UNIVERSIDAD
NACIONAL
DE COLOMBIA

MÉTODO BASADO EN TÉCNICAS DE APRENDIZAJE DE MÁQUINAS PARA ESTIMAR LA PROBABILIDAD DE VICTORIA DE EQUIPOS EN PARTIDOS DE LA ASOCIACIÓN DE ULTIMATE FRISBEE (UFA)

José David Gutiérrez Espinosa

Facultad de Minas

Universidad Nacional de Colombia

Medellín, Colombia

2025

MÉTODO BASADO EN TÉCNICAS DE APRENDIZAJE DE MÁQUINAS PARA ESTIMAR LA PROBABILIDAD DE VICTORIA DE EQUIPOS EN PARTIDOS DE LA ASOCIACIÓN DE ULTIMATE FRISBEE (UFA)

José David Gutiérrez Espinosa

Trabajo final de maestría presentado como requisito parcial para optar al título de:

Magíster en Ingeniería – Ingeniería de Sistemas

Directores:

Pedro Atencio Ortiz, Ph.D.

John Willian Branch Bedoya, Ph.D.

Facultad de Minas

Universidad Nacional de Colombia

Medellín, Colombia

2025

Agradecimientos

Gracias.

Gracias.

Gracias a todo el que creyó en mí desde el primer día.

Cada cosa que logro a ustedes se las dedico, aunque a veces no me entiendan y hayan cosas que no explico.

Abuela, bendición, hoy cumplí otro goal… me fue cabrón.

Gracias a pá y a mami por todos los regaños, gracias a ustedes soy el mismo todos los años.

Resumen

Método basado en técnicas de aprendizaje de máquinas para estimar la probabilidad de victoria de equipos en partidos de la Asociación de Ultimate Frisbee (UFA)

Este trabajo presenta el desarrollo de un método basado en técnicas de aprendizaje de máquinas para estimar la probabilidad de victoria de equipos en partidos de la Asociación de Ultimate Frisbee (UFA). La investigación abarca la recopilación y estructuración de datos históricos de partidos, la selección de características relevantes, la implementación y optimización de modelos de clasificación, y la evaluación de su desempeño mediante métricas adecuadas. Se determinó la ventana de partidos óptima para el cálculo de estadísticas avanzadas y se compararon diversos modelos de aprendizaje de máquinas. Los resultados indican que el modelo regresión logística ofrece un equilibrio entre interpretabilidad, exactitud y calibración, logrando estimaciones confiables de la probabilidad de victoria. Este trabajo representa una contribución significativa a la analítica deportiva en el Ultimate Frisbee y sienta las bases para futuras investigaciones en el área.

Palabras clave: Analítica deportiva, probabilidad de victoria, aprendizaje de máquinas, Ultimate Frisbee.

Abstract

Method based on machine learning techniques to estimate the win probability of teams in Ultimate Frisbee Association (UFA) games.

This study presents the development of a method based on machine learning techniques to estimate the win probability of teams in Ultimate Frisbee Association (UFA) games. The research covers the collection and structuring of historical games data, the selection of relevant features, the implementation and optimization of classification models, and the evaluation of their performance using appropriate metrics. The optimal window of games for calculating advanced statistics was determined, and various machine learning models were compared. Results indicate that the logistic regression model provides a balance between interpretability, accuracy and calibration, achieving reliable win probability estimates. This study represents a significant contribution to sports analytics in Ultimate Frisbee and lays the groundwork for future research in the field.

Keywords: Sports analytics, win probability, machine learning, Ultimate Frisbee.

Tabla de contenido

Agradecimientos 3
Resumen 4
Abstract 5
Tabla de contenido 6
Lista de Figuras 8
Lista de Tablas 9
1. INTRODUCCIÓN 10
1.1. Motivación 10
1.2. Definición del problema 12
1.3. Objetivos 13
1.3.1. Objetivo general 13
1.3.2. Objetivos específicos 13
1.4. Alcance 13
1.5. Estructura del documento 14
2. MARCO TEÓRICO DE REFERENCIA 15
2.1. Asociación de Ultimate Frisbee 15
2.2. Analítica deportiva 15
2.3. Aprendizaje de máquinas 16
2.3.1. Modelos aplicados a problemas de clasificación 17
2.3.2. Métricas de evaluación para modelos de clasificación 19
2.4. Articulación del marco teórico con el método propuesto 21
3. REVISIÓN SISTEMÁTICA DE LA LITERATURA 22
3.1. Resultados de la búsqueda 23
4. MÉTODO PARA ESTIMACIÓN DE PROBABILIDAD DE VICTORIA 27
4.1. Creación del conjunto de datos 28
4.1.1. Recopilación de datos crudos 28
4.1.2. Transformación y depuración de los datos 30
4.1.3. Ingeniería de características 32

4.1.4. Descripción del conjunto de datos final 35
4.2. Diseño experimental 36
4.2.1. Selección de los modelos y sus hiperparámetros 36
4.2.2. Descripción del procedimiento de entrenamiento y validación de los modelos 38
4.3. Resultados y análisis 41
4.3.1. Determinación experimental de la ventana de partidos 41
4.3.2. Entrenamiento y validación de los modelos 41
5. CONCLUSIONES Y TRABAJO FUTURO 45
5.1. Conclusiones 45
5.2. Trabajo futuro 46
6. REFERENCIAS 47

Lista de Figuras

Figura 1: Documentos sobre analítica deportiva por año de publicación. 23
Figura 2: Documentos sobre analítica deportiva por área de trabajo. 24
Figura 3: Diagrama de flujo del método propuesto para la estimación de probabilidad de victoria. 28
Figura 4: Diagrama de flujo del proceso para la creación del conjunto de datos. . 31
Figura 5: Diagrama de cajas y bigotes para las variables de tipo promedio del conjunto final de datos. 32
Figura 6: Diagrama de cajas y bigotes para las variables de tipo porcentaje del conjunto final de datos. 33
Figura 7: Diagrama de correlación de Pearson para las variables del conjunto final de datos..... 34
Figura 8: Resultados (métricas y cantidad de registros) del análisis para la determinación de la ventana de partidos más adecuada. 40
Figura 9: Evaluación de modelos con base en diferentes métricas..... 42
Figura 10: Comparación de las matrices de confusión para los modelos regresión logística y SVM..... 44

Lista de Tablas

Tabla 1: Resumen de artículos relevantes en analítica deportiva..... 25
Tabla 2: Estadísticas básicas recolectadas para cada equipo en un partido..... 29
Tabla 3: Descripción detallada de las estadísticas avanzadas..... 30
Tabla 4: Valores a evaluar en el proceso de optimización de hiperparámetros para cada uno de los modelos..... 37

1. INTRODUCCIÓN

1.1. Motivación

La analítica deportiva se ha convertido en una herramienta esencial en el ecosistema deportivo, proporcionando beneficios que van desde la toma de decisiones estratégicas hasta la mejora del rendimiento y la prevención de lesiones, contribuyendo a elevar el nivel competitivo y la experiencia integral en el deporte [1]. Esta revolución también incluye a los aficionados y espectadores, quienes son la base económica de la profesionalización deportiva, y requiere un enfoque que mejore también su experiencia. La integración de datos y estadísticas en tiempo real durante los eventos deportivos ha mejorado la comprensión y percepción de los fanáticos [2].

El auge de las apuestas deportivas ha cambiado la forma en que los aficionados se involucran con el deporte. Los espectadores se ven motivados no solo por el fanatismo, sino también por el interés en los resultados en relación con las apuestas realizadas. Este mayor compromiso no solo beneficia a los fanáticos y apostadores, sino que también aumenta el interés general en los eventos deportivos, generando ingresos adicionales para las organizaciones deportivas.

Las apuestas deportivas legales representan una importante fuente de ingresos para las entidades deportivas, impulsando la generación de infraestructuras, programas de desarrollo y el respaldo económico para atletas destacados. La disponibilidad de apuestas también puede incrementar el interés y la asistencia a eventos deportivos, beneficiando tanto a las ligas y federaciones como a los patrocinadores. Además, el sector de las apuestas deportivas fomenta la innovación tecnológica, especialmente en el análisis de datos y la inteligencia artificial aplicada al deporte [3].

En los últimos años, la UFA ha dado pasos significativos hacia su consolidación como una liga profesional con mayor visibilidad y proyección comercial. En 2019,

algunas casas de apuestas por fuera de Estados Unidos comenzaron a aceptar apuestas sobre partidos de la UFA, marcando un primer indicio del interés del mercado en la liga como un producto con valor comercial más allá del espectáculo deportivo [4]. Este fenómeno se intensificó en 2021, cuando DraftKings, una de las plataformas de apuestas más reconocidas en Estados Unidos, incluyó oficialmente a la UFA en su oferta en determinados estados, ampliando el alcance del deporte a nuevos públicos y fortaleciendo su perfil competitivo [5].

En una de sus iniciativas más recientes, en 2022, la UFA anunció una asociación estratégica con LSports, valorada en tres millones de dólares, para la distribución de datos deportivos y el desarrollo conjunto de soluciones tecnológicas. Esta alianza no solo refleja el creciente interés por los datos en el ámbito del deporte profesional, sino que también resalta el papel que puede desempeñar la analítica avanzada en la evolución de deportes emergentes como el Ultimate Frisbee [6].

Estos desarrollos refuerzan la pertinencia de investigaciones académicas orientadas a aplicar modelos analíticos en el contexto de la UFA, con miras a aportar valor tanto a los equipos como a los distintos actores involucrados en la cadena de valor del deporte.

En este contexto de creciente profesionalización e interés comercial por la UFA, la generación de modelos que permitan estimar la probabilidad de victoria adquiere un papel estratégico. Estos métodos no solo son fundamentales en el ámbito de las apuestas deportivas, donde proporcionan las bases para establecer cuotas, sino que también resultan valiosos para los propios equipos, analistas y gestores, al ofrecer herramientas objetivas para la evaluación del rendimiento y la toma de decisiones.

La presente investigación se enmarca en esta confluencia de necesidades, buscando contribuir con soluciones analíticas aplicables al Ultimate Frisbee profesional, en línea con las tendencias observadas en otras disciplinas deportivas.

1.2. Definición del problema

La estimación de la probabilidad de victoria en Ultimate Frisbee representa un desafío analítico de gran relevancia, cuyo impacto puede transformar significativamente su ecosistema deportivo. Un método adecuado para la determinación de dicha probabilidad no solo facilitaría la toma de decisiones tácticas y estratégicas por parte de equipos y atletas, sino que también contribuiría al desarrollo de un mercado estructurado para las apuestas deportivas legales. Esto, a su vez, podría incrementar el interés de los aficionados, generando nuevas fuentes de financiación que impulsarían el crecimiento, la profesionalización y la visibilidad del deporte a nivel competitivo.

La disponibilidad de datos históricos de partidos permite explorar la construcción de un conjunto de datos estructurado y la identificación de las estadísticas más relevantes para generar predicciones sobre el posible resultado de un partido.

Un aspecto clave de este proceso es la determinación de la ventana de partidos óptima para el cálculo y análisis de las estadísticas. Esta selección debe permitir capturar de manera precisa los patrones reales de rendimiento de los equipos, garantizando así una representación fiel de su desempeño reciente. Al definir una ventana adecuada, es posible optimizar la estimación de la probabilidad de victoria.

Con base en lo anterior, el problema que se pretende abordar en este trabajo se puede definir a través de las siguientes preguntas de investigación:

- ¿Es posible construir un conjunto de datos estructurado a partir de las estadísticas de juego del histórico de partidos que se encuentra publicado en el sitio web de la Asociación de Ultimate Frisbee?
- ¿Cuáles son las estadísticas de juego más relevantes para la determinación de la probabilidad de victoria y cuál es la ventana de partidos óptima para calcularlas?
- ¿Cuál es el método que mejor permite estimar la probabilidad de victoria de un equipo en un partido de la UFA?

Se espera dar respuesta a estas preguntas de investigación a través del desarrollo y consecución de los objetivos específicos planteados a continuación.

1.3. Objetivos

1.3.1. Objetivo general

Implementar un método basado en técnicas de aprendizaje de máquinas para estimar la probabilidad de victoria de equipos en partidos de la Asociación de Ultimate Frisbee (UFA).

1.3.2. Objetivos específicos

1. Construir un conjunto de datos a partir del histórico de estadísticas de los partidos de la Asociación de Ultimate Frisbee, mediante la aplicación de técnicas de ingeniería de datos.
2. Realizar una exploración de diferentes métodos de aprendizaje de máquinas que permitan determinar la probabilidad de victoria de equipos en partidos de la Asociación de Ultimate Frisbee.
3. Validar el método seleccionado con base en métricas de desempeño.

1.4. Alcance

Este trabajo desarrolla un método basado en aprendizaje de máquinas para estimar la probabilidad de victoria en partidos de la Asociación de Ultimate Frisbee (UFA). Su principal contribución radica en la identificación de las estadísticas avanzadas más relevantes y en la determinación de la ventana de partidos óptima para el cálculo de dichas estadísticas, con el fin de mejorar la precisión y confiabilidad de las estimaciones.

Si bien existen métodos de probabilidad de victoria que se van actualizando continuamente conforme avanza la competición deportiva, este será un método estático calculado previo a la competición.

A pesar de que el Ultimate Frisbee es un deporte federado a nivel mundial por la World Flying Disc Federation (WFDF), el alcance se limitará a la UFA debido a que es la única organización en este deporte que ha recolectado datos de forma exhaustiva y sistemática durante la última década.

1.5. Estructura del documento

Este documento está organizado en seis capítulos. En el segundo capítulo se establece el marco teórico de referencia, abordando los conceptos fundamentales del Ultimate Frisbee, la analítica deportiva y los modelos de aprendizaje de máquinas. Posteriormente, el tercer capítulo presenta una revisión sistemática de la literatura en la que se analizan estudios previos relevantes sobre la estimación de probabilidades de victoria en distintos deportes. A partir del cuarto capítulo se describe el método propuesto, detallando la recopilación y estructuración de datos, la selección de características, la implementación de los modelos y la determinación de la ventana óptima de partidos para el cálculo de estadísticas. Además, se muestran los resultados obtenidos, analizando la calidad del conjunto de datos final y el desempeño de los modelos evaluados. Finalmente, el quinto capítulo presenta las conclusiones y propone líneas de investigación futuras que podrían extender el alcance y mejorar la aplicabilidad del método propuesto en este trabajo.

2. MARCO TEÓRICO DE REFERENCIA

2.1. Asociación de Ultimate Frisbee

El Ultimate Frisbee es un deporte de no-contacto y auto-arbitrado que utiliza un disco volador (Frisbee) como implemento principal. Dos equipos de 7 jugadores compiten en una cancha similar a la de fútbol americano, defendiendo cada uno su zona de anotación al final del campo. El disco se mueve por el campo a través de pases, mientras que el equipo defensor intenta impedir que llegue a su zona de anotación. El jugador que tiene el disco no puede moverse con él y tiene un máximo de 10 segundos para hacer un pase. Los puntos se marcan al recibir el disco en la zona de anotación del equipo contrario [7].

A medida que el Ultimate Frisbee ha venido ganado popularidad, se ha hecho necesario contar con organizaciones deportivas que regulen y promuevan su desarrollo y crecimiento. En este sentido, la Asociación de Ultimate Frisbee (UFA por sus siglas en inglés) se fundó en 2012 con el objetivo de ofrecer un entorno deportivo profesional para los atletas que pertenecen a ella. Actualmente, la UFA está conformada por 24 equipos y su misión es aumentar la visibilidad de uno de los deportes con mayor crecimiento a nivel mundial en las últimas décadas [8]. Además de fomentar la práctica del deporte, la UFA ha contribuido a la recopilación y análisis de datos sobre el rendimiento de los equipos y jugadores, lo que ha fomentado el interés en la aplicación de múltiples métodos de analítica deportiva.

2.2. Analítica deportiva

La analítica deportiva es una disciplina que aplica métodos estadísticos y computacionales al estudio del rendimiento de los atletas y equipos en diversas disciplinas. Su objetivo es mejorar la toma de decisiones estratégicas mediante el análisis de datos históricos y en tiempo real. Con el crecimiento de la tecnología y el acceso a grandes volúmenes de datos, los equipos deportivos han comenzado a

adoptar herramientas avanzadas para optimizar el rendimiento individual y colectivo, evaluar estrategias de juego y mejorar la preparación de los deportistas [2].

En paralelo, el auge de la analítica deportiva ha impulsado el desarrollo de apuestas deportivas legales, en las cuales se emplean modelos probabilísticos para estimar los posibles resultados de eventos deportivos. Estas apuestas se basan en el análisis de datos históricos, estadísticas de rendimiento y factores contextuales para calcular la probabilidad de distintos resultados. La aplicación de modelos de predicción permite a las casas de apuestas y a los apostadores minimizar la incertidumbre y mejorar la precisión de sus estimaciones, convirtiendo el análisis de datos en una herramienta clave dentro de este ámbito [3].

La probabilidad de victoria es una medida estadística que evalúa las probabilidades de que un equipo o atleta obtenga el éxito en un evento deportivo o competencia. Se basa en la aplicación de métodos matemáticos que consideran diversos factores relevantes, como el rendimiento histórico del equipo o atleta en eventos similares, su nivel de experiencia y desempeño en comparación con los contrincantes, las condiciones específicas del juego como el clima o la ubicación, el estado físico de los deportistas, la estrategia empleada y otros aspectos específicos de la competencia que puedan influir en el resultado final.

2.3. Aprendizaje de máquinas

Para calcular esta probabilidad de manera precisa, se requieren herramientas que sean capaces de identificar patrones en los datos y generar predicciones basadas en ellos. En este contexto, el aprendizaje de máquinas se ha convertido en una herramienta fundamental, ya que permite desarrollar modelos capaces de aprender de los datos históricos y generar estimaciones más precisas sobre la probabilidad de victoria. El aprendizaje de máquinas es un campo de la inteligencia artificial que se enfoca en desarrollar algoritmos y métodos computacionales capaces de aprender patrones y tomar decisiones a partir de datos. Estos métodos son

entrenados con conjuntos de datos históricos para identificar relaciones y estructuras que luego pueden aplicarse a nuevos datos para hacer predicciones o tomar decisiones automatizadas [9].

Dentro del aprendizaje de máquinas, los métodos supervisados se caracterizan por el uso de datos etiquetados, lo que permite que el modelo aprenda la relación entre las variables de entrada y salida. Uno de los enfoques más empleados en este tipo de problemas es la clasificación, en la cual se busca predecir a qué clase pertenece una observación, a partir de un conjunto de características.

Un modelo de clasificación es un tipo de modelo matemático que permite asignar una etiqueta categórica con base en un conjunto de variables. Estos modelos se entrenan utilizando un conjunto de datos histórico que contiene tanto las variables de entrada como las etiquetas reales, lo cual permite que el algoritmo identifique patrones en los datos [9]. Dependiendo de la naturaleza de los datos y la complejidad del problema, existen diferentes tipos de modelos de clasificación que pueden ser empleados.

2.3.1. Modelos aplicados a problemas de clasificación

A continuación, se introducen y se definen brevemente algunos de los modelos clásicos de aprendizaje de máquinas aplicados a problemas de clasificación.

La regresión logística es un modelo estadístico utilizado para predecir la probabilidad de pertenencia de una observación a una clase binaria. Este modelo se basa en la transformación logística de una combinación lineal de las variables independientes, lo cual permite obtener una salida en el rango [0, 1]. La predicción final se realiza asignando la clase correspondiente según un umbral definido sobre dicha probabilidad [10].

Las máquinas de soporte vectorial (SVM, por sus siglas en inglés) son modelos de clasificación supervisada que construyen un hiperplano de separación entre clases. El modelo busca maximizar la distancia entre este hiperplano y las observaciones

más cercanas, conocidas como vectores de soporte. En problemas no linealmente separables, se utilizan funciones *kernel* que transforman los datos originales a un espacio de mayor dimensión donde es posible realizar una separación lineal [10].

Los árboles de decisión son modelos que representan una estructura jerárquica compuesta por nodos de decisión y nodos hoja. En cada nodo se establece una condición sobre una variable de entrada, y a partir de ella se divide el conjunto de datos. Este proceso se repite de forma recursiva hasta llegar a una predicción en las hojas del árbol. En clasificación, cada hoja representa una clase determinada [10]. Uno de los desafíos de los árboles de decisión es el sobreajuste, ya que los árboles profundos tienden a ajustarse demasiado a los datos de entrenamiento. Para evitar esto, se pueden emplear técnicas como limitar la profundidad máxima del árbol.

Random Forest es un modelo de ensamble que combina múltiples árboles de decisión con el fin de mejorar la estabilidad y precisión de las predicciones. A diferencia de un solo árbol de decisión, que puede ser altamente sensible a los datos de entrenamiento, Random Forest reduce la varianza del modelo mediante la agregación de múltiples árboles. En este algoritmo se generan múltiples subconjuntos de datos a partir del conjunto original utilizando muestreo con reemplazo. Para cada subconjunto, se entrena un árbol de decisión independiente, pero en cada división del árbol, solo se considera un subconjunto aleatorio de las variables en lugar de todas las disponibles [10].

La clase final predicha corresponde a la clase más frecuente entre los árboles. La combinación de árboles independientes permite que el modelo sea más robusto frente a valores atípicos y menos propenso al sobreajuste. Sin embargo, una desventaja de Random Forest es que puede ser computacionalmente costoso, especialmente cuando se entrenan un gran número de árboles.

El modelo de *k* vecinos más cercanos (*k*-NN, por sus siglas en inglés) clasifica una observación según la clase predominante entre sus *k* observaciones más próximas en el espacio de características. La proximidad entre observaciones se define mediante una métrica de distancia, generalmente la distancia euclidiana [10].

Gradient Boosting es un método de ensamble que combina secuencialmente múltiples clasificadores débiles, como árboles de decisión de baja profundidad. En cada iteración, se entrena un nuevo clasificador que intenta corregir los errores del conjunto anterior. La predicción final se obtiene mediante la suma ponderada de las salidas de todos los clasificadores construidos [10].

2.3.2. Métricas de evaluación para modelos de clasificación

Para evaluar la efectividad de los modelos en la clasificación de observaciones, se utilizan métricas que cuantifican el grado de acierto del modelo con respecto a las clases reales observadas.

La exactitud (*accuracy*) se define como el número de predicciones correctas dividido por el número total de observaciones evaluadas. Esta métrica refleja el porcentaje global de aciertos del modelo. En la Ecuación 1 se muestra la definición del *accuracy*, donde *TP* representa los verdaderos positivos, *TN* los verdaderos negativos, *FP* los falsos positivos y *FN* los falsos negativos.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \textbf{\textit{Ecuación 1}}$$

El *F1 score* es una métrica que representa la media armónica entre la precisión (*precision*) y la sensibilidad (*recall*). La precisión se define como la proporción de verdaderos positivos entre las observaciones clasificadas como positivas por el modelo, mientras que la sensibilidad se define como la proporción de verdaderos positivos entre las observaciones que realmente pertenecen a la clase positiva.

En las Ecuaciones 2, 3 y 4 se pueden observar las formulaciones del *F1 score*, *precision* y *recall*, respectivamente.

$$F1\ score = 2 * \frac{precision * recall}{precision + recall} \qquad \textbf{\textit{Ecuación 2}}$$

$$precision = \frac{TP}{TP + FP} \qquad \textbf{\textit{Ecuación 3}}$$

$$recall = \frac{TP}{TP + FN} \qquad \textbf{\textit{Ecuación 4}}$$

El *Brier score* se define como la media del error cuadrático entre las probabilidades predichas por el modelo y los valores reales de la clase, codificados como 0 o 1. Esta métrica evalúa la calidad de la calibración de las probabilidades asignadas por el modelo. En la Ecuación 5 se muestra la definición del *Brier score* donde N es el número de observaciones, y_i es el valor de la probabilidad predicha y $\hat{y}_i$ es el valor real correspondiente, codificado como 0 o 1.

$$Brier\ score = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad \textbf{\textit{Ecuación 5}}$$

La métrica ROC AUC (área bajo la curva ROC) se basa en la construcción de una curva que relaciona la tasa de verdaderos positivos y la tasa de falsos positivos para diferentes umbrales de clasificación. El valor del área bajo esta curva indica la capacidad del modelo para distinguir entre clases.

2.4. Articulación del marco teórico con el método propuesto

El contenido expuesto en este capítulo proporciona los fundamentos conceptuales necesarios para el desarrollo metodológico presentado en los capítulos siguientes. La descripción del Ultimate Frisbee y su contexto competitivo dentro de la UFA permite entender las particularidades del deporte y justifica la pertinencia de centrar el análisis exclusivamente en esta liga. Adicionalmente, se establece el marco teórico de la analítica deportiva y su vínculo con las apuestas legales y la toma de decisiones estratégicas, elementos claves que motivan la estimación de la probabilidad de victoria en este entorno.

Por otro lado, la introducción al aprendizaje de máquinas, junto con la descripción de modelos de clasificación y métricas de evaluación, ofrece una base sólida para la implementación técnica del método propuesto. Estos elementos teóricos permiten sustentar las decisiones metodológicas adoptadas en la construcción del conjunto de datos, la selección de variables, la elección de modelos y la interpretación de los resultados. En conjunto, este marco teórico de referencia articula los elementos deportivos, analíticos y computacionales que dan forma al enfoque integral de este trabajo.

3. REVISIÓN SISTEMÁTICA DE LA LITERATURA

Para garantizar un proceso riguroso y transparente en la recopilación y selección de estudios relevantes, se adoptó la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). PRISMA es un marco metodológico ampliamente reconocido que orienta la realización de revisiones sistemáticas mediante una estructura clara, reproducible y comprensible. Esta metodología propone un flujo de trabajo que incluye la identificación de documentos, la eliminación de duplicados, la selección según criterios predefinidos, la evaluación de elegibilidad y, finalmente, la inclusión de los estudios pertinentes para el análisis. El uso de PRISMA permite no solo mejorar la calidad de la revisión realizada, sino también facilitar la trazabilidad del proceso de búsqueda y asegurar que los resultados sean consistentes con los objetivos del presente trabajo.

Los criterios utilizados para seleccionar los documentos incluidos en esta revisión sistemática fueron los siguientes:

- Áreas de trabajo: Se consideraron publicaciones relacionadas con analítica deportiva, estimación de probabilidad de victoria, aprendizaje de máquinas y Ultimate Frisbee.
- Idioma: Se incluyeron únicamente documentos escritos en inglés y español, por ser los idiomas manejados en el desarrollo de este trabajo y los más frecuentes en la literatura académica disponible sobre el tema.
- Año de publicación: Se limitó la búsqueda a publicaciones realizadas en la última década, comprendida entre los años 2015 y 2024, con el objetivo de asegurar la pertinencia y actualidad de los enfoques metodológicos y tecnológicos analizados.

Se consultaron bases de datos académicas indexadas de alta calidad como Scopus, Web of Science y Science Direct, las cuales proporcionan literatura revisada por pares. También se utilizaron fuentes no indexadas, como Google Scholar, con el fin de ampliar el campo de revisión y obtener estudios relevantes que puedan haber sido publicados en conferencias, repositorios institucionales o literatura técnica no indexada.

3.1. Resultados de la búsqueda

Los resultados obtenidos reflejan una tendencia creciente en la publicación de investigaciones relacionadas con la analítica deportiva durante los últimos diez años. Esta evolución se puede observar claramente en la Figura 1, donde se evidencia un incremento sostenido en el número de publicaciones anuales, lo cual sugiere un interés creciente por parte de la comunidad científica en el aprovechamiento de datos y herramientas computacionales aplicadas al ámbito deportivo.

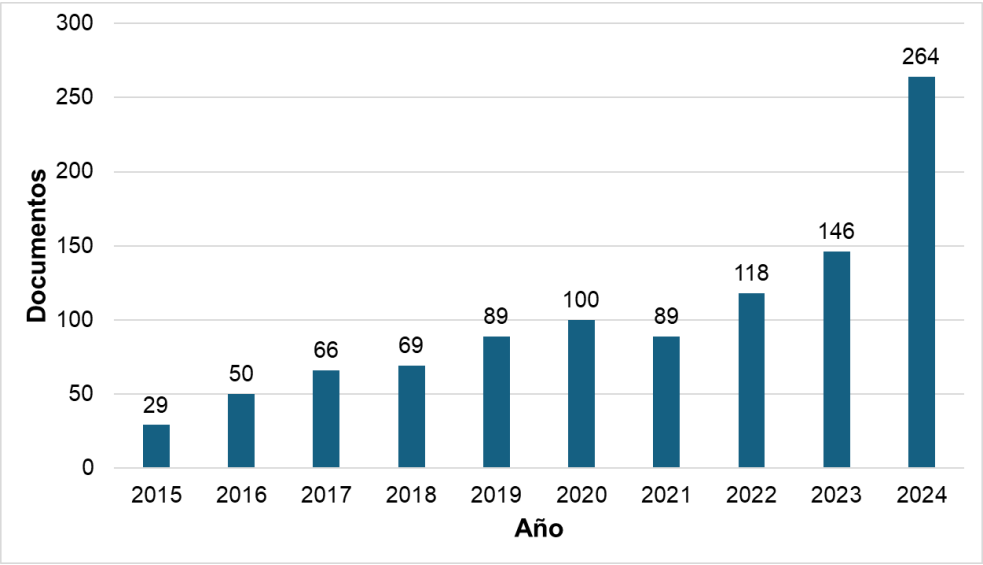

Figura 1: Documentos sobre analítica deportiva por año de publicación.

Adicionalmente, como se muestra en la Figura 2, la mayoría de los estudios identificados sobre analítica deportiva se ubican dentro del campo de las ciencias de la computación, lo que valida la hipótesis de que los avances tecnológicos están desempeñando un rol central en la transformación del deporte. En particular, el uso de técnicas de aprendizaje de máquinas y estadística avanzada permite abordar preguntas complejas sobre el rendimiento, la predicción de resultados y la toma de decisiones estratégicas. Esta evolución apunta hacia un ecosistema deportivo cada vez más orientado a la toma de decisiones basadas en datos.

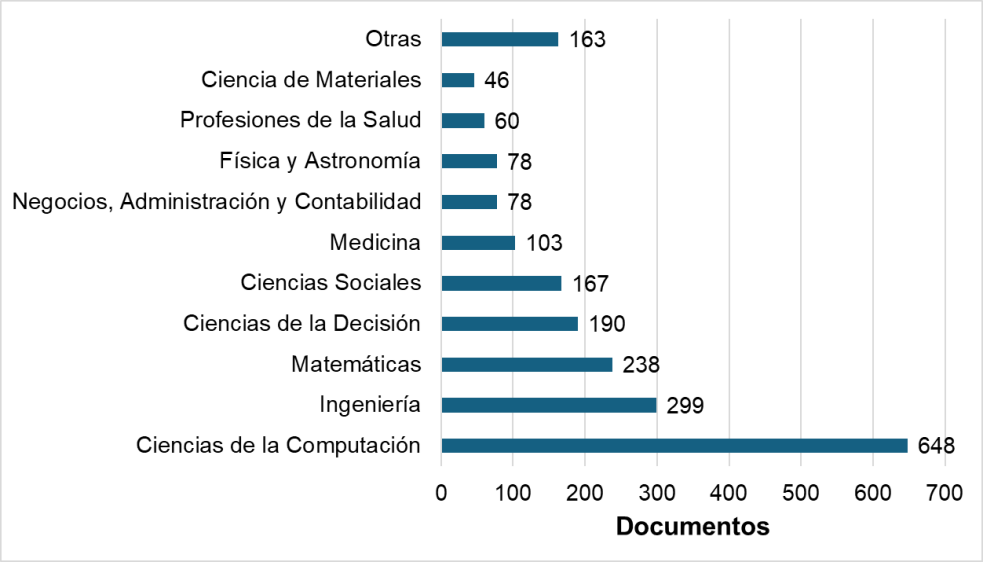

Figura 2: Documentos sobre analítica deportiva por área de trabajo.

A continuación, en la Tabla 1, se describen los artículos más relevantes para el presente trabajo que resultaron de esta revisión sistemática de la literatura:

Tabla 1: Resumen de artículos relevantes en analítica deportiva.

Título	Objetivo	Metodología	Técnicas	Resultados
Using Conformal Win Probability to Predict the Winners of the Canceled 2020 NCAA Basketball Tournaments (2024) [11]	Estimar las probabilidades de victoria en torneos NCAA cancelados por COVID-19.	Análisis retrospectivo con métodos conformales y comparativos con regresión lineal y regresión logística.	Regresión logística, regresión lineal, distribuciones predictivas conformales.	Las probabilidades basadas en métodos conformales ofrecen mayor flexibilidad y menos supuestos que los métodos tradicionales.
Big Ideas in Sports Analytics and Statistical Tools for Their Investigation (2023) [12]	Exponer conceptos clave de analítica deportiva en múltiples disciplinas.	Revisión conceptual y comparativa entre enfoques en diferentes deportes.	Modelos estadísticos, analítica descriptiva.	Se establecen similitudes y diferencias en enfoques de análisis.
The Influence of Tactical and Match Context on Player Movement in Football (2022) [13]	Analizar métricas de movimiento de jugadoras considerando el contexto del juego.	Análisis estadístico de datos ópticos de partidos del Mundial Femenino FIFA 2019.	Prueba de Kolmogorov–Smirnov, distancia de Wasserstein.	Existen diferencias significativas en patrones de movimiento según la fase del juego.
A Bayesian Approach to In-Game Win Probability in Soccer (2021) [14]	Desarrollar un modelo de probabilidad de victoria en tiempo real para fútbol.	Modelado bayesiano sobre datos de 8 temporadas de ligas europeas.	Enfoque bayesiano, predicción probabilística con variables contextuales.	El modelo mejora la experiencia del aficionado y el análisis táctico en tiempo real.
Development of Individual Value of Offense/Defense Contribution in Basketball Teams (2023) [15]	Cuantificar el impacto individual de jugadores en la probabilidad de victoria en baloncesto.	Cálculo de un índice basado en la evolución de la probabilidad de victoria durante jugadas.	Modelos estadísticos de contribución, análisis comparativo de alineaciones.	Se propone el indicador “Valor de Contribución” (VC), validado con datos de la liga B.LEAGUE.

Exploring Predictors of Team Success in Ultimate Frisbee: An Analysis of Game Statistics for Stanford Women's Ultimate (2017) [16]	Identificar patrones y predecir resultados en Ultimate Frisbee femenino universitario.	Clasificación de jugadoras y oponentes con K-Means; regresión logística para predicción.	K-Means, regresión logística.	Características como precisión en lanzamientos influyen significativamente en las probabilidades de victoria.

Del análisis de los estudios revisados pueden extraerse varias conclusiones relevantes para el desarrollo del presente trabajo. En primer lugar, se identifica un uso generalizado de las técnicas de aprendizaje de máquinas en la predicción de resultados deportivos, especialmente en disciplinas donde se dispone de conjuntos de datos históricos estructurados. En este sentido, el Ultimate Frisbee representa una oportunidad valiosa para aplicar estos enfoques, dada la creciente disponibilidad de datos generados por ligas como la UFA.

En segundo lugar, se observa que muchas de las investigaciones revisadas se enfocan en modelos de predicción para deportes con gran volumen de datos o alta popularidad mediática, como el fútbol, el baloncesto o el béisbol. En contraste, disciplinas emergentes como el Ultimate Frisbee han sido escasamente exploradas desde un enfoque computacional, lo que representa una limitación importante en la literatura actual y al mismo tiempo una oportunidad para el presente trabajo.

Por último, los trabajos analizados destacan la importancia de contar con conjuntos de datos bien estructurados y relevantes para lograr predicciones precisas. Elementos como la selección de características, la ventana temporal de análisis, y la validación de modelos, son aspectos recurrentes en las metodologías más exitosas, y han sido considerados en la propuesta de este trabajo.

4. MÉTODO PARA ESTIMACIÓN DE PROBABILIDAD DE VICTORIA

En este capítulo se describe el método propuesto para la estimación de la probabilidad de victoria en partidos de la UFA mediante técnicas de aprendizaje de máquinas. Se presentan los procesos de recopilación y procesamiento de los datos, la selección de las características más relevantes, la implementación de distintos modelos y la evaluación de su desempeño. Además, se discute la importancia de definir una ventana de partidos óptima para el cálculo de estadísticas avanzadas y su impacto en la calidad de las estimaciones.

En la Figura 3 se presenta un diagrama de flujo que sintetiza el método desarrollado en este capítulo para la estimación de probabilidad de victoria en partidos de la UFA.

El proceso inicia con la consulta de estadísticas básicas provenientes del sitio oficial de la liga, las cuales son utilizadas como insumo para generar estadísticas avanzadas. Esta fase incluye la depuración de datos y la selección de características, enmarcadas dentro de una ventana de partidos previamente definida. A partir de esta estructura se implementan modelos base que permiten evaluar tanto el desempeño de los algoritmos como la influencia del tamaño de la ventana de partidos, lo cual conduce a la selección de una configuración óptima.

Posteriormente, se realiza la selección de modelos destacados, sobre los cuales se lleva a cabo una optimización de hiperparámetros. Finalmente, se implementan y evalúan los modelos optimizados, cerrando así un ciclo metodológico fundamentado en la integración de analítica deportiva, ingeniería de características y aprendizaje de máquinas.

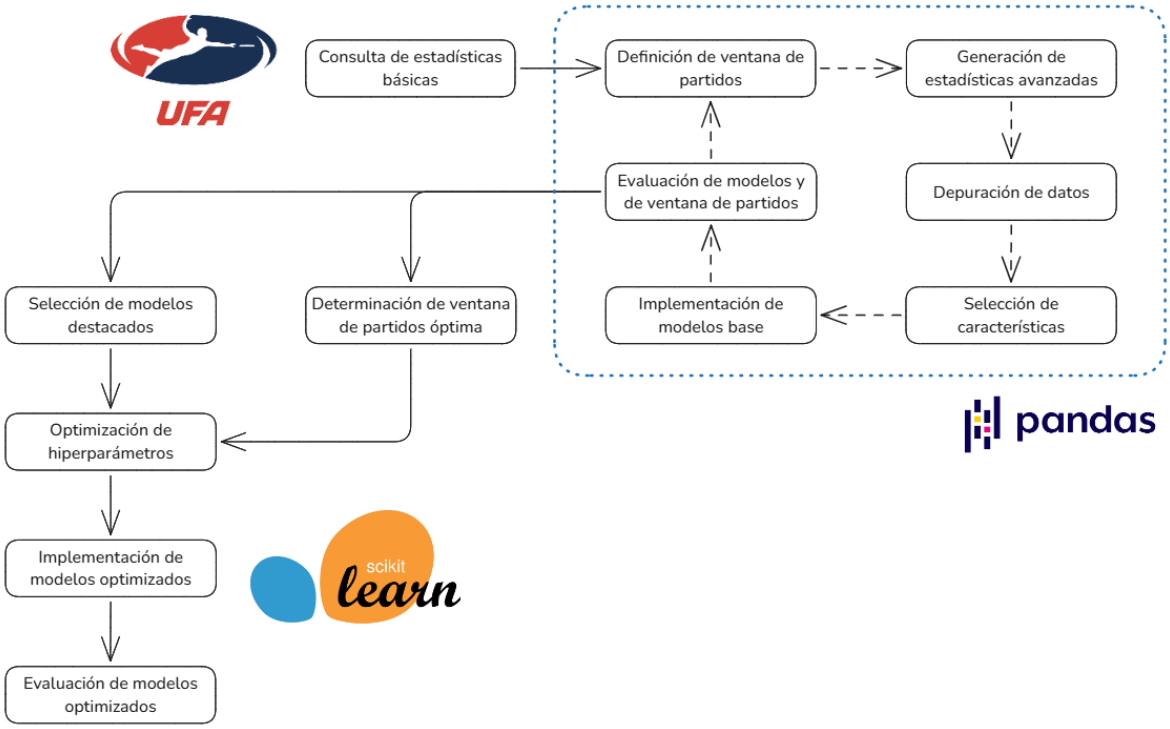

Figura 3: Diagrama de flujo del método propuesto para la estimación de probabilidad de victoria.

4.1. Creación del conjunto de datos

Este conjunto de datos se basa en las estadísticas históricas de los partidos disputados en la Asociación de Ultimate Frisbee en el período comprendido entre 2014 y 2024, excluyendo la temporada 2020, la cual no pudo llevarse a cabo debido a la pandemia de COVID-19.

4.1.1. Recopilación de datos crudos

El proceso inicia con el consumo iterativo del microservicio de información de partidos (https://www.backend.ufastats.com/web-v1/games) alojado en el backend del sitio web de la UFA, a través del cual se recupera el identificador de cada partido junto con información básica de este.

A continuación, se realiza otro consumo iterativo del microservicio de estadísticas y resultados de cada partido (https://www.backend.ufastats.com/web-v1/game-stats) teniendo como base de consulta cada uno de los identificadores obtenidos en el paso anterior. De este proceso se obtiene la información detallada sobre el desempeño, reflejado en sus estadísticas de juego, de los equipos en cada partido.

Dado que estos datos se presentan en un formato donde cada registro corresponde a las **estadísticas de ambos equipos** en un partido, es necesario realizar una fase de manipulación y organización, utilizando técnicas de ingeniería de datos, que permita construir un conjunto de datos estructurado, en el cual cada registro represente las **estadísticas de un solo equipo** en un determinado partido.

Las estadísticas básicas que se obtienen a partir de la información recuperada de los servidores de la UFA se presentan en la Tabla 2.

Tabla 2: Estadísticas básicas recolectadas para cada equipo en un partido.

Estadística básica	Descripción	Rango
completions	Cantidad de pases completados	[13 – 480]
throwingAttempts	Cantidad de pases intentados	[41 – 487]
hucksCompleted	Cantidad de pases largos completados	[0 – 19]
hucksAttempted	Cantidad de pases largos intentados	[2 – 26]
blocks	Cantidad de bloqueos (defensas)	[0 – 39]
turnovers	Cantidad de pérdidas de posesión	[3 – 73]
oLineScores	Anotaciones realizadas por la línea ofensiva	[3 – 28]
oLinePoints	Cantidad de puntos jugados por la línea ofensiva	[6 – 47]
oLinePossessions	Cantidad de posesiones jugadas por la línea ofensiva	[8 – 71]
dLineScores	Anotaciones realizadas por la línea defensiva	[0 – 34]
dLinePoints	Cantidad de puntos jugados por la línea defensiva	[6 – 47]
dLinePossessions	Cantidad de posesiones jugadas por la línea defensiva	[1 – 53]
redZoneScores	Anotaciones realizadas desde las últimas 15 yardas	[3 – 26]
redZonePossessions	Cantidad de posesiones en las últimas 15 yardas	[5 – 30]
win	1 = Victoria | 0 = Derrota	No aplica

4.1.2. Transformación y depuración de los datos

Una vez establecida esta estructura, se procede con el cálculo de estadísticas avanzadas a partir de las estadísticas básicas previamente obtenidas, con el objetivo de obtener información más representativa del rendimiento y desempeño de los equipos en una ventana de tiempo (ventana de partidos previos) determinada. Este procedimiento implica el cálculo de promedios y porcentajes que permitan agregar las estadísticas totales de modo que puedan compararse en un marco de referencia común.

En la Tabla 3 se presenta una descripción detallada de cada una de las estadísticas avanzadas, especificando su significado, forma en que fueron calculadas y rango de valores.

Tabla 3: Descripción detallada de las estadísticas avanzadas.

Estadística avanzada	Descripción	Fórmula de cálculo	Rango
scores_per_game	Promedio de anotaciones por partido	$\frac{\sum oLineScores + \sum dLineScores}{Ventana\ de\ partidos}$	[11,4 – 32,2]
completion_perc	Porcentaje de pases completados	$\frac{\sum completions}{\sum throwingAttempts}$	[0,74 – 0,97]
huck_perc	Porcentaje de pases largos completados	$\frac{\sum hucksCompleted}{\sum hucksAttempted}$	[0,38 – 0,94]
blocks_per_game	Promedio de bloqueos (defensas) por partido	$\frac{\sum blocks}{Ventana\ de\ partidos}$	[5,4 – 21,8]
turnovers_per_game	Promedio de pérdidas de posesión por partido	$\frac{\sum turnovers}{Ventana\ de\ partidos}$	[9,0 – 46,4]
hold_perc	Porcentaje de eficacia ofensiva	$\frac{\sum oLineScores}{\sum oLinePoints}$	[0,36 – 0,87]
OLC_perc	Porcentaje de eficiencia ofensiva	$\frac{\sum oLineScores}{\sum oLinePossessions}$	[0,22 – 0,75]
break_perc	Porcentaje de eficacia defensiva	$\frac{\sum dLineScores}{\sum dLinePoints}$	[0,05 – 0,52]
DLC_perc	Porcentaje de eficiencia defensiva	$\frac{\sum dLineScores}{\sum dLinePossessions}$	[0,08 – 0,79]

RZC_perc	Porcentaje de eficiencia en zona roja	$\frac{\sum redZoneScores}{\sum redZonePossessions}$	[0,57 – 0,95]
winning_perc	Porcentaje de victorias	$\frac{\sum win}{Ventana\ de\ partidos}$	[0,00 – 1,00]

Finalmente, se lleva a cabo un proceso de limpieza y preprocesamiento del conjunto de datos, que incluye la detección de valores inconsistentes, la imputación de datos faltantes y la normalización de variables. Como resultado de este proceso, el tamaño del conjunto de datos pasa de 3174 registros y 12 atributos a 2019 registros y 10 atributos.

Dentro del proceso de consolidación del conjunto de datos, fue necesario descartar los atributos ***huck_perc*** y ***RZC_perc***, ya que las estadísticas base utilizadas para su cálculo solo comenzaron a recolectarse a partir del año 2021. Mantener estas variables habría implicado la eliminación de una cantidad significativa de registros previos, reduciendo considerablemente el tamaño del conjunto de datos y afectando la capacidad de los modelos para generalizar sobre un mayor rango de temporadas.

En la Figura 4 se muestra un diagrama de flujo del proceso descrito anteriormente para la creación del conjunto de datos.

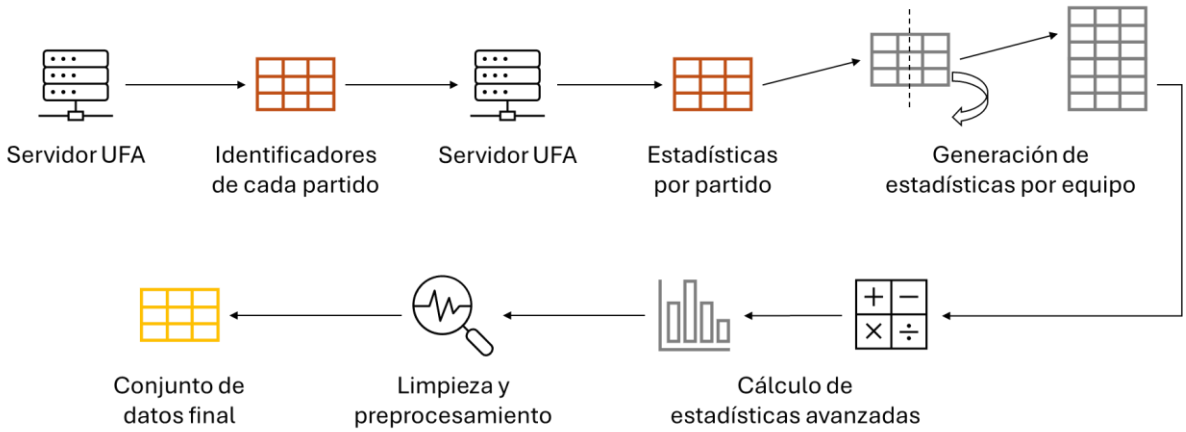

Figura 4: Diagrama de flujo del proceso para la creación del conjunto de datos.

4.1.3. Ingeniería de características

Para garantizar la relevancia de las variables en los modelos, se llevó a cabo un análisis de varianza mediante un diagrama de cajas y bigotes para cada variable. Este procedimiento permitió identificar aquellas variables con una baja variabilidad, las cuales podrían no contribuir significativamente al desempeño de los modelos. Al descartar dichas variables, se busca mejorar la capacidad de los modelos para capturar patrones relevantes en los datos y evitar la inclusión de información poco significativa.

A continuación, en la Figura 5 y en la Figura 6, se presentan los diagramas de cajas y bigotes para cada una de las variables consideradas, facilitando el análisis de su distribución y variabilidad.

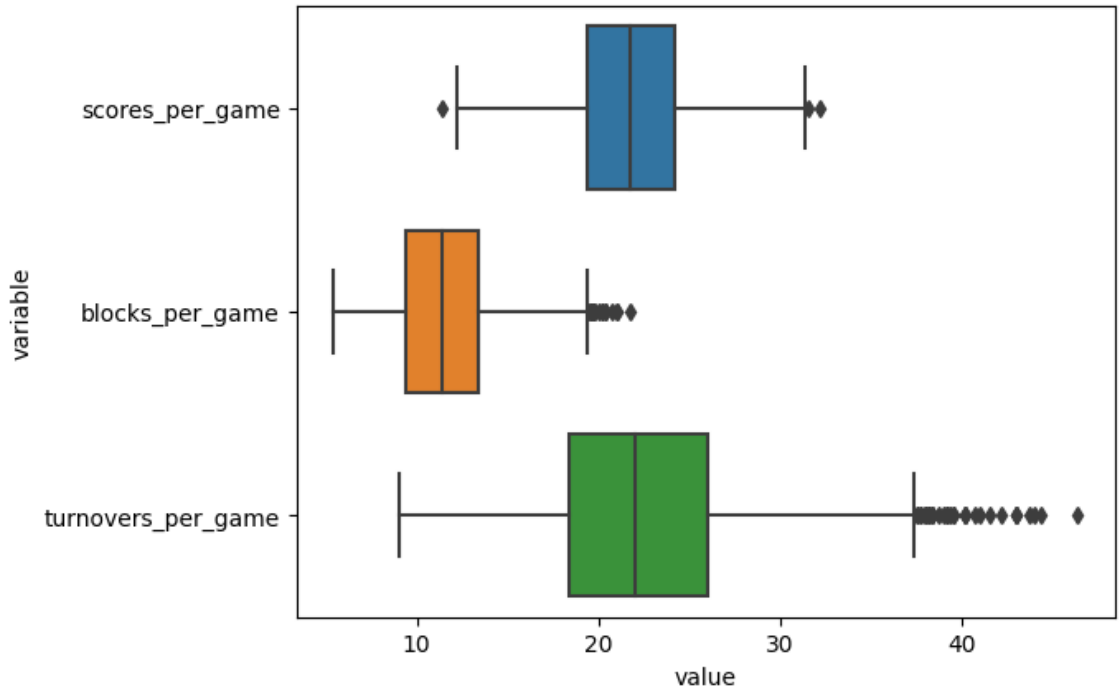

Figura 5: Diagrama de cajas y bigotes para las variables de tipo promedio del conjunto final de datos.

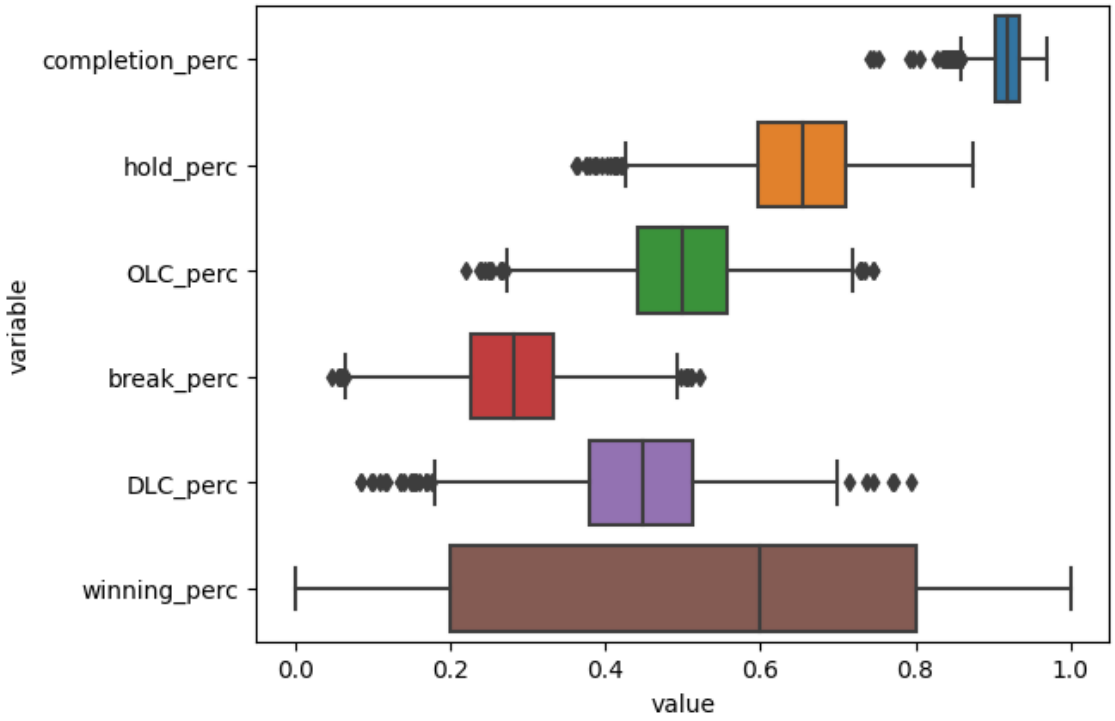

Figura 6: Diagrama de cajas y bigotes para las variables de tipo porcentaje del conjunto final de datos.

A partir del diagrama de cajas y bigotes mostrado en la Figura 6, es posible observar que la variable ***completion_perc*** presenta un rango intercuartílico reducido y bigotes cortos, lo que indica una baja variabilidad en sus valores. Debido a esta falta de dispersión, se concluyó que su inclusión en el modelo no aportaría información relevante para la predicción, por lo que se procedió a descartarla.

Adicionalmente, se realizó un análisis de correlación de Pearson con el propósito de evaluar la relación entre cada par de variables y detectar posibles problemas de multicolinealidad. Se descartaron aquellas variables más complejas cuya correlación superara un umbral de 0.8. Lo anterior significa que, al presentarse una correlación superior al umbral de 0.8 entre un determinado par de variables, se preservó la variable más intuitiva que permitiera seguir explicando el juego de forma simple.

En la Figura 7 se presenta el diagrama de correlación de Pearson para las variables del conjunto de datos.

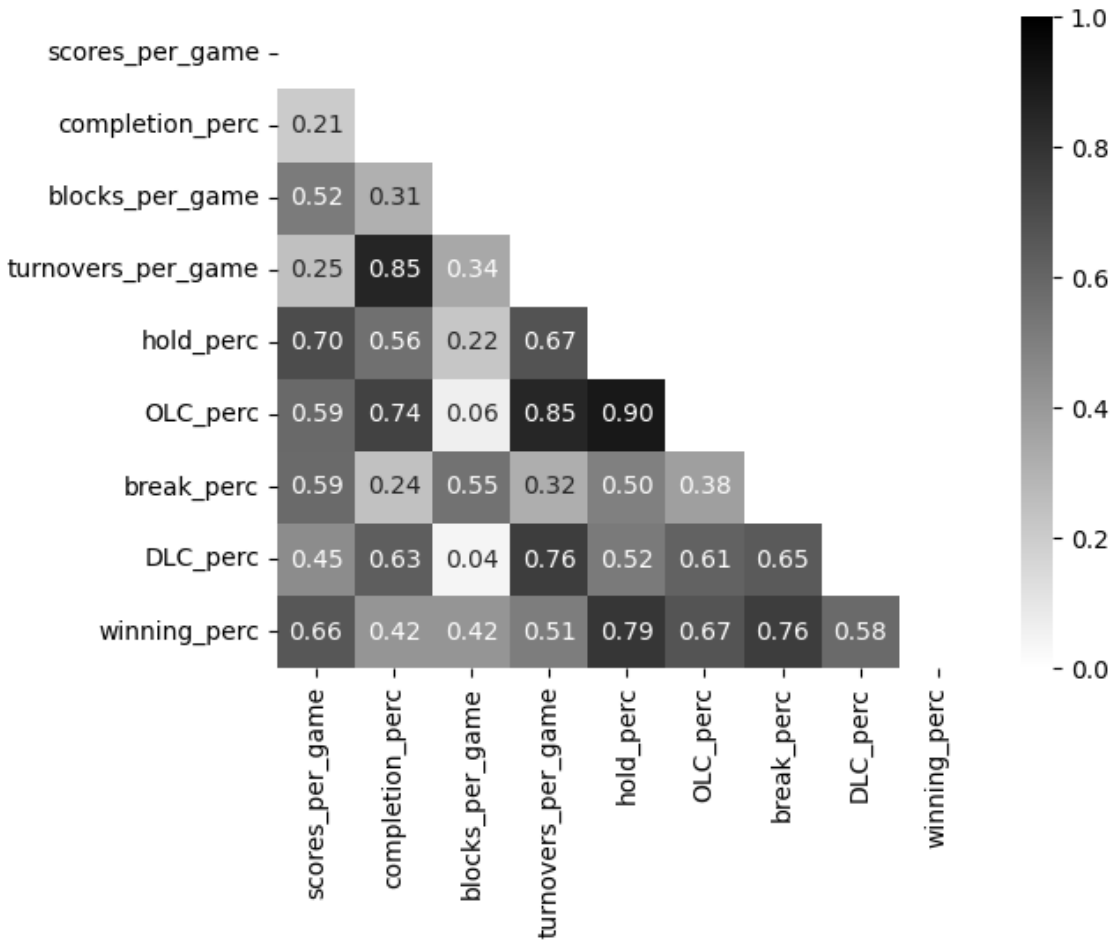

Figura 7: Diagrama de correlación de Pearson para las variables del conjunto final de datos.

Este análisis de correlación permitió evidenciar que las variables ***turnovers_per_game*** y ***OLC_perc*** presentan valores de correlación superiores al umbral establecido, lo que sugiere una redundancia en la información que aportan. Dado que estas variables resultan más complejas y menos interpretables en el contexto del juego en comparación con aquellas con las que mostraron alta

correlación, también fueron descartadas del conjunto de datos final con el propósito de mejorar la eficiencia e interpretabilidad del modelo.

Este proceso de selección de características es esencial ya que permite reducir la dimensionalidad del conjunto de datos y conservar únicamente las variables más informativas y representativas para el entrenamiento de un modelo de aprendizaje de máquinas.

4.1.4. Descripción del conjunto de datos final

El proceso metodológico descrito anteriormente garantiza que el conjunto de datos resultante sea adecuado para su posterior uso en el entrenamiento y evaluación de los modelos de aprendizaje de máquinas implementados.

El conjunto de datos resultante fue diseñado con el objetivo de capturar de manera efectiva el rendimiento de los equipos en función de su historial reciente de partidos. Para ello, cada registro dentro del conjunto de datos representa las estadísticas avanzadas calculadas con base en la ventana de 5 partidos previos, definida en la sección 4.3.1. Además, cada registro incluye la etiqueta correspondiente al resultado obtenido en el partido asociado, es decir, si el equipo en cuestión obtuvo una victoria o una derrota. Esta etiqueta representa la variable objetivo del modelo de clasificación y es de naturaleza binaria, donde los posibles valores son “1” para una victoria y “0” para una derrota.

Finalmente, tras completar el proceso de selección de características, se estableció la estructura definitiva del conjunto de datos, el cual quedó conformado por un total de **2019 registros y 7 atributos**: *scores_per_game*, *blocks_per_game*, *hold_perc*, *break_perc*, *DLC_perc*, *winning_perc* y ***win*** (variable objetivo).

Esta versión final del conjunto de datos garantiza un equilibrio entre representatividad y eficiencia, proporcionando la información necesaria para la estimación de la probabilidad de victoria con base en estadísticas clave del rendimiento de los equipos.

Antes de la implementación de los modelos, el conjunto de datos se divide en dos subconjuntos: uno para el entrenamiento de los modelos y otro para su evaluación. Esta división se realiza utilizando una proporción del 70 % para entrenamiento y 30 % para evaluación, asegurando así una validación adecuada del desempeño de los modelos.

4.2. Diseño experimental

4.2.1. Selección de los modelos y sus hiperparámetros

Una vez obtenido un conjunto de datos debidamente estructurado, limpio y preprocesado, se procede con la implementación, entrenamiento, optimización y evaluación de los modelos de clasificación introducidos en el marco teórico de referencia del presente trabajo:

- Regresión logística
- Máquinas de Soporte Vectorial (SVM)
- Árboles de decisión
- Random Forest
- *k* vecinos más cercanos (*k*-NN)
- Gradient Boosting

El proceso inicia con una fase exploratoria en la que todos los modelos fueron instanciados utilizando sus hiperparámetros predeterminados según la implementación disponible en la librería *scikit-learn* de Python. Esta etapa tiene como propósito obtener una evaluación temprana del desempeño de cada modelo sobre el conjunto de datos, permitiendo identificar aquellos que presentan un mejor ajuste inicial y, por lo tanto, justifican una posterior optimización de hiperparámetros (ver Figura 8).

Para el óptimo ajuste de los modelos seleccionados, se implementó un proceso de optimización de hiperparámetros mediante el método de búsqueda en cuadrícula con validación cruzada (*GridSearchCV*). Este enfoque permite explorar exhaustivamente diferentes combinaciones de hiperparámetros y evaluar su desempeño con base en una métrica de evaluación específica, que en este caso fue la exactitud (*accuracy*). La validación cruzada desempeña un papel fundamental en este procedimiento, ya que mitiga el riesgo de sobreajuste al garantizar que el modelo generalice adecuadamente a datos no vistos. Para este proceso se utilizó validación cruzada con k = 10 particiones (*folds*), lo que implica que el conjunto de entrenamiento se dividió en diez subconjuntos del mismo tamaño, utilizando nueve de ellos para el entrenamiento y uno para la validación en cada iteración.

A continuación, en la Tabla 4, se presentan las combinaciones de hiperparámetros evaluadas durante el proceso de optimización para cada uno de los modelos considerados en este trabajo.

Tabla 4: Valores a evaluar en el proceso de optimización de hiperparámetros para cada uno de los modelos.

Modelo	Hiperparámetro	Valores
Regresión logística	C	0.01, 0.1, 1, 10, 100
	solver	liblinear
	penalty	l1, l2
Máquina de Soporte Vectorial (SVM)	C	0.01, 0.1, 1, 10, 100
	probability	True
	kernel	poly, rbf, sigmoid
Random Forest	max_depth	50, 100
	max_features	sqrt, log2
	min_samples_leaf	1, 10
	n_estimators	100, 1000

Gradient Boosting	max_depth	50, 100
	learning_rate	0.01, 0.1
	max_features	sqrt, log2
	min_samples_leaf	1, 10
	n_estimators	100, 1000

Una vez concluido el proceso de optimización, se realizó una comparación detallada del desempeño de los distintos modelos, tanto en su configuración predeterminada como con sus hiperparámetros optimizados. Este análisis permitió observar diferencias en la capacidad predictiva de los modelos bajo distintas combinaciones de hiperparámetros, lo cual fue esencial para identificar aquellos enfoques que ofrecían una mayor robustez y precisión en la estimación de la probabilidad de victoria. La evaluación se basó en la exactitud (*accuracy*), aunque también se consideraron otras métricas complementarias, como el *F1 score*, el área bajo la curva ROC y el *Brier score*, para lograr una valoración más integral del rendimiento de cada modelo.

Todas las métricas fueron calculadas sobre el subconjunto de evaluación, previamente separado del conjunto de datos original, con el fin de estimar el comportamiento de los modelos ante datos no vistos durante el entrenamiento.

Cabe resaltar que, para que los modelos puedan aprender de manera efectiva estos patrones y reflejar el estado competitivo de los equipos, es fundamental definir adecuadamente el contexto temporal sobre el cual se calculan las estadísticas avanzadas.

4.2.2. Descripción del procedimiento de entrenamiento y validación de los modelos

Uno de los aspectos fundamentales para el correcto desarrollo de este trabajo es la determinación de la ventana de partidos más adecuada para la implementación de

los modelos descritos anteriormente. La selección de esta ventana es crucial, ya que debe permitir capturar de manera efectiva los patrones de desempeño de los equipos sin comprometer en exceso la cantidad de registros disponibles en el conjunto de datos final.

El tamaño de la ventana de partidos conlleva un equilibrio entre representatividad y volumen de datos. Si se elige una ventana demasiado pequeña, se obtendrá un mayor número de registros en el conjunto de datos final. Sin embargo, es probable que la cantidad de información recopilada sobre el rendimiento real del equipo sea insuficiente para reflejar fielmente su desempeño y evolución a lo largo de la temporada. Por el contrario, una ventana de partidos demasiado amplia permitirá capturar con mayor precisión los patrones de rendimiento de los equipos, pero podría reducir significativamente el número de registros disponibles, lo que podría afectar la capacidad de los modelos de aprendizaje de máquinas para generalizar sobre los datos.

La cantidad de partidos previos considerados para cada registro influye directamente en la calidad de la información disponible para la predicción, así como en el tamaño total del conjunto de datos. Por tanto, antes de avanzar con la evaluación final de los modelos, se hace necesario analizar cuál es la ventana de partidos óptima que permite lograr un equilibrio entre representatividad del desempeño reciente de los equipos y un volumen suficiente de datos históricos.

Dada la naturaleza propia del deporte y el contexto competitivo de la Asociación de Ultimate Frisbee, donde los equipos disputan un mínimo de 9 partidos y un máximo de 14 en caso de avanzar hasta la final, se planteó la necesidad de evaluar diferentes tamaños de ventana de partidos. Para ello, se realizó un análisis considerando valores de ventana desde 3 hasta 9 partidos.

Este procedimiento consistió en evaluar el desempeño de los modelos introducidos en el numeral anterior, utilizando sus hiperparámetros predeterminados, con el fin de identificar qué tamaños de ventana conducían a mejores resultados. Además, se

contrastaron estos desempeños con la cantidad de registros obtenidos en el conjunto de datos final para cada tamaño de ventana evaluado.

En la Figura 8 se presentan los resultados obtenidos a partir de este análisis.

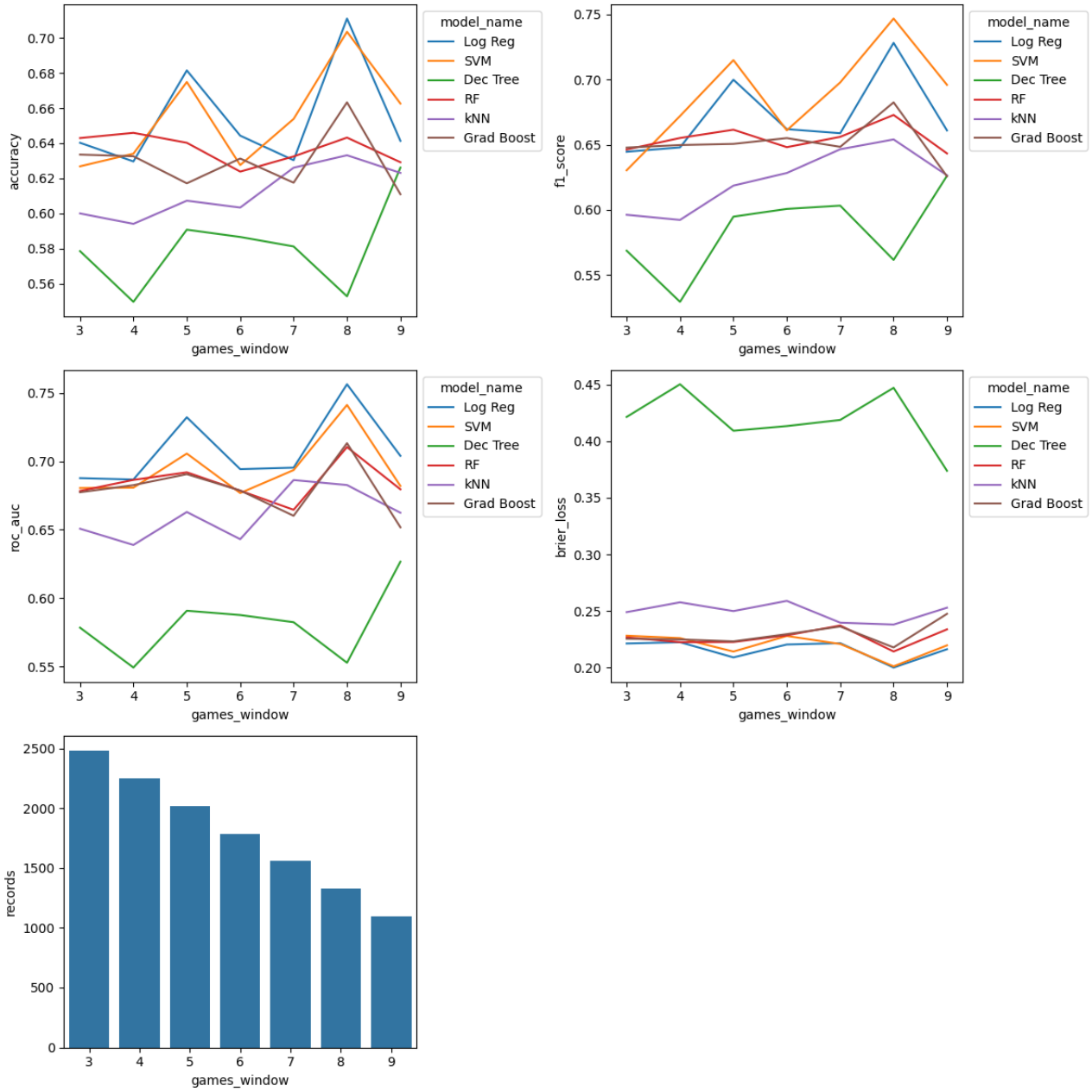

Figura 8: Resultados (métricas y cantidad de registros) del análisis para la determinación de la ventana de partidos más adecuada.

4.3. Resultados y análisis

4.3.1. Determinación experimental de la ventana de partidos

Como parte del diseño experimental, se evaluó el efecto del tamaño de la ventana de partidos sobre el desempeño de los modelos de clasificación. Esta exploración permitió analizar la sensibilidad de cada modelo a distintas configuraciones temporales y tomar una decisión informada sobre cuál ventana utilizar en la implementación final.

A partir de los resultados presentados en la Figura 8, se puede observar que las ventanas de 5 y 8 partidos fueron las que obtuvieron los mejores desempeños en términos de las métricas de evaluación. Sin embargo, para la implementación final de los modelos se optó por utilizar una ventana de 5 partidos. Aunque las métricas de evaluación obtenidas con esta configuración presentan una ligera disminución en comparación con la ventana de 8 partidos, de aproximadamente un 6%, este impacto es compensado por un incremento significativo en la cantidad de registros disponibles en el conjunto de datos final, de aproximadamente un 54%.

4.3.2. Entrenamiento y validación de los modelos

Los resultados obtenidos evidenciaron que algunos modelos presentaban un comportamiento más estable frente a variaciones en los datos, mientras que otros requerían ajustes más específicos para alcanzar niveles competitivos de desempeño. Esta observación resalta la importancia de contar con un conjunto de datos bien estructurado y con características relevantes, que permita a los modelos capturar de forma efectiva los patrones de rendimiento de los equipos. En particular, modelos como la regresión logística y las máquinas de soporte vectorial (SVM) mostraron un desempeño significativo en diferentes escenarios, consolidándose como opciones destacadas para su posterior análisis.

Una vez finalizado el proceso de entrenamiento y optimización de hiperparámetros, se procedió a evaluar el desempeño de los modelos de clasificación implementados.

La comparación se realizó tanto para las versiones base como para las versiones optimizadas de cada modelo, con el objetivo de identificar aquellos enfoques que ofrecieran un mejor ajuste al problema planteado.

La evaluación se basó en las cuatro métricas introducidas en el marco teórico de referencia del presente trabajo: exactitud (*accuracy*), *F1 score*, área bajo la curva ROC y *Brier score*. Estas métricas fueron calculadas sobre el subconjunto de evaluación, con el propósito de estimar el comportamiento real de los modelos frente a datos no vistos durante la fase de entrenamiento.

En la Figura 9 se presenta un resumen gráfico de los resultados obtenidos para cada métrica y cada modelo considerado.

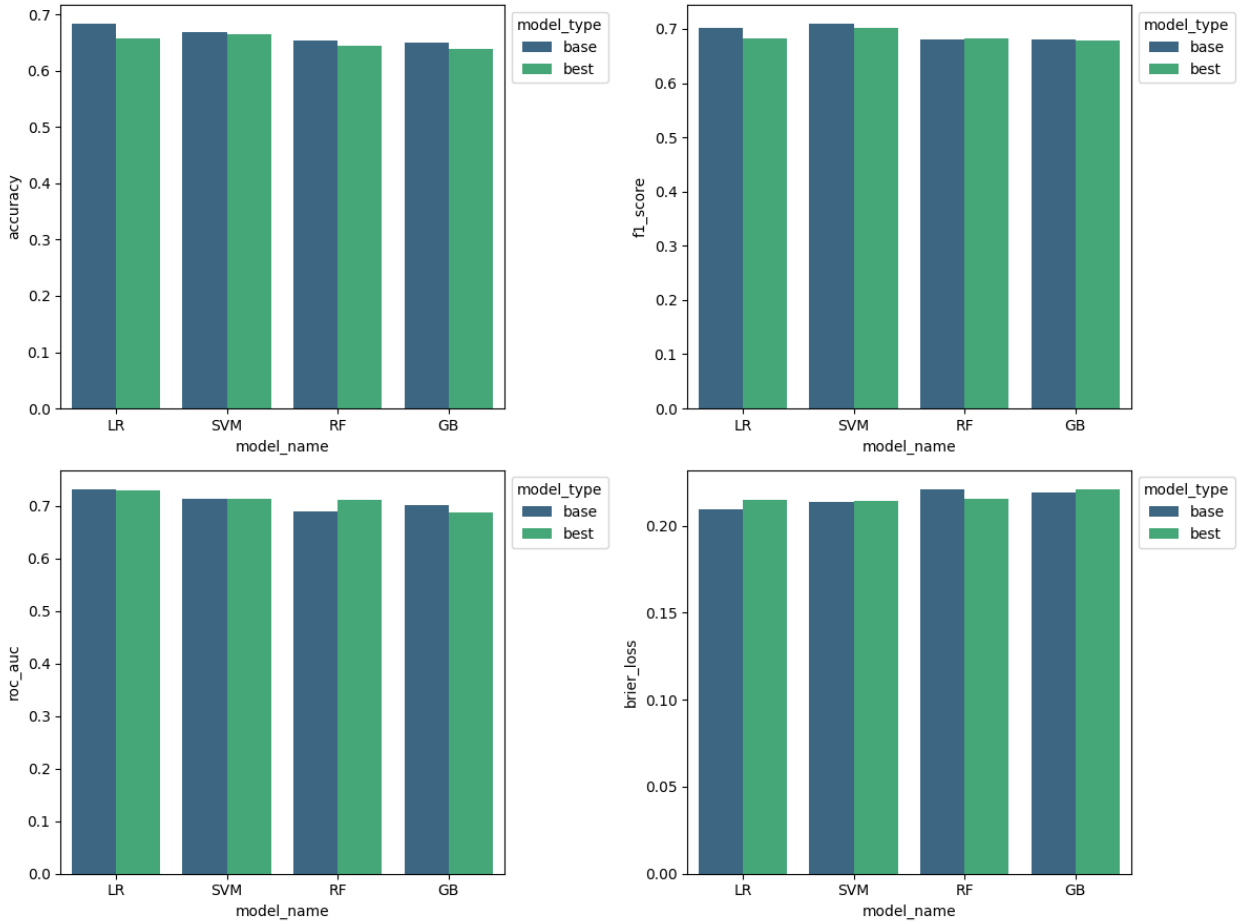

Figura 9: Evaluación de modelos con base en diferentes métricas.

Al comparar el desempeño de los modelos en sus versiones base y optimizada, se observa que las diferencias en las métricas evaluadas son mínimas. En la mayoría de los casos, los valores obtenidos por los modelos optimizados no representan una mejora significativa frente a sus configuraciones por defecto. Esta evidencia sugiere que, para el presente conjunto de datos y el problema abordado, el ajuste de hiperparámetros no tuvo un impacto sustancial en el rendimiento general. No obstante, es importante resaltar que la optimización de hiperparámetros constituye una etapa fundamental en la implementación de modelos de aprendizaje de máquinas, ya que permite explorar el espacio de configuración del modelo y garantizar que su desempeño sea evaluado bajo condiciones óptimas. En este caso particular, se optó por utilizar las versiones base de los modelos, favoreciendo la simplicidad y facilitando su implementación práctica.

Los resultados obtenidos muestran un desempeño similar entre los diferentes modelos evaluados en todas las métricas consideradas. Las diferencias observadas son leves, lo cual sugiere que todos los clasificadores implementados presentan una capacidad predictiva comparable sobre el conjunto de evaluación. Sin embargo, dado el enfoque probabilístico de este trabajo, se otorga una mayor relevancia a las métricas de exactitud (*accuracy*) y *Brier score*, ya que permiten evaluar de forma complementaria tanto la capacidad del modelo para clasificar correctamente las observaciones como la calidad de las probabilidades estimadas.

En términos de exactitud (*accuracy*), los modelos de regresión logística (LR) y máquinas de soporte vectorial (SVM) alcanzaron valores ligeramente más altos. Un comportamiento similar se observa en la métrica *Brier score*, donde LR y SVM obtuvieron valores ligeramente más bajos, lo que indica una mejor calibración en las probabilidades asignadas a las clases.

También vale la pena mencionar que con respecto al *F1 score*, las diferencias entre modelos también son mínimas, lo que refleja un buen balance entre precisión y sensibilidad en todos los casos. En cuanto al área bajo la curva ROC, las puntuaciones obtenidas muestran una capacidad discriminativa similar entre

modelos, con valores superiores a 0.7, lo que indica una adecuada separación entre clases.

Con base en los resultados anteriores, se seleccionaron los modelos de regresión logística y SVM como finalistas para un análisis más detallado. En la Figura 10 se presentan las matrices de confusión correspondientes a estos dos modelos, construidas a partir del conjunto de evaluación.

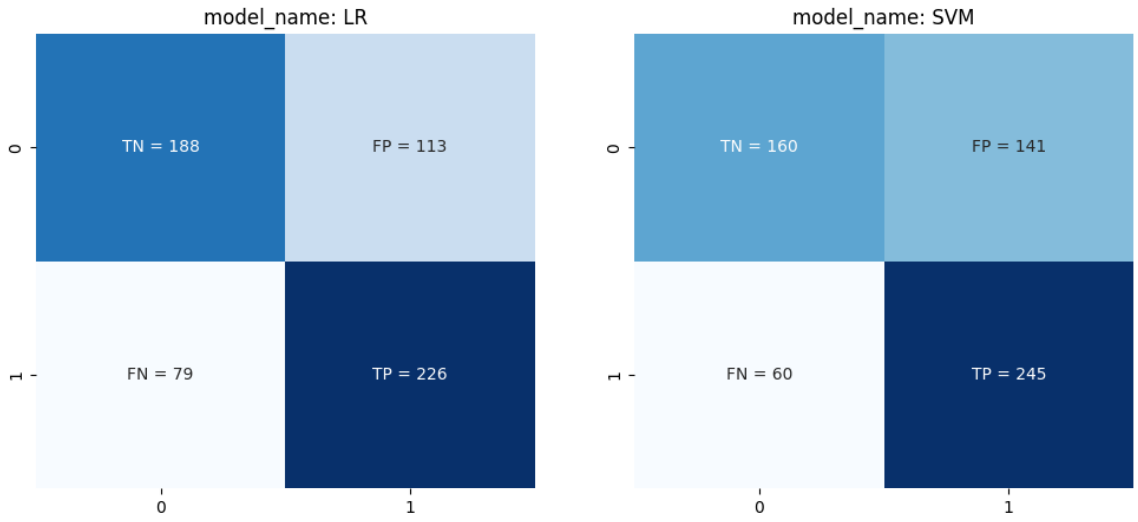

Figura 10: Comparación de las matrices de confusión para los modelos regresión logística y SVM.

A partir de las matrices de confusión se evidencia nuevamente que ambos modelos mantienen un desempeño adecuado y similar entre sí. Sin embargo, dado que el objetivo principal de este trabajo consiste en estimar la probabilidad de victoria de los equipos de forma clara y comprensible para todos los actores del ecosistema deportivo, se consideró pertinente priorizar la interpretabilidad como criterio complementario. En consecuencia, se seleccionó la regresión logística como modelo final para este trabajo, destacando su equilibrio entre interpretabilidad, exactitud y calibración.

5. CONCLUSIONES Y TRABAJO FUTURO

5.1. Conclusiones

El presente trabajo ha demostrado que la disponibilidad de datos históricos deportivos debidamente estructurados y almacenados es un factor clave para el desarrollo y aplicación de metodologías avanzadas de analítica deportiva. La sistematización de la información permite una exploración continua de distintas estrategias analíticas que, a su vez, contribuyen a la optimización de tácticas y enfoques de juego dentro de las entidades deportivas. Además, el acceso a estos datos no solo beneficia a los equipos y entrenadores, sino que también enriquece la experiencia de los aficionados y refuerza la industria del entretenimiento deportivo.

El Ultimate Frisbee, siendo un deporte que ha experimentado un crecimiento y desarrollo significativo en los últimos años, puede beneficiarse ampliamente de estas herramientas analíticas. Un aspecto clave identificado en este trabajo es el potencial que tienen las apuestas deportivas legales, equilibradas y bien estructuradas, no solo como una fuente importante de ingresos económicos para las ligas y los equipos, sino también como un mecanismo que fomenta la consolidación y masificación del deporte. Al proporcionar modelos predictivos sólidos y confiables, se genera un entorno más transparente y atractivo.

El desarrollo de este trabajo ha puesto en evidencia la importancia de contar con un entendimiento profundo de las dinámicas deportivas para la construcción de un conjunto de datos representativo. Desde la selección de características más relevantes hasta la definición de la ventana de partidos óptima, cada etapa del proceso requirió un análisis detallado para garantizar que los datos reflejaran de manera fiel el desempeño real de los equipos. Este enfoque permitió generar un conjunto de datos sólido y bien estructurado, lo que a su vez facilitó la implementación y evaluación de los modelos de aprendizaje de máquinas empleados en este trabajo.

Los resultados obtenidos muestran que los modelos de menor complejidad pueden ser altamente efectivos en la captura de patrones en los datos y la generalización de resultados sin incurrir en un alto costo computacional. Esta conclusión resalta la importancia de evaluar el balance entre la robustez del modelo y la naturaleza del problema abordado, ya que, en muchos casos, modelos más simples pueden ofrecer un rendimiento competitivo sin la necesidad de recurrir a técnicas con un alto nivel de complejidad.

Por último, aunque no es posible establecer una comparación directa entre los resultados obtenidos en este trabajo y otros estudios previos, debido a la escasez de investigaciones formales en el ámbito de la analítica deportiva aplicada al Ultimate Frisbee, este hecho también resalta la relevancia de la presente investigación. La metodología empleada y los hallazgos aquí obtenidos constituyen una contribución significativa al desarrollo de propuestas académicas en este campo.

5.2. Trabajo futuro

A modo de trabajo futuro, existen diversas oportunidades para expandir el alcance y aplicabilidad de este trabajo a partir de estas posibles líneas de investigación:

Implementación de redes neuronales del tipo Long Short-Term Memory (LSTM), que permitirían capturar mejor la evolución temporal del desempeño de los equipos, aprovechando patrones en secuencias de estadísticas históricas [17].

Otra oportunidad prometedora es la construcción de un conjunto de datos más detallado, donde se registre la información de cada jugada dentro de un partido. Esto facilitaría el desarrollo de un modelo de probabilidad de victoria dinámico, capaz de actualizar sus predicciones en tiempo real conforme avanza el juego.

La incorporación de estadísticas individuales y rendimiento de cada uno de los jugadores permitiría mejorar la precisión del modelo al considerar el impacto de cada atleta en el desempeño del equipo [18].

6. REFERENCIAS

[1] V. Sarlis y C. Tjortjis, “Sports analytics - Evaluation of basketball players and team performance”, Inf Syst, vol. 93, p. 101562, nov. 2020, doi: 10.1016/J.IS.2020.101562.

[2] Y. El-Maghrabi y M. Sharif, “Game Changers or Game Predictors? Big Data Analytics in Sports for Performance Enhancement and Fan Engagement”, Journal of Contemporary Healthcare Analytics, vol. 6, núm. 6, pp. 19–39, jun. 2022.

[3] D. Cason, M. Lee, J. Lee, E. J. Arner, y I. S. Yeo, “The Impact of Legalization of Sports Gambling: How Motivation, Fandom, and Gender Influence Sport-Related Consumption”, International Journal of Sport Communication, vol. 13, núm. 4, pp. 643–654, nov. 2020, doi: 10.1123/IJSC.2020-0022.

[4] “Offshore Sportsbook To Start Taking AUDL Bets - Ultiworld”.

[5] “DraftKings Now Taking AUDL Bets in Some States - Ultiworld”.

[6] “American Ultimate Disc League and Lsports Data Enter Into $3 Million Data Distribution and Co-development Strategic Partnership | WatchUFA”.

[7] “Ultimate Frisbee - Federación Colombiana de Disco Volador”. [En línea]. Disponible en: https://fecodv.ultimatecentral.com/ultimate-frisbeemodalidades

[8] “About the UFA”. [En línea]. Disponible en: https://watchufa.com/league/about

[9] K. Murphy, Machine learning: A probabilistic perspective. MIT Press, 2012.

[10] A. Géron, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow. O’Reilly Media, 2019.

[11] C. Johnstone y D. Nettleton, “Using Conformal Win Probability to Predict the Winners of the Canceled 2020 NCAA Basketball Tournaments”, Am Stat, vol. 78, núm. 3, pp. 304–317, 2024, doi: 10.1080/00031305.2023.2283199.

[12] B. S. Baumer, G. J. Matthews, y Q. Nguyen, “Big ideas in sports analytics and statistical tools for their investigation”, Wiley Interdiscip Rev Comput Stat, vol. 15, núm. 6, p. e1612, nov. 2023, doi: 10.1002/WICS.1612.

[13] S. Gregory, S. Robertson, R. Aughey, y G. Duthie, “The influence of tactical and match context on player movement in football”, J Sports Sci, vol. 40, núm. 9, pp. 1063–1077, may 2022, doi: 10.1080/02640414.2022.2046938.

[14] P. Robberechts, J. Van Haaren, y J. Davis, “A Bayesian Approach to In-Game Win Probability in Soccer”, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 3512–3521, ago. 2021.

[15] K. Sugie y E. Konaka, “Development of Individual Value of Offense/Defense Contribution in Basketball Teams”, IEEJ Transactions on Electronics, Information and Systems, vol. 143, núm. 12, pp. 1137–1144, dic. 2023, doi: 10.1541/IEEJEISS.143.1137.

[16] C. Go, “Exploring Predictors of Team Success in Ultimate Frisbee: An Analysis of Game Statistics for Stanford Women’s Ultimate”, 2017.

[17] Q. Zhang, X. Zhang, H. Hu, C. Li, Y. Lin, y R. Ma, “Sports match prediction model for training and exercise using attention-based LSTM network”, Digital Communications and Networks, vol. 8, núm. 4, pp. 508–515, ago. 2022, doi: 10.1016/J.DCAN.2021.08.008.

[18] J. Stübinger, B. Mangold, y J. Knoll, “Machine Learning in Football Betting: Prediction of Match Results Based on Player Characteristics”, Applied Sciences 2020, Vol. 10, Page 46, vol. 10, núm. 1, p. 46, dic. 2019, doi: 10.3390/APP10010046.