



Universidad de Valladolid

Escuela de Ingeniería Informática

MEMORIA DEL PROYECTO DE I+D+i A REALIZAR EN GIR/EMPRESA

Máster en Ingeniería Informática
Modalidad No Presencial

Comparación y evaluación de diferentes técnicas de IA para un modelo de rotación de empleados en empresas aplicado a equipos de fútbol

Alumno:
José María Lozano Olmedo
Tutor:
Joaquín Adiego Rodríguez

Resumen

Este proyecto se centra en evaluar diversas técnicas de inteligencia artificial para predecir el rendimiento de equipos de fútbol mediante el análisis de la rotación de jugadores. El principal objetivo es detectar cómo las diferentes estrategias de rotación aplicadas por los equipos afectan al desempeño del equipo y cómo la inteligencia artificial puede realizar predicciones en base a ellas para ayudar a aplicar las mejores estrategias. Este estudio abarca desde la recopilación y el análisis de datos asociados a las ligas seleccionadas, la creación de modelos de inteligencia artificial y la evaluación de su eficacia. Este documento presenta un plan que incluye la metodología y la planificación del proyecto. Se espera que este proyecto proporcione resultados significativos para optimizar la gestión de equipos de fútbol y así poder facilitar el trabajo a sus dirigentes.

Palabras clave: Fútbol, jugadores, inteligencia artificial, predicciones, rotación, rendimiento

Índice general

Resumen	I
Lista de figuras	V
Lista de tablas	VII
1. Introducción	1
1.1. Contexto	1
1.2. Objetivos de desarrollo	2
1.3. Motivación	2
1.4. Aplicaciones similares	3
2. Metodo	5
2.1. Planificación	5
2.2. Análisis de riesgos	6
2.3. Alcance	12
2.4. Metodología	13
2.5. Obtención de los datos	14
2.6. Tecnologías utilizadas	14
3. Conclusiones	17
3.1. Conclusiones	17

3.2. Líneas de trabajo futuras	18
Bibliografía	19

Lista de Figuras

Lista de Tablas

2.1. R-01. Problemas al obtener los datos. 7

2.2. R-02. Poco tiempo para trabajar en el proyecto. 8

2.3. R-03. Fallo en el equipo de trabajo donde se realiza el proyecto. 9

2.4. R-04. Problemas con la conexión a internet. 10

2.5. R-05. Enfermedad del alumno. 11

2.6. Matriz de Riesgos. 12

2.7. Situación de los riesgos del proyecto en la matriz de riesgos. 12

Capítulo 1

Introducción

1.1. Contexto

Este proyecto se desarrollará durante la estancia del proyecto a realizar para la asignatura de I+D+I en un GIR para el Máster en Ingeniería Informática no presencial en la Universidad de Valladolid.

En los últimos años, el campo de la inteligencia artificial ha sufrido un crecimiento considerable, alterando diferentes aspectos de la sociedad moderna [1]. En este contexto, el deporte, y en especial el fútbol, no se ha mantenido al margen. La capacidad de la inteligencia artificial para analizar grandes volúmenes de datos y extraer patrones útiles ha encontrado una aplicación cada vez más importante en el ámbito deportivo, proporcionando nuevas herramientas para optimizar el rendimiento de los equipos y la toma de decisiones por parte de los directivos y entrenadores [2].

El fútbol, es algo más que un simple juego, se ha convertido en un fenómeno global que supera todas las fronteras. Los clubes de fútbol son empresas con mucho dinero y todo lo relacionado con este deporte, de manera general, mueve grandes cantidades de dinero. En este contexto altamente competitivo, la presión por obtener resultados positivos es máxima, tanto en términos deportivos como financieros [3].

En este escenario, la gestión eficiente de los recursos humanos, como en este caso los jugadores, se ha vuelto prioritaria para el éxito de un equipo. Los entrenadores y directivos se enfrentan al desafío de optimizar el rendimiento de sus jugadores tratando de minimizar el riesgo de lesiones y el cansancio físico. La inteligencia artificial ofrece herramientas potentes para abordar este desafío, permitiendo el análisis de datos relacionados con el estado físico de los jugadores, el rendimiento en partidos anteriores, las lesiones previas y otros factores relevantes.

La gestión de la rotación de jugadores es uno de los aspectos más críticos de la estrategia de un equipo a lo largo de una temporada y que tiene una mayor repercusión sobre su éxito.

La inteligencia artificial puede ayudar a los entrenadores a tomar decisiones documentadas sobre cuándo dar descanso a un jugador, cuándo alinear a un futbolista, que cambios realizar y cómo mantener un equilibrio entre la competitividad y la salud de la plantilla. Por lo tanto, el uso de la inteligencia artificial en el fútbol no solo es una oportunidad para mejorar el rendimiento deportivo, sino también una necesidad para mejorar sobre los rivales en un entorno cada vez más competitivo y exigente. Los equipos que puedan lograr aprovechar de manera efectiva estas herramientas tendrán una ventaja considerable en la consecución de sus objetivos deportivos y financieros.

1.2. Objetivos de desarrollo

El principal objetivo de desarrollo es crear un modelo con inteligencia artificial que ayude a los entrenadores a tomar mejores decisiones sobre qué jugadores utilizar en un partido mediante los datos obtenidos en los partidos anteriores. Para ello, los principales objetivos de desarrollo para este proyecto son:

1. **Obtener los datos de los partidos de fútbol de varias ligas.** Para ello, mediante el *scraping* se extraerán los datos de todos los partidos jugados en diferentes ligas y la información asociada a los jugadores.
2. **Limpiar, transformar y analizar los datos obtenidos.** Se deberán limpiar y transformar los datos obtenidos para que puedan ser utilizados por los modelos que se pretenden crear. Además, se debe realizar un análisis previo sobre los datos para detectar posibles patrones.
3. **Crear diferentes modelos de inteligencia artificial con los datos obtenidos.** En este punto, se debe evaluar el rendimiento que tienen los diferentes modelos sobre los datos obtenidos y se debe realizar una optimización de sus parámetros para mejorar la precisión obtenida.
4. **Seleccionar el mejor modelo y analizar su precisión obtenida.** Sobre todos los modelos evaluados, se deberá seleccionar el que mejor se comporte y se deberá de analizar que precisión tiene sobre partidos que estén pendientes por jugarse.

1.3. Motivación

El crecimiento en los últimos años de la inteligencia artificial ha despertado un interés en su aplicación en diversos campos como en el deporte. En el ámbito del fútbol, la capacidad de utilizar la inteligencia artificial para analizar datos complejos y tomar decisiones estratégicas concretas ofrece un gran potencial para incrementar el rendimiento de los equipos. Esta motivación se debe a la necesidad de los clubes por mantenerse competitivos en un entorno en constante evolución, donde la línea entre el éxito y el fracaso es muy estrecha.

El proyecto aparece como respuesta al incremento en la demanda de herramientas que permitan a los clubes mejorar en la gestión de sus recursos humanos, en concreto de sus jugadores. La inteligencia artificial tiene la capacidad de analizar grandes cantidades de datos sobre los jugadores y equipos, detectando patrones y tendencias que pueden pasar desapercibidos para las personas. Al integrar estas conclusiones en la toma de decisiones, los equipos pueden mejorar la eficiencia de su rotación de jugadores, incrementando así sus posibilidades de éxito en el campo.

Por último, la motivación detrás de este proyecto también se debe a su potencial para marcar un cambio significativo en la manera en que se realiza la gestión deportiva en el fútbol moderno. Al ofrecer a los clubes herramientas avanzadas de análisis y toma de decisiones, se espera que este proyecto ayude no solo a mejorar los resultados deportivos, sino también a fortalecer la posición competitiva y el rendimiento financiero de los equipos en un mercado cada vez más exigente y competitivo.

1.4. Aplicaciones similares

A continuación, se detallan aplicaciones y proyectos similares a lo que se pretende desarrollar y que pueden servir de referencia.

- **LaLiga Beyond Stats:** esta es una iniciativa de LaLiga que tiene como objetivo emplear las últimas tecnologías, relacionadas con el análisis de datos y la inteligencia artificial, para proporcionar una comprensión más profunda y completa de los partidos. Esta plataforma busca ofrecer a los aficionados, entrenadores, jugadores y clubes herramientas innovadoras para analizar y entender el rendimiento en el fútbol, más allá de las estadísticas habituales, a través de datos en tiempo real y visualizaciones interactivas, proporcionando así un enfoque más inteligente e interesante hacia el deporte [4].
- **Aplicación de la inteligencia artificial en la Premier League:** esta liga utiliza la inteligencia artificial para determinar las probabilidades de que un equipo gane un partido mediante el análisis de un amplio rango de datos. Estos factores abarcan datos históricos de partidos anteriores, como el rendimiento del equipo en casa y fuera de casa, su posición en la tabla de clasificación, su forma actual y lesiones de jugadores clave entre otros. Además, se tienen en cuenta variables más específicas, como la posesión de balón, los tiros a puerta, las oportunidades creadas y la efectividad en la defensa y el ataque. Estos datos son proporcionados a algoritmos de aprendizaje automático que son capaces de analizar patrones complejos y entrenar modelos predictivos para estimar las probabilidades de resultados de los partidos. De esta manera, la inteligencia artificial proporciona una herramienta poderosa para predecir resultados de partidos de fútbol con un alto grado de precisión, lo que puede ser utilizado por equipos, aficionados y casas de apuestas para tomar decisiones justificadas [5] [6].
- **Opta:** es una empresa líder en análisis y datos deportivos que es capaz de proporcionar información detallada y estadísticas sobre una amplia gama de eventos deportivos,

incluyendo fútbol, rugby, cricket y otros. Para ello, utiliza tecnologías avanzadas de recopilación y análisis de datos donde recopila datos en tiempo real durante los eventos deportivos y los convierte en información valiosa y estadísticas significativas que son utilizadas por equipos, entrenadores, medios de comunicación y aficionados para comprender mejor el juego, evaluar el rendimiento de los jugadores y equipos, y tomar decisiones justificadas. Opta se ha convertido en un recurso fundamental en el mundo del deporte para análisis de datos y seguimiento de estadísticas [7].

Capítulo 2

Metodo

2.1. Planificación

Este proyecto se pretende realizar en 5 semanas, comenzando el 6 de mayo de 2024 y finalizando el 9 de junio de 2024, empleando 38 horas por semana lo que dará un total de 190 horas que es el tiempo que se propone por parte de los profesores para realizar este trabajo. La planificación de este proyecto se hará por cada semana detallando el trabajo a realizar en cada una de ellas. En los siguientes puntos se detalla el trabajo y las tareas que se pretenden realizar en cada semana. En cada una de estas tareas se incluye la parte asociada a documentar el trabajo realizado para elaborar el informe final.

- **Semana previa al inicio: 29 de abril a 5 de mayo.** Esta semana previa al inicio del proyecto pretende dedicarse a la elaboración de los *scripts* que realizarán el *scraping* para obtener los datos. Además, en esta semana se debe empezar a analizar artículos relacionados con el trabajo que se pretende desarrollar para poder recoger buenas prácticas e ideas básicas. Un desglose más detallado de las tareas es:
 - Diseñar el modelo de datos.
 - Realizar los *scripts* que obtengan los datos mediante el *scraping*.
 - Validar que los datos obtenidos son correctos.
 - Revisar artículos relacionados con el trabajo que se pretende realizar.
- **Semana 1: 6 de mayo a 12 de mayo.** Esta semana se dedicará a la limpieza, transformación y análisis de los datos obtenidos mediante *scraping* para que los modelos puedan utilizarlos. Por otro lado, en esta semana también se deben de terminar de analizar los artículos que queden pendientes de leer. Un desglose más detallado de las tareas es:
 - Realizar una limpieza sobre los datos obtenidos.

- Realizar las transformaciones necesarias sobre los datos obtenidos para que puedan ser utilizados por los modelos.
 - Analizar los datos obtenidos mediante un dashboard para detectar patrones.
 - Terminar de revisar los artículos relacionados con el trabajo que se pretende realizar y resumir los conocimientos obtenidos.
- **Semana 2: 13 de mayo a 19 de mayo.** Esta semana se dedicará a establecer y crear los modelos que se pretenden probar. Un desglose más detallado de las tareas es:
 - Definir de manera teórica los diferentes modelos que se pretenden evaluar.
 - Definir las métricas que se utilizarán para evaluar los modelos creados.
 - Crear y entrenar los modelos sobre el conjunto de datos obtenido previamente.
 - **Semana 3: 20 de mayo a 26 de mayo.** Esta semana se dedicará a optimizar los parámetros de los modelos creados. Un desglose más detallado de las tareas es:
 - Optimizar los parámetros sobre los modelos creados.
 - Obtener los valores de las métricas seleccionadas para los modelos creados.
 - **Semana 4: 27 de mayo a 2 de junio.** Esta semana se dedicará a la selección del mejor modelo de entre todos los creados y evaluar su precisión sobre partidos que se vayan a jugar. Un desglose más detallado de las tareas es:
 - Analizar la precisión y diferentes métricas obtenidas sobre los diferentes modelos.
 - Justificar y seleccionar el modelo con una mayor precisión.
 - Evaluar la precisión y el rendimiento del modelo sobre partidos por jugarse.
 - **Semana 5: 3 de junio a 9 de junio.** Esta semana se dedicará a la revisión de todos los entregables del proyecto y corrección de fallos detectados. Un desglose más detallado de las tareas es:
 - Revisar la documentación a entregar e incorporar las partes que queden pendientes.
 - Revisar el código desarrollado, comprobar si existen errores y realizar su limpieza para que sea limpio.

Los días 10, 11 y 12 de junio se dejan como margen de tiempo hasta la entrega para si surgen imprevistos que hagan que el proyecto se retrase, partiendo de la premisa de que el informe final de este proyecto se debe entregar el 13 de junio.

2.2. Análisis de riesgos

A continuación, se detallan los riesgos más importantes que pueden aparecer durante el desarrollo del proyecto, estos son los que tienen más probabilidades de suceder y que más impacto sobre el proyecto tendrían. Para cada uno de ellos, se establecen planes de mitigación

para reducir la probabilidad de que ocurran y planes de contingencia para reducir su impacto en el proyecto si finalmente se ponen de manifiesto.

Las Tablas 2.1 a 2.5 muestran los riesgos encontrados para el proyecto.

Identificador	R-01
Título	Problemas al obtener los datos.
Descripción	Es posible que aparezcan problemas al extraer los datos de la página web seleccionada mediante <i>scraping</i> .
Probabilidad	Media.
Impacto	Alto.
Plan de mitigación	<ul style="list-style-type: none">■ Formarse en la tecnología que se utiliza para realizar el <i>scraping</i>.■ Buscar páginas web que tengan los mismos datos y que puedan servir como suplentes en caso de que la página web considerada como primera opción no permita la extracción de los datos.
Plan de contingencia	<ul style="list-style-type: none">■ Aplicar alternativas en el código para obtener los datos.■ Obtener los datos deseados desde otra de las páginas que se propusieron como suplentes.

Tabla 2.1: R-01. Problemas al obtener los datos.

2.2. ANÁLISIS DE RIESGOS

Identificador	R-02
Título	Poco tiempo para trabajar en el proyecto.
Descripción	Debido a que el único responsable del proyecto también trabaja en una empresa y otras causas, es posible que no se disponga del suficiente tiempo para desarrollar este proyecto.
Probabilidad	Alta.
Impacto	Alto.
Plan de mitigación	<ul style="list-style-type: none">■ Comenzar el desarrollo del proyecto lo antes posible.■ Identificar las tareas más importantes a realizar.■ Planificar que tareas se deben realizar en cada semana y como debe ir avanzando el proyecto. <i>sprints</i>.
Plan de contingencia	<ul style="list-style-type: none">■ Centrarse en realizar las tareas más importantes.■ Detectar partes que conlleven un esfuerzo y tiempo importante y que proporcionen menor beneficio y posponerlas.

Tabla 2.2: R-02. Poco tiempo para trabajar en el proyecto.

Identificador	R-03
Título	Fallo en el equipo de trabajo donde se realiza el proyecto.
Descripción	El equipo de trabajo con el que se desarrolla el proyecto, tanto el código como la documentación, deja de funcionar o su funcionamiento se ve afectado lo que dificulta trabajar con él.
Probabilidad	Baja.
Impacto	Alto.
Plan de mitigación	<ul style="list-style-type: none">■ Tener preparado un equipo suplente por si el equipo principal sufre fallos.■ Realizar copias de seguridad habitualmente para mantener los datos del proyecto a salvo.
Plan de contingencia	<ul style="list-style-type: none">■ Obtener un equipo de trabajo con el que se pueda seguir trabajando y arreglar el averiado.■ Hacer las tareas que se puedan realizar sin el equipo original y después cuando se obtenga un nuevo equipo, incorporar este contenido.

Tabla 2.3: R-03. Fallo en el equipo de trabajo donde se realiza el proyecto.

2.2. ANÁLISIS DE RIESGOS

Identificador	R-04
Título	Problemas con la conexión a internet.
Descripción	La conexión a internet es probable que sea lenta o se pierda en alguna ocasión debido a la localización donde el alumno desarrolla el proyecto.
Probabilidad	Baja.
Impacto	Medio.
Plan de mitigación	<ul style="list-style-type: none">■ Establecer lugares alternativos donde se tenga conexión a internet.■ Establecer tareas que se puedan realizar sin conexión a internet.
Plan de contingencia	<ul style="list-style-type: none">■ Realizar tareas que no necesiten conexión a internet.■ Retrasar las tareas que requieran conexión a internet.■ Desplazarse otro lugar con conexión a internet.

Tabla 2.4: R-04. Problemas con la conexión a internet.

Identificador	R-05
Título	Enfermedad del alumno.
Descripción	El alumno puede sufrir problemas de salud debido a su historial pasado lo que puede retrasar la consecución del proyecto.
Probabilidad	Media.
Impacto	Medio.
Plan de mitigación	<ul style="list-style-type: none">■ Llevar buenos hábitos cuidando la salud y evitar lugares con virus o personas que estén contagiadas.■ Planificar márgenes de tiempo y flexibilidad en la planificación para que si aparecen imprevistos se tenga tiempo de sobra para finalizar el trabajo.
Plan de contingencia	<ul style="list-style-type: none">■ Adaptar las tareas que no se hayan podido finalizar a las siguientes semanas.■ Reducir el alcance del proyecto y la funcionalidad que se quiere implementar.

Tabla 2.5: R-05. Enfermedad del alumno.

Una vez descritos los riesgos y definido para cada uno su impacto, probabilidad y planes de mitigación y contingencia, se puede crear una matriz de riesgos. Esta matriz es una herramienta que se define para determinar la probabilidad e impacto de un riesgo, lo que ayuda a definir la prioridad asociada a cada riesgo y para así gestionarlos de la mejor manera posible [8].

La Tabla 2.6 muestra una Matriz de Riesgos.

Probabilidad / impacto	Bajo	Medio	Alto
Baja	Monitorizar	Monitorizar	Aplicar los planes de mitigación
Media	Monitorizar	Aplicar los planes de mitigación	Aplicar los planes de mitigación y tener preparados los planes de contingencia para si son necesario aplicarlos
Alta	Aplicar los planes de mitigación	Aplicar los planes de mitigación y tener preparados los planes de contingencia para si son necesario aplicarlos	Aplicar los planes de mitigación y tener preparados los planes de contingencia para si son necesario aplicarlos

Tabla 2.6: Matriz de Riesgos.

De esta forma, se pueden situar los riesgos definidos para este proyecto en la matriz de riesgos y así se podrá definir la prioridad de cada uno de ellos para establecer cómo actuar en función del impacto y la probabilidad que tienen.

La Tabla 2.7 sitúa los riesgos del proyecto en la Matriz de Riesgos.

Probabilidad / impacto	Bajo	Medio	Alto
Baja		R4	R3
Media		R5	R1
Alta			R2

Tabla 2.7: Situación de los riesgos del proyecto en la matriz de riesgos.

2.3. Alcance

El alcance de este proyecto abarca la evaluación y aplicación de diversas técnicas de inteligencia artificial para predecir el rendimiento de equipos de fútbol basándose en la rotación de jugadores. En primer lugar, se definirá la metodología y se seleccionarán las técnicas más correctas para el análisis de datos relacionados con la rotación de jugadores y el rendimiento deportivo. Este apartado incluirá la recopilación, preprocesamiento y análisis de los datos de las ligas, equipos y jugadores de fútbol seleccionados. Para la parte del análisis de los datos,

se realizará un pequeño dashboard que permita sacar conclusiones y detectar patrones sobre los datos para así facilitar su comprensión.

Las ligas sobre las que se obtendrán y utilizarán los datos serán LaLiga EA Sports (primera división española), LaLiga Hypermotion (segunda división española), Premier League (primera división inglesa), Bundesliga (primera división alemana) y Serie A (primera división italiana).

Además, el alcance del proyecto se pretende que también implique la implementación y ajuste de modelos de inteligencia artificial para la predicción del rendimiento deportivo en función de la rotación de jugadores. Para ello, se explorarán diversas técnicas de inteligencia artificial y *machine learning*, como redes neuronales, árboles de decisión, y métodos de aprendizaje automático supervisado y no supervisado, con el objetivo de detectar aquellas que mejor se adapten a las características de este problema. Sobre cada una de ellas, se realizará una optimización de parámetros para mejorar todo lo posible su precisión. Finalmente, se realizará una evaluación de los modelos desarrollados, utilizando métricas de rendimiento para definir su eficacia y precisión en la predicción del rendimiento de los equipos. Después de esto, se seleccionará el mejor modelo y se evaluará su rendimiento en la actualidad aplicándolo sobre partidos que estén por jugarse.

Además de todos estos aspectos comentados, se documentarán y analizarán todas las tareas realizadas en el proyecto, con el objetivo de ofrecer recomendaciones para la gestión de la rotación de jugadores en equipos de fútbol, así como posibles áreas de mejora y futuras investigaciones para este proyecto.

2.4. Metodología

La metodología de este proyecto tiene como base un enfoque que comprende varias etapas clave. En primer lugar, se realizará una revisión de diferentes artículos sobre técnicas de inteligencia artificial aplicadas al análisis de datos deportivos, centrándose especialmente en la predicción del rendimiento de equipos de fútbol basándose en la rotación de los jugadores. Esta revisión ayudará a identificar las mejores prácticas y los enfoques que pueden ser más relevantes para el desarrollo del proyecto.

Posteriormente, se llevará a cabo la recopilación y preparación de datos, donde se recogerán conjuntos de datos históricos que abarquen información relevante sobre la rotación de jugadores y el rendimiento deportivo de equipos de fútbol en las ligas seleccionadas. Esta etapa incluye la limpieza de datos, la selección de características pertinentes, la preparación de los datos para su análisis posterior y un breve análisis sobre ellos para detectar patrones.

Una vez preparados los datos, se realizará la implementación y evaluación de modelos de inteligencia artificial. Se probarán diversas técnicas de modelado, incluyendo redes neuronales, árboles de decisión, y métodos de aprendizaje supervisado y no supervisado. Los modelos se entrenarán y ajustarán utilizando los datos que hemos obtenido previamente, y se evaluará su rendimiento utilizando diferentes métricas. Esta fase permitirá detectar los modelos más eficaces y precisos para predecir el rendimiento deportivo basado en la rotación

de jugadores.

2.5. Obtención de los datos

Para la obtención de los datos, se creará un *script* en Python que extraiga los datos de la página de Resultados De Fútbol [9] mediante *scraping*.

El *scraping* [10] es una técnica utilizada para extraer automáticamente información de sitios web de forma automatizada. Consiste en el análisis y la recopilación de datos de páginas web. Estos programas acceden a la página web de la que se desean obtener los datos, identifican los componentes clave dentro del código HTML y extraen su información para su posterior procesamiento o análisis. El *scraping* es una herramienta útil para obtener datos en gran volumen de manera rápida y eficiente, y es aplicada en variedad de aplicaciones. En este proyecto se ha utilizado esta técnica para obtener los datos ya que no se ha podido encontrar ningún conjunto de datos que recoja información sobre los datos históricos de los partidos en las ligas seleccionadas. Además, mediante el *scraping*, se puede fácilmente ir incorporando a los datos utilizados para crear los modelos los nuevos datos asociados a los últimos partidos jugados.

2.6. Tecnologías utilizadas

A continuación, se detallan las tecnologías base que se utilizaran en el proyecto. Sin embargo, estas tecnologías durante el desarrollo del proyecto pueden variar:

- **PowerBI:** es una plataforma de análisis empresarial que ha sido desarrollada por Microsoft que permite a los usuarios visualizar y extraer conclusiones de los datos de manera intuitiva y efectiva. Permite la conexión con una amplia variedad de fuentes de datos para que luego se puedan transformar y modelar estos datos en paneles interactivos. Con Power BI, se pueden crear visualizaciones personalizadas, como gráficos, tablas y mapas, para explorar y analizar los datos de manera dinámica lo que permite detectar patrones de manera visual. En este proyecto se utiliza esta tecnología para realizar un primer análisis sobre los datos obtenidos y detectar posibles patrones [11].
- **Python:** es un lenguaje de programación versátil y de alto nivel que tiene una enorme popularidad en diversos campos, destacando en la ciencia de datos. Este lenguaje incorpora diferentes bibliotecas que permiten realizar diferentes tareas lo que le convierte en uno de los lenguajes con más funcionalidades diferentes [12]. Una de las bibliotecas más utilizadas en Python es BeautifulSoup para realizar tareas de *scraping*. Esta biblioteca permite analizar y extraer datos de páginas web de manera sencilla y eficiente, ayudando al programador a realizar la manipulación de la estructura HTML de los sitios web para obtener la información que se desee sobre el sitio. Con BeautifulSoup, se pueden crear *scripts* que naveguen por el contenido de una página web, identifiquen elementos específicos y que permitan extraer datos de manera automatizada [13]. Por

otro lado, Python es mundialmente utilizado en el campo de la inteligencia artificial y el *machine learning* por bibliotecas como scikit-learn. Esta es una biblioteca que ofrece una diversa gama de herramientas para la creación de algoritmos de *machine learning* como se pretende en este proyecto. Con scikit-learn, se pueden crear y entrenar modelos de *machine learning* de forma eficiente, utilizando algoritmos ya definidos y técnicas avanzadas de análisis de datos [14]. Además, Python ofrece la biblioteca pandas, que facilita la manipulación y el análisis de datos estructurados mediante la introducción de los DataFrames. Estos son estructuras de datos bidimensionales que tienen la capacidad de almacenar y manipular datos de manera eficiente, de manera similar a una tabla de base de datos o una hoja de cálculo. Con pandas, se pueden cargar datos desde multitud de fuentes, realizar operaciones de limpieza y transformación de datos, y realizar análisis estadísticos y exploratorios de manera rápida. Esto hace que pandas sea una herramienta indispensable para el almacenamiento y la manipulación de datos en proyectos de ciencia de datos y análisis de datos en Python como es en este caso [15].

- **Keras:** es una API de alto nivel para la construcción, entrenamiento y evaluación de modelos de redes neuronales mediante Python. Destaca por su facilidad de uso y su enfoque en la creación rápida y sencilla de modelos de aprendizaje profundo. Ofrece una sintaxis simple y una abstracción de alto nivel que permite crear modelos complejos de manera rápida, lo que lo convierte en una herramienta excelente que brinda flexibilidad y potente para trabajar en una amplia gama de proyectos de inteligencia artificial y aprendizaje profundo. Esta tecnología se utiliza en este proyecto para crear modelos de redes neuronales que pueden tener un buen rendimiento sobre el conjunto de datos proporcionado [16].
- **Tensorflow:** es una biblioteca de aprendizaje automático de código abierto que ha sido desarrollada por Google que proporciona una plataforma flexible y escalable para construir, entrenar y desplegar modelos de aprendizaje profundo. Esta tecnología destaca por su capacidad para trabajar con inmensos volúmenes de datos y su eficiencia en la ejecución en variedad de plataformas. TensorFlow ofrece una amplia gama de herramientas y funcionalidades, incluyendo la construcción de redes neuronales convolucionales, recurrentes y generativas, así como la experimentación con técnicas avanzadas. Por lo tanto, TensorFlow es una opción popular para proyectos de inteligencia artificial y aprendizaje automático en diversas industrias. Esta tecnología se utiliza en este proyecto para la creación de modelos más avanzados que pueden tener un rendimiento elevado sobre los datos proporcionados [17].

Capítulo 3

Conclusiones

3.1. Conclusiones

Las conclusiones de este proyecto destacan la trascendencia y el alcance que provoca la aplicación de diversas técnicas de inteligencia artificial en el contexto del fútbol, específicamente en la predicción del rendimiento de los equipos mediante el análisis de la rotación de jugadores. Este proyecto puede ser capaz de revelar que la implementación de herramientas de inteligencia artificial, como modelos de aprendizaje automático y análisis de datos, puede ayudar en la toma de decisiones a los entrenadores y directivos. En este proyecto se va a poder observar cómo estas tecnologías pueden ofrecer información crucial que tengan una gran repercusión en la toma de decisiones estratégicas de entrenadores y directivos de equipos, permitiéndoles optimizar la rotación de jugadores de manera más precisa y efectiva.

Además, en este proyecto destaca la importancia de disponer de conjuntos de datos completos y de calidad para proporcionar a estos modelos de inteligencia artificial de manera adecuada. La recopilación y preparación de datos precisos y relevantes sobre la rotación de jugadores y el rendimiento deportivo se ha establecido como un componente fundamental para el éxito de este proyecto.

Este proyecto puede ayudar a destacar la necesidad de desarrollar herramientas y metodologías específicas que faciliten la integración de la inteligencia artificial en la gestión deportiva, lo que implicaría una colaboración conjunta entre expertos en deportes y científicos de datos.

En última instancia, este proyecto pretende mostrar el potencial de la inteligencia artificial para transformar y mejorar la gestión y el desempeño de los equipos de fútbol analizando los datos sobre la rotación de sus jugadores. Los hallazgos de este proyecto pretenden invitar a continuar investigando y desarrollando este campo, explorando nuevas tecnologías y metodologías que puedan maximizar el impacto positivo de la inteligencia artificial en el mundo del fútbol.

3.2. Líneas de trabajo futuras

En este proyecto se pretenden analizar el rendimiento de diferentes modelos de inteligencia artificial sobre el desempeño de los equipos de futbol basándose en los datos sobre la rotación de sus jugadores. Sin embargo, por la naturaleza del proyecto, debido a que es un proyecto académico, no se pretende profundizar al máximo en estos aspectos y por tanto a continuación se definen posibles mejoras que puede tener el proyecto en el futuro y que no se pretenden realizar en este trabajo.

- **Incorporación de más ligas:** este aspecto podría incrementar la utilidad del sistema desarrollado de manera que sea capaz de ayudar a dirigentes y entrenadores de más clubes y países. Al abarcar más ligas más usuarios podrían utilizar el sistema.
- **Incorporación de más parámetros relacionados con la rotación de los jugadores:** este aspecto podría ayudar a mejorar el rendimiento de los modelos creados y por tanto proporcionar mejores resultados. En este proyecto desarrollado se pretenden utilizar los parámetros y variables más útiles, pero como mejora futura, se podría considerar analizar más parámetros que analicen diferentes datos.
- **Automatizar todo el código para que actualice los datos con los resultados de los últimos partidos:** en este proyecto, de manera inicial, se ha planteado que se deban ejecutar de manera manual los *scripts* para la obtención de los datos de los últimos partidos, pero sin embargo, esta tarea sería importante automatizarla para el futuro.
- **Desarrollar una aplicación web para mostrar los datos obtenidos:** como mejora final, se podría desarrollar una aplicación web que muestre de una forma más amigable los datos obtenidos de los modelos y que puedan ayudar a los entrenadores y directivos.

Bibliografía

- [1] Kristalina Georgieva. La economía mundial transformada por la inteligencia artificial ha de beneficiar a la humanidad. <https://www.imf.org/es/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity>. Accessed: 2024-4-23.
- [2] MejorConSalud. ¿Cómo se usa la inteligencia artificial en el fútbol profesional? <https://mejorconsalud.as.com/inteligencia-artificial-futbol-profesional/>. Accessed: 2024-4-23.
- [3] José Luis del Olmo Arriaga. El gran negocio del futbol. <https://www.theeconomyjournal.com/texto-diario/mostrar/1525487/gran-negocio-futbol>. Accessed: 2024-4-23.
- [4] LaLiga. LaLiga Beyond Stats. <https://www.laliga.com/beyondstats>. Accessed: 2024-4-23.
- [5] Oracle. Premier League y Match Insights, con tecnología Oracle Cloud: Reimaginando la experiencia de los aficionados. <https://www.oracle.com/es/premier-league/>. Accessed: 2024-4-23.
- [6] Somos Futboleros. Inteligencia artificial predice el campeón de la Premier League. <https://onefootball.com/es/noticias/inteligencia-artificial-predice-el-campeon-de-la-premier-league-38801122>. Accessed: 2024-4-23.
- [7] Opta. OPTA DATA. <https://www.statsperform.com/opta/>. Accessed: 2024-4-23.
- [8] Team Asana. Matriz de riesgos: cómo evaluar los riesgos para lograr el éxito del proyecto. <https://asana.com/es/resources/risk-matrix-template>. Accessed: 2023-6-12.
- [9] ResultadosFutbol. ResultadosFutbol. <https://www.resultados-futbol.com/>. Accessed: 2024-4-23.
- [10] Kinsta. ¿Qué Es el Web Scraping? Cómo Extraer Legalmente el Contenido de la Web. <https://kinsta.com/es/base-de-conocimiento/que-es-web-scraping/>. Accessed: 2024-4-23.

- [11] Microsoft. ¿Qué es Power BI? <https://learn.microsoft.com/es-es/power-bi/fundamentals/power-bi-overview>. Accessed: 2024-4-23.
- [12] Pablo Londoño. Qué es Python, para qué sirve y cómo se usa (+ recursos para aprender). <https://blog.hubspot.es/website/que-es-python>. Accessed: 2024-4-23.
- [13] datascientest. Beautiful Soup : ¿cómo aprender a hacer web scraping en Python? <https://datascientest.com/es/beautiful-soup-aprender-web-scraping>. Accessed: 2024-4-23.
- [14] master-data-scientist. SCIKIT-LEARN, HERRAMIENTA BÁSICA PARA EL DATA SCIENCE EN PYTHON. <https://www.master-data-scientist.com/scikit-learn-data-science/>. Accessed: 2024-4-23.
- [15] datascientest. Pandas : La biblioteca de Python dedicada a la Data Science. <https://datascientest.com/es/pandas-python>. Accessed: 2024-4-23.
- [16] datascientest. Keras: todo sobre la API de Deep Learning. <https://datascientest.com/es/keras-la-api-de-deep-learning>. Accessed: 2024-4-23.
- [17] Jon Larkin Alonso. ¿Qué es TensorFlow y para qué sirve? <https://www.incentro.com/es-ES/blog/que-es-tensorflow>. Accessed: 2024-4-23.