

Trabajo fin de grado

Desarrollo de un modelo predictivo de apoyo a la toma de decisiones en fútbol mediante inteligencia artificial



Laura Vázquez García

Escuela Politécnica Superior
Universidad Autónoma de Madrid
C/ Francisco Tomás y Valiente nº 11

**UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR**



Grado en Ingeniería Informática

TRABAJO FIN DE GRADO

**Desarrollo de un modelo predictivo de apoyo a la
toma de decisiones en fútbol mediante
inteligencia artificial**

**Autor: Laura Vázquez García
Tutor: Alejandro Bellogín Kouki**

junio 2026

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución comunicación pública y transformación de esta obra sin contar con la autorización de los titulares de la propiedad intelectual.

La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (*arts. 270 y sgts. del Código Penal*).

DERECHOS RESERVADOS

© 20 de Junio de 2026 por UNIVERSIDAD AUTÓNOMA DE MADRID

Francisco Tomás y Valiente, n.º 1

Madrid, 28049

Spain

Laura Vázquez García

Desarrollo de un modelo predictivo de apoyo a la toma de decisiones en fútbol mediante inteligencia artificial

Laura Vázquez García

C\ Francisco Tomás y Valiente N.º 11

IMPRESO EN ESPAÑA – PRINTED IN SPAIN

AGRADECIMIENTOS

Este trabajo pone fin a una etapa muy importante de mi vida. Han sido años de aprendizaje, esfuerzo y muchas experiencias compartidas. Por eso, antes de cerrar esta etapa, me gustaría agradecer a quienes me han acompañado durante este camino.

En primer lugar, quiero agradecer a mi tutor, Alejandro, por la paciencia, la confianza y el apoyo que me ha dado durante todo este proceso. Gracias por haberme animado a seguir adelante en los momentos en los que me estancaba o me bloqueaba, y por haber sido también uno de esos profesores que consiguen despertar interés más allá de una asignatura.

A mi familia, porque sabéis mejor que nadie que este camino no ha sido fácil. Gracias por apoyarme siempre. En especial, a mi hermana Marta, simplemente por ser mi hermana y por ocupar ese lugar que nadie más puede ocupar. Y a mi abuelo, por transmitirme esa pasión por el fútbol que, de una forma u otra, también ha acabado formando parte de este trabajo.

A los amigos que me ha regalado esta etapa universitaria. A Silvia, por haber sido tan buena compañera en prácticas, en clase y fuera de ella. A Sofía, por entenderme como pocas personas y por ser una amiga imprescindible en esta etapa y en mi vida. Y a Lola, porque además de compartir clase y casa, he tenido la suerte de crecer a su lado durante estos años.

También quiero agradecer a Helena, por haberme acompañado especialmente en estas últimas semanas y hacer que el final de este camino se sintiera un poco más ligero. Ojalá sigas acompañándome en muchos caminos más.

Por último, y no menos importante, a quienes son la razón de este TFG: mis compañeras de equipo, mis Sangas. Gracias por haber sido hogar durante estos últimos cuatro años y refugio en los momentos menos buenos. Este trabajo nace también de todo lo que hemos compartido juntas.

Y sobre todo gracias a Elvis, mi entrenador, por facilitarme los datos y hacer posible que este trabajo pudiera llevarse a cabo, pero también por haber confiado siempre en mí, tanto dentro como fuera del campo, y por todo lo que me ha enseñado durante este tiempo. Gracias por ayudarme a crecer como jugadora y como persona. Sin la pasión por este deporte que me has sabido contagiar, este proyecto no habría tenido el mismo sentido.

A todos vosotros, gracias por haber formado parte de este camino. Este trabajo lleva mi nombre, pero también un poco de cada una de las personas que me han acompañado durante estos años. Todo lo aprendido, dentro y fuera de la universidad, me lo llevo conmigo para lo que venga a partir de ahora.

RESUMEN

La creciente disponibilidad de datos deportivos ha abierto nuevas oportunidades para apoyar la toma de decisiones en fútbol. Este trabajo presenta el desarrollo de un modelo predictivo basado en inteligencia artificial para asistir al entrenador en el análisis de la carga física y la intensidad en competición. El estudio se basa en datos recogidos de un equipo senior femenino de fútbol 7 durante sesiones de entrenamiento y partidos.

El enfoque propuesto combina métricas de carga externa obtenidas mediante GPS y datos subjetivos de percepción del esfuerzo (RPE). Tras el preprocesado de los datos, se utilizan ventanas temporales de entrenamientos previos para representar la carga reciente de cada jugadora antes de un partido. La tarea predictiva se formula como un problema de clasificación, cuyo objetivo es estimar la zona de intensidad esperada de cada jugadora en competición.

Se evalúan distintos modelos de clasificación y configuraciones de ventanas temporales mediante una estrategia de validación que respeta el orden cronológico de los datos. El modelo seleccionado se elige en función de su rendimiento predictivo y estabilidad. Además, se analiza la relación entre la carga objetiva y la percepción subjetiva del esfuerzo, con el fin de obtener información adicional sobre la respuesta individual de las jugadoras.

Los resultados se integran en una aplicación orientada al entrenador, que permite consultar predicciones, indicadores de carga y perfiles de jugadoras. El sistema se plantea como una herramienta de apoyo a la toma de decisiones para facilitar la planificación del entrenamiento y el seguimiento individual de las jugadoras.

PALABRAS CLAVE

Inteligencia artificial, modelo predictivo, analítica deportiva, fútbol, GPS, RPE, monitorización de carga, clasificación, apoyo a la toma de decisiones

ABSTRACT

The increasing availability of sports data has created new opportunities to support decision-making in football. This work presents the development of an artificial intelligence-based predictive model to assist coaches in the analysis of physical load and match intensity. The study is based on data collected from a senior women's seven-a-side football team during training sessions and matches.

The approach combines GPS-derived external load metrics and subjective perceived exertion data (RPE). After data preprocessing, temporal windows of previous training sessions are used to represent each player's recent workload before a match. The prediction task is formulated as a classification problem, aiming to estimate the expected intensity zone of each player in competition.

Different classification models and temporal window configurations are evaluated using a time-aware validation strategy. The selected model is chosen according to its predictive performance and stability. Additionally, the relationship between objective workload and perceived exertion is analysed to provide further insight into individual player responses.

The results are integrated into an application for coaches, enabling them to access predictions, workload indicators and player profiles. The system is intended as a decision-support tool to facilitate training planning and player monitoring.

KEYWORDS

Artificial intelligence, predictive model, football analytics, GPS, RPE, workload monitoring, classification, decision support

ÍNDICE

1	Introducción	1
1.1	Motivación	1
1.2	Objetivos	2
1.3	Estructura	2
2	Estado del arte	3
2.1	Analítica deportiva y aprendizaje automático en fútbol	3
2.2	Monitorización de la carga: GPS y RPE	4
2.3	Modelos de clasificación, evaluación, validación e interpretación	6
2.3.1	Modelos de clasificación	6
2.3.2	Métricas de evaluación en clasificación y validación temporal	9
2.3.3	Interpretabilidad del modelo	11
2.4	Tecnologías para el desarrollo de aplicaciones de análisis deportivo	12
3	Diseño e implementación	13
3.1	Arquitectura general del sistema	13
3.1.1	Capa de presentación	14
3.1.2	Capa de aplicación	14
3.1.3	Capa de datos	15
3.2	Diseño	16
3.2.1	Requisitos del sistema	16
3.2.2	Diseño de datos	17
3.2.3	Diseño de procesos principales	19
3.2.4	Diseño de la interfaz de usuario	20
3.3	Implementación y desarrollo	21
3.3.1	Tecnologías utilizadas	21
3.3.2	Implementación de la ingesta y almacenamiento de datos	22
3.3.3	Implementación del backend y servicios principales	23
3.3.4	Integración del modelo predictivo	24
4	Pruebas y resultados	25
4.1	Metodología experimental	25
4.1.1	Preprocesado de los datos	25
4.1.2	Construcción de ventanas y planteamiento del análisis	26

4.2	Análisis exploratorio de datos	27
4.2.1	Distribución de variables	27
4.2.2	Correlación entre métricas	28
4.3	Selección de variables de entrada	28
4.4	Comparación de modelos y ventanas temporales	30
4.4.1	Configuración experimental y comparación global	31
4.4.2	Análisis de errores y comportamiento de las predicciones	33
4.4.3	Selección del modelo final	34
4.5	Interpretación del modelo	35
4.6	Análisis del desajuste entre carga objetiva y percepción del esfuerzo	36
4.6.1	Construcción del índice de desajuste	36
4.6.2	Resultados y aplicación práctica	37
5	Conclusiones y trabajo futuro	39
5.1	Conclusiones	39
5.2	Trabajo futuro	40
	Bibliografía	43
	Acrónimos	45
	Apéndices	47
A	Diseño e implementación	49
B	Análisis exploratorio de datos	59
C	Comparación de modelos y ventanas temporales	65

LISTAS

Lista de ecuaciones

2.1	Cálculo del s-RPE	5
2.2a	Cálculo de la accuracy	9
2.2b	Cálculo de la precisión	9
2.2c	Cálculo del recall	9
2.3	Cálculo del F1-score por clase	9
2.4	Cálculo del Macro-F1	9
2.5	Cálculo de la balanced accuracy	9
2.6	Cálculo conceptual de la importancia por permutación	11
4.1	Definición de ventana temporal	26
4.2	Cálculo de variables agregadas	26

Lista de figuras

2.1	Funcionamiento conceptual de la regresión logística	7
2.2	Comparación entre bagging y boosting	7
2.3	Frontera de decisión en SVM	8
2.4	Validación temporal mediante TimeSeriesSplit	10
3.1	Arquitectura general simplificada del sistema	14
3.2	Diagrama entidad-relación	18
4.1	Distribución del target <i>Int. de la sesión</i> en GAME	27
4.2	Comparativa de subconjuntos	30
4.3	Evolución del rendimiento	32
4.4	Heatmap de Macro-F1	32
4.5	Matriz de confusión normalizada	33
4.6	Distribución de clases	34
4.7	Desajuste medio por jugadora	38
A.1	Arquitectura general del sistema	50
A.2	Diagrama de secuencia del proceso de carga y procesamiento de una sesión GPS. ...	51
A.3	Diagrama de secuencia del proceso de actualización periódica de partidos	52

A.4	Diagrama de secuencia del proceso de consulta de resultados en el dashboard	52
A.5	Capturas del dashboard principal de la aplicación	53
A.6	Pantalla de carga de sesiones	54
A.7	Listado de jugadoras de la aplicación	55
A.8	Perfil individual de jugadora	56
A.9	Pantalla de partidos de la aplicación	57
A.10	Detalle del próximo partido	58
B.1	Distribución de las métricas físicas (histogramas y KDE)	61
B.2	Distribución de las métricas físicas mediante diagramas de caja	62
B.3	Heatmap de correlación global	63
C.1	Ranking de variables	66

Lista de tablas

4.1	Mejores resultados experimentales	31
4.2	Interpretación del índice	37
4.3	Resumen por jugadora	38
B.1	Descripción de los atributos del dataset	60
C.1	Resultados de modelos	65

INTRODUCCIÓN

El fútbol es un deporte complejo en el que la toma de decisiones depende de factores físicos, tácticos y contextuales. En este escenario, el uso de datos y técnicas de inteligencia artificial permite complementar la experiencia del entrenador con información objetiva sobre la carga de entrenamiento, la exigencia competitiva y la evolución de las jugadoras. Este trabajo se centra en el desarrollo de un modelo predictivo capaz de estimar la zona de intensidad esperada en competición a partir de datos históricos de entrenamientos y partidos.

1.1. Motivación

La motivación principal de este trabajo surge de la posibilidad de aplicar técnicas de inteligencia artificial a un caso real de fútbol femenino, utilizando datos registrados durante entrenamientos y partidos para generar información útil para el cuerpo técnico.

Aunque las métricas *Global Positioning System* (GPS) permiten medir de forma objetiva la carga física de las jugadoras, estos datos no siempre son fáciles de interpretar de manera aislada ni de transformar en decisiones prácticas. Por ello, se plantea el desarrollo de un modelo predictivo que aproveche el histórico disponible para estimar la zona de intensidad esperada en competición.

Esta predicción puede servir como apoyo para planificar la carga semanal, anticipar escenarios de mayor exigencia física y realizar un seguimiento más individualizado de las jugadoras. Además, el trabajo busca acercar este tipo de herramientas a un entorno deportivo no profesional, donde los recursos suelen ser más limitados. El objetivo no es sustituir el criterio del entrenador, sino ofrecer una herramienta complementaria que transforme los datos recogidos durante la temporada en información interpretable y útil para la toma de decisiones.

1.2. Objetivos

El objetivo principal de este Trabajo de Fin de Grado es desarrollar y evaluar un modelo predictivo de apoyo a la toma de decisiones en fútbol mediante técnicas de inteligencia artificial. A partir de datos históricos de entrenamientos y partidos, el modelo estima la zona de intensidad esperada en competición, ofreciendo una predicción interpretable para el entrenador.

Para ello, se integran, limpian y transforman datos procedentes de distintas fuentes, como métricas GPS, información contextual de los partidos y valores de percepción subjetiva del esfuerzo. A partir de estos datos se construyen variables agregadas mediante ventanas temporales de entrenamientos previos, con el fin de representar la carga reciente de cada jugadora antes de un partido.

El problema se formula como una tarea de clasificación por zonas de intensidad, favoreciendo una salida más comprensible y útil para el cuerpo técnico. Finalmente, se comparan distintos modelos, subconjuntos de variables y tamaños de ventana temporal para seleccionar la configuración más adecuada. Además, se analiza la interpretabilidad del modelo final y el desajuste entre la carga objetiva registrada mediante GPS y la percepción subjetiva del esfuerzo indicada mediante *Rating of Perceived Exertion* (RPE). Los resultados se integran en una aplicación orientada al entrenador, que permite consultar predicciones, indicadores de carga y perfiles individuales de forma accesible.

1.3. Estructura

El documento se organiza en cinco capítulos principales, además de los anexos complementarios:

Capítulo 1. Introducción. Presenta el contexto, la motivación, los objetivos y la estructura del trabajo.

Capítulo 2. Estado del arte. Revisa los fundamentos de analítica deportiva, monitorización de carga, aprendizaje automático y tecnologías utilizadas.

Capítulo 3. Diseño e implementación. Describe la arquitectura del sistema, el diseño de datos, los procesos principales, la interfaz de usuario, las tecnologías utilizadas y la integración del modelo predictivo.

Capítulo 4. Pruebas y resultados. Presenta la metodología experimental, el análisis exploratorio de los datos, la selección de variables, la comparación de modelos y ventanas temporales, la interpretación del modelo final y el análisis del desajuste entre carga objetiva y percepción subjetiva del esfuerzo.

Capítulo 5. Conclusiones y trabajo futuro. Resume las conclusiones principales, las limitaciones del sistema y posibles mejoras futuras.

ESTADO DEL ARTE

Este capítulo presenta los fundamentos necesarios para contextualizar el sistema desarrollado. Se revisa el uso de la analítica deportiva y el aprendizaje automático en fútbol, la monitorización de la carga mediante métricas GPS y RPE, y las técnicas de clasificación, validación temporal e interpretabilidad utilizadas para estimar zonas de intensidad y apoyar la toma de decisiones del entrenador.

2.1. Analítica deportiva y aprendizaje automático en fútbol

La analítica deportiva puede entenderse como el uso sistemático de datos para describir, evaluar y apoyar la toma de decisiones en contextos deportivos. En el fútbol, su desarrollo se ha visto impulsado por la disponibilidad creciente de información procedente de partidos, entrenamientos, eventos de juego y sistemas de seguimiento. Estos datos permiten complementar la observación tradicional del cuerpo técnico con indicadores objetivos que facilitan el análisis del rendimiento desde una perspectiva más estructurada [1, 2].

El análisis del rendimiento en fútbol presenta una dificultad especial debido a la naturaleza dinámica y contextual del juego. A diferencia de otros deportes con acciones más discretas o repetitivas, el fútbol se caracteriza por la interacción continua entre jugadoras, la influencia del rival, el marcador, la localización del partido, la posición ocupada en el campo y las decisiones tácticas adoptadas durante el encuentro. Por ello, los datos no deben interpretarse de forma aislada, sino dentro del contexto en el que se producen [1–3].

En este escenario, el aprendizaje automático constituye una herramienta útil para identificar relaciones entre variables y generar predicciones a partir de datos históricos. En el ámbito del fútbol, estas técnicas se han aplicado a problemas diversos, como la predicción del resultado de partidos, la valoración de acciones individuales, la estimación del rendimiento de jugadores o la clasificación de situaciones de juego.

Dentro de esta línea, algunos estudios se han centrado en valorar acciones individuales más allá de los goles, considerando la contribución de pases, conducciones u otras acciones al rendimiento

del equipo [4]. Otros enfoques han propuesto sistemas de evaluación de jugadores a partir de datos de eventos y rendimiento, incorporando información sobre el rol del jugador y su influencia dentro del equipo [5]. También existen aproximaciones orientadas a la predicción de resultados de partidos mediante algoritmos de clasificación supervisada y variables históricas o contextuales [6].

En el caso del fútbol femenino, el uso de analítica deportiva y aprendizaje automático resulta especialmente relevante debido al crecimiento reciente de la competición y al aumento progresivo de la disponibilidad de datos. Algunos trabajos recientes han utilizado datos espaciotemporales y de comportamiento en fútbol femenino para construir modelos predictivos asociados al resultado competitivo o al comportamiento del equipo en momentos concretos del partido [7]. Estos estudios muestran que las técnicas de aprendizaje automático pueden aportar información útil en contextos femeninos, aunque también ponen de manifiesto la necesidad de adaptar los análisis al nivel competitivo, a la disponibilidad de datos y a las características concretas de cada equipo.

Un aspecto especialmente importante en aplicaciones deportivas es la interpretabilidad de los resultados. En un entorno aplicado, no basta con que un modelo produzca una predicción numérica; dicha predicción debe poder explicarse de forma clara al usuario final. En este sentido, trabajos recientes destacan la necesidad de transformar las salidas de los modelos en información comprensible para entrenadores, analistas y miembros del cuerpo técnico [8]. Esta idea resulta fundamental en sistemas de apoyo a la decisión, donde el objetivo no es sustituir el criterio del entrenador, sino proporcionarle información adicional que facilite la planificación y el análisis del equipo.

El presente trabajo se enmarca dentro de esta línea de analítica deportiva aplicada al fútbol, aunque con un objetivo específico. En lugar de predecir directamente el resultado de un partido o valorar acciones técnicas y tácticas concretas, se propone utilizar la información procedente de los entrenamientos previos para estimar la zona de intensidad esperada en el siguiente partido. De esta forma, el aprendizaje automático se aplica como una herramienta de apoyo al entrenador, transformando datos históricos de carga en una salida discreta e interpretable.

2.2. Monitorización de la carga: GPS y RPE

La monitorización de la carga de entrenamiento y competición es una práctica habitual en fútbol, ya que permite cuantificar el trabajo realizado por las jugadoras y analizar su respuesta ante dicho esfuerzo. Esta información resulta útil para apoyar decisiones relacionadas con la planificación de sesiones, la recuperación, la disponibilidad de las deportistas y la gestión de la carga a lo largo de la temporada [9]. En este sentido, distintos trabajos han destacado la importancia de controlar la carga para comprender mejor la fatiga y adaptar el entrenamiento al estado de las deportistas [10, 11].

Habitualmente se diferencia entre carga externa e interna [12]. La carga externa hace referencia al trabajo físico realizado por la deportista durante una sesión o partido, independientemente de su res-

puesta fisiológica o perceptiva. En fútbol, esta carga se mide habitualmente mediante dispositivos GPS o sistemas de seguimiento, que permiten registrar variables como la distancia recorrida, la velocidad máxima, los esfuerzos a alta intensidad, las aceleraciones y las desaceleraciones [13]. Estas variables permiten caracterizar de forma objetiva la demanda física de entrenamientos y partidos, aunque su interpretación debe realizarse teniendo en cuenta el contexto deportivo y las características de cada sesión [14].

Por otro lado, la carga interna representa la respuesta individual de la deportista ante el esfuerzo realizado. Esta respuesta puede evaluarse mediante variables fisiológicas, como la frecuencia cardíaca, o mediante escalas subjetivas como el *Rating of Perceived Exertion*. El RPE permite recoger de forma sencilla la percepción de esfuerzo de la jugadora, por lo que resulta especialmente útil en contextos deportivos donde se busca una medida rápida, económica y aplicable de forma sistemática [15]. En fútbol, el uso del RPE ha sido propuesto como una herramienta práctica para cuantificar la carga interna asociada al entrenamiento [16].

Una variante ampliamente utilizada es el *Session Rating of Perceived Exertion* (s-RPE), que combina la percepción subjetiva del esfuerzo con la duración de la sesión. Este método permite estimar la carga interna total de una sesión a partir de una medida perceptiva sencilla [16, 17]. Su cálculo se define como:

$$s\text{-}RPE = RPE \times t_{\text{sesión}} \quad (2.1)$$

donde RPE representa la percepción subjetiva del esfuerzo reportada por la jugadora y $t_{\text{sesión}}$ la duración total de la sesión.

De esta forma, el s-RPE incorpora tanto la intensidad percibida como el tiempo total de exposición al esfuerzo. En fútbol femenino, el uso de métodos basados en RPE también ha sido estudiado como una forma práctica de cuantificar la carga interna en contextos de entrenamiento [18].

La combinación de carga externa (GPS) e interna (RPE) permite obtener una visión más completa del estado de la jugadora, ya que deportistas con cargas objetivas similares pueden percibir el esfuerzo de forma diferente debido a factores como la fatiga, la recuperación o la adaptación al entrenamiento [12]. Analizar conjuntamente ambas medidas permite detectar posibles discrepancias entre la carga registrada y la percibida, aportando información útil para interpretar la respuesta individual de cada jugadora [9, 10].

En este trabajo, esta relación se estudia mediante un análisis de desajuste entre la carga objetiva obtenida con métricas GPS y la carga subjetiva medida mediante RPE. Además, se considera la carga acumulada antes de cada partido mediante ventanas temporales construidas a partir de los entrenamientos previos [19, 20]. Estas ventanas permiten generar variables agregadas que se utilizan como entrada del modelo predictivo.

De este modo, las métricas GPS y las variables derivadas del RPE aportan información complementaria tanto para caracterizar el estado reciente de las jugadoras como para analizar posibles diferencias entre la carga objetiva y la percibida. La elección del tamaño de ventana y la construcción del índice de desajuste se justifican experimentalmente en el Capítulo 4.

2.3. Modelos de clasificación, evaluación, validación e interpretación

El aprendizaje supervisado permite construir modelos capaces de aprender relaciones entre un conjunto de variables de entrada y una variable objetivo conocida. En función del tipo de salida que se desea estimar, el problema puede formularse como una tarea de regresión, si la variable objetivo es continua, o como una tarea de clasificación, si se desea predecir una categoría discreta [21].

En el contexto de este trabajo, aunque inicialmente se consideró la predicción de la intensidad del partido como un valor continuo, finalmente se optó por un enfoque de clasificación. Esta decisión permite transformar la intensidad en zonas discretas, facilitando la interpretación de los resultados por parte del cuerpo técnico. En una herramienta orientada a entrenadores, una predicción expresada como una categoría o nivel de exigencia resulta más comprensible y accionable que un valor numérico aislado.

2.3.1. Modelos de clasificación

La clasificación resulta especialmente útil en este trabajo porque permite transformar una predicción continua de intensidad en categorías operativas para el análisis deportivo. A partir de estas categorías es posible identificar con antelación si un partido se espera dentro de una zona de baja, moderada, alta o muy alta exigencia física, facilitando la planificación de cargas y la interpretación de los resultados.

Para abordar este problema se seleccionaron modelos pertenecientes a familias diferentes, con el objetivo de comparar enfoques que presentan características complementarias. En concreto, se evaluaron modelos lineales, modelos basados en árboles, modelos basados en márgenes máximos y clasificadores de referencia o *baseline*.

La regresión logística se incluye como representante de los modelos lineales y como referencia por su simplicidad e interpretabilidad [21]. Su funcionamiento se basa en estimar una probabilidad de pertenencia a una clase a partir de una combinación lineal de las variables de entrada. En su formulación más simple, aplicada a clasificación binaria, el modelo transforma dicha combinación en una probabilidad mediante una función sigmoide, como se muestra en la Figura 2.1.

Aunque dicha figura representa el caso binario de la regresión logística, este modelo también puede

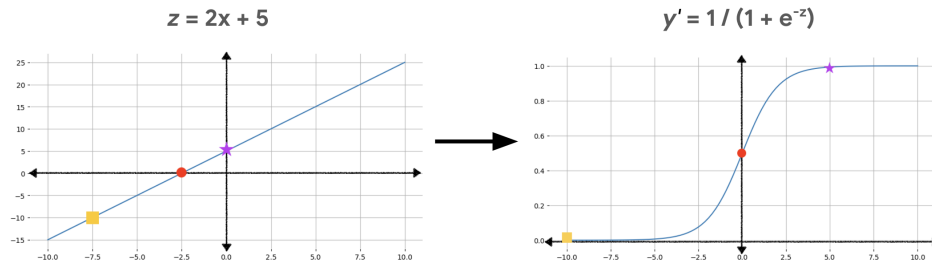


Figura 2.1: Transformación de una salida lineal en una probabilidad mediante una función sigmoide en regresión logística [22].

extenderse a problemas multiclase. En ese caso, se estima la probabilidad de pertenencia a cada una de las clases posibles y se asigna finalmente aquella con mayor probabilidad. En el contexto de este trabajo, esto equivale a estimar la probabilidad de que una observación pertenezca a cada zona de intensidad definida.

Random Forest representa los métodos basados en conjuntos de árboles de decisión. Este modelo se basa en una estrategia de *bagging*, en la que se entrenan múltiples árboles de forma independiente sobre diferentes subconjuntos de los datos. En problemas de clasificación, las predicciones de estos árboles se combinan normalmente mediante votación para obtener la clase final. Al no depender de un único árbol, Random Forest suele ofrecer una mayor capacidad de generalización y permite capturar relaciones no lineales entre variables sin requerir supuestos previos sobre la forma de los datos [23].

XGBoost también se basa en árboles de decisión, pero utiliza una estrategia diferente. En este caso, el entrenamiento sigue un enfoque de *boosting*, donde los árboles se construyen de forma secuencial. Cada nuevo árbol intenta corregir los errores cometidos por los anteriores, ajustando progresivamente el modelo. Esta estrategia permite obtener modelos con elevada capacidad predictiva, especialmente en problemas con datos tabulares e interacciones complejas entre variables [24].

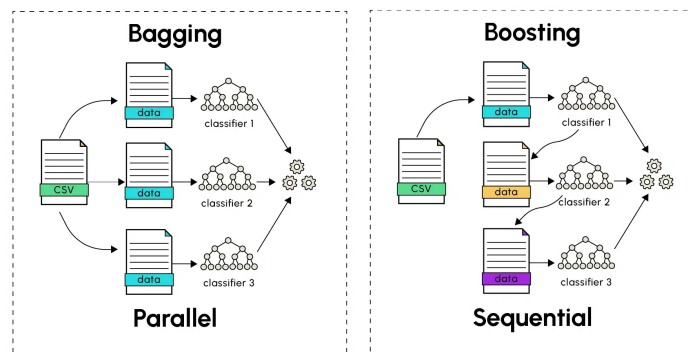


Figura 2.2: Diferencia conceptual entre métodos de ensamblado basados en *bagging* y métodos basados en *boosting* [25].

La Figura 2.2 muestra de forma conceptual la diferencia entre ambos enfoques. Mientras que el *bagging*, asociado en este trabajo a Random Forest, combina modelos entrenados de forma independiente, el *boosting*, asociado a XGBoost, construye los modelos de manera secuencial para mejorar progresivamente las predicciones.

Las máquinas de vectores soporte (*Support Vector Machine (SVM)*) constituyen una alternativa basada en la construcción de fronteras de decisión. Su objetivo es separar las clases maximizando el margen entre la frontera y los ejemplos más cercanos, denominados vectores soporte [26]. Este planteamiento permite construir una separación lo más robusta posible entre clases, aportando una perspectiva diferente respecto a los modelos basados en árboles o en relaciones lineales.

La Figura 2.3 muestra de forma conceptual el funcionamiento de una SVM en un caso de clasificación binaria. Aunque en este trabajo el problema se plantea como una clasificación multiclase, el esquema permite visualizar la idea principal del método: encontrar una frontera de decisión que maximice el margen entre clases.

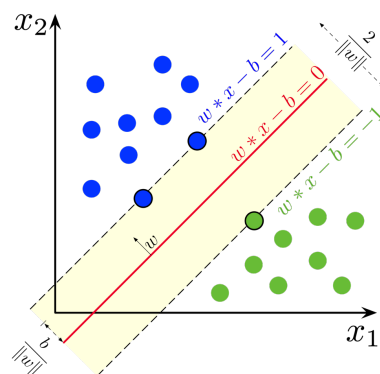


Figura 2.3: Representación conceptual de una máquina de vectores soporte [27].

Finalmente, se incluyen clasificadores **Dummy** como referencia experimental. Estos modelos no aprenden patrones de los datos, sino que realizan predicciones siguiendo estrategias simples. En este trabajo se consideran dos variantes: *most_frequent*, que predice siempre la clase más frecuente del conjunto de entrenamiento, y *stratified*, que genera predicciones aleatorias respetando la distribución de clases observada. Su utilidad consiste en establecer un nivel mínimo de rendimiento que los modelos entrenados deben superar para considerarse útiles.

La implementación experimental se realizó principalmente mediante la biblioteca *scikit-learn*, que proporciona herramientas para el entrenamiento, validación y evaluación de modelos de aprendizaje automático [28]. Además, XGBoost se implementó utilizando su librería específica debido a las funcionalidades adicionales que ofrece para este tipo de algoritmos.

2.3.2. Métricas de evaluación en clasificación y validación temporal

La evaluación de un modelo de clasificación requiere seleccionar métricas adecuadas al tipo de problema y a la distribución de las clases. En una tarea multiclase, la exactitud o *accuracy* permite medir la proporción global de predicciones correctas, pero puede resultar insuficiente si se desea evaluar el comportamiento del modelo de forma equilibrada entre todas las clases [29, 30].

Para evaluar el rendimiento de un clasificador se parte de los aciertos y errores cometidos por el modelo. A partir de ellos pueden definirse métricas básicas como la *accuracy*, la precisión y el *recall*:

$$Accuracy = \frac{N_{correctas}}{N_{total}} \quad (2.2a)$$

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \quad (2.2b)$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (2.2c)$$

donde $N_{correctas}$ representa el número de predicciones acertadas y N_{total} el número total de predicciones realizadas. Para una clase i , TP_i representa los verdaderos positivos, FP_i los falsos positivos y FN_i los falsos negativos. La precisión indica qué proporción de las predicciones asignadas a una clase son correctas, mientras que el *recall* mide qué proporción de los ejemplos reales de dicha clase han sido identificados correctamente por el modelo.

A partir de la precisión y el *recall*, el F1-score combina ambas medidas mediante la media armónica:

$$F1Score_i = 2 \cdot \frac{Precision_i \cdot Recall_i}{Precision_i + Recall_i} \quad (2.3)$$

En problemas multiclase, el Macro-F1 se obtiene calculando el F1-score de cada clase y promediando posteriormente los resultados:

$$Macro-F1 = \frac{1}{C} \sum_{i=1}^C F1Score_i \quad (2.4)$$

donde C representa el número total de clases. Esta métrica otorga el mismo peso a todas las categorías, independientemente del número de ejemplos de cada una. Por este motivo, resulta especialmente útil cuando se desea evaluar si el modelo mantiene un comportamiento equilibrado entre las distintas zonas de intensidad.

Otra métrica utilizada en problemas multiclase es la *balanced accuracy*, que calcula la media del *recall* obtenido en cada clase:

$$Balanced\ Accuracy = \frac{1}{C} \sum_{i=1}^C Recall_i \quad (2.5)$$

En este trabajo, el Macro-F1 se utiliza como métrica principal, ya que el objetivo es evaluar el rendimiento del modelo en todas las zonas de intensidad de forma equilibrada. De forma complementaria, se analizan la *balanced accuracy* y la *accuracy*, con el fin de obtener una visión más completa del comportamiento predictivo. En la implementación experimental, estas métricas se calculan mediante las herramientas disponibles en el módulo *sklearn.metrics* de scikit-learn [28, 31].

Además del valor medio de las métricas, también resulta importante estudiar la estabilidad del modelo. Para ello, se analiza la desviación estándar del Macro-F1 entre los distintos folds de validación. Una desviación estándar reducida indica que el modelo presenta un comportamiento más consistente entre particiones temporales, mientras que valores elevados reflejan una mayor sensibilidad al periodo concreto utilizado para evaluar.

La estrategia de validación también es un aspecto fundamental. En conjuntos de datos con componente temporal, una partición aleatoria puede producir una evaluación poco realista, ya que el modelo podría entrenarse con información posterior a los casos utilizados en prueba. En escenarios donde el orden temporal es relevante, resulta más adecuado emplear una validación que respete la secuencia cronológica de los datos [32].

En este sentido, scikit-learn proporciona la clase *TimeSeriesSplit*, diseñada para generar particiones de entrenamiento y prueba sobre datos ordenados temporalmente [33]. En este esquema, cada fold utiliza información pasada para entrenar el modelo y datos posteriores para evaluarlo, reproduciendo mejor un escenario real de predicción futura.

Este esquema puede representarse visualmente mediante una secuencia de particiones temporales en la que, en cada iteración, el conjunto de entrenamiento se construye a partir de observaciones pasadas y el conjunto de prueba utiliza observaciones posteriores. La Figura 2.4 muestra un ejemplo conceptual de este procedimiento.

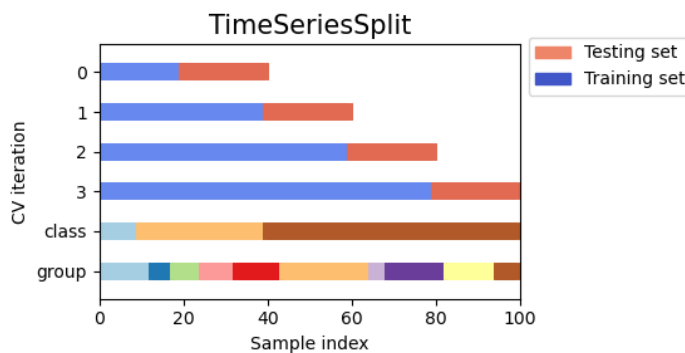


Figura 2.4: Ejemplo conceptual de validación temporal mediante *TimeSeriesSplit* [33].

Como se observa en la Figura 2.4, el conjunto de entrenamiento crece progresivamente en cada fold, mientras que el bloque de prueba se desplaza hacia adelante en el tiempo. Esta estrategia resulta adecuada cuando se desea simular un escenario realista de predicción futura, evitando utilizar información posterior durante el entrenamiento.

En el caso de este trabajo, la validación temporal resulta especialmente relevante porque los registros proceden de entrenamientos y partidos distribuidos a lo largo de la temporada. Además, varias observaciones pueden estar asociadas a un mismo partido, por lo que es necesario evitar que información de un mismo encuentro aparezca simultáneamente en entrenamiento y prueba. Este criterio reduce el riesgo de fuga de información y permite estimar de forma más realista el comportamiento del modelo ante partidos futuros.

2.3.3. Interpretabilidad del modelo

Además del rendimiento predictivo, en los sistemas de apoyo a la toma de decisiones es importante que los resultados del modelo puedan ser interpretados por el usuario final. En el contexto deportivo, una predicción resulta más útil si el cuerpo técnico puede comprender qué variables influyen en ella y cómo se relacionan con aspectos como la planificación, la carga de entrenamiento o el seguimiento individual de las jugadoras [34].

En este trabajo, la interpretabilidad se aborda mediante *permutation importance*, una técnica que permite estimar la relevancia de cada variable analizando cómo cambia el rendimiento del modelo cuando sus valores se permutan aleatoriamente [35]. Si al alterar una variable el rendimiento disminuye de forma notable, se interpreta que dicha variable aporta información relevante al modelo.

De forma conceptual, la importancia por permutación de una variable j puede expresarse como:

$$I_j = S_{\text{base}} - S_{\text{perm}(j)} \quad (2.6)$$

donde I_j representa la importancia de la variable j , S_{base} el rendimiento del modelo original y $S_{\text{perm}(j)}$ el rendimiento obtenido tras permutar dicha variable. En este trabajo, el rendimiento se mide mediante Macro-F1, por lo que una mayor caída en esta métrica indica una mayor contribución de la variable a la predicción de la zona de intensidad.

No obstante, esta técnica debe interpretarse con cautela, ya que no establece relaciones causales entre las variables de entrada y la intensidad posterior del partido. Su objetivo es aportar una explicación complementaria sobre qué dimensiones de la carga reciente utiliza el modelo, facilitando así que los resultados puedan ser comprendidos por el entrenador y utilizados como apoyo a la toma de decisiones.

2.4. Tecnologías para el desarrollo de aplicaciones de análisis deportivo

El desarrollo de herramientas de análisis deportivo se apoya en la combinación de datos procedentes de distintas fuentes y en tecnologías capaces de transformarlos en información útil para el cuerpo técnico. En este tipo de aplicaciones, el objetivo no es únicamente recopilar datos, sino presentarlos de forma clara, interpretable y adaptada a las necesidades del usuario final.

En el contexto del fútbol federado, existen plataformas que permiten consultar información competitiva de los partidos. En este trabajo, la Real Federación de Fútbol de Madrid (RFFM) y aplicaciones como MatchApp constituyen fuentes de información contextual, ya que permiten acceder a calendarios, resultados, clasificaciones o actas de los encuentros [36,37]. Esta información complementa los datos físicos registrados por las jugadoras y permite relacionar la carga observada con el contexto competitivo de cada partido.

Por otro lado, los dispositivos GPS se han convertido en una herramienta habitual para registrar la actividad física en entrenamientos y partidos. En este proyecto, las métricas principales proceden del sistema Oliver, que proporciona información sobre distancia recorrida, velocidad, intensidad, aceleraciones, desaceleraciones y otras variables asociadas a la carga externa [38]. Estos datos constituyen la base cuantitativa sobre la que se construyen los análisis posteriores.

Para el tratamiento de este tipo de información, Python ofrece un ecosistema ampliamente utilizado en análisis de datos y aprendizaje automático. Librerías como pandas y NumPy permiten trabajar con datos tabulares, aplicar transformaciones y calcular variables agregadas [39,40]. Asimismo, scikit-learn proporciona herramientas para el entrenamiento y evaluación de modelos predictivos, lo que facilita su aplicación en problemas de clasificación como el planteado en este trabajo [28].

Además del análisis de datos, una herramienta aplicada necesita presentar los resultados de forma accesible. En este sentido, tecnologías de desarrollo web o móvil como React Native y Expo permiten crear interfaces orientadas a la consulta rápida de indicadores, alertas y perfiles individuales [41,42]. Este aspecto es especialmente importante cuando el usuario final es un entrenador, ya que la información debe poder interpretarse sin requerir conocimientos técnicos avanzados.

En conjunto, las tecnologías utilizadas en aplicaciones de análisis deportivo permiten conectar fuentes de datos deportivas, procesamiento analítico y visualización de resultados. En este trabajo, estas herramientas se emplean con el objetivo de transformar datos GPS y contexto competitivo en información útil para apoyar la planificación y la toma de decisiones del entrenador.

DISEÑO E IMPLEMENTACIÓN

El sistema desarrollado tiene como finalidad servir como herramienta de apoyo a la toma de decisiones para entrenadores de fútbol. A partir de datos históricos de entrenamientos y partidos, el sistema permite analizar la carga reciente de las jugadoras, estimar la zona de intensidad esperada de un partido y estudiar posibles desajustes entre la carga objetiva registrada y la percepción subjetiva del esfuerzo indicada mediante RPE.

El usuario final de la aplicación es el entrenador, por lo que el diseño del sistema se ha orientado a la interpretación práctica de la información. Aunque internamente incluye procesos de preparación de datos, entrenamiento del modelo y generación de predicciones, la interfaz presenta los resultados mediante indicadores comprensibles que apoyan la planificación del entrenamiento, el control de la carga, la recuperación y el seguimiento de las jugadoras.

3.1. Arquitectura general del sistema

La arquitectura general del sistema se ha diseñado siguiendo una estructura por capas, con el objetivo de separar responsabilidades y facilitar tanto el mantenimiento como la evolución futura de la aplicación. Esta organización permite diferenciar claramente la interfaz utilizada por el entrenador, la lógica interna encargada del procesamiento y modelado de datos, y la capa responsable del almacenamiento e integración de la información.

La Figura 3.1 muestra una versión simplificada de la arquitectura propuesta, compuesta por tres capas principales: la capa de presentación, la capa de aplicación y la capa de datos. La capa de presentación contiene la interfaz utilizada por el usuario final; la capa de aplicación centraliza la comunicación con la interfaz y la lógica principal del sistema; y la capa de datos agrupa tanto las fuentes originales de información como la base de datos integrada.

Esta representación resume los bloques principales del sistema sin detallar todos los servicios internos del backend, con el fin de facilitar una visión global de la solución. La versión completa de la arquitectura, donde se muestran con mayor detalle los módulos internos, la relación entre ellos y las tecnologías asociadas a cada componente se incluye en la Figura A.1 del Anexo A. La descripción

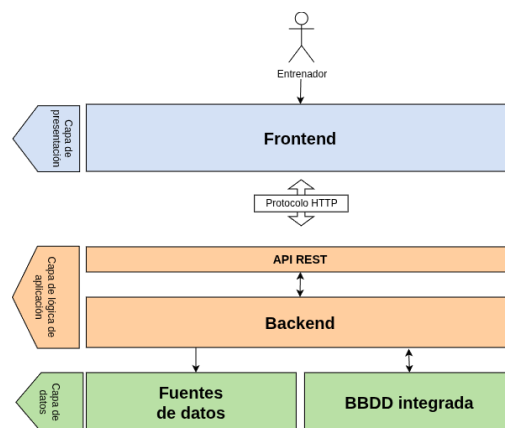


Figura 3.1: Vista simplificada de la arquitectura general del sistema.

detallada del stack tecnológico se desarrolla posteriormente en la Sección 3.3.

3.1.1. Capa de presentación

La capa de presentación corresponde a la interfaz utilizada por el entrenador para consultar la información generada por el sistema. Su objetivo principal es mostrar los resultados de forma clara, visual e interpretable, evitando que el usuario tenga que interactuar con detalles técnicos del modelo o del procesamiento interno.

Esta capa incluye las pantallas principales de la aplicación, como el dashboard general del equipo, la predicción del próximo partido, el perfil individual de cada jugadora, el análisis del desajuste entre carga objetiva y percepción subjetiva del esfuerzo y las alertas interpretables. Desde estas vistas, el entrenador puede consultar el estado del equipo, revisar posibles situaciones de carga elevada y analizar la evolución individual de las jugadoras.

La capa de presentación no realiza el procesamiento principal de los datos ni ejecuta el modelo predictivo. Su función es actuar como punto de interacción con el usuario, enviando peticiones al backend mediante la API REST y mostrando las respuestas recibidas de forma comprensible para el cuerpo técnico.

3.1.2. Capa de aplicación

La capa de aplicación centraliza la lógica principal del sistema. Su función es recibir las peticiones procedentes de la interfaz, coordinar los casos de uso, procesar los datos y ejecutar los módulos necesarios para generar los resultados que posteriormente se muestran al entrenador.

Dentro de esta capa se incluye la *Application Programming Interface* (API) *Representational State Transfer* (REST), que actúa como punto de entrada entre la aplicación cliente y el servidor. Esta API

expone los servicios del backend mediante endpoints *Hypertext Transfer Protocol* (HTTP), valida las solicitudes recibidas y devuelve respuestas estructuradas a la capa de presentación.

El backend se organiza en distintos servicios internos, cada uno con una responsabilidad concreta:

- **Servicios de aplicación:** coordinan los principales casos de uso del sistema, como la gestión de sesiones, la consulta de jugadoras y partidos, la generación de predicciones y la creación de alertas. Funcionan como intermediarios entre la API REST y el resto de módulos internos.
- **Servicio de ingesta de datos:** se encarga de incorporar información procedente de las fuentes externas del sistema. Incluye tanto la obtención de datos de partidos mediante *scraping* como la carga de archivos exportados desde el sistema GPS Oliver.
- **Lógica de negocio y analítica:** agrupa los procesos principales de tratamiento de datos, como la limpieza, validación e integración de la información, la vinculación de métricas GPS con el partido correspondiente, la construcción de ventanas temporales y el cálculo de variables agregadas.
- **Servicio de predicción:** aplica el modelo final seleccionado en la fase experimental. Este modelo utiliza como entrada las variables agregadas de los últimos cinco entrenamientos previos al partido y devuelve una de las cuatro zonas predicha.
- **Servicio de acceso a datos:** centraliza las operaciones de lectura y escritura sobre la base de datos, permitiendo que el resto de módulos accedan a la información de forma controlada y manteniendo una separación clara entre la lógica del sistema y la persistencia de datos.

Aunque no se representa como un bloque independiente en la arquitectura general simplificada, el backend contempla la ejecución de una tarea programada de actualización periódica. Esta tarea se encarga de obtener y actualizar la información contextual de partidos procedente de la página de la RFFM, como calendarios, resultados, rivales y goles. La incorporación de métricas GPS, en cambio, se realiza cuando el entrenador carga manualmente los archivos exportados desde Oliver. El funcionamiento detallado de estos flujos se describe posteriormente en el Apartado 3.2.3.

3.1.3. Capa de datos

La capa de datos agrupa tanto las fuentes originales de información como el almacenamiento persistente del sistema. En este proyecto se utilizan dos fuentes principales: los datos contextuales de partidos obtenidos mediante *scraping* desde la página de la RFFM y los archivos exportados desde el sistema GPS Oliver en formato *Comma-Separated Values* (CSV).

Los datos scrapeados aportan información del contexto competitivo, como la fecha del partido, el rival, la jornada, el resultado o los goles. Esta información se utiliza principalmente con fines de visualización dentro de la aplicación, permitiendo mostrar al entrenador datos de contexto y generar indicadores descriptivos. No obstante, estos datos no forman parte del modelado predictivo, que se basa en las métricas de carga procedentes del GPS. Por otro lado, los archivos procedentes del GPS Oliver contienen las métricas físicas registradas durante entrenamientos y partidos, así como información asociada al RPE, s-RPE e intensidad de sesión.

Un aspecto clave de esta capa es la integración entre ambas fuentes. Las métricas GPS correspondientes a partidos deben vincularse con el partido correspondiente almacenado en la base de datos, de forma que cada encuentro quede representado tanto por su contexto competitivo como por la carga física registrada. Esta integración permite construir un histórico coherente sobre el que se apoyan los módulos de análisis y predicción.

La información integrada se almacena en una base de datos que contiene las entidades necesarias para el funcionamiento del sistema, como jugadoras, sesiones, entrenamientos, partidos, métricas GPS, predicciones generadas y valores de desajuste entre carga objetiva y percepción subjetiva. Esta persistencia permite actualizar el histórico conforme se incorporan nuevas sesiones y consultar la información necesaria desde el backend.

3.2. Diseño

En este apartado se describe el diseño funcional del sistema a partir de los objetivos definidos y de la arquitectura general presentada anteriormente. Mientras que la arquitectura muestra la organización global de la aplicación en capas, el diseño del sistema concreta qué funcionalidades debe ofrecer, cómo se organizan los datos principales y cómo interactúan los distintos módulos en los procesos más relevantes.

3.2.1. Requisitos del sistema

A partir del análisis del problema y del usuario final, se definieron una serie de requisitos funcionales y no funcionales. Estos requisitos tienen como objetivo orientar el diseño de la aplicación hacia una herramienta útil para el cuerpo técnico, priorizando la interpretación de los resultados, la integración de datos y la actualización progresiva del histórico.

Requisitos funcionales

- RF-1.**— El sistema debe permitir registrar partidos obtenidos mediante *scraping* desde la página de la RFFM, almacenando al menos la fecha, el rival, el resultado y las goleadoras.
- RF-2.**— El sistema debe permitir importar archivos exportados desde el sistema GPS Oliver en formato CSV o Excel.
- RF-3.**— El sistema debe validar los datos importados, detectando registros incompletos, inconsistentes o duplicados antes de su almacenamiento.
- RF-4.**— El sistema debe asociar cada sesión GPS de tipo partido con el partido correspondiente almacenado en la base de datos.
- RF-5.**— El sistema debe almacenar de forma persistente la información relativa a jugadoras, sesiones, partidos y métricas GPS.
- RF-6.**— El sistema debe calcular variables agregadas utilizando los cinco entrenamientos previos a cada partido.

- RF-7.**— El sistema debe predecir el nivel de intensidad esperado del siguiente partido, clasificándolo en una de las categorías: baja, media, alta o máxima.
- RF-8.**— El sistema debe calcular indicadores de desajuste entre la carga objetiva registrada y la percepción subjetiva del esfuerzo mediante RPE.
- RF-9.**— El sistema debe mostrar al entrenador las predicciones generadas por el modelo junto con los indicadores asociados.
- RF-10.**— El sistema debe recalcular las variables agregadas y las predicciones cuando se incorporen nuevos datos al sistema.

Requisitos no funcionales

- RNF-1.**— La interfaz debe ser utilizable por usuarios sin conocimientos técnicos avanzados.
- RNF-2.**— La arquitectura debe mantener una separación clara entre presentación, lógica de negocio y acceso a datos.
- RNF-3.**— La información almacenada debe persistir entre ejecuciones del sistema.
- RNF-4.**— Las predicciones deben ir acompañadas de información que facilite su interpretación por parte del entrenador.
- RNF-5.**— El sistema debe permitir incorporar nuevas métricas, fuentes de datos y modelos predictivos sin modificar significativamente la arquitectura existente.
- RNF-6.**— El sistema debe garantizar la consistencia de los datos, evitando asociaciones erróneas entre sesiones, partidos y jugadoras.
- RNF-7.**— El tiempo de respuesta para la generación de una predicción debe ser inferior a cinco segundos.

3.2.2. Diseño de datos

El diseño de datos se ha planteado con el objetivo de integrar de forma coherente la información procedente de las distintas fuentes del sistema y permitir su uso en los módulos de visualización, análisis y predicción. Para ello, se define un modelo centrado en las jugadoras, las sesiones registradas mediante GPS, los partidos y los indicadores derivados calculados por el sistema.

Aunque el sistema distingue entre entrenamientos y partidos, ambos generan registros GPS con la misma estructura al ser exportados desde Oliver. Por este motivo, se define una entidad genérica `Sesion` para representar cualquier actividad registrada mediante GPS, diferenciando posteriormente si se trata de un entrenamiento o de un partido mediante el atributo `tipo_sesion`. A partir de estas sesiones se almacenan las métricas físicas en la entidad `Metricas_GPS`, asociada tanto a la jugadora como a la sesión correspondiente. Esta decisión permite gestionar de forma unificada toda la información procedente de los dispositivos GPS, evitando duplicar estructuras de almacenamiento y facilitando el mantenimiento del sistema.

La información competitiva específica se almacena en la entidad `Partido`, que recoge datos como fecha, jornada, rival, equipo local, equipo visitante, resultado y goles. Cuando existe una sesión GPS correspondiente a un partido, esta se vincula con el registro de partido asociado, permitiendo relacionar

la carga física registrada con el contexto competitivo del encuentro.

La entidad `Gol` almacena información sobre goles marcados y encajados. En el caso de los goles a favor puede incluir una referencia opcional a la jugadora que marca el gol siempre que sea posible identificarla dentro del sistema. En los goles en contra, o cuando la asociación no sea fiable, esta referencia puede quedar vacía.

Para representar la entrada utilizada por el modelo predictivo se introduce la entidad `Ventana_Temporal`. Esta almacena, para cada jugadora y partido, las métricas agregadas calculadas a partir de una ventana de entrenamientos previos. De forma general, el tamaño de la ventana viene determinado por el número N de sesiones anteriores consideradas. En la configuración final del sistema, este valor se fija en $N = 5$, de acuerdo con el análisis experimental realizado posteriormente. De esta forma, el sistema conserva los valores agregados utilizados en cada predicción sin necesidad de almacenar explícitamente todos los entrenamientos que componen la ventana.

Por último, la entidad `Prediccion` guarda la zona de intensidad estimada por el modelo, asociada al partido, a la jugadora y a la ventana temporal utilizada. La entidad `Desajuste_RPE` recoge los valores calculados de la comparación entre la carga objetiva registrada y la percepción subjetiva del esfuerzo indicada mediante RPE.

La Figura 3.2 muestra el modelo entidad-relación simplificado del sistema. En el diagrama se representan únicamente los atributos principales de cada entidad, omitiendo algunas métricas GPS específicas para mantener la claridad de la figura.

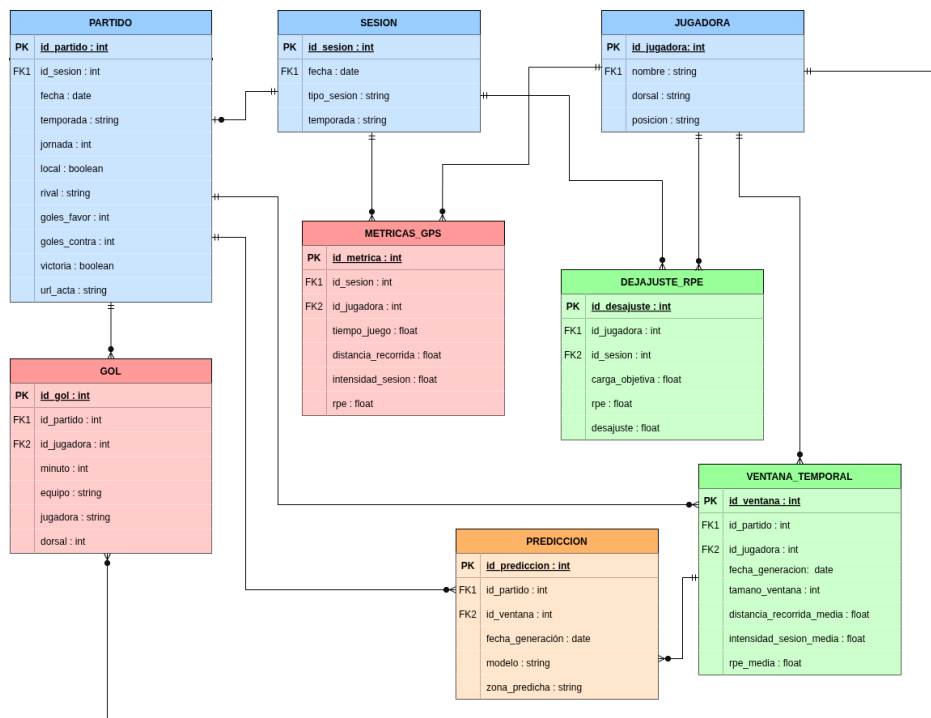


Figura 3.2: Diagrama entidad-relación de la base de datos del sistema.

Este diseño permite separar los datos originales procedentes de las fuentes externas de los datos derivados generados por el sistema. Las entidades `Sesion`, `Partido`, `Jugadora`, `Metricas_GPS` y `Go1` almacenan información de entrada, mientras que `Ventana_Temporal`, `Prediccion` y `Desajuste_RPE` almacenan resultados calculados a partir del histórico. De esta forma el sistema mantiene una estructura clara, evita duplicidades y permite consultar tanto los datos brutos como los indicadores utilizados para apoyar la toma de decisiones del entrenador.

3.2.3. Diseño de procesos principales

Una vez definido el modelo de datos, se describen los procesos principales que permiten transformar la información almacenada en resultados útiles para el entrenador. El sistema se organiza en torno a tres flujos: la incorporación de sesiones GPS, la actualización de indicadores y predicciones, y la consulta de resultados desde la aplicación.

El primer proceso corresponde a la incorporación de nuevas sesiones registradas mediante GPS. Cuando se carga un archivo exportado desde Oliver, el servicio de ingesta valida su estructura, aplica las reglas de limpieza necesarias y almacena la sesión junto con las métricas asociadas a cada jugadora. La incorporación de una nueva sesión actualiza el histórico disponible y puede modificar las ventanas temporales asociadas a próximos partidos, por lo que el sistema recalcula los indicadores dependientes de la carga reciente. Estas ventanas se utilizan como entrada del modelo predictivo, que estima la zona de intensidad del partido y almacena el resultado en la entidad `Prediccion`. Además, se calculan indicadores derivados, como el desajuste entre carga objetiva y RPE.

El segundo proceso corresponde a la actualización periódica del sistema. En este flujo se incorporan o actualizan los datos contextuales de partidos obtenidos mediante scraping y se vinculan con las sesiones GPS de tipo partido cuando existe correspondencia. Su finalidad es mantener actualizada la información de los encuentros disputados, incorporando nuevos partidos o modificando aquellos cuyos datos hayan cambiado desde la última actualización.

El tercer proceso corresponde a la consulta de resultados. El entrenador accede a la aplicación y solicita información como el dashboard general, la predicción del próximo partido o el perfil individual de jugadora. La interfaz se comunica con el backend mediante la API REST, y este recupera de la base de datos la información necesaria para mostrarla de forma estructurada y comprensible.

La interacción entre los principales componentes del sistema se resume en tres diagramas de secuencia. La Figura A.2 representa la carga y procesamiento de una sesión GPS, incluyendo la validación del archivo, el almacenamiento de métricas, el recálculo de ventanas temporales y la generación de predicciones e indicadores derivados. La Figura A.3 muestra la actualización periódica de partidos mediante scraping y su vinculación con sesiones GPS. Finalmente, la Figura A.4 recoge la consulta de resultados por parte del entrenador, que accede desde la aplicación a la información procesada por el

backend.

Este diseño refleja que el entrenador no interactúa directamente con el modelo ni con la base de datos, sino que consulta desde la aplicación información ya procesada. Asimismo, la actualización de ventanas temporales, indicadores y predicciones no implica el reentrenamiento automático del modelo, sino la aplicación del modelo previamente validado sobre los nuevos datos calculados.

3.2.4. Diseño de la interfaz de usuario

El diseño de la interfaz de usuario se ha planteado teniendo en cuenta que el destinatario principal del sistema es el entrenador. Por este motivo, la aplicación prioriza pantallas de consulta rápida, indicadores visuales y mensajes interpretables, evitando mostrar detalles técnicos del modelo predictivo o del procesamiento interno de los datos.

La interfaz se organiza en torno a varias pantallas principales. El **dashboard** actúa como vista inicial del sistema y ofrece una visión general del estado del equipo. En esta pantalla se muestran estadísticas generales del equipo, alertas relevantes como jugadoras en posible situación de riesgo y un resumen inicial de la plantilla. Desde esta vista, el entrenador puede acceder a información más detallada mediante enlaces a las páginas específicas de jugadoras y sus perfiles individuales. Las capturas correspondientes a esta pantalla se incluyen en la Figura A.5.

Una de las pantallas más importantes es la de **carga de sesiones**. Esta vista se diseña en torno a un calendario, con posibilidad de consultar la información por mes. Desde el calendario, el entrenador puede seleccionar un día concreto y subir una sesión GPS asociada a esa fecha. Además, los días que ya tienen una sesión registrada muestran un indicador visual, facilitando la identificación rápida de las sesiones disponibles y del histórico de carga incorporado al sistema. Véase la Figura A.6, donde se muestra una captura de esta pantalla.

La **página de jugadoras** permite consultar la plantilla completa y acceder al perfil individual de cada una. En el perfil de jugadora se muestran indicadores relacionados con su estado reciente, la evolución del RPE, las zonas de intensidad asociadas a partidos y entrenamientos, y el desajuste entre carga objetiva y percepción subjetiva del esfuerzo. Esta pantalla también incluye posibles alertas y recomendaciones orientativas, con el objetivo de facilitar el seguimiento individual y apoyar la toma de decisiones del entrenador. La pantalla de partidos y el detalle del próximo encuentro se muestran en las Figuras A.7 y A.8, respectivamente.

La **página de partidos** recoge el histórico de encuentros disputados junto con sus resultados. En la parte inicial se destaca el próximo partido, mostrando información relevante del encuentro, como el rival, la condición de local o visitante y elementos visuales como los escudos de los equipos. Desde esta pantalla se puede acceder al detalle de un partido. En el caso del próximo encuentro, se muestra información específica relacionada con la preparación del partido, incluyendo la zona de intensidad

estimada para cada jugadora. Para el resto de partidos, el detalle presenta un resumen del encuentro obtenido a partir del acta oficial. Además, en la parte inferior de la pantalla se incluye un acceso a una sección de estadísticas generales de la temporada, donde es posible seleccionar la temporada de interés y consultar diferentes indicadores agregados, como la distribución temporal de goles marcados y encajados o clasificaciones de máximas goleadoras diferenciando entre jugadoras del propio equipo y de los equipos rivales. Las capturas correspondientes a estas pantallas se incluyen en las Figuras A.9 y A.10.

En conjunto, el diseño de la interfaz busca transformar los resultados analíticos en información accionable para el entrenador. La organización por pantallas permite separar las tareas principales del sistema: revisar el estado general del equipo, incorporar nuevas sesiones, consultar la evolución individual de las jugadoras y analizar la información asociada a los partidos.

3.3. Implementación y desarrollo

3.3.1. Tecnologías utilizadas

La implementación del sistema se ha realizado utilizando un conjunto de tecnologías orientadas al desarrollo de aplicaciones, procesamiento de datos, almacenamiento persistente y aprendizaje automático. La elección de estas herramientas responde a la necesidad de construir una solución modular, mantenible y capaz de integrar información procedente de distintas fuentes.

El uso de Python como lenguaje principal de este sistema ha permitido integrar en un mismo entorno el procesamiento de datos, la lógica del backend y la ejecución del modelo predictivo. FastAPI se emplea para exponer los servicios del sistema mediante una API REST, facilitando la comunicación con la aplicación cliente. PostgreSQL se utiliza como sistema de almacenamiento persistente, mientras que SQLAlchemy permite abstraer las operaciones de acceso a datos y mantener una separación clara entre la lógica de negocio y la base de datos.

Para el tratamiento de los datos se utilizan pandas y NumPy, que permiten limpiar los archivos GPS, transformar las variables y calcular las métricas agregadas necesarias para las ventanas temporales. Por su parte, scikit-learn se emplea para integrar el modelo predictivo final dentro del backend. Para la obtención de datos contextuales de partidos se utiliza Selenium, permitiendo automatizar la extracción de información como calendario, rivales, resultados y goles de la página de la RFFM.

La interfaz de usuario se desarrolla mediante React Native y Expo, permitiendo construir una aplicación orientada a un uso táctil y centrada en las necesidades del entrenador. La organización funcional de las pantallas se ha descrito previamente en el Apartado 2.3.3 de diseño de la interfaz de usuario.

3.3.2. Implementación de la ingesta y almacenamiento de datos

La ingesta de datos se implementa mediante procesos encargados de incorporar al sistema las dos fuentes principales de información: los archivos GPS exportados desde Oliver y los datos contextuales de partidos obtenidos desde la web de la RFFM. El objetivo de esta fase es transformar los datos originales en registros estructurados que puedan almacenarse en la base de datos y ser utilizados posteriormente por los módulos de análisis, predicción y visualización.

En el caso de los archivos GPS, el sistema recibe ficheros en formato CSV y comprueba que contengan las columnas necesarias para su procesamiento. Durante esta fase se normalizan los nombres de las variables, se valida la información asociada a jugadoras y sesiones, y se aplican reglas de limpieza para descartar registros no válidos. Los criterios concretos de limpieza aplicados sobre el conjunto de datos se describen posteriormente en el Apartado 4.1.1.

Una vez procesado el archivo, se crea o actualiza el registro correspondiente en la entidad `Sesión`. Cada sesión queda identificada por su fecha, temporada y tipo de sesión, permitiendo diferenciar entre entrenamientos y partidos. Posteriormente, las métricas GPS de cada jugadora se almacenan en la entidad `Metricas_GPS`, asociando cada registro a la sesión y a la jugadora correspondiente.

La información contextual de los partidos se obtiene mediante un proceso de scraping implementado con Selenium. Este proceso automatiza la navegación por la web de la RFFM permitiendo extraer detalles del encuentro. Esta información se almacena en las entidades `Partido` y `Gol`, separando el contexto competitivo del registro físico procedente del GPS.

Cuando existe una sesión GPS correspondiente a un partido, el sistema vincula ambos registros almacenando el identificador de la sesión en la entidad `Partido`. En una primera fase, esta asociación se realiza a partir de la coincidencia entre la fecha de la sesión GPS y la fecha del encuentro. De este modo, la entidad `Partido` mantiene su clave primaria `id_partido`, mientras que la clave foránea `id_sesion` permite acceder a las métricas físicas registradas durante el partido. Esta vinculación permite relacionar la información competitiva obtenida mediante scraping con las métricas GPS almacenadas para esa sesión. En el caso de los goles marcados por el propio equipo, el sistema puede asociar el registro con una jugadora interna cuando la identificación es fiable; en los goles encajados o en casos ambiguos, se conserva únicamente la información descriptiva disponible.

El almacenamiento se realiza en PostgreSQL, utilizando SQLAlchemy para centralizar las operaciones de lectura y escritura. Esta capa de acceso a datos permite mantener separada la lógica de ingesta de la estructura concreta de la base de datos, facilitando el mantenimiento del sistema y futuras ampliaciones.

3.3.3. Implementación del backend y servicios principales

El backend se implementó con FastAPI, siguiendo la separación de responsabilidades definida en la arquitectura general. En lugar de concentrar toda la lógica en los endpoints, la aplicación se organiza mediante routers y servicios internos, de forma que la API actúa como punto de entrada y delega el procesamiento en los módulos correspondientes.

Durante el arranque de la aplicación se inicializan los componentes necesarios para el funcionamiento del sistema. Entre ellos se incluye la preparación del esquema de base de datos, la sincronización del conjunto de datos limpio cuando es necesario y la carga del modelo predictivo. Además, se configura *Cross-Origin Resource Sharing* (CORS) para permitir la comunicación entre el backend y la aplicación cliente.

La exposición HTTP se organiza mediante routers que agrupan las distintas funcionalidades del sistema y actúan como punto de entrada para las peticiones realizadas desde la aplicación cliente. Esta organización permite mantener una estructura modular, facilitando la separación de responsabilidades y el mantenimiento del código. Cada router delega el procesamiento en los servicios internos correspondientes, que se encargan de acceder a los datos, ejecutar los cálculos necesarios y construir las respuestas que posteriormente consume la interfaz de usuario.

La lógica interna del backend se implementa en servicios especializados, que son invocados desde los routers. Estos servicios se encargan de construir las respuestas que consume la aplicación, combinando información procedente de la base de datos con indicadores derivados calculados por el sistema. Por ejemplo, el servicio de perfil de jugadora agrega métricas de entrenamientos y partidos, evolución temporal, valores de RPE, zonas de intensidad y goles por temporada. Del mismo modo, el servicio de dashboard consolida información global del equipo para ofrecer una visión resumida al entrenador.

También se implementan procesos analíticos asociados al seguimiento de la carga. Entre ellos se encuentra el cálculo del desajuste entre la carga objetiva registrada y la percepción subjetiva del esfuerzo, que permite identificar posibles situaciones de sobreesfuerzo o infraesfuerzo. Estos resultados se utilizan para generar alertas interpretables y clasificaciones de riesgo que posteriormente se muestran en la interfaz.

El backend centraliza además la persistencia de resultados derivados, como ventanas temporales, predicciones e indicadores de desajuste. Esta decisión permite conservar trazabilidad sobre los valores utilizados y los resultados generados, evitando que la aplicación cliente tenga que recalcular información compleja. De esta forma, el frontend recibe respuestas ya estructuradas y orientadas a la visualización.

En conjunto, la implementación del backend permite coordinar la ingesta de datos, el acceso a la base de datos, la ejecución de procesos analíticos y la generación de respuestas para la interfaz. Esta

organización mantiene separada la lógica de presentación de la lógica de negocio y facilita futuras ampliaciones del sistema, como la incorporación de nuevas métricas, nuevos indicadores o modelos predictivos alternativos.

3.3.4. Integración del modelo predictivo

El modelo predictivo se integró como un componente interno del backend, como servicio de predicción de forma que queda encapsulado dentro de la propia API y no depende de la ejecución manual de scripts externos. Esta integración permite que las predicciones puedan ser reutilizadas por distintos módulos del sistema.

Durante el arranque de la aplicación, el backend comprueba la disponibilidad del artefacto del modelo y carga en memoria los componentes necesarios para realizar predicciones. De esta forma, el modelo no tiene que inicializarse en cada petición, lo que facilita una respuesta más eficiente por parte del servicio. En caso de que el artefacto no esté disponible o no cumpla el formato esperado, el sistema contempla un mecanismo de recuperación para mantener la capacidad de predicción de la API. Este proceso no forma parte del flujo ordinario de actualización, sino que actúa como respaldo ante la ausencia de un modelo persistido válido.

La generación de una predicción parte del histórico de sesiones almacenado. Para cada partido objetivo y jugadora, el sistema consulta los entrenamientos previos disponibles y calcula las métricas agregadas necesarias para construir la entrada del modelo. Estos valores se almacenan en la entidad `Ventana_Temporal`, que conserva la representación utilizada en la generación de predicciones y permite mantener trazabilidad sobre los datos empleados.

Antes de ejecutar el modelo, las variables de entrada se adaptan al mismo formato utilizado durante la fase experimental. Esto incluye el tratamiento de valores ausentes y la transformación de las variables numéricas cuando resulta necesario. De esta manera, se garantiza que los datos procesados por el backend sean coherentes con los empleados durante el entrenamiento y evaluación del modelo. Una vez preparada la entrada, el servicio de predicción ejecuta el modelo integrado y obtiene una zona de intensidad estimada

La selección del modelo final, la definición de las zonas de intensidad y la comparación entre distintas ventanas temporales se desarrollan en el Capítulo 4. En este apartado se describe únicamente cómo el modelo seleccionado se incorpora al backend y se utiliza dentro del funcionamiento general del sistema.

PRUEBAS Y RESULTADOS

En este capítulo se presentan los resultados obtenidos de las pruebas realizadas para alcanzar el objetivo propuesto. En primer lugar se explica la metodología usada para evaluar el rendimiento de los modelos. A continuación se realiza un análisis exploratorio de los datos para poder comprender su distribución y la naturaleza de las variables. Posteriormente se comparan los distintos modelos *Machine Learning* (ML) utilizados, así como el impacto de las distintas variaciones de ventanas temporales. Finalmente se presentan varios casos de uso reales que permiten evaluar la utilidad del sistema.

4.1. Metodología experimental

Para poder alcanzar el objetivo de comprender el rendimiento de las jugadoras y su desempeño dentro del equipo, se necesitó seguir una metodología experimental centrada en la preparación, organización y análisis de los datos históricos de sesiones de entrenamiento y partidos.

4.1.1. Preprocesado de los datos

Antes de iniciar cualquier análisis se llevó a cabo un proceso de limpieza en el que se aplicaron los siguientes criterios:

- Eliminación de registros con tiempo de juego igual a 0: correspondientes a jugadoras registradas en el sistema pero ausentes en la sesión.
- Eliminación de registros con una distancia total recorrida inferior a 1000 m. Se asume que corresponden a sesiones no completadas o a roles específicos (p. ej., portería), pudiendo introducir ruido en el modelado.
- Eliminación de columnas con más del 50 % de valores iguales a 0. Se eliminan por su escasa capacidad informativa.
- Sustitución de valores no registrados por NaN en la variable RPE y, por consiguiente en su derivada s-RPE.

Este proceso de limpieza garantiza un dataset consistente tanto para analizar patrones individuales como a nivel colectivo para la toma de decisiones sobre este.

4.1.2. Construcción de ventanas y planteamiento del análisis

Para estudiar la evolución de la carga a lo largo del tiempo y su posible relación con la intensidad posterior en competición, se definieron ventanas temporales formadas por 2, 3, 4, 5 y 6 sesiones de entrenamiento previas a cada partido. Para cada jugadora i y partido p , la ventana temporal de tamaño N se define como el conjunto de los N entrenamientos inmediatamente anteriores al partido:

$$W_{i,p}^{(N)} = \{T_{i,p,1}, T_{i,p,2}, \dots, T_{i,p,N}\} \quad (4.1)$$

En esta expresión, $W_{i,p}^{(N)}$ representa la ventana temporal formada por los N entrenamientos previos al partido p para la jugadora i , y $T_{i,p,k}$ representa el entrenamiento k de dicha ventana.

A partir de cada ventana, las variables de carga se agregaron mediante la media, con el objetivo de obtener una representación resumida y estable del comportamiento de la jugadora durante el periodo considerado. Para una variable de carga x , el valor agregado se calcula como:

$$\bar{x}_{i,p}^{(N)} = \frac{1}{N} \sum_{k=1}^N x_{i,p,k} \quad (4.2)$$

En esta fórmula, $x_{i,p,k}$ representa el valor de la variable x para la jugadora i en el entrenamiento k de la ventana previa al partido p . Por su parte, $\bar{x}_{i,p}^{(N)}$ representa el valor agregado utilizado como entrada del modelo para la jugadora i y el partido p , calculado como la media de los N entrenamientos anteriores.

Pese a que las ventanas temporales se construyeron a partir del conjunto completo de las métricas, únicamente un subconjunto de ellas fue seleccionado como entrada para modelo. La justificación de esta selección se detalla en el Apartado 4.3. Por su parte, la variable objetivo se define como la intensidad del partido, entendida como una medida de la exigencia del contexto competitivo.

A partir de esta formulación, se plantea un enfoque predictivo orientado a estimar la intensidad de un partido en función de la carga acumulada en entrenamientos previos. Aunque inicialmente se consideró un enfoque de regresión, finalmente se optó por discretizar la variable objetivo en zonas de intensidad para mejorar la interpretabilidad de los resultados.

De forma complementaria, se incluye un análisis del desajuste entre la percepción subjetiva de la carga y la demanda física real. Este análisis se realiza a nivel de jugadora, utilizando aquellos registros que incluyen información de RPE junto con las métricas objetivas obtenidas mediante sus dispositivos GPS. El objetivo es identificar posibles discrepancias entre la carga percibida y la carga medida, aportando información relevante sobre la adecuación y toma de decisiones en cada sesión.

4.2. Análisis exploratorio de datos

El conjunto de datos proviene de registros propios del equipo de fútbol 7 femenino S.A.D. San Gabriel en categoría Senior. Estos datos fueron obtenidos mediante el dispositivo GPS Oliver con un total de 38 atributos. Para el análisis exploratorio de datos o *Exploratory Data Analysis* (EDA) se consideraron únicamente las métricas numéricas relacionadas con el rendimiento físico, excluyendo aquellas columnas que no aportan información cuantitativa para el análisis estadístico. La descripción completa de todas las métricas se incluye en la Tabla B.1 del Anexo B.

4.2.1. Distribución de variables

Para comprender la estructura de los datos, se analizaron las distribuciones de todas las métricas numéricas relacionadas con el rendimiento físico. Esta examinación se hizo mediante histogramas con estimación de densidad o *Kernel Density Estimation* (KDE) y boxplots, diferenciando entre sesiones de tipo GAME y tipo TRAINING (Figuras B.1 y B.2). Adicionalmente, se incluye en el cuerpo de este apartado la distribución de la variable objetivo (Int. de la sesión) en contexto de partido (Figura 4.1), al tratarse de la variable de referencia del problema y ser clave para la posterior discretización del objetivo en el proceso de modelado.

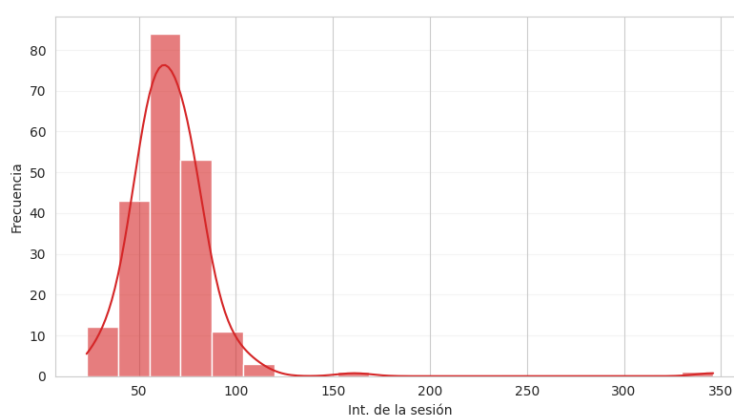


Figura 4.1: Distribución del target *Int. de la sesión* en GAME.

En general, se observan diferencias sistemáticas entre dichas sesiones, reflejando perfiles de cargas distintos en cada contexto: los partidos tienden a exhibir valores superiores y mayor variabilidad en comparación con los entrenamientos. Este patrón refleja mayor exigencia física en competiciones oficiales.

No obstante, el solapamiento observado entre curvas y cajas indica que la separación no es absoluta cuando se considera cada métrica de forma individual. La comparación mediante boxplots permite además examinar de forma robusta la mediana e identificar la cantidad de valores atípicos. La diferencia en mediana y dispersión entre tipos de sesión evidencia la heterogeneidad de la carga física y

refuerza la necesidad de un enfoque multivariable para la interpretación de la información. Por ejemplo, en la distancia recorrida se observa una mediana aproximada de 3100 m en entrenamientos frente a unos 4000 m en partidos. Esta diferencia también aparece en variables de alta intensidad, como las carreras de alta intensidad, cuya mediana se sitúa alrededor de 250-300 m en entrenamientos y aumenta hasta aproximadamente 650 m en partidos. Del mismo modo, la velocidad máxima presenta valores centrales superiores en competición, con una mediana cercana a 20 km/h en partidos frente a unos 17-18 km/h en entrenamientos. Además, la interpretación de las métricas acumulativas debe contextualizarse según la duración de la sesión, ya que los entrenamientos presentan una extensión temporal superior a la de los partidos.

Estas observaciones serán relevantes para comprender la consistencia de cada una de las métricas y su capacidad de reflejar de manera confiable la carga física real.

4.2.2. Correlación entre métricas

Para comprender las relaciones entre las métricas, se calculó la matriz de correlación de Spearman para todas las variables numéricas, visualizada en la Figura B.3. Este coeficiente permite identificar métricas altamente correlacionadas ($\rho \geq 0,8$) que aportan información redundante, así como métricas independiente que conservan variabilidad útil.

Del análisis se observan varios pares de métricas con correlaciones muy elevadas, lo que indica la presencia de información redundante en el conjunto de datos. Algunas de estas relaciones son esperables, como ocurre entre las *Carreras de Máx. Int. (m)* y las *Carreras de Máx. Int. (#)* (0,97), o entre las *Ace. Máx. Int. (m)* y las *Ace. Máx. Int. (#)* (0,99), al medir una misma acción de forma diferente.

También se observan correlaciones altas entre variables relacionadas con la carga física. Por ejemplo, la *Dist. Recorrida (m)* y el *Trote (m)* presentan una correlación de 0,93, lo que indica que una parte importante de la distancia total procede de este rango de desplazamiento. Asimismo, el *Vol. de la sesión* y la *Int. de la sesión - Velocidad* muestran una correlación de 0,80. Estas relaciones sugieren que algunas variables describen componentes próximos de la carga registrada durante la sesión y, por tanto, podrían aportar información parcialmente redundante para el modelado.

Este filtro no supervisado por correlación servirá como base para definir uno de los conjuntos preliminares de métricas candidatas para construir las variables de entrada de los modelos.

4.3. Selección de variables de entrada

La selección de variables de entrada se plantea como un problema de evaluación del impacto que distintos subconjuntos de variables tienen sobre la capacidad predictiva del modelo en un contexto

temporal. En lugar de asumir a priori un único conjunto óptimo de variables, se comparan tres configuraciones representativas

- **Subconjunto basado en criterios metodológicos** Se construye a partir del análisis de redundancia (ver Apartado 4.2.2) y de la estabilidad. En este proceso se busca identificar aquellas métricas que presentan mayor consistencia en diferentes ejecuciones y menor correlación media con el resto de variables, manteniendo únicamente aquellas que aportan información complementaria al modelo.
- **Subconjunto definido por conocimiento experto**: Se define a partir del criterio del cuerpo técnico, con métricas asociadas a carga externa, intensidad, locomoción y algunas acciones con balón. El objetivo de este subconjunto es evaluar si una selección basada en criterios deportivos y fácilmente comprensibles para el entrenador resulta competitiva frente a otros enfoques.
- **Conjunto completo de variables agregadas**: Este conjunto incluye todas las variables agregadas disponibles tras la construcción de ventanas temporales. Su finalidad es analizar el rendimiento del modelo cuando se dispone de toda la información generada a partir de los entrenamientos previos, sin aplicar una selección previa de variables. No obstante, este enfoque puede incrementar la redundancia entre variables y aumentar el riesgo de sobreajuste, especialmente en contextos con un número limitado de observaciones en relación con el número de características disponibles.

La evaluación de estos subconjuntos se realiza mediante un modelo Random Forest combinado con validación cruzada temporal (TimeSeriesSplit), garantizando que la evaluación simula un escenario real de predicción futura. Como métricas de evaluación se utilizan macro F1, balanced accuracy y accuracy, siendo la macro F1 la métrica principal debido al posible desbalance entre clases. Adicionalmente, se analiza la estabilidad del modelo mediante la desviación estándar del macro F1 entre folds.

Para facilitar la comparación global, se construye un score compuesto que integra rendimiento medio, estabilidad y simplicidad del modelo. La **estabilidad** se define a partir de la desviación estándar del macro F1 entre folds temporales, de forma que valores más bajos indican un comportamiento más consistente del modelo mientras que la **simplicidad** se calcula en función del número de variables utilizadas en cada subconjunto, favoreciendo configuraciones con menor dimensionalidad. Este score permite comparar los subconjuntos no solo en términos de precisión, sino también en términos de robustez y complejidad.

$$\text{Score} = 0,7 \cdot F1_{\text{norm}} + 0,2 \cdot \text{Estabilidad}_{\text{norm}} + 0,1 \cdot \text{Simplicidad}_{\text{norm}} \quad (4.3)$$

A partir de esta formulación se obtiene una métrica agregada que permite ordenar los distintos subconjuntos de variables en función de un compromiso entre rendimiento predictivo, estabilidad temporal y simplicidad del modelo. No obstante, el valor del score final no debe interpretarse de forma aislada sino en conjunto con la evolución de las métricas base a lo largo de los folds temporales, tal y como se muestra en la Figura 4.2. La gráfica de la izquierda muestra el rendimiento medio de cada subconjunto mediante macro-F1, balanced accuracy y accuracy. En estas tres métricas, valores más altos indican mejor rendimiento predictivo. La gráfica central muestra la evolución del macro-F1 en los distintos folds temporales. En este caso, no solo interesa alcanzar valores altos, sino también observar una evolución

estable entre particiones, ya que grandes variaciones indicarían menor robustez temporal del subconjunto. Por último, la gráfica de la derecha muestra el score final compuesto que integra rendimiento medio, estabilidad y simplicidad. En este score, valores más altos indican un mejor compromiso global entre capacidad predictiva, estabilidad temporal y complejidad del conjunto de variables.

En este análisis de subconjuntos de variables se fija la ventana temporal en $N = 2$, con el objetivo de mantener constante la estructura temporal de entrada y evaluar únicamente el efecto de la selección de variables. Es importante destacar que esta etapa es independiente del posterior estudio de comparación entre modelos y tamaños de ventana, en el que la longitud de la ventana será tratada como un hiperparámetro adicional.

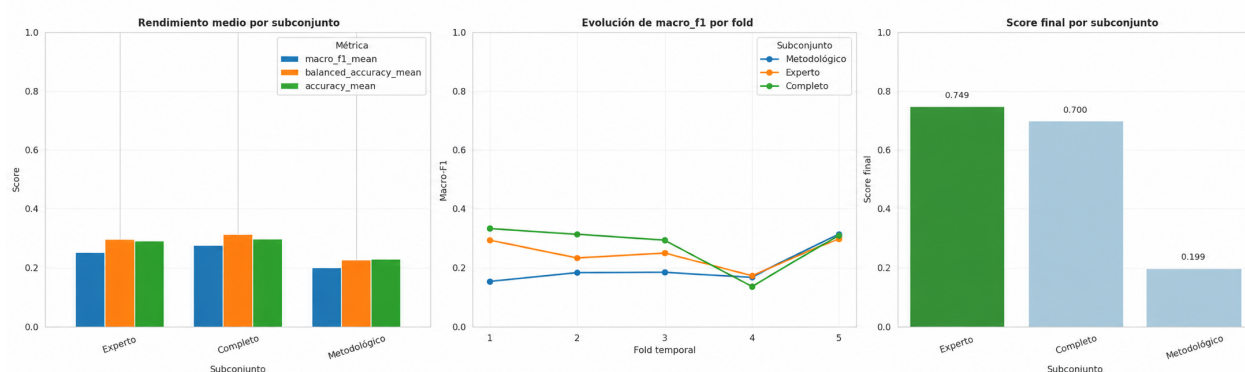


Figura 4.2: Comparativa de subconjuntos (RF, ventana=2).

En este sentido, aunque el subconjunto metodológico presenta una mayor simplicidad, su rendimiento predictivo es sensiblemente inferior y muestra menor consistencia temporal. Por otro lado, el conjunto completo de variables alcanza un mayor rendimiento medio, pero introduce una mayor inestabilidad y complejidad, lo que reduce su utilidad práctica en un contexto aplicado.

Finalmente, el subconjunto definido por el equipo técnico consigue el mejor equilibrio global entre las tres dimensiones evaluadas, manteniendo un rendimiento competitivo (Macro-F1), una estabilidad temporal consistente en los folds de validación y un número de variables suficientemente reducido como para facilitar su interpretación y uso operativo.

4.4. Comparación de modelos y ventanas temporales

Una vez seleccionado el conjunto de variables de entrada se comparó tanto modelos de clasificación como longitudes de ventana temporal para poder identificar la configuración más adecuada para la predicción de zonas de intensidad. Este análisis es relevante porque la carga acumulada en los entrenamientos previos puede influir de forma diferente según el periodo temporal considerado. En este apartado se evalúan ventanas de 2, 3, 4, 5 y 6 entrenamientos previos al partido. El objetivo no

es únicamente identificar la configuración con mayor rendimiento, sino también analizar su estabilidad temporal y su utilidad práctica dentro de una herramienta orientada al cuerpo técnico.

4.4.1. Configuración experimental y comparación global

Para la configuración experimental se evaluaron modelos representativos de distintas familias de aprendizaje supervisado: Logistic Regression, como modelo lineal e interpretable; Random Forest, como modelo basado en árboles de decisión; XGBoost, como técnica de boosting para datos tabulares; y SVM, como modelo basado en márgenes máximos. Además, se incluyeron modelos Dummy como referencia base para comprobar si los modelos entrenados aportan información predictiva real frente a estrategias simples. Cada modelo se evaluó utilizando diferentes tamaños de ventana temporal, con el objetivo de analizar cómo influye la cantidad de entrenamientos previos considerados en la predicción de la zona de intensidad del partido. La Tabla 4.1 recoge las configuraciones con mejor rendimiento según Macro-F1. El conjunto completo de combinaciones evaluada se incluye en el Anexo C.1

Modelo	Ventana	Macro-F1	Std	Balanced Acc.	Accuracy
LogisticRegression	5	0.3541	0.0418	0.3882	0.3867
LogisticRegression	4	0.3453	0.0412	0.3542	0.3582
LogisticRegression	6	0.3299	0.0739	0.3499	0.3800
LogisticRegression	3	0.3145	0.0356	0.3374	0.3315
SVM	6	0.3097	0.1142	0.3556	0.3247
XGBoost	5	0.2999	0.0420	0.3268	0.3136
RandomForest	6	0.2971	0.0865	0.3781	0.3382

Tabla 4.1: Configuraciones con mejor rendimiento según la métrica Macro-F1.

Los resultados muestran que Logistic Regression con una ventana de cinco entrenamientos previos obtiene el mejor rendimiento global, alcanzando un Macro-F1 medio de 0.3541 y una desviación estándar de 0.0418. Aunque el valor absoluto de la métrica es moderado, el modelo supera a los baselines utilizados. Esta comparación es especialmente relevante frente al baseline estratificado, más adecuado en un escenario con clases aproximadamente balanceadas, ya que obtiene un Macro-F1 cercano a 0.24–0.25 frente al 0.3541 del modelo final.

Este resultado indica que las variables agregadas de carga contienen información útil para anticipar la zona de intensidad del partido, aunque la tarea continúa siendo compleja debido a la variabilidad propia del contexto competitivo.

La Figura 4.3 muestra la evolución del Macro-F1 medio en función de la ventana temporal utilizada para cada modelo. Se observa que Logistic Regression mantiene el mejor comportamiento en la mayoría de configuraciones evaluadas, especialmente en ventanas intermedias y largas. El rendimiento aumenta desde las ventanas más cortas hasta alcanzar su máximo con cinco entrenamientos previos,

descendiendo ligeramente al considerar seis sesiones.

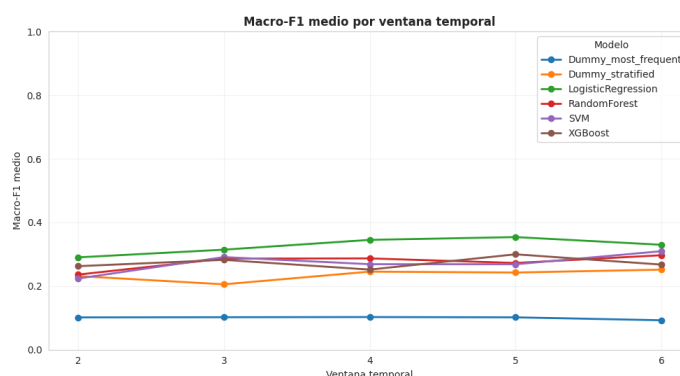


Figura 4.3: Evolución del rendimiento Macro-F1 según modelo y tamaño de ventana temporal.

Este comportamiento sugiere que una ventana intermedia puede capturar mejor la carga acumulada relevante antes del partido. Por el contrario, ventanas demasiado cortas pueden perder información contextual sobre la carga reciente mientras que ventanas excesivamente largas pueden introducir ruido o diluir el efecto de las sesiones más próximas al partido.

En el caso de SVM el mejor resultado aparece con una ventana de seis entrenamiento pero con una desviación estándar elevada, lo que refleja una menor estabilidad temporal. Random Forest y XGBoost presentan resultados más irregulares y no consiguen superar el rendimiento obtenido por Logistic Regression.

La Figura 4.4 presenta el heatmap de Macro-F1 para todas las combinaciones modelo-ventana evaluadas. Esta representación permite visualizar de forma global cómo Logistic Regression concentra los mejores resultados en varias ventanas consecutivas, mostrando además una evolución más consistente que el resto de modelos. Por el contrario, los modelos más complejos presentan patrones menos estables y una mayor dispersión en el rendimiento.

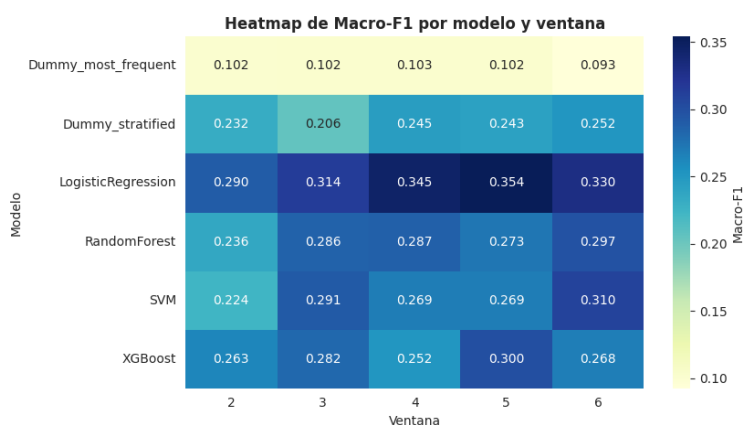


Figura 4.4: Heatmap de Macro-F1 por modelo y ventana temporal.

La ausencia de mejoras claras en Random Forest, XGBoost y SVM puede estar relacionada con el

tamaño limitado del dataset y con la elevada variabilidad del contexto deportivo. Además, las variables utilizadas son agregaciones compactas y estructuradas lo que puede favorecer que un modelo lineal capture gran parte de la señal predictiva sin necesidad de relaciones altamente no lineales.

Desde una perspectiva aplicada, este resultado resulta especialmente relevante ya que el modelo más simple es también el que ofrece mejor comportamiento global y mayor estabilidad temporal. Logistic Regression facilita la interpretación posterior de las predicciones y de la contribución de las variables, aspecto especialmente importante en una herramienta orientada al cuerpo técnico.

4.4.2. Análisis de errores y comportamiento de las predicciones

Una vez identificada la mejor combinación modelo-ventana se analizó con mayor detalle el comportamiento de Logistic Regression con ventana de cinco entrenamientos previos mediante predicciones out-of-fold. Este análisis permite estudiar los errores del modelo sobre observaciones no utilizadas durante el entrenamiento de cada fold temporal.

La Figura 4.5 muestra la matriz de confusión obtenida para la configuración seleccionada. La matriz normalizada por filas permite observar qué proporción de ejemplos de cada zona real son clasificados correctamente o confundidos con otras zonas.

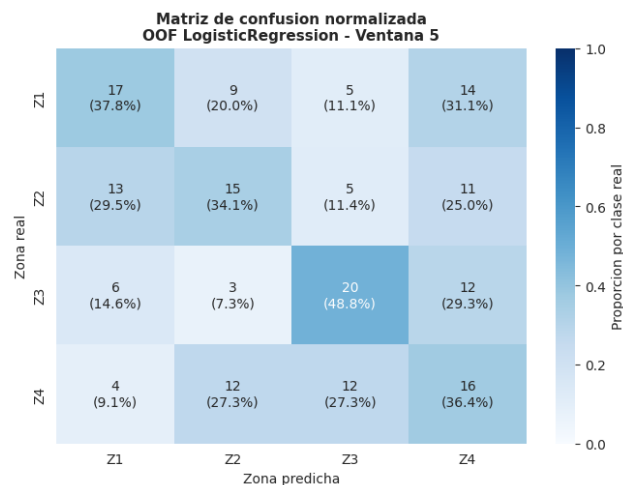


Figura 4.5: Matriz de confusión normalizada obtenida sobre el conjunto de evaluación.

En general, se observa que el modelo no concentra todas sus predicciones en una única clase, sino que distribuye las predicciones entre las distintas zonas de intensidad. No obstante, el rendimiento varía entre clases. La zona Z3 es la que presenta un mejor comportamiento relativo, con un porcentaje de acierto cercano al 49 %. En cambio, el resto de zonas muestran una mayor dispersión en las predicciones.

Las principales confusiones entre zonas reflejan la dificultad del problema. La intensidad de un partido no depende únicamente de la carga acumulada en entrenamientos previos, sino también de

factores contextuales no incluidos en el modelo. Por ello, es esperable que las fronteras entre zonas de intensidad no sean completamente separables a partir de las variables disponibles.

La Figura 4.6 compara la distribución de zonas reales frente a las zonas predichas por el modelo. Se observa que la distribución predicha mantiene una estructura relativamente similar a la distribución real, aunque existe una tendencia a sobreestimar la frecuencia de la zona Z4 y a infraestimar algunas zonas de menor intensidad.

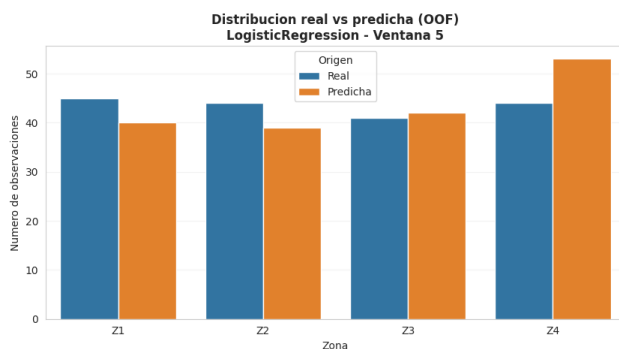


Figura 4.6: Distribución de clases reales frente a las predicciones del modelo.

Este comportamiento indica que el modelo no se limita a reproducir una única clase dominante, pero también muestra que tiene dificultades para representar con precisión todas las zonas. Desde un punto de vista aplicado, esta tendencia a predecir zonas de intensidad elevada puede interpretarse con cautela como una señal conservadora, útil para alertar al cuerpo técnico de posibles escenarios exigentes.

4.4.3. Selección del modelo final

A partir de los resultados obtenidos, se seleccionó Logistic Regression con ventana temporal de cinco entrenamientos previos como configuración final del sistema.

La elección se basa en varios criterios. En primer lugar, esta combinación obtiene el mayor valor medio de Macro-F1 entre todas las configuraciones evaluadas. En segundo lugar, presenta una desviación estándar reducida entre folds, lo que indica una mayor estabilidad temporal en comparación con otros modelos. En tercer lugar, al tratarse de un modelo lineal, ofrece una mayor simplicidad e interpretabilidad, aspectos especialmente relevantes en una aplicación cuyo usuario final es un entrenador y no un analista de datos.

Aunque el rendimiento absoluto del modelo es moderado, debe interpretarse teniendo en cuenta la complejidad del problema. La intensidad de un partido es una variable influida por numerosos factores externos al historial de entrenamiento, por lo que no se espera una predicción perfecta únicamente a partir de métricas de carga acumulada. No obstante, el hecho de superar claramente a los modelos

baseline indica que el sistema captura patrones útiles y puede proporcionar una estimación orientativa de la exigencia competitiva.

Por tanto, el modelo final no debe entenderse como un sustituto del criterio técnico del entrenador, sino como una herramienta complementaria que permite anticipar posibles escenarios de intensidad y apoyar la planificación de entrenamientos.

Una vez seleccionada la configuración final, en el siguiente apartado se analiza la contribución de las variables de entrada al modelo, con el objetivo de aportar interpretabilidad a las predicciones y facilitar su comprensión por parte del cuerpo técnico.

4.5. Interpretación del modelo

Una vez seleccionada la configuración final del modelo, se analizó la influencia de las variables de entrada con el objetivo de aportar interpretabilidad al sistema. Para ello, se empleó la técnica *permutation importance*, utilizando la implementación proporcionada por *scikit-learn*. Esta técnica evalúa la disminución del rendimiento del modelo al permutar aleatoriamente los valores de cada variable, rompiendo así su relación con la variable objetivo.

En este análisis la importancia se calculó sobre el modelo final, correspondiente a Logistic Regression con una ventana temporal de cinco entrenamientos previos. La disminución del rendimiento se midió mediante la métrica Macro-F1, de forma que una mayor caída en dicha métrica indica una mayor relevancia de la variable para el proceso predictivo. No obstante, esta medida debe interpretarse con cautela, ya que no representa relaciones causales directas, sino la contribución relativa de cada variable dentro del modelo entrenado.

La Figura C.1 muestra la ordenación de las variables según su importancia global en el modelo final. Las variables con mayor contribución se relacionan principalmente con métricas locomotoras y esfuerzos de alta intensidad, destacando las carreras de alta intensidad, la distancia recorrida, la intensidad asociada a la velocidad y las aceleraciones de alta intensidad. En cambio, las variables técnicas con balón presentan una menor importancia relativa, lo que indica que aportan menos información discriminativa para este objetivo concreto.

Este resultado sugiere que el modelo utiliza una combinación de indicadores de volumen, intensidad y distribución del esfuerzo para clasificar la intensidad competitiva posterior. Además, la presencia de variables como el trote o la caminata indica que no solo resultan relevantes las acciones explosivas, sino también el patrón general de actividad durante los entrenamientos previos.

Desde el punto de vista del entrenador este análisis permite relacionar la predicción del modelo con posibles decisiones de planificación. Identificar qué dimensiones de la carga previa están más asociadas a una mayor exigencia competitiva puede ayudar al cuerpo técnico a decidir si conviene reforzar

determinados aspectos físicos, controlar la carga de trabajo o priorizar la recuperación antes del partido. Por ejemplo, si las variables más influyentes están relacionadas con acciones de alta intensidad, aceleraciones o volumen acumulado, esta información puede utilizarse para ajustar el contenido de los entrenamientos previos.

Sin embargo, la interpretación del modelo presenta ciertas limitaciones ya que la intensidad competitiva está condicionada por factores contextuales no incluidos en el dataset, como el rival, el planteamiento táctico, los minutos disputados, el marcador o el estado físico y emocional de la jugadora.

El objetivo de este análisis de interpretabilidad no es explicar por completo el rendimiento competitivo, sino aportar información adicional sobre qué aspectos de la carga acumulada utiliza el modelo para diferenciar entre zonas de intensidad.

Finalmente, además de interpretar qué variables influyen en la predicción de intensidad, resulta relevante analizar cómo perciben las jugadoras la carga realizada. Por ello, el siguiente apartado estudia el desajuste entre la carga objetiva registrada y la percepción subjetiva del esfuerzo mediante RPE.

4.6. Análisis del desajuste entre carga objetiva y percepción del esfuerzo

Tras la predicción de intensidad realizada en los apartados anteriores, se planteó un análisis complementario para estudiar cómo responde cada jugadora ante la carga objetiva registrada. El objetivo no es volver a predecir la intensidad sino comprobar si dicha intensidad se corresponde con la percepción subjetiva del esfuerzo reportada mediante el RPE. Este análisis permite detectar posibles desacoplamientos entre la carga externa y la respuesta subjetiva individual. Esto es relevante porque dos jugadoras pueden presentar una carga objetiva similar y sin embargo, percibir la sesión de forma diferente.

4.6.1. Construcción del índice de desajuste

Para poder calcular este desajuste se seleccionaron las sesiones con RPE válido y suficiente información objetiva. Las variables de carga se detectaron automáticamente entre las columnas numéricas del dataset, excluyendo identificadores y variables de contexto. Además, se descartaron columnas con pocos datos válidos, sin variabilidad o con demasiados ceros, para evitar que variables poco informativas afectasen al índice.

Como las variables objetivas tienen escalas distintas, se estandarizaron mediante z-score robusto por tipo de sesión. De este modo, se redujo la influencia de valores extremos y se evitó comparar directamente contextos diferentes, como entrenamientos y partidos.

A partir de la media de las variables objetivas estandarizadas se calculó una carga global de sesión (*objective_load_z*) y posteriormente, el RPE se estandarizó con el mismo procedimiento, obteniendo *rpe_z*. El desajuste puntual entre ambas dimensiones se definió como:

$$acute_mismatch = rpe_z - objective_load_z \quad (4.4)$$

Además, para incorporar el comportamiento individual previo de cada jugadora se calculó una componente histórica del desajuste que indica si el desajuste actual se aleja del patrón histórico de la jugadora. El índice se obtuvo como:

$$mismatch_score = 0,70 \cdot acute_mismatch + 0,30 \cdot mismatch_vs_hist_z \quad (4.5)$$

La interpretación del índice se resume en la Tabla 4.2

Resultado del índice	Interpretación
Valor positivo	La jugadora percibe más esfuerzo del que sugiere la carga objetiva registrada.
Valor negativo	La jugadora percibe menos esfuerzo del que sugiere la carga objetiva registrada.
Valor cercano a cero	Existe mayor alineación entre carga objetiva y percepción subjetiva.

Tabla 4.2: Interpretación de los valores del índice *mismatch_score* calculado.

4.6.2. Resultados y aplicación práctica

Para presentar los resultados del análisis se generó una tabla resumen (ver Tabla 4.3) por jugadora, incluyendo el número de sesiones analizadas, el desajuste medio y mediano, la variabilidad del índice, la proporción de sesiones con sobreesfuerzo e infraesfuerzo percibido, el RPE medio, la carga objetiva media y la clasificación final de tendencia. Esta tabla permite analizar no solo la dirección media del desajuste, sino también su consistencia y variabilidad a lo largo de las sesiones.

Además, en la Figura 4.7 se representó gráficamente el desajuste medio por jugadora para facilitar la comparación visual entre perfiles individuales. Los valores positivos indican una percepción subjetiva superior a la carga objetiva registrada, mientras que los valores negativos reflejan una percepción inferior.

Los resultados muestran que la mayoría de jugadoras presentan una respuesta alineada o mixta, es decir, no mantienen una tendencia suficientemente consistente hacia el sobreesfuerzo o el infraesfuerzo percibido. Sin embargo, se identifican dos perfiles especialmente claros. La jugadora V.G.R. presenta el mayor desajuste medio positivo, con un valor de 0,437, lo que indica una tendencia a percibir un esfuerzo superior al esperado respecto a la carga objetiva registrada. Esta tendencia se refuerza con el valor de % Over mostrado en la Tabla 4.3, que alcanza 0,629. Es decir, en el 62,9% de sus sesiones la jugadora percibió un esfuerzo superior al esperado.

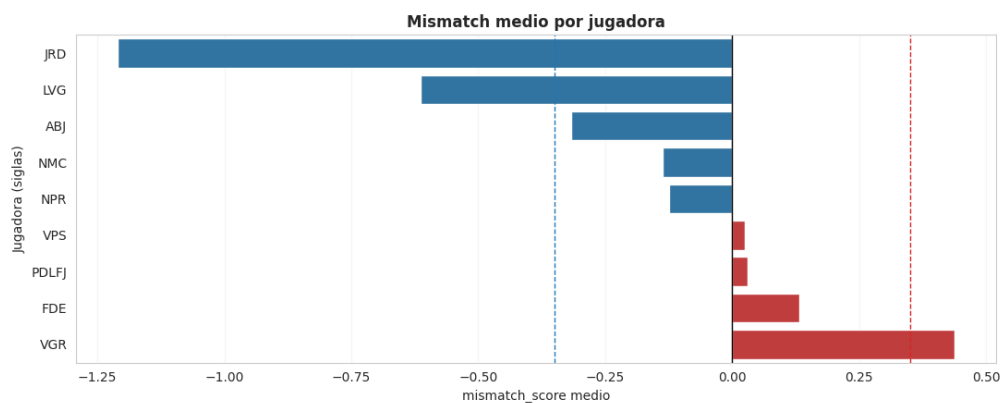


Figura 4.7: Desajuste medio entre carga objetiva y percepción subjetiva por jugadora.

En el extremo contrario, la jugadora J.R.D. presenta el desajuste medio más negativo, con un valor de -1,209, lo que refleja una tendencia a percibir un esfuerzo inferior al esperado respecto a la carga objetiva registrada. Esta interpretación también se confirma mediante el valor de % *Under*, que alcanza 0,725, indicando que en el 72,5 % de sus sesiones la percepción de esfuerzo fue inferior a la esperada.

Desde un punto de vista práctico, esta información puede ayudar al cuerpo técnico a interpretar mejor la respuesta individual de las jugadoras. Por ejemplo, una jugadora que percibe de forma repetida más esfuerzo del que sugieren sus métricas objetivas podría requerir una revisión de su recuperación, descanso, molestias o carga acumulada. Por el contrario, una jugadora que percibe sistemáticamente menos esfuerzo ante cargas objetivas elevadas podría mostrar una mayor tolerancia a la intensidad o una menor percepción del esfuerzo realizado.

De esta forma, este análisis puede servir como apoyo para ajustar la carga de entrenamiento, individualizar el seguimiento y detectar casos que convenga revisar con mayor atención.

Siglas	N	Media	Mediana	Std	% Over	% Under	RPE	Carga obj.	Tendencia
VGR	35	0.437	0.789	1.353	0.629	0.171	7.029	-0.416	Sobreesfuerzo
FDE	22	0.132	0.010	1.349	0.364	0.318	7.045	0.447	Alineada/mixta
PDLFJ	45	0.030	0.060	0.861	0.311	0.311	7.089	0.395	Alineada/mixta
VPS	77	0.026	0.035	1.069	0.351	0.312	6.494	0.008	Alineada/mixta
NPR	51	-0.125	-0.101	1.046	0.294	0.412	6.667	0.195	Alineada/mixta
NMC	65	-0.137	-0.168	0.917	0.231	0.323	7.200	0.755	Alineada/mixta
ABJ	69	-0.317	-0.285	0.758	0.145	0.449	5.797	-0.045	Alineada/mixta
LVG	51	-0.614	-0.477	1.314	0.255	0.490	5.922	0.481	Alineada/mixta
JRD	69	-1.209	-1.275	1.057	0.072	0.725	4.087	0.251	Infraesfuerzo

Tabla 4.3: Resumen del desajuste entre carga objetiva y percepción subjetiva por jugadora, ordenado de mayor a menor valor medio de desajuste.

CONCLUSIONES Y TRABAJO FUTURO

En este capítulo se presentan las principales conclusiones obtenidas tras el desarrollo del trabajo, valorando el grado de cumplimiento de los objetivos planteados. Además, se recogen las principales limitaciones encontradas y se proponen posibles líneas de mejora y ampliación futura.

5.1. Conclusiones

El trabajo desarrollado ha permitido comprobar que es posible construir un modelo predictivo de apoyo a la toma de decisiones en fútbol a partir de datos reales de entrenamientos y partidos. Aunque el contexto deportivo presenta una elevada variabilidad, los resultados muestran que las métricas de carga acumulada contienen información útil para estimar la zona de intensidad esperada en competición.

La formulación del problema como una tarea de clasificación por zonas de intensidad ha resultado adecuada para el objetivo planteado, ya que ofrece una salida más interpretable y útil para el entrenador que una predicción numérica aislada. Además, los resultados experimentales muestran que, en un conjunto de datos reducido y variable, un modelo más simple puede ofrecer un mejor equilibrio entre rendimiento, estabilidad temporal e interpretabilidad.

También se ha puesto de manifiesto la utilidad de combinar métricas objetivas, como las obtenidas mediante GPS, con información subjetiva como el RPE. El análisis del desajuste entre ambas dimensiones aporta una visión más completa de la respuesta individual de las jugadoras y puede servir como apoyo para un seguimiento más personalizado.

No obstante, el sistema presenta limitaciones. La intensidad de una jugadora en competición no depende únicamente de la carga acumulada en los entrenamientos previos, sino también de factores como el rival, el planteamiento táctico, los minutos disputados, la posición ocupada o el desarrollo del partido. Además, el tamaño del conjunto de datos y el uso de información procedente de un único equipo limitan la generalización de los resultados.

En conjunto, el trabajo constituye una primera aproximación funcional a un sistema predictivo apli-

cado al fútbol femenino. El modelo no pretende sustituir el criterio del entrenador, sino actuar como una herramienta complementaria para interpretar los datos disponibles, anticipar posibles escenarios de exigencia física y apoyar la planificación del equipo.

5.2. Trabajo futuro

Como trabajo futuro, una de las principales líneas de mejora consiste en incorporar nuevas variables contextuales como entradas del modelo predictivo. En este trabajo, la estimación de la zona de intensidad de cada jugadora se basa principalmente en la carga acumulada durante los entrenamientos previos. Sin embargo, la exigencia física experimentada en competición también puede depender del contexto del partido.

En este sentido, variables como la posición del rival en la clasificación antes del encuentro, la diferencia de nivel entre ambos equipos o la condición de local o visitante podrían utilizarse para entrenar el modelo y aportar una representación más completa del escenario competitivo. La hipótesis es que enfrentarse a un rival de mayor nivel puede implicar una mayor demanda física para determinadas jugadoras, debido a un aumento de acciones defensivas, transiciones, desplazamientos a alta intensidad o esfuerzos repetidos. Estas variables no tendrían como objetivo predecir el resultado del partido, sino mejorar la estimación de la intensidad física esperada para cada jugadora.

Otra línea de ampliación sería el desarrollo de perfiles de jugadoras a partir de sus métricas físicas y de rendimiento. Esta aproximación permitiría agrupar jugadoras con características similares y analizar qué perfiles se ajustan mejor a determinadas posiciones o roles dentro del campo. De esta forma, el sistema no solo serviría para estimar la intensidad esperada de un partido, sino también como apoyo al entrenador en la toma de decisiones relacionadas con la posición más adecuada para cada jugadora, la rotación del equipo o la planificación individualizada.

También sería interesante ampliar el conjunto de datos disponible, incorporando nuevas temporadas o registros de más partidos. Esto permitiría evaluar si los patrones observados se mantienen a lo largo del tiempo y mejorar la capacidad de generalización del modelo. Además, disponer de un mayor volumen de datos facilitaría el entrenamiento de modelos más específicos, por ejemplo diferenciando por posición, tipo de partido o perfil de jugadora.

Por último, desde el punto de vista de la aplicación, se podrían incorporar nuevas funcionalidades para facilitar el trabajo del entrenador, como alertas personalizadas, comparativas entre jugadoras, informes automáticos o consultas en lenguaje natural. Estas mejoras convertirían el sistema en una herramienta más completa para el seguimiento del equipo y la toma de decisiones durante la temporada.

BIBLIOGRAFÍA

- [1] R. Mackenzie and C. Cushion, "Performance analysis in football: A critical review and implications for future research," *Journal of Sports Sciences*, vol. 31, no. 6, pp. 639–676, 2013. doi: 10.1080/02640414.2012.746720.
- [2] H. Sarmento, R. Marcelino, M. T. Anguera, J. Campaniço, N. Matos, and J. C. Leitão, "Match analysis in football: A systematic review," *Journal of Sports Sciences*, vol. 32, no. 20, pp. 1831–1843, 2014. doi: 10.1080/02640414.2014.898852.
- [3] F. R. Goes, L. A. Meerhoff, M. J. O. Bueno, D. M. Rodrigues, F. A. Moura, M. S. Brink, M. T. Elferink-Gemser, A. J. Knobbe, and S. A. Cunha, "Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review," *European Journal of Sport Science*, vol. 21, no. 4, pp. 481–496, 2021. doi: 10.1080/17461391.2020.1747552.
- [4] T. Decroos, L. Bransen, J. Van Haaren, and J. Davis, "Actions speak louder than goals: Valuing player actions in soccer," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1851–1861, 2019. doi: 10.1145/3292500.3330758.
- [5] L. Pappalardo, P. Cintia, P. Ferragina, E. Massucco, D. Pedreschi, and F. Giannotti, "Playerank: Data-driven performance evaluation and player ranking in soccer via a machine learning approach," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 5, pp. 59:1–59:27, 2019. doi: 10.1145/3343172.
- [6] L. Carloni, A. De Angelis, G. Sansonetti, and A. Micarelli, "A machine learning approach to football match result prediction," in *HCI International 2021 - Posters*, pp. 473–480, 2021. doi: 10.1007/978-3-030-78642-7_63.
- [7] A. Olaizola, I. Errekagarri, E. Fernández, J. Castellano, J. Suckling, and K. Lopez-de Ipina, "Predicting female football outcomes by machine learning: behavioural analysis of goals as high stress events," *Humanities and Social Sciences Communications*, vol. 12, no. 1302, p. 1302, 2025. doi: 10.1057/s41599-025-05490-8.
- [8] P. Rahimian, J. Flisar, and D. Sumpter, "Automated explanation of machine learning models of footballing actions in words," 2025. arXiv:2504.00767.
- [9] P. C. Bourdon, M. Cardinale, A. Murray, P. Gastin, M. Kellmann, M. C. Varley, T. J. Gabbett, and A. J. Coutts, "Monitoring athlete training loads: Consensus statement," *International Journal of Sports Physiology and Performance*, vol. 12, no. Suppl 2, pp. S2–161–S2–170, 2017. doi: 10.1123/IJSPP.2017-0208.
- [10] S. L. Halson, "Monitoring training load to understand fatigue in athletes," *Sports Medicine*, vol. 44, no. Suppl 2, pp. 139–147, 2014. doi: 10.1007/s40279-014-0253-z.
- [11] R. Akenhead and G. P. Nassis, "Training load and player monitoring in high-level football: Current practice and perceptions," *International Journal of Sports Physiology and Performance*, vol. 11,

- no. 5, pp. 587–593, 2016. doi: 10.1123/ijsp.2015-0331.
- [12] M. Miguel, R. Oliveira, N. Loureiro, J. García-Rubio, and S. J. Ibáñez, “Load measures in training/-match monitoring in soccer: A systematic review,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 5, p. 2721, 2021. doi: 10.3390/ijerph18052721.
- [13] G. Ravé, U. Granacher, D. Boulosa, A. C. Hackney, and H. Zouhal, “How to use global positioning systems data to monitor training load in the real world of elite soccer,” *Frontiers in Physiology*, vol. 11, p. 944, 2020. doi: 10.3389/fphys.2020.00944.
- [14] J. J. Malone, R. Lovell, M. C. Varley, and A. J. Coutts, “Unpacking the black box: Applications and considerations for using gps devices in sport,” *International Journal of Sports Physiology and Performance*, vol. 12, no. Suppl 2, pp. S2–18–S2–26, 2017. doi: 10.1123/ijsp.2016-0236.
- [15] G. A. V. Borg, “Psychophysical bases of perceived exertion,” *Medicine and Science in Sports and Exercise*, vol. 14, no. 5, pp. 377–381, 1982.
- [16] F. M. Impellizzeri, E. Rampinini, A. J. Coutts, A. Sassi, and S. M. Marcora, “Use of rpe-based training load in soccer,” *Medicine and Science in Sports and Exercise*, vol. 36, no. 6, pp. 1042–1047, 2004. doi: 10.1249/01.MSS.0000128199.23901.2F.
- [17] M. Haddad, G. Stylianides, L. Djaoui, A. Dellal, and K. Chamari, “Session-rpe method for training load monitoring: Validity, ecological usefulness, and influencing factors,” *Frontiers in Neuroscience*, vol. 11, p. 612, 2017. doi: 10.3389/fnins.2017.00612.
- [18] H. Alexiou and A. J. Coutts, “A comparison of methods used for quantifying internal training load in women soccer players,” *International Journal of Sports Physiology and Performance*, vol. 3, no. 3, pp. 320–330, 2008. doi: 10.1123/ijsp.3.3.320.
- [19] T. J. Gabbett, “The training-injury prevention paradox: should athletes be training smarter and harder?,” *British Journal of Sports Medicine*, vol. 50, no. 5, pp. 273–280, 2016. doi: 10.1136/bjsports-2015-095788.
- [20] T. Soligard, M. Schwellnus, J.-M. Alonso, R. Bahr, B. Clarsen, H. P. Dijkstra, T. Gabbett, M. Gleeson, M. Hägg, and L. Engebretsen, “How much is too much? international olympic committee consensus statement on load in sport and risk of injury,” *British Journal of Sports Medicine*, vol. 50, no. 17, pp. 1030–1041, 2016. doi: 10.1136/bjsports-2016-096581.
- [21] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, second ed., 2009.
- [22] Google Developers, “Logistic regression: Calculating a probability with the sigmoid function,” 2026. Google Machine Learning Crash Course.
- [23] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. doi: 10.1023/A:1010933404324.
- [24] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016. doi: 10.1145/2939672.2939785.
- [25] DataScientest, “Bagging vs boosting,” 2025. Imagen consultada online.

-
- [26] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995. doi: 10.1007/BF00994018.
- [27] Larhmam, "Svm margin.png," 2018. Imagen en Wikimedia Commons, licencia CC BY-SA 4.0.
- [28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011. Artículo.
- [29] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, 2009. doi: 10.1016/j.ipm.2009.03.002.
- [30] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.
- [31] scikit-learn developers, "Metrics and scoring: quantifying the quality of predictions," 2026. Documentación oficial.
- [32] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Information Sciences*, vol. 191, pp. 192–213, 2012. doi: 10.1016/j.ins.2011.12.028.
- [33] scikit-learn developers, "Timeseriesplit," 2026. Documentación oficial.
- [34] C. Molnar, *Interpretable Machine Learning*. Independently published, 2022. Libro disponible online.
- [35] scikit-learn developers, "Permutation feature importance," 2026. Documentación oficial.
- [36] Real Federación de Fútbol de Madrid, "Sitio web oficial de la real federación de fútbol de madrid," 2026. Sitio web oficial.
- [37] MatchApp, "Matchapp," 2026. Aplicación oficial en Google Play.
- [38] O. Sports, "Oliver sports gps tracker," 2026. Página oficial.
- [39] pandas development team, "pandas documentation," 2026. Documentación oficial.
- [40] N. developers, "Numpy documentation," 2026. Documentación oficial.
- [41] Meta Platforms, "React native documentation," 2026. Documentación oficial.
- [42] Expo, "Expo documentation," 2026. Documentación oficial.

ACRÓNIMOS

- API** *Application Programming Interface.*
- CORS** *Cross-Origin Resource Sharing.*
- CSV** *Comma-Separated Values.*
- EDA** *Exploratory Data Analysis.*
- GPS** *Global Positioning System.*
- HTTP** *Hypertext Transfer Protocol.*
- KDE** *Kernel Density Estimation.*
- ML** *Machine Learning.*
- REST** *Representational State Transfer.*
- RFFM** *Real Federación de Fútbol de Madrid.*
- RPE** *Rating of Perceived Exertion.*
- SQL** *Structured Query Language.*
- s-RPE** *Session Rating of Perceived Exertion.*
- SVM** *Support Vector Machine.*

APÉNDICES

DISEÑO E IMPLEMENTACIÓN

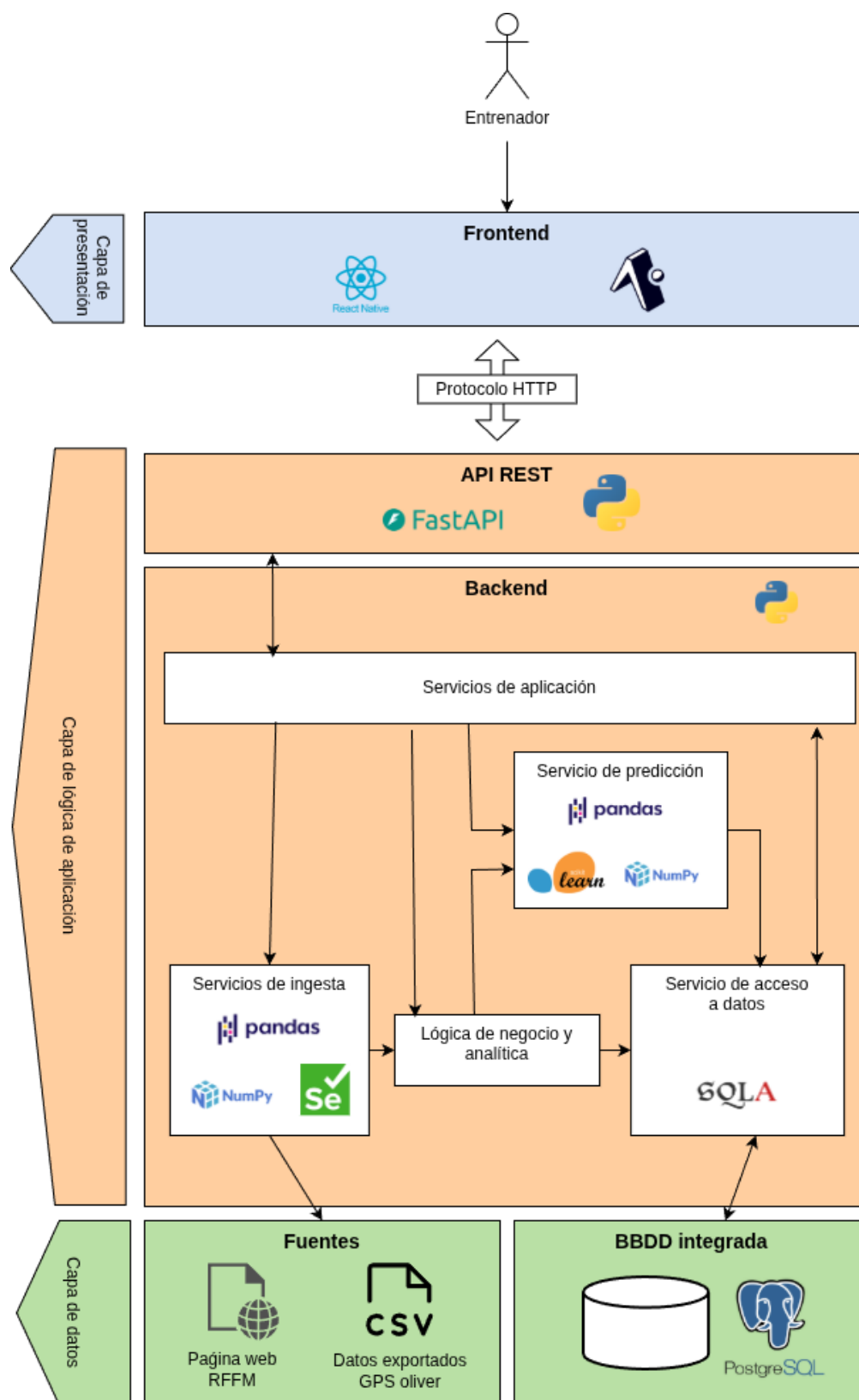


Figura A.1: Arquitectura general del sistema.

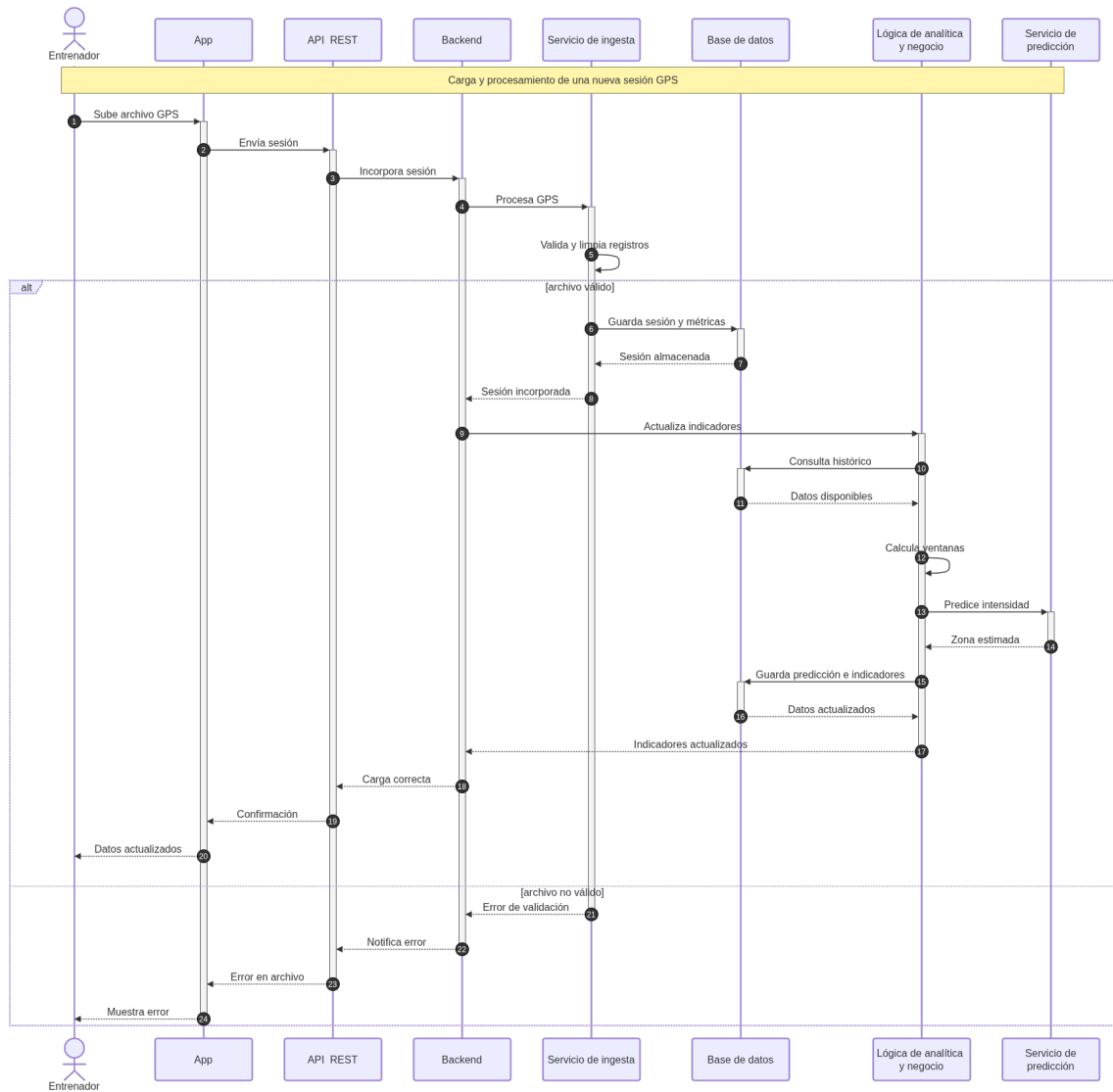


Figura A.2: Diagrama de secuencia del proceso de carga y procesamiento de una sesión GPS.

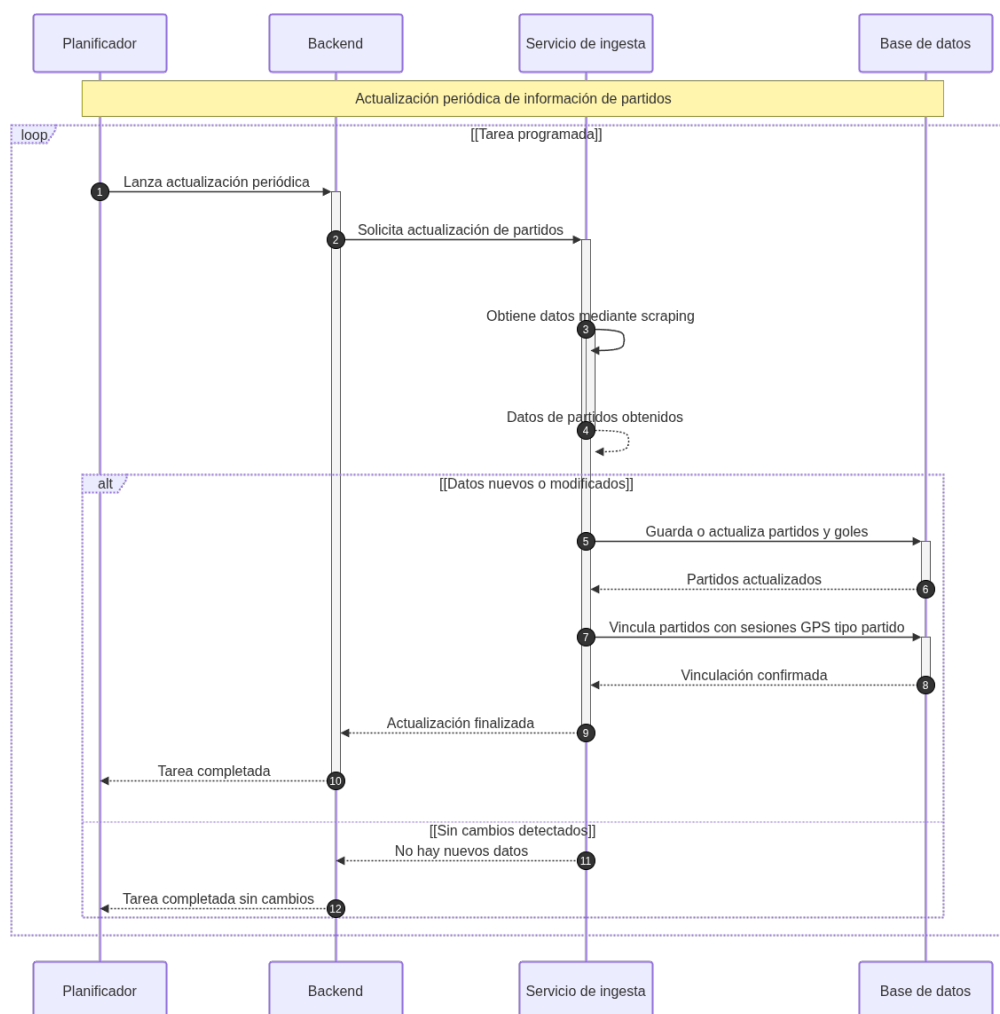


Figura A.3: Diagrama de secuencia del proceso de actualización periódica de partidos.

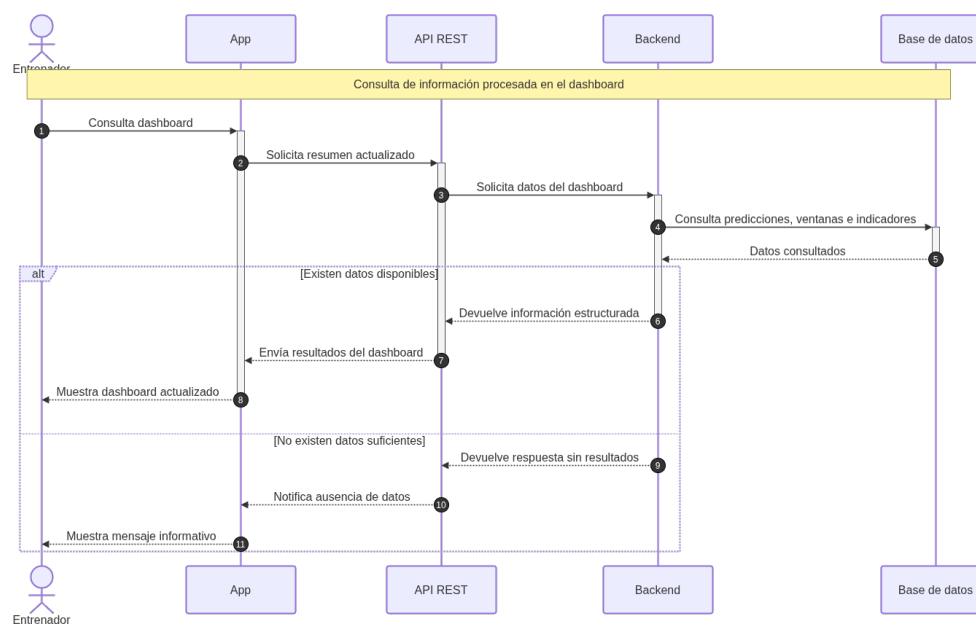


Figura A.4: Diagrama de secuencia del proceso de consulta de resultados en el dashboard.

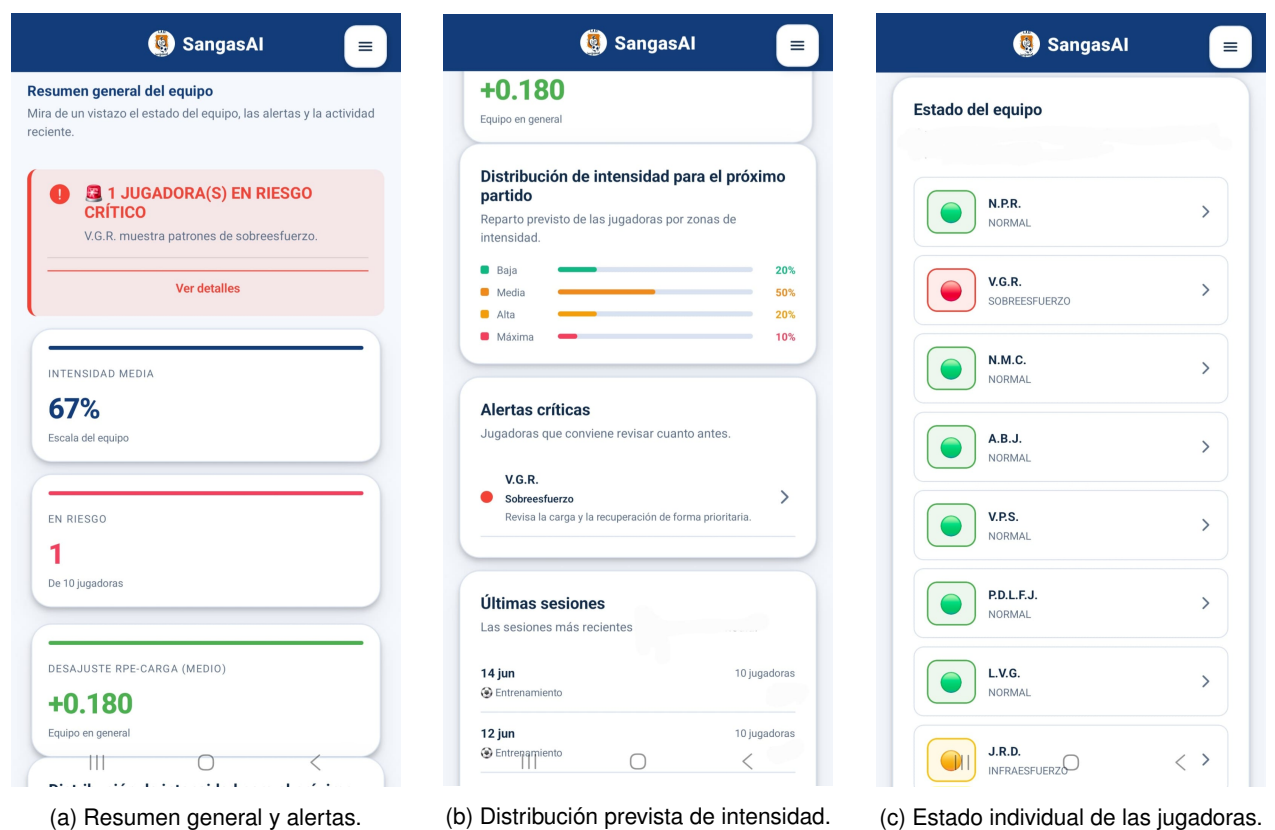


Figura A.5: Capturas del dashboard principal de la aplicación.

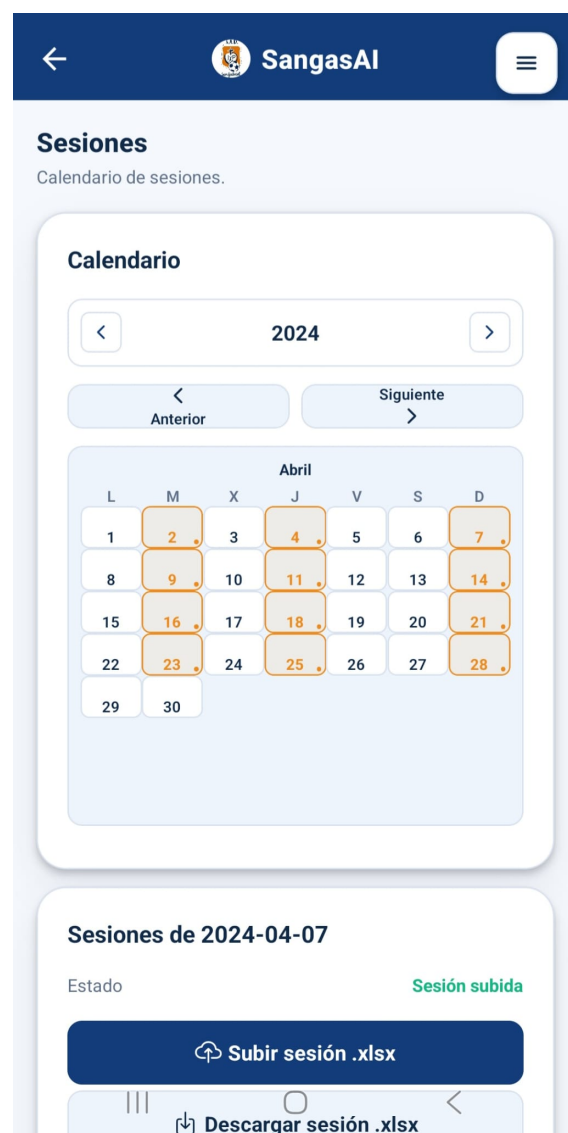


Figura A.6: Captura de la pantalla de carga de sesiones.



Figura A.7: Captura del listado de jugadoras de la aplicación.

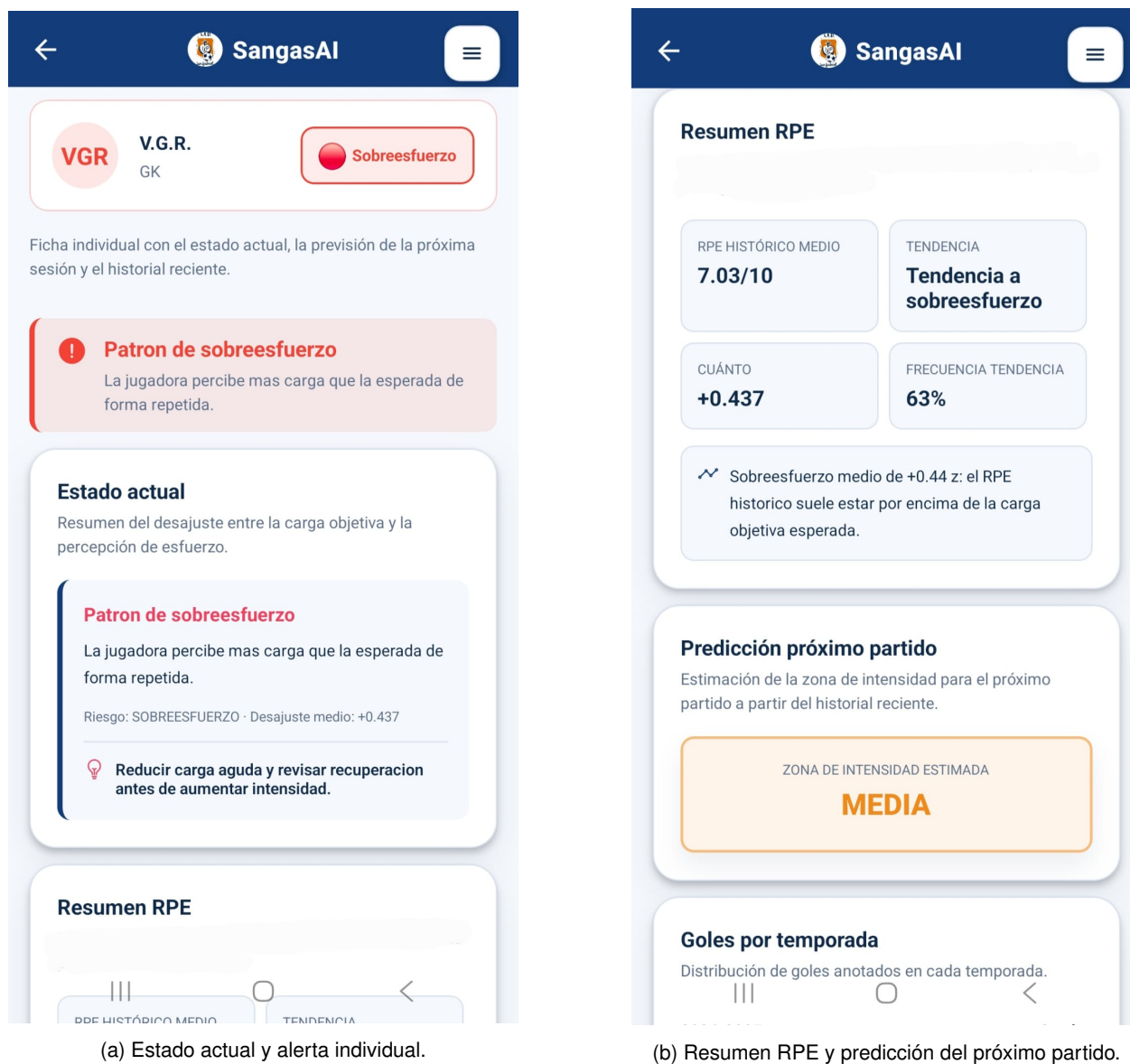


Figura A.8: Capturas del perfil individual de una jugadora.



(a) Próximo partido e histórico de encuentros.



(b) Estadísticas agregadas de temporada.

Figura A.9: Capturas de la pantalla de partidos.



Figura A.10: Captura del detalle del próximo partido.

ANÁLISIS EXPLORATORIO DE DATOS

Métrica	Tipo	Unidad	Descripción	Usada en modelado	Estado en preprocesamiento
Nombre del Jugador	Categórica	No aplica	Nombre completo de la jugadora	No	Conservada
Tiempo de Juego	Numérica	min	Tiempo total activo	Sí	Conservada
Vel. Max.	Numérica	km/h	Velocidad máxima alcanzada	Sí	Conservada
Dist. Recorrida	Numérica	m	Distancia total recorrida	Sí	Conservada
Golpes de baja int. al balón	Numérica	#	Toques suaves al balón	No	Conservada
Golpes de media int. al balón	Numérica	#	Toques moderados al balón	No	Conservada
Golpes de alta int. al balón	Numérica	#	Toques fuertes al balón	No	Conservada
Golpes de Balón	Numérica	#	Número total de contactos con el balón	Sí	Conservada
Fuerza de Pegada	Numérica	km/h	Velocidad media del balón al golpear	Sí	Conservada
Dribblings	Numérica	#	Número total de regateos realizados	No	Eliminada
Dist. en Dribbling	Numérica	m	Distancia total recorrida en regates	No	Eliminada
Carreras de Máx. Int.	Numérica	m	Distancia total recorrida en carreras a máxima intensidad	Sí	Conservada
Carreras de Máx. Int.	Numérica	#	Número de carreras realizadas a máxima intensidad	No	Conservada
Carreras de Alta Int.	Numérica	m	Distancia total recorrida en carreras a alta intensidad	Sí	Conservada
Carreras de Alta Int.	Numérica	#	Número de carreras realizadas a alta intensidad	No	Conservada
Trote	Numérica	m	Distancia total recorrida trotando	Sí	Conservada
Caminata	Numérica	m	Distancia total recorrida caminando	Sí	Conservada
Ace. Máx. Int.	Numérica	m	Distancia recorrida durante aceleraciones a máxima intensidad	No	Conservada
Ace. Máx. Int.	Numérica	#	Número de aceleraciones realizadas a máxima intensidad	No	Conservada
Desac. Máx. Int.	Numérica	m	Distancia recorrida durante desaceleraciones a máxima intensidad	No	Conservada
Desac. Máx. Int.	Numérica	#	Número de desaceleraciones realizadas a máxima intensidad	No	Conservada
Ace. Alta Int.	Numérica	m	Distancia recorrida durante aceleraciones a alta intensidad	Sí	Conservada
Ace. Alta Int.	Numérica	#	Número de aceleraciones realizadas a alta intensidad	No	Conservada
Desac. Alta Int.	Numérica	m	Distancia recorrida durante desaceleraciones a alta intensidad	Sí	Conservada
Desac. Alta Int.	Numérica	#	Número de desaceleraciones realizadas a alta intensidad	No	Conservada
Vol. de la sesión	Numérica	u.a.	Volumen total de trabajo físico	Sí	Conservada
Int. de la sesión	Numérica	u.a.	Intensidad promedio de la sesión, combinando velocidad y aceleración	Sí	Conservada
Int. de la sesión - Velocidad	Numérica	u.a.	Intensidad según velocidad	Sí	Conservada
Int. de la sesión - Aceleración	Numérica	u.a.	Intensidad según aceleraciones y desaceleraciones	Sí	Conservada
RPE	Numérica	Escala 1–10	Esfuerzo percibido subjetivo por la propia jugadora, escala 1 (ligero) - 10 (máximo)	No	Conservada
s-RPE	Numérica	RPE-min	RPE multiplicado por la duración de la sesión, estimando la carga percibida total	No	Conservada
ID del Jugador	Categórica	No aplica	Identificador único de la jugadora	No	Conservada
Posición Principal	Categórica	No aplica	Posición principal de la jugadora	No	Conservada
Fecha	Fecha	No aplica	Fecha de la sesión	No	Conservada
ID de Sesión	Categórica	No aplica	Identificador único de la sesión	No	Conservada
Nombre de Sesión	Categórica	No aplica	Nombre asignado a la sesión	No	Conservada
Tipo de Sesión	Categórica	No aplica	Clasificación en TRAINING o GAME	No	Conservada

Tabla B.1: Descripción de atributos del dataset. La abreviatura *u.a.* indica unidades arbitrarias en métricas que no presentan una unidad física directa.

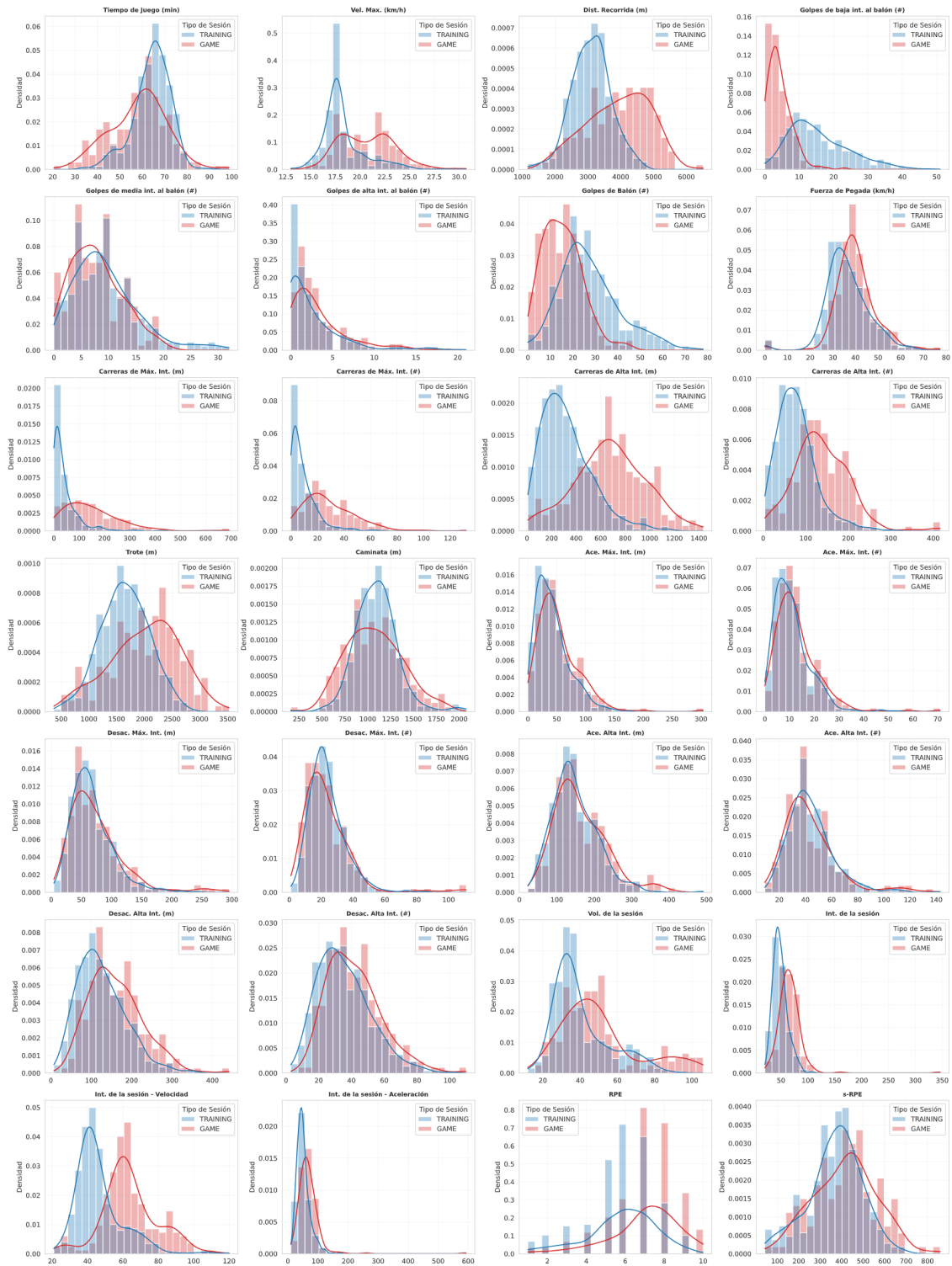


Figura B.1: Distribución de las métricas (histogramas y KDE).

Comparación GAME vs TRAINING por variable (Boxplots)

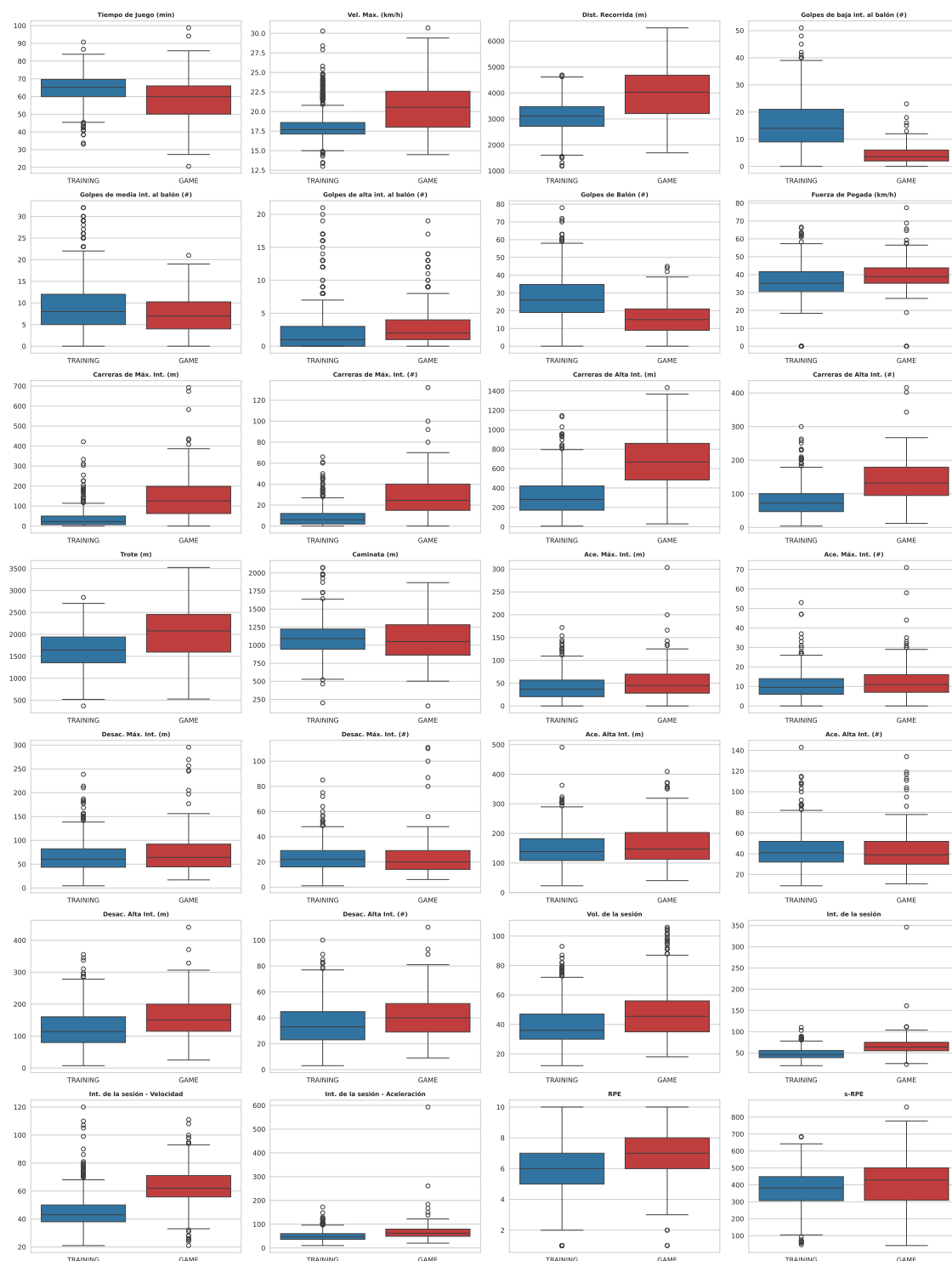


Figura B.2: Distribución de las métricas físicas mediante diagramas de caja.

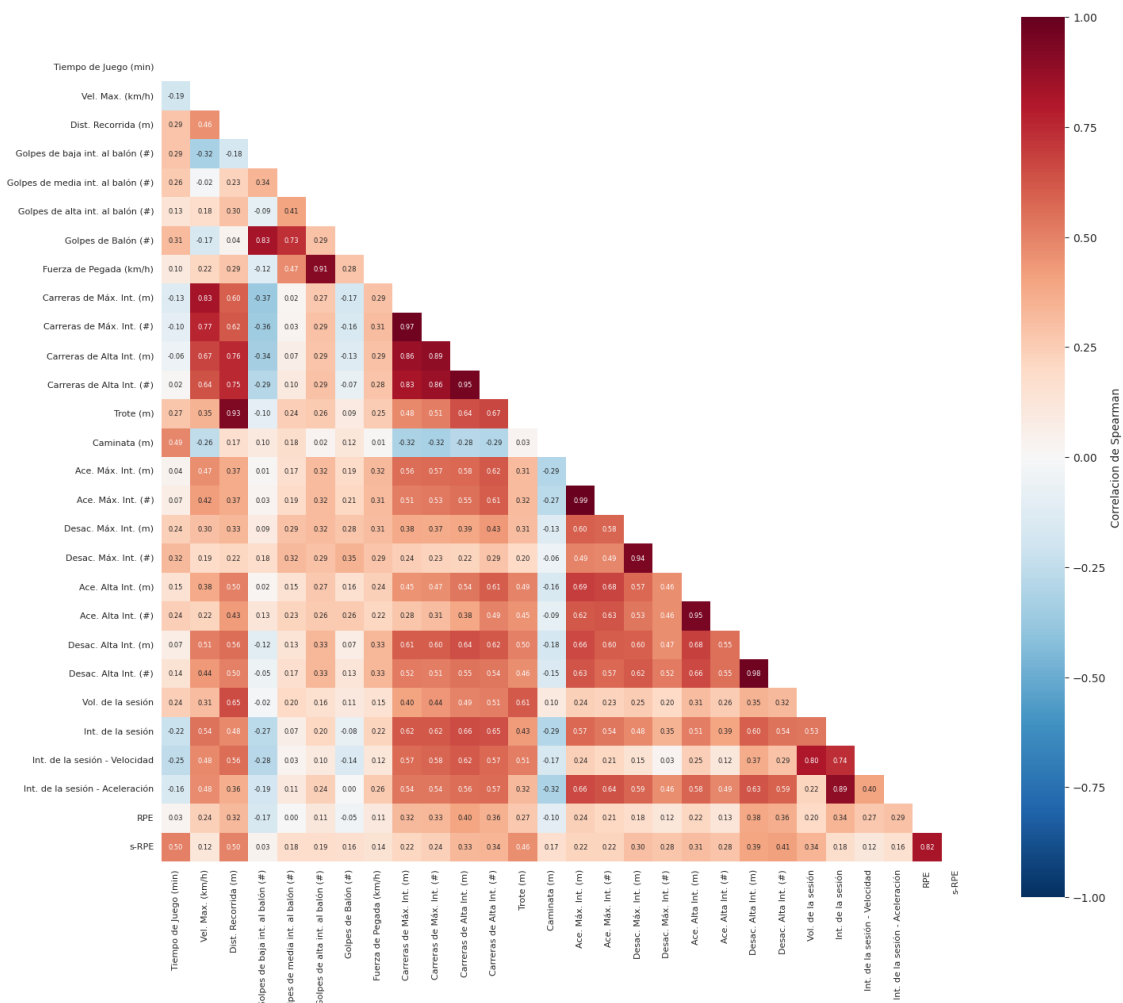


Figura B.3: Heatmap de correlación global.

COMPARACIÓN DE MODELOS Y VENTANAS TEMPORALES

Modelo	Ventana	Macro-F1	Std	Balanced Acc.	Accuracy
LogisticRegression	5	0.3541	0.0418	0.3882	0.3867
LogisticRegression	4	0.3453	0.0412	0.3542	0.3582
LogisticRegression	6	0.3299	0.0739	0.3499	0.3800
LogisticRegression	3	0.3145	0.0356	0.3374	0.3315
SVM	6	0.3097	0.1142	0.3556	0.3247
XGBoost	5	0.2999	0.0420	0.3268	0.3136
RandomForest	6	0.2971	0.0865	0.3781	0.3382
SVM	3	0.2911	0.1223	0.3353	0.3170
LogisticRegression	2	0.2903	0.0459	0.3106	0.3184
RandomForest	4	0.2871	0.0999	0.3287	0.3085
RandomForest	3	0.2864	0.0803	0.3208	0.3101
XGBoost	3	0.2824	0.0639	0.3197	0.2975
RandomForest	5	0.2727	0.0741	0.3153	0.3098
SVM	4	0.2688	0.0707	0.3042	0.2924
SVM	5	0.2687	0.1264	0.3048	0.3166
XGBoost	6	0.2676	0.0640	0.2857	0.2802
XGBoost	2	0.2626	0.0687	0.2780	0.2790
XGBoost	4	0.2519	0.0320	0.3000	0.2839
Dummy_stratified	6	0.2517	0.0478	0.2739	0.2783
Dummy_stratified	4	0.2455	0.0552	0.2630	0.2520
Dummy_stratified	5	0.2427	0.0667	0.2575	0.2530
RandomForest	2	0.2361	0.0631	0.2858	0.2769
Dummy_stratified	2	0.2315	0.0467	0.2544	0.2431
SVM	2	0.2237	0.0701	0.2745	0.2505
Dummy_stratified	3	0.2056	0.0434	0.2341	0.2203
Dummy_most_frequent	4	0.1027	0.0226	0.2500	0.2626
Dummy_most_frequent	3	0.1022	0.0218	0.2500	0.2606
Dummy_most_frequent	5	0.1018	0.0231	0.2500	0.2599
Dummy_most_frequent	2	0.1017	0.0217	0.2500	0.2590
Dummy_most_frequent	6	0.0926	0.0319	0.2500	0.2357

Tabla C.1: Resultados obtenidos para los distintos modelos y tamaños de ventana temporal, ordenados de mayor a menor valor de Macro-F1.

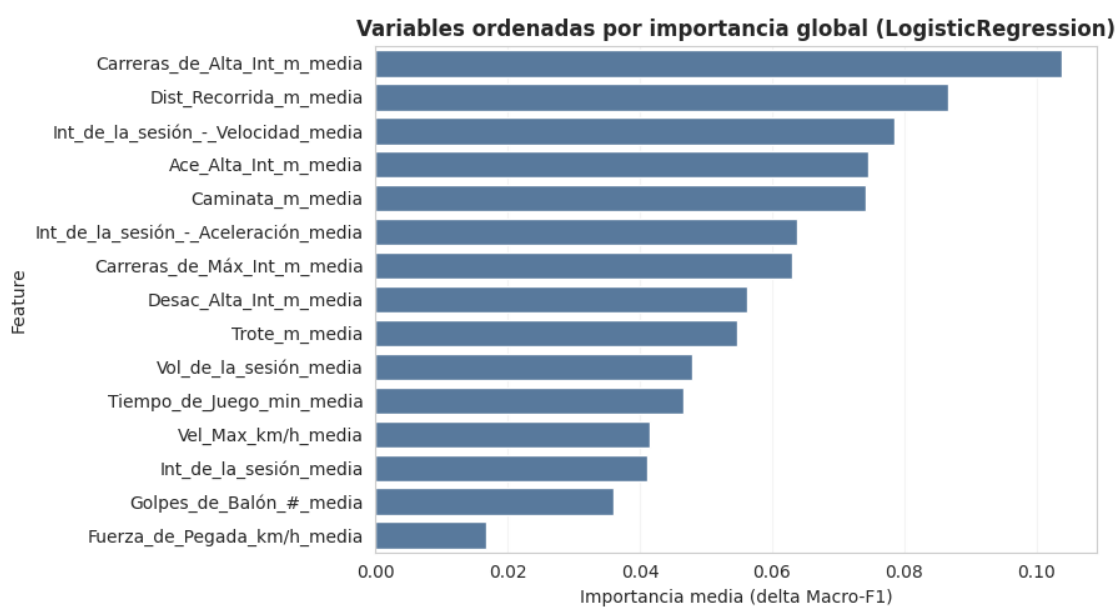


Figura C.1: Variables ordenadas por importancia global en el modelo final.



Universidad Autónoma
de Madrid