



UNIVERSIDAD AUTÓNOMA METROPOLITANA
Unidad Azcapotzalco

División de Ciencias Básicas e Ingeniería
Licenciatura en Ingeniería en Computación



Predicción de resultados de fútbol profesional con redes neuronales

Proyecto tecnológico

Tercera versión

Trimestre 2025-Primavera

Alumno

Daniel García Segura

2213000716

al2213000716@azc.uam.mx

Asesor

Dr. Rodrigo Alexander Castro Campos

Profesor Asociado

Departamento de Sistemas

racc@azc.uam.mx

28 de agosto de 2025

Declaratoria

En caso de que el Comité de Estudios de la Licenciatura en Computación apruebe la realización de la presente propuesta, otorgamos nuestra autorización para su publicación en la página de la División de Ciencias Básicas e Ingeniería.

Daniel García Segura
Alumno

Dr. Rodrigo Alexander Castro Campos
Asesor

1. Introducción

Según el *International Football Association Board* (IFAB), el fútbol es un deporte en el que participan dos equipos que intentan anotar goles introduciendo el balón en la portería rival, sin utilizar brazos o manos, salvo el portero dentro de su área [1]. En su versión profesional, el fútbol es una actividad organizada en la que los jugadores y el personal reciben una remuneración. Está regido por organismos como la FIFA y las federaciones nacionales, desarrollando competiciones oficiales cuyos resultados determinan títulos, ascensos o descensos. Este deporte representa una industria con ingresos derivados de transmisiones, publicidad, entradas y productos.

Los resultados históricos y las estadísticas juegan un papel importante para predecir o incluso para influir en el resultado de un partido de fútbol en las ligas profesionales. Factores como el rendimiento previo de un equipo, los enfrentamientos directos entre ambos rivales, la cantidad de goles marcados y recibidos, el porcentaje de victorias como local o visitante, además del estado físico de los jugadores, son datos que se analizan antes de cada encuentro. Estos registros ayudan a tomar decisiones tácticas y a entender qué equipo llega con ventaja, aunque en el fútbol siempre existe un margen para la sorpresa.

Una red neuronal artificial es un modelo computacional inspirado en el funcionamiento del cerebro humano, diseñado para que las computadoras aprendan a reconocer patrones y tomar decisiones. Una red está formada por capas de unidades computacionales denominadas neuronas artificiales, las cuales individualmente realizan cálculos sencillos, pero que en la red se encuentran conectadas entre sí para poder realizar cálculos complejos. Las interconexiones de la red tienen pesos y umbral que se pueden ajustar hasta lograr que ésta aprenda a realizar las tareas específicas de interés, tales como reconocer rostros, traducir textos o predecir resultados. Se le denomina entrenamiento al proceso de ajustar las interconexiones de la red hasta volverla funcional. Esto normalmente se hace presentándole la red un conjunto de entradas y salidas esperadas. [2]

La Liga Premier de Inglaterra es una fuente ideal de datos para entrenar redes neuronales, al ser una de las más competitivas, además de ofrecer registros detallados y accesibles de cada partido, incluyendo estadísticas como goles, posesión, alineaciones, tarjetas y tiros a puerta. La frecuencia regular de encuentros contribuye a facilitar el entrenamiento de una red neuronal.

En este documento se propone desarrollar un sistema web que genere pronósticos de partidos de fútbol profesional a partir de una red neuronal artificial que será entrenada con resultados históricos de partidos de fútbol de la Liga Premier de Inglaterra a partir de 1993. Esta información le permitirá a la red aprender patrones de comportamiento de y entre equipos para así poder predecir los resultados de partidos futuros.

2. Justificación

El fútbol, siendo el deporte más popular del mundo, tiene un impacto que va más allá de la competencia. Constituye una actividad de profundo valor emocional, social, cultural y económico, capaz de reunir a personas de diversos orígenes y forjar una identidad colectiva poderosa.

Los pronósticos de partidos son una parte esencial del entorno que rodea al fútbol. Por una parte, generan emoción y expectativa entre los aficionados, fomentan debates en redes sociales y medios, y son fundamentales para industrias como las apuestas deportivas. También son herramientas valiosas para analistas y técnicos que evalúan posibles escenarios de juego. Pronosticar no solo entretiene, también tiene un alto valor práctico y económico. Cabe mencionar que las empresas de televisión y los propios equipos gastan millones de dólares al año tanto en herramientas computacionales como en analistas humanos.

El uso de redes neuronales artificiales responde a la necesidad de aprovechar inteligentemente grandes volúmenes de datos disponibles, tales como bases de datos de resultados históricos. Al ser aproximadores universales, las redes neuronales permiten identificar patrones complejos y generar predicciones automáticas con base en los datos vistos. Las redes neuronales se usan en campos como la inteligencia artificial, la medicina, las finanzas y la industria, porque permiten que las máquinas aprendan sin necesidad de programar cada paso manualmente. De este modo, ofrecen una solución tecnológica capaz de procesar información dispersa y convertirla en pronósticos útiles.

3. Objetivos

3.1. Objetivo general

Desarrollar un sistema web que pronostique resultados de partidos de fútbol profesional evaluando una red neuronal artificial entrenada a partir de resultados históricos, para ofrecer a los usuarios una herramienta de análisis predictivo para este deporte.

3.2. Objetivos específicos

1. Recopilar los datos históricos relevantes de los partidos de la Liga Premier de Inglaterra, y almacenar la información de manera estructurada en una base de datos relacional.
2. Diseñar y entrenar una red neuronal artificial que aprenda patrones a partir de los datos históricos y pronostique los resultados de partidos futuros.
3. Desarrollar una aplicación web que le muestre al usuario la información histórica almacenada y la agenda de partidos futuros, y que evalúe la red neuronal para generar un pronóstico de resultado para estos.
4. Evaluar y validar el rendimiento del sistema mediante pruebas con datos reales, midiendo la precisión de los pronósticos.

4. Trabajos relacionados

A continuación, se presentan los trabajos relacionados con el proyecto a desarrollar.

4.1. Deep multilayer neural network for predicting the winner of football matches [3].

Este estudio propone una red neuronal profunda multicapa para predecir el ganador de partidos de fútbol. Utiliza características estadísticas como goles, tiros, tarjetas y faltas, aplicando diferentes combinaciones para analizar su impacto en la precisión. Se entrenó y evaluó sobre datos históricos estructurados. La red logró una precisión superior al 60 % en varias configuraciones.

4.2. Is Football Unpredictable? Predicting Matches Using Neural Networks [4].

Este artículo investiga la predictibilidad del fútbol mediante una red neuronal profunda con datos de trece ligas, incluyendo la Liga Premier de Inglaterra. El modelo emplea estadísticas agregadas para definir la potencia relativa de ataque, defensa y mediocampo. El artículo también estudia la complejidad de predecir ciertos tipos de partidos.

4.3. A Deep Learning Framework for Football Match Prediction [5].

Este artículo desarrolla un marco predictivo para partidos de fútbol utilizando redes neuronales profundas. El autor aplica técnicas específicas para capturar la secuencia temporal de eventos en los partidos, utilizando un conjunto de datos con resultados históricos, clasificaciones de equipos e información de rendimiento. El modelo se probó en la Copa Mundial de la FIFA 2018, donde predijo correctamente el 63.3 % de los partidos de la fase de grupos. Concluye que el aprendizaje profundo es una herramienta prometedora para la predicción de resultados deportivos.

4.4. Generación y resolución de niveles de un videojuego de plataformas usando redes neuronales [6].

Este proyecto tecnológico se enfoca en la generación y resolución de niveles para un videojuego de plataformas mediante el uso de dos redes neuronales distintas. Implementa una red neuronal para generar los niveles del juego con base en una dificultad especificada y una segunda red que funciona como un agente capaz de resolverlos. El sistema se complementa con un motor que define las reglas y un módulo de visualización para pruebas gráficas.

4.5. Prediction of On-Base Percentage (OBP) of Baseball Players Using Neural Networks [7].

Este artículo se enfoca en la predicción del porcentaje de embasado (OBP) de jugadores de béisbol mediante redes neuronales. Se propuso una arquitectura de perceptrón multicapa compuesta por capas de entrada, ocultas y de salida. Las características de entrada incluyeron estadísticas clave como hits, bases por bolas y turnos al bate. El modelo fue entrenado con datos históricos y el preprocesamiento de datos se consideró un paso fundamental.

4.6. Redes neuronales artificiales en la producción de tecnología educativa para la enseñanza de la diagonalización [8]

Este trabajo presenta el desarrollo de un sistema basado en redes neuronales como apoyo al proceso de enseñanza-aprendizaje del tema de diagonalización en álgebra lineal. El modelo se estructura en tres etapas: diseño del sistema, entrenamiento de la red y validación pedagógica. La investigación demuestra el potencial de las redes neuronales como herramienta complementaria para mejorar la comprensión de conceptos matemáticos complejos.

En la Tabla 1 se resumen las similitudes y diferencias de los trabajos citados y el propuesto.

Tabla 1. Comparación cualitativa de los trabajos relacionados con el proyecto.

Ref.	Similitudes	Diferencias
[3]	<ul style="list-style-type: none"> ■ Predicción de resultados de fútbol. ■ Aplicación en la Liga Premier de Inglaterra. ■ Uso de redes neuronales. ■ Entrenamiento basado en datos históricos y estadísticas de los partidos. 	<ul style="list-style-type: none"> ■ El trabajo referido emite la salida en crudo. El proyecto propuesto desarrollará un sistema web que muestre las predicciones de la red. ■ El trabajo referido describe el diseño y la evaluación de la red, pero no proporciona la implementación.

Continuación de la Tabla 1

Ref.	Similitudes	Diferencias
[4]	<ul style="list-style-type: none"> ■ Predicción de resultados de fútbol. ■ Uso de redes neuronales. ■ Entrenamiento basado en datos históricos y estadísticas de los partidos. 	<ul style="list-style-type: none"> ■ El trabajo referido no analiza la Liga Premier de Inglaterra. ■ El trabajo referido profundiza en cómo formular matemáticamente la relación de diversas estadísticas en el modelo. El proyecto propuesto sólo contempla el entrenamiento de la red con datos obtenidos directamente del desarrollo de los partidos.
[5]	<ul style="list-style-type: none"> ■ Predicción de resultados de fútbol. ■ Uso de redes neuronales. ■ Entrenamiento basado en datos históricos y estadísticas de los partidos. 	<ul style="list-style-type: none"> ■ El trabajo referido no analiza la Liga Premier de Inglaterra. En este proyecto se propone trabajar sobre ésta. ■ El trabajo referido considera información dinámica tal como los eventos ocurridos dentro de los propios partidos a pronosticar. El proyecto propuesto contempla sólo estadísticas globales históricas de los equipos implicados.
[6]	<ul style="list-style-type: none"> ■ Uso de redes neuronales. 	<ul style="list-style-type: none"> ■ El trabajo referido entrena una red generadora, mientras que mientras que el proyecto propuesto entrena una red predictora. ■ El trabajo referido trata de redes neuronales aplicadas en videojuegos de plataforma, mientras que el trabajo propuesto las aplica en juegos deportivos.
[7]	<ul style="list-style-type: none"> ■ Uso de redes neuronales para la predicción de resultados deportivos. 	<ul style="list-style-type: none"> ■ El trabajo referido predice resultados de béisbol. ■ Las variables consideradas por el trabajo referido se centran en el rendimiento de los jugadores, mientras que el proyecto propuesto considera sólo los resultados de partidos.

Continuación de la Tabla 1

Ref.	Similitudes	Diferencias
[8]	<ul style="list-style-type: none"> ■ Uso de redes neuronales. ■ Aplicación de inteligencia artificial para facilitar la toma de decisiones. 	<ul style="list-style-type: none"> ■ Uso en tecnología educativa.

5. Descripción técnica

El proyecto estará estructurado de forma modular, integrando procesamiento de datos, predicción con redes neuronales y publicación de resultados mediante un servicio web. Esta organización, mostrada en la Figura 1, busca facilitar el desarrollo y mantenimiento del sistema.

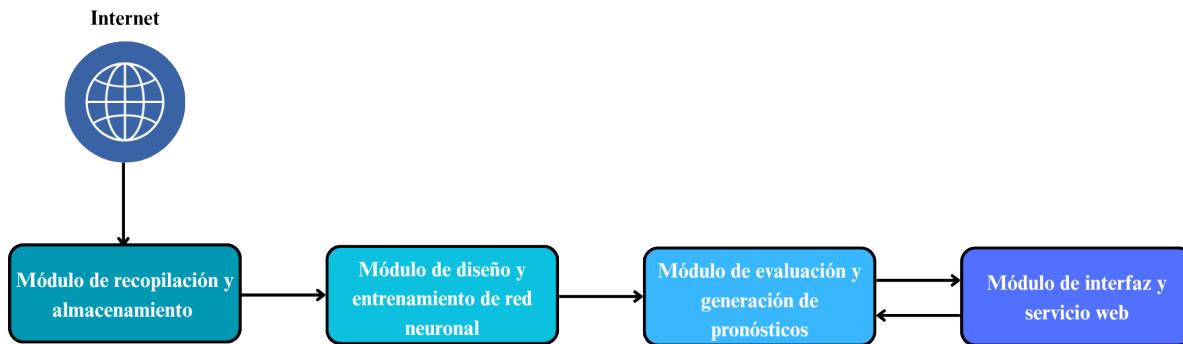


Figura 1. Módulos del sistema y sus interacciones.

5.1. Módulo de recopilación y almacenamiento de datos.

Este módulo es responsable de obtener la información histórica de los partidos pasados y futuros de la Liga Premier de Inglaterra a partir de archivos públicos en internet en formato CSV [9]. En particular, se encarga de descargar de forma automática los archivos disponibles, descartar los datos no relevantes contenidos en las fuentes consultadas y de almacenar los datos útiles de forma estructurada y tipada en una base de datos relacional.

5.2. Módulo de diseño y entrenamiento de red neuronal.

Este módulo construye una red neuronal del tipo perceptrón multicapa, la cual es adecuada para problemas de clasificación multiclase. La red se entrena con los partidos históricos recopilados y sus resultados, con el objetivo de que aprenda a pronosticar el resultado de nuevos encuentros. A partir de dichos datos, se construyen los conjuntos de entrenamiento y validación. La entrada de la red incluirá al menos la fecha del partido y los identificadores de los equipos que juegan. La salida consistirá en una probabilidad para cada uno de los tres posibles resultados: gana el equipo local, gana el equipo visitante o empate.

5.3. Módulo de evaluación de la red neuronal y generación de pronósticos.

Esta parte del sistema es la encargada de recuperar la red neuronal entrenada que se almacenó en el módulo anterior, tomar la lista de partidos futuros agendados tal como se almacenaron en la

base de datos, ejecutar el proceso de inferencia de la red neuronal para cada uno de estos partidos futuros, para finalmente construir un archivo en formato JSON con la información de los partidos, así como la probabilidad estimada por la red para sus posibles resultados.

5.4. Módulo de interfaz y servicio web.

Se expondrán en un sistema web los resultados históricos almacenados en la base de datos, así como los pronósticos generados por la red para partidos futuros mediante una plataforma accesible desde cualquier navegador a través de internet. Se implementará un servicio web utilizando REST para el acceso a los datos a través de *endpoints* bien definidos. Esta arquitectura facilita la integración con otras aplicaciones y la escalabilidad del sistema en caso de futuras ampliaciones.

6. Especificación técnica

La fase inicial en el desarrollo del proyecto consiste en recopilar los datos que conformarán el conjunto de entrenamiento. Para ello, la información será extraída de la página `football-data.co.uk` [9], una fuente reconocida en el registro de partidos correspondientes a las principales competiciones futbolísticas.

La información histórica de los partidos se almacenará en una base de datos relacional basada en MariaDB [10], elegida por su estabilidad, facilidad de uso y compatibilidad con múltiples entornos de despliegue. Para el desarrollo del sistema se utilizará el lenguaje de programación Python, debido a su flexibilidad y extensa comunidad de soporte en el ámbito de la programación web en el *backend*, el análisis de datos y la inteligencia artificial. La conexión entre el sistema y la base de datos se gestionará mediante el uso de la biblioteca PyMysql [11].

El modelo predictivo se implementará utilizando la biblioteca PyTorch [12], la cual permite diseñar redes neuronales de forma modular y eficiente, facilitando tanto la etapa de entrenamiento como la inferencia sobre nuevos datos. El resultado de la evaluación de la red neuronal para una lista de partidos de fútbol agendados se emitirá en formato JSON, el cual es uno de los formatos con mayor interoperabilidad entre aplicaciones, en particular con aplicaciones web.

El sitio web que visualizará los datos históricos recopilados, la agenda de partidos futuros y sus predicciones, se desarrollará utilizando tecnologías estándar del lado del cliente: HTML para la estructura del contenido, CSS para su presentación visual y JavaScript para la interacción dinámica. Las peticiones desde el cliente o *frontend* al *backend* se harán siguiendo el estilo REST, con amplio soporte tanto en JavaScript como en Python. El sistema se desarrollará de modo tal que se pueda actualizar su base de datos de forma periódica fácilmente, recopilando y almacenando los datos más recientes en la base de datos, procesándolos mediante el modelo predictivo y generando el archivo JSON con las nuevas predicciones.

El proyecto se considerará finalizado cuando el sistema sea capaz de realizar pronósticos con base en la información almacenada, logrando una precisión de al menos el 60 % en partidos con resultados conocidos y generando pronósticos realistas para partidos futuros agendados. Además, deberá ser capaz de generar los resultados en formato JSON y presentarlos en la interfaz web, todo ello funcionando correctamente con datos reales de al menos cinco mil partidos históricos y treinta partidos futuros agendados.

Al concluir el proyecto de integración se entregará a la Coordinación de Estudios de Ingeniería en Computación una carpeta digital que incluirá el reporte final del proyecto en un archivo PDF (sin restricciones)¹, el código fuente de la aplicación en un archivo comprimido (sin restricciones)².

¹Debe poder visualizarse sin solicitar contraseña

²Debe poder descomprimirse sin solicitar contraseña

Además, la sección de apéndices del reporte final contendrá al menos un listado del código fuente desarrollado.

7. Calendario de actividades

Las actividades a realizar durante el Trimestre 2025-O en la UEA Proyecto de Integración en Ingeniería en Computación I (1100113) se presentan en la Tabla 2.

Tabla 2. Listado de actividades a realizar durante el Trimestre 2025-O.

No.	Actividad	Horas	Entregable
1	Recopilar los datos históricos y la agenda de partidos futuros	9	Datos en formato no necesariamente estructurado.
2	Diseñar e implementar el modelo entidad-relación de la base de datos para almacenar los datos recopilados	18	Modelo entidad-relación y programa de creación de tablas en MariaDB
3	Importar los datos recopilados a la base de datos relacional	9	Programa de procesamiento de datos recopilados e inserción en la base de datos relacional
4	Diseñar la arquitectura de la red neuronal para predicción de resultados de fútbol profesional	27	Descripción de la entrada y salida de la red, número de capas, número de neuronas por capa y funciones de activación
5	Entrenar la red neuronal artificial con los datos recopilados	36	Red neuronal entrenada y validada
6	Implementar el módulo de generación de pronósticos a partir de la lista de partidos futuros agendados	9	Programa que evalúa la red con la lista de partidos futuros y exporta los pronósticos en formato JSON
7	Implementar los servicios web REST para solicitar la información de la base de datos y los pronósticos generados	27	Implementación del <i>backend</i> del sistema web
8	Implementar la interfaz web que consuma los servicios del <i>backend</i>	27	Implementación del <i>frontend</i> del sistema web
9	Realizar pruebas del sistema integrado y documentarlo	18	Lista de pruebas realizadas y documentación del sistema
10	Redactar el reporte final del proyecto	18	Reporte final en PDF
Total		198	

8. Factibilidad técnica y operativa

8.1. Factibilidad técnica

El proyecto es viable ya que el alumno cuenta con conocimiento y experiencia laboral en tecnologías de desarrollo web, así como en el manejo de bases de datos relacionales como MariaDB. En cuanto al desarrollo de redes neuronales con PyTorch [12] el alumno llevó cursos específicos

de inteligencia artificial e inteligencia computacional, además de que el asesor también cuenta experiencia con el uso de esta tecnología.

Se cuenta con equipo de cómputo suficiente para entrenar la red neuronal. Una vez entrenada, su evaluación es poco costosa y es viable alojar el sistema web en un servidor con pocos recursos. En cuanto a software y licencias, no se contempla gasto alguno puesto que las tecnologías presupuestadas son de uso libre.

8.2. Factibilidad operativa

La factibilidad operativa del proyecto es alta, ya que se han realizado esfuerzos similares con éxito, aunque para contextos, ligas o torneos distintos. Mediante el uso de redes neuronales y datos históricos, se busca ofrecer un modelo predictivo confiable que beneficie a analistas, casas de apuestas y aficionados, mejorando la toma de decisiones y proporcionando una herramienta de valor agregado en el análisis de deportivo.

El sistema será fácil de usar, ya que se implementará en una plataforma web accesible e intuitiva, permitiendo su adopción sin requerir conocimientos técnicos. Además, su estructura modular y basada en tecnologías ampliamente utilizadas garantiza su adaptabilidad ante nueva información disponibles o cambios tecnológicos, asegurando así su sostenibilidad, facilidad de soporte y posibilidad de extensión a futuro.

9. Estimación de costos

La estimación de costos del proyecto se muestra en la Tabla 3.

Tabla 3. Estimación de costos del proyecto.

Rubro	Costo mensual	Costo trimestral
Laptop Dell Latitude 7390, 16 GB de RAM (reacondicionada) [13]	\$5089.82 MXN (pago único)	
Sueldo de desarrollador <i>full stack</i> (<i>frontend</i> y <i>backend</i>) [14]	\$26000 MXN	\$78000 MXN
Hospedaje web y base de datos [15]	\$704 MXN	\$2112 MXN
Internet [16]	\$ 499 MXN	\$ 1497 MXN
Electricidad [17]	\$ 143 MXN	\$ 429 MXN
Costo total		\$85015.82 MXN

El asesor se responsabiliza de guiar al alumnado y de que todos los recursos mencionados en la Factibilidad Técnica estarán disponibles para el alumnado, de modo que el proyecto de integración se pueda concluir en tiempo y forma.

Dr. Rodrigo Alexander Castro Campos
Asesor

Referencias

- [1] International Football Association Board, "Laws of the game 2024/25," 2025, accessed: 2025-08-06. [Online]. Available: <https://www.theifab.com/laws/latest/>
- [2] P. Isasi Viñuela and I. M. Galván León, *Redes de neuronas artificiales: un enfoque práctico*. Pearson Educación, 2004, consultado el 6 de agosto de 2025 via Google Libros.
- [3] S. Anfilets, S. Bezobrazov, V. Golovko, A. Sachenko, M. Komar, R. Dolny, V. Kasyanik, P. Bykovyy, E. Mikhno, and O. Osolinskyi, "Deep multilayer neural network for predicting the winner of football matches," *International Journal of Computing*, vol. 19, no. 1, mar 2020.
- [4] L. Luiz, G. Fialho, and J. Teixeria, "Is football unpredictable? predicting matches using neural networks," *Forecasting*, vol. 6, no. 4, dec 2024.
- [5] M. Rahman, "A deep learning framework for football match prediction," *Springer Nature Switzerland*, vol. 2, no. 165, jan 2020.
- [6] O. A. G. García, "Generación y resolución de niveles de un videojuego de plataformas usando redes neuronales," 2023, proyecto terminal, División de Ciencias Básicas e Ingeniería, Universidad Autónoma Metropolitana, Azcapotzalco, México.
- [7] B. Kalera, S. M. N. K. Naik, A. Mohiuddin, M. Moizuddin, and S. Mustafa, "Prediction of on-base percentage (obp) of baseball players using neural networks," *International Journal of Scientific Research in Engineering and Management*, vol. 8, no. 6, jun 2024.
- [8] L. Coello Blanco, L. Casas, O. L. Pérez González, and A. Caballero Mota, "Redes neuronales artificiales en la producción de tecnología educativa para la enseñanza de la diagonalización," *Academia y Virtualidad*, vol. 8, no. 2, 2015, universidad de Camagüey, Cuba. [Online]. Available: <https://www.researchgate.net/publication/283847824>
- [9] Football-Data.co.uk, "England Football Results Betting Odds." [Online]. Available: <https://www.football-data.co.uk/englandm.php>
- [10] MariaDB Foundation. MariaDB Server: the innovative open source database. [Online]. Available: <https://mariadb.org/>
- [11] PyMySQL Developers, "PyMySQL: Pure-Python MySQL Client," <https://pymysql.readthedocs.io>, 2025.
- [12] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and L. Antiga, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [13] Amazon. Dell Latitude 7390 13.3" FHD, Core i7-8650U 1.9GHz, 16GB de RAM, unidad de estado sólido de 256GB, Windows 11 Pro 64Bit, CAM . [Online]. Available: <https://www.amazon.com.mx/Dell-Latitude-7390-i7-8650U-reacondicionado/dp/B0BYKQT4JP>
- [14] Glassdoor. Salario medio para Full Stack Developer en México 2025. [Online]. Available: https://www.glassdoor.com.mx/Sueldos/full-stack-developer-sueldo-SRCH_KO0,20.htm
- [15] Hostinger. Servicio de Hosting para el sistema web y base de datos con Hostinger (KVM 8). [Online]. Available: <https://www.hostinger.com.mx/vps-hosting>

- [16] Telmex. Internet sin teléfono - Telmex Hogar. [Online]. Available: <https://telmex.com/web/hogar/internet-sin-telefono>
- [17] Comisión Federal de Electricidad. Tarifa DAC - Tarifas para Casa. [Online]. Available: <https://app.cfe.mx/Aplicaciones/CCFE/Tarifas/TarifasCRECasa/Tarifas/TarifaDAC.aspx>