. La idea principal del aprendizaje automático es

establecer una relación matemática o estadística entre el conjunto de datos de entrada (obtenido de forma manual o de forma automática a través de un proceso de extracción de datos que se comentará más adelante en este trabajo) y los datos de salida, sin conocer previamente la relación entre los conjuntos de datos. Estos algoritmos, junto a una gran capacidad computacional de los aparatos que se utilizan hoy en día, permiten generar resultados con una mayor precisión y con un gran potencial interpretativo de los mismos.

Según (Badillo, y otros, 2020), aunque los algoritmos de Machine Learning e inteligencia artificial se estén desarrollando en los últimos años, la idea de aprendizaje computerizado no es nueva y se rondaba alrededor de la década de los 50. Con la creciente capacidad de síntesis de datos y de potencia computacional, se ha conseguido numerosos avances en la computación en áreas de conocimiento muy diversas. Por ejemplo, en el artículo citado, se trata la utilización de algoritmos de aprendizaje automático aplicados a la farmacología clínica.

A continuación, se presentan algunos ejemplos destacados de la aplicación de estas tecnologías en diferentes ámbitos:

- 1) **Salud:** En este respecto, está el caso de Google Health. Según (Google, s.f.), Google ha desarrollado modelos y algoritmos para ayudar en la detección temprana de enfermedades como el cáncer a partir de la secuenciación de los genomas del ADN, asistir en la interpretación de imágenes médicas y proporcionar recomendaciones de tratamientos personalizados.
- 2) **Finanzas y servicios bancarios:** En el sector financiero, está el caso de PayPal. Según el artículo (Paypal, 2021), esta empresa utiliza técnicas de Machine Learning para detectar fraudes y actividades sospechosas en tiempo real, mejorar la precisión en la evaluación crediticia y brindar recomendaciones de inversión personalizadas.
- 3) **Comercio electrónico:** Según (Emerj, 2023), el consorcio de comercio electrónico chino Alibaba utiliza inteligencia artificial en varios aspectos de su estrategia operativa, como pueden ser:
 - Ofrecimiento de sugerencias de productos relevantes y adaptadas a los gustos y preferencias de cada usuario.
 - Mejora de la eficiencia de la cadena de suministro y la logística, optimizando rutas de entrega y agilizando los procesos de almacenamiento y distribución.
 - Desarrollo de asistentes virtuales y chatbots impulsados por IA para mejorar la atención al cliente.
 - En términos de seguridad, protección a sus usuarios de actividades fraudulentas y amenazas cibernéticas mediante soluciones como: detección de fraude, identificación facial y análisis de comportamiento para identificar patrones sospechosos y tomar medidas preventivas.
- 4) **Industria automovilística:** En este sector, destaca el desarrollo tecnológico de Tesla. En su propia página web (Tesla, s.f.), Tesla hace referencia a algunos de los desarrollos que están haciendo en este sentido:

- Uno de los elementos destacados, y que ya se está testeando en fases beta incluso aquí en Europa (estando más avanzado en Estados Unidos) es el Autopilot, un sistema avanzado de asistencia al conductor que emplea técnicas de IA y aprendizaje automático para proporcionar características de conducción semiautónoma. El Autopilot permite que los vehículos de Tesla controlen el acelerador, los frenos y la dirección de forma autónoma en ciertas condiciones, mejorando así la seguridad y la comodidad del conductor.
- Otro elemento, igualmente con mayor vigencia en Estados Unidos, es el Full Self-Driving (FSD por sus siglas en inglés). Esta funcionalidad, en continuo desarrollo, tiene como objetivo lograr la conducción completamente autónoma. Tesla utiliza técnicas de IA, redes neuronales y la implementación de sensores para entrenar y mejorar continuamente los algoritmos de FSD, permitiendo que los vehículos sean capaces de conducir de manera autónoma en cualquier entorno y situación de tráfico.
- Otro elemento importante es la gestión y optimización de la energía de la batería. Al ser vehículos 100% eléctricos, la capacidad de ser eficientes energéticamente es vital en el comportamiento del coche, ya que otorga mayor capacidad de desplazamiento sin realizar cargas adicionales y alarga la vida útil de la batería.

1.2.1. Tipos de aprendizaje

Según (IBM, s.f.), en el campo del Machine Learning, existen diferentes tipos de aprendizaje que permiten a los modelos y algoritmos adquirir conocimiento y realizar tareas específicas. A continuación, se describen los principales tipos de aprendizaje utilizados en el Machine Learning: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo.

- 1) **Aprendizaje supervisado:** El aprendizaje supervisado es uno de los sistemas más comunes en Machine Learning. En este tipo de aprendizaje, se proporciona a los algoritmos un conjunto de datos de entrenamiento que incluye ejemplos etiquetados con la respuesta correcta. El objetivo es que el modelo aprenda a establecer relaciones de las entradas a las salidas correctas y sea capaz de predecir la respuesta correcta para nuevas entradas. Se denominan sistemas de regresión a aquellos que se utilizan cuando la variable respuesta es cuantitativa; mientras los sistemas de clasificación se emplean cuando la variable respuesta es cualitativa. Algunos ejemplos de algoritmos supervisados incluyen regresión lineal, regresión logística, árboles de decisión, máquinas de vectores de soporte (SVM) y redes neuronales.

Un caso de uso de aprendizaje supervisado puede ser una clasificación de correos, en la cual el modelo etiqueta a los mensajes entrantes como spam o no spam. Se entrena al modelo con un conjunto de correos ya etiquetados como deseados o no deseados y, posteriormente, se dan entrada a nuevos correos para que el modelo los clasifique según a las etiquetas que se han recibido como entrada previamente.

- 2) Aprendizaje no supervisado:** A diferencia del aprendizaje supervisado, el aprendizaje no supervisado implica el análisis de datos que no están etiquetados ni categorizados previamente. El objetivo principal es encontrar patrones o formaciones de grupos no determinados explícitamente en los datos. Los algoritmos de aprendizaje no supervisado pueden utilizarse para la segmentación de datos, la reducción de dimensionalidad, la detección de anomalías y la recomendación de sistemas. Algunos ejemplos de algoritmos no supervisados son el análisis de componentes principales (PCA), el clustering y las redes neuronales autoencoder.

Un caso de uso de aprendizaje no supervisado es el que se utilizó para el Trabajo Fin de Grado que realicé (Soria Polo, 2021), en el cual se segmentaban o categorizaban jugadores por grupos (clústers) a partir de los datos de un conjunto de métricas relacionadas con el fútbol.

- 3) Aprendizaje por refuerzo:** Según (Mathworks, s.f.), el aprendizaje por refuerzo se basa en un enfoque diferente al de los aprendizajes supervisado y no supervisado. En este tipo de aprendizaje, un agente informático se encarga de realizar una tarea de forma repetitiva, recibiendo retroalimentación en forma de recompensas o castigos en función de las acciones realizadas. El objetivo del agente es aprender a tomar decisiones que maximicen la recompensa acumulada a lo largo del tiempo. Los algoritmos de aprendizaje por refuerzo se utilizan en aplicaciones como juegos, robótica y optimización de recursos.

El caso de uso más típico de aprendizaje por refuerzo es un modelo capaz de jugar al ajedrez. Como se ha comentado previamente, un agente interactúa con el modelo de manera que, en función de las acciones que toma en la experimentación, se recompensa o no al modelo, de manera que este aprende cuáles son las mejores jugadas gracias a la interacción con el propio agente, con el fin de maximizar la recompensa obtenida. En la Figura 1 se muestra un esquema con los principales tipos de sistemas de aprendizaje automático mencionados previamente.

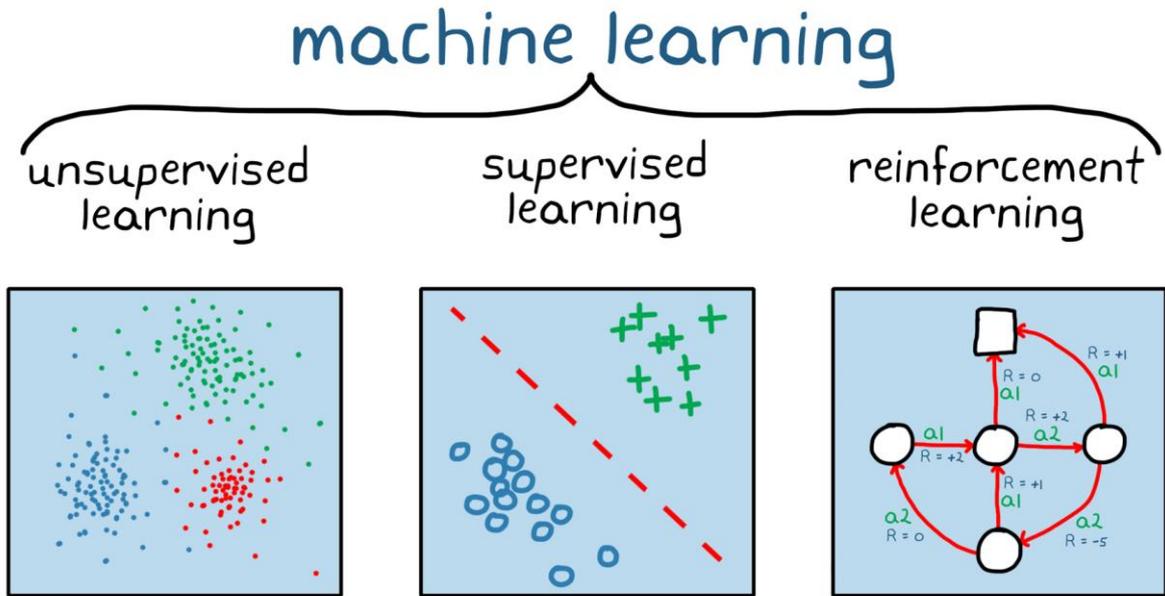


Figura 1: Diferencia gráfica entre los tres tipos de aprendizaje en ML (Mathworks, s.f.)

2 ESTADO DEL ARTE SOBRE DIRECCIONES DEPORTIVAS DATA DRIVEN

Hasta hace relativamente poco tiempo, esta práctica se ha implementado principalmente en clubes de élite que pueden permitirse económicamente estas novedades. Sin embargo, en la actualidad, debido a la democratización de los costes y la mayor accesibilidad de los datos en el mundo del fútbol, cada vez más, los clubes con menor capacidad económica están comenzando a implementar elementos de analítica de datos en la toma de decisiones en materia de adquisiciones y ventas.

En el pasado, las direcciones deportivas tradicionales se basaban únicamente en la experiencia de los scouts y del director deportivo para tomar decisiones. En la actualidad, las direcciones deportivas que utilizan el análisis de datos (conocidas como Direcciones Deportivas Data Driven) tienen a su disposición información objetiva y ajustada a la realidad, lista para ser analizada según los requerimientos de cada club. Esto enriquece la capacidad de los clubes para tomar decisiones más fundamentadas en la búsqueda, selección e incorporación de jugadores.

Esta información puede ayudar a los equipos a identificar las fortalezas y debilidades de los jugadores en estudio, así como descubrir nuevos talentos no considerados por los scouts en los visionados de partidos tradicionales. Además, las herramientas de análisis de datos también permiten estimar el valor de mercado de los jugadores, lo que abre la posibilidad de identificar oportunidades interesantes en el mercado para el club.

Los equipos de fútbol y, particularmente, las direcciones deportivas han experimentado una evolución significativa en la implementación de avances tecnológicos en los últimos años. Uno de los ámbitos en los que esta evolución ha destacado con mayor intensidad es en el scouting, una función vital para identificar y seleccionar talentos emergentes.

En la actualidad, ya no solo es suficiente con el visionado de partidos de fútbol por parte de los scouts como habitualmente se ha hecho en las direcciones deportivas hasta hace poco tiempo, sino que también es necesario recolectar y tratar datos para dotar de mayor objetividad al análisis. Gracias a estos datos, se pueden realizar procesos automáticos de detección de talento tempranas en base a las necesidades concretas de un club, necesidades concretas o debilidades del equipo en base a los datos, entre otros.

Como ejemplificación del desarrollo de las técnicas de Machine Learning en los clubes de fútbol, se presentan dos casos de éxito de clubes que han conseguido materializar la inversión en analítica de datos en éxitos deportivos.

2.1 Caso de éxito: Toulouse F.C.

El primer caso de éxito trata sobre el Toulouse F.C. de la liga francesa (Ligue 1). Según lo comentado

por (Razablan, 2020), este equipo, en vez de usar proveedores tradicionales de datos (Opta, Instat, Wyscout, etc.), hizo uso de la base de datos del videojuego Football Manager. Esto, en cierta manera, es una estrategia de diferenciación clara con respecto a otros clubes, ya que se busca una fuente de datos distinta para poder encontrar talento. Teniendo en cuenta que en el momento de implementación estaban en Segunda División francesa (Ligue 2), necesitaban una fuente de información asequible y fidedigna para poder implementar estas técnicas sin acometer una gran inversión.

Es por ello por lo que Damien Comolli, en ese momento presidente del Toulouse, decidió apostar por esta estrategia para la adquisición de jugadores. En unas declaraciones de 2020 que recoge el Diario AS (Onrubia, 2022), Comolli decía lo siguiente: "Por ejemplo, como ellos (Football Manager) tienen miles de personas introduciendo datos, tienen información en su base de datos sobre el historial de lesiones de los jugadores que sólo ellos tienen, en cierto modo. Es algo que nos interesa enormemente porque antes de contratar a un jugador, lo primero que miramos es el número de minutos que ha estado en el campo. Y si se trata de un jugador que ha estado pocos minutos sobre el terreno de juego, analizamos el motivo: ¿es una decisión del entrenador o es una lesión? Nos remontamos a cinco u ocho años atrás para ver el historial de lesiones antes de traerlo, y a veces eso hace que lo descartemos, aunque sea un jugador muy bueno. Sabemos que la base de datos del historial de lesiones de los jugadores en Football Manager es única, así que probablemente recurramos a ella".

En ese mercado de fichajes estival, se ficharon a jugadores de muchas ligas y nacionalidades distintas, solo por el hecho de que se había determinado, a partir de los datos, que los jugadores a fichar eran idóneos para el estilo de juego del entrenador, Philippe Montanier. Esta nueva estrategia se llevó a cabo en un momento del club muy delicado, acababan de descender de la primera división francesa (Ligue 1) con 13 puntos, una puntuación irrisoria. Posterior a ese descenso, estuvieron dos temporadas en Ligue 2, ascendiendo a Ligue 1 al segundo año, ganando la liga.

En el presente año (2023), siendo su primer año de vuelta en Ligue 1 no solo están matemáticamente salvados, sino que han conseguido ganar la Copa de Francia (1-5 en la final contra el Nantes) y, consecuentemente, conseguir la clasificación para la UEFA Europa League, en la que no competían desde el año 2010.

2.2 Caso de éxito: Sevilla F.C.

Otro club que desde hace mucho tiempo viene innovando en todas las áreas de trabajo del club y, particularmente, en la dirección deportiva, es el Sevilla F.C. A raíz de la cosecha de éxitos desde 2006 hasta la actualidad, el club ha mejorado su situación deportiva y, por ende, económica. Esta mejora en su situación económica ha llevado al equipo a tener capacidad para acometer fichajes más costosos y de invertir en nuevas necesidades o novedades que hay en el panorama actual.

Esta necesidad de inversión en nuevas tecnologías propicia que el Sevilla F.C. y, particularmente, el Consejo de Administración del club, decida invertir en la generación de proyectos que haga al Sevilla F.C. un

club más digitalizado y con capacidad de aumentar su rendimiento deportivo y económico de la mano de los datos. Algunos de los proyectos más relevantes llevados a cabo por el club, desde el año 2019, son:

1) Proyecto 'Business Intelligence Center'

Según (La Liga, 2020), la función de esta plataforma se basa en integrar todas las interacciones con los aficionados, recopilando datos como compras de alimentos y bebidas en el estadio, compras en la tienda física del club y adquisición de entradas a través de la web. Estos datos se utilizan para elaborar perfiles de los aficionados y ofrecerles un trato personalizado para satisfacer sus necesidades y que esto reporte un beneficio al club. Para la creación de esta plataforma, el proyecto se ha dividido en fases:

- a) Fase de análisis de datos existentes y puntos de contacto, entre los cuales se encuentran la página web del club, e-commerce, el Wi-Fi en las tiendas, el área de promociones y el área digital para socios y abonados.
- b) En una etapa posterior, se añadirán la aplicación móvil en desarrollo, los bares del estadio y las tiendas físicas como puntos adicionales para recopilar información.

El director de Marketing del club en ese momento (2020), Ramón Loarte, acerca de la digitalización del club, comentó: "Ha sido difícilísimo, pero ha enriquecido a todo el club, desde el área de taquillas, área social, IT (Tecnología Informática), estrategia digital o nuestro oficial de protección de datos, todos han terminado la primera fase de manera muy satisfactoria por ver cómo ha evolucionado todo".

A través de campañas como el caso del Black Friday, mencionado en el artículo, Loarte comenta que el club ha conseguido triplicar las ventas online respecto al año anterior, gracias a la segmentación y particularización en las recomendaciones hechas a los aficionados.

2) Proyecto Brand Center

Según (Aggylon BrandCenter, s.f.), el Sevilla F.C. buscaba generar una imagen de marca para redefinir la estrategia y crear un entorno visual acorde a la idiosincrasia y valores del club, de la ciudad y de sus aficionados. Sin embargo, después de completar el rebranding, surge el desafío de gestionar los nuevos activos y percepciones de la marca.

Para coordinar esto, el Sevilla F.C., en colaboración con la consultora de branding Summa y la empresa Aggylon, creó un proyecto denominado Brand Center, cuya función principal es la clasificación, organización y sincronización de los equipos de trabajo para adecuarse a la nueva imagen de marca. Esto es importante para reunir a todos los colaboradores internos y externos, de manera que se asegure que todos los elementos visuales de la marca, tanto físicos como digitales, estén actualizados y sean coherentes con la imagen renovada.

En el sector del fútbol, donde la pasión juega un papel importante, el sentimiento de pertenencia a los colores del club es clave. Es necesario que el club esté presente donde están sus aficionados, en los

momentos necesarios y con un mensaje coherente pero adaptado. La gestión global de la marca será fundamental para lograr el éxito, según señaló el director general del Sevilla FC en la presentación de la nueva identidad de marca.



Figura 2: Crecimiento de la imagen de marca del Sevilla F.C. (Sevilla F.C., 2021)

3) Proyecto ‘Sevilla F.C. Innovation Center’

Esta plataforma aúna todas las propuestas de innovación que el club ofrece a distintas entidades que buscan colaborar en proyectos relacionados con el ámbito futbolístico de cara al desarrollo de estudios que consigan incentivar y divulgar el conocimiento en este aspecto y el emprendimiento entre los jóvenes.

Algunos ejemplos podrían ser:

- Cátedra Universitaria de la Universidad Pablo de Olavide de Sevilla, centrada, mayormente, en la investigación en medicina y ciencias del deporte.
- Creación del Máster en Big Data aplicado al Scouting en Fútbol en colaboración con la Universidad Católica de Murcia (UCAM).



Figura 3: Máster en Big Data aplicado al Scouting en Fútbol (Sevilla FC, 2021)

4) Creación del Departamento I+D+i

Según este artículo, (Contreras, 2020), este departamento creado en 2019 y anexo a la dirección deportiva, se encarga de dar soporte y complementar el trabajo de los miembros de esta área, en tres vertientes:

- a) **Scouting:** Este departamento, en este aspecto, ayuda a realizar un seguimiento de todos los jugadores con potencial para ser adquiridos por el club (en total, unos 18000 aproximadamente), filtrando aquellos jugadores que, en base a unos parámetros ajustados a las necesidades del club en cada una de las posiciones del campo, cumplan con los requerimientos del entrenador. Sin la ayuda del dato, para los scouts, sería imposible analizarlos a todos, lo que llevaría a, posiblemente, dejar sin analizar a algunos jugadores que pueden resultar de interés al club.

Además, en un análisis posterior cuando ya se ha realizado un primer filtrado, el Departamento de I+D+i complementa el análisis subjetivo por parte del scout con un análisis objetivo en base a los datos que posee el club, otorgando así otro punto de vista para tener en cuenta en los informes de cada uno de los jugadores, dotándolos de mayor objetividad y certitud.

- b) **Prevención de lesiones:** A partir de los datos recogidos de minutos jugados, el desgaste de cada jugador, el desempeño en los entrenamientos y la gestión de esfuerzos, este departamento otorga información a la secretaría técnica para evitar potenciales lesiones que el algoritmo puede predecir.
- c) **Valor de mercado:** Gracias a los datos, el club puede valorar tanto a los jugadores que son

objeto de adquisición por parte del club como jugadores que pueden salir, de tal manera que se pueda valorar de una forma objetiva cada una de las ofertas tanto realizadas por el club de cara a la compra de jugadores como a la venta de jugadores a otros clubes, identificando oportunidades de mercado.

5) Creación del Departamento de Datos

Este departamento va en consonancia con el Departamento de I+D+i del Sevilla F.C., con la principal salvedad de que el Departamento de Datos no solo da soporte a la dirección deportiva, sino que da soporte a todas las áreas del club, siendo un departamento completamente transversal a todas las áreas.

El nacimiento de este departamento surgió a partir del interés del Sevilla F.C. por desarrollar una aplicación completamente propia para el club de cara a facilitar la labor de búsqueda y detección de talento. Para ello, buscó la colaboración del Doctor en Física Elías Zamora, que por aquel entonces trabajaba en otra empresa, para poder desarrollar dicha aplicación basada en algoritmos de inteligencia artificial y Machine Learning. Fruto de esta colaboración, nació la aplicación AIFootball.

El éxito de esta aplicación llevó al Sevilla a la creación del Departamento de Datos, liderado por Zamora, compuesto por un equipo multidisciplinario de ingenieros informáticos, físicos y matemáticos. Para poder aplicar algoritmos basados en modelos matemáticos y estadísticos especializados para cada área del club, y según menciona (Rocha, 2023), el Departamento de Datos creó varias aplicaciones de gestión de datos para cada una de las áreas de gestión de la entidad, como son:

- **AIFootball:** Como se ha comentado previamente, es la aplicación que llevó al Sevilla F.C. a crear un departamento específico y especializado en la gestión de datos. Esta aplicación tiene tres funcionalidades principales:
 - a) Complementar el punto de vista subjetivo del scout con una valoración objetiva a partir de la recolección de datos y su posterior interpretación, enriqueciendo y dotando de mayor certitud al informe de valoración de cada uno de los futbolistas, de cara a reducir el riesgo en las incorporaciones de futbolistas del club.

Este estudio de los datos lo realiza el Departamento de I+D+i una vez se ha realizado el primer filtrado por parte de los scouts, en el cual ya se cuenta con una lista de jugadores de unos 25-30 por puesto. Para este estudio de datos, el Departamento de I+D+i caracteriza, en función de las métricas, al jugador y al entrenador que esté en ese momento en el club, decidiendo qué jugadores de la lista aportada son los más adecuados para el estilo de juego del entrenador y, por ende, del equipo.
 - b) Calcular el valor de mercado de los jugadores objeto de adquisición del club y de los jugadores que pertenecen al club, de cara a la objetivación del cálculo de este valor, detectando así ofertas fuera de mercado tanto en compras como en ventas de jugadores.

- c) Prevenir posibles lesiones, haciendo caso a los históricos de lesiones de los jugadores, gestión de esfuerzos y minutos, rendimiento en los entrenamientos, entre otros.
- **AIScouting:** A partir del estudio y la extracción de los datos de todos los informes que los scouts han generado en los últimos 10 años aproximadamente, esta herramienta es capaz de determinar cuándo y dónde ver a un determinado jugador del cual se necesita un informe pormenorizado y actualizado, y, además, quién es el técnico que va a ir al estadio a ver a dicho jugador, en función de los informes que haya realizado previamente. De tal manera, se automatizan y se mejoran los procesos de scouting de los jugadores, haciendo esta gestión de forma automática en vez de coordinada por parte del jefe del departamento de scouting (Emilio de Dios).
- **AIRadar:** Se centra en el descubrimiento de talento prematuro, permitiendo, mediante la activación de alertas en función de determinados parámetros, a los técnicos de la dirección deportiva, el detectar de forma rápida y efectiva a partir de los datos, el talento para su posterior análisis por parte del departamento competente.
- **AIFans:** Según (Sevilla FC, 2023), a partir de la colaboración del Sevilla con la empresa DruID resultó la creación de AIFans, que es una herramienta de segmentación de aficionados basada en datos. Utilizando los datos de aproximadamente 220.000 aficionados que han otorgado permisos para el manejo de su información, el club, con la creación de esta aplicación, busca comprender mejor los intereses de los aficionados y ofrecerles productos personalizados. Esta iniciativa demuestra una evolución significativa en la forma en que el club interactúa con sus seguidores en comparación con una segmentación más básica utilizada en el pasado.
- **AITicketing:** Permite un análisis profundo de los patrones de asistencia y la predicción de asistencia a los próximos partidos.
- **Transfer Tracker:** En los traspasos de los jugadores, el club de formación posee derechos de traspasos de los jugadores. En muchos casos, es tan grande la cantidad de jugadores que el club forma que es imposible realizar un seguimiento de todos y cada uno de ellos y de sus correspondientes movimientos en el mercado. Es por ello que, y según la información recogida en (Sevilla FC, 2022), el Sevilla F.C., de la mano de la alianza con LaLiga Tech (sección de desarrollo tecnológico de La Liga), se creó esta aplicación, cuyo principal objetivo es rastrear los mecanismos de solidaridad en los traspasos de jugadores, generando beneficios económicos para el club. Según esta noticia, el club ha conseguido, en un año, generar una plusvalía cercana al millón de euros.

Estas aplicaciones, aparte de servir de ayuda al Sevilla F.C. de cara a la automatización de procesos, sirve también para abrir una nueva línea de negocio, estableciendo un servicio B2B. Este tipo de servicios, básicamente, tienen como principal objetivo que otros clubes puedan hacer uso de estas aplicaciones en su propio

organigrama, repercutiendo en un beneficio económico para el club vendedor.

3 CONFECCIÓN DE UNA DIRECCIÓN DEPORTIVA DATA DRIVEN

Una dirección deportiva en un club de fútbol es un componente fundamental de su estructura organizativa. Como se ha comentado a lo largo de la sección introductoria del proyecto, la dirección deportiva es el área de gestión y planificación de todas las cuestiones relacionadas con el ámbito deportivo del club, con el objetivo de asegurar su éxito a largo plazo. La principal función de una dirección deportiva es diseñar y ejecutar una estrategia deportiva integral para el club. Esto implica la planificación a largo plazo de todas las áreas relacionadas con el fútbol, centrándose en dos aspectos fundamentales:

- **Planificación de la plantilla:** La dirección deportiva tiene la responsabilidad de identificar y fichar a jugadores que se ajusten a las necesidades del club, realizando un seguimiento y análisis exhaustivo del mercado, tanto a nivel nacional como internacional. Los directores deportivos y los ojeadores que trabajan en el departamento (scouts, en inglés) evalúan a los jugadores y entrenadores mediante la observación en partidos, la revisión de estadísticas y el análisis de su rendimiento general. Además, se encargan de las negociaciones y contrataciones teniendo en cuenta aspectos como el presupuesto del club y las necesidades específicas del entrenador y de la plantilla.

Además, la dirección deportiva se encarga de establecer y mantener relaciones con otros clubes, agentes y organismos relacionados con el fútbol. Esto implica negociar traspasos de jugadores, establecer acuerdos de colaboración y participar en redes de intercambio de información y conocimientos.

- **Coordinación de categorías inferiores:** Otra función crucial de la dirección deportiva es la de desarrollar y coordinar la formación de talento joven. Esto implica la creación de programas de captación de talentos, la supervisión de las categorías inferiores y la promoción de jugadores jóvenes hacia el primer equipo. Los directores deportivos buscan maximizar el potencial de los jóvenes talentos, brindándoles las oportunidades y el apoyo necesarios para su desarrollo, con el fin de hacerlos promocionar hasta llegar al primer equipo. Invertir y trabajar en cantera es importantísimo para la dirección deportiva y para el club en general, en los siguientes aspectos:
 - o La inversión para que un futbolista llegue al primer equipo desde las categorías inferiores, comparativamente con la inversión para adquirir a un futbolista de otro equipo es netamente inferior, por lo que la promoción de jugadores al primer equipo desde cantera es mucho más beneficioso para el club económicamente hablando.
 - o Los jugadores de cantera conocen la idiosincrasia del club desde muy pequeños, facilitando la labor de adaptación al primer equipo y creando mayor sentimiento de pertenencia en la primera plantilla.

- La dirección deportiva y los entrenadores en categorías inferiores pueden ir trabajando individualmente con cada futbolista, atendiendo a sus necesidades concretas de acuerdo con los intereses del club, para poder hacer una carrera de largo plazo y prepararlo para el salto al primer equipo desde una edad muy temprana.

Estas dos funciones son comunes a todas las direcciones deportivas en todos los clubes de fútbol, pero no es menos cierto que hay clubes en los que la dirección deportiva tiene un mayor protagonismo que en otros. A continuación, se comentan los diferentes tipos de gestión deportiva en los clubes, según lo citado en (Soria Polo, 2021):

- a) **Modelo presidencialista:** Es un modelo que gira entorno al presidente de la entidad, que, en colaboración y con ayuda de otras áreas del club, es el que tiene la última palabra en las decisiones más relevantes del club, entre las que se incluye la adquisición y venta de futbolistas. En este sentido, la dirección deportiva se encarga de analizar el mercado para encontrar futbolistas objeto de adquisición por parte del club, de presentar opciones al presidente y de coordinar las negociaciones con el club vendedor y sus agentes e intermediarios. Como ejemplo principal, aquí se puede destacar el Real Madrid C.F. de Florentino Pérez.
- b) **Modelo anglosajón:** Este modelo da todo el peso de la confección de la plantilla al entrenador (manager, en inglés), determinando cuáles son los perfiles necesarios a cubrir y los jugadores a fichar, así como las ventas que tiene que acometer el club. Este modelo se sigue en la Premier League inglesa, que es una liga que destaca, aunque cada vez menos, por su total fidelidad y confianza hacia los entrenadores, aunque el rendimiento sea malo. El hecho de que cada vez haya menos confianza en los entrenadores puede deberse a la europeización del fútbol inglés, dejando al entrenador en un papel más secundario y prescindiendo de sus servicios de una manera más rápida en caso de que el club no esté consiguiendo resultados. Como ejemplos (pasados, al ser una filosofía en desuso, como se ha comentado) están el Arsenal F.C. de Arsène Wenger o el Manchester United F.C. de Alex Ferguson.
- c) **Modelo mixto:** Es el modelo más utilizado en la élite futbolística europea y que predomina en la mayoría de los clubes. En este modelo, la gestión del club se divide en varios responsables. Por una parte, está el Consejo de Administración del club, con el presidente a la cabeza, que se encarga de gestionar la parte económica del club y de decidir la inversión y ventas de jugadores, pero nunca de decidir quién se ficha o se vende. Esta última función corresponde a la dirección deportiva, cuyo principal objetivo es proponer jugadores en función a las necesidades y restricciones del club y de la plantilla, consensuadas tanto con el cuerpo técnico como con el Consejo de Administración de la Entidad. A continuación, en la Figura 4, se muestra un ejemplo de organigrama del Sevilla F.C., en el cual se puede ver la composición de cada departamento y, en particular, de la dirección deportiva.

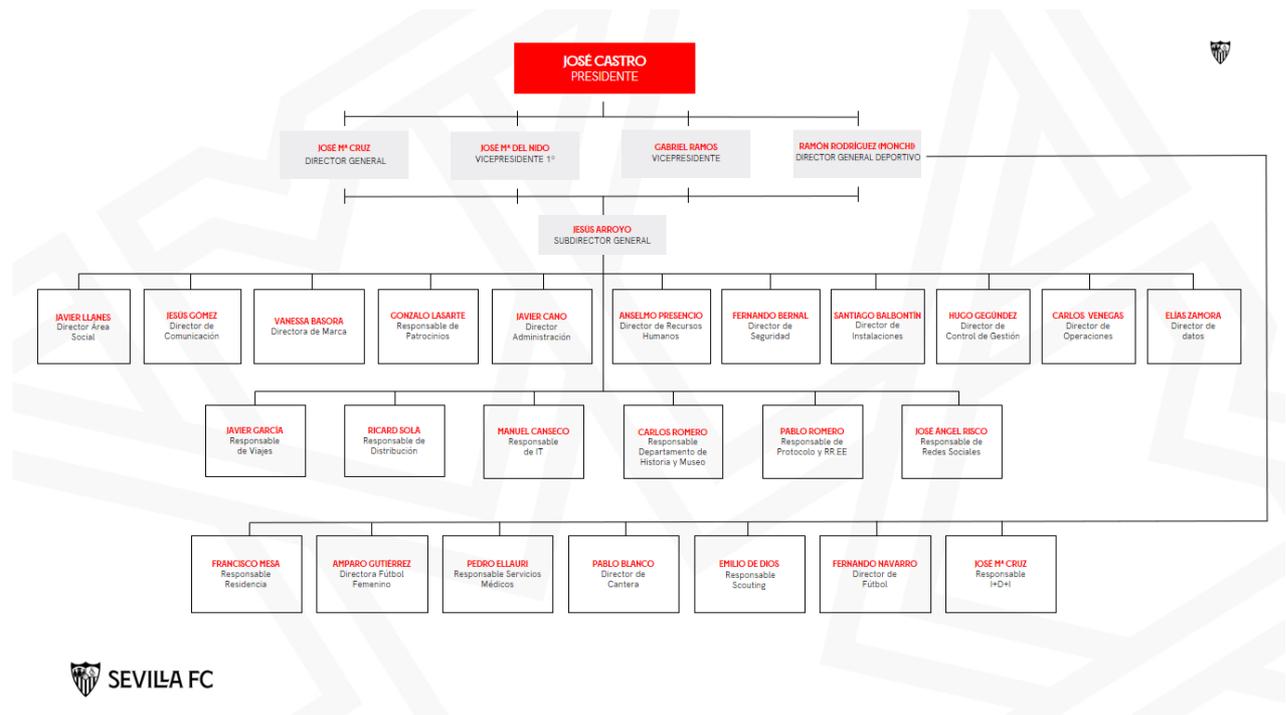


Figura 4: Organigrama del Sevilla F.C. (Sevilla F.C., s.f.)

3.1. Organización y costes asociados

Para establecer una Dirección Deportiva Data Driven en fútbol, es esencial contar con el personal y los recursos adecuados añadidos a los que una dirección deportiva tradicional tiene. En esta sección, se describirán los roles clave y los elementos necesarios para que esta dirección funcione correctamente, añadiendo adicionalmente unas estimaciones de inversión que supone la implementación de este tipo de dirección deportiva.

3.1.1 Personal clave

- **Director deportivo:** Es el encargado, al igual que en las direcciones deportivas tradicionales, de liderar la dirección deportiva y de tomar las decisiones finales referentes a qué perfil de jugadores y, posteriormente, qué jugadores concretos se van a adquirir y, por otro lado, qué jugadores se van a vender. Además, debe también coordinar las labores de cantera, como se ha comentado en epígrafes anteriores. Como miembro más importante y representativo de la dirección deportiva, área estratégica clave en el desarrollo de la entidad, debe ser una persona capaz de desarrollarse en un entorno multidisciplinar, teniendo conocimientos en los siguientes aspectos:
 - o Comprensión del fútbol muy alta, al nivel de los analistas de fútbol y scouts.
 - o Nociones básicas de matemática y estadística, para entender el trasfondo de los algoritmos aplicados en la confección de listas por parte de los analistas, científicos e ingenieros de datos.

- Buenas capacidades de negociación con agentes de futbolistas y clubes, para poder defender los intereses del club.
- Conocimiento del mercado, para poder detectar posibles oportunidades y precios inflados e irreales de cara a una adquisición.
- **Ayudante del director deportivo:** Estos trabajadores se encargan de dar soporte en sus labores al director deportivo, en las funciones que requiera este. Es labor del director deportivo determinar las tareas que quiere delegar en el ayudante de cara a disminuir la carga de trabajo.
- **Scouts:** Estos profesionales se encargan de identificar y evaluar a los jugadores potenciales, realizando visualizaciones de partidos de las diferentes ligas de todo el mundo que la dirección deportiva precisa analizar. Su trabajo en el seguimiento de partidos para encontrar talentos que se ajusten a las necesidades y requerimientos del club.
- **Expertos en tratamiento de datos:** En esta cuestión es donde se da la principal diferencia entre una dirección deportiva clásica, compuesta únicamente por scouts, con respecto a una Dirección Deportiva Data Driven, en la que también se precisan de técnicos especialistas en datos. Estos profesionales están enfocados en el análisis de datos, aplicando algoritmos de Machine Learning e inteligencia artificial para poder extraer conclusiones de los datos recogidos.

Antes de determinar los miembros necesarios para el desarrollo de la analítica de datos dentro de una Dirección Deportiva Data Driven, es importante definir cuáles son los diferentes tipos de perfiles que pueden trabajar en un club de fútbol en general, que aplique este tipo de técnicas en su operativa. Posteriormente, se determinará cuáles de estos perfiles, en función de los requerimientos concretos de la dirección deportiva, serán necesarios para poder realizar las labores de scouting aplicando analítica de datos:

- **Jefe de datos (Chief Data Officer):** Coordinador general de los expertos en el tratamiento de datos, siempre bajo el mando y las pautas definidas por el director deportivo, máximo responsable de esta área. Se encargará de ser el nexo de unión entre el director deportivo y los scouts con los expertos en datos, para que la coordinación de los trabajos sea más efectiva. Además, será el encargado de contactar con el director deportivo para que este le comunique los puntos más importantes extraídos de la reunión con el entrenador previa al mercado de fichajes (esta cuestión se particularizará luego). Este puesto debe estar ocupado por una persona con gran experiencia tanto en áreas puramente de negocio como en aspectos más técnicos relacionados con matemáticas y estadística.
- **Ingeniero de datos (Data Engineer):** Miembro encargado de desarrollar los sistemas que permitan canalizar el gran volumen de datos asumidos por la dirección deportiva a partir de proveedores externos, así como el limpiado de aquellos datos que no son de utilidad. Este rol lo deben desempeñar programadores e ingenieros informáticos ya que tienen una gran

capacidad de diseño de bases de datos y de interfaces de usuario capaces de integrar todas las funciones que determine el Chief Data Officer.

- **Científico de datos (Data Scientist):** Miembro encargado de diseñar los algoritmos de interpretación de los datos proporcionados. Estos algoritmos, apoyados en tecnología de Inteligencia Artificial, Machine Learning y redes neuronales, están basados en principios matemáticos, estadísticos y de programación y, además, atienden a las necesidades concretas determinadas por el Chief Data Officer como interesantes para el modelo de negocio del Club. Este puesto lo deben desempeñar ingenieros, matemáticos y estadísticos, que son personas con unas grandes aptitudes técnicas para desarrollar los algoritmos en base a criterios matemáticos y estadísticos, además de conocimientos de programación.
- **Analista de datos (Data Analyst):** Miembro encargado de extraer conclusiones a partir de los resultados obtenidos del procesamiento de los datos en los algoritmos elaborados por el científico de datos y de realizar informes y dashboards para facilitar la comprensión de las partes interesadas (principalmente, el director deportivo y los scouts). Este rol, entre los tres que se han mencionado, es el de mayor relevancia en una Dirección Deportiva Data Driven, ya que, por una parte, debe tener conocimientos técnicos suficientes para poder extraer conclusiones de los resultados obtenidos (para la posterior elaboración de informes) y, por otra parte, debe tener conocimientos futbolísticos para poder contextualizar dicho análisis con aspectos futbolísticos. Por lo tanto, en este rol podrían encajar personas del mundo del fútbol (entrenadores, preparadores físicos e incluso scouts) con nociones de programación o personas del ámbito técnico (ingenieros, estadísticos, matemáticos, entre otros) que tengan nociones de fútbol. De esta búsqueda de perfiles que combinen ambos aspectos y debido a la importancia de tener personal de este perfil, nace el previamente mencionado ‘Máster en Big Data aplicado al Scouting en fútbol’ por la UCAM en colaboración con el Sevilla F.C.

3.1.1.1 Determinación de número de profesionales necesarios por cada rol

Una vez conocidos los roles necesarios para que la Dirección Deportiva Data Driven a implantar funcione correctamente, se deberá determinar el número de trabajadores de cada uno de los roles necesarios para cubrir el mercado de manera efectiva.

Para ello, se deberán definir, de forma aproximada, las ligas que se pretenden analizar junto a la importancia de cada una de ellas, haciendo un reparto de estas para cada uno de los trabajadores del club.

- **Competiciones VIP:** Grupo que contiene las cinco grandes ligas del fútbol europeo, de mayor competitividad y de las cuales procederán un gran porcentaje de los jugadores objeto de adquisición por parte del club. Estas ligas son: La Liga española, la Premier League inglesa, la Bundesliga alemana, la Ligue 1 francesa y la Serie A italiana.
- **Competiciones tipo A:** En este grupo se consideran las competiciones europeas de clubes, como son

la Champions League, la UEFA Europa League y la Conference League. Además, se añaden también los campeonatos de naciones de primer nivel, en caso de que los hubiere, como pueden ser el Mundial de fútbol, la Eurocopa y la Copa América. Por lo tanto, harían un total de 6 competiciones, siendo 3 de ellas competiciones veraniegas que no precisarán de asignación durante la temporada.

- **Competiciones tipo B:** Ligas de segundo nivel internacional como podrían ser la Liga NOS portuguesa, la Jupiler Pro League belga, la Eredivisie holandesa, la Bundesliga austríaca, la Süper Lig turca, el Brasileirao brasileño, la Primera División argentina, la Ekstraklasa polaca, la Liga HNL croata, la Superliga suiza, la 1. Liga checa. Por consiguiente, sería un total de 11 competiciones.
- **Competiciones tipo C:** Ligas de nivel más bajo, entre las que se encuentran la Eliteserien noruega, la Allsvenskan sueca y la Superliga danesa, la Liga MX mexicana y la MLS estadounidense. Adicionalmente, se añaden los campeonatos entre naciones como la Copa de África, la Copa Asia, los campeonatos clasificatorios de Europa y América además de los amistosos entre naciones y los torneos internacionales de categorías inferiores. Por lo tanto, sería un total de 11 competiciones, de las cuales serían únicamente 5 las que se analizarían durante la temporada, siendo las demás competiciones que no discurren durante la temporada.

Por lo tanto, teniendo en cuenta la división realizada y las competiciones que sí habría que analizar durante la temporada, se tendrían un total de 24 competiciones.

Realizada la organización de las ligas, para determinar el número de trabajadores, se necesita realizar la asignación de ligas a cada uno de los trabajadores. Cabe puntualizar que las competiciones europeas e internacionales no se disputan de forma semanal durante toda la temporada, por lo que puede darse el caso de que algunos trabajadores tengan asignado un gran número de ligas sin incurrir en una sobrecarga de trabajo. Las reglas definidas para cada rol son las siguientes:

- **Director deportivo y Chief Data Officer:** Como son roles de gran responsabilidad, solo se precisaría uno por puesto.
- **Scout:** Para determinar el número de scouts suficiente para poder analizar de forma correcta las ligas definidas previamente, se realizará el siguiente reparto:
 - Los scouts sólo analizarán competiciones tipo VIP, A y B, dejando el análisis de las competiciones de tipo C a los expertos en datos, con el fin de liberar a los scouts de carga de trabajo adicional innecesaria. Por lo tanto, los scouts analizarán un total de 19 competiciones.
 - A cada scout se le asignarán 3 ligas, asegurando siempre que se asigne a cada scout una de las ligas tipo VIP y que, de las ligas de demás tipos, no se asignen más de 2 de cada tipo. Además, quedarían por asignar 4 competiciones, que se asignarán de forma aleatoria a cada scout.
 - Teniendo en cuenta este reparto, se precisa de 5 scouts para cubrir todas las competiciones de los tipos definidos previamente.

- **Analista de datos:** La carga de trabajo del analista de datos es menor que la del scout, por lo que, a la hora de definir el número de analistas de datos necesarios para poder cubrir las necesidades del club, el reparto se hará de forma análoga al de los scouts, pero asignando un mayor número de ligas a cada analista, por lo que quedaría de la siguiente forma:
 - Los analistas de datos, como se ha comentado previamente, estarán encargados de supervisar ligas de todos los tipos, siendo responsables en exclusiva, además, de hacer un seguimiento de las ligas tipo C.
 - Cada analista se ocupará, como mucho, de 2 ligas tipo VIP, 1 liga tipo A, 4 ligas tipo B y 2 ligas tipo C; asegurando siempre que el número de ligas a analizar por parte de cada uno de los analistas sea 8.
 - Teniendo en cuenta este reparto, se precisaría de 3 analistas de datos.
- **Científicos e ingenieros de datos:** Para estos dos casos, al trabajar de forma genérica para todas las ligas, no se necesitará determinar el número de profesionales en este ámbito de forma dependiente al número de ligas a analizar. Por tanto, considerando la carga de trabajo general, se precisará 1 profesional en cada uno de los puestos.

3.1.1.2 Inversión salarial

A continuación, en la Tabla 1 se especifican los salarios aproximados a percibir por parte de los trabajadores de la Dirección Deportiva Data Driven, para poder realizar una estimación del coste incurrido por la contratación de personal especializado en esta área.

Por un lado, los salarios correspondientes al grupo de profesionales expertos en datos se han obtenido a partir de las tablas salariales del Convenio Colectivo Nacional de empresas de ingeniería, cuyas disposiciones se encuentran definidas en (Boletín Oficial del Estado, 2023). Los sueldos de estos profesionales, al no estar íntimamente relacionados con el fútbol, se ajustan en mayor medida a la realidad de lo que se podría encontrar en cualquier empresa de otro sector, por lo que consultar las tablas salariales del Convenio Colectivo se ajusta a los salarios que podrían percibir este grupo de trabajadores.

Para este caso, al haber consultado el Boletín Oficial del Estado, se deben particularizar ciertos aspectos que dependerán de la titulación académica y las funciones desempeñadas en el puesto de trabajo:

- **Grupo profesional:** Grupo profesional I. En este grupo profesional, el criterio mínimo de acceso es contar con un título universitario (grado, máster o doctorado) en ingeniería (entre otras titulaciones que no son de interés para este caso particular), siendo esta cuestión de vital importancia para los roles definidos para los expertos en datos.
- **Nivel salarial:** En función del rol a desempeñar, la responsabilidad de cada técnico y la titulación necesaria para desarrollar las labores correspondientes a cada puesto, se determinan los siguientes niveles salariales:

- **Chief Data Officer:** Nivel salarial 1. Este perfil se encuadra en este nivel salarial debido a que, entre sus funciones, se encuentra la “cadena de operaciones a seguir”, según lo dispuesto en las funciones de dicho nivel salarial, al tratarse de un puesto en el que se debe coordinar toda la labor de los expertos de datos con el resto de la dirección deportiva.

Por las obligaciones y responsabilidades que conlleva el desempeño de este puesto tales como la iniciativa y autonomía para ser resolutivo en la identificación y resolución de problemas y la complejidad del puesto en sí a la hora de la toma de decisiones, el trabajador percibirá un 20% adicional al salario bruto percibido por los miembros del mismo nivel salarial respecto a lo dispuesto en el Convenio colectivo.

- **Ingenieros de datos:** Nivel salarial 1. Este perfil se encuadra en este nivel salarial debido a que, entre sus funciones, se encuentra la “puesta a punto de las aplicaciones”, según lo dispuesto en las funciones de dicho nivel salarial, al tratarse de un puesto en el que se deben programar la extracción, recogida y almacenamiento de datos.
- **Científicos de datos:** Nivel salarial 1. Este perfil se encuadra en este nivel salarial debido a que, entre sus funciones, se encuentra la de “puesta a punto de las aplicaciones”, según lo dispuesto en las funciones de dicho nivel salarial, al tratarse de un puesto en el que se deben programar los algoritmos de aprendizaje automático para que la aplicación, a partir de los datos, otorgue resultados lógicos y fidedignos.
- **Analista de datos:** Nivel salarial 2. Este perfil se encuadra en este nivel salarial debido a que, entre sus funciones, no se encuentra ninguna de las expuestas en las disposiciones del nivel salarial superior, pero, se precisa de titulación universitaria especializada en este ámbito.

Por otro lado, para el puesto de director deportivo, ayudante y los scouts, se ha partido de la base de que la Dirección Deportiva Data Driven se va a confeccionar para un club de Primera División española. Al ser un puesto relacionado íntimamente con el mundo del fútbol, el salario a percibir por este tipo de trabajadores no suele ir en consonancia con lo dispuesto en las tablas salariales consultadas previamente, por la simple razón de que, por lo general, perciben mayor salario comparativamente con el grupo profesional al que pertenecerían por su titulación académica. Es por ello por lo que se han consultado fuentes más especializadas en el ámbito futbolístico que tratan los salarios aproximados de estas profesionales en (Canle, s.f.) y (Unisport Management School, 2022), se han obtenido los sueldos de estos profesionales, definidos en la Tabla 1.

PRESUPUESTO I – PERSONAL ESPECIALIZADO		
Cargo del trabajador	Cantidad	Salario anual base en bruto (€)
Director deportivo	1	70.000,00
Ayudante del director deportivo	1	35.000,00
Scout	5	25.000,00
Chief Data Officer	1	32.536,58
Ingeniero de datos	1	27.113,82
Científico de datos	1	27.113,82
Analista de datos	3	21.031,52
TOTAL INVERTIDO		379.858,72

Tabla 1: Presupuesto I – Estimación de inversión anual en personal especializado

3.1.2 Recursos necesarios

Además del personal específico y particular de una ‘Dirección Deportiva Data Driven’, se necesitan una serie de recursos tecnológicos para poder combinar el análisis de los scouts en el visionado de partidos con el análisis de datos. Para ello, se proponen los siguientes recursos necesarios:

- **Licencia de plataforma de análisis:** Es fundamental que cada scout cuente con una plataforma de análisis futbolístico que permita visualizar y analizar la mayor cantidad de partidos y ligas posible a los scouts para poder abarcar la mayor cantidad de competiciones posible y, así, poder ampliar la base de datos de los ojeadores del club. En este sentido, la mejor (y más utilizada por los equipos) aplicación en el ámbito futbolístico es Wyscout. Esta aplicación, según la información de su sitio web (Wyscout, s.f.), es una plataforma que aporta los siguientes servicios a los clubes:
 - Realizar análisis de partidos propios y de rivales, que es un componente importante tanto para los scouts, el director deportivo y el cuerpo técnico.
 - Poner en contacto a agentes de jugadores con los clubes compradores o vendedores para facilitar las negociaciones.

En la Figura 5 se pueden observar distintos gráficos y representaciones que aparecen en la aplicación sobre el jugador del Atlético de Madrid, Koke Resurrección.

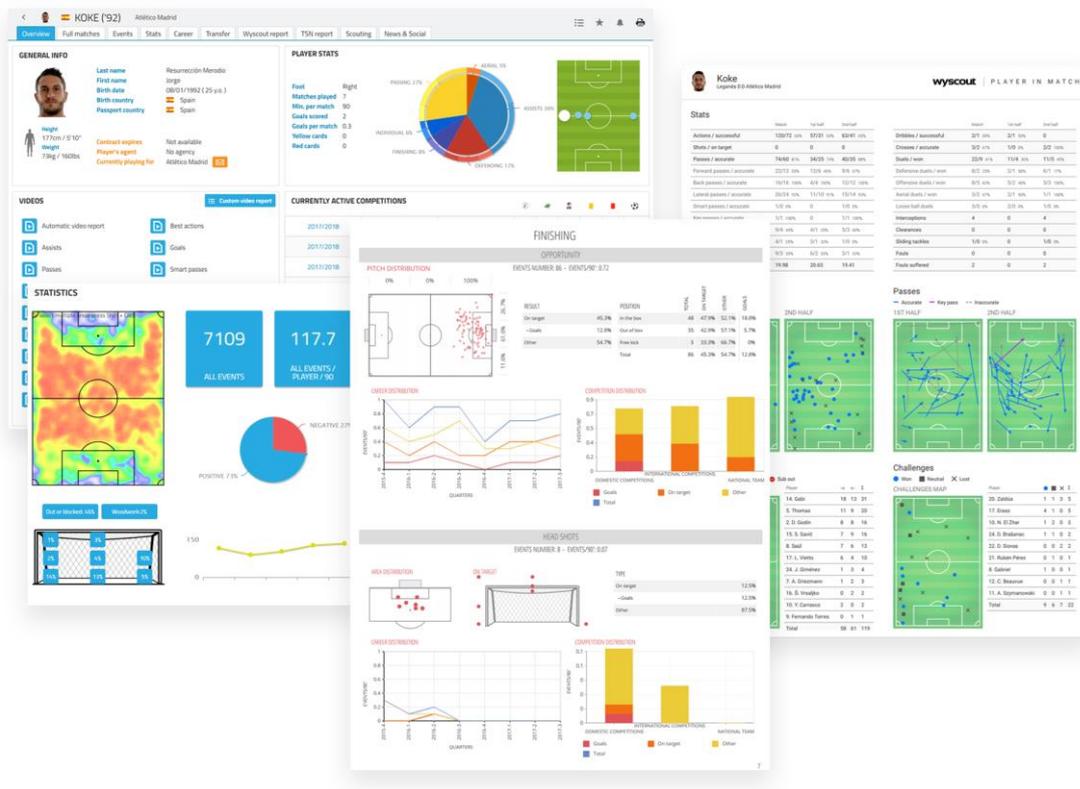


Figura 5: Visualizaciones de la aplicación Wyscout (Wyscout, s.f.)

- **Licencia de servicio de datos:** Es importante que cada trabajador especializado en datos cuente con una base de datos actualizada que incluya información detallada sobre jugadores de todo el mundo. En este sentido, hay varias alternativas en el mercado como StatsBomb, Instat, entre otros, pero el servicio más utilizado por los clubes para disponer de datos es Stats Perform (o, con su denominación anterior y más conocida, Opta).

Este software, según su página web (Stats Perform, s.f.), cuenta con una base de datos de 500.000 partidos, más de 20 acuerdos con distintas entidades, casi 4.000 competiciones cubiertas y, lo más importante, más de mil millones de datos de tracking de jugadores. Además, cuenta con servicios y aplicaciones programadas por ellos basadas en inteligencia artificial, en las cuales ya se ofrecen resultados y análisis pormenorizados a partir de los datos. A continuación, en la Figura 6, se puede observar un ejemplo de análisis de rendimiento realizado por la división dedicada a analítica basada en inteligencia artificial de Opta, en el cual se realiza una comparativa de rendimiento entre el reciente exdelantero del Real Madrid C.F. (Karim Benzema), con el que, a fecha de hoy y de acuerdo con los rumores, es el principal objetivo del club (Harry Kane). Posteriormente, Harry Kane ha acabado fichando por el Bayern de Múnich de la liga alemana (Bundesliga).

En la Figura 6 se pueden observar las diferencias de rendimiento de los dos jugadores, gracias a la implementación de dos gráficos de radar. Opta Analyst, en este caso, ha optado por presentar las

métricas de los jugadores valorándolas con percentiles de rendimiento (que posteriormente se explicarán en qué consisten desde un punto de vista teórico). Estos percentiles, desde el punto de vista futbolístico, miden el rendimiento de un jugador contextualizándolo con los demás jugadores que forman parte de la muestra de datos, de tal forma que, cuanto mayor es el valor del percentil, mejor es el rendimiento del jugador en una determinada métrica. Estas métricas de rendimiento inciden sobre 3 características o atributos de juego: ataque, posesión y físico.

Por una parte, en las métricas de ataque (en rojo), el rendimiento de ambos delanteros es muy bueno, aunque destaca en mayor medida Karim Benzema, estando en el percentil 97 para la métrica ‘goles por 90 minutos’, en el percentil 98 para ‘tiros por 90 minutos’ y en el percentil 98 para ‘toques en área rival por 90 minutos’. Esto significa que el 97/98% (según el caso) de jugadores pertenecientes a la muestra están por debajo en rendimiento en dichas métricas.

Por otro lado, en las métricas físicas, sin destacar particularmente ninguno de los dos jugadores, destaca más Harry Kane. Al contrario ocurre con las métricas de posesión, en las cuales destaca más Benzema.

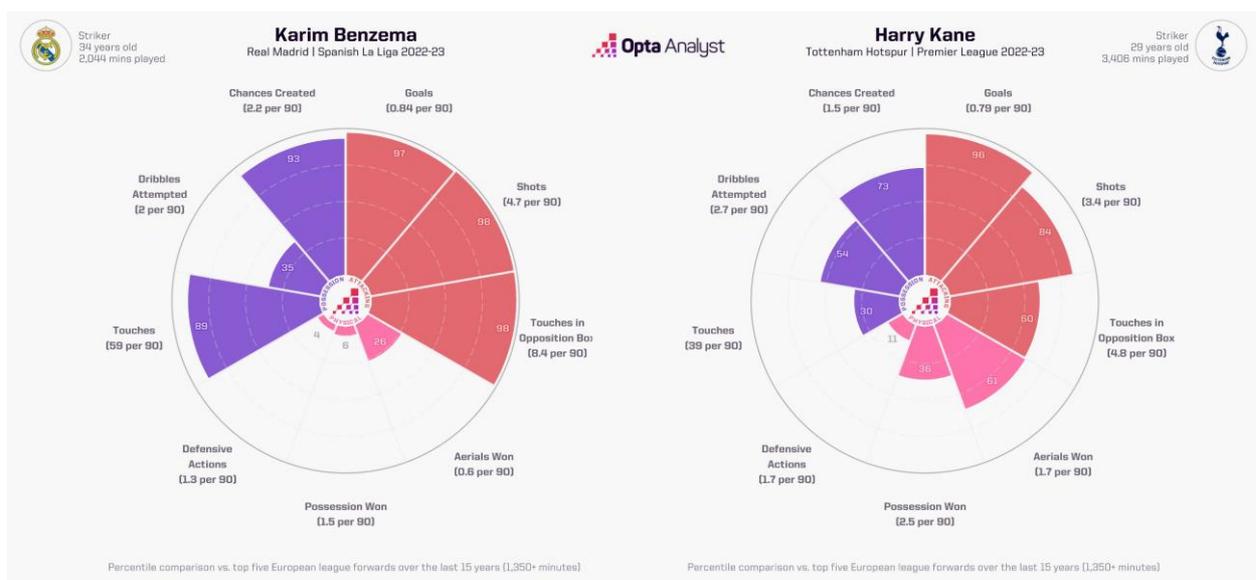


Figura 6: Comparativa de rendimiento entre Karim Benzema y Harry Kane (OptaAnalyst, 2023)

Tanto para la licencia de plataforma de análisis (en este caso, se elegiría Wyscout), como para la licencia del proveedor de datos (Opta), se tiene un coste parecido. Esto es debido a que, en ambos casos, los precios están sujetos a los servicios a contratar y las ofertas que tienen los propios clubes, son aproximaciones obtenidas a partir de fuentes de personas que trabajan con ambas entidades, pero en ningún caso deben tomarse como referencia precisa. Para obtener una mejor estimación del precio, se deberá consultar, en nombre del club en cuestión, a cada una de las entidades, para que otorguen un precio adecuado en cada caso.

- **Herramientas de visualización de datos:** Estas herramientas permiten presentar los resultados del análisis de manera clara y comprensible, para que los miembros de la dirección deportiva o del cuerpo

técnico que no tengan nociones estadísticas y matemáticas tengan una comprensión clara del análisis llevado a cabo por el analista correspondiente. Estas aplicaciones cuentan con gráficos, tablas y paneles interactivos ayudarán a transmitir la información relevante de manera efectiva. En este sentido, las dos aplicaciones más utilizadas (que, además, son gratuitas) son PowerBI y Tableau.

- **Equipos tecnológicos:** Para implementar una Dirección Deportiva Data Driven, es fundamental contar con un sistema robusto de adquisición y almacenamiento de datos. Esto implica, entre otras cuestiones, disponer de tecnologías que permitan recopilar, procesar e interpretar métricas de rendimiento de los jugadores. En este sentido, para desarrollar los trabajos programados en este tipo de dirección deportiva, se necesitarán:

- o **Equipos informáticos:** Los miembros de la dirección deportiva deben contar con ordenadores de alta capacidad de procesamiento (CPU y GPU de alta capacidad) y potencia de cálculo para el procesamiento de datos y la posterior ejecución de los algoritmos de Machine Learning (enfocados a los expertos en datos) y también con gran potencia para el procesamiento de vídeos (enfocado a los scouts). Como estos equipos son necesarios para el trabajo diario de cada miembro de la dirección deportiva, debe haber un ordenador por cada trabajador.

Para asegurar el buen rendimiento de los equipos en cada uno de los roles necesarios dentro de la dirección deportiva, se opta por adquirir equipos con un hardware potente, para poder realizar con solvencia las labores de analítica y procesamiento de datos, así como su gestión en la nube, visualización de partidos en las plataformas de análisis, elaboración de informes en las plataformas de visualización de datos, entre otras funciones.

Se han escogido tres opciones distintas en precio y en especificaciones (consultadas el 3 de mayo de 2023), teniendo en cuenta que todas tengan, como mínimo, el procesador más potente del mercado, así como una memoria RAM lo suficientemente potente como para poder trabajar en varias tareas en simultáneo (como mínimo, 16GB):

- Opción asequible: Lenovo IdeaPad 3 15ITL6 (699€). Para más información: [Lenovo IdeaPad 3 15ITL6 Intel Core i7-1165G7/16GB/512 GB SSD/15.6" | PcComponentes.com](https://www.pccomponentes.com/lenovo-idea-pad-3-15itl6-intel-core-i7-1165g7-16gb-512gb-ssd-15-6/)
- Opción intermedia: Asus Zenbook 13 UX325EA-KG448W (949€). Para más información: [ASUS ZenBook 13 OLED UX325EA-KG448W Intel Core i7-1165G7/16GB/512GB SSD/13.3" | PcComponentes.com](https://www.pccomponentes.com/asus-zenbook-13-oled-ux325ea-kg448w-intel-core-i7-1165g7-16gb-512gb-ssd-13-3/)
- Opción cara: MSI Katana 15 B12VGK-094XES (1349€). Para más información: [MSI Katana 15 B12VGK-094XES Intel Core i7-12650H/16GB/1TB SSD/RTX 4070/15.6" | PcComponentes.com](https://www.pccomponentes.com/msi-katana-15-b12vgk-094xes-intel-core-i7-12650h-16gb-1tb-ssd-rtx-4070-15-6/)

Se ha optado por escoger una opción con tarjeta gráfica no dedicada (opción cara), una opción

de ordenador más portable y de menor peso (opción intermedia) y una opción barata en caso de que el coste sea la restricción más importante. De cara al presupuesto, se ha determinado que se va a adquirir la opción intermedia.

Para los equipos informáticos, al ser bienes de utilización para más de un año, se debe proponer una amortización, en este caso lineal, por los años de uso de dichos equipos. Se estima una vida útil de dichos equipos de 8 años, con un valor residual nulo. Por lo tanto, el gasto anual incurrido sería el coste de cada equipo entre el número de años de vida útil estimada.

Una vez definidos todos los recursos necesarios para poder articular la confección de la Dirección Deportiva Data Driven, es necesario definir la inversión requerida para cada uno de los recursos definidos como indispensables para desarrollar la actividad de este departamento, lo cual se define en la Tabla 2.

PRESUPUESTO II – RECURSOS			
Concepto	Cantidad	Coste unitario (€)	Coste total (€)
Licencia de plataforma de análisis	1	50.000,00	50.000,00
Licencia de servicio de datos	1	50.000,00	50.000,00
Herramienta de visualización de datos	1	Gratuito	
Equipos tecnológicos	17	118,63	2016,71
TOTAL INVERTIDO			102.016,71

Tabla 2: Presupuesto II - Estimación de inversión anual en recursos

3.2. Cronograma de las principales fases para la planificación de una plantilla

En el mundo del fútbol, para realizar cualquier cuestión que esté íntimamente relacionada con la organización de un departamento, hay que tener en cuenta los detalles acerca de cuándo se desarrolla la temporada. Particularmente, en el caso de las direcciones deportivas, no solo hay que tener en cuenta cuándo se juega la temporada, sino que además hay que tener en cuenta las ventanas de fichajes disponibles para hacer movimientos en el mercado. Por lo tanto, lo primero a definir para poder organizar un cronograma de planificación anual es cómo están estructurados la temporada y los mercados de fichajes.

Como referencia para desgranar los diferentes estadios de la temporada, se escogerá el calendario de la próxima temporada de La Liga española (2023/2024), la cual no se ve afectada en su comienzo ni final por una competición internacional de naciones, como pueda ser un Mundial de fútbol, una Eurocopa, entre otros. Para la extracción de fechas, se ha utilizado la página web de fútbol Transfermarkt (Transfermarkt, s.f.).

- 1) **Finalización de la temporada 2022/2023:** 4 de junio de 2023.
- 2) **Apertura del mercado de fichajes de verano:** 1 de julio de 2023.
- 3) **Comienzo de la temporada 2023/2024:** 12 de agosto de 2023.
- 4) **Cierre del mercado de fichajes de verano:** 1 de septiembre de 2023.
- 5) **Apertura del mercado de fichajes de invierno:** 2 de enero de 2024.
- 6) **Cierre del mercado de fichajes de invierno:** 1 de febrero de 2024.
- 7) **Finalización de la temporada 2023/2024:** 26 de mayo de 2024.

Cabe puntualizar que los mercados de fichajes (tanto el estival como el invernal) sirven para poder inscribir a los jugadores adquiridos, pero realmente, se puede negociar con clubes las compras y ventas de jugadores fuera de dicho período. Igualmente, para la realización del cronograma, se supone que sólo se puede negociar y adquirir o vender jugadores en el período de mercado.

Otra puntualización u observación que se puede realizar acerca de este calendario es que, en el caso del mercado de verano, hay un período en el cual la temporada ya ha comenzado (entre el 12 de agosto y 1 de septiembre de 2023), por lo que la temporada comienza con el mercado aún abierto. En el caso del mercado de invierno, se desarrolla íntegramente durante la temporada, aunque también, en gran parte de este, se encuentran las vacaciones de invierno, por lo que prácticamente no se juega durante el mercado.

El cronograma o esquema a seguir durante sendas planificaciones (verano e invierno) es equivalente, ya que para acometer cada una de las planificaciones, los procesos necesarios para llevarlas a cabo son los mismos.

Una vez definidas las fechas más relevantes del calendario de la temporada, se definen los períodos en los que se divide el calendario de trabajo anual de la Dirección Deportiva Data Driven, divididos en fases. Para clarificar el flujo de información necesario de cara a cada una de las fases, se van a definir las entradas y salidas de información necesarias.

3.2.1. Fase I – Análisis de jugadores

Esta fase trata, como su propio nombre indica, de recopilar jugadores a seguir durante la temporada y a engrosar la base de datos de jugadores del club. Es la base de la planificación de la temporada, ya que es el proceso en el cual se analiza el mercado y se determinan qué jugadores pueden ser objeto de adquisición por parte del club, ajustado al presupuesto destinado para la planificación. En la Tabla 3, se pueden observar las entradas y salidas de flujo de información de esta fase, así como las fechas en las que discurre.

Fechas	Entradas	Salidas
<ul style="list-style-type: none"> - Mercado de verano: enero 2023 – abril 2023. - Mercado de invierno: agosto – octubre 2023. 	<ul style="list-style-type: none"> - Base de datos de jugadores en el mercado - Base de datos de jugadores propios - Base de datos de equipos - Partidos de la temporada - Asignación de ligas 	<ul style="list-style-type: none"> - Evaluación de jugadores por parte de los scouts

Tabla 3: Relación entrada-salida de información y fechas de la Fase I

3.2.1.1. Primer paso: Mapeo

Este primer paso es el análisis preliminar de las competiciones que sirve para recabar información de los jugadores de las distintas ligas asignadas a cada uno de los scouts y profesionales del ámbito del dato. A continuación, se describen las funciones generales que realiza cada uno de los miembros de la dirección deportiva:

- **Scouts:** Los ojeadores tienen como objetivo principal el visionado de partidos sin un objetivo claro. Este visionado se producirá, en primera instancia, en plataformas digitales o en directo por televisión. Cada scout, en las ligas que tiene asignada, es responsable de recabar información y analizar, de forma preliminar, a aquellos jugadores que, en un visionado de 3-4 partidos (como mínimo) le haya resultado de interés. Dentro de los jugadores elegidos por parte del scout, solo pueden estar aquellos que, por capacidad económica, puedan entrar dentro de los parámetros económicos del club, para evitar realizar análisis de aquellos jugadores que son inalcanzables para el club en primera instancia.

Todos los jugadores que hayan sido seleccionados en este primer paso por alguno de los scouts se introducirán dentro de la base de datos propia del club. Para añadir esta información, el scout realizará una pequeña encuesta de valoración (puntuación de 1-5) acerca de las capacidades técnico-tácticas del jugador que se hayan podido observar en las visualizaciones de partidos, entre las cuales se encuentran:

- 1) Valoración del general del futbolista.
 - 2) Valoración del jugador en aspectos ofensivos (capacidad goleadora, capacidad asistencial, técnica de regate, capacidad asociativa, entre otros) y en aspectos defensivos (juego aéreo, posicionamiento en el campo, capacidad de robo de balón, entre otros).
 - 3) Valoración del jugador en relación con la capacidad de adaptabilidad al modelo de juego del equipo.
- **Personal experto en datos:** En este paso, tendrán dos principales cometidos:
 - 1) Programación de los algoritmos necesarios para analizar el rendimiento de los jugadores objeto

de adquisición por parte del club para asegurar que el análisis sea lo más fidedigno, actualizado y ajustado a la realidad posible.

- 2) Obtención, tratamiento y actualización constante de los datos de los proveedores y posterior análisis para, al igual que los scouts, insertar a los jugadores más interesantes a ojos de los datos dentro de la base de datos del club.
- 3) Análisis del rendimiento de los jugadores de la plantilla para complementar la información del cuerpo técnico en lo relativo a gestión de esfuerzos, prevención de lesiones, estudio del rendimiento general de los jugadores, entre otras cuestiones.

3.2.1.2. Segundo paso: Cribado

En este paso, el objetivo principal es profundizar en el análisis de aquellos jugadores que, según lo observado y analizado por los scouts y los profesionales en datos, son más interesantes para el club, de cara a la generación de informes que se realizará en etapas posteriores, una vez definidas las necesidades concretas del club. En este sentido, se concretarán las observaciones realizadas en el mapeo, con el fin de ofrecer al director deportivo la mayor cantidad de información posible.

- **Scouts:** En esta fase los scouts realizarán un análisis más profundo de aquellos jugadores que figuren en la base de datos recogida en el mapeo, ahondando en las siguientes cuestiones relativas al rendimiento del futbolista tanto dentro como fuera del terreno de juego, extraídas de (Soria Polo, 2021):
 - o **Perfil físico:** En función del requerimiento concreto en dicha posición, analizar las capacidades físicas del jugador a la hora de repetir esfuerzos, capacidad de cambio de ritmo, rapidez de movimientos, capacidad de esprintar, entre otros. Estos aspectos tienen más importancia en las posiciones con una mayor demanda física, como pueden ser los laterales, los centrocampistas defensivos o futbolistas de banda.
 - o **Perfil técnico:** En función de la posición a analizar, las características técnicas variarán, pero, por lo general, en este aspecto se evalúan aspectos como la visión de juego y la capacidad de filtrar pases entre líneas, el control y la habilidad del jugador con balón, precisión de pase y tiro, rapidez en la toma de decisiones, habilidad para perfilarse correctamente, etcétera. Estas cuestiones son más predominantes y de mayor importancia en los centrocampistas de corte más creativo y en los jugadores atacantes en general.
 - o **Perfil táctico:** En esta parte se analizará, sobre todo, la capacidad del futbolista de situarse correctamente en el campo en función de lo que demande la jugada, evaluando el nivel y la capacidad de anticipación de las situaciones del juego. Esta cuestión es vital para los jugadores de corte más defensivo, como pueden ser los centrales y los pivotes (como se conoce a los centrocampistas más defensivos), que necesitan tener buenas capacidades tácticas para poder

defender de manera más férrea el área propia.

- **Perfil psicológico:** Según lo que comenta (Pamos, s.f.), en un artículo sobre la relación psicología-scouting, los parámetros psicológicos relevantes a la hora de analizar a un futbolista en este aspecto son los siguientes:
 - **Características del jugador:** Por una parte, se debe analizar el comportamiento general del jugador en el terreno de juego durante el transcurso del partido. Es importante analizar estas características en diferentes partidos, que sean de mayor y/o menor importancia, para tener un análisis más exhaustivo en el que se tengan en cuenta circunstancias de juego distintas. Se debe incidir sobre las siguientes cuestiones:
 - **Personalidad:** Afectación del jugador teniendo en cuenta lo que pasa durante el partido y cómo le afecta a su juego.
 - **Actitud:** Capacidad de gestión de momentos de tensión o estrés durante el partido y cómo responde ante las mismas, si con personalidad (participando más) o sin ella (participando menos e intentando retirarse de las acciones de peligro).
 - **Estados de ánimo:** Gestión de emociones del jugador en momentos en los que el equipo está jugando bien o mal, analizando su rendimiento en cada caso.
 - **Sentido de responsabilidad:** Capacidad de “echarse el equipo a la espalda” (tomar responsabilidad) del futbolista en un momento en el que el equipo no está bien anímica o futbolísticamente.
 - **Entorno del jugador:** Por otra parte, es primordial analizar el entorno más cercano del futbolista, para saber si este es sano y que aporta al jugador o, por el contrario, es contraproducente para el rendimiento del futbolista. Para ello, se debe contactar con personas del club vendedor, agentes o incluso familiares, que sean de confianza y que puedan hablar de cómo es la vida del futbolista fuera de los terrenos de juego.
- **Condiciones económicas:** Esta información, con la ayuda del director deportivo y del personal de datos de la dirección deportiva, se debe tener en cuenta a la hora de analizar un futbolista, para ver si está dentro de las posibilidades económicas del club. Para ello, se deben considerar los siguientes factores:
 - **Valor de mercado aproximado:** Montante de dinero referencia (no es el precio final del futbolista) que sirve de punto de partida para valorar al jugador. Para estimar su valor de mercado, se pueden considerar las siguientes fuentes de información:
 - Webs especializadas como Transfermarkt.

- Aplicaciones y/o servicios informáticos que apliquen algoritmos predictivos para valorar al futbolista en base a criterios y datos objetivos, como hacen algunos clubs a nivel interno.
- **Valor de la cláusula de rescisión:** En la mayoría de los contratos de futbolistas, es la cantidad que libera al jugador del contrato que tiene con el club en el que milita. Cobra mayor importancia en el caso de que el club vendedor no quiera negociar y el jugador tenga un gran interés en fichar por el club comprador.
- **Precio demandado por el club vendedor:** En caso de que se tenga esta información (a veces, de difícil acceso), es la cifra más importante a tener en cuenta, ya que es el valor por el cual el club vendedor está dispuesto a desprenderse de los servicios del jugador en cuestión. Para acceder a esta información, se puede consultar al propio club o al agente del jugador.
- **Salario:** Además del montante a pagar al club vendedor, es importante conocer cuáles son las pretensiones económicas de retribución del propio futbolista ya que, en algunos casos, aunque la cantidad del traspaso solicitada por el club vendedor sea baja, puede darse el caso de que el salario del futbolista sea prohibitivo para la masa salarial del club. La masa salarial del club debe estar ajustada a los criterios establecidos por la liga en la que se compite, ya que un incumplimiento en la misma puede incurrir en una multa o incluso en una posible descalificación.
- **Personal experto en datos:** Su principal objetivo es la visualización de los datos de los jugadores que pertenezcan a la liga asignada a cada analista de datos para tenerlo en consideración en etapas posteriores de la planificación.

3.2.2. Fase II – Definición de perfiles

Es la fase donde se da comienzo a la planificación de la temporada, en la cual se definen las necesidades concretas del club en base al criterio del entrenador y se elaboran listas de jugadores por posición en base a lo referido en el informe de necesidades. En la Tabla 4, se definen las entradas y salidas de flujo de información de la presente fase, así como las fechas en las que discurre para cada mercado:

Fechas	Entradas	Salidas
- Mercado de verano: abril – junio 2023	- Base de datos de jugadores - Base de datos de clubes	- Informe de necesidades - Listas finales de jugadores
- Mercado de invierno: octubre – noviembre 2023.	- Presupuesto de la planificación	- Informes finales de jugadores

Tabla 4: Relación entrada-salida de información y fechas de la Fase II

3.2.2.1. Primer paso: Reunión con el entrenador

Este primer paso es responsabilidad del director deportivo que, como miembro más relevante de la dirección deportiva, va a ser el máximo responsable de la planificación. Por lo tanto, se le otorga la tarea más importante y de la que depende el trabajo del resto del departamento, la reunión con el entrenador. En esta reunión, se concretarán los siguientes aspectos:

- **Demarcaciones a reforzar:** Definiendo el número de jugadores por posición necesarios para llevar a cabo su idea de juego durante toda la temporada, otorgándole la suficiente profundidad de plantilla para poder contar con efectivos sanos durante todo el transcurso de la temporada. El número de jugadores va a depender, en gran parte, del número de competiciones que dispute el equipo en cuestión, siendo mayor en caso de que se jueguen competiciones europeas.
- **Perfiles del jugador a adquirir:** En este respecto, el entrenador debe determinar con la mayor exactitud posible el perfil de jugador que necesitaría adquirir para poder desarrollar su idea de juego. Esto ocurre porque, en función de la filosofía de juego del entrenador, se pueden requerir características distintas en cada una de las posiciones. Por ejemplificar, se añade un caso concreto en el que, para la misma demarcación (en este caso, la de central), hay varias opciones según el perfil del futbolista:
 - **Central con buen juego aéreo:** En este caso, si el entrenador solicitara un central con estas características, es muy probable que se trate de un entrenador que busque defender muy cerca de su portería. Esto propicia que el rival, en la mayoría de las ocasiones, busque llegar a portería a través de centros al área, lo que hace que sea indispensable que el central a adquirir tenga buen juego aéreo, priorizando en este caso el rendimiento defensivo del jugador.
 - **Central con buena salida de balón:** Este caso suele darse en equipos que realizan juego de posición. El juego de posición, según (Idoate, s.f.), es un estilo de juego que busca crear superioridades desde la posición, sacando la pelota desde área propia, avanzando conjuntamente hasta el área rival. Para esto, se necesita un central que tenga un trato de balón lo suficientemente bueno como para poder superar la primera línea de presión rival, dándole mayor peso del que se presupone para un central a las métricas relativas a la posesión de balón.
 - **Central rápido:** En este caso, quizá, el entrenador puede buscar, en su disposición en el campo, que el equipo sea lo más corto posible, entendiendo por corta la distancia entre la línea defensiva y la línea atacante, para poder obstaculizar en mayor nivel la salida y progresión del balón del rival. La contraprestación que tiene este estilo de juego es que la línea defensiva está muy adelantada respecto al área propia, lo que propicia que el entrenador busque que sus centrales sean lo más rápido posible, para poder recuperar su posición cerca del área en caso de que el rival consiga superar las líneas de presión propias, en cuyo caso se priorizarán las métricas relativas al físico del jugador, así como las métricas defensivas.

- **Jugadores a vender:** En esta cuestión, el entrenador debe comentar al director deportivo los jugadores con los que no cuenta o que no casan con su filosofía de juego. El director deportivo intentará atender todas las peticiones del entrenador, siempre y cuando haya compradores ofreciendo una cantidad justa y suficiente por dichos jugadores. Puede darse el caso, también, de que haya jugadores de los que el entrenador no quiera prescindir, pero, a raíz de una oferta fuera de mercado, el director deportivo plantee su venta. En este caso, el director deportivo deberá suplir esta baja con otro futbolista atendiendo, de nuevo, a las peticiones específicas del entrenador.

El director deportivo realizará un informe de necesidades en el cual figurarán todas las cuestiones comunicadas por el entrenador en este paso, concretando las necesidades específicas en cada una de las posiciones y los jugadores que se pretenden vender a otro club.

3.2.2.2. Segundo paso: Selección

El director deportivo entregará el informe de necesidades al Chief Data Officer (para que lo comunique con el personal experto en datos) y a los scouts, para que conozcan las necesidades de la plantilla comunicadas por el entrenador. En este paso, las funciones de cada uno de los trabajadores de la plantilla serían:

- **Scouts:** Los ojeadores, a partir de los informes generales realizados en la fase anterior y atendiendo a las instrucciones dadas por el director deportivo en el informe de necesidades sobre los perfiles necesarios para fichar, elaborarán una lista de jugadores en la cual, bajo el criterio de los ojeadores, se encuentren aquellos jugadores con un mayor parecido al perfil requerido ajustados al presupuesto del club, dando prioridad a aquellos jugadores con mejor valoración general.
- **Personal experto en datos:** Tienen dos objetivos marcados:
 - o **Caracterización:** El personal específico dedicado al análisis de datos, cuando recibe la información del informe de necesidades, debe comenzar su trabajo de caracterización de los jugadores que se encuentran en la base de datos. La caracterización, entendida en el argot del fútbol y en este proyecto, es la determinación de los jugadores que más se asemejan a los perfiles que ha determinado el entrenador en la reunión con el director deportivo. Para ello, se deben priorizar aquellas métricas más relevantes para que el jugador cumpla las funciones determinadas por el entrenador y, aplicando algoritmos de Machine Learning, se estima quienes son los jugadores más parecidos a lo requerido en cada caso.
 - o **Selección:** A partir de los resultados obtenidos en la caracterización y los informes realizados en la fase anterior, los analistas de datos elaborarán una lista definitiva en la cual se determinarán cuáles son los jugadores que más se ajustan al perfil requerido por el entrenador ajustados al presupuesto del club, de igual forma que realizan los scouts. En esta lista, figurarán jugadores por cada una de las posiciones por reforzar junto a la puntuación que obtiene cada jugador a partir del algoritmo (objetivo).

Como producto de la conclusión de la Fase II, se obtienen dos listas de jugadores, una realizada por los responsables del área de datos y otra por parte de los scouts. A continuación, algún miembro de la dirección deportiva debe ponderar estas listas de manera que se obtenga una nota final ponderando la opinión de las dos vertientes de información por igual (50% de la puntuación a partir del criterio de los scouts y el otro 50% a partir de los resultados obtenidos de los datos). Ponderadas las listas, simplemente quedaría trasladar al director deportivo estos resultados para que evalúe, en la Fase II, la confección de la plantilla.’

A estas listas de jugadores se le añaden también los informes de valoración de los futbolistas (los realizados en la Fase I con actualizaciones en la Fase II), para que el director deportivo y el entrenador los analicen.

3.2.3. Fase III – Confección de la plantilla

De igual forma que para las fases anteriores, en la Tabla 5 se concreta el flujo de información de entrada y de salida que tienen lugar en esta fase y las fechas en las que se desarrolla:

Fechas	Entradas	Salidas
<ul style="list-style-type: none"> - Mercado de verano: junio – agosto 2023. - Mercado de invierno: diciembre 2023 – enero 2024. 	<ul style="list-style-type: none"> - Listas finales de jugadores - Informes de valoración de jugadores actualizados - Presupuesto de la planificación 	<ul style="list-style-type: none"> - Compraventa de jugadores

Tabla 5: Relación entrada-salida de información y fechas de la Fase III

3.2.3.1. Primer paso: Priorización

En este punto, el director deportivo, con la ayuda del entrenador deben ordenar por prioridad cada uno de los jugadores que aparecen en las listas e informes que proceden de la fase anterior. Para ello, deberá tener en cuenta el gasto incurrido en fichar a cada uno de los futbolistas, ya que puede darse el caso, si se fichan a todas las primeras opciones de la dirección deportiva por cada posición, de que el gasto total supere al presupuesto del club en materia de fichajes para ese mercado.

El director deportivo, además, deberá considerar en mayor medida a aquellos jugadores que aparezcan en la lista y el entrenador conozca de primera mano, ya que esto facilitará su adaptación al club y a la manera de trabajar del entrenador. Si ocurriese que el entrenador, de la lista proporcionada por el director deportivo, no ha trabajado con ninguno de los jugadores, el director deportivo bajo su criterio decide qué futbolista se convierte en la primera opción del club.

3.2.3.2. Segundo paso: Negociación y compraventa de jugadores

- **Director deportivo:** Una vez las listas de jugadores estén ordenadas por prioridad, el director deportivo debe comenzar la negociación con los clubes vendedores. En este punto, sobre todo, debe

tener en cuenta que no se esté pagando un sobreprecio respecto al valor de mercado estimado por el club como máximo a pagar por dicho futbolista. En caso de no poder llevar a cabo la contratación del futbolista prioridad (opción A), se llevará a cabo la contratación de la opción B, y así sucesivamente hasta poder adquirir a un futbolista.

- **Scout y Datos – Reserva:** En esta Fase III, toda la dirección deportiva (salvo el director deportivo), está a la expectativa de cualquier novedad que haya en el mercado o en las competiciones veraniegas, en caso de que las hubiera. Su función es vigilar estas competiciones por si hubiera algún futbolista que destaque en su rendimiento y que pudiera ser potencial objeto de adquisición del club y no estuviera contemplado previamente. Además, servirán de soporte al director deportivo en caso de que les haga cualquier petición si fuera necesario.

Una vez definidas todas las fases del cronograma de la planificación, se ha realizado un esquema general de todas las fases en el cual se pueden observar las fechas concretas de cada uno de los pasos dentro de cada fase de forma gráfica, teniendo en cuenta además que muchos pasos dentro de cada fase se solapan, estableciéndose, de esa forma, un flujo de trabajo continuo. Se puede observar este cronograma en la Figura 7.

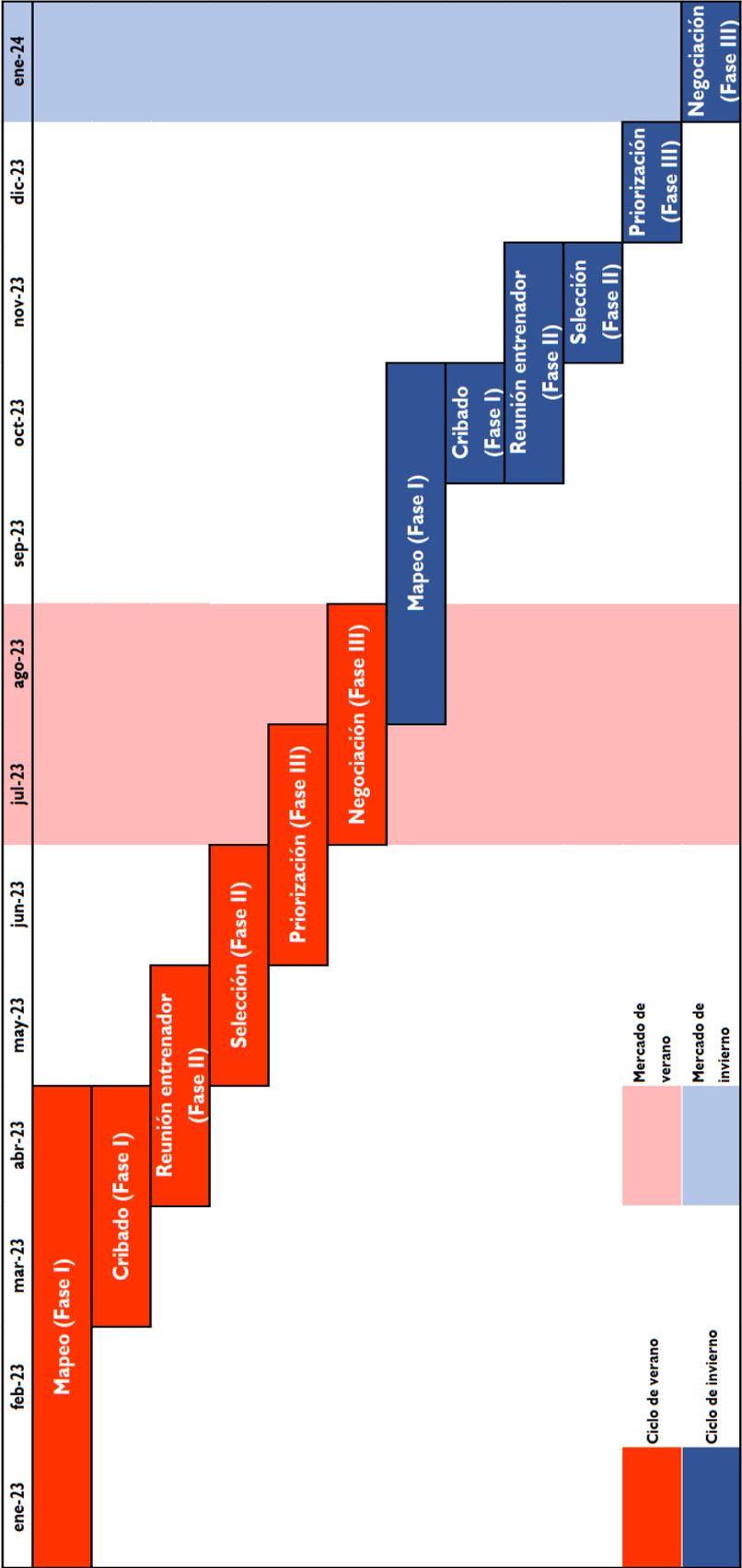


Figura 7: Cronograma de las principales fases para la planificación de una plantilla (Fuente: Elaboración propia)

4 PLANIFICACIÓN DE LA PLANTILLA

Definida la confección de la Dirección Deportiva Data Driven y cómo se organizaría, en tiempo y recursos, el próximo paso es el desarrollo de los procesos necesarios para poder realizar las planificaciones de la plantilla, cuya base se apoye en el uso de los datos que recojan el rendimiento de los futbolistas. La planificación de la plantilla, en este caso enfocada en el primer equipo, es el proceso más importante y diferencial en un club de fútbol, ya que, si el equipo tiene una plantilla completa y ajustada a las necesidades del entrenador, es más probable que el club coseche éxitos deportivos y, por ende, consiga un mejor rendimiento económico.

Este proyecto tiene como objetivo optimizar, automatizar y aportar más información al análisis de los trabajadores de la dirección deportiva. Para ello, es importante tener en cuenta los procesos a seguir, análogos a cualquier proyecto que haga uso de la gestión de datos.

Según (Big Data International Campus, s.f.), existen algunas metodologías para realizar proyectos basados en analítica de datos con el fin de extraer información, analizarla y llegar a conclusiones de cara a la toma de decisiones. Dentro de las metodologías expuestas en la referencia utilizada, se opta por elegir la KDD o Knowledge Discover Database (en español, descubrir conocimiento de una base de datos).

Este proceso se centra en extraer conocimiento de un conjunto de datos a partir de lo que se conoce como minería de datos, que no es más que el proceso de tratamiento del conjunto de datos para poder extraer conclusiones a partir de la información arrojada. Se trata de un ciclo iterativo e interactivo y consta de cinco etapas:

- 1) **Selección:** Es el primer paso del ciclo e implica seleccionar la fuente de datos y, posteriormente, elegir qué métricas o variables son de interés para el análisis que se pretende aplicar. Para un análisis desde un punto de vista futbolístico, es muy importante diferenciar de manera clara la naturaleza de las variables (ofensivas, defensivas, de posesión de balón, entre otras) porque, en función del objetivo concreto que se pretenda, se debe dar mayor o menor peso a unas métricas u otras.
- 2) **Procesamiento:** En esta etapa se procede a limpiar y transformar los datos obtenidos en la fase de selección, con el fin de asegurar que se elimina el sesgo y que se cuente con datos uniformes, que es un factor vital para poder aplicar análisis estadísticos y modelos de Machine Learning.
- 3) **Transformación:** En esta fase se busca transformar la base de datos obtenida a partir de las dos primeras etapas en una base de datos de menor dimensión, pero sin perder información, lo que se conoce como reducción de la dimensionalidad, con el fin de facilitar la aplicación de los modelos de Machine Learning.
- 4) **Minería de datos:** Es el punto en el cual se aplican los algoritmos de aprendizaje automático que, según el conjunto de datos del que se dispone, va a extraer la mayor cantidad de información útil de cara a satisfacer las necesidades concretas del proyecto, en este caso, realizar la planificación de la

plantilla.

- 5) **Interpretación o evaluación:** Es la última etapa del proceso de análisis de datos y su función es extraer conclusiones del estudio y evaluar la adecuación de los modelos implementados en la etapa anterior con el conjunto de datos que se tiene.

En la Figura 8, se resume de forma esquemática los procesos a seguir para realizar una planificación en una Dirección Deportiva Data Driven.

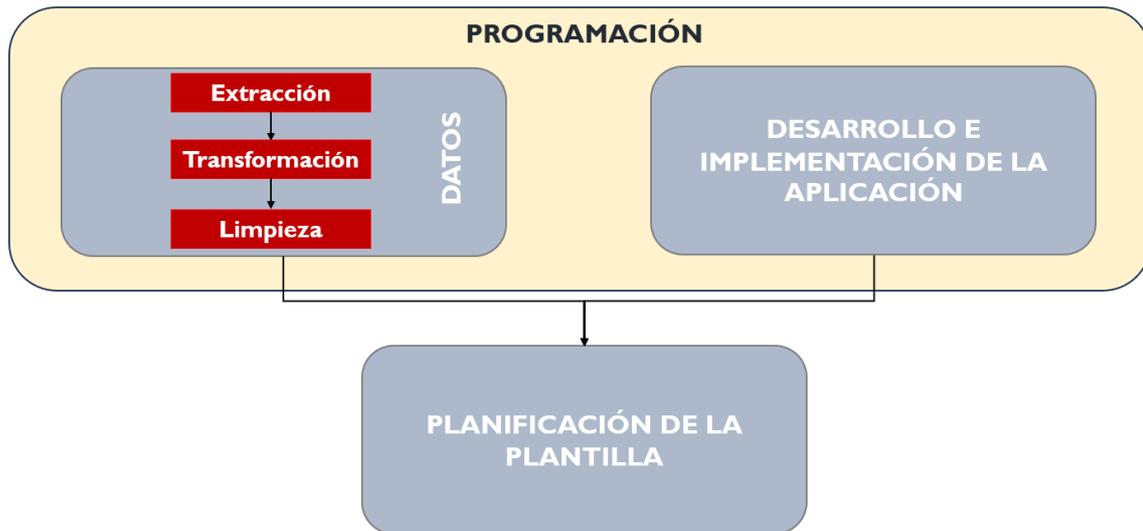


Figura 8: Esquema de procesos para realizar una planificación de plantilla (Fuente: Elaboración propia)

4.1. Lenguajes de programación y software utilizado

Para la extracción y transformación de los datos, la implementación de código y la generación de la aplicación, es necesario utilizar lenguajes de programación que dispongan de librerías que permitan realizar una gestión adecuada de bases de datos. Esta interpretación de los conjuntos de datos se realiza mediante herramientas de aprendizaje automático.

Según (Santander, 2021), Python es un lenguaje de programación de alto nivel, cada vez más demandado en el ámbito laboral y utilizado en el ámbito académico. Además, es ampliamente reconocido por la comunidad de programadores gracias a su facilidad de uso y legibilidad del código, lo que hace que sea un lenguaje de programación bastante común para personas que se quieren iniciar en el mundo de la codificación. La Figura 9 muestra el logo oficial de este lenguaje.



Figura 9: Logo oficial de Python

Debido a su naturaleza de código abierto, el desarrollo de Python ha experimentado un crecimiento exponencial gracias a la colaboración de la comunidad de programadores. Esto ha llevado a muchas empresas de diversos sectores a adoptar Python debido a su versatilidad. Python se utiliza para el desarrollo de software y aplicaciones web básicas, como la desarrollada para este trabajo, y, además, es el lenguaje preferido para implementar algoritmos de inteligencia artificial, aprendizaje automático y manejo de grandes volúmenes de datos (lo que, en inglés, se conoce como Big Data).

Entre las características clave de Python se destaca su facilidad de codificación y legibilidad del código, su enfoque en la programación orientada a objetos y funcional, su capacidad para ser utilizado en múltiples plataformas y, no menos importante, su amplia variedad de librerías disponibles.

En cuanto al entorno de desarrollo, se ha hecho uso de Spyder, que es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) obtenido a partir de Anaconda Navigator. El logo oficial de Spyder se muestra en la Figura 10.



Figura 10: Logo oficial de Spyder

Anaconda Navigator es una interfaz gráfica de usuario (GUI) y gestor de paquetes de software de programación. Esta plataforma es muy popular para los proyectos relacionados con la ciencia y la analítica de datos, así como el aprendizaje automático. Es una interfaz relativamente sencilla de utilizar y que permite visualizar con facilidad las distintas herramientas y entornos disponibles dentro del ecosistema que ofrece Anaconda. El logo oficial de Anaconda Navigator se muestra en la Figura 11.



Figura 11: Logo oficial de Anaconda Navigator

Este entorno de programación cuenta con un editor de texto, una consola interactiva y un explorador de variables. Estos entornos facilitan el proceso de desarrollo, proporcionando características como resaltado de sintaxis, autocompletado de código y herramientas de control de versiones, lo que mejora la productividad del desarrollador.

A continuación, se procede a dar información extraída de (Immune Technology Institute, 2022) acerca de las librerías a utilizar para la codificación de la aplicación:

- 1) **Streamlit:** Streamlit es un framework de código abierto diseñado específicamente para la creación rápida y sencilla de aplicaciones web interactivas en Python. Con su sintaxis sencilla y su enfoque en la visualización de datos, Streamlit permite construir interfaces de usuario dinámicas y elegantes con solo unas pocas líneas de código y sin necesidad de conocer lenguajes específicos de desarrollo web como JavaScript u otros. Esta librería se utilizará como el núcleo de la aplicación para crear la interfaz y gestionar la interacción con el usuario, en este caso, el director deportivo o los scouts que compongan la dirección deportiva. En la Figura 12, se puede ver el logo oficial de la librería Streamlit.



Figura 12: Logo oficial de Streamlit

- 2) **Pandas:** Pandas es una herramienta fundamental para el análisis y manipulación de datos en Python. A partir del conjunto de datos, obtenido desde un archivo Excel o CSV o incluso extrayendo tablas desde páginas web, es capaz de proporcionar estructuras de datos flexibles como DataFrames. Estas estructuras de datos permiten la carga, limpieza y transformación de datos de forma sencilla. Pandas será empleada para la gestión y procesamiento de la base de datos de jugadores, así como para realizar operaciones estadísticas y de filtrado. En la Figura 13, se puede observar el logo oficial de Pandas.



Figura 13: Logo oficial de Pandas

- 3) **Scikit-learn:** Scikit-learn es una biblioteca ampliamente utilizada en el ámbito del Machine Learning. Ofrece una amplia variedad de algoritmos de aprendizaje automático, herramientas de preprocesamiento de datos y funciones de evaluación de modelos. En esta aplicación, Scikit-learn se empleará para desarrollar el sistema de recomendación de jugadores, aplicando algoritmos como el filtrado colaborativo para generar recomendaciones personalizadas. La Figura 14 muestra el logo oficial de Scikit-Learn.



Figura 14: Logo oficial de Scikit-Learn

- 4) **Matplotlib:** Matplotlib es una librería de visualización de datos ampliamente utilizada en Python que permite la creación de gráficos y visualizaciones interactivas, como gráficos de dispersión, gráficos de barras o líneas, entre otros. Con sus numerosas funciones y opciones de personalización, Matplotlib proporciona una forma flexible y potente de presentar los resultados del análisis de datos y mostrar información relevante al usuario. En esta aplicación, Matplotlib se empleará para visualizar gráficos estadísticos y resultados de recomendaciones. La Figura 15 muestra el logo oficial de Matplotlib.



Figura 15: Logo de Matplotlib

- 5) **Numpy:** Su nombre significa "Numerical Python", es una biblioteca de Python ampliamente utilizada en el campo de la informática científica y la computación numérica. Proporciona soporte para trabajar con arreglos multidimensionales y matrices, junto con una colección de funciones matemáticas de alto rendimiento para realizar operaciones en estos arreglos. NumPy, cuyo logo oficial se encuentra en la Figura 16, es una de las bibliotecas más esenciales en el ecosistema de Python para ciencia de datos, aprendizaje automático y computación científica en general.



Figura 16: Logo de NumPy

- 6) **SQLite3:** La biblioteca SQLite3 es un módulo de la biblioteca estándar de Python que proporciona una API para interactuar con bases de datos SQLite. Esta librería permite:
- **Conexión a bases de datos SQLite:** Permite conectar y gestionar bases de datos SQLite

directamente desde Python. Puedes crear una conexión a una base de datos SQLite existente o crear una nueva base de datos en el sistema de archivos local.

- **Creación de tablas:** Se pueden ejecutar instrucciones SQL para crear tablas en una base de datos SQLite. Esto incluye definir la estructura de la tabla, sus columnas y restricciones.
- **Manipulación de datos:** Permite utilizar instrucciones SQL para insertar nuevos datos en tablas existentes, actualizar registros existentes o eliminar registros según se requiera en cada caso.
- **Consultas SQL:** Permite ejecutar consultas SQL para recuperar datos de la base de datos. Puedes utilizar SQL estándar para filtrar, ordenar y agrupar datos según tus requisitos.

Esta librería, cuyo logo se encuentra en la Figura 17, que precisa de conocimientos básicos de programación en SQL, se usará principalmente para habilitar a los trabajadores de la Dirección Deportiva Data Driven para poder realizar informes de rendimiento de jugadores, como se explicará más adelante en la programación de la aplicación.



Figura 17: Logo de SQLite

- 7) **Requests y BeautifulSoup:** Según (Gómez, s.f.), Requests y BeautifulSoup son librerías de Python que permiten realizar labores de 'webscraping'. El webscraping, es una técnica para extraer datos de una página web con Python sin tener que hacer uso de archivos csv o excels previamente cargados de forma local en el entorno de trabajo, sino que directamente se accede a la página web. Por una parte, Requests es una librería que realiza peticiones web para extraer la información, mientras que BeautifulSoup permite extraer contenido en formato HTML. El logo de esta librería se puede observar en la Figura 18.



Figura 18: Logo de BeautifulSoup

Por último, para poder lanzar la aplicación con el fin de visualizar e interactuar con la misma, se debe

hacer uso de Jupyter Lab. Jupyter Lab, según su web oficial (Jupyter, s.f.), es un entorno de código abierto de desarrollo web que permite a los usuarios acceder a la interfaz desarrollada en entornos de programación de forma práctica e interactiva, como en este caso, para proyectos desarrollados en Python.

Algunas de las ventajas del uso de Jupyter Lab son:

- **Gran variedad de elección de lenguajes:** Soporta más de 40 lenguajes de programación distintos como Python o R. Además, admite una amplia variedad de tipos de archivos como scripts de Python, documentos Markdown, imágenes, entre otros.
- **Compartir cuadernos:** Jupyter Lab ofrece la capacidad de compartir cuadernos con otros usuarios mediante Github, Jupyter Notebook Viewer e, incluso, por correo.
- **Interactividad:** Da la oportunidad de ofrecer resultados interactivos con los que el usuario puede visualizar de manera rápida e intuitiva el feedback otorgado por la aplicación.
- **Integración con sistemas de Big Data:** Puede integrarse con herramientas de Big Data como Spark o Scala. Además, se puede hacer exploración de datos, como se pretende hacer en este proyecto, con librerías específicas de Python, entre las que se encuentran Pandas, Scikit-Learn, TensorFlow, entre otras.

Se ha optado por utilizar Jupyter Lab porque, para lanzar la aplicación con la ayuda de la librería Streamlit de Python, es el entorno de desarrollo que mejor permite visualizar el contenido de la aplicación e interactuar con ella de forma sencilla y amigable para cualquier miembro de la dirección deportiva. En la Figura 19 se puede visualizar el logo de Jupyter Lab.



Figura 19: Logo de Jupyter Lab

Para la generación de informes y visualización de datos, se ha hecho uso de Power BI. Según su propia página web (Microsoft, n.d.), Power BI es una plataforma de inteligencia empresarial (Business Intelligence, en inglés) que funciona como herramienta de visualización y análisis de datos que se utiliza ampliamente en diversos campos. Permite a los usuarios conectarse a los datos, visualizarlos e incorporarlos en objetos visuales o informes.

En el contexto de un trabajo relacionado con la gestión deportiva, Power BI se convierte en una herramienta esencial para analizar datos relacionados con el desempeño de los jugadores, el rendimiento del equipo y otros aspectos clave, que se definirán posteriormente.

Los usuarios pueden conectar múltiples fuentes de datos, como bases de datos, hojas de cálculo y servicios en la nube, y luego utilizar la interfaz intuitiva de Power BI para crear gráficos, informes interactivos y paneles de control personalizados. Lo que distingue a Power BI del resto de softwares de visualización de datos es su capacidad para proporcionar una visión en tiempo real de los datos, lo que permite a los profesionales del deporte tomar decisiones basadas en información actualizada y relevante.

En la Figura 20, se puede ver el logo de esta aplicación:



Figura 20: Logo de Power BI

4.2. Proceso ETL

El procesamiento de datos es una componente fundamental en el ciclo de vida de la información y gestión de la analítica de datos. Consiste en una serie de operaciones que se aplican a los datos para convertirlos en información útil, significativa y confiable. El proceso abarca desde la adquisición y extracción de datos crudos hasta su transformación y preparación para su posterior análisis y uso.

Una de las fases clave del procesamiento de datos es lo que se conoce como ETL, que representa las siglas de "Extracción, Transformación y Carga" (Extract, Transform, Load en inglés). El proceso ETL es un conjunto de técnicas y operaciones diseñadas para gestionar datos de manera efectiva, y es especialmente relevante en el ámbito de la gestión de bases de datos y el análisis de datos.

El proceso ETL es esencial para garantizar la integridad y la calidad de los datos antes de que sean utilizados en aplicaciones analíticas, informes, visualizaciones u otros fines. Además, permite la integración de datos de diversas fuentes en un único repositorio, lo que facilita la toma de decisiones basadas en datos y el descubrimiento de información valiosa.

4.2.1. Extracción de los datos

La etapa de extracción de datos se encarga de recopilar información de una amplia variedad de fuentes, que pueden incluir bases de datos propias o procedentes de servicios web, aplicaciones, entre otros. El proceso de extracción puede implicar diferentes métodos y técnicas, dependiendo de la naturaleza de las fuentes de datos. Estas pueden ser estructuradas, como bases de datos relacionales, o no estructuradas, como archivos de texto sin formato o contenido web. Además, los datos pueden estar distribuidos y almacenados en diferentes formatos, lo

que agrega complejidad al proceso de extracción.

El factor más importante a la hora de aplicar cualquier algoritmo o programa basado en datos es contar con una base de datos con información actualizada y fidedigna. Un buen conjunto de datos permite recopilar información completa y útil sobre el rendimiento ofensivo y defensivo de los jugadores durante la temporada, de cara al posterior mercado de fichajes. Contar con datos precisos y actualizados es esencial para obtener recomendaciones que se ajusten a la realidad.

En los clubes de fútbol, como se indicó en la gestión de recursos de la Dirección Deportiva Data Driven, en la mayoría de los casos se utilizan proveedores de datos, los cuales se encargan de recopilar y estructurar los datos para que sean los clubes, bajo los requerimientos y necesidades de cada uno, los que se encarguen de realizar la interpretación de estos. Al ser estas herramientas de pago, para realizar este proyecto, se han buscado alternativas para poder acceder a datos de calidad sin realizar un desembolso económico. Es aquí donde entra en juego la página web FBref.

FBref, cuyo logo se puede ver la Figura 21, es una página web que ofrece métricas de rendimiento de jugadores de fútbol de todo el mundo. Concretamente, como se puede extraer de su propia página web (FBref, s.f.), esta página cuenta con datos que cubren:

- 47 países
- 144 competiciones
- 5.169 equipos
- 196.489 jugadores
- 334.153 partidos



Figura 21: Logo de FBref

FBref utiliza datos proporcionados por Data Sports Group y por Opta (proporcionando esta última la mayoría de las métricas avanzadas). Además de contar con un número muy importante de datos almacenados en su web, ofrece la posibilidad de descargar los datos en formato csv o, incluso, poder hacer la descarga directamente implementando líneas de código de extracción html (o *webscraping*, como se denomina la extracción de datos directamente desde páginas web en inglés) como se podrá observar posteriormente, lo que facilitará, en gran medida, el proceso de extracción de los datos.

Para la realización de este proyecto, se ha hecho uso de los datos extraídos de FBref de jugadores y

equipos de la temporada pasada de las cinco grandes ligas.

Para dimensionar estas bases de datos extraídas y contextualizar la cantidad de datos que se va a utilizar, se particularizan ciertos datos de las dimensiones:

- **Ligas consideradas:** La Liga (España), Premier League (Inglaterra), Bundesliga (Alemania), Ligue 1 (Francia), Serie A (Italia), Brasileirao (Brasil), Eredivisie (Holanda), Liga NOS (Portugal).
- **Jugadores considerados:** 4549 (4209 jugadores de campo y 340 porteros).
- **Métricas analizadas:** 68 (51 para jugadores de campo y 17 para porteros).
- **Total de datos considerados:** 309.332 datos.

4.2.1.1. Métricas relevantes

Debido a la importancia de la calidad de los datos, uno de los procesos más relevantes a la hora de diseñar la estrategia de recolección de datos es definir qué métricas son más importantes a la hora de analizar a un futbolista. Dentro de la página web FBref, hay una gran cantidad de métricas disponibles de diferente naturaleza, por lo tanto, es importante realizar un estudio previo de concreción de qué métricas son las más importantes a analizar pormenorizadamente.

En este punto, es importante contextualizar, para los lectores no familiarizados con el mundo del fútbol, las distintas demarcaciones del terreno de juego. Esto es relevante porque, a la hora de definir las métricas objeto de estudio en la base de datos a configurar, en función de la naturaleza de estas, van a ser más importantes para un jugador u otro en función del rol que desempeñan en el campo. Se toma como referencia una disposición táctica de 1-4-2-3-1, como se puede observar en la Figura 22.



Figura 22: Esquema táctico 1-4-2-3-1 (Luque, 2013)

Siguiendo la numeración dispuesta en la Figura 22, se definen las posiciones que servirán como base para,

posteriormente, definir las métricas que se encontrarán en la base de datos extraída de la web FBref y, también se determina el conjunto de métricas más importantes a considerar para cada uno de ellos.

- **1 (Portero):** Jugador encargado en defender la portería propia de los ataques del rival, evitando así recibir gol. Las métricas más relevantes para esta demarcación son las específicas de portero.
- **2-3 (Lateral):** Jugador que forma parte de la línea defensiva (línea de 4) y que juega en el costado. Tiene como objetivo principal defender los ataques exteriores del equipo rival. Cada vez más, en el fútbol moderno, se da a los laterales el rol de tener más presencia ofensiva ya que dotan al equipo de una mayor amplitud a la hora de atacar, lo que hace que los rivales tengan más dificultades a la hora de defender. Por tanto, para analizar el rendimiento de los laterales, hay que tener en cuenta todas las métricas salvo las de portero (las que se conocen como métricas de campo), dando más importancia a las métricas defensivas u ofensivas en función de los requerimientos del entrenador para dicha demarcación.
- **4-5 (Central):** Jugador que, al igual que el lateral, pertenecen a la línea defensiva. Su principal cometido es defender la zona central de la línea defensiva, habitualmente del delantero del equipo rival. Deben destacar en juego aéreo, tener potencia física y, habitualmente, contar con buena salida de balón (también en función de los requerimientos concretos del entrenador). Las métricas más relevantes en esta posición son las defensivas y, adicionalmente, en función de las funciones que le quiera otorgar el entrenador, las métricas con balón.
- **6-7 (Mediocentro):** Posición más versátil dentro de todas las que hay en el campo ya que, en función del rol a desempeñar, se pueden definir los siguientes tipos de mediocentros, según (Herrera Gómez-Platero, 2019):
 - **Mediocentro defensivo:** También denominado pivote, es el encargado de, por una parte, referenciar la salida de balón del equipo y servir de apoyo a los demás jugadores y, por otra, tener una comprensión de la presión y los espacios suficiente para defender de forma correcta los ataques del rival. Es por ello por lo que, para evaluar su rendimiento, hay que incidir sobre todo en las métricas defensivas y con balón.
 - **Mediocentro posicional:** Es el encargado de apoyar la salida de balón junto al mediocentro defensivo. Su principal función es realizar la transición de balón entre la línea defensiva y la línea ofensiva, sirviendo de enlace entre los defensas y los atacantes. Por lo tanto, las métricas más relevantes para este tipo de mediocentro son las que tienen que ver con la posesión de balón, teniendo en cuenta también tanto las métricas defensivas como las ofensivas.
 - **Volante:** También llamado interior o, en su denominación inglesa, box-to-box. Sus funciones son parecidas a las del mediocentro posicional, pero la principal diferencia entre uno y otro es la manera en la que enlazan el juego entre defensa y ataque. Por una parte, el mediocentro posicional prioriza la posesión de balón para poder enlazar la defensa con el ataque, mientras

que el volante basa mucho más su juego en el apartado físico (con conducciones de balón, llegada desde segunda línea, entre otros), abarcando mucha más cantidad de campo que un mediocentro puro. Por lo tanto, las métricas relevantes para este tipo de jugador son las ofensivas y, en menor medida, las relativas a la posesión de balón.

- **8-11 (Extremo):** Jugador ofensivo que desarrolla su juego por las bandas en el campo, cuya misión principal es surtir de balones al delantero e, incluso, producir goles y asistencias. Por su naturaleza, las métricas más importantes a la hora de evaluar su rendimiento son las ofensivas y las relativas a la posesión de balón.
- **10 (Mediapunta):** Jugador que se sitúa directamente por detrás del delantero, siendo el centrocampista más adelantado del equipo. Debe tener una gran capacidad para leer el juego de manera rápida para poder elegir la mejor opción de pase que, en muchas ocasiones, será el último pase del ataque. Por lo tanto, tienen predominancia las métricas de posesión de balón y, en mayor medida que los jugadores del centro del campo, las métricas ofensivas.
- **9 (Delantero centro):** Jugador más adelantado del equipo y responsable de marcar el mayor número de goles posible. Al ser el jugador más determinante, lo más natural es que sea el jugador más caro de la plantilla, lo que hace su adquisición en el mercado una operación relativamente complicada. Las métricas ofensivas, en este caso, son las más relevantes.

Esta definición posicional se basa en una disposición 1-4-2-3-1, pero, como es natural, en función de la disposición táctica y las particularidades posicionales que el entrenador pretenda imponer en su estilo de juego, la naturaleza y rol de los jugadores puede variar. Por ello es importante, como primer paso de la planificación, definir exactamente cuál es la idea de juego del entrenador y plasmarla en un sistema táctico concreto, lo que ayudará al director deportivo a analizar de manera más rápida y sencilla, posición por posición, las necesidades concretas del equipo.

A continuación, con el fin de comprender de mejor forma las métricas a las que se ha hecho referencia en la definición de las posiciones del campo, se va a realizar una descripción detallada de todas las métricas que se van a usar en el proyecto a partir del servidor de datos utilizado (se ha respetado la nomenclatura de las métricas de la web FBref), agrupándolas en función de la naturaleza de cada una:

- **Métricas estándar:** Estas métricas resumen la información principal y básica del futbolista. En este grupo, se encuentran las siguientes métricas: ‘Jugador’, ‘País de nacimiento’, ‘Posición de juego’, ‘Competición’, ‘Equipo actual’, ‘Edad’, ‘Partidos jugados’, ‘Partidos como titular’, ‘Minutos jugados’, ‘Tarjetas amarillas’ y ‘Tarjetas rojas’. Estas métricas son de igual relevancia para todas las demarcaciones del campo, ya que se limitan a definir las características principales sin hacer hincapié en la naturaleza del juego.
- **Métricas de porteros:** Estas métricas son específicamente dedicadas al rendimiento de los porteros, ya que, por la propia naturaleza de juego, las estadísticas a analizar para un portero no son las mismas con

respecto a las de un jugador de campo (entiéndase por jugador de campo todos los demás jugadores).

Entre las métricas analizadas dentro de este grupo, se encuentran:

- **GC:** Goles en contra. No es una métrica cuyo mérito vaya a recaer en el portero enteramente, ya que una buena defensa colabora en el hecho de que se reciban menos goles, pero igualmente, el rendimiento del portero en este aspecto es lo suficientemente relevante como para que se considere esta métrica.
- **Salvadas:** Número de paradas realizadas.
- **% Salvadas:** Es una métrica compuesta de otras métricas (en adelante, se definen este tipo de métricas como métrica compuesta). Se define de la siguiente forma:

$$\% \text{ Salvadas} = \frac{\text{Disparos a puerta en contra} - \text{Goles en contra}}{\text{Disparos a puerta en contra}}$$

- **PaC:** Porterías a cero. Ocurre lo mismo que con la métrica 'GC' (Goles en contra).
- **PaC%:** Porcentaje de porterías a cero.
- **PD:** Penaltis parados.
- **% Salvadas_Pen:** Porcentaje de penaltis parados sobre el total de penaltis en contra.
- **PSxG:** Es una métrica avanzada ofrecida por Opta, cuyas siglas en inglés significan Post-Shot Expected Goal, que, en español, significa goles esperados después de tiro. Al igual que el Expected Goal (métrica que se explicará posteriormente), el PSxG es la probabilidad de que un tiro sea gol en función de variables situacionales en el momento del disparo (zona a la que se dirige el balón, colocación del portero en el momento del tiro, velocidad del balón, entre otros).
- **PSxG/SoT:** Métrica compuesta que relaciona el PSxG de un portero respecto a los tiros a puerta recibidos.
- **PSxG+/-:** Métrica compuesta que relaciona el PSxG de un portero con los goles recibidos. Esta métrica evalúa, entonces, la capacidad del portero de recibir menos goles de los esperados, de tal forma que, si esta métrica arroja un valor positivo, el portero estará recibiendo menos goles de los esperados y viceversa.
- **Cmp:** Pases largos (más de 35 metros) completados.
- **%Cmp:** Porcentaje de pases largos (más de 35 metros) completados.
- **Long. Prom.:** Longitud promedio de los pases en metros.
- **Stp:** Centros al área atajados por el portero.
- **% de Stp:** Porcentaje de centros atajados al área atajados por el portero sobre el total de centros.
- **Núm. de OPA:** Número de acciones defensivas realizadas fuera del área.

- **DistProm.:** Distancia promedio desde la portería propia de las acciones defensivas realizadas fuera del área.
- **Métricas defensivas:** Estas métricas están centradas en analizar el rendimiento defensivo de los jugadores. Son de especial relevancia para los jugadores de la parcela defensiva, o lo que es lo mismo,
 - **Tkl:** Número de tackles (entradas, en español) efectuadas.
 - **TklG:** Número de tackles (entradas, en español) ganados. Una entrada es un intento de defensa de arrebatarle la posesión de balón al jugador del equipo rival.
 - **Bloqueos:** Métrica compuesta que suma los disparos y los pases bloqueados.
 - **Tkl+Int:** Métrica compuesta que suma la cantidad de entradas más el número de intercepciones de balón intentadas.
 - **Fls:** Número de faltas cometidas.
 - **FR:** Número de faltas recibidas.
 - **PA (inversa):** Número de veces que un futbolista cae en fuera de juego. Se trata de una métrica inversa, cuanto mayor es su valor, peor es el rendimiento defensivo del jugador.
 - **Recup.:** Número de balones sueltos recuperados.
 - **% de ganados:** Porcentaje de duelos aéreos ganados.
 - **Penal concedido (inversa):** Número de penaltis concedidos por un jugador.
- **Métricas con balón:** Estas métricas se centran en analizar la capacidad técnica del jugador con balón, especialmente relevantes para centrocampistas. Valoran la capacidad asociativa del jugador, así como su visión de juego y rapidez mental para ver espacios en la línea defensiva del rival.
 - **PrgP:** Número de pases completados que desplazan el balón, como mínimo, 9 metros hacia la portería rival (en adelante, pases progresivos).
 - **Cmp:** Número de pases totales completados.
 - **%Cmp:** Porcentaje de pases totales completados.
 - **Dist. Tot.:** Distancia total recorrida, medida en metros, por el balón a través de los pases realizados.
 - **Dist. Prg:** Distancia total recorrida, medida en metros, por el balón hacia la portería rival a través de los pases realizados. Para entender esta métrica, se debe entender qué es un pase progresivo. Un pase progresivo es aquel que no es en horizontal ni hacia atrás, o, lo que es lo mismo, que reduce la distancia a la portería rival.
 - **Cmp_Cor:** Pases cortos completados (entre 4,5 y 14 metros).

- **% Cmp_Cor:** Porcentaje de pases cortos completados (entre 4,5 y 14 metros).
 - **Cmp_Med:** Pases medios completados (entre 14 y 27,4 metros).
 - **% Cmp_Med:** Porcentaje de pases de media distancia completados (entre 14 y 27,4 metros).
 - **Cmp_Lar:** Pases largos completados (más de 27,4 metros).
 - **% Cmp_Lar:** Porcentaje de pases largos completados (más de 27,4 metros).
 - **PC:** Número de pases clave efectuados o, lo que es lo mismo, pases que preceden a un tiro.
 - **Dist. Prg. (Circulación):** Distancia progresiva recorrida en metros con el balón en los pies.
 - **1/3 (Circulación):** Distancia recorrida en el último tercio del campo en metros con el balón en los pies.
 - **Errores de control (inversa):** Métrica inversa que evalúa el número de errores a la hora de controlar el balón.
 - **Des (inversa):** Número de veces que un jugador pierde el control del balón a partir de una entrada del rival.
 - **Camb.:** Cambios de orientación realizados por un jugador. Un cambio de orientación es un pase largo (35 metros o más) que va de un lado del campo a otro, con el fin de hacer bascular al equipo rival y propiciar la generación de huecos en la defensa.
 - **SE:** Saques de esquina acometidos por el jugador.
 - **Att:** Número de regates intentados.
 - **Succ:** Número de regates satisfactorios, superando al rival.
 - **PPA:** Número total de pases al área rival.
 - **PrgC:** Número de acarreos o conducciones de balón que lo acercan, al menos, 9 metros respecto a la posición donde se recibe el balón.
- **Métricas ofensivas:** Estas métricas están centradas en medir la capacidad del jugador de producir acciones de ataque frente al equipo rival. Son de suma importancia, como es de esperar, para los jugadores de perfil ofensivo, cuyo cometido principal es generar la mayor cantidad de acciones de peligro posible.
- **Gls:** Número de goles marcados.
 - **Ass:** Número de asistencias proporcionadas.
 - **xG:** La métrica xG (Expected Goals o Goles Esperados) se ha convertido en la métrica avanzada más utilizada en el mundo del fútbol y del periodismo deportivo debido a su capacidad para proporcionar un análisis detallado y objetivo del desempeño de los equipos y

jugadores en el terreno de juego. Según (Whitmore, 2021), esta métrica fue introducida por Sam Green de Opta en 2012.

El xG calcula la probabilidad de que un disparo a portería se convierta en gol a partir del contexto situacional de la jugada. Las variables más comunes que se tienen en cuenta para el análisis son la posición del disparo, el ángulo del disparo, la distancia a la portería, el tipo de jugada, la parte del cuerpo con la que se realiza el disparo, la cantidad de jugadores rivales entre el balón y la portería, entre otros.

Una de las razones principales por las cuales la xG se ha popularizado en el periodismo deportivo es su capacidad para agregar valor a las narrativas de los partidos. En lugar de centrarse únicamente en los goles marcados, la xG permite analizar la calidad de las oportunidades de gol creadas y desperdiciadas por cada equipo. De esta manera, un equipo puede haber ganado un partido por un margen estrecho, pero si su xG fue significativamente mayor que el del rival, esto indicaría que tuvieron más y mejores oportunidades de gol, lo cual proporciona un contexto valioso para entender el rendimiento real de ambos equipos.

Particularmente aplicado al rendimiento individual del jugador, la métrica xG ayuda a identificar a aquellos jugadores que pueden estar teniendo un buen desempeño ofensivo, a pesar de no haber marcado muchos goles, o viceversa. Por ejemplo, un delantero que ha tenido varias oportunidades de gol con una alta xG, pero no ha logrado convertirlas, podría estar mostrando un buen rendimiento que no se refleja en sus estadísticas tradicionales de goles anotados.

Es importante tener en cuenta que cada proveedor de datos aplica su propia metodología de cálculo de xG, teniendo en cuenta y/o dando mayor o menor peso a ciertas variables, lo que hace que pueda variar ligeramente el valor de xG calculado por un proveedor de datos respecto de otro.

En la Figura 22, se puede observar un ejemplo aplicado por Opta Analyst del cálculo del xG del jugador Gabriel Jesús durante la temporada 2019-2020 jugando en el Manchester City F.C. En este gráfico se pueden observar, en morado, la situación en el terreno de juego de todos los tiros de Gabriel Jesús durante la temporada que no acabaron en gol y, en rojo, los tiros que sí fueron gol. El tamaño del círculo representa el xG del tiro en cuestión, teniendo mayor diámetro aquellos tiros con un xG mayor y viceversa.

A partir de los datos, se puede observar que este jugador marcó menos goles que los esperados, lo que implica que, en esa temporada, tuvo una eficacia goleadora baja (este aspecto se concretará en mayor medida, posteriormente, en el detalle de la métrica compuesta G-xG).

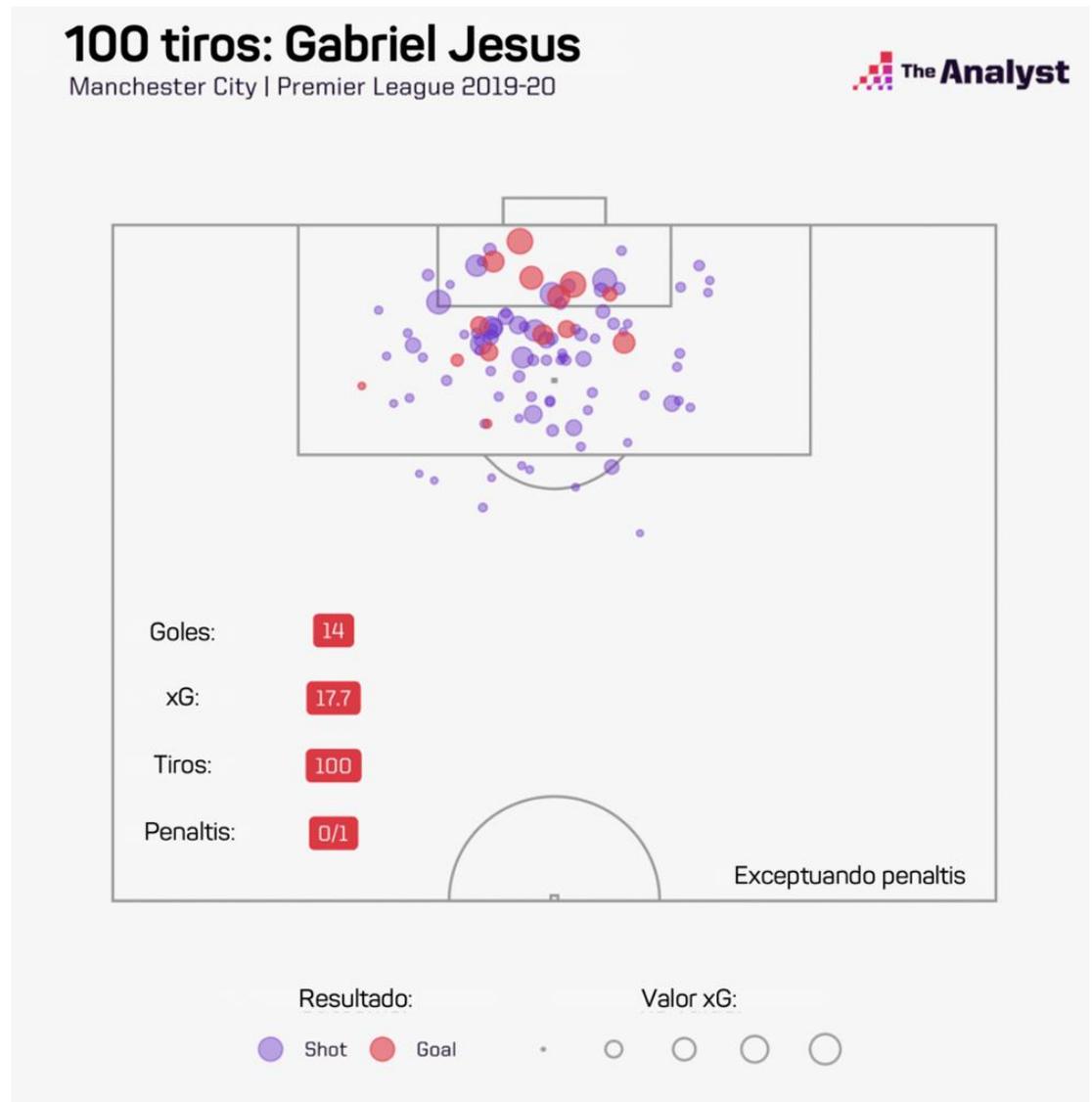


Figura 23: Cálculo de xG de Gabriel Jesús en la temporada 2019-2020 (Whitmore, 2021)

- **xA:** Esta métrica avanzada es equivalente al xG, pero aplicada a la capacidad asistencial del jugador. Evalúa la probabilidad de que un pase o un centro resulte en una asistencia de gol. De igual manera que el xG, para su cálculo, se tienen en cuenta variables situacionales para su estimación, como pueden ser el tipo de pase, la posición del receptor del balón, la distancia y el ángulo del pase, la organización defensiva del equipo rival, la habilidad del receptor, entre otros factores.
- **% de TT:** Porcentaje de tiros a puerta respecto al total de disparos efectuados. Esta métrica evalúa la efectividad del jugador a la hora de dirigir correctamente sus tiros hacia la portería rival.
- **Dis:** Disparos efectuados.
- **DaP:** Disparos a puerta efectuados.

- **Penal ejecutado:** Número de penaltis lanzados por el jugador.

Todas las variables o métricas consideradas son numéricas, salvo las siguientes: ‘Jugador’, ‘País’, ‘Posición 1’, ‘Posición 2’, ‘Equipo’ y ‘Competición’. Teniendo en cuenta que estas variables son descriptivas y se van a utilizar como herramientas de filtrado en el análisis, no se va a realizar ningún tratamiento de estas. Este tratamiento, que se verá en el siguiente epígrafe, estará reservado a las variables numéricas.

4.2.2. Transformación de los datos

Definida la base de datos, el siguiente proceso definido en las etapas de un proceso basado en analítica de datos es preprocesar y transformar el conjunto de datos. Este proceso es importante por varios factores:

- **Eliminación de datos irrelevantes o ruidosos:** Es importante realizar una limpieza adecuada para eliminar datos irrelevantes o, más comúnmente conocido en el argot de los datos como ruido, que podrían afectar negativamente la calidad de las recomendaciones. Los datos irrelevantes pueden introducir sesgos a la hora de aplicar los algoritmos, lo que puede conducir a conclusiones erróneas.
- **Consistencia y coherencia de los datos:** Esto implica aplicar técnicas de normalización y estandarización para que los datos estén en un formato uniforme y para que se pueda establecer comparaciones entre los mismos, de cara a la identificación de patrones. Sin esta normalización, puede darse el caso de que haya variables cuyos datos tengan más peso que otras simplemente por la propia naturaleza de los valores. La consistencia en la estructura y formato de los datos es esencial para garantizar que los algoritmos de machine learning puedan interpretar y analizar correctamente la información.
- **Eliminación de errores:** Los errores en la extracción de datos, como valores incorrectos o datos faltantes, pueden llevar a conclusiones inexactas y afectar la confiabilidad de las recomendaciones, lo que podría afectar negativamente la imparcialidad y objetividad de estas.

Entendida la importancia de realizar un trabajo previo con los datos para asegurar la coherencia del conjunto, se debe explicar qué técnicas de estandarización, normalización y reducción de la dimensionalidad se utilizan comúnmente en la analítica de datos.

4.2.2.1. Estandarización de variables

La estandarización de variables, también conocida como normalización estándar (nombre de la distribución estadística en la que se basa), es una técnica fundamental en la analítica de datos que busca transformar las variables de un conjunto de datos para que sigan una distribución con media cero ($\mu = 0$) y desviación estándar uno ($\sigma = 1$). Para realizar el proceso de estandarización de variables, se deben calcular los z-scores para cada dato de la siguiente forma:

$$z = \frac{x - \mu}{\sigma}$$

Esta transformación permite comparar y analizar diferentes variables en un mismo contexto, mitigando los efectos de las unidades de medida (de especial relevancia en este proyecto, ya que hay variables de distinta naturaleza que se buscan comparar) y facilitando la interpretación de los resultados.

Desde una perspectiva estadística, la estandarización tiene varios aspectos fundamentales que respaldan su aplicación en el análisis de datos. En primer lugar, la normalización reduce el impacto de la magnitud de las variables en los análisis multivariados y en la construcción de modelos predictivos. Si las variables presentan unidades de medida distintas, algunas pueden dominar las demás y afectar el resultado del análisis. La estandarización corrige esta disparidad, alineando las variables en una misma escala.

En la Figura 24, se observan dos curvas de campana. Partiendo de unos datos ya normalizados (primera campana de Gauss), se puede observar que el proceso de estandarización de forma gráfica, obteniéndose una distribución de valores con media cero y desviación típica uno.

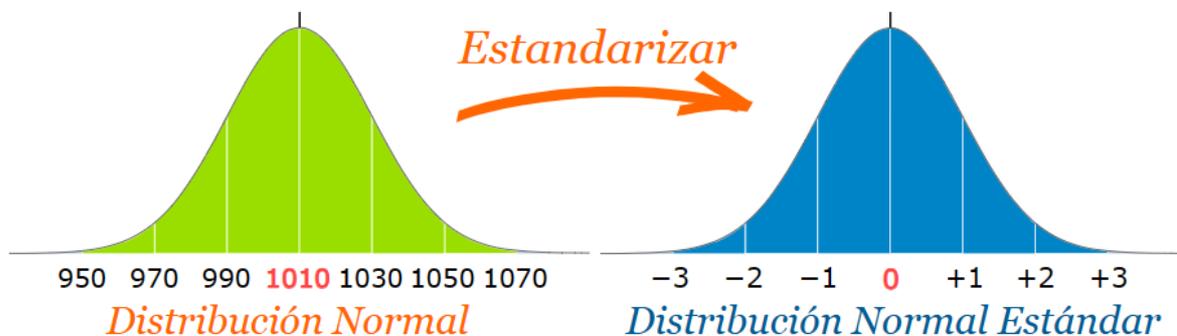


Figura 24: Distribución normal estándar (Disfrutalasmaticas, s.f.)

Además, la estandarización es especialmente relevante cuando se trabajan algoritmos sensibles a la escala, como el gradiente descendente en el aprendizaje automático o algunos métodos de clustering. Estos algoritmos pueden converger más rápido y producir resultados más precisos cuando las variables están estandarizadas.

Otro beneficio de la estandarización es la reducción de la multicolinealidad en el análisis de regresión. La multicolinealidad surge cuando existe una alta correlación entre dos o más variables independientes, lo que puede dificultar la interpretación de los coeficientes y la estabilidad del modelo. La estandarización reduce esta correlación artificial, lo que puede mejorar la calidad de los resultados obtenidos.

Para realizar esta transformación o escalado de datos, se hace uso de la librería Scikit-Learn mencionada previamente en el apartado 4.1., la cual tiene un módulo específico para realizar una estandarización de los datos: StandardScaler. Escalados los datos, lo único que faltaría por asegurar es que no existen datos vacíos en todos los conjuntos de datos, para lo cual, se tiene una función en Python que identifica los valores que se conoce como NaN (Not A Number, en inglés; que no es un número o está vacío, por la traducción al español) y, posteriormente, los considera como 0.

4.2.3. Carga de los datos

A partir del apartado anterior, se tienen los datos transformados y limpios, almacenados en distintos CSVs. Estos CSVs se introducirán en la aplicación a partir de la función `read_csv()` de Pandas, mediante la cual todos los archivos se pueden leer para, posteriormente, almacenarlos en el entorno de programación (en este caso, Spyder) para su posterior uso y manipulación. Al no tener una base de datos de destino o cualquier otro sistema de almacenamiento que sea relevante para el proyecto, no se realiza una interacción con bases de datos complejas u otros sistemas de almacenamiento, lo que simplifica en gran medida la etapa de carga.

4.3. Desarrollo de la aplicación con Streamlit

La herramienta desarrollada tiene como objetivos principales facilitar y ayudar a la dirección deportiva para la interpretación objetiva del rendimiento de los jugadores a través de los datos, el fin de complementar la visión subjetiva de los scouts.

Una vez los datos están almacenados y transformados debidamente, el siguiente paso es la codificación de la herramienta. Para ello, y como se ha hecho hincapié en todo el trabajo, se plantea el uso de técnicas de aprendizaje automático, cuya base es puramente matemática y estadística. A partir de la implementación de estos algoritmos, se obtendrán unos resultados. Como se ha comentado previamente, en función de los datos que se tengan y del resultado que se pretenda, se debe utilizar un tipo de aprendizaje u otro.

Los resultados obtenidos en el procedimiento de aplicación de los algoritmos de Machine Learning se someterán a interpretación por parte del equipo de datos de la dirección deportiva para que, finalmente, dicha interpretación se incluya en los informes finales de los futbolistas que informarán al director deportivo de cuáles son las mejores alternativas a considerar en cada posición del terreno de juego a ojos de los datos. Esta herramienta también dará la capacidad al analista de datos para poder realizar una caracterización del juego del club y, así, poder realizar recomendaciones personalizadas lo más similares posibles al perfil definido por el entrenador.

Como primera aproximación a la aplicación, se deben definir los lenguajes de programación, así como los entornos de desarrollo que se van a utilizar para poder incorporar todas las herramientas a utilizar por parte de la Dirección Deportiva Data Driven. La aplicación propuesta para este proyecto consta de una pestaña de inicio y de 5 pestañas, cuya finalidad no solo está sujeta a la implementación de algoritmos de Machine Learning, sino que también ayudarán a organizar de forma más automatizada las labores de la Dirección Deportiva Data Driven. En este epígrafe se comentará cada una de las pestañas, cuál es el objetivo de cada una de ellas y cómo son soportadas desde un punto de vista de programación. Para visualizar e interactuar con la aplicación, se debe realizar una configuración previa de la misma siguiendo estos pasos:

- 1) **Descarga e implementación de las librerías:** Se debe asegurar que todas las librerías definidas en el apartado anterior se encuentren instaladas debidamente en el entorno de desarrollo, en este caso Anaconda, para poder disponer de todos los algoritmos y funciones necesarios para que la aplicación

funcione correctamente cuando se quiera lanzar en local.

- 2) **Definir ubicación del script de Python de la pestaña ‘Inicio’:** Se debe abrir Anaconda Prompt, que es la línea de comandos utilizada por Anaconda (similar al CMD de Windows), e indicar la carpeta donde se encuentra alojado el script de Python que tiene programada la pestaña ‘Inicio’, que es la que va a lanzar la aplicación y dará la posibilidad de acceder al resto de pestañas.
- 3) **Lanzar Jupyter Lab:** Una vez que Anaconda Prompt verifique que la ruta de la carpeta donde se encuentra el script sea correcta, se debe escribir en otra línea de código: ‘jupyter-lab’. En ese momento, se lanzará en local Jupyter Lab.
- 4) **Lanzar la aplicación:** Por último, dentro de Jupyter Lab, se debe escribir la siguiente línea de código: “streamlit run .\01_Inicio.py”. En ese momento, se lanzará la aplicación en local. En caso de que haya un error de codificación, la ventana que se abre en Jupyter Lab indica cuál es el error de codificación y dónde se encuentra (en qué pestaña y línea de código en concreto).

4.3.1. Pestaña 1: Inicio

En esta pestaña se da información relativa a las funcionalidades de la aplicación, la fuente de datos utilizada, los jugadores, equipos y ligas considerados en la base de datos. Como se puede observar en la Figura 25, la pestaña de inicio se divide en dos partes. En la parte izquierda (en adelante, sidebar, nomenclatura utilizada en Streamlit), se tiene la selección de pestañas, además de información general del contenido de la aplicación.



Figura 25: Visión general (Pestaña 1: Inicio)

En la Figura 26, se observa la información contenida en los llamados por Streamlit como *expander*, que son bloques de contenido que se expanden y se contraen al pulsarlos. En ellos, se incluye información sobre el contenido y el objetivo de cada una de las pestañas, así como de la fuente de extracción de datos y detalles de la base de datos extraída.

¿Qué contiene? ^

Pestaña 1 - Inicio: Página principal de la aplicación que da acceso a las demás.

Pestaña 2 - Asignación de ligas: Pestaña para realizar el reparto de ligas a analizar para cada scout o experto en datos, según corresponda.

Pestaña 3 - Evaluación Scout: Pestaña para realizar, guardar y consultar información relevante sobre los futbolistas objeto de adquisición por parte del club

Pestaña 4 - Rating: Pestaña para analizar los jugadores de la base de datos y rankearlos con un Score, con el fin de definir la mejor opción en cada posición según los requerimientos concretos

Pestaña 5 - Clustering: Pestaña para estudiar parecidos entre futbolistas basados en algoritmos de clasificación de Machine Learning

Extracción de datos ^

Fuente de datos:



Ligas consideradas: La Liga (España), Premier League (Inglaterra), Bundesliga (Alemania), Ligue 1 (Francia), Serie A (Italia), Brasileirao (Brasil), Eredivisie (Holanda), Liga NOS (Portugal)

Jugadores considerados: 4549 (4209 jugadores de campo y 340 porteros)

Métricas analizadas: 68 (51 para jugadores de campo y 17 para porteros)

Figura 26: Información Expanders (Pestaña 1: Inicio)

4.3.2. Pestaña 2: Asignación de Ligas

El objetivo de esta pestaña es realizar el reparto de ligas para cada uno de los trabajadores de la Dirección Deportiva Data Driven, siguiendo las restricciones definidas en el apartado 3.1.1.1. del trabajo (“Determinación de número de profesionales necesarios por cada rol”).

En la Figura 27, se puede observar una visión general de la pestaña:

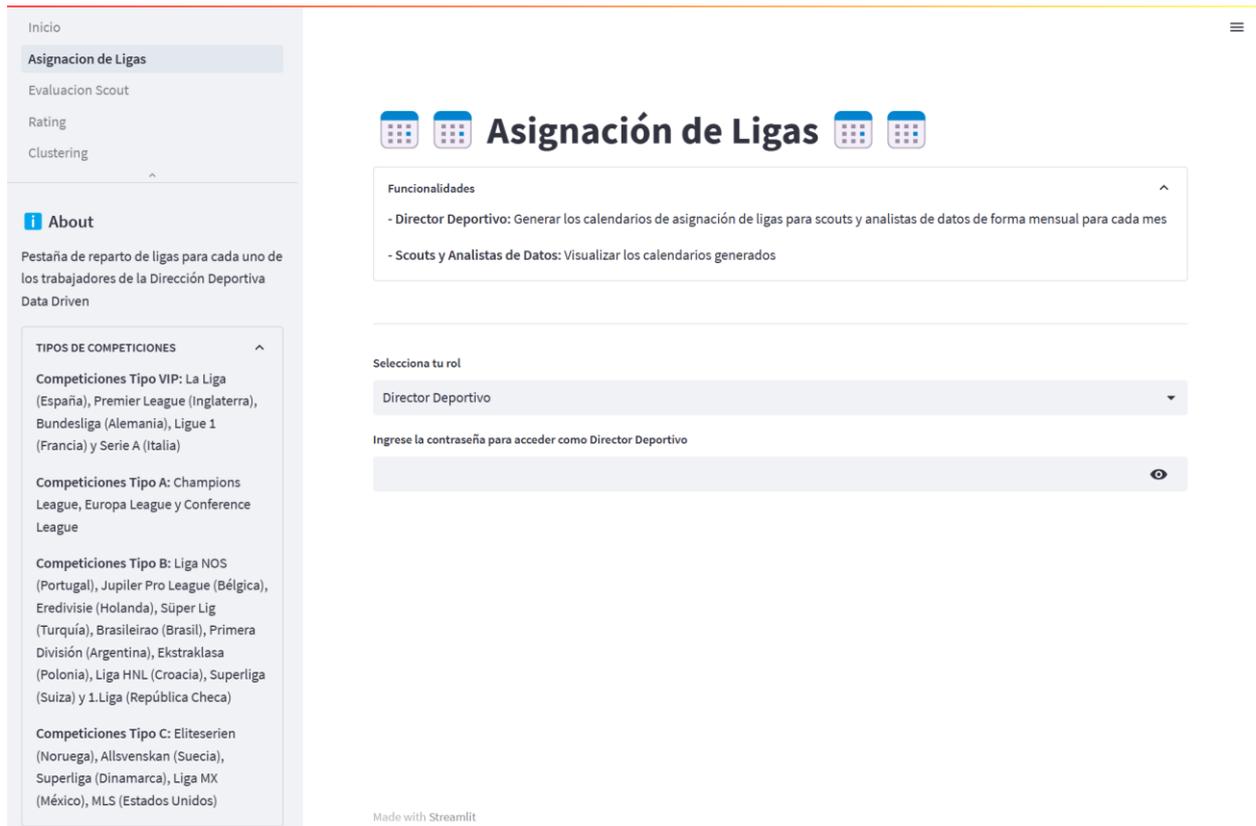


Figura 27: Visión general (Pestaña 2: Asignación de Ligas)

Como se puede observar en la Figura 27, por un lado, en el sidebar, se encuentra información general de la pestaña y un desplegable en el cual se definen los tipos de competiciones, además de las ligas que pertenecen a cada uno de los tipos. Por otro lado, en la parte central, se encuentra el título de la pestaña y, debajo, un selector de rol del trabajador que ha entrado en la pestaña. En este caso, se puede observar que el rol que se ha seleccionado por defecto en el selector es el de director deportivo. Como se puede observar en la Figura 28, el selector consta de tres opciones, que son los trabajadores de la dirección deportiva a los cuales les afecta este reparto de ligas: el director deportivo, el scout y el analista de datos.

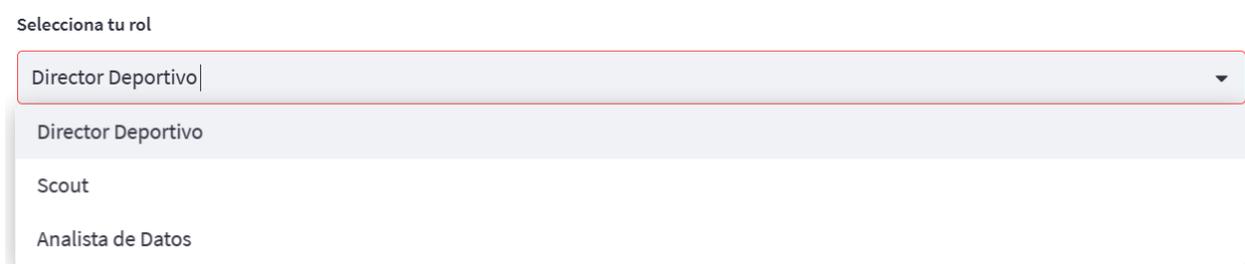
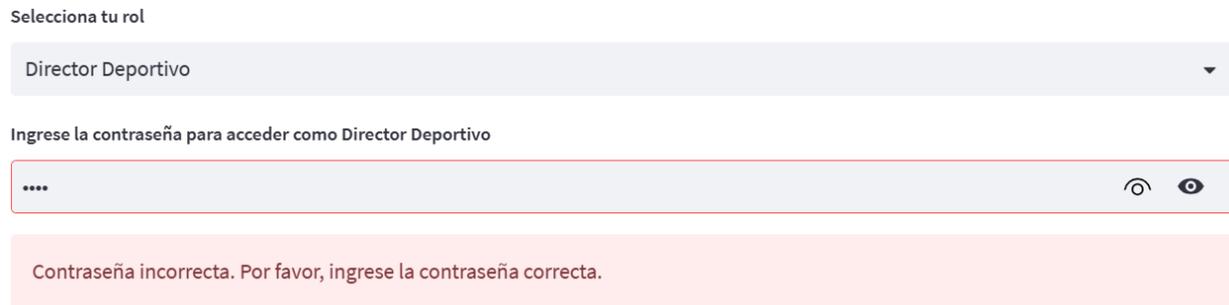


Figura 28: Selector de rol (Pestaña 2: Asignación de Ligas)

El director deportivo es el encargado de simular el reparto de ligas y, además, será el único que tenga acceso al software de asignación. Para cerciorarse de que ningún trabajador tenga acceso a dicho software de

asignación, se ha añadido una contraseña para que solo el director deportivo pueda acceder. En la Figura 29, se ha añadido una captura en la cual se puede observar qué es lo que ocurre cuando se introduce una palabra aleatoria como contraseña.



Selecciona tu rol

Director Deportivo

Ingrese la contraseña para acceder como Director Deportivo

....

Contraseña incorrecta. Por favor, ingrese la contraseña correcta.

Figura 29: Contraseña incorrecta (Pestaña 2: Asignación de Ligas)

Por el contrario, como se observa en la Figura 30, si la contraseña se ha introducido correctamente, se visualizará un mensaje de bienvenida al director deportivo, además de un nuevo selector que preguntará qué asignación se pretende hacer, si la de los scouts o la de los analistas, ya que, como se ha definido previamente, las restricciones para cada uno de los trabajadores a la hora de realizar la asignación son distintas.



Selecciona tu rol

Director Deportivo

Ingrese la contraseña para acceder como Director Deportivo

.....

Contraseña correcta. ¡Bienvenido, Director Deportivo!

Selecciona qué asignación deseas hacer

Scout

Figura 30: Contraseña correcta (Pestaña 2: Asignación de Ligas)

En este caso, como el selector está en la opción 'Scout', se realizará el reparto de ligas para los scouts (con sus correspondientes restricciones). Como esta funcionalidad está dedicada a la fase de 'Mapeo', los meses en los cuales haría falta realizar el reparto son: enero, febrero, marzo, abril, agosto, septiembre y octubre (véase cronograma de planificación: Figura 7). Por último, en las Figuras 31 y 32, se añaden dos ejemplos de repartos de ligas, uno para los scouts y otro para los analistas:

	Scout1	Scout2	Scout3	Scout4	Scout5
Enero	["Premier League_ING"; Jupiler Pro League_BEL; Súper Lig_TUR; Champions League"]	["Ligue 1_FRA"; Primera División_ARG; Europa League; Brasileirao_BRA"]	["La Liga_ESP"; Eredivisie_HOL; Superliga_SUI; Bundesliga_AUS"]	["Bundesliga_ALE"; 1.Liga_CHE; Liga HNL_CRO; Conference League"]	["Serie A_ITA"; Ekstraklasa_POL; Liga NOS_POR"]
Febrero	["La Liga_ESP"; Bundesliga_AUS; Conference League; Ekstraklasa_POL"]	["Serie A_ITA"; Superliga_SUI; 1.Liga_CHE; Brasileirao_BRA"]	["Premier League_ING"; Jupiler Pro League_BEL; Liga HNL_CRO"]	["Ligue 1_FRA"; Europa League; Champions League; Súper Lig_TUR"]	["Bundesliga_ALE"; Primera División_ARG; Liga NOS_POR"]
Marzo	["La Liga_ESP"; Champions League; Superliga_SUI; 1.Liga_CHE"]	["Premier League_ING; Liga HNL_CRO; Eredivisie_HOL; Súper Lig_TUR"]	["Ligue 1_FRA"; Jupiler Pro League_BEL; Primera División_ARG; Bundesliga_AUS"]	["Bundesliga_ALE"; Brasileirao_BRA; Liga NOS_POR; Ekstraklasa_POL"]	["Serie A_ITA"; Europa League; Conference League"]
Abril	["Bundesliga_ALE"; Ekstraklasa_POL; Champions League; Jupiler Pro League_BEL"]	["La Liga_ESP; Bundesliga_AUS; Súper Lig_TUR; Conference League"]	["Ligue 1_FRA; Liga HNL_CRO; Brasileirao_BRA"]	["Serie A_ITA; Primera División_ARG; Superliga_SUI; Europa League"]	["Premier League_ING; 1.Liga_CHE; Liga NOS_POR"]
Agosto	["Serie A_ITA; Liga HNL_CRO; Primera División_ARG; Jupiler Pro League_BEL"]	["Bundesliga_ALE; Liga NOS_POR; Brasileirao_BRA; Eredivisie_HOL"]	["Premier League_ING; Europa League; Ekstraklasa_POL; Champions League"]	["La Liga_ESP; Superliga_SUI; Súper Lig_TUR; Bundesliga_AUS"]	["Ligue 1_FRA; 1.Liga_CHE; Conference League"]
Septiembre	["Bundesliga_ALE"; Jupiler Pro League_BEL; Liga HNL_CRO; 1.Liga_CHE"]	["La Liga_ESP; Bundesliga_AUS; Liga NOS_POR; Súper Lig_TUR"]	["Premier League_ING; Champions League; Superliga_SUI; Primera División_ARG"]	["Ligue 1_FRA; Europa League; Brasileirao_BRA; Ekstraklasa_POL"]	["Serie A_ITA; Eredivisie_HOL; Conference League"]
Octubre	["Serie A_ITA; Jupiler Pro League_BEL; Brasileirao_BRA; Bundesliga_AUS"]	["Premier League_ING; Superliga_SUI; Súper Lig_TUR; Ekstraklasa_POL"]	["Ligue 1_FRA; Liga NOS_POR; Eredivisie_HOL; Champions League"]	["Bundesliga_ALE; Conference League; Europa League; Liga HNL_CRO"]	["La Liga_ESP; 1.Liga_CHE; Primera División_ARG"]

Figura 31: Ejemplo de reparto de ligas para scouts (Pestaña 2: Asignación de ligas)

	Analista 1	Analista 2	Analista 3
Enero	["MLS_USA; Jupiler Pro League_BEL; Champions League; Liga HNL_CRO; Serie A_ITA; Brasileirao_BRA; Primera División_ARG; Conference League"]	["Eredivisie_HOL; Ligue 1_FRA; Bundesliga_ALE; Liga NOS_POR; Eliteserien_NOR; Europa League; Liga MX_MEX; Superliga_SUI"]	["1.Liga_CHE; Superliga_DIN; La Liga_ESP; Bundesliga_AUS; Súper Lig_TUR; Ekstraklasa_POL; Allsvenskan_SUE; Premier League_ING"]
Febrero	["Conference League; La Liga_ESP; Súper Lig_TUR; Europa League; Brasileirao_BRA; Premier League_ING; Bundesliga_AUS; Liga HNL_CRO"]	["Ligue 1_FRA; Liga MX_MEX; Superliga_SUI; Bundesliga_ALE; Liga NOS_USA; Liga NOS_POR"]	["Primera División_ARG; 1.Liga_CHE; Champions League; Allsvenskan_SUE; Ekstraklasa_POL; Jupiler Pro League_BEL; Serie A_ITA; Eliteserien_NOR"]
Marzo	["Superliga_DIN; La Liga_ESP; Bundesliga_ALE; 1.Liga_CHE; Liga NOS_POR; Eredivisie_HOL; Superliga_SUI; Brasileirao_BRA"]	["Bundesliga_AUS; Allsvenskan_SUE; Jupiler Pro League_BEL; Ligue 1_FRA; MLS_USA; Serie A_ITA; Liga MX_MEX; Champions League"]	["Ekstraklasa_POL; Eliteserien_NOR; Europa League; Súper Lig_TUR; Liga HNL_CRO; Primera División_ARG; Conference League; Premier League_ING"]
Abril	["Liga NOS_POR; La Liga_ESP; 1.Liga_CHE; Bundesliga_AUS; Súper Lig_TUR; Serie A_ITA; Premier League_ING"]	["Primera División_ARG; Eredivisie_HOL; Jupiler Pro League_BEL; Superliga_DIN; Ligue 1_FRA; Europa League; Liga HNL_CRO; MLS_USA"]	["Brasileirao_BRA; Liga MX_MEX; Allsvenskan_SUE; Eliteserien_NOR; Champions League; Ekstraklasa_POL; Superliga_SUI; Conference League"]
Agosto	["Liga HNL_CRO; Bundesliga_AUS; Conference League; Champions League; Bundesliga_ALE; MLS_USA; Liga NOS_POR; Jupiler Pro League_BEL"]	["Superliga_SUI; Europa League; Premier League_ING; Superliga_DIN; 1.Liga_CHE; Brasileirao_BRA; Eliteserien_NOR; Primera División_ARG"]	["Ligue 1_FRA; Ekstraklasa_POL; Allsvenskan_SUE; La Liga_ESP; Eredivisie_HOL; Liga MX_MEX; Serie A_ITA; Súper Lig_TUR"]
Septiembre	["Liga MX_MEX; Jupiler Pro League_BEL; Bundesliga_AUS; Liga NOS_POR; Liga HNL_CRO; Conference League; Eredivisie_HOL; 1.Liga_CHE"]	["Superliga_DIN; Allsvenskan_SUE; Eliteserien_NOR; Superliga_SUI; Primera División_ARG; Champions League; Serie A_ITA; Premier League_ING"]	["Ekstraklasa_POL; Ligue 1_FRA; Europa League; Bundesliga_ALE; MLS_USA; Brasileirao_BRA; La Liga_ESP; Súper Lig_TUR"]
Octubre	["Súper Lig_TUR; Superliga_DIN; Superliga_SUI; Premier League_ING; Ligue 1_FRA; Brasileirao_BRA; Eredivisie_HOL; Bundesliga_ALE"]	["Jupiler Pro League_BEL; Liga NOS_POR; Serie A_ITA; Europa League; Liga MX_MEX; Primera División_ARG; Eliteserien_NOR; MLS_USA"]	["Ekstraklasa_POL; 1.Liga_CHE; Allsvenskan_SUE; Bundesliga_AUS; Champions League; Conference League; Liga HNL_CRO; La Liga_ESP"]

Figura 32: Ejemplo de reparto de ligas para analistas de datos (Pestaña 2: Asignación de ligas)

Si en el selector de rol se escoge una opción distinta a la de Director Deportivo (la opción scout o analista de datos), se podrá visualizar el calendario que el Director Deportivo ha generado en su área de trabajo en esta pestaña, para que cada miembro de la Dirección Deportiva pueda consultar su asignación de ligas concreta.

4.3.3. Pestaña 3: Evaluación de los scouts

Esta pestaña se ha diseñado especialmente para el uso de los scouts. Su principal objetivo es la generación y consulta de informes de futbolistas, calificando el rendimiento de cada uno de ellos en diferentes áreas de interés relativas a temas deportivos y extradeportivos. La generación de informes se realizará durante la Fase I (Mapeo y Cribado), ya que es el período en el cual se nutre la base de datos de la Dirección Deportiva con el rendimiento de los futbolistas durante la temporada.

Además, aunque no generen informes de rendimiento, ya que esa cuestión está reservada a los scouts, tanto el Director Deportivo como los analistas de datos pueden consultar la información. Estos últimos, además, la deben consultar para incluirla en el informe final del jugador (cuya estructuración se concretará posteriormente).

En la Figura 32, se visualiza la información contenida en la pestaña. A la izquierda, en el sidebar, se encuentran los diferentes perfiles a considerar. Cada uno de los perfiles tienen asociados unos conceptos que ahora se concretarán. En el centro de la pestaña, se puede observar una breve explicación del funcionamiento de la pestaña y, más abajo, un selector en el que se debe determinar si se busca generar un nuevo informe o, si por el contrario, se busca consultar un informe ya realizado.



Figura 33: Visión general (Pestaña 3: Evaluación Scout)

En la Figura 34, se puede observar la información que resulta de pulsar la opción 'Generar Nuevo Informe'. Antes de empezar a evaluar al futbolista en sí, la única información requerida es el nombre del futbolista del cual se va a realizar el informe y quién es el que lo realiza. En este caso, como en la Dirección Deportiva diseñada consta de 5 scouts, hay 5 opciones que representan cada uno de los scouts.

Selecciona una opción:

- Generar Nuevo Informe
- Consultar Informes Realizados

Nombre del Futbolista:

Selecciona el scout que va a realizar este informe:

Scout 1 ▼

- Scout 1
- Scout 2
- Scout 3
- Scout 4
- Scout 5

Figura 34: Opción 'Generar Nuevo Informe' (Pestaña 3: Evaluación Scout)

Una vez definida esta información, directamente se pasa a calificar al futbolista. Como se ha comentado previamente, se han definido una serie de perfiles para caracterizar al futbolista, deportiva y extradeportivamente, a partir de la información obtenida en (Centro de Estudio y Desarrollo del Fútbol (CEyDF), s.f.), añadiendo además algunos conceptos que no figuraban en esta referencia que se consideraban de interés. Estos perfiles, y sus correspondientes características particulares son:

- 1) **Perfil físico:** Evalúa las capacidades físicas del futbolista, cada vez más importantes en el mundo del fútbol, ya que suponen una ventaja competitiva. Entre ellas, se han definido las siguientes características:
 - **Velocidad de reacción:** Es la capacidad de un jugador para identificar rápidamente una situación en el campo de juego y responder de manera instantánea y efectiva. Esto puede incluir reacciones a pases, cambios de dirección del balón o movimientos de los oponentes.
 - **Explosividad:** Se refiere a la capacidad de un jugador para realizar movimientos de alta

intensidad en un corto período de tiempo. En el fútbol, esto se ve en la capacidad de un jugador para realizar sprints rápidos, saltos explosivos o cambios de dirección bruscos.

- **Agilidad:** Se relaciona con la capacidad de un jugador para cambiar de dirección rápidamente y de manera efectiva, en posesión del balón o sin él. La agilidad es crucial para esquivar a los oponentes y superar obstáculos en el campo.
- **Repetición de esfuerzos:** Este aspecto se refiere a la capacidad de un jugador para realizar esfuerzos intensos y repetitivos a lo largo del partido. Incluye la resistencia cardiovascular y muscular para mantener un alto nivel de rendimiento durante los 90 minutos de juego, lo cual es de relativa importancia en la gestión de las transiciones ofensiva y defensiva.
- **Sprint:** Se basa en analizar el esfuerzo máximo de velocidad de un futbolista en distancias cortas.
- **Fuerza (duelos):** La fuerza es esencial en situaciones de duelos físicos, como enfrentamientos cuerpo a cuerpo con oponentes para disputar el balón o mantener la posesión. La fuerza contribuye a la capacidad de un jugador para ganar esos enfrentamientos y mantener el control.

2) **Perfil técnico:** Mide la capacidad, habilidad y creatividad del jugador, centrándose en las acciones con balón. El entrenamiento de tecnificación, para incidir en este aspecto, es de vital importancia no solo en el fútbol de élite, sino en cualquier categoría o nivel, ya que es el parámetro que diferencia a un jugador de élite respecto a un jugador de una categoría inferior. En esta cuestión, se han evaluado las siguientes características:

- **Control de oposición:** Se refiere a la habilidad de un jugador para mantener el control del balón mientras enfrenta la presión o la marca de un oponente. Esto implica usar el cuerpo y las piernas de manera efectiva para proteger la posesión de balón.
- **Acción técnica pie-cuerpo:** Este aspecto se refiere a la coordinación entre los movimientos del cuerpo y el uso adecuado de los pies para controlar, pasar, hacer un regate o realizar un disparo.
- **Visión de juego:** Se relaciona con la capacidad de un jugador para leer y comprender la dinámica del campo, identificar oportunidades y tomar decisiones inteligentes en el juego. A menudo se relaciona también con la capacidad de un jugador para realizar pases de forma creativa y que generan una situación de peligro para el rival.
- **Habilidad:** La habilidad se relaciona con la destreza técnica de un jugador, incluyendo su capacidad para realizar regates y movimientos de complicación técnica con el balón. Los jugadores habilidosos pueden eludir a los oponentes de cara a generar oportunidades de peligro.
- **Precisión de pase:** La precisión de pase se refiere a la capacidad de un jugador para enviar el balón de manera precisa y efectiva a un compañero de equipo. Incluye pases de media, pases

largos, pases en profundidad y pases filtrados entre líneas.

- **Precisión de tiro:** Este aspecto se relaciona con la capacidad de un jugador para disparar a portería con precisión. Incluye la capacidad de colocar tiros en lugares específicos de la portería y la habilidad para marcar goles.
 - **Toma de decisiones:** Implica la capacidad de un jugador para elegir la acción más adecuada en una situación dada. Habitualmente, las acciones se deciden en décimas de segundo, por lo que tener habilidad para gestionar la toma de decisiones es un aspecto vital que se trabaja particularmente en el mundo del fútbol. Esto puede incluir decidir si pasar, driblar, disparar o mantener la posesión del balón en función de las circunstancias del juego.
 - **Perfil:** Se refiere a la capacidad de un jugador para adoptar la posición adecuada y orientar el cuerpo correctamente en relación con el balón, los compañeros de equipo y los oponentes. Perfilarse adecuadamente es esencial para recibir, pasar e incluso disparar con eficacia, ya que permite al jugador estar en la posición óptima en el momento preciso.
- 3) **Perfil táctico:** Engloba las características relativas a la capacidad de un jugador para entender y ejecutar las estrategias y los roles específicos dentro de un sistema de juego determinado en lo que respecta a colocación, manejo de variantes defensivas y ofensivas, entre otras cuestiones, que se proceden a desglosar a continuación:
- **Colocación:** Se refiere a la capacidad de un jugador para posicionarse estratégicamente en el campo, tanto en defensa como en ataque. Esto implica estar en el lugar adecuado en el momento adecuado para influir y potenciar el juego del equipo, de cara a proporcionar soluciones efectivas.
 - **Robo de balón:** Implica la capacidad de un jugador para quitar al oponente el balón de manera legal y efectiva. Esto a menudo requiere una combinación de anticipación, colocación y habilidad defensiva.
 - **Anticipación:** La anticipación es fundamental en el fútbol y se refiere a la habilidad de un jugador para prever los movimientos del balón, los compañeros de equipo y los oponentes. Un jugador que anticipa bien puede tomar decisiones tácticas más acertadas.
 - **Capacidad defensiva:** La capacidad defensiva abarca una serie de habilidades tácticas utilizadas para evitar que el equipo contrario avance o anote. Esto incluye la marca, el bloqueo de pases, la interceptación de balones y la presión sobre el rival.
 - **Gestión de transiciones:** La gestión de transiciones se refiere a la capacidad de un jugador de colocarse correctamente en una acción defensiva u ofensiva. A continuación, se define qué son las transiciones defensivas y ofensivas:
 - **Transición defensiva:** Según el portal de fútbol (The Coaches' Voice, s.f.), una

transición defensiva es el momento del juego en el que el equipo pierde la posesión de balón y pasa a defender, organizándose colectivamente en función del modelo de juego, para volver a recuperarlo. A no ser que se produzca un robo, una transición defensiva puede enforzarse de dos formas distintas:

- **Repliegue:** Es la agrupación de los futbolistas en el terreno de juego para acortar las distancias entre líneas, impidiendo al equipo rival progresar con el balón de forma sencilla. En función de la filosofía del entrenador, el repliegue se puede producir en bloque bajo, medio o alto, dependiendo de la altura de la línea defensiva del equipo.
 - **Presión tras pérdida:** En vez de agruparse, los jugadores responden a la pérdida del balón intentando recuperarlo inmediatamente, por medio de la presión adelantada al jugador que posee el balón, además de a los jugadores que, potencialmente, pueden ser receptores de él. Esta técnica de transición defensiva es arriesgada ya que las distancias entre líneas son mayores y puede ocurrir que el equipo rival supere la presión y consiga progresar más rápidamente.
- **Transiciones ofensivas:** De nuevo, el portal de fútbol (The Coaches' Voice, s.f.) hace referencia a este aspecto del juego. Las transiciones ofensivas ocurren cuando el rival pierde la posesión de balón, lo que propicia un ataque propio. Estas acciones de juego buscan la profundidad y amplitud del campo para encontrar espacios en el tejido defensivo rival. Las transiciones ofensivas son de vital importancia para los equipos cuyo modelo de juego se basa en verticalizar el juego y dar pocos pases para acabar la jugada en portería rival.
- 4) **Perfil psicológico:** Engloba las características, habilidades y rasgos mentales que influyen en su comportamiento y rendimiento en el terreno de juego durante un partido. Estos aspectos psicológicos son cruciales para el éxito en el fútbol y pueden tener un impacto significativo en el rendimiento de un jugador. Para evaluar este perfil, es recomendable visualizar partidos de diferentes naturalezas: contra un rival débil, contra un rival fuerte, rivales con una grada intimidatoria, jugando con un futbolista menos, entre otras situaciones. Las características a evaluar relativas a este perfil serían:
- **Agresividad:** Se refiere a la dureza con la que juega un futbolista y la disposición a competir de manera intensa y determinada. No se trata necesariamente de comportamiento violento, sino más bien de la voluntad de luchar por el balón y buscar oportunidades para ganar. En algunas ocasiones, es necesario incidir en esta cuestión ya que, una mala gestión de la agresividad puede conducir a una sanción (tarjeta amarilla o roja), lo que perjudicaría al equipo.
 - **Personalidad:** La personalidad de un jugador puede influir en su estilo de juego y su

comportamiento en el campo. Algunos jugadores pueden ser extrovertidos y expresivos, mientras que otros pueden ser más introvertidos y centrados. La personalidad puede afectar la interacción con compañeros de equipo, entrenadores y adversarios.

- **Actitud:** La actitud se relaciona con la mentalidad y el enfoque de un jugador hacia el fútbol. Una actitud positiva y orientada hacia el trabajo puede contribuir al éxito y ayudar a sus compañeros para, en conjunto, poder sobreponerse a situaciones complicadas, mientras que una actitud negativa o complaciente puede ser perjudicial y puede afectar al rendimiento individual y colectivo.
 - **Gestión de las emociones:** La gestión de las emociones implica la capacidad de un jugador para controlar y canalizar sus emociones de manera efectiva. Esto es importante para evitar que las emociones negativas, como la frustración o la ira, afecten el rendimiento y para mantener la calma bajo presión. Está íntimamente relacionado con el resto de las características, debido a que una mala gestión de las emociones puede desencadenar problemas de agresividad, actitud, entre otros.
 - **Sentido de responsabilidad:** Un jugador con un fuerte sentido de responsabilidad toma el papel protagonista del equipo y, en situaciones adversas, es capaz de ser el principal sustento del equipo.
- 5) **Adecuación al club:** Evalúa las posibilidades de integración del jugador tanto en cuestiones deportivas como extradeportivas. Muchos fichajes, a lo largo de la historia del fútbol, no han tenido buen rendimiento en el nuevo club debido a que no ha habido una adaptación adecuada a la idiosincrasia del club o de la ciudad a la que iban, por lo tanto, conocer (dentro de lo posible) la personalidad del jugador dentro y fuera del campo, a partir de su entorno, es clave para poder determinar si la adaptación al club puede llegar a ser un problema en el futuro. En este respecto, las características evaluadas serían:
- **Modelo de juego del equipo:** Implica evaluar la capacidad del jugador para comprender y ejecutar las tácticas y las funciones específicas requeridas por el entrenador y el sistema de juego del equipo.
 - **Ciudad:** Influencia que puede tener en la comodidad del jugador su adaptación a la cultura local y su disposición para vivir y jugar en la ciudad del club que ficha.
 - **Rol:** Se refiere a la adecuación del jugador a la posición y las responsabilidades específicas que se le asignan en el sistema de juego. La adecuación del jugador a su rol implica su capacidad para desempeñar las funciones y tareas requeridas de manera efectiva y en línea con la estrategia del equipo.

Una vez definidos todos los perfiles y todas sus correspondientes características, se añade la Figura 35 en la cual se puede ver que, para evaluar todos estos aspectos, se han añadido lo que se conoce en la librería

Streamlit como 'sliders', que son barras deslizantes que permiten establecer un valor concreto a una variable. En este caso concreto, se añade en la captura únicamente el perfil físico, pero se haría de igual forma para todos los perfiles.

Calificaciones

Otorgue una nota desde el 1 (muy malo) hasta el 5 (muy bueno) en las siguientes cuestiones:

Perfil Físico

Velocidad de reacción



Explosividad



Agilidad



Repetición de esfuerzos



Sprint



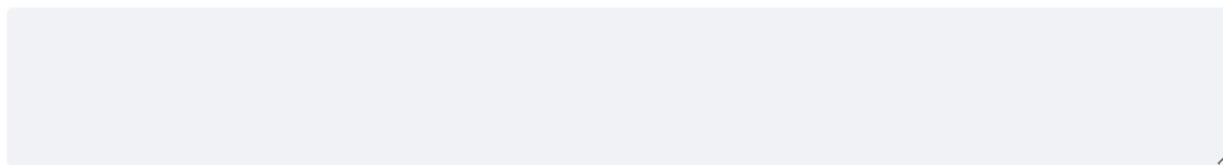
Fuerza (duelos)



Figura 35: Ejemplo de sliders 'Perfil Físico' (Pestaña 3: Evaluación Scout)

Además, se añade un recuadro en el cual el scout puede añadir cualquier puntualización a las calificaciones realizadas o para comentar alguna cuestión o apreciación de interés para la realización del informe final del futbolista, como se puede observar en la Figura 36. Una vez terminado el informe, se presiona el botón 'Almacenar Informe' que guardará la información aportada por el scout en una base de datos programada en SQL.

Observaciones Generales



Almacenar Informe

Figura 36: Cuadro de texto para observaciones generales (Pestaña 3: Evaluación Scout)

Por otro lado, como se ha comentado previamente, en el primer selector de la pestaña se puede seleccionar la opción de consultar informes ya realizados. Para esta opción y como se observa en la Figura 37, aparecen dos selectores, en los cuales se deben elegir el jugador y el scout correspondiente del cual se quiere extraer el análisis. En este caso, se busca extraer la información recogida por el scout 5 del jugador de C.A. Osasuna, Jon Moncayola.

Consultar Informes Realizados

Selecciona el jugador:

Jon Moncayola

Selecciona el scout:

Scout 5

Figura 37: Selectores de consulta de informes (Pestaña 3: Evaluación Scout)

Si se selecciona una combinación de jugador-scout de la cual no se puede rescatar ningún informe en la base de datos de SQL, aparece un mensaje de error que indica al usuario que no existe informe alguno para ese jugador realizado por dicho scout, como se observa en la Figura 38.

Consultar Informes Realizados

Selecciona el jugador:

Jon Moncayola

Selecciona el scout:

Scout 4

No hay informes realizados por Scout 4 para el jugador Jon Moncayola.

Figura 38: Error de informe (Pestaña 3: Evaluación Scout)

En caso contrario, si la combinación existe, se mostrará por pantalla el informe realizado por el scout. En la Figura 39, se muestra una captura del informe realizado por el scout 5 de Jon Moncayola.

Informe de Jon Moncayola

Realizado por: Scout 5

Fecha: 14-09-23

Perfil Físico

- Velocidad de reacción: 4
- Explosividad: 4
- Agilidad: 5
- Repetición de esfuerzos: 4
- Sprint: 5
- Fuerza (duelos): 3

Perfil Técnico

- Control de oposición: 4
- Acción técnica pie-cuerpo: 5
- Visión de juego: 4
- Habilidad: 5
- Precisión de pase: 4
- Precisión de tiro: 4
- Toma de decisiones: 5
- Perfil: 5

Perfil Táctico

- Colocación: 4
- Robo de balón: 5
- Anticipación: 4
- Capacidad defensiva: 5
- Gestión transiciones: 5

Perfil Psicológico

- Agresividad: 3
- Personalidad: 4
- Actitud: 5
- Gestión de las emociones: 5
- Sentido de responsabilidad: 4

Adecuación al club

- Modelo de juego del equipo: 4
- Ciudad: 4
- Rol: 5

Observaciones Generales

Jon Moncayola es un futbolista que se ha destacado por su versatilidad táctica y su capacidad para desempeñar múltiples roles en el centro del campo. Su juego se caracteriza por su inteligencia táctica y su capacidad para leer el juego con precisión, lo que le permite anticipar las acciones del equipo contrario y realizar transiciones efectivas entre las fases defensiva y ofensiva.

Desde una perspectiva táctica, Moncayola exhibe una capacidad excepcional para ocupar espacios clave en el centro del campo y funcionar como un "mediocampista box-to-box". Es hábil en la distribución del balón y es capaz de realizar pases precisos tanto en corto como en largo alcance. Además, su capacidad de toma de decisiones y su visión periférica le permiten actuar como un "mediocampista organizador", dirigiendo el juego de su equipo desde el mediocampo.

En defensa, Moncayola destaca por su capacidad de presión efectiva y su capacidad para recuperar balones en zonas cruciales del campo. Su agresividad controlada y su excelente lectura del juego le permiten interceptar pases y contribuir a la ruptura de las acciones del equipo contrario.

Su estilo de juego se caracteriza por una técnica pulida y un muy buen control del balón, lo que le permite mantener la posesión y construir ataques desde el mediocampo. Además, es un jugador que no teme asumir la responsabilidad de tomar tiros a distancia cuando la oportunidad se presenta, destacando su precisión en los intentos de tiro a portería.

En resumen, Jon Moncayola se destaca como un jugador de fútbol integral con una mentalidad táctica excepcional. Su capacidad para adaptarse a múltiples roles en el centro del campo, su visión de juego y su capacidad para mantener la calma bajo presión hacen de él un activo valioso para su equipo, contribuyendo tanto en la fase defensiva como en la ofensiva. Su versatilidad y su inteligencia táctica son atributos que lo convierten en un futbolista sumamente destacado.

Figura 39: Informe de Jon Moncayola por Scout 5 (Pestaña 3: Evaluación Scout)

4.3.4. Pestaña 4: Rating

La funcionalidad principal de esta pestaña es hacer el cálculo de una calificación para todos los futbolistas (en adelante, referido como ‘rating’ por su denominación en inglés) para determinar, en función de la muestra escogida y los parámetros elegidos, qué jugador es el que mejores métricas presenta.

En la Figura 40 se puede observar, de igual manera que en las demás pestañas, la información general de la pestaña. Por una parte, en el sidebar, se encuentra una breve explicación de en qué consiste la pestaña, además de información sobre las funcionalidades de esta.

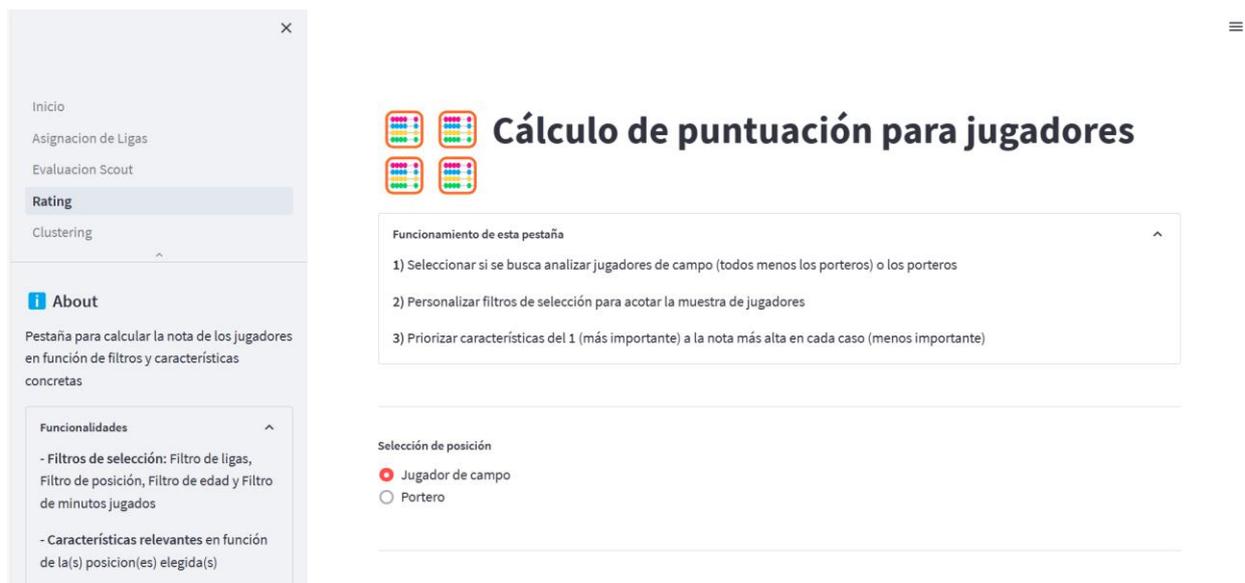


Figura 40: Visión general (Pestaña 4: Rating)

Además, en el cuerpo de la pestaña se encuentra un expander donde se explica, resumidamente, el funcionamiento de la pestaña y, debajo, un selector que pregunta si el jugador o los jugadores que se busca evaluar son jugadores de campo o porteros. Independientemente de la posición elegida, el proceso es análogo en ambos casos, variando únicamente las métricas consideradas, cuestión que se comentará después.

Seleccionada la demarcación que se pretende analizar, si se desliza hacia abajo en la pestaña, hay dos expanders que se procederá a comentar a continuación, como se puede observar en la Figura 41.

Configuración

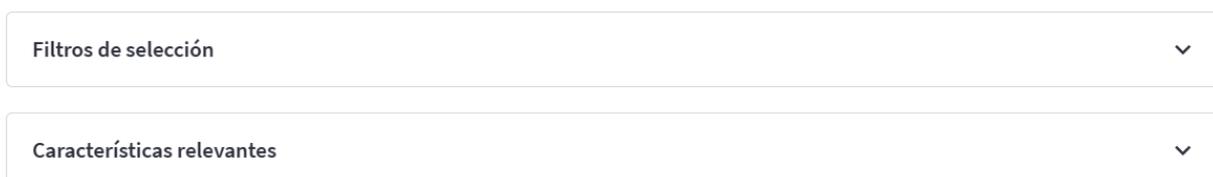


Figura 41: Expanders de configuración (Pestaña 4: Rating)

En el primer expander tiene como finalidad realizar un filtro de selección de jugadores para acotar la muestra. Los filtros de selección son los siguientes:

En la Figura 42 se encuentra el filtro de ligas. Este filtro consta de un ‘multiselect’, que es una funcionalidad de Streamlit para poder seleccionar varias opciones entre las disponibles. Básicamente, la función de este filtro es limitar las ligas que se van a tener en cuenta a la hora de hacer el cálculo de rating. Como se ha comentado previamente en el trabajo, las ligas tenidas en cuenta son 8.

1. Filtro de Ligas

Elige las competiciones a considerar:

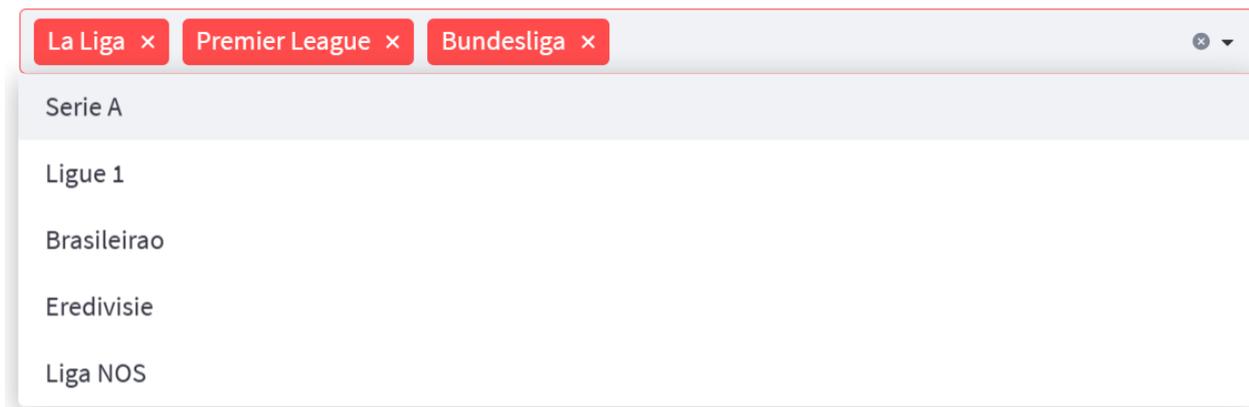


Figura 42: Filtro de ligas (Pestaña 4: Rating)

El siguiente filtro, que aparece en la Figura 43, es el filtro de posición. Este filtro busca limitar la cantidad de jugadores tenidos en cuenta en la muestra en función de la demarcación en la que desarrollan su juego. Como apreciación inicial, cabe comentar que FBref define una o dos posiciones para cada jugador en función de si este juega en una o varias demarcaciones.

Es por ello por lo que, primero, se debe definir si solo se quiere trabajar con la primera posición (preferida) del jugador. Para ello, si solo se quiere considerar la primera posición del jugador, se debe marcar la casilla habilitada para esa función. Si se marca, automáticamente desaparece la opción de definir la segunda posición preferida del jugador.

Si ese botón no se selecciona, se debe seleccionar una segunda posición preferida y solo aparecerán jugadores que cumplan ambas condiciones simultáneamente.

Se propone un ejemplo para terminar de aclarar la explicación: Jon Moncayola, jugador de C.A. Osasuna y cuyo informe se ha realizado como ejemplo en la pestaña anterior, figura en FBref como centrocampista (primera posición) y como defensa (segunda posición). Pueden darse las siguientes situaciones:

- Se marca la opción “Pulsar botón si solo se debe considerar la posición principal del jugador”. En este caso, para que Jon Moncayola figure en la tabla final de jugadores, se debe marcar en primera posición la opción: ‘CC’.

- No se marca la opción “Pulsar botón si solo se debe considerar la posición principal del jugador”. Para que Jon Moncayola aparezca en la tabla final, se marca como primera posición preferida la opción ‘CC’ y como segunda opción ‘DF’. Si, por ejemplo, se marca de forma inversa (primera opción ‘DF’ y segunda opción ‘CC’), el jugador no aparecerá en la tabla final. De igual manera, en cualquier otro caso, el jugador tampoco aparecerá en dicha muestra.

2. Filtro de Posición

Pulsar botón si solo se debe considerar la posición principal del jugador

Selecciona la primera posición del jugador:

CC ▼

Selecciona la segunda posición del jugador:

DL ▼

Figura 43: Filtro de posición (Pestaña 4: Rating)

El tercer filtro diseñado para la aplicación es el filtro de edad. Para este filtro, se coge la mínima y máxima edad de los jugadores que restan en la muestra después de aplicar los dos filtros anteriores para definir los valores extremos del slider. Como se puede ver en la Figura 44, la funcionalidad de este filtro es acotar, tanto inferior como superiormente, mediante un slider, el valor de edad requerido para los jugadores que aparecen en la muestra.

3. Filtro de Edad



Figura 44: Filtro de edad (Pestaña 4: Rating)

El cuarto y último filtro de los filtros de selección confeccionados es el filtro de minutos jugados. Como se puede observar en la Figura 45, está diseñado de igual manera que el filtro de edad. Se trata de un slider cuyo mínimo valor es el valor más bajo de minutos jugados entre todos los jugadores que se encuentran en la muestra después de aplicar los filtros anteriores a este. Ídem para el valor máximo.

A diferencia que el filtro de edad, el filtro de minutos únicamente necesita fijar un valor mínimo de minutos jugados por el jugador para figurar en la tabla final.

4. Filtro de Minutos



Figura 45: Filtro de minutos (Pestaña 4: Rating)

Acotada la muestra a partir de los filtros de selección, los siguientes parámetros de personalización a definir son las características relevantes del jugador. Para definir las características relevantes, a diferencia de los aspectos relevantes definidos en la pestaña de evaluación de jugadores por parte de los scouts, se han definido específicamente para cada posición y de acuerdo con las métricas obtenidas de FBref.

Como las características relevantes mostradas en la aplicación son específicas para cada posición, las características a mostrar van a depender de qué posiciones haya definido el usuario previamente en el filtro de posición. Si el usuario selecciona una única posición en dicho filtro, solo se mostrarán las características relevantes para esa posición. Por el contrario, si el usuario selecciona dos posiciones, se mostrarán las características relevantes para esas dos posiciones.

Teniendo en cuenta que el cálculo del rating se va a realizar a partir de las métricas obtenidas de FBref, las características relevantes a definir en cada caso deben ser adecuadas a las métricas de las que se disponen. Antes de entender cómo funciona esta funcionalidad en la aplicación, se procede a definir las características relevantes específicas para cada posición y cuáles son las métricas, cuya descripción detallada se encuentra en el apartado 4. que deben ser tenidas en cuenta para cada una de dichas características:

- **Características relevantes para porteros:** Inciden en el rendimiento de las paradas y el juego de pies, que es relevante en equipos que buscan tener la posesión de balón y salidas de balón desde área propia con el balón controlado.
 - **Paradas:** Es la característica, por lo general, más importante para los porteros, ya que el nivel de las paradas y la capacidad del portero para encajar el menor número de goles posibles son las cuestiones más relevantes a la hora de adquirir un portero. Referente a esta característica, inciden las siguientes métricas:
 - Goles en contra ('GC').
 - Número de paradas ('Salvadas').
 - Porcentaje de paradas respecto al total de disparos a puerta recibidos ('% Salvadas').
 - Porterías a cero ('PaC').
 - Porcentaje de porterías a cero respecto al total de partidos disputados ('PaC%').

- Penaltis parados ('PD').
- Porcentaje de penaltis parados respecto al total de penaltis concedidos ('%Salvadas_Pen').
- Post-Shot Expected Goal ('PSxG'). Para consultar, de nuevo, el significado de esta métrica avanzada, ir al apartado 4.1.1.1.
- PSxG respecto al total de tiros a puerta recibidos ('PSxG/SoT').
- PSxG con el jugador sobre el campo respecto al PSxG sin él ('PSxG+/-').
- **Juego aéreo:** Es una habilidad esencial en un portero, que se relaciona con su capacidad para interceptar o despejar balones aéreos. Generalmente, si un portero domina esta faceta, hace que la línea defensiva mejore su rendimiento al generar confianza sobre la misma. Para evaluar esta característica o atributo, en FBref se encuentran:
 - Número de centros interceptados ('Stp').
 - Porcentaje de centros interceptados respecto al total de centros efectuados por el rival ('% de Stp').
 - Acciones defensivas fuera del área ('Núm. de OPA').
 - Distancia promedio de las acciones defensivas fuera del área respecto de la portería propia ('DistProm.').
- **Técnica con balón:** Evalúa la habilidad de un portero para distribuir el balón en distancias largas para generar ventajas y superar líneas de presión rival con precisión y control. En FBref, las métricas para evaluar esta característica son:
 - Pases completados de más de 35 metros ('Cmp').
 - Porcentaje de pases completados de más de 35 metros respecto al total de pases ('%Cmp').
 - Longitud promedio del pase ('Long. prom.').
- **Características relevantes para defensas:** Tienen en cuenta el rendimiento defensivo del jugador, la técnica con balón del jugador, así como, sobre todo para valorar a los laterales, el rendimiento ofensivo.
 - **Contundencia:** Se refiere a la habilidad del defensa para abordar duelos físicos con éxito para recuperar la posesión de balón. Un defensa contundente es capaz de ganar balones divididos, realizar entradas efectivas, evitando de esa manera que el equipo contrario progrese y pueda realizar ataques con facilidad. Entre las métricas tenidas en cuenta para este atributo, se encuentran:
 - Porcentaje de duelos ganados sobre el total de duelos ('% de ganados').

- Tackles (entradas) realizados ('Tkl').
 - Tackles realizados con éxito ('TklG').
 - Bloqueos realizados ('Bloqueos').
 - Tackles más intercepciones realizados ('Tkl+Int').
 - Recuperaciones de balón ('Recup.').
 - Faltas cometidas ('Fls').
- **Colocación y anticipación:** La colocación se relaciona con la capacidad del jugador para posicionarse estratégicamente en el campo, cerrando espacios, cortando líneas de pase y marcando a los adversarios de manera efectiva. La anticipación implica la capacidad de prever las jugadas del equipo contrario y actuar en consecuencia, interceptando pases o cortando ataques antes de que se desarrollen por completo. Las métricas para tener en cuenta en este atributo son:
 - Tackles realizados con éxito ('TklG').
 - Recuperaciones de balón ('Recup.').
 - Bloqueos realizados ('Bloqueos').
 - Número de robos de balón ante un intento regate rival ('Succ').
 - Penaltis concedidos, métrica inversa ('Penal concedido_inv').
- **Salida de balón:** Se refiere a la habilidad de un defensa para dar continuidad del juego con balón desde área propia. Un defensa con una buena salida de balón contribuye significativamente a la construcción del juego ofensivo de su equipo. Las métricas para tener en cuenta serían:
 - Número de acarreo o conducciones progresivas ('PrgC').
 - Número de pases progresivos efectuados ('PrgP').
 - Distancia progresiva recorrida por los considerados pases progresivos ('Dist. prg.').
 - Pases completados totales, cortos, medios y largos ('Cmp', 'Cmp_Cor', 'Cmp_Med', 'Cmp_Lar').
 - Porcentaje de pases completados respecto al total de pases intentados totales, cortos, medios y largos ('%Cmp', '%Cmp_Cor', '%Cmp_Med', '%Cmp_Lar').
 - Distancia progresiva recorrida en posesión de balón ('Dist. prg._circ').
 - Distancia progresiva recorrida en posesión de balón en el último tercio del campo ('1/3_circ').

- Errores de control de balón, métrica inversa ('Errores de control_inv').
- **Rendimiento ofensivo:** El rendimiento ofensivo de un defensa se relaciona con su capacidad para sumarse al ataque y crear oportunidades de gol, ya sea mediante la participación en jugadas a balón parado, centros al área desde la banda o tiros desde fuera del área. Un defensa con un buen rendimiento ofensivo agrega una dimensión adicional al juego de su equipo, aumentando la amenaza en el área rival. Por lo general, como se ha comentado previamente, esta métrica es de mayor importancia para los laterales, ya que son los defensas que más se prodigan en acciones ofensivas. Las métricas relevantes en este aspecto son:
 - Número de acarreos o conducciones progresivas ('PrgC').
 - Número de pases progresivos efectuados ('PrgP').
 - Goles anotados ('Gls').
 - Asistencias realizadas ('Ass').
 - Número de pases clave ('PC').
 - Número de pases al área rival ('PPA').
 - Goles esperados ('xG').
 - Asistencias esperadas ('xA').
 - Cambios de orientación realizados ('Camb.').
 - Regates intentados ('Att').
 - Regates conseguidos ('Succ').
- **Características relevantes para centrocampistas:** En función de la demarcación, el rol y el estilo de juego propio del club, las características a buscar en un centrocampista varían enormemente. Es por ello por lo que es importante definir todas las características, para que el usuario tenga la posibilidad de priorizar y personalizar el resultado en función de los requerimientos concretos.
 - **Contundencia:** Definida de igual manera que para los defensas. En este caso, esta característica es de vital importancia para los mediocentros defensivos, que deben destacar en esta cuestión.
 - **Colocación y anticipación:** Ídem que para la característica 'Contundencia'.
 - **Técnica con balón:** Incide sobre la precisión y la técnica del mediocentro respecto al toque de balón, la precisión y el control sobre el mismo. Referente a esta característica, se ha evaluado mediante estas métricas:
 - Pases completados medios y largos ('Cmp_Med', 'Cmp_Lar').

- Porcentaje de pases completados respecto al total de pases intentados medios y largos ('%Cmp_Med', '%Cmp_Lar').
 - Distancia progresiva recorrida por los considerados pases progresivos ('Dist. prg.').
 - Número de pases progresivos efectuados ('PrgP').
 - Número de acarreos o conducciones progresivas ('PrgC').
 - Distancia progresiva recorrida en posesión de balón ('Dist. prg._circ').
 - Distancia progresiva recorrida en posesión de balón ('Dist. prg._circ').
 - Distancia progresiva recorrida en posesión de balón en el último tercio del campo ('1/3_circ').
 - Errores de control de balón, métrica inversa ('Errores de control_inv').
- **Rendimiento ofensivo:** Definida de igual manera que para los defensas. En este caso, esta característica es de vital importancia para los mediocentros ofensivos, que deben destacar en esta cuestión.
 - **Capacidad goleadora:** Evalúa el nivel del jugador de cara a portería y de generar peligro en general, enfocado a la capacidad del jugador de marcar goles. Adquirir a un jugador que tenga mucha habilidad para marcar goles es complicado, ya que cuanto mayor sea la capacidad goleadora del jugador, más precio se va a pagar por él. Entre las métricas a destacar se encuentran:
 - Goles marcados ('Gls').
 - Goles esperados ('xG').
 - Disparos totales ('Dis').
 - Disparos a puerta ('DaP').
 - Penaltis ejecutados totales ('Penal ejecutado').
 - **Capacidad asistencial:** Refleja la habilidad de un jugador para dar pases de gol al compañero, ya sea con pases sencillos o con balones filtrados. Por lo general, dar el último pase en fútbol es una cuestión compleja debido al orden defensivo rival, por lo que esta cuestión es de vital importancia para el juego ofensivo del equipo. Entre las métricas evaluadas se encuentran:
 - Asistencias realizadas ('Ass').
 - Asistencias esperadas ('xA').
 - Pases clave ('PC').
 - **Generación de juego:** Hace hincapié en la capacidad del jugador para dar continuidad a la

jugada y hacerla progresar, ya sea teniendo en cuenta la recepción aérea de balones como acciones de pase que acerquen el balón a la portería rival. Las métricas a tener en cuenta son:

- Fuera de juego, métrica inversa ('PA_inv').
 - Porcentaje de duelos aéreos ganados ('% de ganados').
 - Distancia progresiva recorrida en posesión de balón en el último tercio del campo ('1/3_circ').
 - Número de pases progresivos efectuados ('PrgP').
 - Número de acarreos o conducciones progresivas ('PrgC').
 - Pases clave ('PC').
 - Asistencias realizadas ('Ass').
 - Asistencias esperadas ('xA').
 - Regates intentados ('Att').
 - Regates conseguidos ('Succ').
- **Características relevantes para delanteros:** Las características relevantes para un delantero en el fútbol varían según el estilo de juego del equipo y la posición específica del delantero, ya que puede haber delanteros centro, extremos o segundos delanteros, entre otros. Sin embargo, en general, las características clave que se tienen en cuenta para un delantero incluyen:
- **Capacidad goleadora:** Definida de igual manera que para los mediocentros. En este caso, esta característica es de vital importancia para los mediocentros ofensivos, que deben destacar en esta cuestión.
 - **Capacidad asistencial:** Definida de igual manera que para los mediocentros. En este caso, esta característica es de vital importancia para los mediocentros ofensivos, que deben destacar en esta cuestión.
 - **Técnica con balón:** Definida de igual manera que para los defensas. En este caso, esta característica es de vital importancia para los mediocentros puros (lo que se conoce como regista o distribuidor de balón), que deben destacar en esta cuestión.
 - **Generación de juego:** Definida de igual manera que para los defensas. En este caso, esta característica es de vital importancia para los mediocentros ofensivos, que deben destacar en esta cuestión.

Definidas todas las métricas, se añade una captura de cómo se muestran estas características, en función de la posición definida previamente. En este caso, como se ha definido que el jugador debe ser centrocampista (primera posición) y delantero (segunda posición), se deben combinar las características concretas de ambas

posiciones, como se puede ver en la Figura 46.

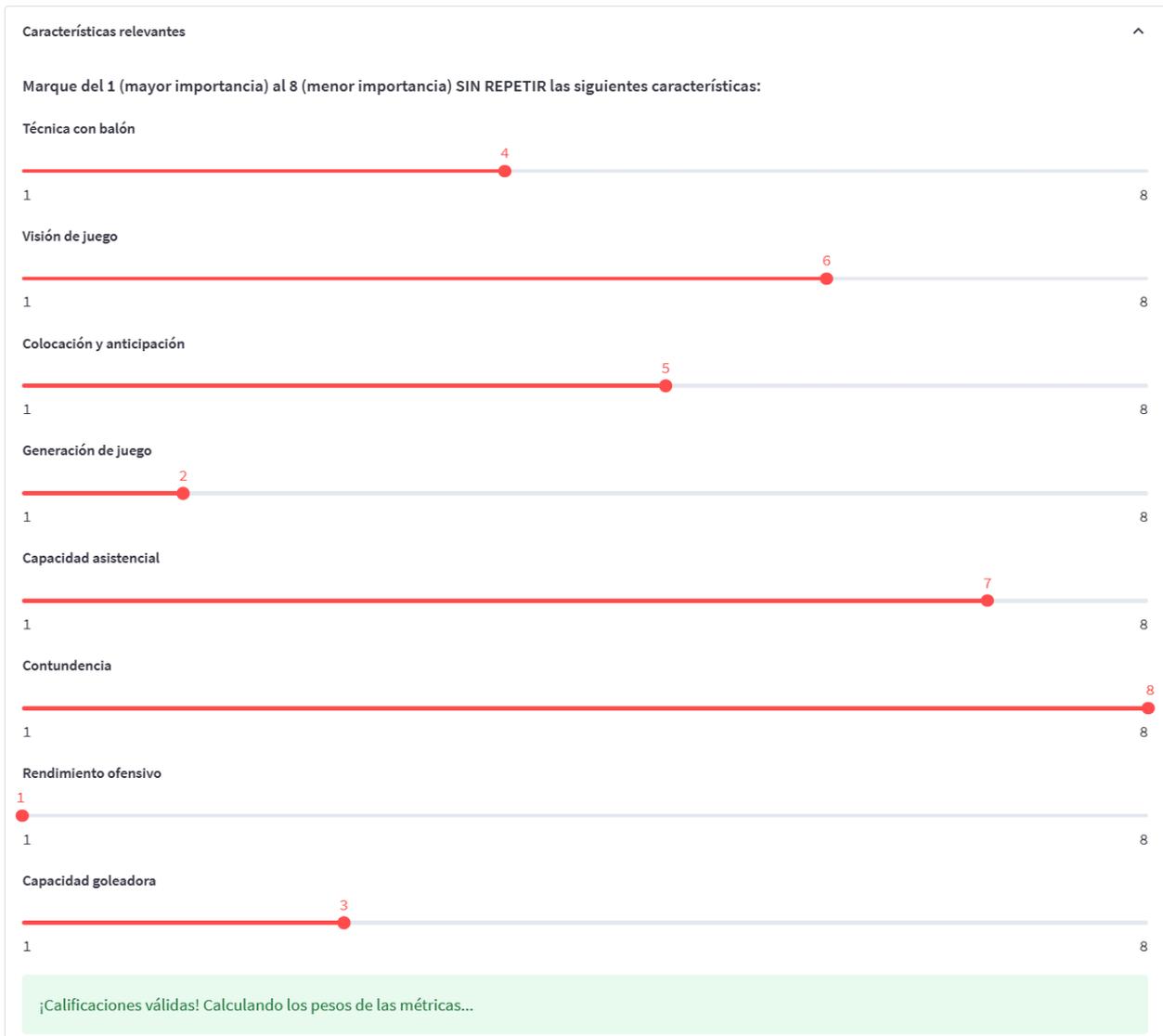


Figura 46: Visualización de características relevantes (Pestaña 4: Rating)

Como se ha comentado previamente, hay algunas características más importantes que otras. Por lo tanto, la mecánica elegida para esta parte de la aplicación es la de priorización. En la Figura 46, se ha dado una combinación correcta de sliders (todos con un valor único asignado) y, por consiguiente, al final del todo, se muestra un mensaje de confirmación de que las calificaciones son correctas y que se han calculado los pesos de las métricas.

El proceso de cálculo del peso de cada una de las métricas es el siguiente:

1. A partir del orden de prioridad definido por el usuario en la aplicación, se calcula el peso relativo de cada característica de la siguiente manera:

$$Peso_{característica} = \frac{\text{Número total de características} - \text{Calificación característica} + 1}{\text{Número total de características}}$$

2. Para cada métrica, se busca si pertenece a alguna de las características anteriores y, en caso positivo, se añade el peso correspondiente a cada característica a la métrica que corresponda. Como hay métricas que se repiten en varias características, el peso relativo de cada característica en la cual se encuentre dicha métrica se adicionará al total del peso relativo de la métrica, entonces:

$$Peso_{métrica} = \sum_{Características} Peso_{característica} \text{ si } Métrica \in Característica$$

Hay algunas métricas que están en la base de datos pero que no pertenecen a ninguna característica en concreto. A dichas métricas se les asigna el valor de peso más bajo obtenido de este proceso definido previamente.

Una vez asignados pesos a todas las métricas del conjunto de datos, se escala el valor de 0 a 1, para no desvirtuar el análisis ni la propia naturaleza de los datos, que previamente habían sido normalizados. Por último, se multiplican los pesos de cada métrica por los valores de dichas métricas para cada uno de los jugadores.

Haciendo esta multiplicación, se obtiene una nueva tabla de valores que tienen en cuenta la priorización del usuario. Para poder calcular, finalmente, el rating para cada jugador basta con adicionar todos los valores obtenidos en esta última tabla. Con el fin de que el rating obtenido sea entendible, se ha escalado la nota de 0 a 10, como se puede ver en la Figura 47:

	Jugador	Equipo	Competición	Rating
0	Lee Kang-in	Mallorca	La Liga	5.31
1	Federico Valverde	Real Madrid	La Liga	5.08
2	Rodrigo Riquelme	Girona	La Liga	4.47
3	Óscar Trejo	Rayo Vallecano	La Liga	4.33
4	Moi Gómez	Osasuna	La Liga	4.20
5	Gavi	Barcelona	La Liga	4.13
6	Alex Baena	Villarreal	La Liga	4.09
7	Carles Pérez	Celta Vigo	La Liga	4.00
8	Iker Muniain	Athletic Club	La Liga	3.87
9	Óliver Torres	Sevilla	La Liga	3.77
10	Aimar Oroz	Osasuna	La Liga	3.54
11	Fidel	Elche	La Liga	3.34
12	Giovani Lo Celso	Villarreal	La Liga	3.30
13	Theo Bongonda	Cádiz	La Liga	3.27
14	Rubén Sobrino	Cádiz	La Liga	3.22
15	Suso	Sevilla	La Liga	3.03
16	Carles Aleñá	Getafe	La Liga	2.90
17	Nico Ribaudó	Espanyol	La Liga	2.64
18	Iván Alejo	Cádiz	La Liga	2.47
19	Álex Fernández	Cádiz	La Liga	2.45

Figura 47: Tabla final de Ratings (Pestaña 4: Rating)

4.3.5. Pestaña 5: Clustering

En esta última pestaña de la aplicación se han introducido herramientas basadas en algoritmos de Machine Learning para la valoración de los futbolistas, concretamente, con algoritmos de clasificación. La finalidad, por tanto, es establecer parecidos entre jugadores con ayuda de los datos.

El clustering, según (Soria Polo, 2021), es un método de agrupamiento de los datos (en este caso, de los jugadores) en función de los atributos (en este caso, de las métricas de rendimiento de cada uno).

En la Figura 48 se muestra una visión general inicial de la pestaña una vez se abre la misma. En ella, se puede observar las funcionalidades añadidas a la pestaña y el funcionamiento de esta, que se comentará más pormenorizadamente en este epígrafe.

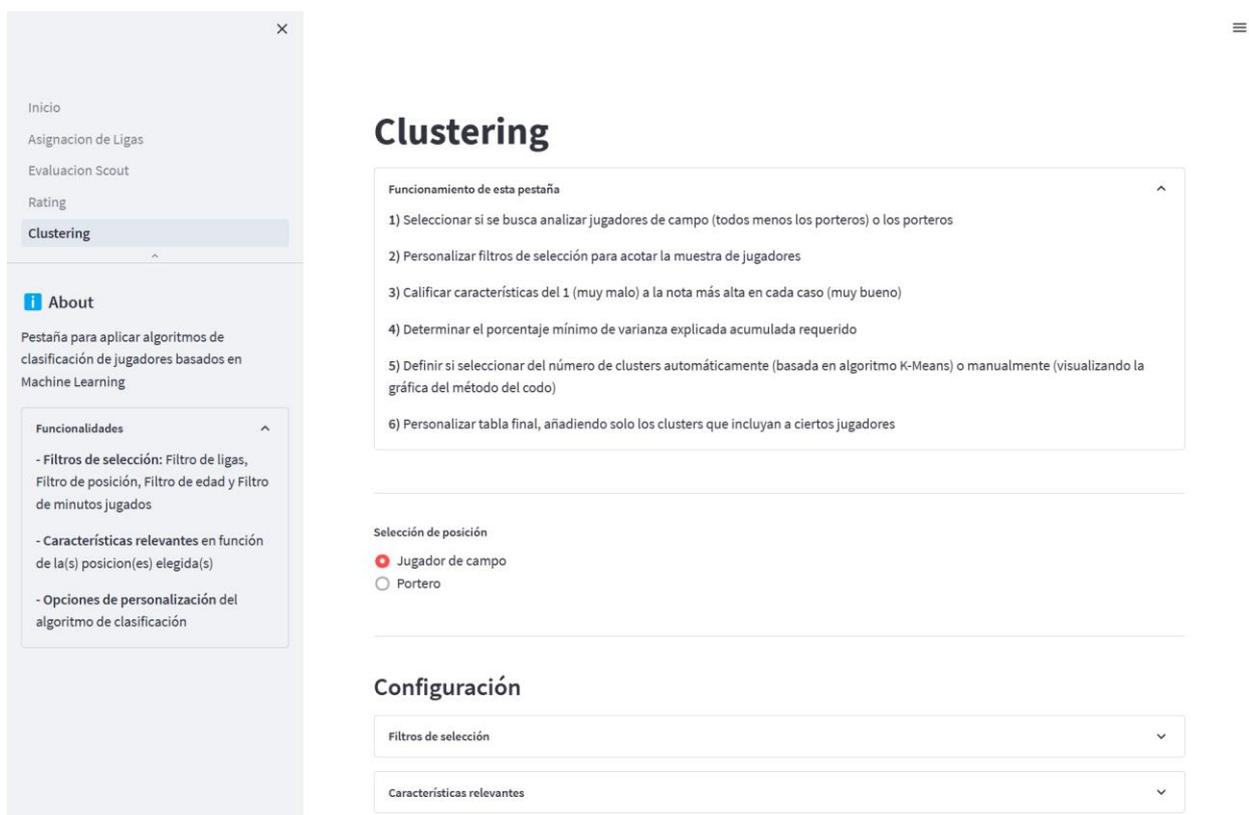


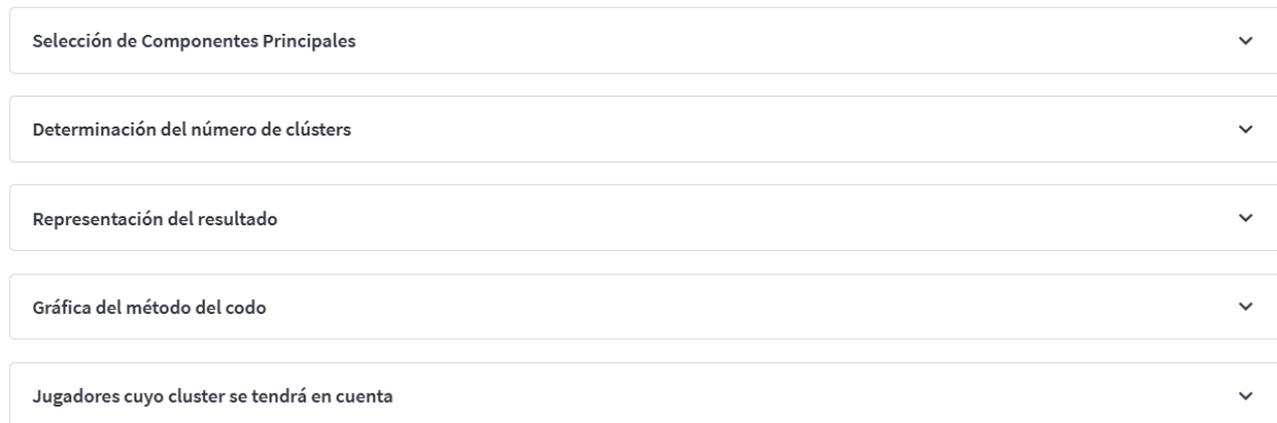
Figura 48: Visión general (Pestaña 5: Clustering)

La personalización disponible para el usuario de la base de datos es equivalente a la dispuesta en la pestaña anterior. Por un lado, se vuelve a añadir un selector en el que el usuario puede elegir la posición a analizar (siendo, de nuevo, análogo el funcionamiento de la pestaña en cualquiera de los dos casos). Por otro lado, también se han añadido los filtros de selección y la selección de características relevantes comentada en la pestaña anterior, con el fin de otorgar el mismo nivel de personalización al usuario que previamente.

Una vez acotada la muestra de jugadores respecto a la que se tenía inicialmente, el siguiente paso es añadir una tercera herramienta de personalización, en este caso, dedicada exclusivamente a la customización apoyada

en conceptos de Machine Learning. En la Figura 49, se pueden ver todas las opciones de personalización disponibles para el usuario, las cuales se explicarán una a una, tanto desde un punto de vista teórico, como su aplicación práctica en la pestaña.

Opciones de personalización basadas en Machine Learning



The image shows a vertical list of five dropdown menu items, each with a downward-pointing chevron icon on the right side. The items are: 'Selección de Componentes Principales', 'Determinación del número de clústers', 'Representación del resultado', 'Gráfica del método del codo', and 'Jugadores cuyo cluster se tendrá en cuenta'.

Figura 49: Opciones de personalización Machine Learning (Pestaña 5: Clustering)

La primera opción de personalización es relativa a la selección de componentes principales, basado en el Análisis de Componentes Principales.

- **Análisis de Componentes Principales (PCA):** Según el artículo (Karamizadeh, Abdullah, Manaf, Zamani, & Hooman, 2013), el Análisis de Componentes Principales (PCA, por sus siglas en inglés) es una técnica estadística fundamental utilizada en diversos campos, cuyo principal objetivo es la reducción de dimensionalidad de datos. Además, es una técnica que se utiliza para transformar un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de valores de variables linealmente no correlacionadas.

En otras palabras, el PCA es una herramienta matemática que nos permite encontrar nuevas dimensiones o "componentes principales" en los datos originales, donde estas nuevas dimensiones están diseñadas para ser ortogonales entre sí (no correlacionadas). Cada uno de estos componentes principales captura diferentes cantidades de variabilidad en los datos.

Al producirse una reducción de la dimensionalidad de los datos, el PCA permite visualizar datos de grandes dimensiones, como es el caso particular de este trabajo, en un espacio de dimensiones más reducidas, sin eliminar gran parte del significado de los datos.

La primera funcionalidad, como se puede ver en la Figura 50, sirve para elegir el número óptimo de componentes principales a partir del porcentaje mínimo de varianza acumulada explicada por las variables. La selección adecuada del número de componentes principales es un paso crítico en el proceso de PCA, y tiene implicaciones significativas para la representación y comprensión de los datos.

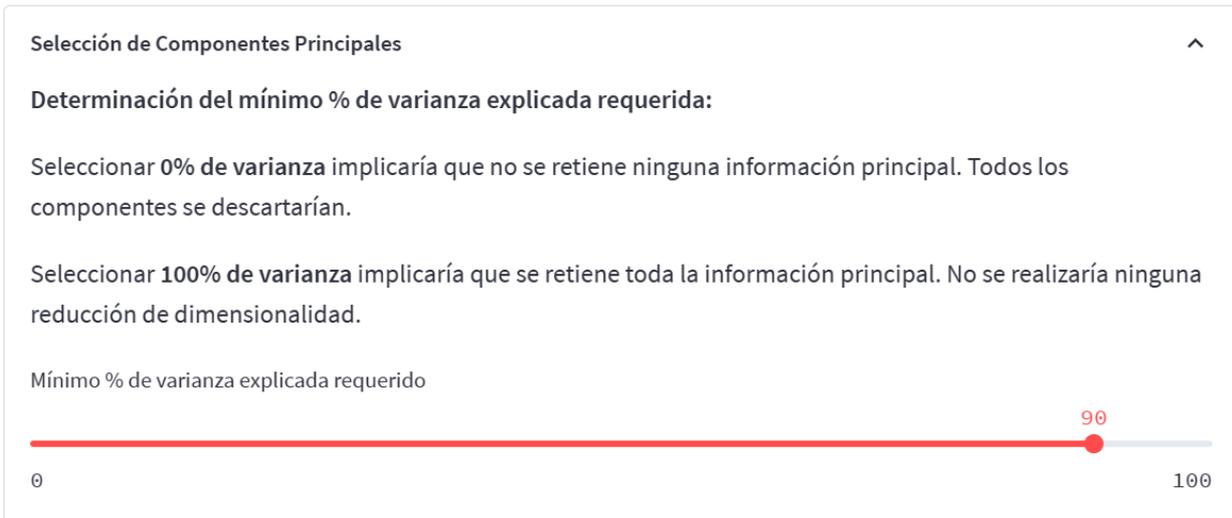


Figura 50: Personalización n°1 Machine Learning (Pestaña 5: Clustering)

El mínimo porcentaje de varianza explicada requerida es un parámetro esencial en la configuración de PCA. Representa el porcentaje mínimo de varianza en los datos que se debe conservar en las nuevas dimensiones (componentes principales). Este parámetro influye en la cantidad de información retenida y, por lo tanto, en la calidad de la representación de los datos.

En la aplicación, se ha diseñado un slider que permite al usuario ajustar este parámetro. El usuario puede seleccionar un valor entre 1 y 100, lo que refleja el porcentaje mínimo de varianza que se requiere en las nuevas dimensiones. Un valor más alto indica una mayor conservación de la varianza y, por lo tanto, una representación más detallada de los datos originales, mientras que un valor más bajo resulta en una reducción de la dimensionalidad con pérdida de información.

Una vez se ha reducido la dimensión del conjunto de datos inicial en un dataframe de menor tamaño basado en componentes principales, el siguiente paso es determinar los clusters para agrupar a los jugadores. Para ello, se añade una segunda funcionalidad, como se puede ver en la Figura 51, que se trata de un selector que da la oportunidad al usuario de elegir cómo hacer la selección de dichos clústers, si de forma manual, siendo el propio usuario el que determina el número de clusters específico que desea o, si por el contrario, quiere que sea el propio algoritmo el que decida el número óptimo de clústers en función de la muestra.

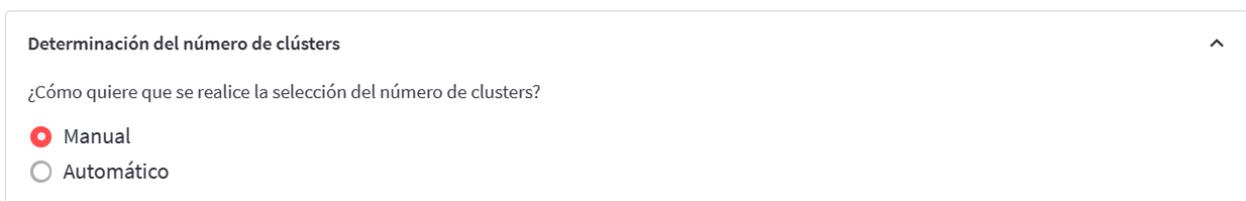


Figura 51: Personalización n°2 Machine Learning (Pestaña 5: Clustering)

Si el usuario decide hacerlo de forma manual, se añade un expander adicional, en el que figura: “Gráfica

del método del codo”. A continuación, se va a definir qué es el método del codo y qué implica en la decisión del número óptimo de clusters en una agrupación de datos.

- **Método del codo:** Según (Soria Polo, 2021), el método del codo es una metodología que busca definir el número idóneo de clusters en función del conjunto de datos que se tiene. Tiene dos principales premisas:

1. Minimizar la varianza intra-cluster (dentro de cada uno de los clústers).
2. Maximizar la varianza inter-cluster (entre los clústers).

La varianza se mide en términos de distancia, para lo cual hace falta definir los centroides de cada cluster. El centroide de un cluster es el punto equidistante a todos los puntos (en este caso, jugadores) que pertenecen a un mismo cluster.

A continuación, en la Figura 52, se añade una foto de cómo se ha implementado el método del codo:

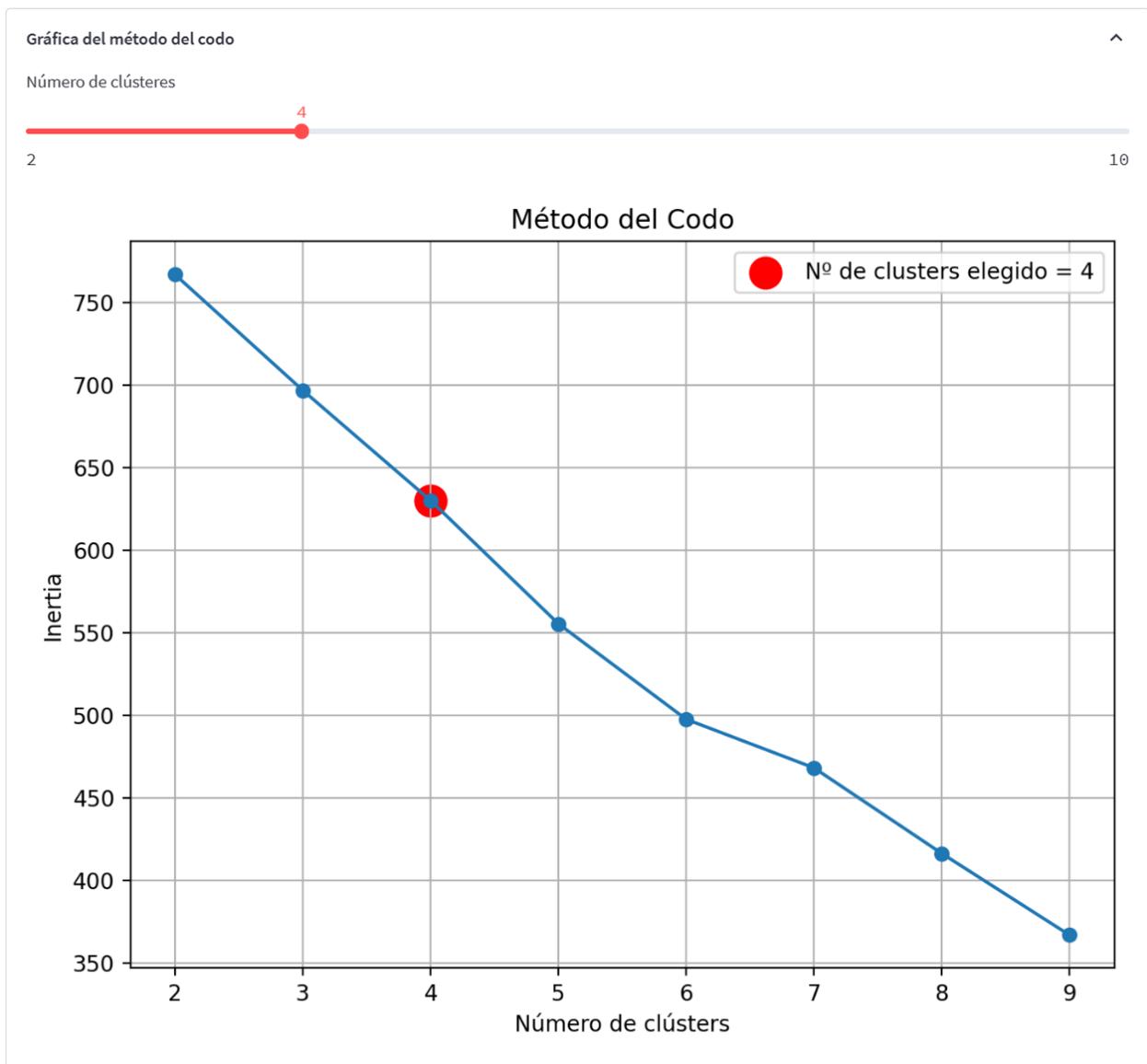


Figura 52: Personalización nº3 Machine Learning (Pestaña 5: Clustering)

Primero, el usuario debe observar la gráfica del método del codo y definir, de forma visual, dónde se observa un codo en la gráfica. Con ayuda del slider que se encuentra justo encima de la gráfica, el usuario decide cuál es el número de clústers que quiere definir y, en la gráfica, se resalta el número de clústers definido en el slider con un círculo rojo.

Si se decide, en el cálculo de clústers, que la selección es automática, la propia aplicación calculará cuál es el valor de clústers, en función de los datos, que propicia que haya un codo en la gráfica del método del codo.

Seguidamente, teniendo el conjunto de datos reducido a través del PCA y, además, habiendo seleccionado el número de clusters a definir en la muestra, lo que se debe determinar es la representación final de los datos. En este caso, se ha otorgado al usuario la posibilidad de definir si se quiere una tabla o un gráfico, como se puede ver en la Figura 53.

Representación del resultado

Elija una de las dos opciones siguientes:

Tabla

Gráfico

Figura 53: Opción nº4 Machine Learning (Pestaña 5: Clustering)

Por un lado, si se selecciona la opción ‘Tabla’, se mostrarán los jugadores que se han obtenido a partir de los filtros de selección previos clasificados en el número de clusters definido previamente. En la tabla, se podrá ver el nombre del jugador, el clúster al que pertenece y, por último, el índice de silueta calculado en cada caso.

El índice de silueta, según (Soria Polo, 2021), es una medición de la calidad del agrupamiento realizado, que se calcula a partir de la distancia intra-cluster. Esta distancia se mide en el rango $[-1, 1]$, de forma que:

- Valores cercanos a 1: Jugador alejado del resto de clústers (pertenece al cluster correcto).
- Valores cercanos a 0: Jugador cercano o incluso en la frontera entre dos clusters.
- Valores cercanos a -1: Jugador ubicado en un cluster erróneo.

Para calcular este valor, se utiliza la siguiente fórmula:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)}$$

En la cual:

- a_i representa la distancia promedio entre un jugador y el resto de los jugadores dentro del mismo clúster.
- b_i representa la distancia promedio entre un jugador y el resto de los jugadores del clúster más cercano diferente al del jugador. Esto mide cómo de lejos está el jugador de otros jugadores en el clúster vecino más cercano.

Por otro lado, si se selecciona la opción ‘Gráfico’, se mostrará un gráfico de dispersión bidimensional de jugadores en el cual se representan las dos componentes principales que recogen una mayor varianza de los datos originales.

Por último, se añade una quinta opción de personalización, como se puede ver en la Figura 54. En dicha opción, se tiene un multiselect de jugadores, el cual recoge todos los jugadores que pertenecen a la muestra después de aplicar los filtros de selección anteriores. El usuario tiene la capacidad de elegir al jugador o jugadores que más le interese representar, ya sea en la tabla o el gráfico final. De esta manera, se van a representar únicamente los clústers que incluyan a uno o más jugadores de los seleccionados en el multiselect dispuesto.

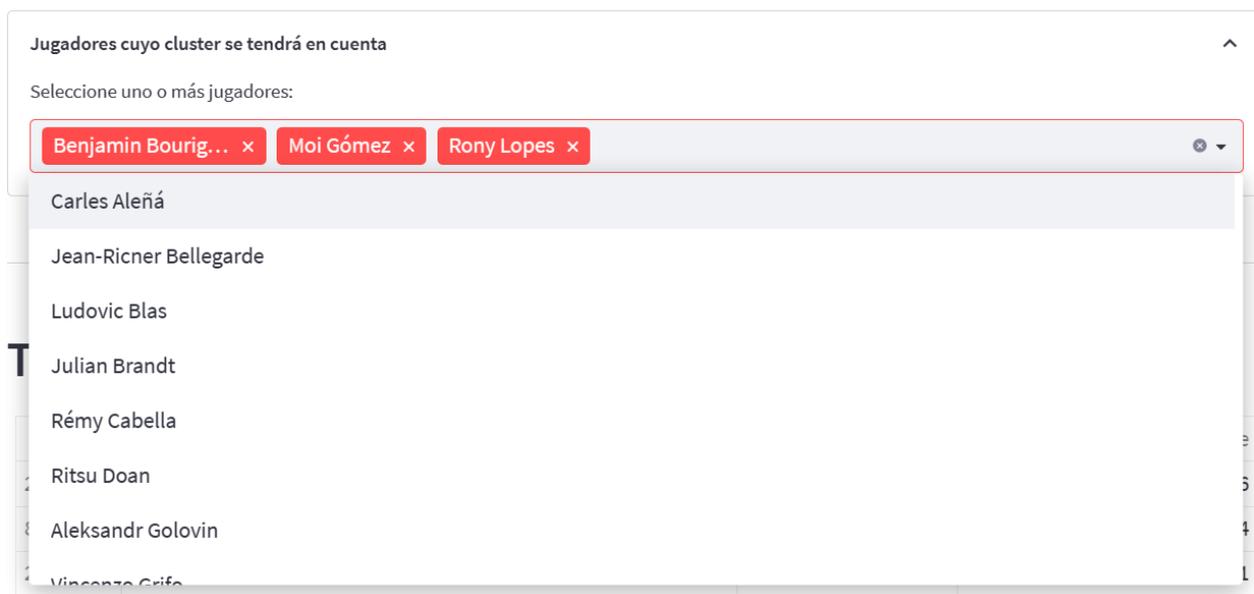


Figura 54: Opción n°5 Machine Learning (Pestaña 5: Clustering)

Una vez comentadas todas las opciones de personalización de esta pestaña, se añaden las Figuras 55 y 56 en las cuales se mostrarán ejemplos de cómo se representarían los resultados finales en función de qué se seleccione en el selector de la opción de personalización 4 (el selector tabla/gráfico).

Tabla Final de Jugadores

	Jugador	Cluster	Silhouette
0	Federico Valverde	1	0.1366
1	Moi Gómez	1	0.2164
2	Óliver Torres	2	-0.0001
3	Rubén Sobrino	2	0.1067
4	Elba Rashani	2	0.1587
5	Carles Pérez	2	0.1141
6	Rony Lopes	2	0.1323
7	Jérémy Le Douaron	2	0.0845
8	Andrej Kramarić	2	0.0544
9	Carles Aleñá	2	0.2326
10	Adrien Hunou	2	0.1189
11	Jean-Ricner Bellegarde	2	-0.0140
12	Óscar Trejo	2	0.1868
13	Vincenzo Grifo	4	0.1149
14	Rémy Cabella	4	0.1894
15	Julian Brandt	4	-0.0084
16	Benjamin Bourigeaud	4	0.1230
17	Ludovic Blas	4	0.0618
18	Florian Sotoca	4	0.1351
19	Florian Kainz	4	0.0962

Figura 55: Resultado final en tabla (Pestaña 5: Clustering)

Gráfico Final de Jugadores

Algoritmo de clusterización (PCA + KMeans) | 23 Jugadores // 4 Clusters

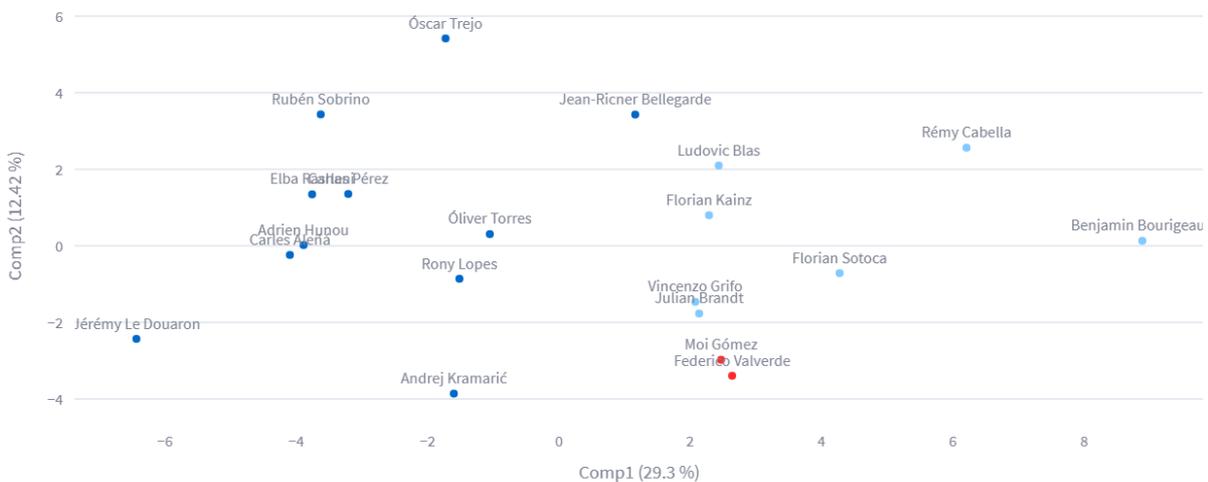


Figura 56: Resultado final en gráfico (Pestaña 5: Clustering)

4.4. Resultado

Este caso de uso se desarrollará de manera cronológica, siguiendo una secuencia lógica y ajustándose a las complejidades y desafíos inherentes a la gestión deportiva. El objetivo principal de este análisis es proporcionar una visión completa y práctica de cómo las directrices definidas a lo largo del proyecto pueden contribuir de manera significativa a la toma de decisiones y la gestión eficaz en el ámbito deportivo.

Con el fin de demostrar la aplicabilidad real del análisis realizado y para poner en práctica, por una parte, el cronograma de trabajo y la organización de la dirección deportiva, y por otra, la aplicación desarrollada, se propone un caso de uso realista de lo que sería una planificación de la plantilla en un club de fútbol, desarrollado de forma cronológica y ajustado a la realidad.

Como es lógico, al proponerse un desarrollo cronológico de los acontecimientos acaecidos en la planificación, se partirá de la base de la cronología definida en el apartado 3.2 (más concretamente, la cronología definida en la Figura 7 del trabajo).

4.4.1. Desarrollo de la Fase I

La Fase I, como se definió de forma más pormenorizada en el apartado 3.2 del trabajo, es un período de recolección y depuración de información relativa al rendimiento de los jugadores, para completar la base de datos del club. En esta fase, se definieron dos procesos principales: ‘Mapeo’ y ‘Cribado’.

Para que, tanto los scouts como los analistas de datos, puedan comenzar a trabajar, lo primero es definir las competiciones que, mensualmente, cada miembro de la Dirección Deportiva Data Driven tiene el deber de visualizar y analizar. Para ello, como se ha explicado debidamente en la definición de la aplicación, se ha habilitado una pestaña en la aplicación (Pestaña 2: Asignación de Ligas) cuya función es definir el reparto de ligas. Este reparto de ligas cobrará relevancia, en este caso de estudio concreto, a la hora de realizar los informes de rendimiento en el proceso de ‘Cribado’.

En las Figuras 57 y 58, se definen los repartos de ligas de este caso de estudio para scouts y analistas de datos.

	Scout1	Scout2	Scout3	Scout4	Scout5
Enero	["Premier League_ING"; "Superliga_SUI"; "Europa League"; "Liga_CHE"]	["La Liga_ESP"; "Conférence League"; "Liga NOS_POR"; "Liga HNL_CRO"]	["Liga_1_FRA"; "Basileirao_BRA"; "Jupiler Pro League_BEL"; "Champions League"]	["Serie A_ITA"; "Eredivisie_HOL"; "Ekstraklasa_POL"; "Primera División_ARG"]	["Bundesliga_ALE"; "Super Lig_TUR"; "Bundesliga_AUS"]
Febrero	["Serie A_ITA"; "Jupiler Pro League_BEL"; "Superliga_SUI"; "Liga_CHE"]	["La Liga_ESP"; "Liga NOS_POR"; "Conférence League"; "Liga HNL_CRO"]	["Premier League_ING"; "Champions League"; "Basileirao_BRA"; "Primera División_ARG"]	["Bundesliga_ALE"; "Europa League"; "Bundesliga_AUS"; "Ekstraklasa_POL"]	["Liga_1_FRA"; "Eredivisie_HOL"; "Super Lig_TUR"]
Marzo	["Serie A_ITA"; "Champions League"; "Liga HNL_CRO"; "Bundesliga_AUS"]	["La Liga_ESP"; "Liga_CHE"; "Ekstraklasa_POL"; "Basileirao_BRA"]	["Liga_1_FRA"; "Jupiler Pro League_BEL"; "Eredivisie_HOL"; "Conférence League"]	["Premier League_ING"; "Europa League"; "Super Lig_TUR"; "Primera División_ARG"]	["Bundesliga_ALE"; "Liga NOS_POR"; "Superliga_SUI"]
Abril	["Premier League_ING"; "Ekstraklasa_POL"; "Basileirao_BRA"; "Superliga_SUI"]	["La Liga_ESP"; "Jupiler Pro League_BEL"; "Super Lig_TUR"; "Champions League"]	["Bundesliga_ALE"; "Conférence League"; "Europa League"; "Bundesliga_AUS"]	["Liga_1_FRA"; "Liga_CHE"; "Eredivisie_HOL"; "Liga HNL_CRO"]	["Serie A_ITA"; "Primera División_ARG"; "Liga NOS_POR"]
Agosto	["Bundesliga_ALE"; "Jupiler Pro League_BEL"; "Super Lig_TUR"; "Eredivisie_HOL"]	["La Liga_ESP"; "Primera División_ARG"; "Liga_CHE"; "Europa League"]	["Serie A_ITA"; "Liga HNL_CRO"; "Conférence League"; "Liga NOS_POR"]	["Premier League_ING"; "Superliga_SUI"; "Basileirao_BRA"; "Champions League"]	["Liga_1_FRA"; "Bundesliga_AUS"; "Ekstraklasa_POL"]
Septiembre	["Premier League_ING"; "Bundesliga_AUS"; "Conférence League"; "Primera División_ARG"]	["Liga_1_FRA"; "Superliga_SUI"; "Super Lig_TUR"; "Liga HNL_CRO"]	["La Liga_ESP"; "Basileirao_BRA"; "Ekstraklasa_POL"; "Liga_CHE"]	["Bundesliga_ALE"; "Jupiler Pro League_BEL"; "Eredivisie_HOL"; "Liga NOS_POR"]	["Serie A_ITA"; "Europa League"; "Champions League"]
Octubre	["La Liga_ESP"; "Superliga_SUI"; "Basileirao_BRA"; "Primera División_ARG"]	["Premier League_ING"; "Liga HNL_CRO"; "Ekstraklasa_POL"; "Super Lig_TUR"]	["Bundesliga_ALE"; "Champions League"; "Europa League"; "Jupiler Pro League_BEL"]	["Serie A_ITA"; "Bundesliga_AUS"; "Liga_CHE"; "Liga NOS_POR"]	["Liga_1_FRA"; "Eredivisie_HOL"; "Conférence League"]

Figura 57: Reparto de Ligas para Scouts (Resultado)

	Analista 1	Analista 2	Analista 3
Enero	["Premier League_ING"; "Bundesliga_AUS"; "Ekstraklasa_POL"; "La Liga_ESP"; "Serie A_ITA"; "Europa League"; "Liga NOS_POR"; "Eliteserien_NOR"]	["Liga HNL_CRO"; "Allsvenskan_SUE"; "Jupiler Pro League_BEL"; "Conférence League"; "Bundesliga_ALE"; "Superliga_SUI"; "Super Lig_TUR"; "Liga MX_MEX"]	["Champions League"; "Liga_CHE"; "Superliga_SUI"; "Basileirao_BRA"; "Eredivisie_HOL"; "MLS_USA"; "Primera División_ARG"; "Liga_1_FRA"]
Febrero	["Primera División_ARG"; "Bundesliga_AUS"]	["Conférence League"; "Bundesliga_ALE"; "Superliga_SUI"; "Ekstraklasa_POL"; "Liga_1_FRA"; "Allsvenskan_SUE"; "Super Lig_TUR"]	["Premier League_ING"; "Liga MX_MEX"; "Basileirao_BRA"; "Europa League"; "Eredivisie_HOL"; "La Liga_ESP"; "MLS_USA"; "Eliteserien_NOR"]
Marzo	["Ekstraklasa_POL"; "Eliteserien_NOR"; "Jupiler Pro League_BEL"; "Europa League"; "Primera División_ARG"; "Eredivisie_HOL"; "Super Lig_TUR"; "Liga_1_FRA"]	["Superliga_SUI"; "Basileirao_BRA"; "Bundesliga_ALE"; "Serie A_ITA"; "Premier League_ING"; "La Liga_ESP"; "Bundesliga_AUS"; "Liga MX_MEX"]	["Liga NOS_POR"; "Champions League"; "MLS_USA"; "Liga HNL_CRO"; "Liga_CHE"; "Superliga_SUI"; "Allsvenskan_SUE"; "Conférence League"]
Abril	["Bundesliga_ALE"; "La Liga_ESP"; "Liga NOS_POR"; "Liga_CHE"; "Super Lig_TUR"; "Bundesliga_AUS"; "Eredivisie_HOL"; "Premier League_ING"]	["Liga HNL_CRO"; "Jupiler Pro League_BEL"; "Liga_1_FRA"; "Primera División_ARG"; "Eliteserien_NOR"; "Ekstraklasa_POL"; "MLS_USA"; "Basileirao_BRA"]	["Allsvenskan_SUE"; "Conférence League"; "Europa League"; "Champions League"; "Serie A_ITA"; "Liga MX_MEX"; "Superliga_SUI"; "Superliga_SUI"]
Agosto	["Jupiler Pro League_BEL"; "Premier League_ING"; "Liga NOS_POR"; "Liga_1_FRA"; "Allsvenskan_SUE"; "Eliteserien_NOR"; "Superliga_SUI"; "Champions League"]	["Liga MX_MEX"; "Basileirao_BRA"; "Super Lig_TUR"; "MLS_USA"; "Bundesliga_ALE"; "Liga_CHE"; "Bundesliga_AUS"; "Liga HNL_CRO"]	["Europa League"; "Eredivisie_HOL"; "Serie A_ITA"; "Conférence League"; "Ekstraklasa_POL"; "Superliga_SUI"; "Primera División_ARG"; "La Liga_ESP"]
Septiembre	["Liga MX_MEX"; "Europa League"; "La Liga_ESP"; "Eredivisie_HOL"; "Superliga_SUI"; "Superliga_SUI"]	["Champions League"; "Super Lig_TUR"; "MLS_USA"; "Serie A_ITA"; "Eliteserien_NOR"; "Bundesliga_AUS"; "Ekstraklasa_POL"; "Liga NOS_POR"]	["Premier League_ING"; "Liga_1_FRA"; "Allsvenskan_SUE"; "Basileirao_BRA"; "Conférence League"; "Liga HNL_CRO"; "Jupiler Pro League_BEL"; "Primera División_ARG"]
Octubre	["Allsvenskan_SUE"; "Eredivisie_HOL"; "Primera División_ARG"; "Basileirao_BRA"; "Superliga_SUI"; "Conférence League"; "Liga MX_MEX"; "Liga NOS_POR"]	["Bundesliga_ALE"; "Serie A_ITA"; "Champions League"; "Super Lig_TUR"; "Ekstraklasa_POL"; "Superliga_SUI"; "Liga HNL_CRO"; "MLS_USA"]	["Eliteserien_NOR"; "Europa League"; "Liga_1_FRA"; "La Liga_ESP"; "Bundesliga_AUS"; "Premier League_ING"; "Jupiler Pro League_BEL"; "Liga_CHE"]

Figura 58: Reparto de Ligas para Analistas de Datos (Resultado)

Una vez se ha definido el reparto de ligas para cada miembro, ante la imposibilidad de disponer de una dirección deportiva real en la cual poder aplicar esta metodología, se da por cerrado el período de ‘Mapeo’ en este caso de estudio.

El período de ‘Cribado’ de esta fase, en el presente caso de estudio y al igual que el período de ‘Mapeo’, es complicado de simular de forma realista al no estar inmiscuidos en el día a día de una dirección deportiva para poder realizar análisis de jugadores. En esta parte de la Fase I, el principal cometido, en este caso de los scouts, es evaluar, en función de los diferentes perfiles definidos en el trabajo, a los jugadores que, potencialmente, pueden ser objeto de adquisición por parte del club. Esta información, en la siguiente fase, se añadirá al informe final de cada uno de los jugadores.

4.4.2. Desarrollo de la Fase II

El desarrollo de la Fase II comienza con la reunión del Director Deportivo con el entrenador, con el fin de definir las necesidades concretas de la primera plantilla para el mercado. En esta reunión, se generará un informe de necesidades que marcará la hoja de ruta a seguir para todos los miembros de la Dirección Deportiva Data Driven.

Para llevar a cabo este análisis de forma relativamente realista, se deberá tener en cuenta el apartado económico, de gran relevancia a la hora de acometer una planificación. Por lo tanto, se partirá de una planificación de verano ya realizada por un club y, a partir de la información recogida de dicha planificación, se definirá un presupuesto aproximado como ejemplo. Esta información permitirá realizar una nueva planificación, haciendo uso del análisis de los miembros de la Dirección Deportiva Data Driven. Esta planificación de ejemplo servirá, además, para poder definir el informe de necesidades citado previamente.

4.4.2.1. Generación de informe de necesidades

Como resumen preliminar, el informe de necesidades es un análisis basado en criterios puramente deportivos que determina las posiciones a reforzar y los perfiles que deben de cumplir los jugadores que serán objeto de adquisición por parte del club.

La planificación que servirá de ejemplo para el resto de la simulación es la planificación del mercado de verano del Sevilla F.C. de la temporada 2023/2024. A continuación, se concretará la plantilla que tenía a comienzos de mercado y los fichajes y las salidas acometidas por la dirección deportiva del club. Esta información definirá el presupuesto aproximado para el mercado de fichajes a desarrollar en la simulación.

A continuación, se ha elaborado una infografía con la aplicación LINEUP11 (Figura 59) de la plantilla del Sevilla F.C. previa al mercado de fichajes de verano. Esta información se ha obtenido del portal web Transfermarkt (Transfermarkt, s.f.).



Figura 59: Plantilla Sevilla F.C. (Temporada 22/23)

A continuación, se definen las operaciones de entradas y salidas, únicamente teniendo en cuenta los jugadores definidos en la Figura 59. El resto de las operaciones no mencionadas en las tablas a definir son movimientos referentes a canteranos, que no reportan ningún ingreso a la entidad ni eran jugadores asiduos en la primera plantilla, por lo que carecen de interés en este caso.

Hay algunos jugadores que pertenecían a la plantilla, como se puede ver en la Figura 59, pero no son propiedad del club al estar cedidos y, terminada la temporada, vuelven a sus equipos de origen. Estos jugadores no se tendrán en cuenta como ventas del club, por lo que, simplemente, no se tendrán en cuenta en el análisis. Estos jugadores son: Alex Telles (Manchester United, Premier League), Pape Gueye (Olympique de Marsella, Ligue 1) y Bryan Gil (Tottenham Hotspur, Premier League).

Por un lado, en la Tabla 6, se pueden observar los jugadores adquiridos por parte del club junto a su valor de mercado. En el caso de las compras, por lo general, la mayoría de los clubes (entre ellos el Sevilla F.C.) adquiere jugadores libres y/o cedidos para compensar el gasto que sí pueda acometer en el fichaje de otros jugadores. En este trabajo, al no tener el conocimiento del mercado ni la cantidad de información que trabaja el personal del club sobre la situación particular de cada futbolista, se ha optado por coger como referencia el valor de mercado de cada jugador en vez del importe pagado por el club. De esta manera, se puede comparar y simular de manera más realista una planificación que teniendo en cuenta el coste real incurrido por el club.

COMPRA DE JUGADORES TEMPORADA 23/24			
Jugador	Posición	Procedencia	Valor de mercado (€)
Loïc Badé	Central	Stade Rennais (Ligue 1)	12.000.000
Djibril Sow	Centrocampista	Eintracht Frankfurt (Bundesliga)	22.000.000
Dodi Lukébakio	Extremo izquierdo	Hertha Berlín (2. Bundesliga)	12.000.000
Adrià Pedrosa	Lateral izquierdo	R.C.D. Espanyol (La Liga Hypermotion)	5.000.000
Sergio Ramos	Central	Agente Libre	4.500.000
Mariano Díaz	Delantero	Agente Libre	2.500.000
Orjan Nyland	Portero	Agente Libre	600.000
Boubakary Soumaré	Centrocampista	Leicester City (Championship)	20.000.000
TOTAL INVERTIDO			78.600.000

Tabla 6: Compra de jugadores Sevilla F.C. (Mercado de verano 2023/2024)

La cifra total de las incorporaciones, teniendo en cuenta el valor de mercado, es 78.600.000 €. Esta cifra, como se ha comentado previamente, es la cifra que se tendrá en consideración a la hora de simular una nueva planificación.

Definida esta cuestión, se estaría en condiciones de generar el informe de necesidades producto de la reunión con el entrenador por parte del director deportivo. Además, para personalizar aún más la planificación, se va a simular también la venta de jugadores. Para valorar la venta de los jugadores, al igual que ocurre con la

compra, al no tener conocimiento real de mercado, se va a tomar como referencia el valor de mercado de cada uno de los jugadores vendidos, para adicionarlos al presupuesto final de compra de la planificación simulada.

Para poder dar un informe de necesidades completo, se definirán primero las ventas a acometer, teniendo en cuenta aquellos jugadores que no son del agrado del entrenador y que son prescindibles desde su punto de vista.

Una vez conocidas las ventas pronosticadas, teniendo en cuenta la plantilla restante, se define el informe de necesidades para cubrir tanto las posiciones a reforzar producto de la venta definida en la Tabla 7, como las necesidades concretas del club que ya tenía antes de producirse dichas ventas.

VENTA DE JUGADORES TEMPORADA 23/24		
Jugador	Posición	Valor de mercado (€)
Marko Dmitrovic	Portero	4.000.000
Tanguy Nianzou	Central	7.500.000
Gonzalo Montiel	Lateral derecho	12.000.000
Joan Jordán	Centrocampista	9.000.000
Óliver Torres	Centrocampista	9.000.000
Papu Gómez	Mediapunta	3.000.000
Adnan Januzaj	Extremo izquierdo	6.000.000
Rafa Mir	Delantero	10.000.000
TOTAL INVERTIDO		60.500.000

Tabla 7: Venta de jugadores Sevilla F.C. (Mercado de verano 2023/2024)

Por lo tanto, el presupuesto total con el que se cuenta en la planificación a simular, producto de la suma de la compraventa de jugadores, es de 139.100.000€.

Para definir de manera realista las necesidades y, por ende, la plantilla, hay que tener en cuenta la normativa UEFA, al estar el Sevilla F.C. en competiciones europeas. La normativa de La Liga española determina que la composición de la plantilla debe ser de 25 jugadores, pero la normativa UEFA en este respecto es más concreta. A partir de la información obtenida de la página web oficial de UEFA (UEFA, 2016), se definen las normas para tenerla en cuenta de cara a la simulación de la planificación:

- **Lista A:** Máximo 25 jugadores, dos de los cuales deben ser porteros. Además, hay un cupo de ocho plazas reservadas a “jugadores formados localmente”, habiendo dos categorías dentro de este cupo:
 - Jugadores canteranos (formados en el club): Futbolistas que han estado en la cantera del club durante mínimo tres años entre los 15 y los 21 años.
 - Jugadores formados en la misma federación del club: Futbolistas que se han formado durante los 15 y 21 años en un club de la misma federación que el club origen. Máximo 4 jugadores dentro de los 8 del cupo.

Si un club tiene menos de ocho jugadores de estas dos categorías mencionadas previamente, el club debe restar jugadores a la lista A, por lo que debería tener una plantilla de menos de 25 jugadores.

- **Lista B:** Jugadores menores de 21 años que han jugado en el club durante un período ininterrumpido de dos años desde los 15 años desde su inscripción en la UEFA, o durante tres años ininterrumpidamente en un club de la misma federación (con un máximo de un año de cesión).

Como no se dispone de información actualizada del rendimiento de la cantera, solo se va a definir la primera plantilla del equipo o, lo que es lo mismo, solo se definirá la lista A.

El entrenador y el director deportivo, ahora sí, se encuentran en condiciones de definir las necesidades del club en cada una de las posiciones. Las necesidades de cada jugador van a depender, en gran parte, de las características del otro jugador o jugadores con los que comparte posición el jugador a detallar, ya que, por lo general, se busca que todas las posiciones estén bien cubiertas tanto en número como en variedad de atributos de los jugadores.

Se detalla, a continuación, cada una de las posiciones a reforzar y el perfil, de manera resumida, de cada uno de los jugadores:

1. **Portero:** Se requiere un portero que destaque, sobre todo, en reflejos y en juego aéreo, con el fin de equilibrar las características de la plantilla, ya que el juego de pies es una característica a destacar de Yassine Bounou, el otro portero en plantilla. Además, se busca un portero relativamente joven y con potencial, con el fin de que se pueda obtener una plusvalía en el futuro si el rendimiento es positivo.
2. **Lateral derecho:** Se necesita un lateral derecho con clara vocación ofensiva y buen rendimiento con balón, capaz de generar acciones de peligro a través de centros al área. Como se requiere un buen rendimiento ofensivo del jugador y también se requiere que defienda, el jugador debe tener buenas capacidades físicas para repetir esfuerzos constantemente. Teniendo en cuenta las limitaciones respecto a las fichas y, además, sabiendo que el club ha optado por incluirlo en la primera plantilla, el lateral derecho a adquirir será Juan Luis Sánchez (Juanlu).
3. **Central (1):** Se precisa un central con buena salida de balón, al no disponer de un central con buen toque en plantilla. Además, debe tener una gran capacidad de anticipación y de fortaleza en duelos defensivos debido a que, al tener la línea defensiva tan avanzada en el terreno de juego hacia campo

rival, el central deberá tener capacidad de recuperación de balón y de dominancia en segundas jugadas.

4. **Central (2):** Se busca un central potente en el duelo, con gran capacidad de anticipación y dominante en el juego aéreo. Debe tener una muy buena capacidad de colocación y ser aseado en salida de balón, sin ser esto último primordial en sus características. Por la misma razón que ocurre con el lateral, de cara a confeccionar una plantilla lo más amplia posible para hacer frente de forma holgada en el aspecto físico a todas las competiciones (3, en el caso del Sevilla F.C.), se va a replicar, en este caso, el fichaje de Sergio Ramos, que también figuraría como canterano en la lista UEFA y, por tanto, permitiría adquirir a otro futbolista más.
5. **Lateral izquierdo:** Se requiere un lateral izquierdo de gran recorrido y de vocación claramente ofensiva, ya que el entrenador basa gran parte de su estilo de juego en la apertura del campo a lo ancho, con la finalidad de generar espacios en la línea defensiva rival y, por ende, dificultarles las labores defensivas. Debe tener, además, una buena salida de balón, sobre todo importante por la habilidad del jugador con el balón en conducción.
6. **Centrocampista (1):** El primero de los centrocampistas a adquirir debe ser un pivote defensivo, para alternar con Fernando Reges que, debido a su edad (36 años), no puede jugar la mayoría de los partidos (de hecho, según el propio futbolista afirmó para el periódico Estadio Deportivo (Sáez, 2022) las siguientes declaraciones: "Yo creo que el Sevilla FC tiene que buscar un recambio para mí, pues yo tengo 35 años").

En este caso, se necesitaría un jugador con muy buenas características defensivas: excelente contundencia, colocación y anticipación. Además, se le pide tener buenas capacidades con balón, aunque son menos importantes que los atributos defensivos.

7. **Centrocampista (2):** El segundo centrocampista a fichar debería ser un mediocentro de ida y vuelta, lo que se conoce en la jerga futbolística, por su terminología en inglés, como box-to-box (de área a área). Este tipo de centrocampistas deben ser muy dinámicos y tener muy buenas capacidades con balón (técnica y visión de juego) y de esfuerzo. Además, deben ser relativamente buenos en conceptos defensivos de colocación y en rendimiento ofensivo. Son los centrocampistas más completos porque, por lo general, deben desarrollar su juego en varias zonas del campo, lo que hace que deban tener buen rendimiento tanto defendiendo como atacando.
8. **Centrocampista (3):** El tercer centrocampista requerido debe tener una gran capacidad técnica y de control de balón, además de una muy buena visión de juego para dar pases filtrados y permitir al equipo avanzar con la posesión de balón en el campo. Debe tener nociones defensivas, pero no van a ser el requerimiento principal.

Como apunte, este tercer centrocampista sería necesario debido al cambio de posición de Iván Rakitic en la confección de la plantilla, que ha pasado de ser mediocentro a ser mediapunta. Esto se debe a que cualquier jugador del mediocentro dentro de la filosofía de juego del club, se le debe exigir una

capacidad de esfuerzo y briega relativamente alta, cuestión que Iván no cumple con suficiencia debido a su edad.

9. Mediapunta: Como no se dispone de un perfil de las siguientes características en la plantilla, se busca un mediapunta que pueda jugar tanto en el último tercio de campo como delantero. Sus principales atributos deben ser: el rendimiento ofensivo, capacidad goleadora y asistencial y la visión de juego. En menor medida, se precisa que el jugador tenga buena técnica con balón y que sepa colocarse y anticiparse, con el fin de que sea intenso en la segunda jugada cuando el equipo vaya a presionar alto al rival.

10. Extremo izquierdo: Se requiere un extremo izquierdo con gran desequilibrio de cara a portería y con capacidad de generar tanto goles como asistencias. Además, debe tener buen trato de balón y capacidad para ganar duelos.

11. Delantero: Se busca un delantero con características distintas a Youssef En-Nesyri, el otro delantero en plantilla. Por lo tanto, aparte de una gran capacidad goleadora, característica primordial en el delantero, debe tener buen rendimiento en métricas relativas a la posesión de balón, en términos de generación de asistencias a otros compañeros, generación de juego y técnica con balón, para ayudar al equipo a dar continuidad al juego.

Esto hace un total de 11 fichajes, con dos de ellos ya definidos (Sergio Ramos y Juanlu), cuyo coste (en este caso, valor de mercado), se restará del total del presupuesto. Este informe de necesidades se trasladará a los miembros de la Dirección Deportiva Data Driven para que, finalmente, den la lista de jugadores finales para cada una de las posiciones a reforzar.

4.4.2.2. Análisis del informe de necesidades en la aplicación

Como apunte inicial, hay que tener en cuenta que el trabajo de los scouts es puramente de observación de partidos y en base a criterios subjetivos, en la simulación que se va a realizar, únicamente se va a mostrar cómo se haría la selección por parte de los expertos en datos, que harán uso de la aplicación para escoger, en base a los requerimientos concretos en cada posición, los jugadores que mejor se adaptan a las características definidas por el entrenador. Por lo tanto, el próximo paso en la simulación es trasladar estas instrucciones a la aplicación, para poder, después, extraer las conclusiones que nos otorga.

Lo importante a la hora de traducir las características definidas en el informe de necesidades a la aplicación, es realizar una priorización de características lo más parecida a la descripción dada por el entrenador para cada uno de los jugadores a fichar. En la memoria se va a reflejar un ejemplo para uno de los futbolistas a adquirir, siendo este proceso equivalente en el resto de los casos.

Se toma, como ejemplo, el fichaje del mediapunta. Revisando la descripción realizada previamente en el informe de necesidades, se incide en la importancia del rendimiento ofensivo, la capacidad goleadora y asistencial y, también, la posibilidad de jugar de delantero. Conociendo esta información, lo primero que se debe hacer es acceder a la primera de las pestañas de evaluación de rendimiento, que es 'Rating'.

Lo primero, en esta pestaña, es definir los filtros de selección:

- **Filtro de ligas:** Se eligen todas las competiciones que se recogen en la aplicación: La Liga (España), Premier League (Inglaterra), Bundesliga (Alemania), Ligue 1 (Francia), Serie A (Italia), Brasileirao (Brasil), Eredivisie (Holanda), Liga NOS (Portugal).
- **Filtro de posición:** Debe poder jugar de centrocampista (primera posición) y delantero (segunda posición).
- **Filtro de edad:** Entre 19 y 27 años.
- **Filtro de minutos:** Por lo menos, debe haber jugado un 60% respecto al máximo de minutos posibles en el filtro, en este caso, 2027 minutos.

Definidos los filtros de selección, el siguiente paso es realizar la traducción de las características a partir de la descripción definida en el informe de necesidades, como se puede ver en la Figura 60.

Características relevantes ^

Marque del 1 (mayor importancia) al 8 (menor importancia) SIN REPETIR las siguientes características:

Generación de juego

1 8

Técnica con balón

1 8

Colocación y anticipación

1 8

Capacidad goleadora

1 8

Rendimiento ofensivo

1 8

Visión de juego

1 8

Capacidad asistencial

1 8

Contundencia

1 8

¡Calificaciones válidas! Calculando los pesos de las métricas...

Figura 60: Priorización de características en la pestaña 'Rating'

4.4.2.3. Simulación de la planificación

Calculados los pesos de las métricas, a partir de los jugadores que cumplen los filtros de selección, se pueden observar los jugadores con mejor rendimiento a partir de la priorización definida para la posición mediapunta, que se pueden ver en la Figura 61. En este caso, hay bastantes jugadores que exceden el presupuesto recomendable a invertir en un futbolista, por lo que habrá que filtrar esta lista con los datos de valor de mercado buscados en Transfermarkt.

	Jugador	Equipo	Competicion	Rating
0	Bruno Fernandes	Manchester Utd	Premier League	7.38
1	Dominik Szoboszlai	RB Leipzig	Bundesliga	5.90
2	Jamal Musiala	Bayern Munich	Bundesliga	5.82
3	Ludovic Blas	Nantes	Ligue 1	5.71
4	James Maddison	Leicester City	Premier League	5.69
5	Julian Brandt	Dortmund	Bundesliga	5.43
6	Lee Kang-in	Mallorca	La Liga	5.28
7	Jean-Ricner Bellegarde	Strasbourg	Ligue 1	5.23
8	Christoph Baumgartner	Hoffenheim	Bundesliga	4.69
9	Aleksandr Golovin	Monaco	Ligue 1	4.67
10	Alex Iwobi	Everton	Premier League	4.55
11	Federico Valverde	Real Madrid	La Liga	4.46
12	Jacob Ramsey	Aston Villa	Premier League	4.32
13	Nikola Vlašić	Torino	Serie A	4.28
14	Dejan Kulusevski	Tottenham	Premier League	4.24
15	Rony Lopes	Troyes	Ligue 1	4.18
16	Eberechi Eze	Crystal Palace	Premier League	4.17
17	Rodrigo Riquelme	Girona	La Liga	4.14
18	Bernardo Silva	Manchester City	Premier League	4.01
19	Emi Buendía	Aston Villa	Premier League	3.98
20	Joelinton	Newcastle Utd	Premier League	3.96

Figura 61: Tabla final para el puesto mediapunta

Para la pestaña ‘Clustering’, se van a aplicar los mismos filtros y características que los definidos en la pestaña ‘Rating’. Además, como se explicó en el epígrafe dedicado a la aplicación, la pestaña ‘Clustering’ ofrece distintas opciones de personalización del algoritmo de Machine Learning. Para estas opciones, se va a definir:

- Selección de Componentes Principales: 90% porcentaje mínimo de varianza explicada por las componentes principales.
- Determinación del número de clusters: automático. Se obtienen 4 clusters.
- Jugadores cuyo cluster se tendrá en cuenta: en este filtro, se añaden los tres jugadores con mejor nota

obtenidos en la pestaña 'Rating', que, como se puede ver en la Figura 61 son: Bruno Fernandes (Manchester United, Premier League), Dominik Szoboszlai (RB Leipzig, Bundesliga) y Jamal Musiala (Bayern Múnich, Bundesliga).

En este caso, se obtienen dos clusters. Bruno Fernandes pertenece al cluster 1, mientras que los otros dos jugadores comparten cluster (en este caso, el número 3). A continuación, se presentan los resultados en forma de tabla y de gráfico.

	Jugador	Cluster	Silhouette
0	Bruno Fernandes	1	0.3199
1	Alex Iwobi	1	0.1070
2	Jean-Ricner Bellegarde	3	0.0739
3	Ludovic Blas	3	0.1854
4	Julian Brandt	3	0.1732
5	Eberechi Eze	3	0.0591
6	Lee Kang-in	3	0.1308
7	James Maddison	3	0.1998
8	Jamal Musiala	3	0.1694
9	Dominik Szoboszlai	3	0.2064
10	Federico Valverde	3	0.0551
11	Nikola Vlašić	3	-0.0169

Figura 62: Tabla final de jugadores para la posición mediapunta obtenido de la pestaña Clustering

Por una parte, en la Figura 62, se puede observar como en el cluster de Bruno Fernandes sólo se incluye un jugador: Alex Iwobi, del Arsenal F.C. de la Premier League. Por otro lado, en el cluster de Szoboszlai y de Musiala se tienen otros 8 jugadores.

Analizando el índice silueta, se puede ver que Iwobi y Bruno Fernandes, aunque pertenecen al mismo cluster, no tienen un índice silueta parecido, por lo que están bastante distanciados uno de otro respecto al parecido de características.

En el cluster al que pertenecen los otros dos jugadores, se pueden ver algunos jugadores que tienen un índice silueta parecido a lo que tienen Musiala y Szoboszlai, como pueden ser: Ludovic Blas (Nantes, Ligue 1), Jean-Ricner Bellegarde (Strasbourg, Ligue 1), Nikola Vlasic (Torino, Serie A), Kang-in Lee (RCD Mallorca,

La Liga), Eberechi Eze (Crystal Palace, Premier League) y Julian Brandt (Borussia Dortmund, Bundesliga).

En la Figura 62, se puede ver la disposición de estos jugadores en un gráfico cuyos ejes representan las dos componentes principales que recogen una mayor varianza entre todas las componentes principales. Se debe puntualizar que esta distribución gráfica no es del todo significativa ya que recoge muy poca varianza de la total recogida por las componentes principales (aproximadamente un 40%), pero sirven para situar a los jugadores en el contexto del clustering de forma aproximada.

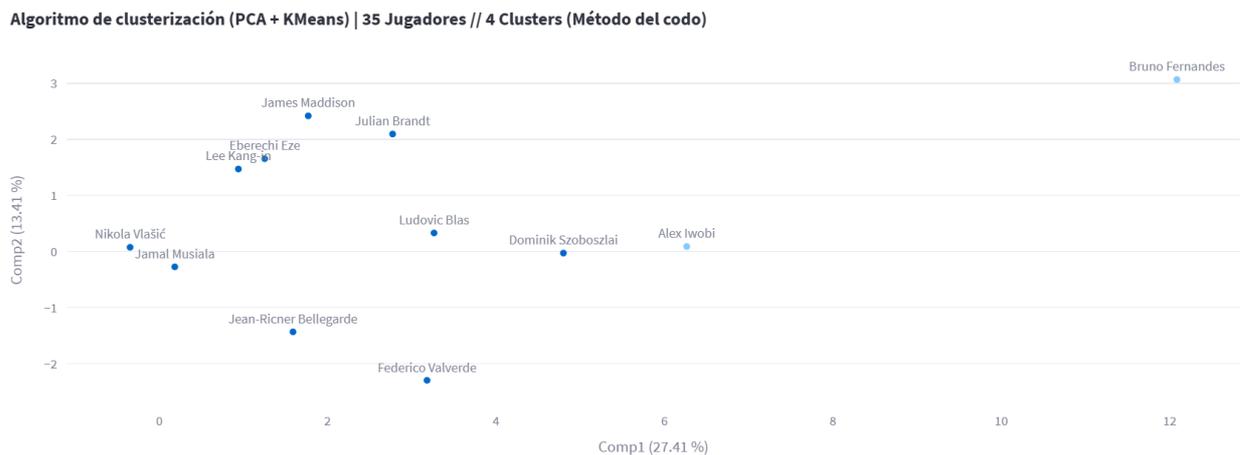


Figura 63: Gráfico final de jugadores para la posición mediapunta obtenido de la pestaña Clustering

A partir de la información recogida de las pestañas ‘Rating’ y ‘Clustering’, y buscando jugadores cuyo valor de mercado sea accesible para la planificación a realizar, se definen 5 alternativas (en cada posición se definirán el mismo número de alternativas), ordenadas de mejor a menor rendimiento esperado (de la A, mejor opción; a la E, opción menos buena), de cara a su futura adquisición. Para la posición de mediapunta, combinando los resultados obtenidos en ambas pestañas, se ha optado por los siguientes jugadores:

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 9 (MEDIAPUNTA)				
Opción	Jugador	Equipo	Competición	VM (€)
A	Ludovic Blas	Nantes F.C.	Ligue 1	20.000.000
B	Kang-in Lee	R.C.D. Mallorca	La Liga	22.000.000
C	Jean-Ricner Bellegarde	Strasbourg	Ligue 1	10.000.000
D	Nikola Vlasic	Torino	Serie A	17.000.000
E	Christoph Baumgartner	Hoffenheim	Bundesliga	24.000.000

Tabla 8: Lista final de jugadores para la posición mediapunta

A continuación, se va a proceder de igual manera para todas las posiciones en las que se debe adquirir un futbolista. Desde la Tabla 9 hasta la Tabla 18, se definen las opciones desde la A a la E para cada uno de los fichajes a acometer:

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 1 (PORTERO)

Opción	Jugador	Equipo	Competición	VM (€)
A	David Soria	Getafe C.F.	La Liga	8.000.000
B	Alexander Nübel	VfB Stuttgart	Bundesliga	8.000.000
C	Mory Diaw	Clermont Foot 63	Ligue 1	2.000.000
D	Ivan Provedel	SS Lazio	Serie A	12.000.000
E	Jeremías Ledesma	Cádiz C.F.	La Liga	8.000.000

Tabla 9: Lista final de jugadores para la posición portero

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 3 (CENTRAL)

Opción	Jugador	Equipo	Competición	VM (€)
A	Facundo Medina	RC Lens	Ligue 1	22.000.000
B	Arthur Theate	Stade Rennes	Ligue 1	22.000.000
C	Anthony Rouault	Toulouse F.C.	Ligue 1	10.000.000
D	Jhon Lucumí	Bolonia	Serie A	12.000.000
E	Rasmus Nicolaisen	Toulouse F.C.	Ligue 1	6.000.000

Tabla 10: Lista final de jugadores para la posición central

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 5 (LATERAL IZQUIERDO)

Opción	Jugador	Equipo	Competición	VM (€)
A	Javi Galán	R.C. Celta de Vigo	La Liga	12.000.000
B	Fabiano Parisi	F.C. Empoli	Serie A	10.000.000
C	Neto Borges	Clermont Foot 63	Ligue 1	4.000.000
D	Rogério	Sassuolo	Serie A	5.500.000
E	Tommaso Augello	Cagliari	Serie A	2.000.000

Tabla 11: Lista final de jugadores para la posición lateral izquierdo

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 6 (CENTROCAMPISTA: PIVOTE)

Opción	Jugador	Equipo	Competición	VM (€)
A	Branco Van den Boomen	Toulouse F.C.	Ligue 1	12.000.000
B	Salis Abdul Samed	RC Lens	Ligue 1	18.000.000
C	Stijn Spierings	Toulouse F.C.	Ligue 1	6.000.000
D	Walace	Udinese	Serie A	7.000.000
E	Azor Matusiwa	Stade de Reims	Ligue 1	7.500.000

Tabla 12: Lista final de jugadores para la posición centrocampista (pivote)

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 7 (CENTROCAMPISTA: BOX-TO-BOX)

Opción	Jugador	Equipo	Competición	VM (€)
A	Jean-Ricner Bellegarde	Strasbourg	Ligue 1	10.000.000
B	Lucas Robertone	U.D. Almería	La Liga	7.000.000
C	Angel Gomes	LOSC Lille	Ligue 1	18.000.000
D	Nabil Bentaleb	LOSC Lille	Ligue 1	7.000.000
E	Razvan Marin	F.C. Empoli	Serie A	8.000.000

Tabla 13: Lista final de jugadores para la posición centrocampista (box-to-box)

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 8 (CENTROCAMPISTA: TOQUE)

Opción	Jugador	Equipo	Competición	VM (€)
A	Enzo Le Fée	Lorient	Ligue 1	20.000.000
B	Sergi Darder	R.C.D. Espanyol	La Liga	16.000.000
C	Aleix García	Girona F.C.	La Liga	10.000.000
D	Téji Savanier	Montpellier	Ligue 1	10.000.000
E	André Almeida	Valencia C.F.	La Liga	20.000.000

Tabla 14: Lista final de jugadores para la posición centrocampista (toque)

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 10 (EXTREMO IZQUIERDO)				
Opción	Jugador	Equipo	Competición	VM (€)
A	Jonathan Bamba	LOSC Lille	Ligue 1	20.000.000
B	Dodi Lukebakio	Hertha Berlín	Bundesliga	12.000.000
C	Armand Laurienté	US Sassuolo	Serie A	18.000.000
D	Zakaria Aboukhlal	Toulouse F.C.	Ligue 1	9.000.000
E	Chris Führich	VfB Stuttgart	Bundesliga	4.000.000

Tabla 15: Lista final de jugadores para la posición extremo izquierdo

OPCIONES DE FICHAJE PARA LA ADQUISICIÓN 11 (DELANTERO)				
Opción	Jugador	Equipo	Competición	VM (€)
A	Folarin Balogun	Stade de Reims	Ligue 1	30.000.000
B	Habib Diallo	Strasbourg	Ligue 1	18.000.000
C	Sheraldo Becker	Unión Berlín	Bundesliga	17.000.000
D	Gaëtan Laborde	OGC Niza	Ligue 1	16.000.000
E	Marvin Ducksch	SV Werder Bremen	Bundesliga	7.500.000

Tabla 16: Lista final de jugadores para la posición delantero

4.4.2.4. Generación de informes finales

Los jugadores que aparecen en las listas finales definidas en el epígrafe anterior son los jugadores obtenidos a partir del análisis de los analistas de datos. Esta información, para poder decidir con fundamentos y con las opiniones de todos los miembros de la Dirección Deportiva Data Driven, se aglutinará en el informe final de rendimiento de cada jugador, realizado en PowerBI. Este informe sería análogo a todos los informes que debe realizar la dirección deportiva para que el Director Deportivo tenga información actualizada de cada jugador.

A continuación se ofrece la información de cada una de las pestañas diseñadas en Power BI, así como la estructura del informe y la información contenida en el mismo. Como jugador referencia, se va a mostrar la información de Folarin Balogun (Stade de Reims, Ligue 1), opción A para la delantera a partir de las listas finales elaboradas en el apartado anterior. Como en cada pestaña, a la izquierda de la visualización, se dispone de un buscador de jugadores, se puede elegir y, por consiguiente, mostrar, la información de cualquier jugador que se encuentre en la base de datos. Las pestañas serán las siguientes:

- **Información general:** Nombre, edad, fecha de nacimiento, posición(es) preferida(s), equipo y competición actuales, pie preferido, valor de mercado (consultado en Transfermarkt, web referencia para valorar a jugadores de fútbol), entre otros. Adicionalmente, a la izquierda de la visualización, se incluye un buscador de jugadores, como se puede observar en la Figura 64.

The screenshot displays a Power BI report for Folarin Balogun. On the left, there is a search bar labeled 'Jugador' with the text 'baloqun' and a selection box containing 'Folarin Balogun'. The main content area features a large blue header 'INFORMACIÓN GENERAL' and a white card with the player's name 'Folarin Balogun'. Below the name are three stacked cards: 'Equipo' with 'Reims', 'Competición' with 'Ligue 1', and a central portrait of the player. To the right of the portrait is a vertical column of attribute cards: 'País de nacimiento' (USA), 'Edad' (21), 'Posición 1' (DL), and 'Posición 2' (0).

Figura 64: Información general de Folarin Balogun (Pestaña 1 del informe final en Power BI)

- **Estadísticas generales de rendimiento:** Goles marcados, asistencias otorgadas, xG y xA, minutos jugados, tarjetas amarillas y rojas, entre otros. La información sobre esto se puede consultar en la

Figura 65, a continuación.



Figura 65: Estadísticas generales de Folarin Balogun (Pestaña 2 del informe final en Power BI)

- **Información recogida de la pestaña ‘Evaluación Scout’ de la aplicación:** Esta información hace hincapié en cada uno de los perfiles definidos tanto en el período ‘Cribado’ de la Fase I como en la propia definición de la pestaña en la aplicación. Tanto las calificaciones recogidas de esta pestaña por parte de los scouts como las observaciones generales se incluirán en el informe final. La información se representará en gráficos de radar, para visualizar fácilmente, para cada uno de los perfiles, las capacidades del jugador por parte del scout que haya analizado su rendimiento.

En las Figuras 66, 67 y 68 se podrán ver las tres pestañas en las que se determina esta información, que de un vistazo permite ver el rendimiento del jugador analizado desde distintas perspectivas, en este caso, perfiles.

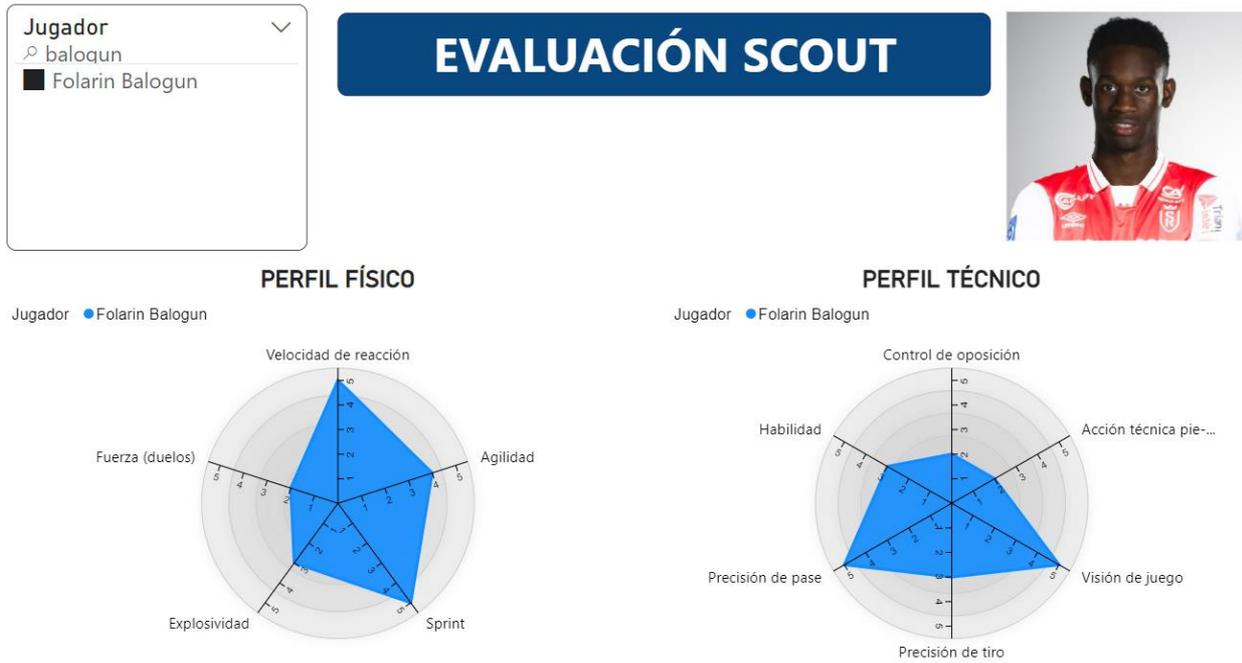


Figura 66: Información (I) 'Evaluación Scout' (Pestaña 3 del informe final en Power BI)

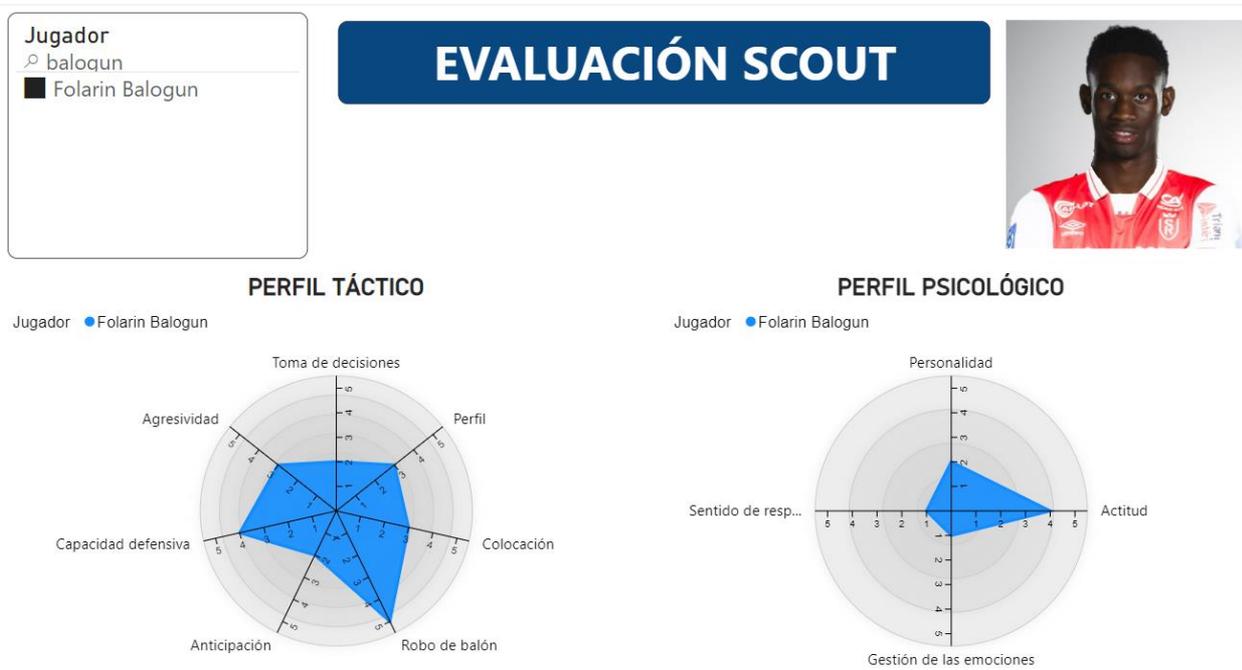


Figura 67: Información (II) 'Evaluación Scout' (Pestaña 4 del informe final en Power BI)



Figura 68: Información (III) 'Evaluación Scout' (Pestaña 5 del informe final en Power BI)

Además, en estas pestañas, al incorporarse gráficos de radar, se puede añadir varios futbolistas a una comparación, para que, en un gráfico de radar, se pueda ver la información de varios jugadores a la vez, a modo de comparativa, como se puede ver en la Figura 69. En este caso, se ha añadido la primera pestaña representando los valores de todos los jugadores que son opción en la lista final de delanteros.

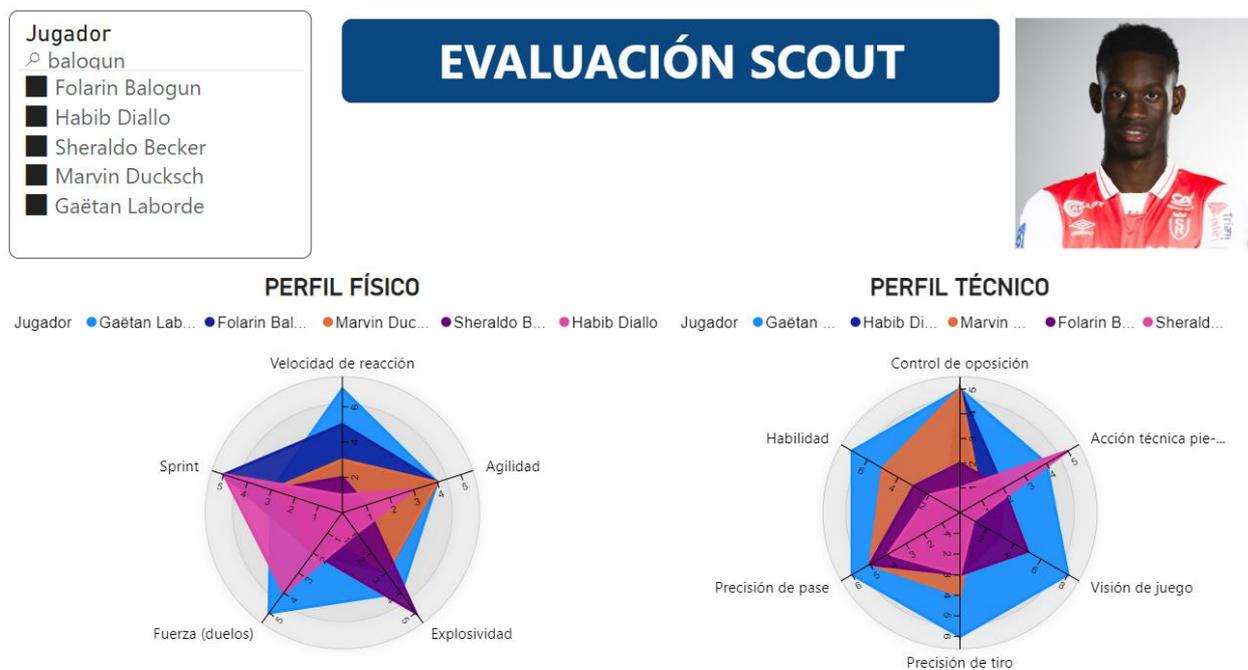


Figura 69: Información 'Evaluación Scout' comparativa delanteros (Pestaña 5 del informe final en Power BI)

- **Estadísticas avanzadas de rendimiento:** Los analistas de datos analizarán y añadirán información del rendimiento del jugador, para contextualizar si, para una métrica en concreto, un jugador destaca comparativamente respecto al resto de jugadores de su liga. En este apartado, se pueden considerar gráficas de rendimiento, comparación de métricas con respecto a otros jugadores, elaboración de percentiles de rendimiento, entre otros.
 - **Percentiles:** Según (Superprof, s.f.), Los percentiles son una herramienta estadística fundamental que divide a un conjunto de datos en 100 partes iguales. Esto, en un contexto futbolístico, permite evaluar y comparar el rendimiento de los jugadores en relación con un grupo de referencia más amplio. En esencia, lo que permite es identificar la posición relativa de un jugador en una métrica específica en comparación con el resto de los jugadores en la muestra de referencia. Por ejemplo, si evaluamos la velocidad de un futbolista y determinamos que se encuentra en el percentil 80 en comparación con otros jugadores de la misma competición, posición, etc., esto significa que ese jugador es más rápido que aproximadamente el 80% de los jugadores comparables a él. En cambio, si un jugador está en el percentil 30 en habilidades de regate, esto indica que está por debajo del 70% de los jugadores en términos de habilidad en esta área específica.

En la Figura 70, se puede ver la visualización que integra el análisis del rendimiento usando percentiles. En este caso, también se ha hecho uso de los gráficos de radar, haciendo uso algunas de las métricas extraídas de FBref a modo de visualización rápida del rendimiento del jugador.

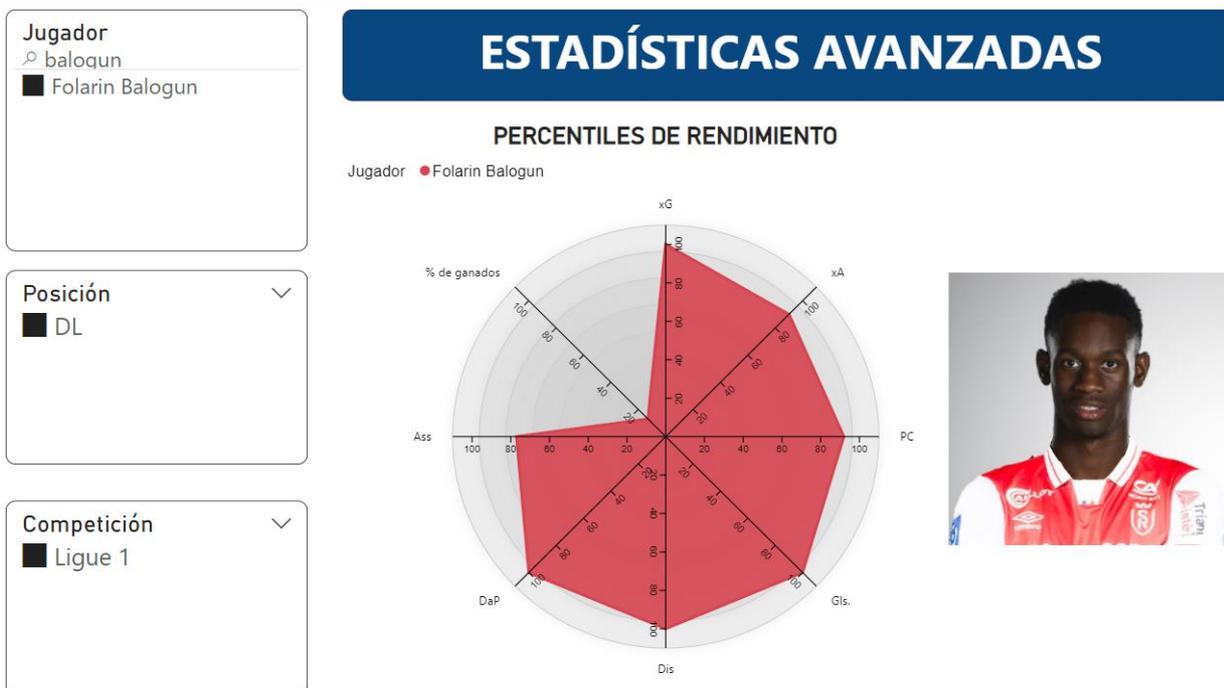


Figura 70: Representación de percentiles de Folarin Balogun (Pestaña 6 del informe final en Power BI)

En esta figura se puede ver como, a diferencia de las otras visualizaciones en la que el único filtrado era el de nombre de jugadores, en esta se ha añadido un filtrado por posición y competiciones. Sobre el rendimiento en particular del jugador, se puede ver como Balogun destaca sobremanera, estando en el percentil más alto en casi todas las métricas, salvo en el porcentaje de duelos aéreos ganados, en cuyo caso el nivel del jugador es relativamente más bajo. Esto puede ocurrir debido a que

De igual forma que en el caso de las pestañas de ‘Evaluación Scout’, se puede realizar una multiselección de jugadores a modo de comparativa, como se puede ver en la Figura 71. En este caso, también se va a plantear una comparativa de todos los delanteros que forman parte de la lista final para la posición delantero.

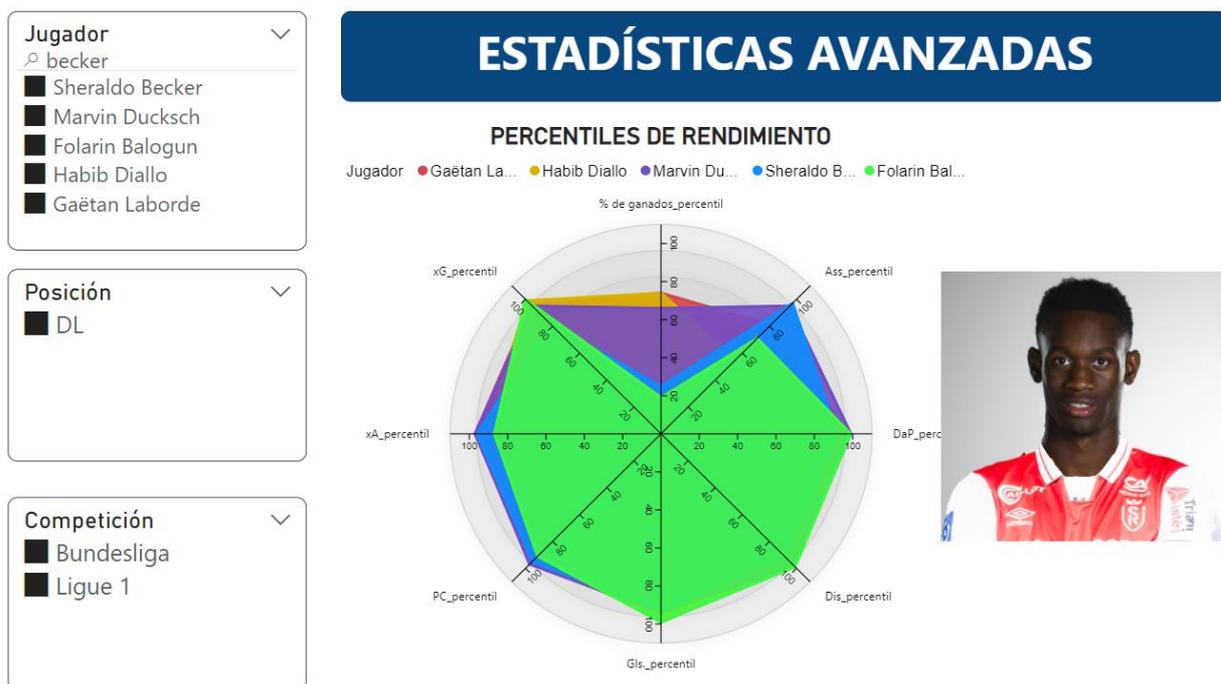


Figura 71: Comparativa de rendimiento a través de la representación de los percentiles (Pestaña 6 del informe final en Power BI)

En esta figura se puede observar cómo todos los jugadores elegidos tienen un rendimiento muy parecido en lo que respecta a las métricas de cara a portería (goles, goles esperados, disparos a puerta, entre otras), mientras que la gran diferencia entre ellos reside en el porcentaje de duelos aéreos ganados, en la cual los jugadores que más destacan, con un percentil alrededor de 80 son Gaëtan Laborde y Habib Diallo. Los percentiles calculados son en base a, en este caso, el resto de los delanteros de la base de datos, por lo que se puede observar que el rendimiento de los delanteros escogidos en la lista final es bastante bueno, ya que tienen un valor de percentil muy alto respecto al resto de delanteros de la muestra.

4.4.2.5. Planificación final

Para concluir el trabajo, se debe dar una planificación final, de la cual su máximo responsable va a ser siempre el Director Deportivo. Aunque tanto el cuerpo técnico, con el entrenador a la cabeza, como los miembros de la Dirección Deportiva Data Driven colaboren, al final el responsable y el que debe rendir cuentas ante el Consejo de Administración de la entidad es el Director Deportivo.

De forma resumida, se concretan las restricciones habidas para realizar la planificación final de la plantilla:

- **Restricciones debido a la lista UEFA para la lista A:**
 - 4 jugadores, mínimo, formados en canteras españolas, fuera del propio club.
 - 4 jugadores, máximo, formados en la cantera del Sevilla F.C. Si se tienen menos de 4, la cifra de jugadores en la lista A decrecerá desde 25 hasta el número de jugadores canteranos que no figuren en dicha lista. En este caso, como se tienen 3 jugadores formados en la cantera del club, la lista A deberá constar de 24 fichas profesionales.
- **Restricciones económicas:** 139.100.000€.

Como es lógico, se pueden hacer varias combinaciones de jugadores producto de las listas finales, siempre teniendo en cuenta que se cumplen las restricciones definidas previamente. Se va a dar una planificación final, pero se puede elegir cualquier otra combinación siempre y cuando se cumplan las restricciones mencionadas. En la Tabla 17, se puede ver la planificación final elegida.

PLANIFICACIÓN FINAL EN EL SEVILLA F.C. PARA EL MERCADO DE VERANO 2023/2024					
Opción	Jugador	Posición	Equipo	Competición	VM (€)
A	David Soria	Portero	Getafe C.F.	La Liga	8.000.000
-	Juanlu	Lateral derecho	Agente Libre	-	800.000
C	Anthony Rouault	Central	Toulouse F.C.	Ligue 1	10.000.000
-	Sergio Ramos	Central	Agente Libre	-	4.500.000
A	Javi Galán	Lateral izquierdo	R.C. Celta de Vigo	La Liga	12.000.000
A	Branco Van den Boomen	CC (Pivote)	Toulouse F.C.	Ligue 1	12.000.000
A	Jean-Ricner Bellegarde	CC (Box-to-box)	Strasbourg	Ligue 1	10.000.000
A	Sergi Darder	CC (Toque)	R.C.D. Mallorca	La Liga	16.000.000
A	Ludovic Blas	Mediapunta	Nantes F.C.	Ligue 1	20.000.000
B	Dodi Lukebakio	Extremo izquierdo	Hertha Berlín	Bundesliga	12.000.000
A	Folarin Balogun	Delantero	Stade de Reims	Ligue 1	30.000.000
TOTAL A INVERTIR					135.300.000

Tabla 17: Planificación final en el Sevilla F.C. para el mercado de verano 2023/2024

Se ha priorizado coger las opciones A en la zona del campo en la que el equipo está peor reforzado en la actualidad: del mediocampo hacia delante. En la zona defensiva y en la portería, ya se tenían mimbres para competir, pero se necesitaba completar plantilla, es por ello por lo que se ha optado por fichajes menos costosos.

Se cumplen las restricciones:

- **4 jugadores formados en canteras españolas:** David Soria, Javi Galán, Sergi Darder y Suso.
- **3 jugadores formados en la cantera del Sevilla F.C.:** Juanlu, Sergio Ramos y Jesús Navas.
- **Presupuesto a invertir:** 135.300.000€, no superando el presupuesto de 139.100.000€.

5 CONCLUSIONES

En el transcurso de este proyecto, se ha avanzado significativamente en la creación de una dirección deportiva basada en datos. La implementación de esta estrategia ha demostrado ser un paso crucial en la evolución del enfoque de la digitalización en la gestión deportiva. A medida que se ha avanzado en este proceso, se ha logrado identificar los elementos esenciales que constituyen una Dirección Deportiva Data Driven sólida y eficaz.

Primeramente, se han definido claramente los roles y responsabilidades de los miembros que compondrán esta dirección deportiva. Esta estructura organizativa proporciona una base sólida para que el equipo funcione de manera eficiente y efectiva, asegurando que cada miembro contribuya de manera significativa a la unión entre la toma de decisiones mediante la estrategia de visualización de partidos tradicional por parte de los ojeadores, como a la toma de decisiones basadas en datos.

Otro de los aspectos más destacados de este proyecto ha sido la concreción de la cronología de la operativa de la dirección deportiva, para llevar una mayor organización y gestión de los tiempos y la documentación necesaria para poder proceder de forma ordenada y efectiva, desde el primer análisis preliminar de partidos sin un objetivo claro hasta el mercado de fichajes donde se realiza la compraventa de jugadores.

Sin embargo, teniendo en cuenta la naturaleza de innovación y automatización de procesos que tiene este proyecto, el logro más importante y de mayor utilidad de este proyecto es la creación de la herramienta basada en análisis de datos aplicados al scouting en fútbol. Esta herramienta no solo simplifica el proceso de recopilación y análisis de datos, sino que también mejora la precisión y la calidad de las decisiones tomadas en el ámbito del scouting. Su implementación ha sido un gran paso adelante en la búsqueda de la excelencia en la gestión deportiva basada en datos.

Por último, a partir del análisis obtenido de esta aplicación, la generación de informes de jugadores y teniendo en cuenta las restricciones económicas y de contratos, se han definido distintas alternativas para cada uno de los roles que requería el entrenador y se ha dado una planificación final basada en la implementación de los datos en la operativa de la dirección deportiva.

Como futuras líneas de investigación, se proponen:

- Tener en consideración varias temporadas para contextualizar el análisis con una mayor cantidad de datos.
- Conseguir datos de más ligas para poder incorporar al análisis un mayor número de jugadores.
- No limitar el uso de datos al scouting y adaptarlo a otras áreas del club.

En resumen, este proyecto ha logrado con éxito su objetivo inicial de establecer una Dirección Deportiva Data Driven. Se ha avanzado significativamente en la integración de datos en nuestro enfoque de toma de decisiones, se ha contextualizado la operativa de gestión, se han definido roles de trabajo claros y se ha

proporcionado una herramienta de utilidad para la dirección deportiva.

Sin lugar a dudas, este proyecto incide en una línea en constante evolución en la gestión deportiva como es la utilización de los datos en el ámbito del fútbol para poder encontrar ventajas competitivas que hagan progresar al club deportiva y económicamente. Este trabajo y todo el estudio realizado en este ámbito hará avanzar a los clubes hacia la modernidad y la autogestión, que sin duda son metas no solo del fútbol en particular, sino de la industria en general.

REFERENCIAS

- Aggylon BrandCenter. (s.f.). *Rebranding del Sevilla FC - Aggylon..* Retrieved May 4, 2023, from <https://aggylon.com/blog/sevilla-fc-unidos-por-una-pasion-y-bajo-una-marca-bien-gestionada/>
- Amazon. (s.f.). *Amazon Web Services.* Retrieved May 16, 2023, from <https://aws.amazon.com/es/what-is/machine-learning/>
- Badillo, S., Banfai, B., Birzele, F., Davydov, I., Hutchinson, L., Kam-Thong, T., . . . Zhang, J. (2020). An Introduction to Machine Learning. *Clinical Pharmacology & Therapeutics*, 107, 871-885. Retrieved May 5, 2023, from doi:10.1002/cpt.1796.
- Big Data International Campus. (s.f.). *Procesos de Análisis de Datos..* Retrieved July 2, 2023, from <https://www.campusbigdata.com/big-data-blog/item/104-procesos-de-analisis-de-datos>
- Boletín Oficial del Estado. (27 de Febrero de 2023). *Disposición 6346 del BOE núm. 59 de 2023..* Retrieved April 21, 2023, from <https://www.boe.es/boe/dias/2023/03/10/pdfs/BOE-A-2023-6346.pdf>
- Canle, E. (s.f.). *¿Cuánto gana un director deportivo de fútbol? ¡Descúbrelo!..* Retrieved May 16, 2023, from <https://masters.es/blog/director-deportivo-futbol-salario/>
- Centro de Estudio y Desarrollo del Fútbol (CEyDF). (s.f.). Retrieved August 13, 2023, from Aspectos del análisis a desarrollar en el jugador: <https://www.ceydef.com/analisis/del-jugador>
- Contreras, E. (8 de Febrero de 2020). *Así ficha Monchi: departamento I+D, data y algoritmo entre 18.000 futbolistas | Marca.com..* Retrieved April 29, 2023, from <https://www.marca.com/primeraplana/2020/02/08/5e3709e822601d8a5b8b458e.html>
- Disfrutalasmaticas. (s.f.). *Distribución Normal Estándar..* Retrieved June 26, 2023, from <https://www.disfrutalasmaticas.com/datos/distribucion-normal-estandar.html>
- Emerj. (1 de Mayo de 2023). *Artificial Intelligence at Alibaba – Two Current Use-Cases | Emerj Artificial Intelligence Research..* Retrieved June 12, 2023, from <https://emerj.com/ai-sector-overviews/artificial-intelligence-at-alibaba/>
- FBref. (s.f.). *Estadísticas e Historia del Fútbol | FBref.com..* Retrieved June 23, 2023, from <https://fbref.com/es/>
- Gómez, J. J. (s.f.). *Web scraping con Python. Extraer datos de una web. Guía de inicio de BeautifulSoup.* Retrieved August 12, 2023, from <https://j2logo.com/python/web-scraping-con-python-guia-inicio-beautifulsoup/>
- Google. (s.f.). *Advancing Healthcare Research & AI in Medicine - Google Health.* Retrieved April 28, 2023, from <https://health.google/health-research/>

- H., A. E.-S., Ahmed Hendawi, A. M., & El Bastawissy, A. H. (1 de Julio de 2011). A proposed model for data warehouse ETL processes. Retrieved July 24, 2023, from *Computer and Information Sciences*, 23(2), 91-104.
- Herrera Gómez-Platero, V. (6 de Julio de 2019). *Medio centros (tipos de mediocentros y posicionamiento en diferentes sistemas) | Futbol+Qpartidos..* Retrieved July 12, 2023, from <https://futbolmasquepartidos.wordpress.com/2019/07/06/medio-centros-tipos-de-mediocentros-y-posicionamiento-en-diferentes-sistemas/>
- IBM. (s.f.). *IBM Cloud*. Retrieved May 16, 2023, from <https://www.ibm.com/es-es/topics/supervised-learning>
- Idoate, G. (s.f.). *Juego de Posición: Qué es, Objetivos e Información..* Retrieved July 13, 2023, from <https://www.misentrenamientosdefutbol.com/diccionario/juego-de-posicion>
- Immune Technology Institute. (17 de Marzo de 2022). *Librerías de Python, ¿qué son y cuáles son las mejores?* Retrieved June 12, 2023, from <https://immune.institute/blog/librerias-python-que-son/>
- Jupyter. (s.f.). *Project Jupyter | Home..* Retrieved August 3, 2023, from <https://jupyter.org/>
- Karamizadeh, S., Abdullah, S. M., Manaf, A. A., Zamani, M., & Hooman, A. (2013). An Overview of Principal Component Analysis. Retrieved August 27, 2023, from *An Overview of Principal Component Analysis*, 173-175.
- La Liga. (31 de Enero de 2020). *La transformación digital del Sevilla FC: una revolución basada en mejorar la relación con sus aficionados | Fútbol Global..* Retrieved April 22, 2023 from <https://newsletter.laliga.es/futbol-global/la-transformacion-digital-del-sevilla-fc-una-revolucion-basada-en-mejorar-la-relacion-con-sus-aficionados-1>
- Luque, F. (13 de Abril de 2013). *Sistema de juego 1-4-2-3-1, el sistema de moda | Fútbol en Positivo..* Retrieved September 2, 2023, from <https://futbolenpositivo.com/sistema-de-juego-1-4-2-3-1-el-sistema-de-juego-de-moda/>
- Mathworks. (s.f.). *MathWorks*. Retrieved May 16, 2023, from <https://es.mathworks.com/discovery/reinforcement-learning.html#:~:text=Reinforcement%20learning%20es%20una%20t%C3%A9cnica,error%20con%20un%20entorno%20din%C3%A1mico.>
- Microsoft. (s.f.). *¿Qué es Power BI?* Retrieved September 20, 2023, from <https://powerbi.microsoft.com/es-es/what-is-power-bi/>
- Microsoft. (s.f.). *Extracción, transformación y carga de datos (ETL) - Azure Architecture Center | Microsoft Learn..* Retrieved June 18, 2023, from <https://learn.microsoft.com/es-es/azure/architecture/data-guide/relational-data/etl>
- Onrubia, A. (26 de Abril de 2022). *El Toulouse asciende a Ligue 1 usando la base de datos del Football*

- Manager* - *AS.com*.. Retrieved April 11, 2023, from https://as.com/futbol/2022/04/26/internacional/1650962414_446462.html#:~:text=El%20gran%20exp%20del%20Big,plazo%20a%20la%20Ligue%201.
- OptaAnalyst. (5 de Junio de 2023). *Opta Analyst (@OptaAnalyst) / Twitter*.. Retrieved August 6, 2023, from <https://twitter.com/OptaAnalyst/status/1665677072218374144/photo/1>
- Pamos, J. M. (s.f.). *Qué factores psicológicos analizar del futbolista en el proceso de scouting*.. Retrieved August 1, 2023, from https://objetivoanalista.com/que-factores-psicologicos-analizar-futbolista-scouting/#A_El_propio_jugador
- Paypal. (21 de Diciembre de 2021). *Leveraging Machine Learning for Fraud Prevention | PayPal US*.. Retrieved May 5, 2023, from <https://www.paypal.com/us/brc/article/paypal-machine-learning-stop-fraud>
- Razablan. (8 de Octubre de 2020). *El Toulouse usará Football Manager para fichar y confiará su futuro a las bases de datos - Millenium*.. Retrieved April 12, 2023, from <https://www.millenium.gg/noticias/27902.html>
- Rocha, C. (21 de Marzo de 2023). *El 'big data' que sí le funciona Monchi y que ahora el Sevilla quiere llevar a las gradas*.. Retrieved April 22, 2023, from https://www.elconfidencial.com/deportes/futbol/liga/2023-03-21/el-big-data-que-si-le-funciona-monchi-y-que-ahora-el-sevilla-quiere-llevar-a-las-gradas_3596201/
- Sáez, A. (30 de Septiembre de 2022). *Fernando revela la verdadera razón de su bajón físico (y no es el tobillo)*.. Retrieved September 9, 2023, from <https://www.estadiodeportivo.com/futbol/sevilla-fc/fernando-revela-verdadera-razon-bajon-fisico-tobillo-20220929-405269.html>
- Santander. (9 de Abril de 2021). *Santander Becas*.. Retrieved May 16, 2023, from <https://www.becas-santander.com/es/blog/python-que-es.html>
- Sevilla F.C. (18 de Mayo de 2021). *La marca Sevilla FC, 23ª más valiosa del mundo | Sevilla FC*.. Retrieved April 22, 2023, from <https://sevillafc.es/actualidad/noticias/informe-brand-finance-2021>
- Sevilla F.C. (s.f.). *Estructura Organizativa | Sevilla FC*.. Retrieved April 22, 2023, from <https://sevillafc.es/el-club/la-entidad/estructura-organizativa>
- Sevilla FC. (7 de Julio de 2021). *Máster en Big Data aplicado al scouting en fútbol | Sevilla FC*.. Retrieved April 23, 2023, from <https://sevillafc.es/actualidad/noticias/master-big-data-sport-data-campus-ucam>
- Sevilla FC. (25 de Diciembre de 2022). *El año del liderazgo tecnológico | Sevilla FC*.. Retrieved April 22, 2023, from <https://sevillafc.es/actualidad/noticias/ano-confirmacion-liderazgo-tecnologico-sevilla-fc>
- Sevilla FC. (23 de Marzo de 2023). *El Sevilla FC y DruID colaboran en el proyecto AiFans | Sevilla FC*.. Retrieved April 22, 2023, from <https://www.sevillafc.es/actualidad/noticias/proyecto-aifans-marzo-2023>

- Soria Polo, C. (2021). *idUS - Diseño y aplicación de técnicas de machine learning para optimizar el Scouting en clubes de fútbol*.. Retrieved July 18, 2023, from <https://idus.us.es/handle/11441/125973>
- Stats Perform. (s.f.). *Sports Data - Sports AI, Technology, Data Feeds | Stats Perform*.. Retrieved June 16, 2023, from <https://www.statsperform.com/>
- Superprof. (s.f.). *Saberlo todo sobre los percentiles*. Retrieved September 19, 2023, from <https://www.superprof.es/apuntes/escolar/matematicas/estadistica/descriptiva/percentiles.html>
- Tesla. (s.f.). *AI & Robotics | Tesla*.. Retrieved April 28, 2023, from <https://www.tesla.com/AI>
- The Coaches' Voice. (s.f.). *Análisis - Transiciones defensivas*. Retrieved June 12, 2023, from <https://es.coachesvoice.com/cv/transiciones-defensivas-explicacion-tuchel-barcelona/#:~:text=Las%20transiciones%20defensivas%20hacen%20referencia,equipo%20de%20atacar%20a%20defender>.
- The Coaches' Voice. (s.f.). *Conceptos - Transiciones ofensivas*. Retrieved August 12, 2023, from <https://es.coachesvoice.com/cv/transiciones-klopp-mourinho-simeone/>
- Transfermarkt. (s.f.). *Período de fichajes de verano | Transfermarkt*.. Retrieved September 17, 2023, from https://www.transfermarkt.com.ar/transfer-statistik/transferfenster/statistik/stat/plus/0?art=s&kontinent_id=6&option=2
- Transfermarkt. (s.f.). *Sevilla FC - Plantilla detallada 22/23*. Retrieved September 9, 2023, from https://www.transfermarkt.es/sevilla-fc/kader/verein/368/plus/0/galerie/0?saision_id=2022
- UEFA. (31 de Agosto de 2016). *Requisitos de las plantillas para la UEFA Champions League*. Retrieved September 21, 2023, from <https://es.uefa.com/uefachampionsleague/news/0230-0e94d96fa719-b31973ed3c2d-1000--requisitos-de-las-plantillas-para-la-uefa-champions-league/>
- Unisport Management School. (29 de Noviembre de 2022). *¿Cuánto cobra un ojeador de fútbol? ¡Te lo contamos! | Unisport*.. Retrieved May 16, 2023, from <https://unisport.es/cuanto-cobra-ojeador-futbol/>
- Whitmore, J. (27 de Junio de 2021). *¿Qué son los goles esperados (xG)? | The Analyst*.. Retrieved July 6, 2023, from <https://theanalyst.com/eu/2021/06/que-son-los-goles-esperados-xg/>
- Wyscout. (s.f.). *La Plataforma Para Clubs Y Selecciones Nacionales - Wyscout*.. Retrieved May 16, 2023, from <https://wyscout.com/es/clubs-y-selecciones/>