****

ANÁLISIS DE SERIES TEMPORALES PARA LA PREDICCIÓN DE RESULTADOS EN PLATAFORMAS DE APUESTAS DEPORTIVAS

MODALIDAD DEL TFG: CONVENCIONAL

CONVOCATORIA: ORDINARIA

ALUMNA: ANA ISABEL GONZÁLEZ SAHAGÚN

TUTORA: MARÍA DEL MAR ANGULO MARTÍNEZ

GRADO: MATEMÁTICAS COMPUTACIONALES

CONTENIDO

RESUMEN 4

ABSTRACT 4

1. INTRODUCCIÓN 5

1.1 Motivación y contexto 5

1.2 Planteamiento del problema 6

1.2.1 Definición de éxito de los equipos 6

1.2.2 Encontrar un Modelo para Reducir al Máximo el Impacto del Azar 7

1.2.3 Definición de Métricas Claras para Evaluar el Desempeño del Modelo 7

1.2.4 Limitaciones de los modelos matemáticos y estadísticos 7

1.3 Objetivos del trabajo 8

2. ESTADO DE LA CUESTIÓN 9

2.1 Plataformas de Apuestas Deportivas 9

2.1.1 Historia y Evolución de las Apuestas Deportivas 9

2.1.2 Funcionamiento de las Apuestas Deportivas 9

2.1.3 Datos Estadísticos del Mercado: Impacto Económico y Social 11

2.2 El Futbol y las Apuestas 12

El fútbol es el deporte más apostado a nivel mundial, representando hasta el 86% de las apuestas deportivas en algunos países. Su popularidad y la gran cantidad de partidos disputados a lo largo del año lo convierten en el mercado más activo dentro de la industria. 12

2.2.1 Apuestas en el Fútbol 12

2.2.2 Impacto Económico de las Apuestas en el Fútbol 12

2.3 Marco teórico del trabajo 13

2.3.1 Procesos estocásticos en el Deporte 13

2.3.2 Introducción a las Series Temporales 13

2.3.4 Estacionariedad 13

2.3.5 Familia ARIMA 13

2.3.6 Validación del Modelo 14

2.3.7 Introducción al modelo Transformer 14

2.3.8 Transformers en Series Temporales 14

2.2 Trabajos Relacionados 14

3. ASPECTOS METODOLÓGICOS 15

3.1 Metodología 15

3.2 Tecnologías empleadas 15

4. DESARROLLO DEL TRABAJO 16

5. CONCLUSIONES 17

6. REFERENCIAS 18

6.1 Bibliografía 18

6.2 Índices de citas e imágenes 18

ANEXOS 19

# RESUMEN

# ABSTRACT

# 1. INTRODUCCIÓN

## 1.1 Motivación y contexto

El deporte es un factor decisivo para el crecimiento económico y el desarrollo social de muchos países. No solo genera ingresos por eventos, derechos de televisión o patrocinios, sino que también impulsa la creación de empleo y fomenta la actividad comercial a su alrededor. Además, al involucrar a millones de personas, despierta un interés global que lo convierte en un fenómeno de gran relevancia social.

Dentro del deporte, el fútbol destaca como una de las disciplinas con mayor audiencia y volumen económico en el mundo. En España, el fútbol profesional generó un impacto económico total de 18.350 millones de euros en la temporada 2021/2022, equivalente al 1,44% del PIB nacional [[5](#_6.1_Bibliografía)]. Al mismo tiempo, las apuestas deportivas han crecido de manera exponencial, con un gasto en apuestas online que superó los 2.950 millones de euros en el mismo período. Las plataformas de apuestas buscan predecir resultados de forma fiable, ya que un pronóstico acertado influye directamente en sus ganancias y en las decisiones de los apostadores. No obstante, la naturaleza impredecible del deporte convierte esta tarea en un desafío complejo. Aunque existe una gran disponibilidad de datos, aún no existen modelos suficientemente establecidos que garanticen una predicción exacta de los resultados, tanto en términos económicos como deportivos. La importancia de estas decisiones en el fútbol resalta la necesidad de un análisis cuantitativo sólido y riguroso.

Por otro lado, las matemáticas y la estadística ofrecen herramientas para abordar problemas de predicción. Permiten estimar probabilidades y comportamientos futuros basándose en la información histórica disponible. En este contexto, el análisis de series temporales se ha convertido en una técnica clave para comprender cómo evolucionan ciertos eventos a lo largo del tiempo. Su aplicación abarca campos tan variados como la economía, la meteorología y la salud. A través de este análisis, es posible identificar patrones y tendencias en los datos históricos, lo que resulta esencial para optimizar la toma de decisiones y anticipar posibles escenarios futuros. Además, se pueden integrar métodos de Machine Learning capaces de procesar grandes volúmenes de información y aplicarlos incluso en contextos con conjuntos de datos más reducidos.

En los últimos años, se han logrado avances significativos en el procesamiento de datos secuenciales. Esto ha impulsado la aparición de modelos más sofisticados, como los Transformers, que han destacado por superar las limitaciones de los métodos tradicionales y por su capacidad de capturar dependencias temporales de largo alcance [[1](#_6.1_Bibliografía)]. Estos modelos no se basan en bucles recurrentes, sino en mecanismos de atención que han demostrado eficacia para manejar grandes volúmenes de información.

Los Transformers se consideran una de las tecnologías más recientes en el ámbito de la inteligencia artificial y son objeto de una intensa investigación. Su novedad se refleja en numerosos trabajos que profundizan en su arquitectura y aplicaciones. Recientemente, se han desarrollado arquitecturas específicas de Transformers adaptadas al análisis de series temporales, logrando así una mejora notable de su rendimiento en tareas de predicción y modelado de datos secuenciales [[2](#_6.1_Bibliografía)].

## 1.2 Planteamiento del problema

La predicción de resultados de fútbol es un problema complejo debido a la necesidad de integrar de múltiples factores que pueden afectar al marcador final de un partido. Como por ejemplo rendimiento de los equipos, condiciones físicas de los jugadores, clima, decisiones arbitrarías, o tácticas empleadas. Esta variedad de factores hace que los modelos matemáticos y/o estadísticos tradicionales tengan limitaciones en su capacidad de predicción.

Uno de los principales desafíos es la disponibilidad y calidad de los datos históricos. Muchas veces estos datos presentan sesgos, información incompleta o registros inconsistentes, haciendo más difícil la construcción de modelos robustos. Para abordar este problema, se ha utilizado un conjunto de datos históricos de resultados de partidos de fútbol de la liga española desde el año 1995. A continuación, se presentan los principales retos que enfrenta la predicción de resultados deportivos.

1.2.1 Definición de Éxito de los Equipos

En el contexto del fútbol, la victoria es la medición inmediata más directa para evaluar el éxito de un equipo. Sin embargo, la definición de éxito toma matices específicos según el interés del estudio. Por ejemplo, para medir el rendimiento de un equipo se pueden considerar factores como la clasificación final en una liga, la diferencia de goles, el porcentaje de victorias de una temporada, y otros indicadores relacionados con el desempeño. Todos estos elementos ofrecen información para evaluar el éxito de un equipo. Además, indican la necesidad de realizar un análisis previo para determinar cuáles de estas variables resultan más útiles a la hora de desarrollar modelos de predicción e interpretar los resultados.

1.2.2 Encontrar un Modelo para Reducir al Máximo el Impacto del Azar

El fútbol está sujeto a un grado inevitable de incertidumbre. Lesiones repentinas, cambios meteorológicos o actuaciones arbitrales pueden alterar de forma significativa el desarrollo de un partido. El objetivo de estos modelos no es suprimir completamente el azar, sino intentar reducir su impacto. Si se logra encontrar un modelo que identifique patrones estables y ofrezca predicciones sólidas, se podría reducir el peso de la imprevisibilidad hasta cierto punto. Así, aun sabiendo que existen limitaciones, se trata de encontrar un modelo que se aproxime lo máximo a la realidad.

1.2.3 Necesidad de Definir de Métricas Claras para Evaluar el Modelo

Uno de los principales problemas en el desarrollo de modelos de predicción es la dificultad para medir de manera precisa el desempeño del modelo. Sin criterios de evaluación bien definidos, es complicado determinar si un modelo realmente mejora o si las diferencias observadas se deben a fluctuaciones aleatorias de los datos. Esto puede llevar a conclusiones erróneas y a modelos que parecen eficaces en el entrenamiento pero fracasan con datos nuevos.

1.2.4 Limitaciones de los modelos matemáticos y estadísticos

Aunque las técnicas tradicionales de predicción proporcionan un marco sólido para el análisis, también presentan limitaciones.

* **Fuerte dependencia de supuestos simplificados:** Se asume que las variables son independientes, o que los equipos mantienen un rendimiento estable sin altibajos significativos.
* **Capacidad de adaptación a cambios drásticos:** Los modelos basados en datos históricos tienen dificultades para reflejar los cambios que pueden ocurrir en un partido.
* **Sesgos en el mercado de apuestas:** Investigaciones recientes señalan que se producen ineficiencias relacionadas con la estimación de la ventaja de localía y con el efecto longshot-favourite. Tales distorsiones implican que el comportamiento real de los equipos y el establecimiento de cuotas en las casas de apuestas no siempre coincide con los supuestos matemáticos de equilibrio y eficiencia [[3](#_6.1_Bibliografía)].

Estas limitaciones explican la necesidad de implementar métodos que combinen técnicas estadísticas y de inteligencia artificial, con el objetivo de mejorar la precisión captando patrones complejos. Sin embargo, estos métodos avanzados también requieren una continua actualización y validación frente a nuevos datos e imprevistos del entorno.

1.2.5 Tema poco desarrollado

La predicción de resultados deportivos es un área con poca investigación debido a la alta imprevisibilidad del deporte. Factores externos difíciles de modelar, como decisiones arbitrales o el estado anímico de los jugadores, dificultan la creación de modelos precisos. Esto ha limitado el desarrollo de modelos fiables en este ámbito

## 1.3 Objetivos del trabajo

El objetivo principal de este trabajo es evaluar la capacidad de los modelos de series temporales y Transformers en la predicción de resultados de fútbol, un problema que se caracteriza por una alta aleatoriedad y múltiples factores influyentes. Para ello, se seguirá una metodología estructurada que permitirá analizar y comparar distintos enfoques de modelado con el fin de determinar su efectividad en este problema. Para lograr este propósito, se plantean los siguientes objetivos específicos:

* **Mejorar la eficiencia de los modelos:** Trabajar con técnicas para ordenar y extraer características clave, facilitando criterios de clasificación claros.
* **Desarrollo y evaluación de modelos:** Construir y probar diferentes esquemas de predicción para medir su desempeño real.
* **Comparar métodos estadísticos tradicionales con técnicas avanzadas**: Comparar enfoques clásicos con algoritmos modernos para hallar ventajas y desventajas en cada caso.
* **Analizar la base matemática de los Transformers:** Examinar los fundamentos teóricos y sus implicaciones en el rendimiento del modelo.
* **Evaluar la capacidad de los Transformers para transformar datos cotidianos en información estadística:** Verificar cómo estos modelos convierten los datos cotidianos en representaciones matemáticas para la predicción.
* **Definir y aplicar métricas de evaluación**: Establecer indicadores concretos para valorar la calidad y precisión de las predicciones.
* **Examinar limitaciones y efectividad:** Identificar puntos débiles y fortalezas de los diferentes métodos propuestos.

# 2. ESTADO DE LA CUESTIÓN

## 2.1 Plataformas de Apuestas Deportivas

Las apuestas deportivas han evolucionado de una práctica informal a una industria global con un impacto económico significativo. La digitalización ha permitido el acceso a una mayor variedad de mercados y ha impulsado el uso de modelos analíticos avanzados para mejorar la precisión de las predicciones. En esta sección se contextualiza el sector de las apuestas deportivas, su evolución, funcionamiento y su impacto económico y social.

2.1.1 Historia y Evolución de las Apuestas Deportivas

Las apuestas en eventos deportivos se remontan a la antigüedad, con registros en la Grecia y Roma clásicas. Sin embargo, su regulación formal comenzó en el siglo XVIII en Inglaterra, con la legalización de las apuestas en carreras de caballos y la aparición de las primeras casas de apuestas.[[1]](#footnote-1)

Durante el siglo XX, la práctica se extendió a deportes como el fútbol y el baloncesto, pero el cambio más relevante ocurrió en los años 90 con la llegada de Internet. Esta innovación facilitó la creación de plataformas digitales y la integración de modelos estadísticos avanzados para la fijación de cuotas de apuestas en tiempo real.

2.1.2 Funcionamiento de las Apuestas Deportivas

Las casas de apuestas operan estableciendo cuotas que reflejan la probabilidad de que ocurra un resultado en un evento deportivo. Para ello, se emplean modelos matemáticos y análisis estadísticos basados en datos históricos y factores relevantes como el rendimiento de los equipos y las condiciones del partido.

Para garantizar su rentabilidad, se incorpora un margen de beneficio en las cuotas ofrecidas, conocido como *vig* o *juice*, para obtener ganancias a largo plazo, independientemente del resultado. Este margen se obtiene ofreciendo cuotas ligeramente inferiores a las probabilidades reales del evento, lo que permite que la suma de las probabilidades implícitas supere el 100%. Además, Las casas incorporan sesgos de mercado, ajustes de liquidez, y consideraciones sobre el volumen total apostado para equilibrar su exposición al riesgo y minimizar pérdidas. Este ajuste se intensifica en apuestas en vivo, donde las cuotas se actualizan en directo según transcurre el evento

**Tipos de Casas de Apuestas**

Existen dos tipos principales de casas de apuestas:

* **Casas de apuestas de contrapartida**: Son las más comunes y operan estableciendo cuotas y aceptando apuestas directamente contra los apostantes. Ejemplos de este tipo incluyen Bet365, William Hill y Bwin.
* **Casas de apuestas de intercambio**: Permiten que los apostantes jueguen entre sí en lugar de hacerlo contra la casa. En este modelo, los usuarios pueden fijar sus propias cuotas y aceptar apuestas de otros jugadores. La casa actúa como intermediaria, obteniendo un beneficio a través de una comisión sobre las ganancias de los jugadores. Un ejemplo destacado es Betfair.

**Ajustes de Cuotas y Margen de Beneficio: Sesgos e Ineficiencias**

Para protegerse de posibles pérdidas, las casas de apuestas incluyen un margen de seguridad en las cuotas. Este margen está basado en la probabilidad estadística de los resultados y en el comportamiento de los apostantes. Existen varios estudios que investigan los sesgos y las ineficiencias de las casas de apuestas [[3](#_6.1_Bibliografía)]. Entre los más importantes se encuentran:

* Longshot Bias: Se refiere a la tendencia de sobrevaloración de las apuestas a resultados poco probables, implicando que la probabilidad real de ganar es subestimada en las cuotas [[4](#_6.1_Bibliografía)].
* Home-Field Advantage Misestimation: Error al estimar la ventaja del equipo local. Tradicionalmente, se considera que el equipo de casa tiene una mayor probabilidad de ganar, lo que se refleja en cuotas más bajas. Pero existen situaciones como la ausencia del público que pueden modificar este efecto.
* Incorporación Tardía de Información: Ocurre cuando las casas de apuestas actualizan de forma lenta las cuotas ante la aparición de nueva información.

**Modelo de rentabilidad**

La rentabilidad de una casa de apuestas se basa en la diferencia entre las cantidades apostadas y los pagos realizados a los ganadores. El margen de beneficio se puede expresar con la formula:

Si el resultado es desfavorable para la casa, su perdida se limita a la cantidad apostada

El modelo de negocio de las casas de apuestas no se basa en acertar los resultados de los eventos, sino de asegurarse ganancias a largo plazo ajustando las cuotas y controlando el riesgo financiero [[6](#_6.1_Bibliografía)].

2.1.3 Datos Estadísticos del Mercado: Impacto Económico y Social

El mercado global de apuestas deportivas genera ingresos superiores a los 90.000 millones de dólares anuales, impulsado por su creciente legalización y digitalización. En Europa y América del Norte, este sector es un componente clave de la industria del entretenimiento. Se estima que el mercado de las apuestas deportivas previsto para 2030 sea de 608.410 millones USD[[2]](#footnote-2).

Sin embargo, la accesibilidad de las apuestas en línea también ha incrementado la adicción al juego, especialmente entre los jóvenes. Se estima que alrededor de 80 millones de adultos en el mundo padecen problemas relacionados con el juego compulsivo[[3]](#footnote-3). Un ejemplo claro es Brasil, donde los informes indican que los ciudadanos gastan más de 3.200 millones de euros al mes en apuestas, representando aproximadamente el 20% de la masa salarial del país[[4]](#footnote-4).

## 2.2 El Futbol y las Apuestas

El fútbol es el deporte más apostado a nivel mundial, representando hasta el 86% de las apuestas deportivas en algunos países[[5]](#footnote-5). Su popularidad y la gran cantidad de partidos disputados a lo largo del año lo convierten en el mercado más activo dentro de la industria.

2.2.1 Apuestas en el Fútbol

Las apuestas deportivas en el fútbol han evolucionado en las últimas décadas con la digitalización y legalización de plataformas de juego en línea. Aunque existen múltiples modalidades de apuestas, este trabajo se centrará en las siguientes:

1. 1-X-2: Apuesta a qué equipo ganará el partido, o si habrá un empate
2. Resultado exacto: Acertar el resultado final del partido

Sin embargo, existen diversas modalidades de apuestas en el fútbol, cada una con características específicas que permiten diferentes estrategias de juego. Algunas de las más destacadas son:

1. Descanso/Final del Partido: Se debe acertar el resultado al término de la primera parte y al final del encuentro.
2. Doble Oportunidad: Permite apostar a dos posibles desenlaces. Ofrecen cuotas más bajas, pero aumentan las probabilidades de acierto.
3. Apuesta con reembolso en caso de empate: Se apuesta por la victoria de un equipo con la ventaja de que, si el partido termina en empate, se devuelve el dinero invertido.
4. Apuestas sobre goles: Incluyen predicciones sobre la cantidad total de goles en el partido.
5. Hándicap Asiático: Se asigna una ventaja o desventaja en el marcador a uno de los equipos antes del inicio del partido.
6. Scorecast: Se debe acertar tanto el primer goleador del partido como el resultado final, lo que genera cuotas elevadas debido a su dificultad.
7. Apuestas especiales: Se pueden realizar predicciones sobre diversos aspectos del partido, como el número de tarjetas mostradas.

2.2.2 Impacto Económico de las Apuestas en el Fútbol

Las competiciones de fútbol de mayor prestigio, como la Copa del Mundo y la UEFA Champions League, generan un enorme impacto económico en el mercado de apuestas deportivas. Durante estos torneos, se observa un incremento significativo en la actividad de apuestas, lo que provoca variaciones en los mercados financieros relacionados con esta industria. En España, las apuestas deportivas representan casi el 1% del PIB, reflejando su influencia en la economía nacional [[5](#_6.1_Bibliografía)].

El fútbol no solo tiene un impacto en la industria del juego, sino que también influye en la economía general del país. La actividad económica generada por el futbol repercute en la generación de empleo, con más de 194.381 empleos a jornada completa, incluyendo puestos directos e indirectos relacionados con el deporte[[6]](#footnote-6).

## 2.3 Marco teórico del trabajo

2.3.1 Procesos estocásticos en el Deporte

* Modelos de Poisson y su aplicación en la predicción de goles.
* Cadenas de Márkov para transiciones entre estados de juego.

2.3.2 Introducción a las Series Temporales

* Definición de series temporales.
* Componentes: Tendencia, Estacionalidad, Ciclos y Ruido.

2.3.4 Estacionariedad

* Importancia de la estacionariedad para modelar series temporales.
* Pruebas de estacionariedad (ADF, KPSS) y transformaciones para lograrla.
* Correlación

2.3.5 Familia ARIMA

* Modelos AR, MA y ARMA.
* Introducción a ARIMA y SARIMA.
* SARIMAX
* Box-Jenkins

2.3.6 Validación del Modelo

* Técnicas de validación cruzada en series temporales (walk-forward validation).
* Métricas: MAE, RMSE, MAPE.

2.3.7 Introducción al modelo Transformer

* Origen en procesamiento de lenguaje natural.
* Mecanismo de atención y su importancia en la modelación de secuencias.

2.3.8 Transformers en Series Temporales

* Adaptaciones del Transformer a series temporales: Time Series Transformer (TST).

## 2.2 Trabajos Relacionados

# 3. ASPECTOS METODOLÓGICOS

## 3.1 Metodología

## 3.2 Tecnologías empleadas

# 4. DESARROLLO DEL TRABAJO

# 5. CONCLUSIONES

# 6. REFERENCIAS

## 6.1 Bibliografía

[1] Vaswani, A. et al. (2017). “Attention Is All You Need.” Advances in Neural Information Processing Systems.

[2] Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2023). Transformers in time series: A Survey.

[3] Lora Gutiérrez, J. D., & Macias Burbano, Y. A. (2022). Ineficiencias en el mercado de apuestas del fútbol: Los efectos del COVID-19 [Tesis de grado, Universidad Icesi].

[4] Van Raaij, V. (2019). Favorite-longshot bias in European Football betting market: Differences between popular and non-popular football competitions. Radboud Universiteit.

[5] KPMG Asesores S.L. (2023). Impacto socioeconómico del fútbol profesional en España. KPMG.

[6] Martín Domínguez, D. (2013). Análisis de resultados deportivos y estimación implícita de probabilidades: Fútbol. UC3M

## 6.2 Índices de citas e imágenes

# ANEXOS

1. La evolución histórica de las apuestas deportivas: de los juegos antiguos a la era digital. (2023, 21 junio). Fox Sports. https://www.foxsports.com. [↑](#footnote-ref-1)
2. Fernández, R. (2025, 13 febrero). Las apuestas y los juegos de azar en el mundo: Datos estadísticos. https://es.statista.com/ [↑](#footnote-ref-2)
3. Mouzo, J. (2024, 24 octubre). La amenaza de llevar un casino en el bolsillo: 80 millones de adultos sufren adicción al juego. El País. https://elpais.com [↑](#footnote-ref-3)
4. Zuppello, M. (2024, 30 septiembre). Adicción a las apuestas online en Brasil. Infobae. https://www.infobae.com [↑](#footnote-ref-4)
5. (2024, abril) El Pilón. Los favoritos de los aficionados: Los deportes más populares para apostar. https://elpilon.com.co [↑](#footnote-ref-5)
6. Villar, G. (2023, 13 octubre). La cara B de la riqueza que genera el fútbol: el 43% del gasto de los aficionados va a las apuestas online. Relevo. https://www.relevo.com [↑](#footnote-ref-6)