

UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR



Grado en Ingeniería Informática

TRABAJO FIN DE GRADO

Sistema de recomendación de fichajes para equipos de fútbol basado en inteligencia artificial

Autor: Catalin Rotaru Rotaru

Tutor: Juan Jesús Roldan Gómez

JUNIO 2024

Agradecimientos

No sé muy bien como empezar este apartado de mi TFG, pero sí sé que tiene un fin mayor que agradecer a mi entorno por contribuir de manera positiva al proceso de elaboración.

Estos años de carrera han supuesto un conjunto de altibajos emocionales, donde he experimentado desde tristeza hasta la mayor plena felicidad que he sentido en mi vida. Pero sin duda alguna un sentimiento que ha ido creciendo de manera exponencial durante todo este trayecto es mi pasión por la Ingeniería Informática. Esta carrera se ha ido ganando un espacio en mi vida, demostrándome no solo lo apasionante que es este mundo sino también que soy capaz y es mi vocación.

Me gustaría agradecer en primer lugar a mi yo de hace unos años por su autosuperación, incansable constancia, predisposición y fuerza que depositó en si mismo, pero sobre todo por seguir intentándolo sin descanso.

En segundo lugar, me gustaría agradecer a mi tutor Juan Jesús por darme la oportunidad de elegir con total libertad un tema de mi interés, por apoyarme y acompañarme en esta última etapa.

Por último, me gustaría agradecer a toda mi familia, por su confianza y paciencia, por soportar mis nervios e incertidumbres. Gracias, mamá, papá y hermana.

Resumen

En el mundo del fútbol no existe una aplicación precisa para poder predecir el costo de transferencia de los jugadores, por ello, en este trabajo de Fin de Grado se busca llevar a cabo un software encargado de realizar este tipo de predicciones haciendo uso de modelos de inteligencia artificial recopilando información de diversas páginas webs.

Entre los objetivos se encuentran el desarrollo de un software efectivo para realizar predicciones; recopilación de datos sobre las estadísticas de jugadores de las principales ligas europeas; limpieza y procesamiento de estos datos; búsqueda de información sobre el valor de los jugadores en el mercado; investigación y selección de algoritmos eficaces; construcción de modelos predictivos y entrenamiento y validación de estos modelos.

Se han explorado varios modelos como DecisionTreeClassifier y LinearRegression pero no se han obtenido los resultados esperados por lo que finalmente se han seleccionado los modelos K-Nearest Neighbors para obtener los jugadores más parecidos a un jugador que se le busca reemplazo y una red neuronal para poder predecir el costo de transferencia de los jugadores obteniendo un error cuadrático medio de 0.0357 y de coeficiente de determinación 0.9537.

Tras la realización de esta aplicación se pueden observar que los resultados obtenidos son óptimos, ya que se consigue de forma eficaz que el software lleve a cabo la predicción de datos de los jugadores que se le piden.

Abstract

In the world of football there is no precise application for predicting the transfer cost of players, therefore, this Final Degree project aims to develop software for making this type of predictions using artificial intelligence models by collecting information from various websites.

Among the objectives are the development of an effective software for making predictions; collection of data on the statistics of players in the main European leagues; cleaning and processing of this data; search for information on the value of players in the market; research and selection of effective algorithms; construction of predictive models and training and validation of these models.

Several models have been explored, such as DecisionTreeClassifier and LinearRegression, but the expected results have not been obtained, so finally the K-Nearest Neighbors models have been selected to obtain the most similar players to a player who is looking for a replacement and a neural network to predict the transfer cost of the players, obtaining a mean square error of 0.0357 and a coefficient of determination of 0.9537.

After carrying out this application, it can be seen that the results obtained are optimal, as the software effectively predicts the data of the players requested.

Índice

<i>Agradecimientos</i>	III
<i>Resumen</i>	5
<i>Abstract</i>	VII
<i>Índice</i>	IX
<i>Índice de figuras</i>	XI
<i>Índice de ilustraciones</i>	XIII
<i>Glosario</i>	XV
1 Introducción	1
1.1. Motivación	2
1.2. Objetivos	3
1.3. Estructura del documento.....	4
2 Estado del arte	5
2.1. Técnicas aplicadas	7
2.1.1. Correlación	8
2.1.2. K-Nearest Neighbors	9
2.1.3. Red Neuronal	10
2.2. Comparación con trabajos similares.....	11
2.2.1. FIFA 18 transfer value wage model	11
2.2.2. Using Python to predict transfer market values	14
2.2.3. Otros	14
3 Diseño	15
3.1. Obtención de los datos	15
3.1.1. Fbref	15
3.1.1.1. Obtención de los datos Fbref	15
3.1.1.2. Limpieza de datos	16
3.1.2. TRANSFERMARKT	17
3.2. Objetivo de la aplicación.....	17
3.2.1. Implementación del algoritmo k-NN	18
3.2.2. Implementación de la red neuronal	19
3.2.2.1. Activación de Unidad Lineal Rectificada (ReLU)	24
3.2.2.2. Estimación Adaptativa de Momentos (Adam).....	24
4 Pruebas y desarrollos	25
4.1. Ejemplo 1	26

4.2. Ejemplo 2.....	27
4.3. Ejemplo 3.....	29
4.4. Ejemplo 4.....	31
5 Conclusión y trabajo futuro	35
Bibliografía	37
ANEXOS.....	39
Anexo A Ficheros descargados	39
Anexo B Tfg comparados	62

Índice de figuras

Figura 1 Evolución anual de ingresos FIFA por derechos de TV 2007-2022	6
Figura 2 Mapa de correlación FIFA 18	11
Figura 3 Mapa de correlación del fichero de estadísticas estándar	12
Figura 4 Entrenamiento de estadísticas estándar	20
Figura 5 Comparativa de valores reales vs predicciones	20
Figura 6 Valores reales y predichos	21
Figura 7 Valores reales y predichos 2	22
Figura 8 Valores reales y predichos 3	23
Figura 9 Valores reales y predichos 4	24
Figura 10 Función de activación ReLu	24
Figura 11 Sadio Mané Vs Robert Lewandowski.....	27
Figura 12 Aurélien Tchouaméni Vs Casemiro	28
Figura 13 Gabriel Jesús Vs Lacazzette Vs Pépé	30
Figura 14 Wesley Fofana Vs Rüdiger Vs Christensen	33

Índice de ilustraciones

Ilustración 1 Resultados Sadio Mané	26
Ilustración 2 Aurélien Tchousméní.....	28
Ilustración 3 Gabriel Jesús	29
Ilustración 4 Gabriel Jesús 2	30
Ilustración 5 Wesley Fofana	31
Ilustración 6 Wesley Fofana 2	32

Glosario

IA: Inteligencia artificial

ML: Machine Learning

FIFA: Fédération Internationale de Football Association

TFG: Trabajo de Fin de Grado

K-NN: K-Nearest Neighbors

Fee: costo de la transferencia del jugador.

RMSE: Raíz del Error Cuadrático Medio

R²: Coeficiente de determinación

Introducción

En las últimas décadas el mundo del fútbol ha experimentado un crecimiento exponencial en términos de popularidad, tanto a nivel estratégico como económico. La última edición de la Copa Mundial de la FIFA, celebrada entre el 20 de noviembre y el 18 de diciembre de 2022, batió récords de audiencia a escala internacional, superando los 3.500 millones de espectadores de la que tuvo lugar en Rusia en 2018. Solo la final entre Argentina y Francia superó los 1.120 millones. [1]

Este fenómeno ha transformado la realización de fichajes en un aspecto de vital importancia para los clubes y los agentes de los jugadores, por lo que la correcta selección de jugadores y nuevos talentos se puede traducir en la obtención de un gran impacto de un equipo a niveles deportivos y económicos.

Debido a este crecimiento se puede observar un gran avance en las técnicas de recopilación de datos y análisis con el objetivo de crear métricas avanzadas y técnicas de inteligencia artificial (IA) con la finalidad de mejorar la predicción de los fichajes y posible rendimiento de los jugadores de determinados equipos.

Gracias a estas técnicas los equipos especializados son capaces de analizar y procesar gran cantidad de datos, así como también son capaces de identificar patrones ocultos y relaciones complejas entre las variables obtenidas a lo largo de cada partido o temporada.

1.1. Motivación

Esta elección surge porque el fútbol es un deporte dinámico, competitivo y presenta una constante necesidad de mejora y evolución de los equipos para conseguir éxitos.

Este trabajo surge de la necesidad de la sociedad para implementar una herramienta eficaz para la realización de fichajes y sus consecuencias en el ámbito económico y deportivo del club. Una buena toma de decisión a la hora de realizar un buen fichaje se traduce en un buen rendimiento deportivo que implica que el club tenga más ingresos y alcance lograr más títulos, además de la revalorización del propio jugador como la de sus compañeros.

Por ello he decidido implementar una herramienta para la predicción del valor de los fichajes que va a permitir profundizar mis conocimientos técnicos y evaluar el posible impacto de esta en el ámbito deportivo.

Esta herramienta se desarrolla mediante algoritmos Machine Learning (ML), K-Nearest Neighbors (k-NN) y redes neuronales. Estos algoritmos son capaces de identificar patrones ocultos y formar relaciones complejas entre las variables; además dan capacidad de construir modelos más precisos para evaluar las necesidades de un equipo, el verdadero valor de un jugador.

Considero que hacer uso de esta herramienta podría tener un impacto en la estrategia de fichajes, permitiendo identificar y adquirir jugadores de alto potencial y adaptados a las necesidades tácticas del equipo. Además, los resultados podrían servir como inicio de futuros estudios y aplicaciones relacionadas con la predicción en el deporte.

1.2. Objetivos

Los objetivos del TFG llevado a cabo son los siguientes:

1. Desarrollar un software efectivo que se encargue de predecir el costo de la transferencia del jugador.
2. Recopilar datos fiables que incluyen estadísticas de todos los jugadores de las ligas europeas más importantes (Bundesliga, Premier League, Ligue 1, Serie A y La Liga).
3. Limpiar y procesar datos.
4. Buscar información sobre el valor de mercado de los jugadores.
5. Investigar, analizar y seleccionar algoritmos eficaces k-NN y redes neuronales.
6. Construir modelos predictivos con los algoritmos seleccionados.
7. Entrenar y validar estos modelos haciendo uso de técnicas adecuadas para garantizar su precisión y fiabilidad.

El principal objetivo que se persigue consiste en contribuir al avance del conocimiento en el campo de la predicción del costo de transferencia de un jugador, mediante el desarrollo de una herramienta precisa que pueda usarse en la toma de decisiones idóneas en el mundo del fútbol profesional.

En resumen, el objetivo principal es la construcción de un software eficiente para la predicción del costo de transferencia de un jugador en base a sus estadísticas pasadas gracias al uso de técnicas avanzadas de inteligencia artificial y machine learning.

1.3. Estructura del documento

Este documento consta de los siguientes apartados:

En el primer apartado “Introducción” se desarrolla la evolución que ha experimentado el fútbol en las últimas décadas, así como el avance en las técnicas para la recopilación y análisis de los datos a través de la creación de nuevas métricas y técnicas de IA para mejorar la predicción de los fichajes. Dentro de este, también se encuentra la motivación aquí se explica la necesidad que busca cubrir este trabajo para la sociedad, y por último los “Objetivos”, donde a partir de estos se desarrolla el software.

En el segundo “Estado del arte” se explica con más detalles la evolución experimentada en el mundo fútbol en las últimas décadas; los algoritmos empleados para predecir el costo de transferencia de un jugador; y por último una comparativa con TFG similares.

El tercer apartado “Diseño”, se encuentran los datos obtenidos de las páginas: Fbref y Transfermarkt, así como el proceso de obtención de dichos datos. Dentro de este punto se encuentra el “objetivo de la aplicación” con la implementación del algoritmo k-NN y de la red neuronal.

En cuarto lugar “Pruebas y desarrollo” se desglosan las pruebas llevadas a cabo y el desarrollo de los modelos creados para la implementación de esta herramienta. También se encuentran ejemplos del funcionamiento de los modelos de k-NN y de la red neuronal.

En el último apartado “Conclusión y trabajos futuros” se desarrolla la forma en la que se ha cumplido cada objetivo y posibilidades de trabajos futuros al desarrollo de esta herramienta de predicción de fichajes.

Estado del arte

En el contexto actual, el fútbol ha experimentado una gran transformación en las últimas décadas, por lo que la toma de decisiones sobre la incorporación de nuevos jugadores a un equipo se ha convertido en un aspecto fundamental para el éxito de los clubes, tanto a nivel deportivo como a nivel económico. Los equipos aún forman parte del patrimonio cultural y sentimental de las ciudades. Mas allá de esto, las últimas décadas han sido dominadas por la importancia económica y mediática de este deporte.

Existen paralelismos entre la economía del fútbol y las condiciones económicas mundiales. Los pilares fundamentales en los que se sustenta la economía del fútbol se basan en la taquilla, los patrocinadores, los derechos de imagen de los jugadores y los derechos televisivos.

Respecto a la taquilla, según el portal DeportesyFinanzas.com la Premier League y la Bundesliga ocupan año tras año las primeras posiciones de asistencia a sus estadios, por encima de la española, la francesa e italiana. Los modelos de la Bundesliga y la Premier League son diferentes. Los clubes alemanes no incrementan los precios de las entradas a pesar del aumento de la demanda, debido a que tienen que cumplir con la legislación alemana. En cambio, la Premier League tiene un reparto de los derechos televisivos más equilibrado por lo que genera más interés entre sus seguidores, lo que se traduce a una mayor afluencia a los estadios. [4]

Los clubes alemanes fueron los que más dinero recibieron en concepto de traspasos, con un total de 1.106 millones de euros. Portugal llevó a cabo más contrataciones que cualquier otra federación (1017), mientras que los clubes brasileños traspasaron o cedieron al mayor número de futbolistas (1217). [3]

LaLiga sigue perdiendo peso en el mercado de fichajes. La Premier sigue mandando con 2.730 millones de gasto, por los 960 recibidos. Francia, Alemania e Italia están por delante de España, que ocupa el quinto lugar con 432 millones gastados, por los 660 millones recibidos.

Según el Global Transfer Report, los clubes invirtieron 8.883 millones de euros en traspasos internacionales, un aumento del 48.1% con respecto a 2022.

Sobre los patrocinadores y su impacto económico, destaca la influencia del Abu Dhabi United Group, propietario del Manchester City y el New York City entre otros, todos patrocinados por Etihad. Fly Emirates es la marca con más equipos patrocinados, situándose entre los 50 más relevantes de Europa, seguido por Etihad y Qatar Airways, que han apostado por los gigantes futbolísticos europeos. [4]

En España el alto costo de la televisión esta creando una burbuja que no corresponde con los beneficios obtenidos, por ello empresas como Google, Amazon o Facebook se muestran interesadas por comprar los derechos de transmisión televisiva. [4]

En la figura 1, se puede observar la evolución anual de los ingresos de la FIFA, por derechos de televisión de 2007 a 2022. En 2022, la FIFA obtuvo unos ingresos de alrededor 3.000 millones de dólares estadounidenses procedentes de los derechos de televisión, la cifra más alta de las últimas dos décadas [2].

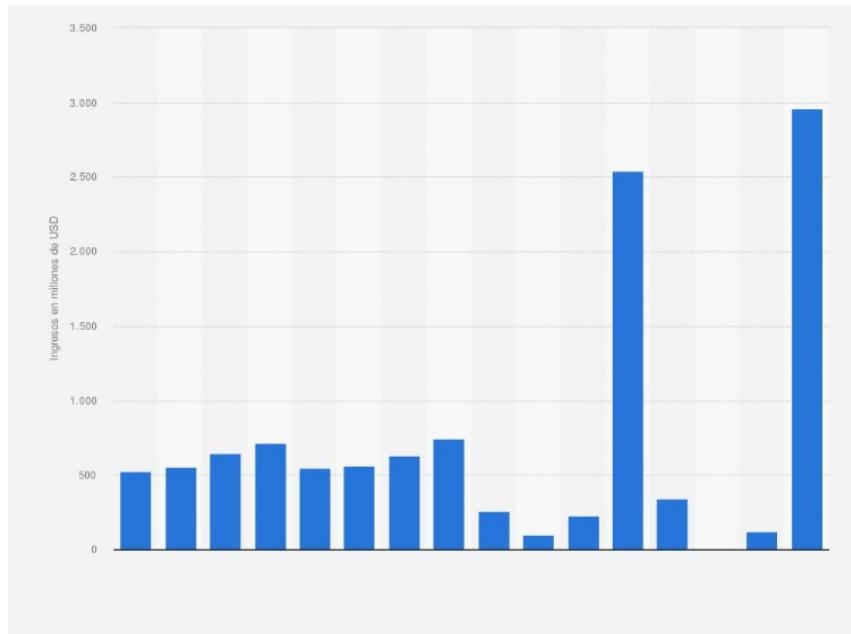


Figura 1 Evolución anual de ingresos FIFA por derechos de TV 2007-2022

Debido a todo esto, surge la necesidad de desarrollar una herramienta basada en inteligencia artificial y machine learning, para la predicción del costo de transferencia de un jugador, en base a sus estadísticas, con el objetivo de pagar un precio acorde al mercado para que el nuevo jugador tenga un rendimiento inmediato y asegure al club éxitos a nivel deportivo y económico.

2.1. Técnicas aplicadas

En este apartado se desarrolla la parte relacionada con la inteligencia artificial de la herramienta, es decir, los algoritmos empleados para predecir el costo de transferencia de un jugador y la preselección de jugadores más parecidos al que vamos a sustituir.

El tipo de problema que se va a abordar es un problema de regresión debido a que se busca predecir una variable de respuesta cuantitativa a partir de varias entradas.

Se ha llevado a cabo el siguiente proceso:

- En primer lugar, se han obtenido los datos más correlacionados con la variable a predecir a través de un mapa de correlación, y se han filtrado las variables que superan un cierto umbral.
- En segundo lugar, se ha procedido a entrenar los modelos con las variables más correlacionadas debido a que son las que más información otorgan y con las que mejor resultado se va a obtener.

Para medir la capacidad de predicción de los algoritmos se ha utilizado dos métricas, que son muy usadas para evaluar el rendimiento en los problemas de regresión:

- RMSE (Raíz del error cuadrático medio): es una medida que indica la magnitud del error en las predicciones.
- R² (Coeficiente de determinación): es una medida que indica la varianza en la variable dependiente que es predecible a partir de las variables independientes.

Todos estos algoritmos han sido probados en un entorno virtual creado a través de Pipenv, haciendo uso del lenguaje de programación Python, y de las librerías que nos ofrece para la creación de modelos de IA y ML.

- **Scikit Learn**: librería enfocada al aprendizaje automático, esta incluye los siguientes algoritmos de regresión:
 - Train test split: se encarga de dividir en dos mitades un conjunto de datos, una de ellas para realizar entrenamientos y otra para validación
 - MinMaxScaler: transforma las características escalándolas a un rango dado.
- **Tensorflow**: permite construir y entrenar redes neuronales para detectar y reconocer patrones. Dentro de esta librería se encuentra **Keras**, esta es una biblioteca de aprendizaje automático de alto nivel que permite construir redes neuronales.

2.1.1. Correlación

La correlación es una medida estadística que indica el grado de relación entre dos variables.

En total se tienen 11 ficheros de datos con distintos conceptos, estos a su vez desglosados en subconceptos que se encuentran en el anexo A Ficheros descargados.

- **Estadísticas estándar:** tiempo jugado, rendimiento, expectativa, progresión y por 90 minutos de juego.
- **Porteros:** tiempo jugado, rendimiento, y tiros penales.
- **Portería avanzada:** goles, expectativa, iniciado, pases, saques de meta, pases cruzados y barredora.
- **Tiros:** estándar, expectativa.
- **Pases:** total, cortos, medios, largos, expectativa.
- **Tipos de pases:** tipos de pases, saques de esquina, resultados.
- **Creación de goles y tiros:** ACT, tipos de SCA, ACG, tipos de GCA
- **Acciones defensivas:** derribos, desafíos, bloqueos.
- **Posesión del balón:** toques, tomas, transportes, recepción.
- **Tiempo jugado:** tiempo jugado, titular, sup., logros del equipo, logros del equipo (xG).
- **Estadísticas diversas:** rendimiento, duelos aéreos.

Para crear esta matriz de correlación se usa el método “Pearson”, este se encarga de realizar la prueba que mide la relación estadística entre dos variables continuas.

Al hacer uso de la matriz de correlación se realiza una selección de las variables que están correlacionadas por encima del umbral que se ha definido (0.15), esto indica una correlación débil pero lo suficientemente significativa para demostrar una cierta relación entre las variables. Por ello solo tendrán en cuenta para la creación del modelo las variables que superen este coeficiente de correlación.

Tras llevar a cabo lo anterior, se puede reducir la entrada de datos a la red neuronal con el objetivo de eliminar las variables que no aportan demasiada información al modelo y mejorar el rendimiento de este. De esta forma se simplifica el modelo y se facilita la interpretación de los resultados obtenidos.

El valor del coeficiente de correlación varía entre -1 y 1, es decir: [5]

- Un valor entre 0 y 1 indica una correlación positiva entre estas variables, esto significa que al aumentar una variable la otra también aumenta en proporción a esta constante obtenida.
- Un valor entre -1 y 0 también indica correlación negativa, esto quiere decir que al aumentar una variable la otra disminuye con esta proporción.
- Un valor 0 o cercano a este indica la falta de correlación entre las variables.

2.1.2. K-Nearest Neighbors

El algoritmo K-Nearest Neighbors también conocido como k vecinos más cercanos, es un clasificador de aprendizaje supervisado (este aprendizaje supervisado utiliza conjuntos de datos etiquetados, es decir, cada columna se corresponde con las distintas características de los jugadores) no paramétrico y se basa en la proximidad para realizar clasificaciones o predicciones sobre la agrupación de un punto de datos individual. [6] y [7]

Este algoritmo tiene 2 usos: para realizar regresiones, o para clasificaciones sobre un conjunto de datos; lo más común es usarlo para esto último, suponiendo que se pueden encontrar puntos similares cerca uno del otro.

Se ha decidido usar este algoritmo como un problema de clasificación ya que el objetivo de este es identificar los vecinos más cercanos de un punto de consulta dado, es decir, dado un jugador que queremos reemplazar, va a buscar los jugadores más parecidos a este en términos estadísticos, de tal forma que el nuevo jugador que se va a incorporar al nuevo equipo sea lo más similar posible.

Para el desarrollo de esta herramienta no se han tenido en cuenta las formaciones que han usado los equipos ni los estilos de juego de cada uno de ellos, esto es un aspecto fundamental para poder evaluar un posible fichaje, puesto que se considera que el contexto táctico donde ese jugador va a poner sus habilidades al servicio del club repercute en su rendimiento.

Para determinar qué puntos están más cerca del punto de consulta hacemos uso del algoritmo ball-tree, encargado de obtener una selección de k jugadores parecidos al de entrada. Este parámetro se puede modificar en cualquier momento para poder obtener más o menos jugadores parecidos al que vamos a buscar sustituto y de esta forma se obtiene una base de jugadores de los cuales se tiene que decidir cuál va a ser el sustituto del jugador y poder realizar predicciones en base a las estadísticas de estos jugadores sobre su posible fee.

2.1.3. Red Neuronal

Una red neuronal es un método de IA que enseña a las máquinas a procesar una gran cantidad de datos y su funcionamiento está inspirado en la forma en la que lo hace el cerebro humano. Consta de varias capas, una de entrada donde se reciben los datos; una o varias capas ocultas interconectadas que realizan el procesamiento de la información y una capa de salida donde se obtiene el resultado del procesamiento. Gracias a esto, las computadoras son capaces de crear sistemas adaptables que usan para aprender de sus propios errores de esta forma, mejoran continuamente y se pueden aplicar para la resolución de problemas complicados. [8]

Las redes neuronales artificiales están formadas por neuronas artificiales que trabajan juntas para resolver un problema, estas neuronas son módulos de software que se denominan nodos, y las redes neuronales artificiales son programas de software o algoritmos que hacen uso de sistemas informáticos para resolver cálculos matemáticos.

A continuación, se explica la arquitectura de una **red neuronal simple** que tiene neuronas artificiales interconectadas en tres capas.

- **Capa de entrada:** en esta capa es donde la información del problema entra en la red neuronal, los nodos de entradas procesan los datos, los analizan o clasifican y los pasan a la siguiente capa.
- **Capa oculta:** tiene como entrada la salida de la capa de entrada o la salida de otra capa oculta. Las redes neuronales artificiales se pueden configurar para tener una gran cantidad de capas ocultas en función de la complejidad del problema que se quiere resolver. Cada capa oculta analiza la salida de la capa anterior, la vuelve a procesar y pasa la información a la siguiente capa.
- **Capa de salida:** en esta capa se obtiene el resultado final del procesamiento de todos los datos que realiza la red neuronal. Aquí se pueden tener uno o varios nodos. En caso de tener un problema de clasificación binaria, donde la salida sólo puede tener 2 valores es suficiente con tener un único nodo de salida, en cambio, si vamos a tratar un problema de clasificación de varias clases, esta capa deberá estar formada por más de un nodo en la capa de salida.
- **Función de activación:** es un filtro con función limitadora o umbral encargado de modificar el valor resultado o imponer un límite que debe superpasar la neurona de salida para proseguir a otra neurona. Dependiendo del tipo de problema planteado se usará una función u otra.

2.2. Comparación con trabajos similares.

2.2.1. FIFA 18 transfer value wage model

Este trabajo FIFA 18 transfer value wage model [9] hace uso del conjunto de datos y estadísticas de los jugadores del videojuego FIFA 18 de la empresa Electronic Arts para predecir el valor de mercado y el salario de los jugadores.

En la primera etapa se realiza un análisis exploratorio de los datos que busca la correlación entre las características de un jugador y el valor de mercado y su salario, como se observa en la figura 2.

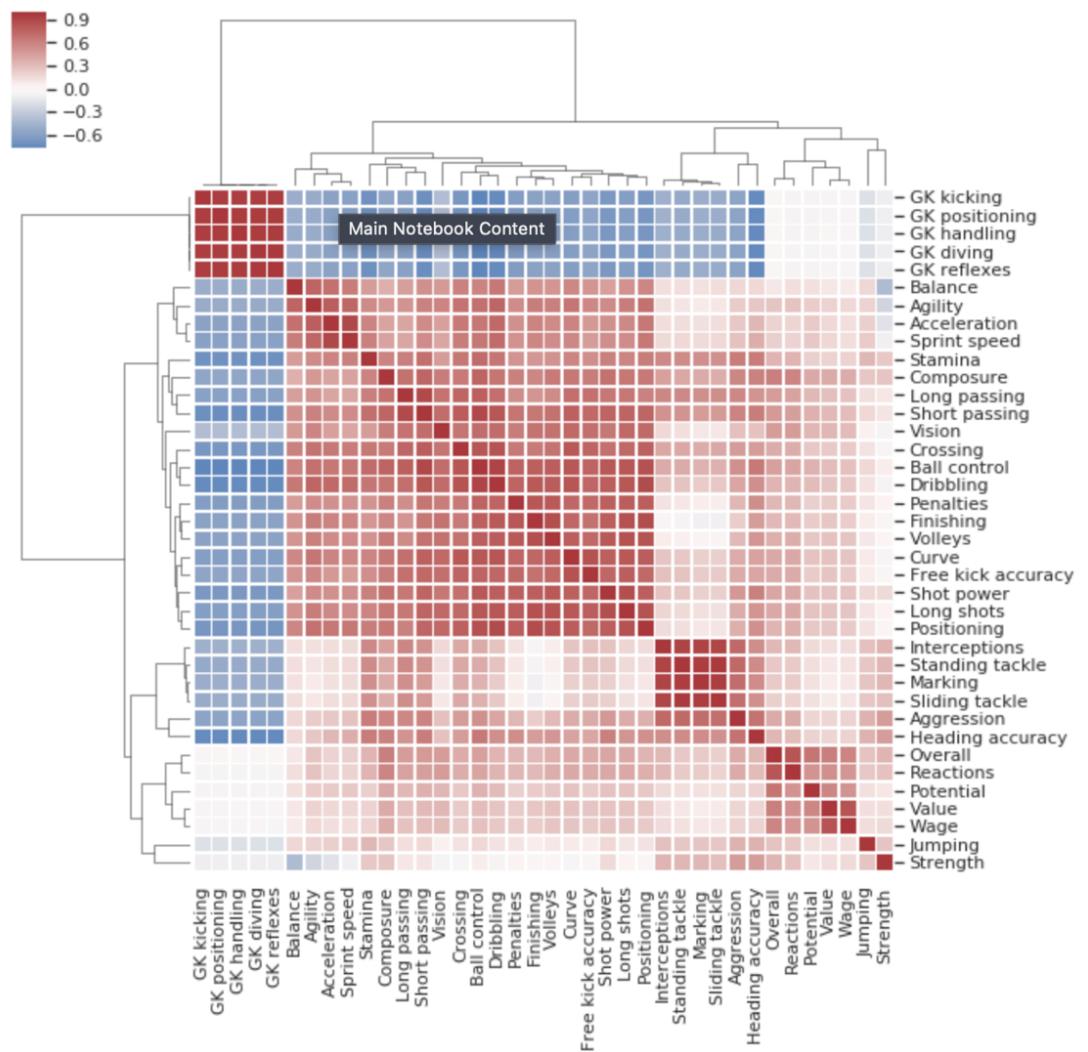


Figura 2 Mapa de correlación FIFA 18

En la segunda etapa realiza un modelo de aprendizaje automático para predecir el valor y el salario de los jugadores. Para determinar los hiperparámetros óptimos hace uso de XGBoost, posterior al entrenamiento mide la precisión del modelo en con un conjunto de pruebas y lo compara con un Error Medio de Referencia.

En la figura 3 se observa el mapa de correlación que se ha obtenido en este TFG, del fichero de “estadísticas_estándar.csv”.

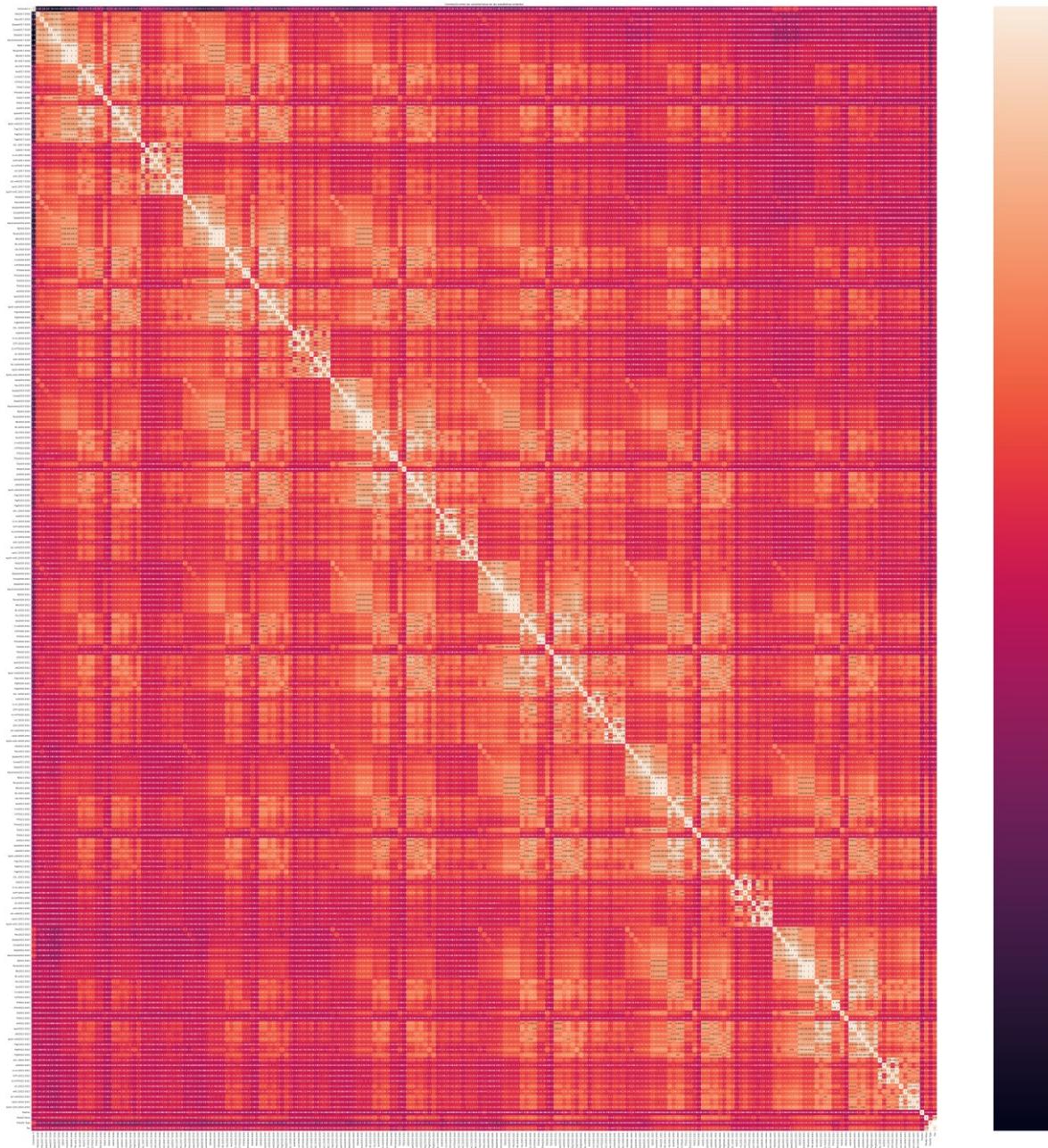


Figura 3 Mapa de correlación del fichero de estadísticas estándar

Se ha realizado la correlación de todos los ficheros y se han guardado en la carpeta /data/correlaciones en nuestro repositorio.

Diagonal principal: los nombres de las filas y las columnas que coinciden, esto significa que se está intentando correlacionar la misma estadística por lo que siempre se va a obtener el valor 1.

Patrones de cuadrados o tablero de ajedrez repetitivo:

De izquierda a derecha los datos van teniendo más intensidad, es decir, están más correlacionados con el “Fee” del jugador

Los datos están ordenados desde la temporada 2017-2018 hasta la temporada 2021-2022 por lo que se pueden observar unos patrones y es que, cuanto más a la izquierda, más antiguas son dichas características, y por lo tanto se correlacionan menos con el objetivo. Se aprecia que los valores de las últimas dos temporadas son los más interesantes para tener en cuenta a la hora de usarlos de cara a la realización de este modelo predicción del valor de un jugador.

Se pueden observar ciertas similitudes y diferencias en los dos trabajos, entre ellas:

Similitudes:

1. Los dos buscamos predecir el valor de transferencia de los jugadores en base a unas estadísticas.
2. Los dos realizamos un estudio previo de los datos y hacemos uso de algoritmos basados en IA.

Diferencias:

1. En su trabajo hace uso de una base de datos del juego FIFA18 donde los datos están sesgados, los jugadores tienen un potential rating que en la mayoría de las veces no se corresponden con el potencial real de los jugadores.
2. El valor de mercado que el busca predecir es el valor del juego y no el valor de la vida real.
3. En mi TFG obtengo datos reales de los jugadores de los últimos 5 años y los aglutino todos en un dataset, además del valor real de transferencia de los jugadores del año pasado temporada 2022/2023 y en base a esto realizamos el entrenamiento y las predicciones.
4. Podemos observar varias diagonales debido a la columna “país”.
5. Origen de datos: los datos que he obtenido son datos reales realizados por los jugadores en los partidos que han disputado. En cambio, los datos del juego se realizan mediante algoritmos que tienen en cuenta ciertas características de los jugadores de temporadas anteriores, estos datos son estimaciones y pueden no reflejar la realidad.
6. Los datos obtenidos en los partidos reales están compuestos por multitud de métricas, en cambio, los datos de los videojuegos están limitados a un conjunto más sencillo.
7. Finalidad: los datos de los partidos reales se usan para poder realizar métricas del rendimiento de los jugadores, en cambio, los datos del juego se usan para crear una experiencia lo más realista posible y divertida para los jugadores.

2.2.2. Using Python to predict transfer market values

En este trabajo “Using Python to predict transfer market values” [21] se hace uso de los datos recopilados por la empresa KPMG, el dataset contiene datos de más de 8000 jugadores de las principales ligas europeas, sudamericana y asiática en su versión premium. En la versión gratuita incluye datos de 220 jugadores con los que realiza el modelo para poder realizar predicciones, pero antes de ello trata los datos para que puedan servir como entrada a los modelos empleados.

Similitudes:

Enmascaramiento de datos: Esto consiste en asignar valores numéricos a las variables de entrada que están formadas por datos categóricos (como pueden ser cadenas de caracteres en el valor del nombre del club, el nombre de la liga a la que pertenece cada jugador, la posición que ocupa cada jugador en el terreno de juego, ó la nacionalidad).

Diferencias:

Este trabajo hace uso de algoritmos de regresión entre los cuales destacamos k-NN, árboles de decisión, bosques aleatorios, gradient boosting. En cambio, en mi herramienta se crea una red neuronal.

2.2.3. Otros

En el TFG “Estudio de técnicas de data science para la predicción de rendimientos deportivos” [Anexo B] se busca predecir el rendimiento del año siguiente de un jugador en base a estadísticas pasadas haciendo uso de los siguientes algoritmos: Grid Search, Elastic Net, Random Forest y XGBoost. En el TFG “Análisis de resultados de encuentros de fútbol mediante técnicas de machine learning” [Anexo B] se crea un sistema de predicción de resultados de fútbol para generar beneficios mediante apuestas deportivas, haciendo uso de técnicas de inteligencia artificial a través de modelos K-NN, regresión logística, máquina de soporte vectorial y clasificación mediante random forest.

Diseño

3.1. Obtención de los datos

Los datos para llevar a cabo esta herramienta se han obtenido mediante la técnica de Web Scraping (extracción de datos de la web) en las siguientes páginas:

3.1.1. Fbref

Fbref [12] es un sitio web dedicado al seguimiento de estadísticas de equipos de fútbol y jugadores de todo el mundo. Se lanzó en 2018 con cobertura para 6 naciones, pero estos países sólo abarcan una pequeña parte del fútbol mundial, poco a poco ha ido añadiendo más países llegando a cubrir más de 40, así como ligas nacionales de copa, super copa y ligas juveniles de los principales países europeos, además incluyen cobertura para los torneos internacionales como la UEFA Champions League o la Copa Libertadores. Colaboran con la empresa Opta Sports para poder incluir datos analíticos avanzados. [11]

3.1.1.1. Obtención de los datos Fbref

Se obtienen los datos juntos de las 5 grandes ligas europeas: Premier League, Bundesliga, Serie A, Ligue 1 y La Liga. Y se encuentran estadísticas de los equipos que han conformado dichas ligas en la temporada 2017-2018, así como las de los jugadores que han formado parte de estos equipos en este año.

Para la obtención de los datos estadísticos los jugadores se descargan los datos desde la temporada 2017-2018 hasta la 2022-2023 ambas incluidas. [13]

Se ha decidido comenzar desde la temporada 2017-2018 debido a que en el apartado de “Estadísticas del Jugador” es cuando se comienza a tener datos más completos para la realización de este estudio. Las estadísticas de los jugadores descargadas son:

- Estadísticas Estándar
- Porteros
- Portería Avanzada
- Tiros
- Pases y tipos
- Creación de goles y tiros
- Acciones defensivas
- Posesión del balón
- Tiempo jugado
- Estadísticas diversas

Se obtienen un total de 11 ficheros por cada temporada, desde la temporada 2017-2018 hasta la temporada 2022-2023 tenemos un total de 66 ficheros con todas las estadísticas realizadas por todos los jugadores que han formado parte de las principales ligas europeas.

En el anexo A se indica el significado de cada una de las columnas que contiene el dataframe descargado.

Todos estos datos son suficientes para crear un dataset completo y así poder encontrar jugadores similares a uno dado y entrenar una red neuronal para predecir el fee de un jugador los estos datos se encuentran en la carpeta /data/original_data/XXXX_YYYY, donde XXXX e YYYY hace referencia a la temporada que contienen los datos.

Todo el contenido de este trabajo en cuanto a código se encuentra en el siguiente repositorio: <https://github.com/CatalinRotaru8/IA-TFG>

3.1.1.2. Limpieza de datos

Para poder usar estos datos en modelos de IA es necesario realizar un preprocesamiento de estos con el objetivo de garantizar que sean los adecuados para usarlos en los modelos elegidos.

Para ello se realizan los siguientes pasos:

- 1) Eliminación de la cabecera, los datos se descargan de 25 en 25 filas y se duplica la cabecera, así como la última columna, que hace referencia a los partidos que ha jugado cada jugador y no es relevante.
- 2) Tras esto se observa que contamos con jugadores duplicados, esto significa que el jugador ha cambiado de equipo a mitad de temporada, por lo que se suman sus estadísticas de ambos equipos.

*Todos estos cambios se pueden observar en los ficheros de la capeta /data/processed_data/XXXX_YYYY

- 3) Una vez obtenidos cada uno de los ficheros de cada temporada, con los cambios mencionados anteriormente, se unen todos los que contienen las mismas características en uno solo, es decir, todos los ficheros que se llaman EstadisticasEstandar.csv de todas las temporadas se juntan en un solo fichero, y así sucesivamente con los 10 restantes. De esta forma se obtiene un dataset único por cada tipo de estadística que será de gran utilidad para poder obtener la correlación entre estas variables y el Fee de cada jugador.

*Estos ficheros se encuentran en la carpeta /data/ processed_data/all_years/.

- 4) Transformación de los datos a tipo numéricos, ya que las columnas que hacen referencia al país, posición, equipo y competición son datos de tipo categórico. Esto se realiza haciendo uso de LabelEncoder (biblioteca de sklearn.preprocessing).

*Esto se encuentra en /data/ processed_data/replace_strings /

- 5) A los datos anteriores se le añaden los necesarios del fichero de transfermarkt (position, Market Value, TransferType, Fee) y de esta forma se tiene un dataset con la

información procesada para poder usarlo como entrada a los métodos de inteligencia artificial.

*Estos nuevos ficheros se encuentran en data/input_data/

3.1.2. TRANSFERMARKT

Transfermarkt [14] es una página de origen alemán que contiene información del mundo del fútbol como puede ser: puntuaciones de los jugadores, resultados, noticias de transferencias de mercado, valores de los jugadores, etc.

Esta web extrae valores de todos los foros y determina una estimación que es la que publica como valor de mercado.

Los datos que contienen la información de los traspasos realizados en la temporada 2022 se encuentran en /data/TransferMarket/2022/transfermarket_2022.csv

Este archivo se ha obtenido del Github del autor Farhan Hasin Saad con título “Transfermarkt Data Analysis” a fecha del 12 de abril del 2024. [15]

Este fichero contiene las siguientes columnas:

- **Name:** nombre del jugador
- **Age:** edad del jugador
- **Position:** posición en la que juega el jugador
- **Player Nationality:** nacionalidad del jugador
- **MarketValue:** Valor de mercado que indica transfermarkt del jugador en base a sus métricas.
- **Club:** club de destino del jugador
- **League:** nombre de la liga a la que pertenece el club destino
- **League Country:** país al que pertenece la liga del club destino
- **TransferType:** tipo de transferencia (Permanent, Loan, Free)
 - o **Permanent:** vendido
 - o **Loan:** cedido
 - o **Free:** gratis
- **Fee:** valor que ha pagado el club destino por el jugador

3.2. Objetivo de la aplicación

El principal objetivo de este TFG es crear una herramienta suficientemente efectiva para realizar la predicción del costo de la transferencia de un jugador para que sea útil a un director deportivo.

Para llevar a cabo esta herramienta es necesario realizar una preselección de jugadores que se quieren incorporar a un equipo y una vez se tenga esta preselección se procede a realizar mediante la red neuronal la estimación fee de dichos jugadores.

Después de realizar el preprocesamiento de los datos obtenidos de la web de Fbref, se encuentra un dataset lo suficientemente amplio y con información muy relevante sobre los aspectos futbolísticos desarrollados por los futbolistas de las 5 grandes ligas europeas en las últimas 6 temporadas. Por lo que se puede comenzar con la implementación de la herramienta.

3.2.1. Implementación del algoritmo k-NN

En primer lugar, se ha decidido implementar el algoritmo k-NN en esta herramienta ya que ofrece numerosas ventajas a la hora de buscar jugadores parecidos al que se quiere reemplazar con los datos procesados. [16]

Implementar y entender el algoritmo es una tarea sencilla, ya que no es necesario usar fórmulas matemáticas complejas o técnicas de optimización, k-NN se basa en medir la distancia entre los datos de la muestra y un valor k que será el número de jugadores parecidos a mostrar.

Otra ventaja que tiene este algoritmo es que no necesita construir modelos a partir de los datos de entrenamiento. También es un algoritmo que se adapta a todos los cambios realizados en los datos y da la posibilidad de trabajar con datos ruidosos y no lineales.

Para no trabajar con todos los datos descargados, sólo se ha tenido en cuenta el fichero de estadísticas estándar, que engloba unos datos más genéricos y obtenemos mejores resultados.

Como se ha mencionado anteriormente, este algoritmo tiene dos posibles usos, clasificador o regresor, para esta herramienta se usa como clasificador debido a que se busca la similitud entre jugadores.

El entrenamiento recibe como datos de entrada las columnas más correlacionadas con el objetivo y son mayores a un umbral predefinido, de tal forma que resulte más sencillo realizar el cálculo de las similitudes entre jugadores y así obtener un mejor rendimiento del modelo. Otra forma de comparar a los jugadores es mediante un gráfico de radar mostrando todas estas variables de entrada.

Una vez se tienen seleccionados los posibles jugadores sustitutos el segundo paso consiste en realizar predicciones de su posible valor de mercado gracias a la implementación de una red neuronal.

3.2.2. Implementación de la red neuronal

Mediante la implementación de la red neuronal se puede predecir el fee de los jugadores que queremos reemplazar en nuestro equipo.

En primer lugar, los datos de entrada a la red se escalan utilizando “MinMaxScaler” para obtener las estadísticas normalizadas entre 0 y 1, de esta forma se facilita el entrenamiento de la red neuronal.

El conjunto inicial de datos se divide en dos subconjuntos, un conjunto de entrenamiento que contiene el 80% de los datos y otro conjunto de prueba que contiene el 20% restante, que sirve para evaluar el rendimiento del modelo.

La red neuronal está formada por tres capas conectadas de 64 neuronas cada una, la función de activación empleada se denomina ReLu. En la capa de salida sólo se encuentra una neurona para realizar la predicción. Una vez definido el modelo de la red neuronal, se compila haciendo uso del optimizador Adam y como función de pérdida “mean_squared_error”, para minimizar el error cuadrático medio.

Tras compilar el modelo, se realiza un entrenamiento de 100 épocas con un tamaño de batch de 16, usando el 20% del conjunto de entrenamiento como conjunto de validación.

Para poder ver si el modelo es óptimo para realizar predicciones sobre el conjunto de prueba y se calcula el RMSE obteniendo un valor de 0.0357, este valor se considera bajo e indica que las predicciones del modelo se desvían muy poco de los valores reales. Es decir, cuanto menor sea el valor, mejor será el rendimiento del modelo. El valor obtenido de R^2 es de 0.9537, este se ajusta muy bien a los datos e indica que las variables independientes explican una gran parte de la variabilidad en la variable objetivo.

Con esto podemos concluir que el modelo es lo suficientemente robusto para realizar predicciones.

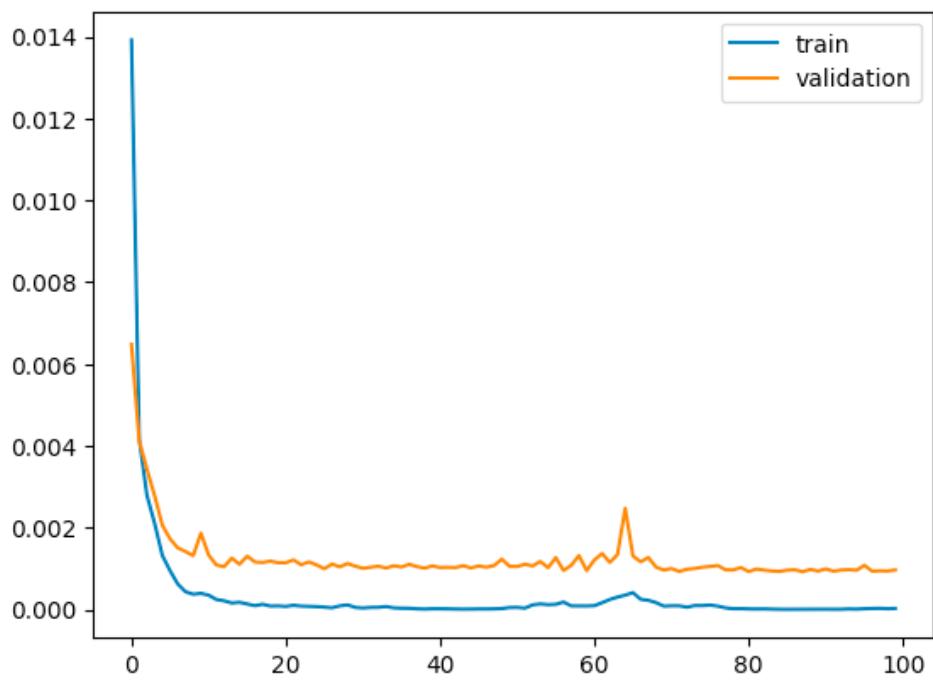


Figura 4 Entrenamiento de estadísticas estándar

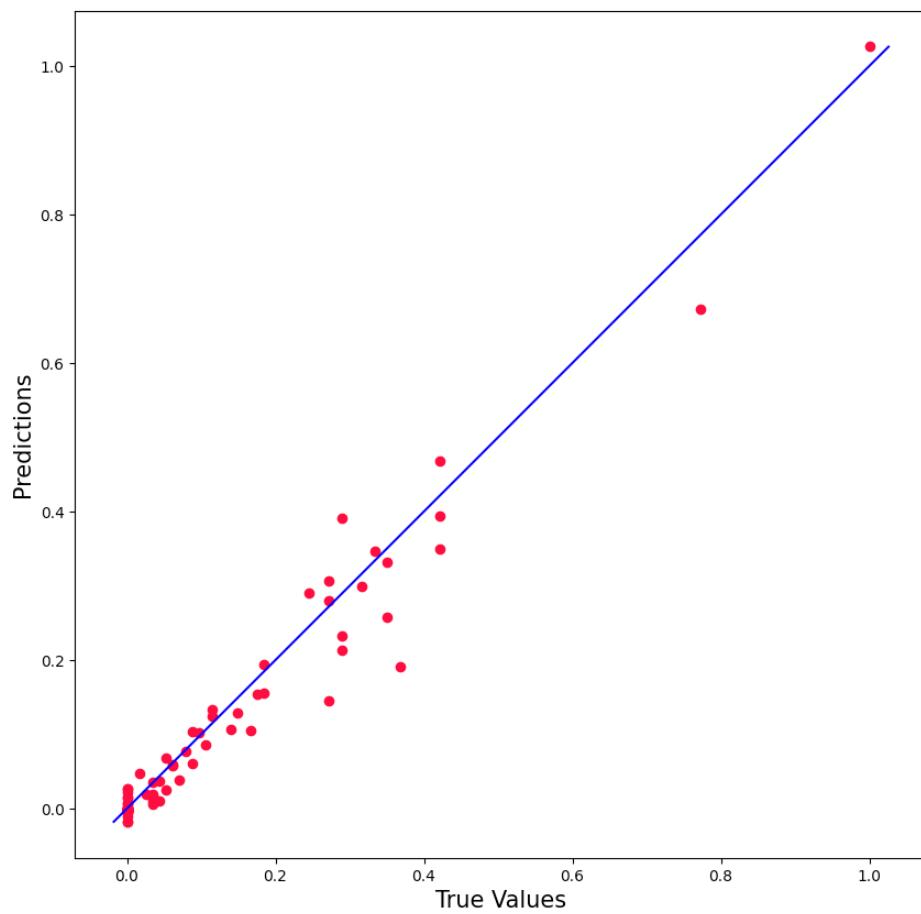


Figura 5 Comparativa de valores reales vs predicciones

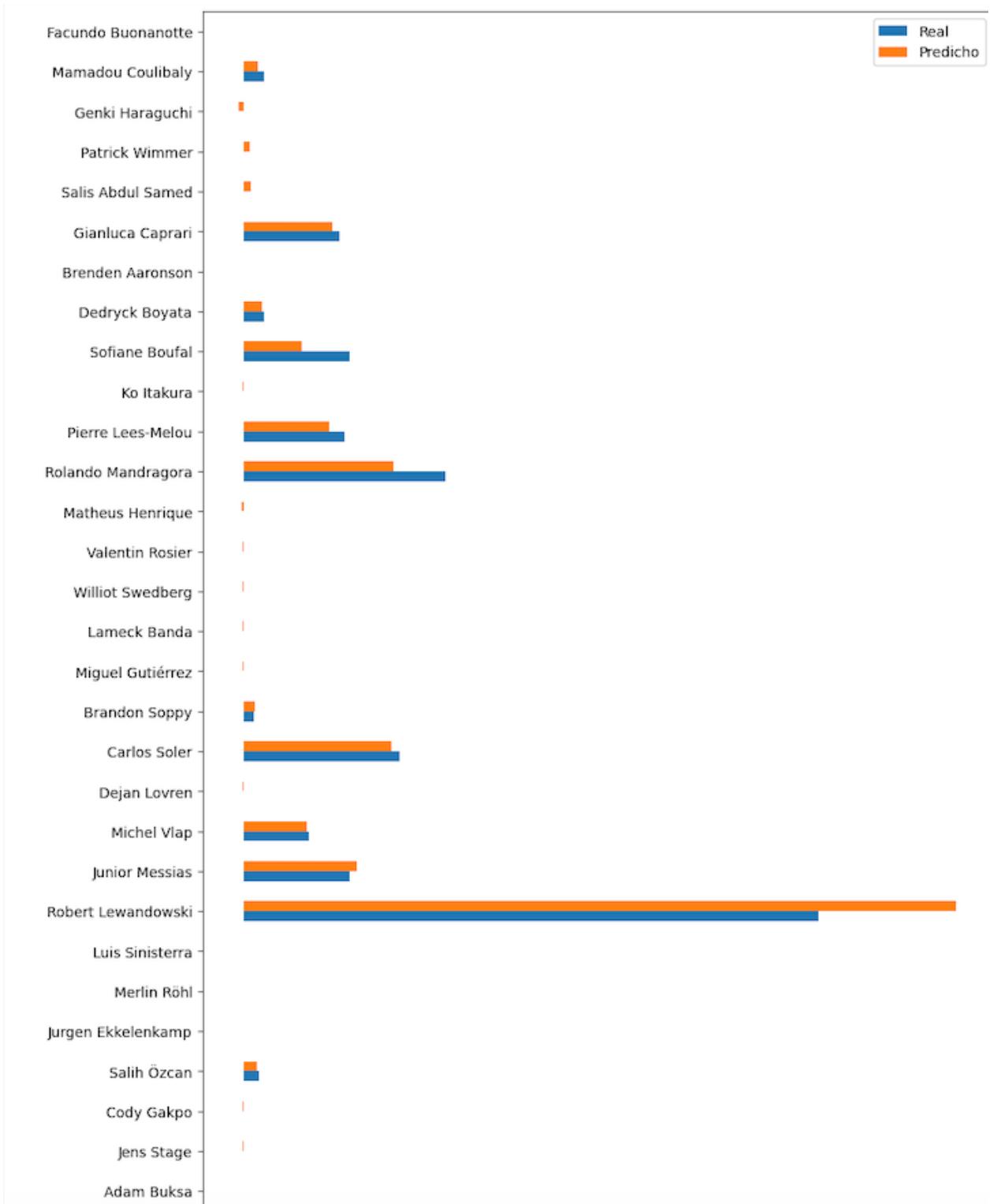


Figura 6 Valores reales y predichos

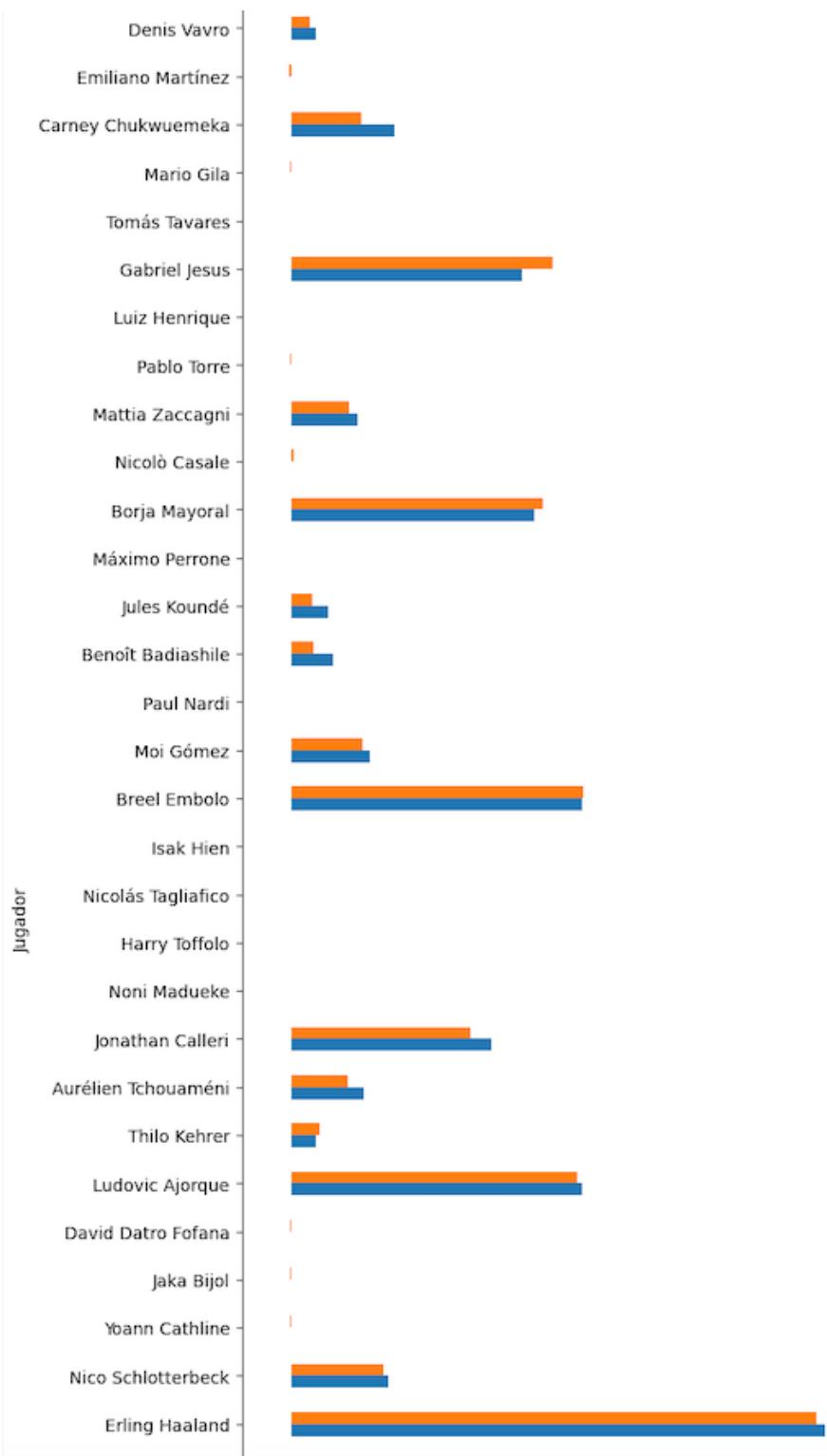


Figura 7 Valores reales y predichos 2

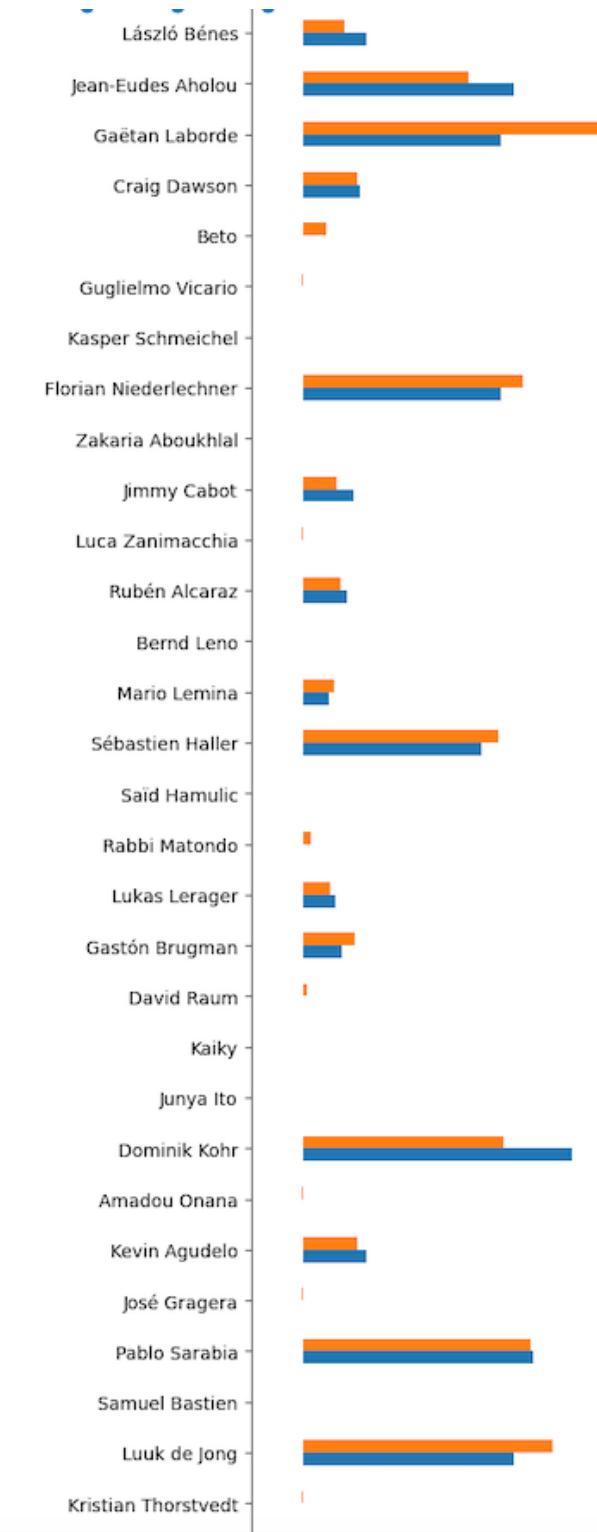


Figura 8 Valores reales y predichos 3

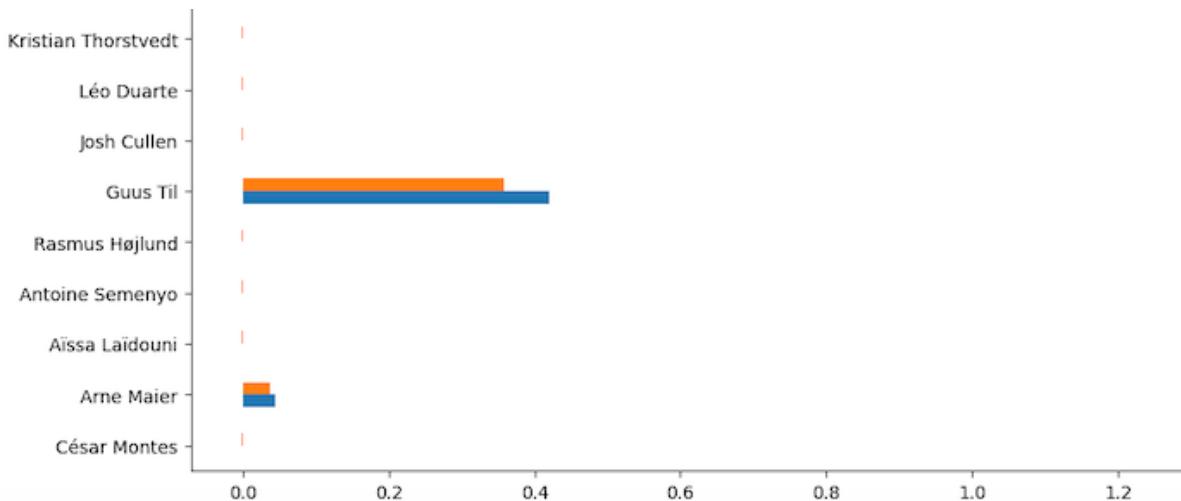


Figura 9 Valores reales y predichos 4

3.2.2.1. Activación de Unidad Lineal Rectificada (ReLU)

Esta función de activación (figura 10) devuelve el valor 0 para valores negativos y la propia entrada para valores positivos.

Para entradas mayores a 0, actúa como una función lineal, lo que significa que no modifica la escala de las entradas positivas. En cambio, produce 0 para todas las entradas negativas, lo que provoca activaciones dispersas, es decir, sólo se activa un subconjunto de neuronas derivando en una computación más eficiente. [17]

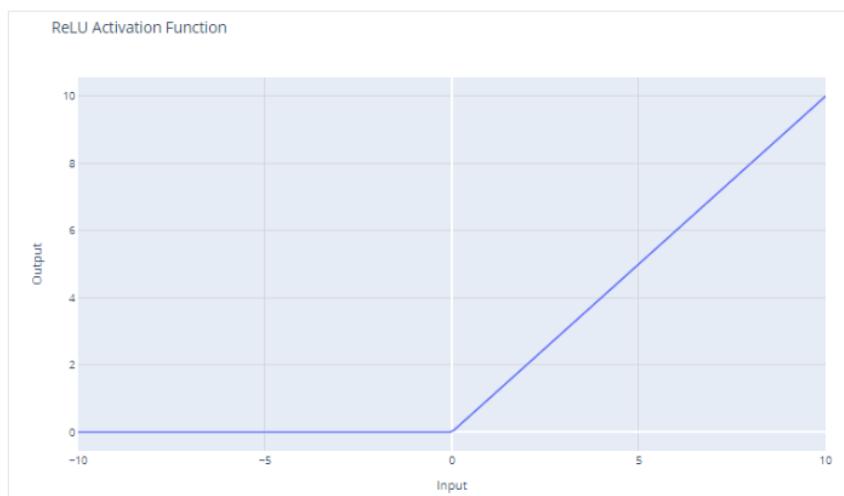


Figura 10 Función de activación ReLu

3.2.2.2. Estimación Adaptativa de Momentos (Adam)

Es un algoritmo de optimización utilizado en aprendizaje automático y redes neuronales, ajusta la tasa de aprendizaje durante el entrenamiento para cada parámetro basándose en gradientes pasados, es decir, la derivada de una función con más de una variable de entrada. Esto ayuda a mejorar la convergencia del entrenamiento. [18]

Pruebas y desarrollos

En este apartado se detallan las pruebas y el desarrollo de los modelos creados para la realización de esta herramienta de predicción, en la que se implementa un modelo de K-Nearest Neighbors (K-NN) y una red neuronal.

El modelo de K-NN tiene el principal objetivo de buscar jugadores similares a un jugador que queremos reemplazar, el modelo recibe como entrada dos parámetros: el nombre del jugador al que se le busca reemplazo y un número entero indicando cuántos jugadores similares queremos obtener. Tras recibir estos parámetros, el modelo se encarga de obtener el número de jugadores más similares y también la distancia entre ambos jugadores, esto es una métrica bastante valiosa para poder cuantificar cuánto se parecen dichos jugadores.

Se realiza una correlación de los datos, obteniendo únicamente las columnas que superan el umbral de correlación 0.15 debido a que son las que más información aportan a este modelo.

Estos indicadores son los encargados de señalar el rendimiento del modelo en cuanto a términos de precisión para poder predecir el valor de mercado de un jugador.

- **RMSE:** es una medida que indica la magnitud del error en las predicciones, obteniendo un valor de 0,0357.
- **R²:** es una medida que indica la varianza de la variable dependiente que es predecible a partir de las variables independientes, obteniendo un valor de 0,9537.

Se han creado cuatro funcionalidades:

- 1) **Compare_players:** Esta función recibe 2 jugadores y la temporada a comparar, el primer jugador es el jugador que el equipo vende, y el segundo jugador es el que el equipo compra para cubrir esa posición. Se encarga de comparar las estadísticas de ambos jugadores, indicando si el jugador que cubre esa posición obtiene un mejor rendimiento o no.
- 2) **Calculate_distance:** la función recibe a los 2 jugadores e indica la distancia que hay entre ellos estadísticamente hablando, es decir, cuanto más bajo sea el número que se recibe, los jugadores se parecen más, lo que puede implicar que el fichaje sea óptimo, aunque no siempre es así. Se puede entender como un indicador para medir si el jugador es un buen candidato para sustituir al jugador vendido.
- 3) **Predict_fee:** usa el modelo creado por la red neuronal. Esta función recibe el nombre del jugador que un equipo quiere comprar y devuelve el costo de la transferencia del jugador.
- 4) **Plot_radar_chart:** la función recibe la lista de jugadores que se quiere comparar. Esta lista de jugadores incluye el jugador vendido y los posibles sustitutos para ver de forma gráfica sus estadísticas.

4.1. Ejemplo 1

En la ilustración 1, se puede ver como la red neuronal predice un valor de 52,7 millones de euros para el jugador Sadio Mané, pero se compró por 32 millones de euros.

En el modelo K-NN obtenemos a Mané como posible sustituto de Robert Lewandowski, pero no es el jugador más recomendado y la distancia entre ellos no es la mínima. Cabe destacar que el Bayern Múnich tras perder a su delantero titular necesita realizar una incorporación de nivel, aprovechan comprando a Mané debido a que no llegó a un acuerdo económico con su club (Liverpool) y le quedaba un año de contrato [19], motivo suficiente para fichar a este jugador a un precio muy asequible sabiendo que Lewandowski se vendió por 45 millones de euros.

Con esto se concluye que ambos modelos funcionan bien para este caso, y que el precio de mercado puede ser aproximado debido a que no se tienen en cuenta otros parámetros que podrían ser relevantes, como los años de contrato que le faltan a un jugador, las negociaciones contractuales o el malestar en un equipo.

En la figura 11 se puede observar como el rendimiento de Lewandowski en la temporada 2022-2023 es mejor que el de Sadio Mané.

```

predict_fee('Sadio Mané')
knn(nombre_jugador_entrada=robert_lewandowski, k=20)
compare_players('Robert Lewandowski', 'Sadio Mané', '2022-2023')
calculate_distance('Robert Lewandowski', 'Sadio Mané')
plot_radar_chart(['Robert Lewandowski', 'Sadio Mané'])

✓ 0.6s

1/1 ━━━━━━ 0s 10ms/step
El valor de mercado predicho para el jugador Sadio Mané es: 52.70118713378906 millones de euros
Los 20 jugadores más similares a Robert Lewandowski son:
1 Cristiano Ronaldo con una distancia de 2.0357996322979286
2 Karim Benzema con una distancia de 2.089364053437654
3 Ciro Immobile con una distancia de 2.110854816076695
4 Kylian Mbappé con una distancia de 2.2732807734885894
5 Mohamed Salah con una distancia de 2.344439131830941
6 Harry Kane con una distancia de 2.4933648189045483
7 Lautaro Martínez con una distancia de 2.5736773094955443
8 Erling Haaland con una distancia de 2.5905502013830075
9 Wissam Ben Yedder con una distancia de 2.600091585138029
10 André Silva con una distancia de 2.7949749381100193
11 Sadio Mané con una distancia de 2.854375274568753
12 Wout Weghorst con una distancia de 2.885548586295429
13 Andy Delort con una distancia de 2.9031473053916326
14 Jamie Vardy con una distancia de 2.9154431972774915
15 Duván Zapata con una distancia de 2.9166483282913074
16 Son Heung-min con una distancia de 2.9254899450652743
17 Romelu Lukaku con una distancia de 2.9783134431218397
18 Dušan Vlahović con una distancia de 3.014353676753966
19 Edin Džeko con una distancia de 3.0307387852768146
20 Raheem Sterling con una distancia de 3.035553177181941

Sadio Mané ha sido peor que Robert Lewandowski en el año 2022-2023.
La distancia entre Robert Lewandowski y Sadio Mané es 2.854375274568753

```

Ilustración 1 Resultados Sadio Mané

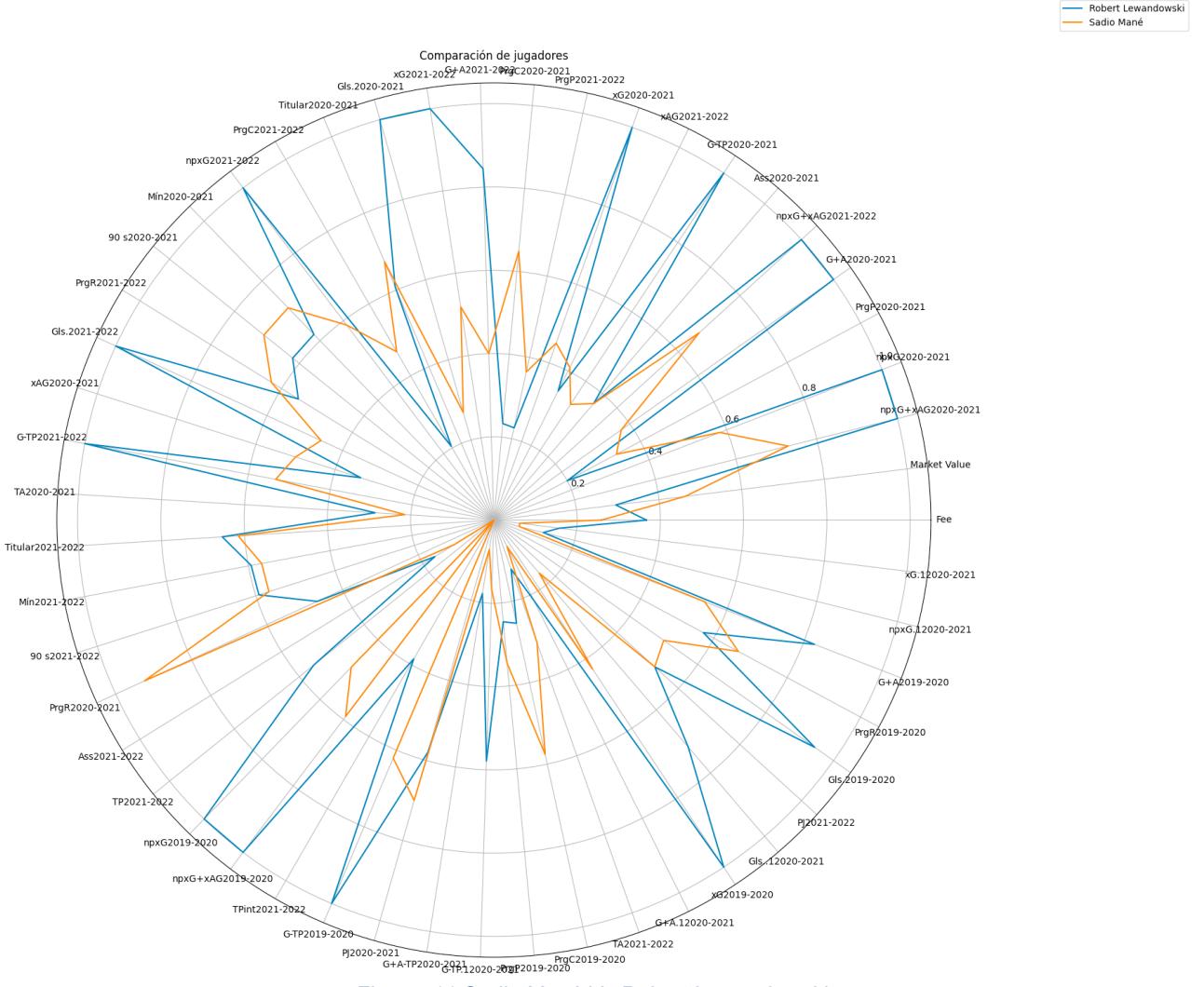


Figura 11 Sadio Mané Vs Robert Lewandowski

4.2. Ejemplo 2

En el ejemplo de la ilustración 2, la red neuronal predice un valor de 11,37 millones de euros para el fichaje de Aurélien Tchouaméni. Cabe destacar que es un jugador muy joven y tiene un gran potencial, es por ello una de las principales razones del elevado precio (80 millones de euros) que realizó el Real Madrid. [20]

En el modelo K-NN, obtenemos a varios jugadores, pero el más parecido a Casemiro es Aurélien Tchouaméni, lo que sugiere que es el sustituto ideal para Casemiro en base a las métricas empleadas. En la figura 12, Tchouaméni ha tenido un rendimiento similar al de Casemiro, por lo que se puede concluir que ha sido un fichaje acertado.

Pruebas y desarrollos

```
knn(nombre_jugador_entrada='Casemiro', k=20)
compare_players('Casemiro', Aurelien, '2022-2023')
calculate_distance('Casemiro', Aurelien)
predict_fee(Aurelien)
plot_radar_chart(['Casemiro', Aurelien])
✓ 0.6s

Los 20 jugadores más similares a Casemiro son:
1 Aurélien Tchouaméni con una distancia de 0.6710953471510185
2 Rúben Neves con una distancia de 0.7664954199082239
3 Joan Jordán con una distancia de 0.8125689034027207
4 Mauro Arambarrí con una distancia de 0.814832940974039
5 Nikola Milenković con una distancia de 0.8245301145004251
6 Douglas Luiz con una distancia de 0.8279019493305713
7 Davide Faraoni con una distancia de 0.8362408510819473
8 Gianluca Mancini con una distancia de 0.8499668776047984
9 John McGinn con una distancia de 0.852073930639622
10 Nicolas Höfler con una distancia de 0.8550071542505894
11 Benjamin André con una distancia de 0.8624619591853564
12 Harry Maguire con una distancia de 0.8722905754551413
13 Bryan Cristante con una distancia de 0.8790931347370246
14 Koke con una distancia de 0.8876840800257331
15 Damián Suárez con una distancia de 0.8914250348573263
16 Baptiste Santamaría con una distancia de 0.893014926152324
17 Maximilian Arnold con una distancia de 0.8990326146980054
18 Abdoulaye Doucouré con una distancia de 0.9021774029967615
19 Marcelo Brozović con una distancia de 0.9032131921714153
20 Morten Thorsby con una distancia de 0.9065178417563565

Aurélien Tchouaméni ha sido mejor que Casemiro en el año 2022-2023.
La distancia entre Casemiro y Aurélien Tchouaméni es 0.6710953471510185
1/1 ━━━━━━ 0s 10ms/step
El valor de mercado predicho para el jugador Aurélien Tchouaméni es: 11.372506141662598 millones de euros
```

Ilustración 2 Aurélien Tchouaméni

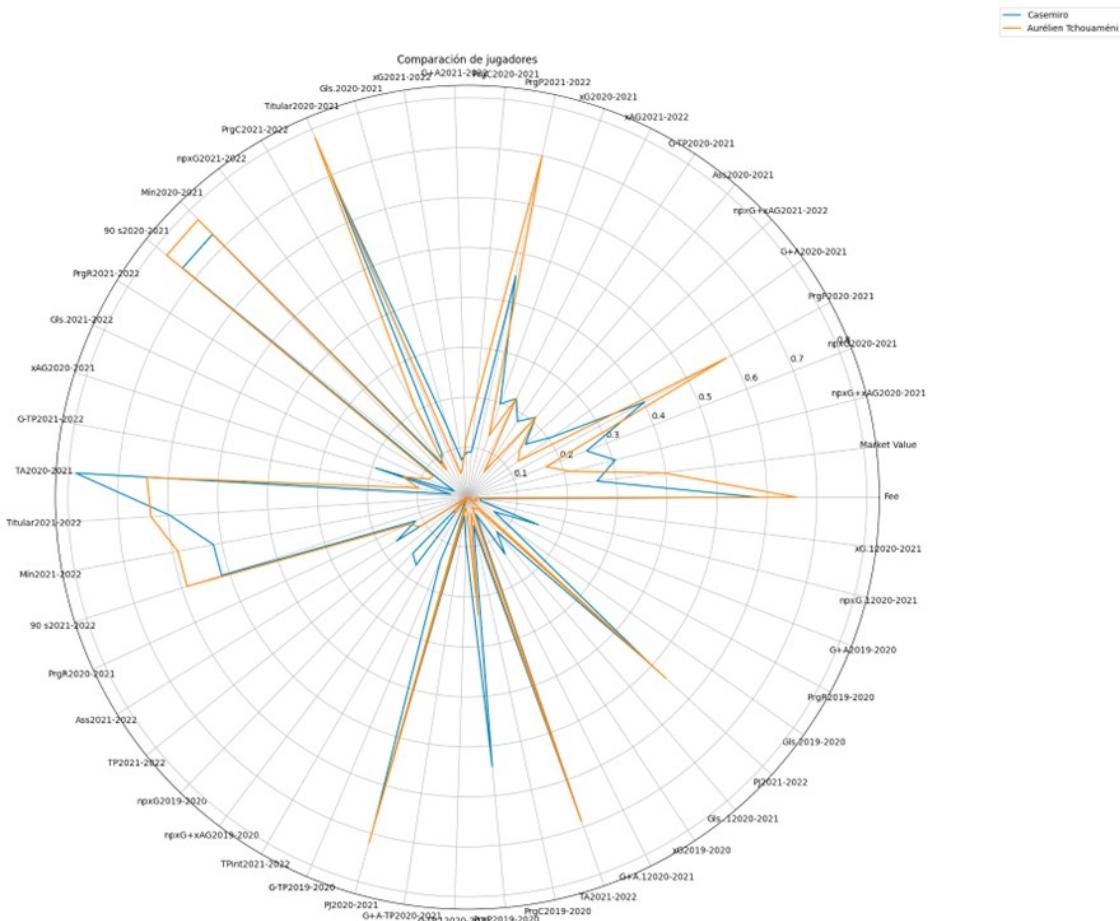


Figura 12 Aurélien Tchouaméni Vs Casemiro

4.3. Ejemplo 3

El Arsenal tiene la necesidad de realizar una inversión en un nuevo delantero para suprir las bajas de Nicolas Pépé, que se marcha cedido y Lacazette, que finaliza su contrato.

La red neuronal [ilustración 3], predice un valor de 42,57 millones de euros para Gabriel Jesús, pero el Arsenal acaba pagando una cifra un poco superior 52,20 millones de euros. Podemos decir que en este caso el valor obtenido se corresponde con la realidad debido a que Gabriel Jesús es un jugador contrastado y no se trata de un jugador joven y con mucho potencial. También indica los jugadores más parecidos a Alexandre Lacazette.

En el uso de k-NN [ilustración 3 y 4], no se obtiene a Gabriel Jesús como potencial relevante de estos jugadores y esto se debe a que los clubs no basan su búsqueda únicamente en este tipo de algoritmos, por lo que no se puede concluir que el modelo tiene un mal funcionamiento, sino que influyen otros factores y criterios que no tiene en cuenta este modelo.

Gabriel Jesús [ilustración 4] ha sido peor en términos estadísticos que Lacazette pero mejor que Nicolas Pépé. Cabe destacar que según la gráfica [figura 13] podemos observar que el rendimiento de los 3 jugadores es muy similar y Gabriel Jesús supera a ambos en varias estadísticas.

```
predict_fee('Gabriel Jesus')
knn(nombre_jugador_entrada='Alexandre Lacazette', k=20)
knn(nombre_jugador_entrada='Nicolas Pépé', k=20)
compare_players('Alexandre Lacazette', 'Gabriel Jesus', '2022-2023')
compare_players('Nicolas Pépé', 'Gabriel Jesus', '2022-2023')

calculate_distance('Alexandre Lacazette', 'Gabriel Jesus')
calculate_distance('Nicolas Pépé', 'Gabriel Jesus')

plot_radar_chart(['Alexandre Lacazette', 'Gabriel Jesus', 'Nicolas Pépé'])
✓ 0.7s

1/1 ━━━━━━━━ 0s 12ms/step
El valor de mercado predicho para el jugador Gabriel Jesus es: 42.576332092285156 millones de euros
Los 20 jugadores más similares a Alexandre Lacazette son:
1 Ante Budimir con una distancia de 0.5407077636906118
2 Fabio Quagliarella con una distancia de 0.5586663818094323
3 Kasper Dolberg con una distancia de 0.6188620672932434
4 Habib Diallo con una distancia de 0.6572042502051972
5 Willian José con una distancia de 0.6691661553802704
6 Kelechi Iheanacho con una distancia de 0.6756392226182535
7 Santi Mina con una distancia de 0.6779667722791656
8 Roger Martí con una distancia de 0.6853123594618858
9 Mama Samba Baldé con una distancia de 0.691507930576618
10 Jorge Molina con una distancia de 0.7058711620295199
11 Zlatan Ibrahimović con una distancia de 0.7235400828677021
12 Borja Mayoral con una distancia de 0.7272602249708054
13 Antonio Sanabria con una distancia de 0.7288119361678052
14 Callum Wilson con una distancia de 0.729013364117835
15 Irvin Cardona con una distancia de 0.7487507668172502
16 Arkadiusz Milik con una distancia de 0.7516755546853983
17 Breel Embolo con una distancia de 0.7648839619236968
18 Youssef En-Nesyri con una distancia de 0.7735866519054745
19 Riccardo Orsolini con una distancia de 0.773766936048764
20 Kevin Lasagna con una distancia de 0.7792563207490616
```

Ilustración 3 Gabriel Jesús

Pruebas y desarrollos

Los 20 jugadores más similares a Nicolas Pépé son:

- 1 Anwar El Ghazi con una distancia de 0.45657740599570346
- 2 Irvin Cardona con una distancia de 0.5962850643888218
- 3 Goran Pandev con una distancia de 0.6097155218671771
- 4 Ayoze Pérez con una distancia de 0.6322147348298214
- 5 Mason Greenwood con una distancia de 0.644257766466905
- 6 Romain Hamouma con una distancia de 0.6474034818985316
- 7 Dodi Lukebakio con una distancia de 0.6492974235372014
- 8 Joaquín Correa con una distancia de 0.6530881923456119
- 9 Yussuf Poulsen con una distancia de 0.6578758400960573
- 10 Ignatius Ganago con una distancia de 0.6592646422766669
- 11 Borja Mayoral con una distancia de 0.6604249684211478
- 12 Marcus Thuram con una distancia de 0.6663915273102234
- 13 Cristian Tello con una distancia de 0.68185475656873
- 14 Christian Pulisic con una distancia de 0.6840835549842996
- 15 Krzysztof Piątek con una distancia de 0.6881933276923252
- 16 Adam Ounas con una distancia de 0.6887993437437835
- 17 Christian Benteke con una distancia de 0.6917611932363587
- 18 Keita Baldé con una distancia de 0.69193058309526
- 19 Joaquín con una distancia de 0.6962171014352447
- 20 Darwin Machís con una distancia de 0.7007670854255279

Gabriel Jesus ha sido peor que Alexandre Lacazette en el año 2022–2023.
 Gabriel Jesus ha sido mejor que Nicolas Pépé en el año 2022–2023.
 La distancia entre Alexandre Lacazette y Gabriel Jesus es 0.9888633233218999
 La distancia entre Nicolas Pépé y Gabriel Jesus es 1.4258739106448404

Ilustración 4 Gabriel Jesús 2

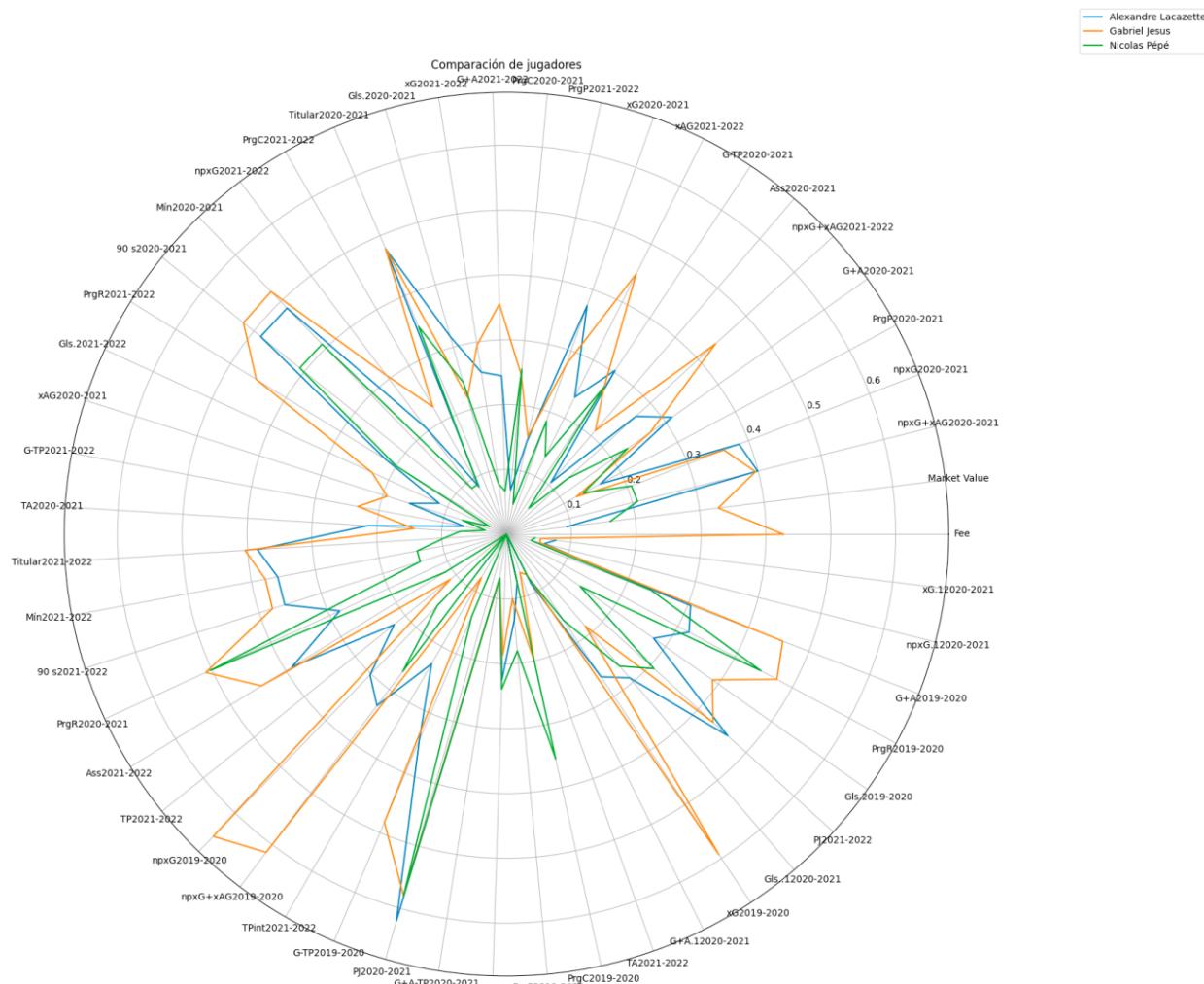


Figura 13 Gabriel Jesús Vs Lacazette Vs Pépé

4.4. Ejemplo 4

El Chelsea se enfrenta a la necesidad de reforzar la defensa tras la salida de Antonio Rüdiger y Andreas Christensen, ambos defensas titulares y dejan el equipo con la carta de libertad, es decir, a coste cero. Para cubrir estas bajas decidieron invertir en Wesley Fofana pagando un total de 80,4 millones de euros, en cambio, la red neuronal predice un valor de 6,09 millones de euros.

Este alto precio se debe a varios motivos, el potencial a largo plazo de Fofana, las condiciones del mercado y ante la necesidad del Chelsea de fichar un defensa central, todo esto ha hecho elevar el precio de este jugador.

En el uso del modelo k-NN [ilustración 5 y 6], no se obtiene a Fofana como relevo de Rüdiger y Christensen, esto significa que los clubs no basan las decisiones de fichajes únicamente en este modelo y también tienen otros aspectos en cuenta como la proyección futura del jugador.

Aunque Fofana ha tenido un rendimiento inferior [ilustración 6] al de ambos jugadores a los que sustituye, no sólo hay que juzgar este fichaje en ese aspecto ya que este equipo ha realizado muy mala temporada y no se puede atribuir la responsabilidad a un único jugador.

```
predict_fee('Wesley Fofana')
knn(nombre_jugador_entrada='Andreas Christensen', k=20)
knn(nombre_jugador_entrada='Antonio Rüdiger', k=20)
knn(nombre_jugador_entrada='Wesley Fofana', k=10)
calculate_distance('Andreas Christensen', 'Wesley Fofana')
calculate_distance('Antonio Rüdiger', 'Wesley Fofana')
compare_players('Andreas Christensen', 'Wesley Fofana', '2022-2023')
compare_players('Antonio Rüdiger', 'Wesley Fofana', '2022-2023')
plot_radar_chart(['Andreas Christensen', 'Antonio Rüdiger', 'Wesley Fofana'])

✓ 0.8s
1/1 ━━━━━━━━ 0s 10ms/step
El valor de mercado predicho para el jugador Wesley Fofana es: 6.09172248840332 millones de euros
Los 20 jugadores más similares a Andreas Christensen son:
1 Jack Cork con una distancia de 0.2848156158942113
2 Davinson Sánchez con una distancia de 0.2975907554193418
3 Nikola Maksimović con una distancia de 0.3069704385223792
4 Luca Ceppitelli con una distancia de 0.3305278103430392
5 Giorgio Chiellini con una distancia de 0.33386812267427945
6 Robin Koch con una distancia de 0.3350707519204015
7 Sean Longstaff con una distancia de 0.3396352852792083
8 Thilo Kehrer con una distancia de 0.3426612826768406
9 Nemanja Radoja con una distancia de 0.35136873846941796
10 Unai García con una distancia de 0.3531425202180284
11 Nathaniel Clyne con una distancia de 0.36364362304318454
12 Martin Agirregabiria con una distancia de 0.3799554695595251
13 Jacob Barrett Laursen con una distancia de 0.3832007450873072
14 Emil Kraft con una distancia de 0.3855228472581744
15 Abdou Diallo con una distancia de 0.38671336499895037
16 Pedro Bigas con una distancia de 0.3909180311641567
17 Arthur Melo con una distancia de 0.3949992322045689
18 Alex Ferrari con una distancia de 0.3966852541281578
19 Davide Biraschi con una distancia de 0.4045102564708064
20 Jack Stephens con una distancia de 0.40770598501347394
```

Ilustración 5 Wesley Fofana

Los 20 jugadores más similares a Antonio Rüdiger son:

- 1 William Saliba con una distancia de 0.47785579781045706
- 2 Aymeric Laporte con una distancia de 0.49339686529215127
- 3 Gabriel Dos Santos con una distancia de 0.5071338815224739
- 4 Diego Llorente con una distancia de 0.5272441459865542
- 5 Lucas Torró con una distancia de 0.5337497878081588
- 6 Ben Davies con una distancia de 0.5398531173676632
- 7 Nico Schlotterbeck con una distancia de 0.5406913749367795
- 8 Wallace con una distancia de 0.5410326910895974
- 9 Marquinhos con una distancia de 0.5565347786690634
- 10 Jean-Clair Todibo con una distancia de 0.5715788751671917
- 11 Oriol Romeu con una distancia de 0.5787684550654203
- 12 Mohammed Salisu con una distancia de 0.5802350899135538
- 13 Jean-Charles Castelletto con una distancia de 0.5849348401561062
- 14 Çağlar Söyüncü con una distancia de 0.5916792715053397
- 15 Romain Saïss con una distancia de 0.5923719105655896
- 16 Fikayo Tomori con una distancia de 0.5986942230391434
- 17 Gerard Piqué con una distancia de 0.602439769254937
- 18 Étienne Capoue con una distancia de 0.6108629019038575
- 19 Axel Disasi con una distancia de 0.612397693078053
- 20 Koray Günter con una distancia de 0.6159250968142368

Los 10 jugadores más similares a Wesley Fofana son:

- 1 Kalvin Phillips con una distancia de 0.6673089477555453
- 2 Ohis Felix Uduokhai con una distancia de 0.6845347073287154
- 3 Andrea Masiello con una distancia de 0.6882279048129022
- 4 Yannick Cahuzac con una distancia de 0.691944890751414
- 5 Loïc Bade con una distancia de 0.7015734029855353
- 6 Jerry St. Juste con una distancia de 0.7115659982051327
- 7 Moreto Cassamá con una distancia de 0.7154724582229999
- 8 Federico Barba con una distancia de 0.7178132223873293
- 9 Iván Marcone con una distancia de 0.7242090965401943
- 10 Unai Núñez con una distancia de 0.7243617581588923

La distancia entre Andreas Christensen y Wesley Fofana es 1.0632347191404323

La distancia entre Antonio Rüdiger y Wesley Fofana es 1.5226781805787843

Wesley Fofana ha sido peor que Andreas Christensen en el año 2022–2023.

Wesley Fofana ha sido peor que Antonio Rüdiger en el año 2022–2023.

Ilustración 6 Wesley Fofana 2

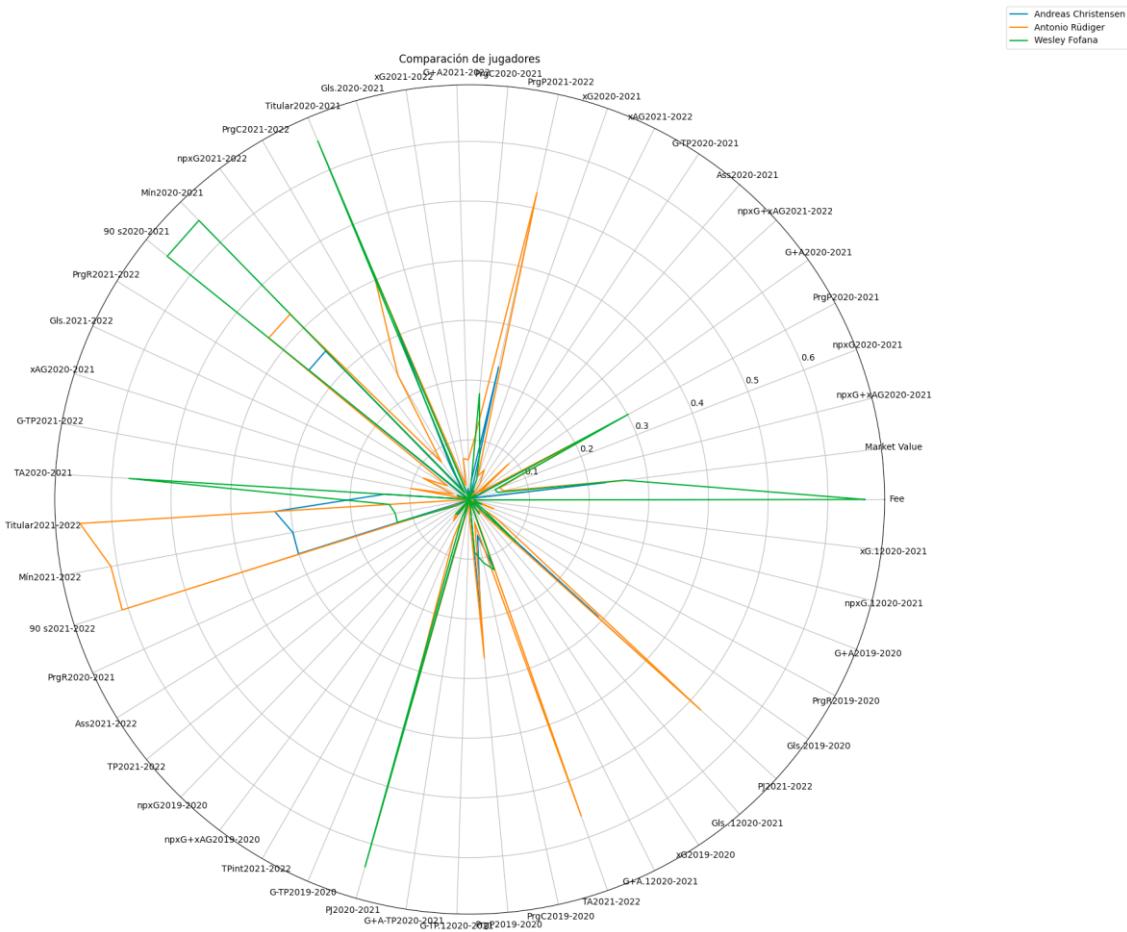


Figura 14 Wesley Fofana Vs Rüdiger Vs Christensen

En resumen, el modelo de k-NN y la red neuronal ofrecen una perspectiva de los posibles fichajes y precios que se pueden pagar por un jugador. Hay que tener en cuenta que estos resultados sirven como apoyo y no como sustitución del trabajo realizado por los empleados y managers de un club a la hora de planificar las nuevas incorporaciones de la siguiente temporada.

Es crucial no juzgar el modelo de k-NN como erróneo debido a que no produce resultados parecidos con las decisiones tomadas por cada club. Esto mismo ocurre con la red neuronal, ya que no todos los casos de uso se asemejan con la realidad, debido a múltiples factores que no se pueden tener en cuenta en este modelo, como puede ser la negociación entre clubs y los intereses de los representantes de los jugadores entre otros factores.

También cabe destacar que no podemos saber el resultado obtenido en el caso en el que los clubs decidieran hacer uso de estos modelos como parte del proceso de selección y negociación de los fichajes de los jugadores.

Conclusión y trabajo futuro

En conclusión, se han alcanzado todos los objetivos propuestos al inicio de la memoria, a continuación, se detalla cómo se ha llevado a cabo cada uno de ellos.

1. **Desarrollar un software efectivo que se encargue de predecir el valor de los fichajes:** Se ha desarrollado un software robusto y funcional que permite realizar predicciones sobre el valor de fichajes de los jugadores de fútbol, está compuesto por dos modelos predictivos, el primero está basado en K-NN para poder obtener jugadores parecidos a uno dado y el segundo, es una red neuronal capaz de realizar la predicción sobre el costo de transferencia de un jugador.
2. **Recopilar datos fiables que incluyen estadísticas de todos los jugadores de las ligas europeas más importantes (Bundesliga, Premier League, Ligue 1, Serie A y La Liga):** Se han obtenido los datos de la página Fbref que colabora con Opta. Este punto ha sido fundamental para implementar un modelo eficiente.
3. **Limpiar y procesar datos recopilados:** Para crear un conjunto de datos óptimo para el entrenamiento de k-NN y red neuronal.
4. **Buscar información sobre el valor de mercado de los jugadores:** investigando en la página de Transfermarkt que ofrece datos oficiales de los traspasos, obteniendo un dataset de los jugadores que han sido traspasados.
5. **Investigar, analizar y seleccionar algoritmos eficaces (Machine Learning, K-Nearest Neighbors (k-NN) y redes neuronales):** mediante el análisis de distintos algoritmos que pueden ser útiles para crear la herramienta, eligiendo K-NN y creando una red neuronal propia.
6. **Construir modelos predictivos con los algoritmos seleccionados:** tras la construcción de estos modelos, se han optimizado ajustando los parámetros de entrada para obtener resultados óptimos y precisos.
7. **Entrenar y validar estos modelos:** realizado con conjuntos de datos de entrenamiento y validación para evaluar el rendimiento de estos modelos.

Este trabajo abre la posibilidad sobre nuevas vías de investigación y mejoras, una de ellas es realizar una red neuronal que obtenga como datos de entrada los datos más relevantes de cada dataframe obtenido en la descarga inicial, con el objetivo de poder tener un modelo mucho más amplio y con más datos de entrada. De tal forma que pueda contemplar muchas más variables relevantes en el valor de mercado de un jugador y sus parecidos.

También abre la posibilidad de realizar una interfaz de usuario, que sea mucho más fácil de usar la herramienta, de esta forma un usuario sin conocimientos en programación puede seleccionar un jugador que quiera sustituir y automáticamente se podrán obtener los jugadores más parecidos. También podrá seleccionar entre estos jugadores para obtener su posible valor de mercado y las gráficas comparativas.

Bibliografía

- [1] D. Garcia. "2022, el año en el que el Mundial de Qatar batió audiencias pese a las polémicas". Palco23 | Noticias económicas del negocio del deporte. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.palco23.com/competiciones/2022-el-ano-en-el-que-el-mundial-de-qatar-batio-audiencias-pese-a-las-polemicas>
- [2] "Ingresos anuales de la FIFA por derechos de televisión | Statista". Statista. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://es.statista.com/estadisticas/504152/fifa-ingresos-por-derechos-television/>
- [3] J. F. DÍAZ. "2023 bate un nuevo récord de gasto total en fichajes internacionales con 8.883 millones de euros". MARCA. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.marca.com/futbol/mercado-fichajes/2024/01/30/65b8f91f22601d2c548b459a.html>
- [4] "La economía del fútbol: mucho más que un deporte". Thinking Heads. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://thinkingheads.com/tendencia-global/economia-del-futbol/>
- [5] "¿Qué es el coeficiente de correlación de Pearson?" QuestionPro. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.questionpro.com/blog/es/coeficiente-de-correlacion-de-peerson/>
- [6] "¿Qué es el algoritmo de k vecinos más cercanos? | IBM". IBM - United States. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.ibm.com/es-es/topics/knn>
- [7] "1.6. Nearest Neighbors". scikit-learn. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://scikit-learn.org/stable/modules/neighbors.html>
- [8] "¿Qué es una red neuronal? - Explicación de las redes neuronales artificiales - AWS". Amazon Web Services, Inc. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://aws.amazon.com/es/what-is/neural-network/>
- [9] "Using Python to Predict Transfer Market Values - Freddy Barnes". Freddy Barnes. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://freddybarnes.com/using-python-to-predict-transfer-market-values/>
- [10] "Football Statistics and History | FBref.com". FBref.com. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://fbref.com/en/>
- [11] "Opta data from Stats Perform". Stats Perform. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.statsperform.com/opta/>
- [12] "Todo sobre FBref.com | FBref.com". FBref.com. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://fbref.com/es/about/#:~:text=Bienvenido,a%20a%20FBref.com,,,-Reference%20y%20Basketball-Reference>.
- [13] "Estadísticas 2017-2018 Las 5 grandes ligas europeas | FBref.com". FBref.com. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://fbref.com/es/comps/Big5/2017-2018/Estadisticas-2017-2018-Las-5-grandes-ligas-europeas>
- [14] "Fichajes de fútbol, rumores, valores de mercado y noticias". Fichajes de fútbol, rumores, valores de mercado y noticias | Transfermarkt. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.transfermarkt.es>

[15]“GitHub - saadism777/Transfermarkt-Data-Analysis: Analyzing and visualizing data of football transfers scraped from <https://www.transfermarkt.com/> of 2022/2023 season.” GitHub. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://github.com/saadism777/Transfermarkt-Data-Analysis>

[16]“¿Cuáles son las ventajas y desventajas de k vecinos más cercanos para la clasificación?” LinkedIn: Log In or Sign Up. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.linkedin.com/advice/1/what-advantages-disadvantages-k-nearest-neighbors?lang=es&originalSubdomain=es>

[17]M. Ali. “Introducción a las funciones de activación en las redes neuronales”. Learn Data Science and AI Online | DataCamp. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.datacamp.com/es/tutorial/introduction-to-activation-functions-in-neural-networks>

[18]“Estimación adaptativa del momento: entender a Adam y utilizarlo correctamente”. Konfuzio. Accedido el 14 de junio de 2024. [En línea]. Disponible: [https://konfuzio.com/es/estimacion-adaptativa-de-momentos/#:~:text=La%20Estimaci%C3%B3n%20Adaptativa%20de%20Momentos%20\(Adam\)%20es%20un%20algoritmo%20de,aprendizaje%20adaptativa%20para%20cada%20par%C3%A1metro](https://konfuzio.com/es/estimacion-adaptativa-de-momentos/#:~:text=La%20Estimaci%C3%B3n%20Adaptativa%20de%20Momentos%20(Adam)%20es%20un%20algoritmo%20de,aprendizaje%20adaptativa%20para%20cada%20par%C3%A1metro).

[19]I. Mentruit. “Por qué Mané se va del Liverpool”. Mundo Deportivo. Accedido el 14 de junio de 2024. [En línea]. Disponible: <https://www.mundodeportivo.com/futbol/fichajes/20220620/1001825198/mane-liverpool.html>

[20]J. Pacheco. “Casemiro vs. Tchouaméni: la comparación definitiva y la carrera por la titularidad con Ancelotti”. El Español. Accedido el 14 de junio de 2024. [En línea]. Disponible: https://www.elspanol.com/elbernabeu/real-madrid/futbol/20220805/casemiro-tchouameni-comparacion-definitiva-carrera-titularidad-ancelotti/693180768_0.html

[21]“Using Python to Predict Transfer Market Values - Freddy Barnes”. Freddy Barnes. Accedido el 17 de junio de 2024. [En línea]. Disponible: <https://freddybarnes.com/using-python-to-predict-transfer-market-values/>

ANEXOS

Anexo A Ficheros descargados

Estadísticas estándar 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno

y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

Tiempo Jugado

PJ -- Partidos jugados

Partidos jugados por el jugador o el equipo

Titular -- Partido o partidos iniciados por el jugador

Mín -- Minutos

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Rendimiento

Gls. -- Goles

Goles marcados o permitidos

Ass -- Asistencias

Asistencias

G+A -- Goles + Asistencias

Goles y asistencias

G-TP -- Goles sin penalización

Goles sin penalización

TP -- Tiros penales ejecutados

Tiros penales ejecutados

TPint -- Tiros penales intentados

Tiros penales intentados

TA -- Tarjetas amarillas

Tarjetas amarillas

TR -- Tarjetas rojas

Tarjetas rojas

Expectativa

xG -- xG: Goles esperados

Goles esperados

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

npxG -- npxG: Goles esperados (xG) sin contar penaltis

Goles esperados sin penalidad

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

xAG -- xAG: Exp. Assisted Goals

Goles esperados con asistencia

Es un xG que acompaña un pase de asistencia a un tiro

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

npxG+xAG -- npxG + xAG

Goles esperados no penalizados más goles asistidos

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Progresión

PrgC -- Acarreos progresivos

Acarreos que mueven el balón hacia la línea de gol del oponente por lo menos 10 yardas desde

su punto más alejado en los últimos seis pases, o cualquier acarreo al área penal. Excluye acarreos que terminan en la defensa del 50% del campo

PrgP -- Pases progresivos

Pases progresivos

Pases completos que mueven el balón hacia la línea de gol del oponente al menos 10 yardas desde su punto más alejado en los últimos seis pases, o cualquier pase completo al área penal. Excluye pases del 40% de la defensa del campo

PrgR -- Pases progresivos Rec

Pases progresivos recibidos

Pases completos que mueven el balón hacia la línea de gol del oponente al menos 10 yardas desde su punto más alejado en los últimos seis pases, o cualquier pase completo al área penal. Excluye pases del 40% de la defensa del campo

Por 90 Minutos

Gls. -- Goles/90

Goles marcados por 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Ast -- Asistencias/90

Asistencias por 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

G+A -- Goles + asistencias/90

Goles y asistencias por 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

G-TP -- Goles sin penalización/90

Goles menos penales ejecutados por cada 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

G+A-TP -- Goles sin penalización + asistencias/90

Goles más asistencias menos tiros de penales ejecutados por 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

xG -- xG/90

Goles esperados por 90 minutos

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

xAG -- xAG/90

Goles esperados con asistencia por 90 minutos

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

xG+xAG -- xG + xAG/90

Goles esperados más goles asistidos por 90 minutos

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

npxG -- npxG/90

Goles esperados sin penalidad por cada 90 minutos

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

npxG+xAG -- npxG + xAG

Goles esperados no penalizados más goles asistidos por 90 minutos

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Portero 2017-2018

RL – Clasificación

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno

y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

Tiempo Jugado

PJ -- Partidos jugados

Partidos jugados por el jugador o el equipo

Titular -- Partido o partidos iniciados por el jugador

Mín -- Minutos

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Rendimiento

GC -- Goles en contra

Goles en contra

GC90 -- Goles en contra/90

Goles en contra por 90 minutos

DaPC -- Disparos a puerta en contra

Disparos a puerta en contra

% Salvadas -- Porcentaje de salvadas

Porcentaje de salvadas

(Disparos a puerta en contra - Goles en contra)/Disparos a puerta en contra

Ten en cuenta que no todos los disparos a puerta son tapados por el portero, algunos pueden ser tapados por los defensores

No incluye los tiros penales

PG -- Partidos Ganados

Partidos Ganados

PE -- Partidos Empatados

Partidos Empatados

PP -- Partidos Perdidos

Partidos Perdidos

PaC -- Portería a cero

Portería a cero

Partidos completos del portero en los que no se permitieron goles.

PaC% -- Porcentaje de portería a cero

Porcentaje de portería a cero

Porcentaje de partidos que dan como resultado una portería a cero.

Tiros penales

TPint -- Tiros penales intentados

Tiros penales intentados

PD -- Penales concedidos

Penales concedidos

PD -- Penales detenidos

Penales detenidos

PC -- Tiros penales fallidos

Tiros penales fallidos

% Salvadas -- % Salvadas (Tiros penales)

Porcentaje de penales salvados

Goles en contra/Intentos de penales

No se incluyen los tiros de penales que no aciertan

Portería avanzada del jugador 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno

y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Goles

GC -- Goles en contra

Goles en contra

PD -- Penales concedidos

Penales concedidos

TL -- Goles de tiro libre en contra

Goles de tiro libre en contra

TE -- Goles de tiro de esquina en contra

Goles de tiro de esquina en contra

GC -- Goles propios marcados contra el portero

Goles propios marcados contra el portero

Expectativa**PSxG -- Goles esperados posteriores al tiro**

Goles esperados posteriores al tiro

PSxG son los goles esperados basados en la probabilidad que tiene el portero de salvar el tiro.

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

PSxG/SoT -- Goles esperados posteriores al tiro por tiros en el arco

No incluye los tiros penales

PSxG son los goles esperados basados en la probabilidad que tiene el portero de salvar el tiro.

Los números más altos indican que los disparos son más difíciles de detener y es más probable que anoten.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

PSxG+/- -- PSxG-GA

Goles esperados posteriores al tiro menos goles permitidos

Los números positivos sugieren una mejor suerte o una capacidad superior al promedio para detener los disparos

PSxG son los goles esperados basados en la probabilidad que tiene el portero de salvar el tiro.

Nota: No incluye los goles propios.

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

/90 -- PSxG-GA/90

Goles esperados posteriores al tiro menos goles permitidos por 90 minutos

Los números positivos sugieren una mejor suerte o una capacidad superior al promedio para detener los disparos

PSxG son los goles esperados basados en la probabilidad que tiene el portero de salvar el tiro.

Nota: No incluye los goles propios.

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Iniciado**Cmp -- Pases completados (Iniciado)****Pases completados**

Pases de más de 40 yardas

Int. -- Pases intentados (Iniciado)**Pases intentados**

Pases de más de 40 yardas

% Cmp -- Porcentaje de cumplimiento de pases (Iniciado)**Porcentaje de cumplimiento de pases**

Pases de más de 40 yardas

Pases

Att (GK) -- Pases Intentados (GK)

Pases Intentados (GK)

Sin incluir los saques de meta

TI -- Tiros intentados

Tiros intentados

%deLanzamientos -- Lanzamiento %

Porcentaje de pases que fueron realizados

Sin incluir los saques de meta

Pases de más de 40 yardas

Long. prom. -- Promedio de longitud del pase

Longitud media de los pases, en yardas

Sin incluir los saques de meta

Saques de meta

Int. -- Saques de meta

Tiros a puerta intentados

%deLanzamientos -- %deLanzamientos (Saques de meta)

Porcentaje de saques de meta que fueron realizados

Pases de más de 40 yardas

Long. prom. -- Prom. de longitud de los saques de meta

Longitud media de los tiros al arco, en yardas

Pases cruzados

Opp -- Cruces superados

Intentos del oponente de cruzar al área de penal

Stp -- Cruces detenidos

Número de centros en el área penal que fueron frenados con éxito por el arquero

% de Stp -- % de cruces detenidos

Porcentaje de centros en el área de penal que fueron frenados con éxito por el arquero

Barredora

Núm. de OPA -- Acciones defensivas fuera del área penal

Núm. de acciones defensivas fuera del área de penal

Núm. de OPA/90 -- Acciones defensivas fuera del área de penal por 90 minutos

DistProm. -- Distancia promedio de las acciones defensivas

Distancia promedio desde la meta (en yardas) de todas las acciones defensivas

Disparos de jugadores 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas
 DL - Delanteros
 LT - Defensores laterales
 LI - Lateral izquierdo
 LD - Lateral derecho
 DC - Defensa central
 MCD - Mediocampistas defensivos
 MC - Mediocampistas centrales
 CI - Centrocampistas izquierdos
 CD - Centrocampistas derechos
 INT - Volantes
 VI - Volantes izquierdos
 VD - Volantes derechos
 MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno
y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento**90 s -- 90 Jugados****90s jugados**

Minutos jugados divididos por 90

Estándar**Gls. -- Goles**

Goles marcados o permitidos

Dis -- Total de disparos

Total de disparos

No incluye los tiros penales

DaP -- Lanzamientos en el Objetivo

Lanzamientos en el Objetivo

Nota: Los tiros al blanco no incluyen los penales.

% de TT -- Disparos en el Objetivo %

Porcentaje de tiros a puerta

Se requiere un mínimo de 0,395 tiros por juego de equipo para calificar como líder.

Nota: Los tiros al blanco no incluyen los penales.

T/90 -- Total de tiros/90

Tiros al arco cada 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

TalArc/90 -- Tiros al blanco/90

Disparos a puerta por 90 minutos

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Nota: Los tiros al blanco no incluyen los penales.

G/T -- Goles/tiros

Goles por tiro

Se requiere un mínimo de 0,395 tiros por juego de equipo para calificar como líder.

G/TalArc -- Goles/Disparo en el Objetivo

Goles por tiros al arco

Se requiere un mínimo de 0,111 tiros al blanco por partido de equipo para calificar como líder.

Nota: Los tiros al blanco no incluyen los penales.

Dist -- Distancia media de disparo

Distancia promedio, en yardas, del total de disparos realizados hacia la portería

Se requiere un mínimo de 0,395 tiros por juego de equipo para calificar como líder.

No incluye los tiros penales

FK -- Lanzamientos desde Tiros Libres

Lanzamientos desde Tiros Libres

TP -- Tiros penales ejecutados

Tiros penales ejecutados

TPint -- Tiros penales intentados

Tiros penales intentados

Expectativa

xG -- xG: Goles esperados

Goles esperados

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

npxG -- npxG: Goles esperados (xG) sin contar penaltis

Goles esperados sin penalidad

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

npxG/Sh -- npxG/Shot

Goles esperados sin penalidad por cada tiro

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 0,395 tiros por juego de equipo para calificar como líder.

G-xG -- Goles - xG

Goles menos los goles esperados

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

np:G-xG -- Goles sin penalización - npxG

Goles que no son de penalidad menos goles esperados que no son de penalidad

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Pases de jugadores 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno

y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Total

Cmp -- Pases completados

Pases completados

Int. -- Pases intentados

Pases intentados

% Cmp -- % de pase completo

Porcentaje de cumplimiento de pases

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Dist. tot. -- Distancia total de pase

La distancia total, en yardas, que han recorrido los pases completados en cualquier dirección

Dist. prg. -- Distancia de paso progresiva

Distancia progresiva

La distancia total, en yardas, que han recorrido los pases completados hacia la meta del oponente. Nota: Los pases alejados de la meta del oponente se cuentan como cero yardas progresivas.

Cortos**Cmp -- Pases completados (Cortos)**

Pases completados

Pases entre 5 y 15 yardas

Int. -- Pases intentados (Cortos)

Pases intentados

Pases entre 5 y 15 yardas

% Cmp -- % de pase completo (Cortos)

Porcentaje de cumplimiento de pases

Pases entre 5 y 15 yardas

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Medios**Cmp -- Pases completados (Medios)**

Pases completados

Pases entre 15 y 30 yardas

Int. -- Pases intentados (Medios)

Pases intentados

Pases entre 15 y 30 yardas

% Cmp -- % de pase completo (Medios)

Porcentaje de cumplimiento de pases

Pases entre 15 y 30 yardas

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Largos**Cmp -- Pases completados (Largos)**

Pases completados

Pases de más de 30 yardas

Int. -- Pases intentados (Largos)

Pases intentados

Pases de más de 30 yardas

% Cmp -- % de pase completo (Largos)

Porcentaje de cumplimiento de pases

Pases de más de 30 yardas

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Ass -- Asistencias

Asistencias

xAG -- xAG: Exp. Assisted Goals**Goles esperados con asistencia**

Es un xG que acompaña un pase de asistencia a un tiro

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Expectativa**xA -- xA: Asistencias Esperadas**

Asistencia esperada

La probabilidad de que cada pase completado se convierta en una asistencia de gol, teniendo en cuenta el tipo de pase, la fase del juego, la ubicación y la distancia.

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

A-xAG -- Asistencias - xAG**Asistencias menos goles esperados con asistencia**

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

PC -- Pases clave

Pases que conducen directamente a un tiro (tiro asistido).

1/3 -- Pases en el último tercio de la cancha

Pases completados que entran en el último tercio del área más cercana a la portería.

No incluye jugada a balón parado

PPA -- Pases al área de penalización

Pases completados en el área de 18 yardas

No incluye jugada a balón parado

CrAP -- Cruce en el área de penalización

Cruces completados en la zona de las 18 yardas

No incluye jugada a balón parado

PrgP -- Pases progresivos

Pases progresivos

Pases completos que mueven el balón hacia la línea de gol del oponente al menos 10 yardas desde su punto más alejado en los últimos seis pases, o cualquier pase completo al área penal. Excluye pases del 40% de la defensa del campo

Tipos de pases de jugadores 2017-2018**RL -- Clasificación**

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales
CI - Centrocampistas izquierdos
CD - Centrocampistas derechos
INT - Volantes
VI - Volantes izquierdos
VD - Volantes derechos
MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno
y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Int. -- Pases intentados

Pases intentados

Tipos de pases

Balón vivo -- Pases de Balón Vivo

Pases de Balón Vivo

Balón muerto -- Pases de Balón Muerto

Pases de Balón Muerto

Incluye tiros libres, saques de esquina, saques de salida, saques de banda y saques de meta.

FK -- Pases de tiros libres

Pases intentados en tiros libres

PL -- Pases largos

Pase completo enviado entre los defensas hacia el espacio abierto.

Camb. -- Cambios

Pasos que recorren más de 40 yardas del ancho del campo.

Pcz -- Pases cruzados

Pases cruzados

Lanz. -- Lanzamientos Tomados

Lanzamientos Tomados

SE -- Saques de esquina

Saques de esquina

Saques de esquina

Dentro -- Saques de esquina hacia adentro

Saques de esquina hacia adentro

Fuera -- Saques de esquina hacia afuera

Saques de esquina hacia afuera

Rect. -- Saques de esquina rectos

Saques de esquina rectos

Resultados

Cmp -- Pases completados

Pases completados

PA -- Pases fuera de juego

Posición adelantada

Bloqueos -- Pases bloqueados

Blocked by the opponent who was standing in the path

Creación de goles y tiros del jugador 2017-2018**RL -- Clasificación**

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno

y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento**90 s -- 90 Jugados****90s jugados**

Minutos jugados divididos por 90

ACT**ACT -- Acciones para la creación de tiros**

Acciones para la creación de tiros

Las dos acciones ofensivas que conducen directamente a un disparo, como pases, tomas y

faltas de dibujo. Nota: un solo jugador puede recibir crédito por múltiples acciones y el lanzador también puede recibir crédito.

SCA90 -- Pases que acaban en un tiro/90

Acciones de creación de tiro por 90 minutos.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Tipos de SCA

PassLive -- ACT (Pase de Balón Vivo)

Pases completos de balón vivo que conducen a un intento de tiro.

PassDead -- ACT (Pase de Balón Muerto)

Pases completos de balón muerto que conducen a un intento de tiro.

Incluye tiros libres, saques de esquina, saques de salida, saques de banda y saques de meta.

HASTA -- ACT (Toma)

Tomas exitosas que conducen a un intento de disparo

Dis -- ACT (Tiro)

Tiros que conducen a otro intento de tiro.

FR -- ACT (Faltas recibidas)

Las faltas cometidas que provocan un intento de tiro.

Def -- ACT (Acción Defensiva)

Acciones defensivas que producen un intento de tiro

ACG

ACG -- Acciones para la creación de goles

Acciones para la creación de goles

Las dos acciones ofensivas que conducen directamente a un gol, como pases, tomas y cometer faltas. Nota: un solo jugador puede recibir crédito por múltiples acciones y el lanzador también puede recibir crédito.

GCA90 -- Pases que acaban en un gol/90

Acciones de creación de goles por 90 minutos.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Tipos de GCA

PassLive -- ACG (Pase de Balón Vivo)

Pases de balón vivo completados que conducen a un gol.

PassDead -- ACG (Pase de Balón Muerto)

Pases completos a balón parado que conducen a un gol. Incluye tiros libres, saques de esquina, saques de salida, saques de banda y saques de meta.

HASTA -- ACG (Toma)

Adquisiciones exitosas que conducen a una meta

Dis -- ACG (Tiro)

Tiros que conducen a otro tiro de gol.

FR -- ACG (Faltas recibidas)

Las faltas cometidas que conducen a un gol.

Def -- ACG (Acción Defensiva)

Acciones defensivas que conducen a un gol

Acciones defensivas del jugador 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno
y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Derribos

Tkl -- Derribos

Número de jugadores tackleados

TkIG -- Derribos conseguidos

Tackleos que le dieron la posesión del balón al equipo que lo realizó.

3.º def. -- Derribos (3.º def.)

Entradas en la defensa 1/3

3.º cent. -- Derribos (3.º cent.)

Entradas en el centro 1/3

3.º ataq. -- Derribos (3.º ataq.)

Entradas en el ataque 1/3

Desafíos

Tkl -- Regateadores tackleados

Número de dribladores tackleados

Att -- Regateos desafiados

Número de desafíos fallidos más número de regateadores abordados

Tkl% -- % de Dribladores Derribados**Porcentaje de regateadores tacleados**

Regateadores tacleados dividido por el número de intentos de desafiar a un regateador contrario

Mínimo .625 regateadores desafiados por juego de equipo para calificar como líder

Pérdida -- Desafíos perdidos

Número de intentos infructuosos de desafiar a un jugador que regatea

Bloqueos

Bloqueos -- El número de veces que se bloquea el balón poniéndose en su camino.

Dis -- Disparos bloqueados

El número de veces que se bloquea un tiro al pararse en su camino.

Pases -- Pases bloqueados

El número de veces que se bloquea un pase por pararse en su camino.

Int -- Intercepciones

Intercepciones

Tkl+Int -- Número de jugadores tackleados más número de intercepciones

Desp. -- Despeje

Despeje

Err -- Errores

Errores que provocan el tiro de un oponente.

Posesión del balón 2017-2018**RL -- Clasificación**

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno
y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento**90 s -- 90 Jugados****90s jugados**

Minutos jugados divididos por 90

Toques

Toques -- El número de veces que un jugador tocó el balón. Nota: Recibir un pase, luego driblar, y luego enviar un pase cuenta como un toque.

Def. pen. -- Toques (Def. pen.)

Toques en el área defensiva del penal

3.º def. -- Toques (3.º def.)

Toques en la defensiva 1/3

3.º cent. -- Toques (3.º cent.)

Toques en el centro 1/3

3.º ataq. -- Toques (3.º ataq.)

Toques en el ataque 1/3

Ataq. pen. -- Toques (Ataq. pen.)

Toques en el área penal de ataque

Balón vivo -- Toques (Pelota activa)

Toques de balón vivo. No incluye saques de esquina, tiros libres, saques de banda, saques de salida, saques de meta o penales.

Tomas**Att -- Tomas intentadas**

Número de intentos de enfrentarse a los defensores mientras regatean

Succ -- Tomas exitosas

Número de defensas enfrentados con éxito, regateándolos

Las tomas fallidas incluyen intentos en los que el regateador retiene la posesión pero no puede pasar al defensor

Exitosa% -- % de toma exitosa

Porcentaje de tomas completadas con éxito

Las tomas fallidas incluyen intentos en los que el regateador retiene la posesión pero no puede pasar al defensor

Mínimo .5 tomas por juego de escuadrón para calificar como líder

Tkld -- Tiempos abordados durante la toma

Número de veces que un defensor ha hecho tacleadas durante un intento de tacleada

Tkld% -- Abordado durante el porcentaje de aceptación

Porcentaje de tiempo tacleado por un defensor durante un intento de enfrentamiento

Mínimo .5 tomas por juego de escuadrón para calificar como líder

Transportes

Transportes -- El número de veces que el jugador controló el balón con sus pies.

Dist. tot. -- Distancia total de trasladados

Distancia total, en yardas, que un jugador movió el balón mientras lo controlaba con sus pies, en

cualquier dirección.

Dist. prg. -- Distancia de traslado progresivo

Distancia progresiva

Distancia total, en yardas, que un jugador movió el balón mientras lo controlaba con sus pies hacia la meta del oponente.

PrgC -- Acarreos progresivos

Acarreos que mueven el balón hacia la línea de gol del oponente por lo menos 10 yardas desde su punto más alejado en los últimos seis pases, o cualquier acarreo al área penal. Excluye acarreos que terminan en la defensa del 50% del campo

1/3 -- Traslados en el último tercio

Traslados que entran en el 1/3 del campo más cercano a la portería

TAP -- Traslados al área de penal

Traslados al área de 18 yardas

Errores de control -- El número de veces que un jugador falló al intentar ganar el control de un balón.

Des -- Despojado

Número de veces que un jugador pierde el control del balón tras ser placado por un jugador contrario. No incluye intentos de tomas

Recepción

Rec -- Pases recibidos

Número de veces que un jugador ha recibido un pase con éxito.

PrgR -- Pases progresivos Rec

Pases progresivos recibidos

Pases completos que mueven el balón hacia la línea de gol del oponente al menos 10 yardas desde su punto más alejado en los últimos seis pases, o cualquier pase completo al área penal. Excluye pases del 40% de la defensa del campo

Tiempo de juego del jugador 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos
CD - Centrocampistas derechos
INT - Volantes
VI - Volantes izquierdos
VD - Volantes derechos
MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno
y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

Tiempo Jugado

PJ -- Partidos jugados

Partidos jugados por el jugador o el equipo

Mín -- Minutos

Mn/PJ -- Minutos por partido jugado

Minutos por partido jugado

% min -- Porcentaje de minutos jugados por la plantilla

Porcentaje de minutos jugados

Porcentaje de minutos totales del equipo en los que el jugador estuvo en el campo

Minutos jugados divididos por el total de minutos jugados por el equipo.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Titular

Titular -- Partido o partidos iniciados por el jugador

Mn/arranque -- Minutos por partido iniciado

Minutos por partido iniciado

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Compl -- Partidos Completos Jugados

Partidos completos jugados

Sup.

Sup -- Presentaciones Sustitutas

Partidos como suplente

Partido o partidos que el jugador no jugó, también como suplente

Mn/Sust -- Minutos por sustitución

Minutos por sustitución

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Partidos como sustituto -- Juegos como Sub no Utilizado

Partidos sin jugar como sustituto

Logros del equipo

PPP -- Puntos por partido

Puntos por partido

Promedio de puntos obtenidos por el equipo en los partidos en los que participó el jugador.
Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

onG -- Goles marcados (en el campo)

Goles marcados por el equipo en el campo

onGA -- Goles permitidos (en el campo)

Goles permitidos por el equipo en el campo

+/- -- Más/Menos

Más/Menos

Los goles marcados menos los goles permitidos por el equipo mientras el jugador estaba sobre el terreno de juego.

+/-90 -- Más/Menos/90

Más/Menos por 90 minutos

Los goles marcados menos los goles permitidos por el equipo mientras el jugador estaba en el terreno de juego por cada 90 minutos jugados.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Dentro-Fuera -- Más/Menos neto por 90 minutos

Goles netos por 90 minutos del equipo mientras el jugador estaba en el terreno de juego menos goles netos permitidos por 90 minutos por el equipo mientras el jugador estaba fuera del terreno de juego.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Logros del equipo (xG)

onxG -- Goles esperados (xG) (en el campo)

Goles esperados por equipo en el terreno de juego

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

onxGA -- Goles anotados (xGA) (en el campo)

Goles esperados permitidos por el equipo mientras está en el terreno de juego

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

xG+/- -- Más/Menos goles esperados (xG)

Más/Menos de xG

Los goles esperados anotados menos los goles esperados permitidos por el equipo mientras el jugador estaba sobre el terreno de juego.

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

xG+/-90 -- Más/Menos/90 goles esperados (xG)

Más/Menos de xG por 90 minutos

Los goles esperados anotados menos los goles esperados permitidos por el equipo mientras el jugador estaba sobre el terreno de juego por cada 90 minutos jugados.

Los totales de xG incluyen los tiros desde el punto penal, pero no incluyen las tandas de penales (a menos que se indique lo contrario).

Ofrecido por Opta.

Un subrayado indica que hay un partido que no contiene datos, pero que se actualizará cuando esté disponible.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Dentro-Fuera -- Intermitencia goles esperados (xG)

Más/Menos de xG neto por 90 minutos

Los goles netos esperados por cada 90 minutos del equipo mientras el jugador estaba en el terreno de juego menos los goles netos esperados por cada 90 minutos del equipo mientras el jugador estaba fuera del terreno de juego.

Se requiere un mínimo de 30 minutos jugados por partido de equipo para calificar como líder.

Estadísticas diversas 2017-2018

RL -- Clasificación

Esto es un conteo de las filas de arriba a abajo.

Se recalcula siguiendo el orden de una columna.

País -- Nacionalidad del jugador.

Primero, comprobamos nuestros registros en competencias internacionales de mayor nivel.

Luego en las juveniles.

Luego la ciudadanía que aparece en Wikipedia.

Por último, usamos su lugar de nacimiento cuando está disponible.

Posc -- Posición

Posición que juega más comúnmente el jugador

PO - Porteros

DF - Defensores

CC - Centrocampistas

DL - Delanteros

LT - Defensores laterales

LI - Lateral izquierdo

LD - Lateral derecho

DC - Defensa central

MCD - Mediocampistas defensivos

MC - Mediocampistas centrales

CI - Centrocampistas izquierdos

CD - Centrocampistas derechos

INT - Volantes

VI - Volantes izquierdos

VD - Volantes derechos

MCO - Centrocampistas de ataque

Comp -- Competencia

Competencia

El número junto a la competencia indica qué nivel ocupa esta liga en la pirámide de la liga del país.

Edad -- Edad al inicio de la temporada

Dado el 1 de agosto para las ligas de invierno

y el 1 de febrero para las ligas de verano.

Nacimiento -- Año de nacimiento

90 s -- 90 Jugados

90s jugados

Minutos jugados divididos por 90

Rendimiento**TA -- Tarjetas amarillas**

Tarjetas amarillas

TR -- Tarjetas rojas

Tarjetas rojas

2a amarilla -- Segunda tarjeta amarilla

Segunda tarjeta amarilla

FIs -- Faltas cometidas

Faltas cometidas

FR -- Faltas recibidas

Faltas recibidas

PA -- Posición adelantada

Posición adelantada

Pcz -- Pases cruzados

Pases cruzados

Int -- Intercepciones

Intercepciones

TkIG -- Derribos conseguidos

Tackleos que le dieron la posesión del balón al equipo que lo realizó.

Penal ejecutado -- Penales ejecutados

Penales ejecutados

Penal concedido -- Penales concedidos

Penales concedidos

GC -- Goles en contra

Goles en contra

Recup. -- Recuperación de pelotas

Número de balones sueltos recuperados

Duelos aéreos**Ganados -- Aéreos Ganados**

Aéreos Ganados

Perdidos -- Aéreos Perdidos

Aéreos Perdidos

% de ganados -- % of Aerials Won

Porcentaje de aéreos ganados

Se requiere un mínimo de 0,97 duelos aéreos por partido de equipo para calificar como líder

Anexo B Tfg comparados

<https://repositorio.uam.es/handle/10486/698384>

<https://repositorio.uam.es/handle/10486/679999>