

5C Online: Plataforma Digital para Conectar Personas y Jugar Fútbol 5

Bustos Sepúlveda Brayan Alexander, Delgado Triana Jeisson Javier, Herrera Macías Juan Sebastián y Pulido Córdoba Juan Pablo

Ingeniería de Sistemas, Fundación Universitaria Compensar

Proyecto Fin de Grado

Msc. Sandra Milena García Córdoba

Bogotá, mayo de 2023



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.

Contenido

Resumen 4

Abstract 4

Introducción 5

Antecedentes 6

Análisis de los Antecedentes 9

Pregunta de Investigación 10

Definición de Objetivos 10

Objetivo General 10

Objetivos Específicos 10

Cronograma de Actividades 11

Justificación 12

Metodología de Investigación 13

Marco Referencial 14

Marco Teórico 14

Marco Normativo 16

Marco Contextual 18

Desarrollo del proyecto 21

Levantamiento de requerimientos 21

Definición de Requerimientos 27

Requerimientos Funcionales 28

Requerimientos No Funcionales 32

Modelado y Diseño 36

Descripción General de la Arquitectura AWS 51

Mockups de la plataforma 56

Funcionamiento del Algoritmo 66

Filtrado Colaborativo 66

Factorización de Matrices 76

Codificación del Algoritmo en Amazon SageMaker 81

Resultados 89

Referencias 91

Resumen

En el presente trabajo de grado se realizó el diseño conceptual de una plataforma digital basada en servicios de la nube que integrará un algoritmo de Machine Learning el cual permitirá, a través de recomendaciones, la conexión social entre nuevos usuarios. El objetivo principal de este trabajo fue analizar un modelo que permitiera a la comunidad del deporte, específicamente a los jugadores de fútbol 5 conocer y conectar con nuevas personas con sus mismos gustos de una manera fácil y efectiva, para ello se realizó el análisis de un algoritmo de recomendación basado en el filtrado colaborativo para su posterior adaptación al caso práctico de la plataforma digital, así como el diseño conceptual completo, el cual abarcó la arquitectura de la plataforma que estará basada en servicios de la nube de AWS, de igual forma se realizó la diagramación correspondiente que permitió ejemplificar el funcionamiento de la plataforma. El resultado final de este trabajo es una combinación del diseño de la plataforma y la comprensión del algoritmo de recomendación seleccionado, permitiendo así, sentar las bases para su posterior desarrollo e implementación.

Palabras Clave: Plataforma digital, Amazon Web Services, Aprendizaje Automático, Algoritmo de recomendación, Filtrado Colaborativo

Abstract

In the present thesis, the conceptual design of a cloud-based digital platform was carried out, which will integrate a Machine Learning algorithm enabling social connections among new users through personalized recommendations. The main objective of this work was to analyze a model that would allow the sports community, specifically futsal players, to meet and connect with new people who share their interests easily and effectively. To achieve this, an analysis of a

collaborative filtering-based recommendation algorithm was conducted, followed by its adaptation to the practical case of the digital platform. Additionally, a comprehensive conceptual design was developed, encompassing the architecture of the platform, which will be built using AWS cloud services. Diagrams were also created to illustrate the functionality of the platform. The outcome of this work is a combination of the platform design and a thorough understanding of the selected recommendation algorithm, laying the groundwork for its subsequent development and implementation.

Key Words: Digital Platform, Amazon Web Services, Machine Learning, Recommendation Algorithm, Collaborative Filtering.

Introducción

Con el auge tecnológico, las relaciones sociales ahora se manejan de una manera diferente a como se solían manejar hace no más de una década, es tan así que, gracias a la tecnología, se han trastornado incluso el desarrollo de actividades físico-deportivas como el fútbol, considerándose esta como una de las que mayores vínculos sociales crean.

El hecho de jugar con un equipo conformado por personas conocidas suele fortalecer los vinculo sociales y ayuda al interior del grupo a mejorar habilidades y destrezas en cada uno de sus integrantes, de tal manera que jugar siempre con un mismo circulo social sería el escenario ideal para la mayoría de las personas que se apasiona por este deporte y lo practica de manera recreativa, sin embargo, para muchos, la realidad es distinta, viéndose afectados por varios desafíos como lo pueden ser la falta de jugadores, ya que no siempre es posible encontrar suficientes jugadores disponibles para completar un equipo de fútbol 5.

Esto puede ser especialmente difícil si se necesitan jugadores específicos, como un portero; además, las lesiones son una posibilidad constante, por lo que si un jugador se lesiona, puede ser difícil encontrar un reemplazo adecuado en el corto plazo; finalmente, el factor que se considera más influyente y que más imposibilita la coordinación de partidos es la disponibilidad horaria de los posibles integrantes de un equipo, ya que, puede ser difícil coordinar los horarios de los jugadores, especialmente si tienen trabajos u otros compromisos que rivalizan con el tiempo de juego.

Ante estos desafíos, surge la necesidad de buscar una solución que facilite y promueva la formación de equipos fútbol 5 de una manera eficiente y conveniente. Es en este contexto que se plantea el presente trabajo de grado, que tiene como objetivo el diseño de una plataforma basada en servicios de nube que permita a aficionados del fútbol 5, conectar y conocer con nuevas personas para la conformación de los equipos cuando se requiera.

Antecedentes

Haciendo una revisión preliminar relacionada con el tema de investigación, se obtuvieron los siguientes antecedentes de proyectos similares, en los que se plantea como solución el desarrollo de plataformas que permitan diversas gestiones relacionadas con el deporte y la conexión de personas.

Fubles.com. Es un primer referente; se trata de " una plataforma de (Social Sport Sharing) con una de las comunidades deportivas más atractivas de Europa". Fubles es una plataforma digital que permite a los usuarios organizar y encontrar partidos de fútbol de manera fácil y rápida. (Fubles.com, s.f.)

Fubles fue desarrollado por un equipo de emprendedores italianos liderado por Mirko Trasciatti quien, junto con su equipo, creó la plataforma Fubles como una solución para un problema común en Italia "Me cansé de tener que andar detrás de mis amigos, en vez de andar detrás de un balón que era lo que realmente quería". (Trasciatti, 2011). Desde entonces, la plataforma ha evolucionado y se ha expandido a otros países, convirtiéndose en una herramienta popular para organizar partidos de fútbol en línea. (Vanguardia, 2011)

S***istema de gestión de eventos deportivos basado en redes sociales, aplicaciones web y móviles***. Este segundo hallazgo tiene como objetivo optimizar la forma en cómo se organizan los eventos deportivos amateurs en Perú basándose en el concepto de una red social, interactuando con las redes ya existentes. De esta manera, los autores plantearon como objetivo específico el desarrollo de un sistema de gestión de eventos deportivos aplicando los conceptos de ingeniería que tiene como fin resolver un problema práctico de la organización en sí haciendo uso de tecnologías actuales. En cuanto a la metodología de desarrollo (MDS) utilizada para este proyecto, se optó por usar Scrum, ya que se adaptaba con las características de este.

Los autores del proyecto tuvieron como público objetivo a los deportistas amateurs y profesionales, de igual forma, a los administradores de los espacios deportivos. “Ambos compartirán información a través de una herramienta social deportiva utilizando sus distintos servicios, siendo su principal valor el organizar de forma efectiva todo encuentro deportivo y mantener una interacción constante entre deportistas y administradores” (Essenwanger, Humberto, Rojas, & Moisés, 2014, p. 7).

Aplicación móvil gol app organizadora de partidos de fútbol. Un Tercer trabajo corresponde a León, Macpherson, Núñez y Olea (2018), quienes realizaron propusieron el desarrollo de una aplicación para organizar partidos de fútbol. En este trabajo la metodología de

investigación realizada fue de tipo cualitativa mediante el método exploratorio, los cuales usaron como herramienta entrevistas a profundidad. El público objetivo fueron hombres de Lima Metropolitana (Perú) que residen en las Zonas 6 y 7, del NSE A y B, entre 18 y 55 años.

Desarrollo de la app android: Footman. Un Cuarto trabajo corresponde a Pérez, Muñoz (2020), quien propuso el desarrollo de una aplicación para el sistema operativo Android. En este trabajo, el objetivo es simplificar la gestión de torneos de fútbol, especialmente aquellos que son organizados por aficionados o de carácter no profesional a través de una aplicación móvil, ya que los métodos tradicionales resultan ser no tan eficientes. De esta manera, la aplicación ayudará a los organizadores de los torneos a gestionarlos de manera más efectiva.

Aplicación para la suscripción a eventos deportivos. Un quinto trabajo corresponde a Ballara y Martin (2020) quienes proponen el desarrollo de una aplicación para la suscripción a eventos deportivos donde el objetivo principal es el "Diseño e implementación de una aplicación móvil que permita a los usuarios practicar diferentes deportes (fútbol, vóley, pádel, básquet, etc.) aún si no disponen del número suficiente de integrantes para conformar un equipo completo" (Ballara & Martín, 2020)

De tal manera que lo que se busca poder unir a las personas para que puedan crear eventos deportivos y de igual forma se puedan unir a los que ya están creados. La metodología usada fue usar un proceso en espiral, el cual consiste, en el que fueron desarrollando el proyecto paso a paso, aumentando el nivel de complejidad por el paso de cada ciclo (Planeación, Análisis de riesgo, Implementación y Evaluación). A su vez, se detallan las diferentes etapas del ciclo, desde la definición de las necesidades y tecnologías, hasta la implementación del código en sí y las pruebas de comprobación. Se utilizan también herramientas como UML, diagramas y

Mockup´s para guiar el desarrollo del proyecto. Al final, se realizan las pruebas o testeos de la aplicación por usuarios externos a esta, para comprobar su funcionamiento a nivel general.

Análisis de los Antecedentes

Los hallazgos obtenidos en la revisión de antecedentes muestran la importancia de la tecnología en el ámbito deportivo y la necesidad de herramientas digitales que permitan a usuarios conectar entre sí como una solución a las limitaciones que presentan los métodos tradicionales de formación de equipos deportivos. Se evidencia que, mediante el uso de tecnología, los usuarios pueden encontrar compañeros de equipo que cumplan con criterios específicos de búsqueda y establecer una red de contactos deportivos.

Adicional a la notable importancia que tiene la tecnología en la modernidad, también hay que tener en cuenta que la mayoría de los usuarios de plataformas buscan experiencias agiles y precisas, por lo que en la actualidad se hace usó de herramientas como los algoritmos de Machine Learning para lograr así, mejorar la experiencia del usuario final en las plataformas, esto se consigue ya que la mayoría de estos algoritmos tienen en cuenta la forma en cómo interactúan los usuarios al interior de estas, es decir, con qué elementos interactúa con mayor frecuencia, que horarios son los preferidos por parte de un usuario para hacer uso de la plataforma, sus gustos, entre muchos otros más factores, para al final, ofrecerle a los usuarios una experiencia más personal en lo posible a través de predicciones que tienen como objetivo recomendarles contenido el cual tiene alta probabilidad de ser de agrado para ellos.

Pregunta de Investigación

Tomando como punto de partida los resultados del análisis de los antecedentes, se formula la siguiente pregunta de investigación.

¿Cómo puede el diseño y modelado de una plataforma digital que haga uso de un algoritmo de machine learning mejorar la experiencia de usuario y permitirle así conectar con nuevas personas basándose sus en intereses y preferencias compartidas de una manera eficaz?

Definición de Objetivos

Objetivo General

Diseñar conceptualmente una plataforma de conexión social, que utilice un algoritmo de Machine Learning para recomendar a los usuarios potenciales compañeros de equipo basándose en sus perfiles deportivos.

Objetivos Específicos

- Modelar el funcionamiento de la plataforma digital con el fin de comprender cómo operará esta en términos de componentes, interacciones y procesos.
- Analizar el algoritmo de recomendación basado en filtrado colaborativo con el fin de evaluar su eficacia y precisión para el contexto de la plataforma.
- Codificar el algoritmo de recomendación en el servicio de AWS SageMaker con el fin de sentar las bases para su futura implementación.

Cronograma de Actividades

En la Ilustración 1 se ilustra el cronograma de las actividades a desarrollar para dar cumplimiento de los objetivos estipulados en el presente proyecto.

Ilustración 1

Cronograma de actividades

Nombre	Comienzo	Fin	Duración	% completado
Plataforma digital 5C online	mar 14/02/23	dom 14/05/23	90 días	100%
Fase preliminar	mar 14/02/23	mié 15/02/23	2 días	100%
Elaboración cronograma	mar 14/02/23	mié 15/02/23	2 días	100%
Fase 1: Planteamiento del problema	jue 16/02/23	mar 14/03/23	27 días	100%
Análisis del problema	jue 16/02/23	mié 22/02/23	7 días	100%
Definición del problema	jue 23/02/23	mar 28/02/23	6 días	100%
Definición alcance proyecto	mié 1/03/23	vie 3/03/23	3 días	100%
Formulación problema investigación	sáb 4/03/23	dom 5/03/23	2 días	100%
Definición de objetivos	lun 6/03/23	lun 6/03/23	1 día	100%
Justificación del proyecto	mar 7/03/23	vie 10/03/23	4 días	100%
Investigación antecedentes	sáb 11/03/23	lun 13/03/23	3 días	100%
Documentació de antecedentes	mar 14/03/23	mar 14/03/23	1 día	100%
Fase 2: Marco referencial	mié 15/03/23	sáb 25/03/23	11 días	100%
Construcción marco teórico	mié 15/03/23	sáb 18/03/23	4 días	100%
Construcción marco normativo	dom 19/03/23	dom 19/03/23	1 día	100%
Construcción marco contextual	lun 20/03/23	mar 21/03/23	2 días	100%
Definición metodología de investigación	mié 22/03/23	jue 23/03/23	2 días	100%
Estimación de costos	vie 24/03/23	sáb 25/03/23	2 días	100%
Fase 3: Diseño del proyecto	dom 26/03/23	dom 14/05/23	50 días	100%
Levantamiento de requerimientos	dom 26/03/23	mar 4/04/23	10 días	100%
Elaboración de encuesta	dom 26/03/23	lun 27/03/23	2 días	100%
Distribución de la encuesta	mar 28/03/23	sáb 1/04/23	5 días	100%
Análisis de la información recolectada	dom 2/04/23	mar 4/04/23	3 días	100%
Definición de requerimientos	mié 5/04/23	vie 7/04/23	3 días	100%
Requerimientos funcionales	mié 5/04/23	jue 6/04/23	2 días	100%
Requerimientos no funcionales	vie 7/04/23	vie 7/04/23	1 día	100%
Diagramación UML	sáb 8/04/23	vie 21/04/23	14 días	100%
Elaboración Modelo Entidad Relación	sáb 8/04/23	dom 9/04/23	2 días	100%
Elaboración Modelo Relacional	lun 10/04/23	lun 10/04/23	1 día	100%
Elaboración diccionario de datos	mar 11/04/23	mar 11/04/23	1 día	100%
Elaboración diagrama de clases	mié 12/04/23	jue 13/04/23	2 días	100%
Elaboración casos de uso	vie 14/04/23	sáb 15/04/23	2 días	100%
Elaboración diagramas de flujo	dom 16/04/23	mar 18/04/23	3 días	100%
Elaboración Arquitectura AWS	mié 19/04/23	vie 21/04/23	3 días	100%
Elaboración Mockups	sáb 22/04/23	sáb 29/04/23	8 días	100%
Codificacion algoritmo en AWS SageMaker	dom 30/04/23	dom 14/05/23	15 días	100%

Justificación

Los jugadores tienen problemas para buscar partidos o completar un equipo de fútbol 5, esto conlleva a la búsqueda de métodos tradicionales e ineficientes para lograr estos objetivos, tales como, la voz a voz, llegando incluso a no poder completar los equipos y terminar así, jugando un partido con el número incompleto de jugadores. Por lo tanto, se propone una plataforma digital que solucione estos inconvenientes, recomendando a los usuarios otros perfiles que se adapten a los intereses de los jugadores.

A continuación, se enlistan los beneficios que supondrían el desarrollo de dicha plataforma para los usuarios finales.

Ahorro de tiempo y esfuerzo: La plataforma digital permitirá a los usuarios organizar partidos de manera más eficiente, eliminando la necesidad de coordinar con cada jugador de manera individual por medio de métodos tradicionales. Esto conllevará a un ahorrar tiempo y esfuerzo de los jugadores que juegan regularmente, ya que, en lugar de tener que buscar manualmente a otros jugadores, la plataforma presentaría automáticamente una lista de recomendaciones basadas en preferencias propias de cada usuario.

Facilitar la organización de partidos: Una plataforma digital bien diseñada permite una facilidad y agilidad mayor a la hora de organizar partidos con diferentes jugadores, permitiendo a los jugadores establecer fechas, horarios y lugares adecuados para los partidos.

Mejora la calidad de los partidos: La plataforma digital también mejorará la calidad de los partidos, ya que los jugadores podrán tener acceso a estadísticas, clasificaciones y habilidades de otros jugadores, lo que les permitirá organizar partidos más equilibrados aumentando la competitividad de los partidos, haciéndolos más divertidos y satisfactorios.

Mayor variedad de oponentes: La plataforma presentará a los usuarios otros jugadores de toda la ciudad, los cuales no conocían con anterioridad, lo que les permitirá jugar tanto juntos, como en contra de oponentes diferentes y así, ampliar la red de contactos deportivos. Esto podrá mejorar la experiencia del juego y ayudar a conocer a más personas con intereses similares.

Mejora continua de las recomendaciones: La plataforma utilizará la retroalimentación que se recibe después de cada partido para mejorar continuamente las recomendaciones de usuarios y hacer que sean más precisas y relevantes para cada uno. De esta manera, la plataforma podrá adaptarse a las necesidades y preferencias cambiantes de los usuarios a medida que evolucionan los juegos y los intereses.

Metodología de Investigación

La metodología de investigación empleada en el presenta proyecto de grado es una metodología mixta, ya que aborda enfoques cuantitativos y enfoques de diseño, lo que permitirá abordar de una manera más profunda el problema de investigación, para la recolección cuantitativa de los datos se hará uso de herramientas como lo es una encuesta estructurada que permitirá conocer con un enfoque estadístico la percepción de la muestra representativa, que en este caso corresponde a jugadores aficionados de fútbol 5 para así identificar luego de un análisis los patrones, tendencias y relaciones existentes entre las diferentes variables.

A su vez, la metodología de diseño posibilitará la creación de un modelo conceptual de la plataforma, dicho proceso se dividirá en varias etapas, incluyendo la exploración y definición del problema, la generación de ideas, recopilación de requisitos que va muy alineado con el enfoque cuantitativo de la investigación, para terminar así, con un proceso de modelamiento; este enfoque mixto permitirá abordar el problema de manera sistemática para definir una solución eficaz y práctica.

Marco Referencial

Marco Teórico

El fútbol es un deporte popular que se juega en todo el mundo, y su popularidad se debe en gran parte a su capacidad para unir a las personas y construir relaciones interpersonales. Los estudios muestran que el fútbol no solo es beneficioso para la salud física, sino que también tiene un impacto positivo en la salud mental y en la construcción de relaciones interpersonales "Lo que sucede cuando jugamos fútbol no solo tiene efecto en el campo de juego, sino que afecta lo que sucede fuera del campo de juego, en la vida social de los jugadores (…)" (Cáceres, 2008).

En primer lugar, el fútbol es un deporte que puede ser jugado por personas de todas las edades y habilidades, lo que significa que puede unir a personas de diferentes orígenes y culturas. A través del fútbol, las personas pueden compartir experiencias y crear vínculos emocionales a través del juego en equipo, la competencia amistosa y el apoyo mutuo. Cáceres (2008), afirmo lo siguiente:

Aquí se pone en operación un nivel superior de configuración cognitiva, la que construye la alteridad, la base relacional de la comunicación, la interacción con los compañeros de equipo y los oponentes (…) El fútbol es una matriz constructiva de la interacción humana, promueve las figuras cognitivas de la colaboración y la competencia al mismo tiempo. (p. 15)

El fútbol puede servir como una herramienta para la integración social y la superación de barreras culturales. En muchos países, el fútbol es un pasatiempo popular que une a las personas más allá de las diferencias culturales y sociales. Los equipos de fútbol pueden servir como un lugar donde las personas pueden reunirse y compartir su amor por el deporte, lo que puede fomentar la comprensión y la tolerancia entre diferentes grupos de personas.

En palabras de Cáceres (2008):

(…) El fútbol se configura como una matriz de interacción y, por tanto, se vincula con formas culturales y sociales como la dinámica familiar, las relaciones entre vecinos y amigos, e incluso las relaciones de trabajo y las políticas. Esto ubica al juego más allá del entretenimiento o de la sólo actividad lúdica o de salud física. El fútbol puede llegar a correlacionarse con las figuras de la democracia (…) (p. 15-16).

Y, en segundo lugar, el fútbol fomenta valores como el trabajo en equipo, la comunicación, la resolución de conflictos y el respeto mutuo. López (2009) Destacó que estos valores son esenciales para la construcción de relaciones interpersonales saludables y duraderas (p. 1-8).

Por otro lado, compartir intereses y aficiones comunes, puede ser una excelente manera de iniciar una conversación y conocerse mejor. A menudo, cuando se comparte intereses similares con alguien, es normal sentirse más cómodo hablando sobre ellos y de esta manera, poder tener conversaciones más profundas y significativas, lo que puede traducirse en relaciones más sólidas fuera del campo de juego.

Afortunadamente, la sociedad actual está experimentado un crecimiento tecnológico que ha transformado la forma en que las personas comparten sus actividades y pasatiempos. En la actualidad, gracias a las redes sociales y otras plataformas digitales, es posible conectarse de manera más sencilla con personas que comparten los mismos intereses.

Se podría decir que dicho avance tecnológico ha impactado de manera positiva en la forma en que las personas pueden relacionarse, pues ahora, la mayoría de las plataformas

digitales, implementan algoritmos de recomendación o coincidencia que pueden ayudar a los usuarios a encontrar personas que tengan intereses similares.

Según Gómez y otros (2019), estos algoritmos utilizan datos sobre las preferencias pasadas de los usuarios para predecir en qué elementos podrían estar interesados en el futuro, lo que aumenta la probabilidad de conexiones significativas entre los usuarios.

Es por esto por lo que el fútbol es un deporte que va más allá de la actividad física y la competencia deportiva. Además de sus beneficios para la salud física y mental, el fútbol tiene la capacidad de construir y fomentar valores significativos en las relaciones interpersonales. En la actualidad, el deporte, las plataformas digitales y los algoritmos de recomendación de manera combinada, pueden ser herramientas útiles para conectar a personas con intereses comunes y potenciar la capacidad del deporte para facilitar la creación de relaciones interpersonales positivas.

Marco Normativo

Propiedad Intelectual. Se entiende por propiedad intelectual la disciplina jurídica que busca la protección de elementos inmateriales de naturaleza intelectual y de contenido creativo. Esta se divide en tres categorías, Propiedad Industrial, Derechos de Autor y Obtención de variedades Vegetales; sin embargo, para el objeto del presente proyecto, la ley aplicable corresponde a la propiedad Industrial, la cual encierra todo lo relacionado a las patentes de invención, marcas, diseños industriales y denominaciones de origen.

La entidad colombiana que se encuentran a cargo de la administración de los derechos de propiedad intelectual en materia de la Propiedad Industrial y que regula diferentes componentes

relacionados a la actividad de innovación es La Super Intendencia de Industria y Comercio (SIC). (Orjuela, Villar, & Herrera, 2017)

Normativa de protección de datos personales. Ley 1581 de 2012 tiene como objeto el desarrollo al derecho constitucional con el que cuenta la población colombiana a conocer, actualizar y rectificar la información que se haya registrado sobre ellas en bases de datos o archivos por parte de entidades públicas o privadas.

De manera tal que, todo tratamiento de datos personales debe realizarse bajo estándares legítimos con previo consentimiento del titular, bien sea por escrito, de manera oral o "mediante conductas inequívocas del titular de la información las cuales permitan concluir a los responsables que se otorgó la autorización" (Super Intendencia de Industria y Comercio (SIC), 2022).

Existen una serie de responsabilidades para el tratamiento de información que se deben cumplir por parte de aquellos que efectúen el rol de encargados de la información.

Normativa de comercio electrónico. Ley 527 de 1999, la cual establece y reglamenta el uso de mensajes de datos, firmas digitales y las bases para la construcción de un marco regulatorio para el comercio electrónico.

Entiéndase como mensajes de datos toda información tratada por medios electrónico, que contempla labores como el envío, recepción, transmisión y almacenamiento. Al tratarse de medios electrónicos abarca las comunicaciones que por su naturaleza se relacionan con la electrónica.

El comercio electrónico es toda relación comercial, contractual o no, donde se logre evidenciar algún tipo de suministro o intercambio de bienes y/o que esté representado por mensajes de datos.

Normativa de consumo. Ley 1480 de 2011, Estatuto del consumidor.

Se considera la normativa de consumo ya que todo aquel usuario que hace uso de una plataforma tecnológica tiene derechos como consumidor, y, por tanto, cualquiera que desarrolle y publique una plataforma digitalizada debe considerar y cumplir con una serie de directrices, entre las principales están:

- Brindar información clara que permita elegir los servicios que se desean adquirir.
- En el caso de una plataforma digital de acceso a todo el público, se debe contar con un control parental que se ofrezca contenidos adecuados a los infantes y adolescentes.
- Proteger contra publicidad engañosa.
- Derecho a la reclamación.

Marco Contextual

Para dar contexto del impacto cultural que genera el desarrollo de actividades deportivas, en este caso, futbol, en espacios públicos y privados en la ciudad de Bogotá, Colombia, se tienen en cuenta las siguientes normas y leyes:

Congreso de Colombia. (1995). Ley 181 de 1995. “Por la cual se establece el régimen especial para la promoción y fomento del deporte, se crea el Sistema Nacional del Deporte y se dictan otras disposiciones. Artículo 1”. El cual indica que se debe fomentar el deporte en

cualquiera de sus ramas, además de crear diferentes entidades y órganos para que regulen estas prácticas, además de velar por la salud de los deportistas y asignación de recursos para la ejecución de los deportes. Esta ley acoge a cualquier colombiano que quiera ejercer su derecho a practicar cualquier deporte, no necesariamente tiene que practicarse el deporte de manera profesional, de hecho, la ley que se establece como derecho fundamental, e indica que …" promueve la práctica en cualquier forma, ya sea de manera recreativa, escolar, comunitaria, amateur o profesional" Congreso de Colombia. (1995). Ley 181 de 1995. Artículo 2. Por ende, es que el gobierno y privados tienen la libertad de promover todo tipo de actividades relacionadas con el deporte, en este caso, el deporte objeto de estudio, el futbol, cuenta como una de estas actividades. Los privados suelen ofrecer facilidades a las personas interesadas en practicar este deporte, facilitando canchas sintéticas que se pueden alquilar, junto con otros servicios que pueden o no ofrecer, dependiendo del establecimiento o lugar del local.

Para los Bogotanos que son apasionados por el deporte, de manera recurrente, suelen pactar cita en dichos lugares con amigos, o amigos de amigos para generar un encuentro amistoso o competitivo, este último acogido por la Ley 1356 de 2009, "Artículo 1. Objeto y principios. La presente ley tiene por objeto prevenir y contrarrestar la violencia y el racismo en los espectáculos deportivos. Tiene como principio fundamental garantizar la seguridad y la convivencia pacífica en los eventos deportivos y establecer un marco regulatorio para la organización y el desarrollo de estos eventos en el país.". Los establecimientos que tienen como función promover este tipo de eventos deportivos se ven en la obligación de promover ambientes sanos, restringiendo medidas que pudieran afectar la convivencia de otros jugadores o competiciones dentro del establecimiento.

La popularidad de dichos establecimientos ha generado que estos no tengan que publicitarse por sí mismos, sino que los mismos jugadores o usuarios lo hagan por medio de la voz a voz, por ende, se podría decir que es una actividad en la cual el lugar siempre va a estar disponible, dependiente claramente de su reserva y demás variables propias de fecha y horario. Lo que no siempre va a estar disponible es la posibilidad de jugar con personas desconocidas, con quien se pueda medir el nivel de juego y sacar a flote la competitividad individual o grupal, porque no hay espacios para generar estos posibles encuentros, debido a que los establecimientos están prestos para arrendar la cancha para juegos individuales o torneos pagos, no hay opción de pactar con desconocidos para jugar posibles futuros encuentros.

Actualmente existen aplicaciones, descritas en los antecedentes de este documento, que permiten crear el espacio para que esta opción sea viable, pero no son aplicaciones que satisfagan por completo las necesidades de los usuarios, no son igual de populares como lo son las propias canchas que se publicitan así mismas. El desarrollo de estas aplicaciones dentro de Colombia se acoge bajo las siguientes leyes:

Congreso de Colombia. La ley 1341 de 2009, la cual…"establece las políticas públicas para el desarrollo de la tecnología de la información y las comunicaciones (TIC) en Colombia. Esta ley tiene como objetivo promover el acceso universal a las TIC, fomentar la innovación y la competitividad del sector y garantizar la protección de los derechos de los usuarios."

Congreso de Colombia. Ley 1581 de 2012, la cual indica…"regular el tratamiento de los datos personales de los usuarios por parte de entidades públicas y privadas. Esta ley establece que el tratamiento de los datos personales debe hacerse con respeto a los derechos de los titulares de la información, tales como el derecho a la privacidad, la intimidad, la autodeterminación informativa y la libertad."

Desarrollo del proyecto

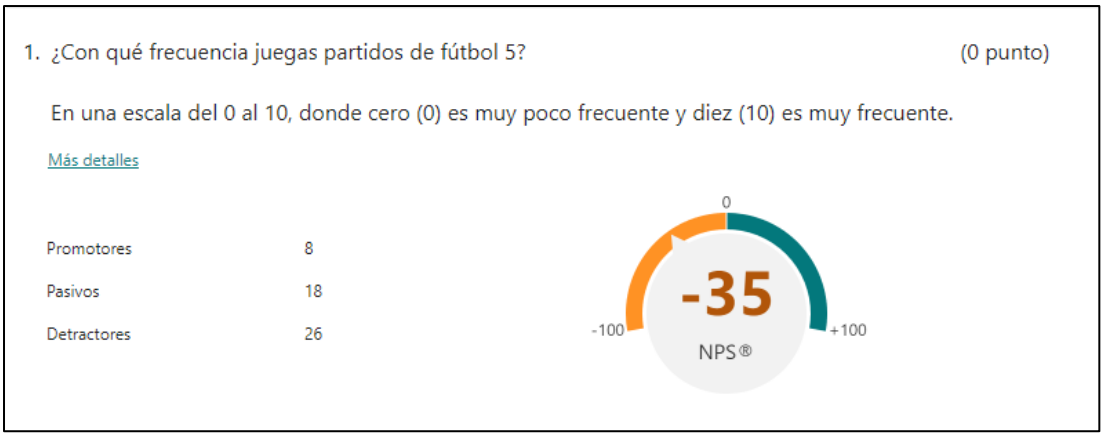
Levantamiento de requerimientos

Para poder definir los requerimientos de la plataforma, se realizó una encuesta estructurada en la que se plantearon preguntas que son de relevancia para recabar las opiniones de la muestra ante la existencia de una posible plataforma que contribuya a conocer nuevas personas y de esa manera a la conformación de equipos de fútbol 5.

La primera de las preguntas de la encuesta establece la frecuencia con la que la población de la muestra juega partidos de fútbol 5, esto es de vital importancia ya que permitirá evidenciar si el uso de una plataforma digital ayudaría a aumentar el juego en la población muestreada, como resultado se obtuvo que del total de la muestra la mayoría no juegan con frecuencia partidos de fútbol 5 como se evidencia en la Ilustración 2.

Ilustración 2

Estructura de la Encuesta 1



La siguiente pregunta, permite evidenciar el punto de vista de cada uno de los integrantes de la población muestreada, en lo referente a la dificultad que tienen a la hora de encontrar integrantes para completar sus equipos de fútbol 5, de igual forma los resultados obtenidos de esta

pregunta proporcionan datos relevantes para la toma de decisiones y que ilustran la necesidad de la existencia de una plataforma digital que ayude a los usuarios a conocer a nuevas personas que estén dispuestas a jugar partidos de fútbol 5. Los resultados muestran que para un 73% de la muestra no es tan sencillo encontrar compañeros para completar sus equipos de juego, como se evidencia en la Ilustración 3.

Ilustración 3

Estructura de la Encuesta 2



La siguiente pregunta de la encuesta es de tipo abierta y el objetivo de esta es identificar los patrones más utilizados por cada integrante de la muestra para conseguir miembros para sus equipos de juego, de forma tal que se pueda observar la ineficiencia de los métodos tradicionales, los resultados de esta pregunta se evidencian en la Ilustración 3.

Ilustración 4

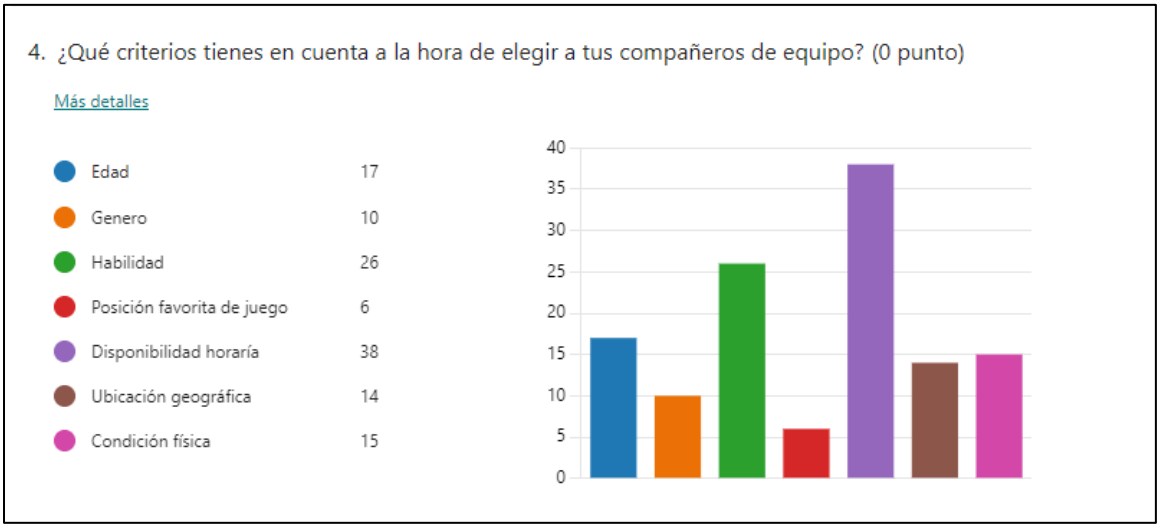
Estructura de la Encuesta 4



La cuarta pregunta de la encuesta tiene como objetivo definir los principales criterios que tienen en cuenta los integrantes de la muestra a la hora de elegir a los otros miembros que conformarán sus equipos de juego, este dato es de vital importancia ya que de aquí dependerán los criterios que el algoritmo de recomendación tendrá en cuenta a la hora de realizar sus predicciones y posteriores recomendaciones, los resultados de esta pregunta se evidencian en la lustración 4.

Ilustración 5

Estructura de la Encuesta 5



Las preguntas de la 5 a la 8 tienen como objetivo conocer la posición de la población muestreada frente al desarrollo de una plataforma digital y la favorabilidad que esta tendría ante un posible desarrollo y puesta en producción, los resultados de estas preguntas se evidencian en las Ilustraciones 6, 7, 8 y 9.

Ilustración 6

Estructura de la Encuesta 5

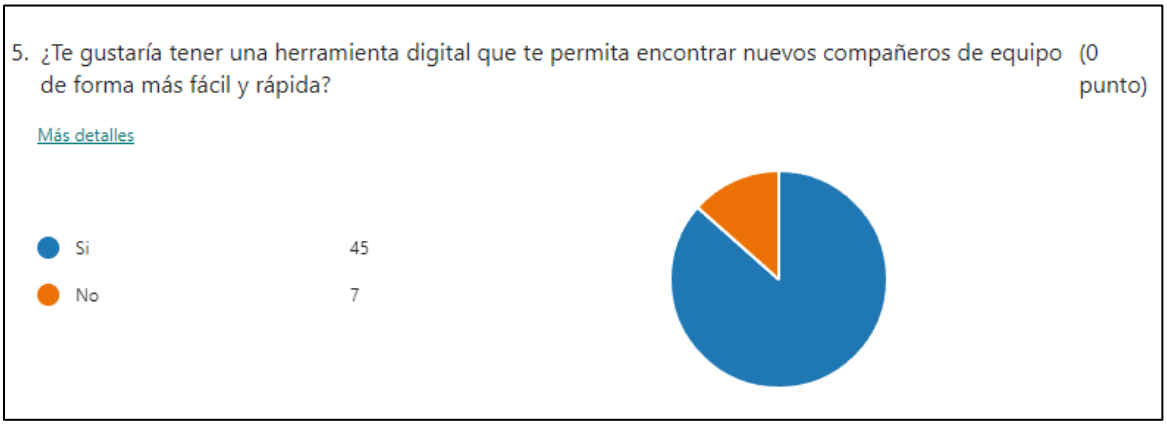


Ilustración 7

Estructura de la Encuesta 6

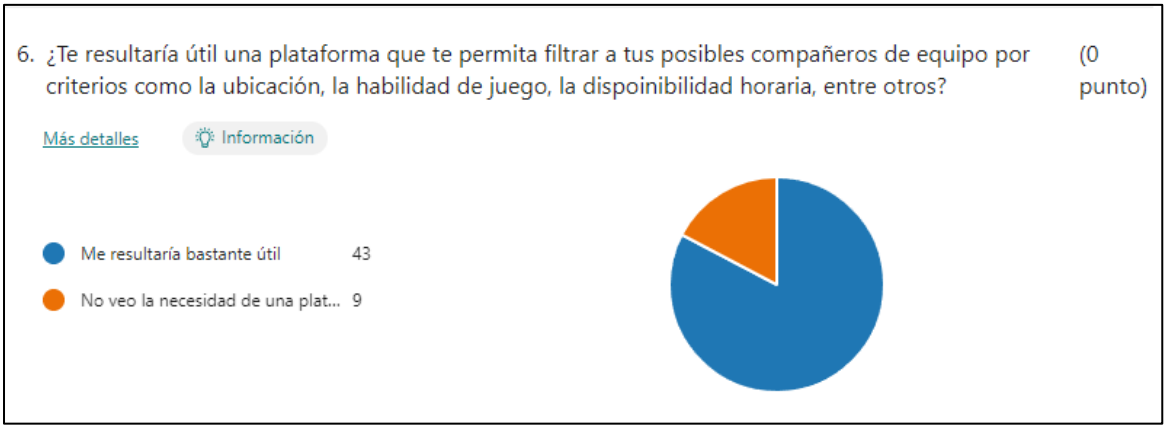


Ilustración 8

Estructura de la Encuesta 7

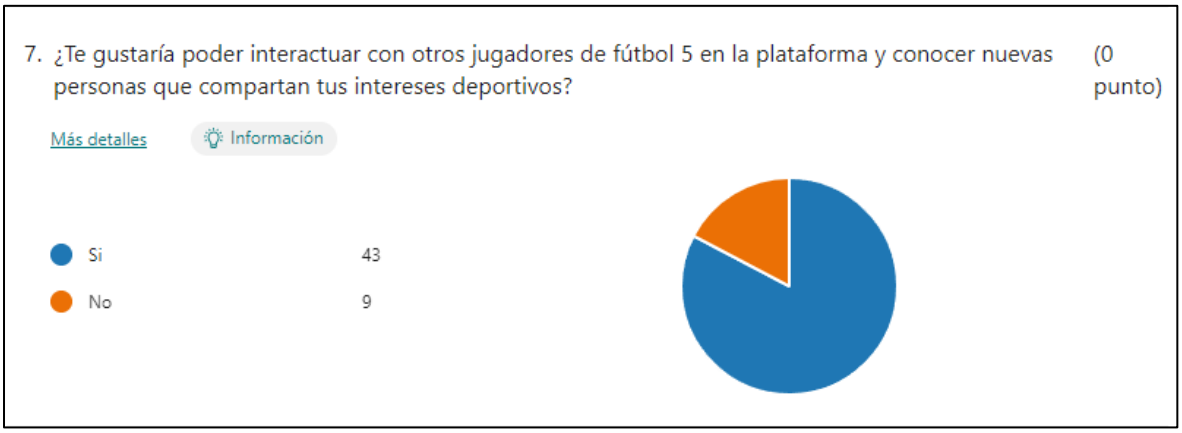
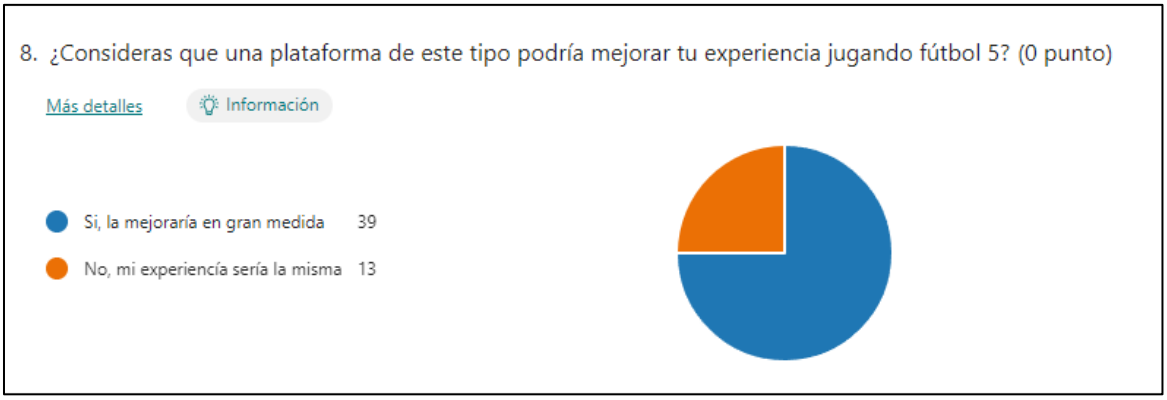


Ilustración 9

Estructura de la Encuesta 8



La novena pregunta es de tipo abierta y tiene como objetivo conocer la voz de cada uno de los integrantes de la población muestreada con el fin de tener en cuenta dichas solicitudes para una posible implementación en la plataforma digital, los resultados de esta pregunta se evidencian en la Ilustración 10.

Ilustración 10

Estructura de la Encuesta 9



Definición de Requerimientos

Luego de analizar los resultados obtenidos de la encuesta, se hace evidente que el desarrollo de una plataforma digital que permita conocer y conectar con personas nuevas y que a su vez integre un algoritmo de recomendación para que la experiencia del usuario sea más agradable y precisa tendría una gran favorabilidad en la comunidad de jugadores de fútbol aficionado, lo que puede traducirse como una justificación empírica para el desarrollo de dicha plataforma. Por lo que, para dar continuidad al desarrollo del proyecto, se establecieron los siguientes requerimientos funcionales y no funcionales de la plataforma que se evidencian en las Tablas de la 1 a la .

Requerimientos Funcionales

Tabla 1

Requerimientos Funcionales RF1

Identificador del requerimiento	RF1
Nombre del requerimiento	Inicio de Sesión y Registro
Descripción del requerimiento	Funcionalidad Log In / Sing In donde se obtenga datos como: Tipo de documento, Número de documento, Nombre completo, Edad, Genero, Correo electrónico, Contraseña, Ubicación, entre otros.
Prioridad del requerimiento	Alta

Tabla 2

Requerimientos Funcionales RF2

Identificador del requerimiento	RF2.
Nombre del requerimiento	Creación de Perfiles de Usuario con Información Personal y Características Deportivas.
Descripción del requerimiento	La plataforma debe contar con perfiles de usuarios para cada uno de estos en donde se pueda completar la información necesaria para el uso de la plataforma, Foto de perfil, Preferencia de juego, Disponibilidad (Ubicación & Franja Horaria), Habilidades (Tiro, Pases, Regates, Defensa y Físico).
Prioridad del requerimiento	Alta

Tabla 3

Requerimientos Funcionales RF3

Identificador del requerimiento	RF3.
Nombre del requerimiento	Sistema de Filtrado.
Descripción del requerimiento	La plataforma debe contar con un sistema de filtrado que permita encontrar jugadores en función de una serie de parámetros relacionados con los perfiles de usuario
Prioridad del requerimiento	Alta

Tabla 4

Requerimientos Funcionales RF4

Identificador del requerimiento	RF4.
Nombre del requerimiento	Algoritmo de Recomendación.
Descripción del requerimiento	La plataforma debe contar con un sistema de recomendación de jugadores que ofrezca sugerencias personalizadas a los usuarios en función de las preferencia y comportamientos previos de cada uno de estos.
Prioridad del requerimiento	Alta

Tabla 5

Requerimientos Funcionales RF5

Identificador del requerimiento	RF5.
Nombre del requerimiento	Chat Interno.
Descripción del requerimiento	La plataforma debe contar con un sistema de comunicación interna entre usuarios con el fin de que puedan gestionar todo lo referente a los encuentros de fútbol.
Prioridad del requerimiento	Alta

Tabla 6

Requerimientos Funcionales RF6

Identificador del requerimiento	RF6.
Nombre del requerimiento	Sistema de Calificación.
Descripción del requerimiento	El sistema debe contar con un sistema de calificación y comentarios sobre los demás jugadores, el cual garantizará la evaluación de habilidades de los jugadores, así como sus conductas.
Prioridad del requerimiento	Alta

Requerimientos No Funcionales

Tabla 7

Requerimiento No Funcional – Seguridad de los Datos RNF1

Identificador del requerimiento	RNF1.
Nombre del requerimiento	Seguridad de los Datos.
Descripción del requerimiento	La plataforma debe garantizar la privacidad y protección de los datos personales de los usuarios, así como prevenir el acceso no autorizado a la información y cuentas de los usuarios.

Tabla 8

Requerimiento No Funcional – Escalabilidad RNF2

Identificador del requerimiento	RNF2.
Nombre del requerimiento	Escalabilidad.
Descripción del requerimiento	La plataforma debe ser capaz de manejar un gran número de usuarios y solicitudes de manera eficiente sin comprometer el rendimiento.

Tabla 9

Requerimiento No Funcional – Disponibilidad RNF3

Identificador del requerimiento	RNF3.
Nombre del requerimiento	Disponibilidad.
Descripción del requerimiento	La plataforma debe estar disponible en todo momento, minimizando el tiempo de inactividad o mantenimiento programado.

Tabla 10

Requerimiento No Funcional – Usabilidad RNF4

Identificador del requerimiento	RNF4.
Nombre del requerimiento	Usabilidad.
Descripción del requerimiento	La plataforma debe ser de fácil uso y navegabilidad para los usuarios, sin requerir de una capacitación extensa para su uso.

Tabla 11

Requerimiento No Funcional – Accesibilidad RNF5

Identificador del requerimiento	RNF5.
Nombre del requerimiento	Accesibilidad.
Descripción del requerimiento	La plataforma debe ser compatible con diferentes navegadores, para garantizar una experiencia de usuario uniforme.

Tabla 12

Requerimiento No Funcional – Rendimiento RNF6

Identificador del requerimiento	RNF6.
Nombre del requerimiento	Rendimiento.
Descripción del requerimiento	La plataforma debe ser rápida y responder de manera eficiente a las solicitudes de los usuarios, proporcionando una experiencia de usuario fluida.

Tabla 13

Requerimiento No Funcional – Confiabilidad RNF7

Identificador del requerimiento	RNF7.
Nombre del requerimiento	Confiabilidad.
Descripción del requerimiento	La plataforma debe ser confiable y estable, evitando errores inesperados o fallos del sistema.

Modelado y Diseño

Ilustración 11

Modelo Entidad – Relación

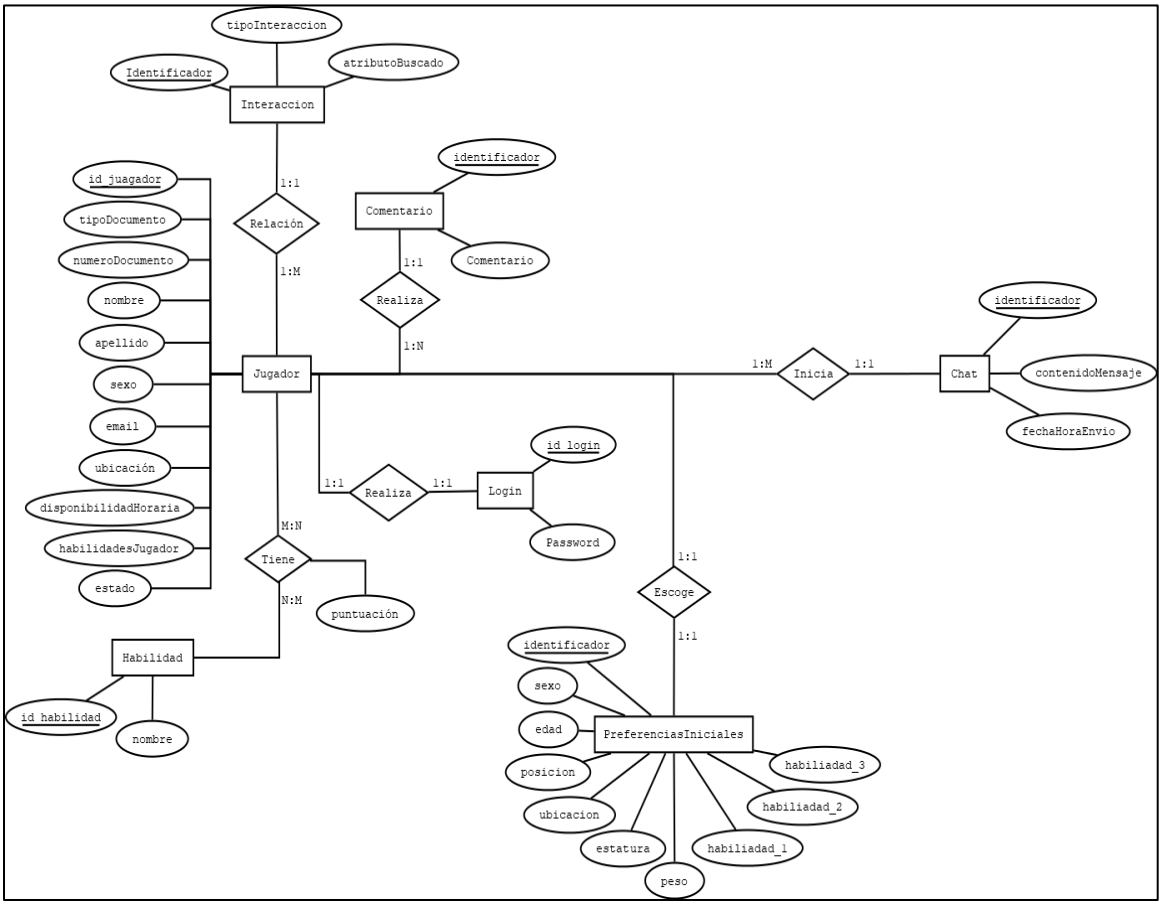


Ilustración 12

Modelo Relacional de la Base de Datos

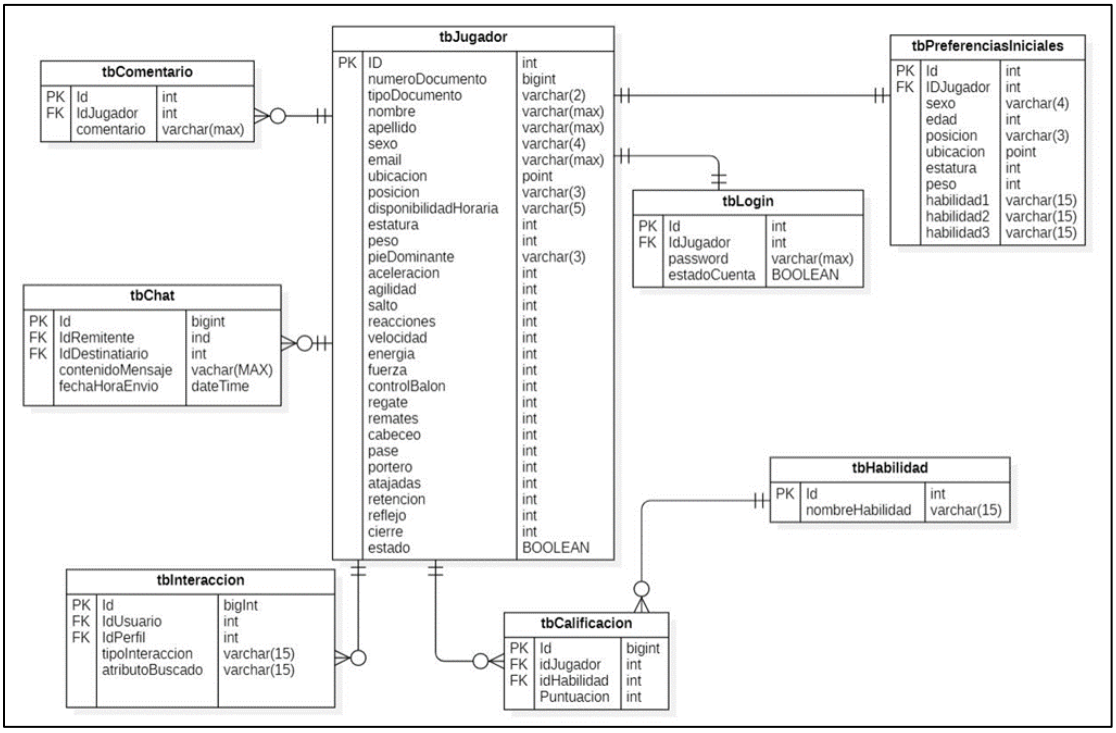


Tabla 14

Diccionario de Datos de la Tabla Login

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica el Logueo	Big Int	PK
password	Clave que se le asigna a su usuario personal	Varchar (max)	
idJugador	Id que representa al usuario en la base de datos	Big Int	FK
estadoCuenta	Estado en la que se encuentra la cuenta del usuario	BOOLEAN	

Tabla 15

Diccionario de Datos de la Tabla Usuario

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica el usuario	Big Int	PK
numeroDocumento	Numero de documento del usuario	Big Int	
tipoDocumento	Tipo de documento del usuario	Varchar(2)	
nombre	Nombre del usuario	Varchar(max)	
fechaNacimiento	Fecha de nacimiento del usuario	Date	
sexo	Sexo del usuario	Varchar(4)	
email	Email del usuario	Varchar(max)	

Entidad	Descripción	Tipo de Dato	Llave
ubicación	Ubicación corresponde a la zona geográfica determinada por las localidades de Bogotá en donde el usuario prefiere jugar	Varchar(30)	
posicion	Posición en la cancha del usuario	Varchar(3)	
disponibilidadHoraria	Disponibilidad horaria del usuario	Varchar(5)	
estatura	Estatura del usuario	Int	
peso	Peso del usuario	Int	
pieDominante	Pie dominante del usuario	Varchar(3)	
aceleración	Media de la habilidad de la aceleración del usuario	Int	
agilidad	Media de la habilidad de la agilidad del usuario	Int	
salto	Media de la habilidad del salto del usuario	Int	
reacciones	Media de la habilidad de la reacción del usuario	Int	
velocidad	Media de la habilidad de la velocidad del usuario	Int	
energía	Media de la habilidad de la energía del usuario	Int	

Entidad	Descripción	Tipo de Dato	Llave
fuerza	Media de la habilidad de la fuerza del usuario	Int	
controlBalon	Media de la habilidad del control de balón del usuario	Int	
regate	Media de la habilidad del regate del usuario	Int	
remate	Media de la habilidad del regate del usuario	Int	
cabeceo	Media de la habilidad del cabeceo del usuario	Int	
pase	Media de la habilidad del pase del usuario	Int	
atajadas	Media de la habilidad de las atajadas del usuario (portero)	Int	
retención	Media de la habilidad de la retención del usuario (portero)	Int	
reflejos	Media de la habilidad de los reflejos del usuario (portero)	Int	
cierre	Media de la habilidad del cierre del usuario (portero)	Int	
estado	Estado del jugador (Activo/Inactivo)	BOOLEAN	

Tabla 16

Diccionario de Datos de la Tabla Calificación

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica la calificación del usuario	Big Int	PK
idJugador	Id que representa al usuario en la base de datos	Big Int	FK
idHabilidad	Identificador de una habilidad en concreto	Int	FK
puntuacion	Puntaje de cada habilidad por cada usuario	Int	

Tabla 17

Diccionario de Datos de la Tabla Comentario

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica el comentario del usuario	Big Int	PK
idJugador	Id que representa al usuario al que se le está realizando el comentario	Big Int	FK
comentario	Comentario realizado por usuario	Varchar(max)	

Tabla 18

Diccionario de Datas de la Tabla Chat

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica el chat del usuario	Big Int	PK
idRemitente	Id que representa al remitente en la base de datos	Big Int	FK
idDestinatario	Id que representa al destinatario en la base de datos	Big Int	FK
contenidoMensaje	Texto o información que se encuentra en el cuerpo de un mensaje	Varchar(max)	
fechaHoraEnvio	fecha y hora en que un mensaje fue enviado desde el remitente hasta el receptor	Date time	

Tabla 19

Diccionario de Datos de la Tabla Habilidad

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica una habilidad en concreto	Int	PK
nombreHabilidad	Nombre característico de una habilidad	Varchar(15)	

Tabla 20

Diccionario de Datos de la Tabla PreferenciasInciales

Entidad	Descripción	Tipo de Dato	Llave
id	Clave Primaria, Id que identifica las preferencias del usuario	Big Int	PK
idJugador	Id que representa al usuario en la base de datos	Big Int	FK
sexo	Sexo del usuario a buscar	Varchar(4)	
posición	Posición del usuario a buscar	Varchar(3)	
ubicación	Ubicación corresponde a la zona geográfica determinada por las localidades de Bogotá en donde el usuario prefiere jugar	Varchar(30)	
estatura	Estatura del usuario a buscar	Int	

Entidad	Descripción	Tipo de Dato	Llave
peso	Peso del usuario a buscar	Int	
edad	Edad del usuario a buscar	Int	
habilidad1	Primera habilidad del usuario a buscar	Int	
habilidad2	Segunda habilidad del usuario a buscar	Int	
habilidad3	Tercera habilidad del usuario a buscar	Int	

Tabla 21

Diccionario de Datos de la Tabla Interacción

Entidad	Descripción	Tipo de Dato	Llave
Id	Clave Primaria, Id que identifica las preferencias del usuario	Big Int	PK
idUsuario	Id que representa al usuario que realiza la interacción	Big Int	FK
idPerfil	Id del perfil visitado	Big Int	FK
tipoInteraccion	Tipo de interacción realizada por el usuario	Varchar(15)	
atributoBuscado	En este campo se almacena el atributo buscado por el usuario	Varchar(15)	

Ilustración 13

Diagrama de Clases

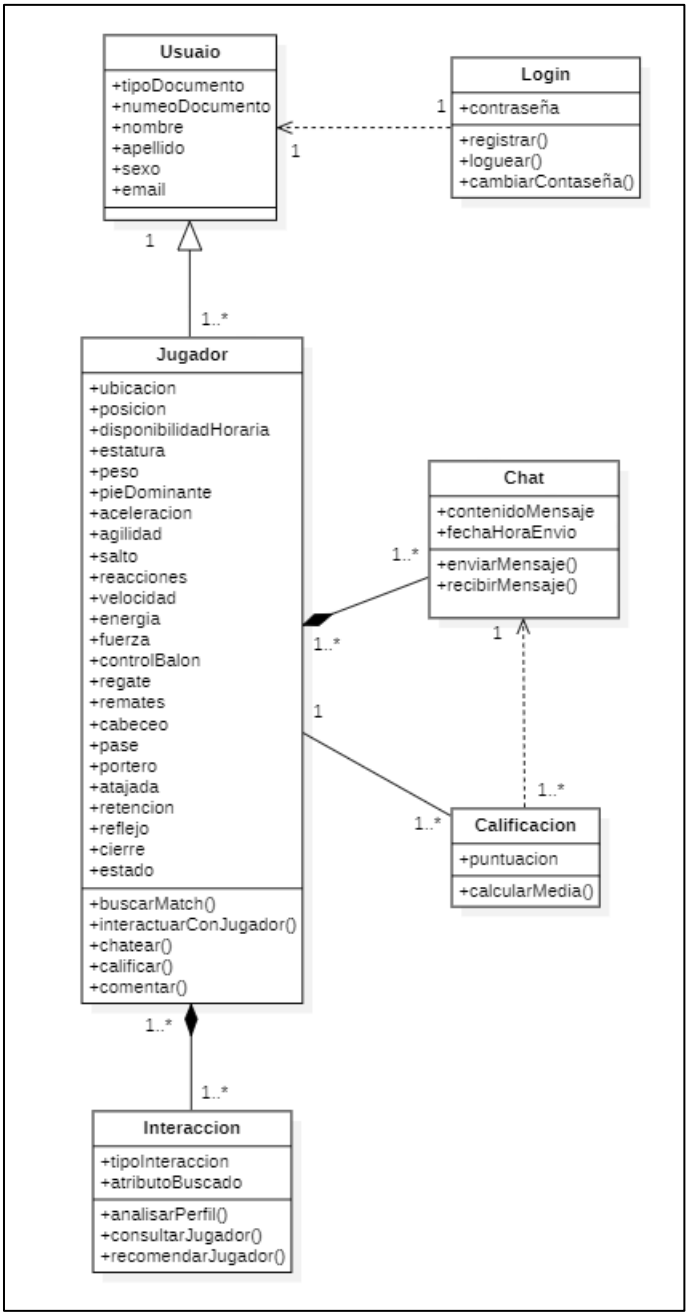


Ilustración 14

Casos de Uso

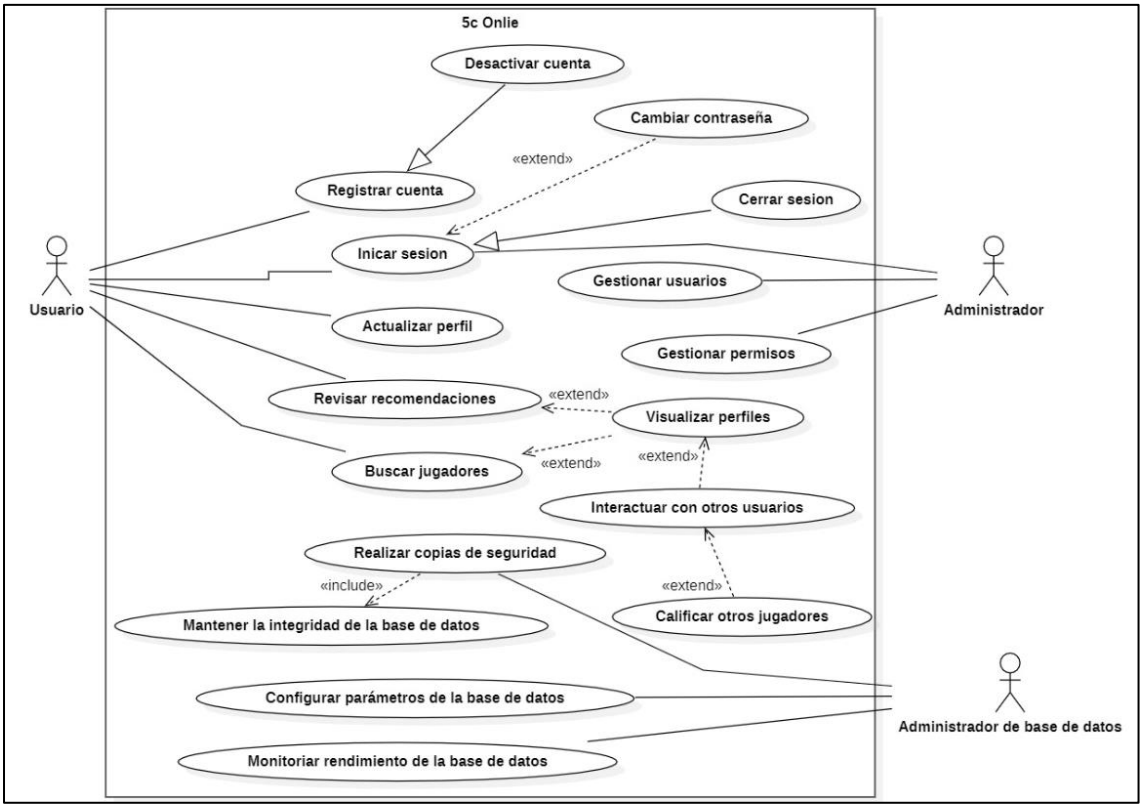


Ilustración 15

Diagrama de Flujo – Inicio de Sesión

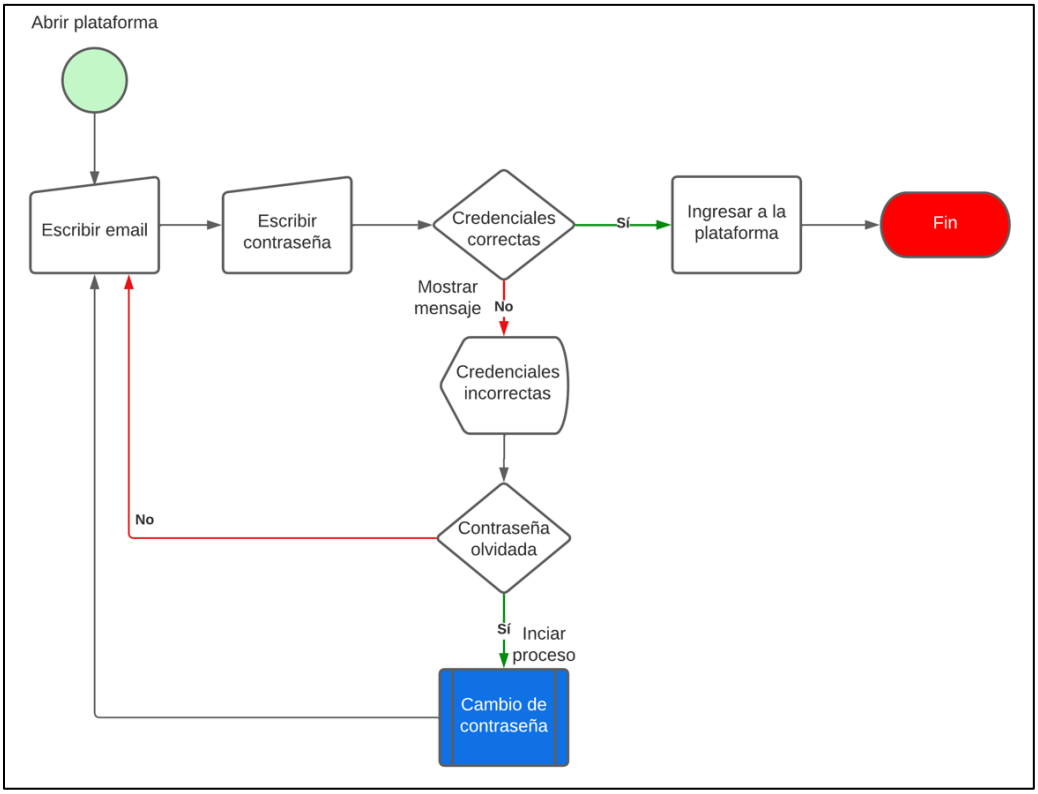


Ilustración 16

Diagrama de Flujo – Creación de Perfil

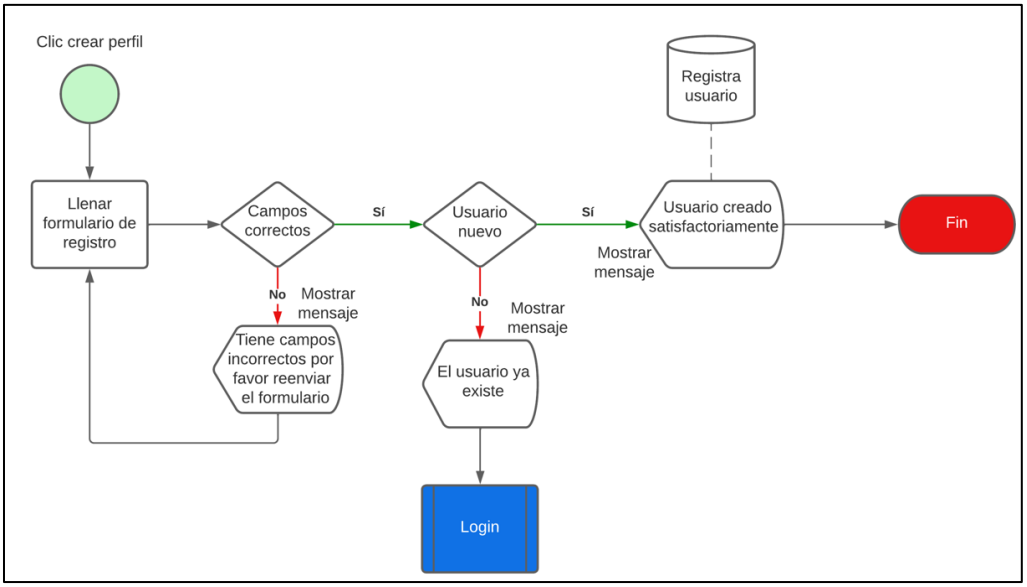


Ilustración 17

Diagrama de Flujo – Edición de Perfil

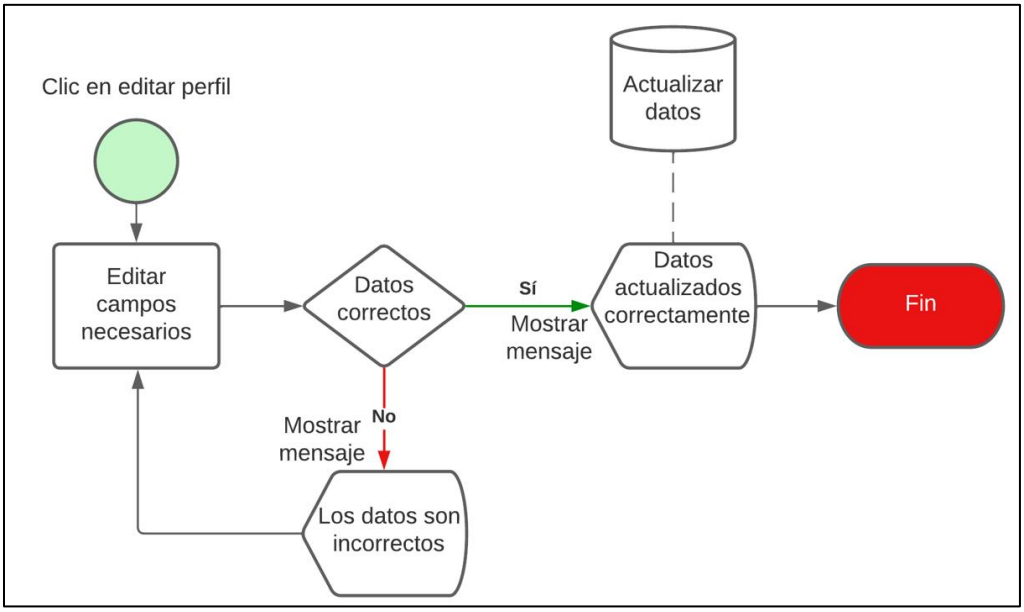


Ilustración 18

Diagrama de Flujo – Cambio de Contraseña

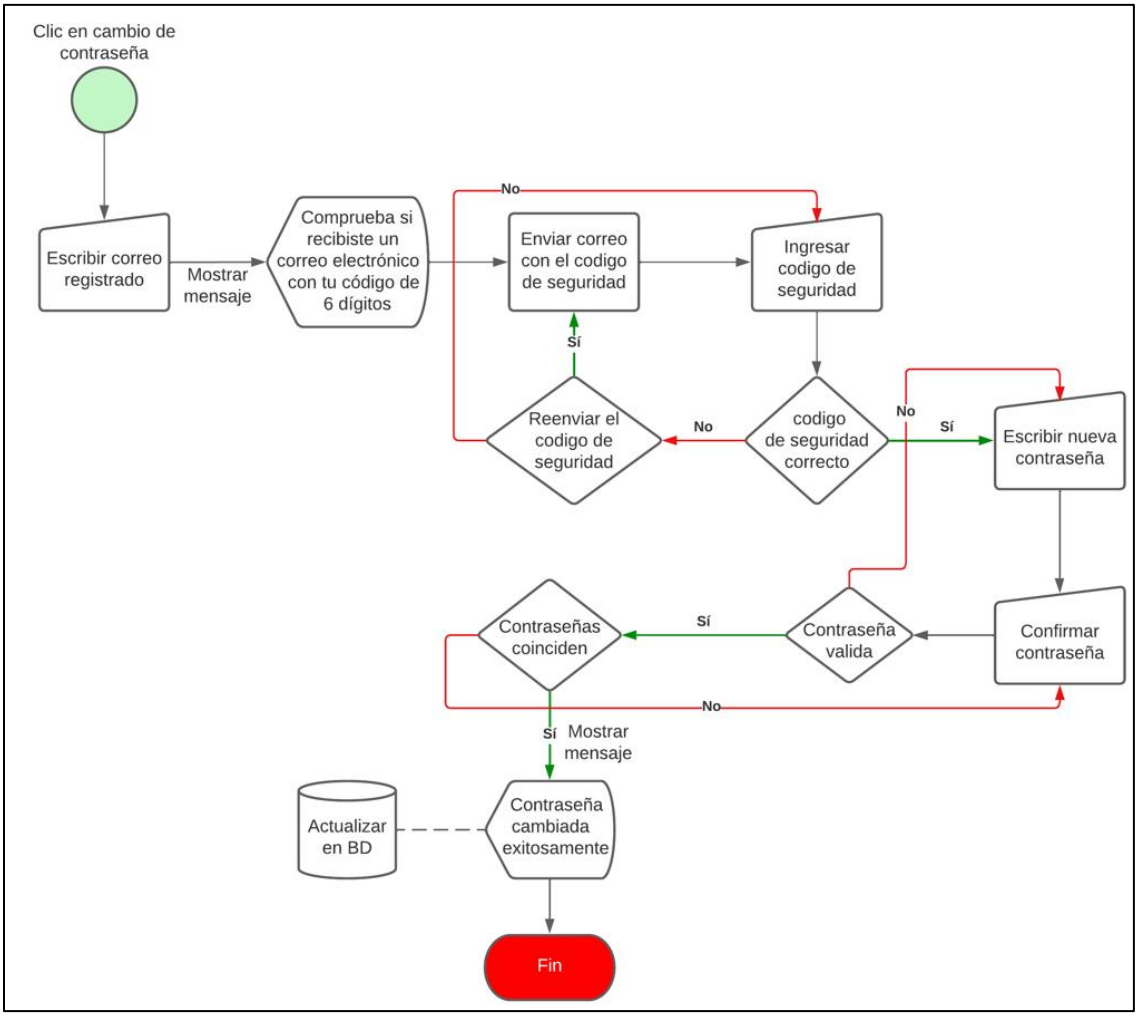


Ilustración 19

Diagrama de Flujo - Inicio de Conversación

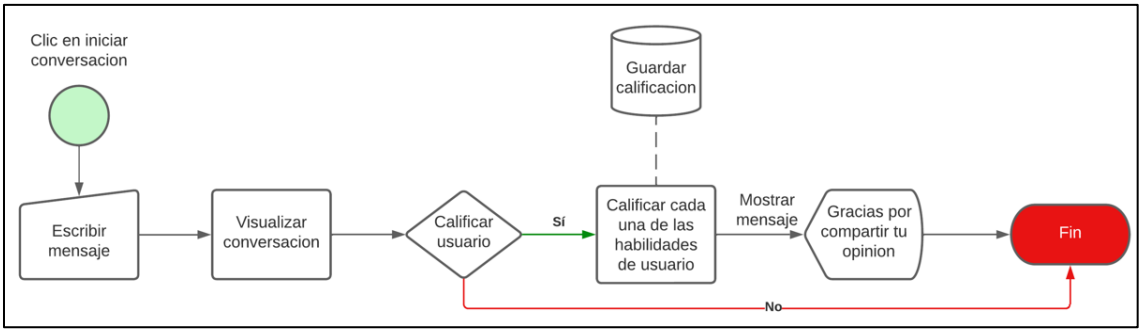


Ilustración 20

Diagrama de Flujo – Búsqueda

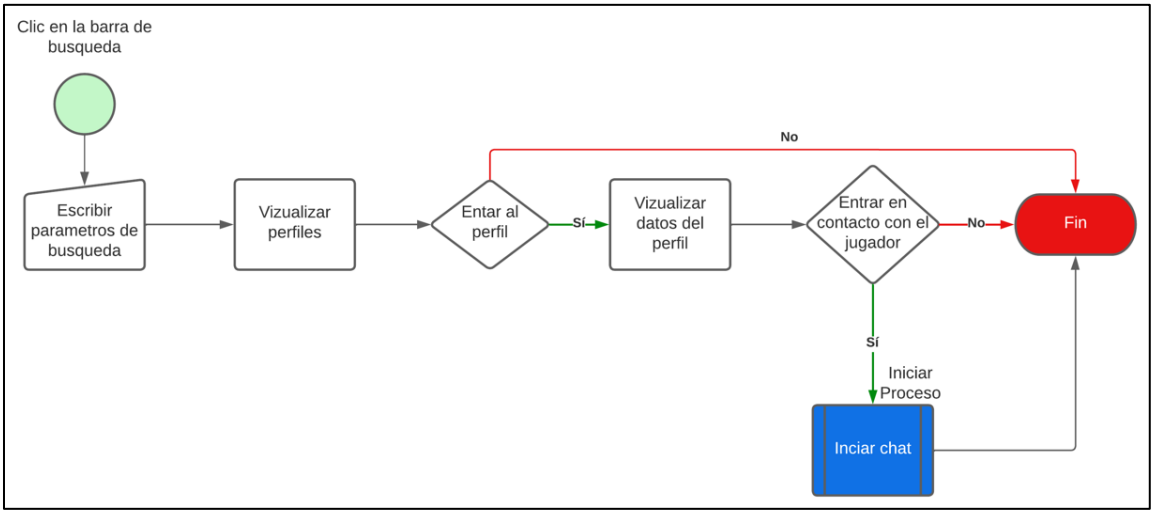
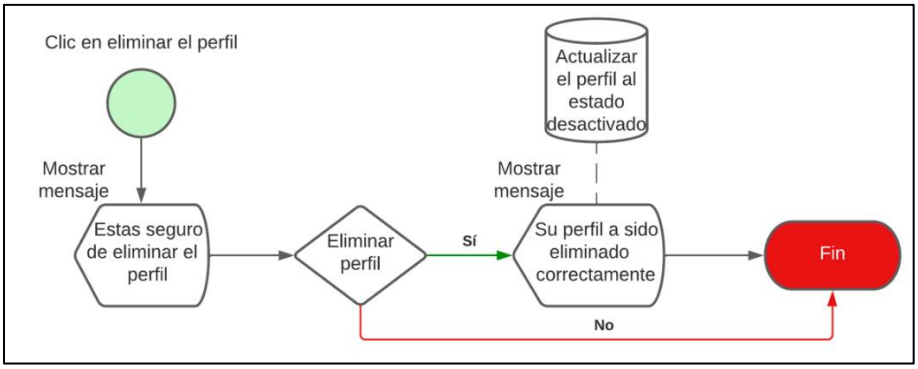


Ilustración 21

Diagrama de Flujo – Eliminación de Perfil

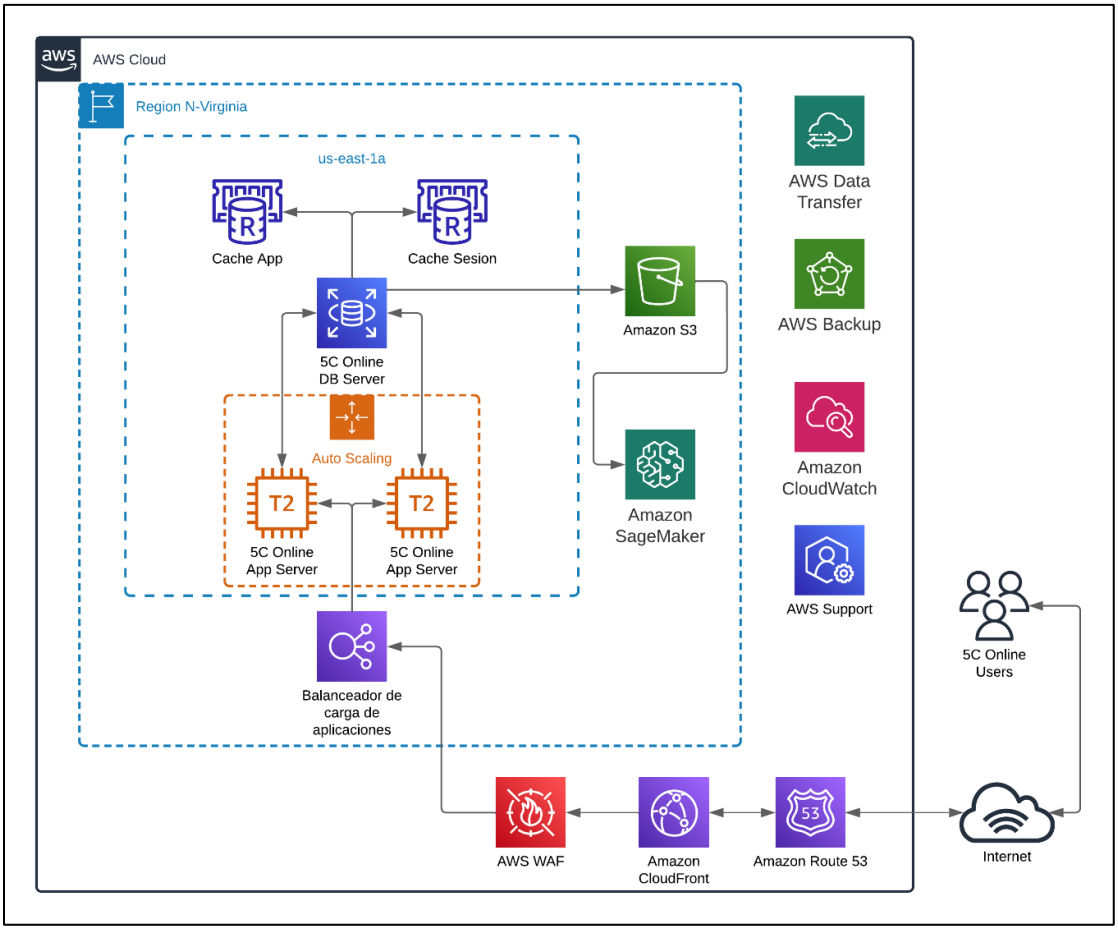


Descripción General de la Arquitectura AWS

En la Ilustración 22 se muestran los componentes o servicios que integran y soportarán la plataforma digital 5C Online, a continuación, se hace una breve descripción de la funcionalidad de cada uno de ellos y cómo se relacionan unos con los otros

Ilustración 22

Diagrama de la Arquitectura en AWS



Amazon Route 53: Es un servicio de resolución de nombre de servicios (DNS) y de registro de dominios desarrollado por AWS, por lo que la función que cumplirá este servicio en la arquitectura global en la infraestructura de la plataforma es el de proveer los dominios pertinentes para que sean legibles para los usuarios de internet y a su vez realice la resolución de DNS y redirija todas las peticiones de los usuarios al recurso de la plataforma 5C Online alojada en AWS.

Amazon CloudFront: Es un servicio de distribución de contenido en cache, el cual proporcionará una experiencia de usuario más agradable ya que se encargará de almacenar en cache el contenido estático de la plataforma digital en ubicaciones estratégicas minimizando así, la latencia y reduciendo así los tiempos de carga y espera a la hora de mostrar el contenido, mejorando la experiencia del usuario final.

AWS WAF: Es un acrónimo de sus siglas que significa Web Application Firewall y como su nombre lo indica, es un servicio de Firewall a nivel lógico, que proporciona seguridad a las aplicaciones desplegadas en la nube de AWS, protegiéndolas contra ataques maliciosos como inyección SQL, ataques DDoS entre otros. Este servicio se relaciona tanto con Amazon CloudFront y Elastic Load Balancing ya que permite crear reglas de seguridad personalizadas protegiendo las aplicaciones web desde los puntos o ubicaciones donde se almacena el cache de estas, de igual forma permite escalar la protección de manera eficiente.

Amazon Elastic Load Balancig: Es un servicio que proporciona balanceo de tráfico entre diferentes destinos como las instancias de Amazon EC2 en las que se alojará la plataforma 5C Online, este servicio funciona a nivel lógico y conceptual en la capa 7 del modelo OSI, lo que puede traducirse en la capacidad de enturar las solicitudes basándose en información más detalladas; se integra de manera nativa con el servicio de WAF de AWS ya que recibe las solicitudes de los usuarios luego de ser filtradas en firewall; de igual manera, se integra con las instancias de EC2, ya que este es uno de los posibles destinos a los que se les distribuye las cargas y solicitudes de tráfico.

Amazon EC2: Es un servicio que proporciona capacidad de cómputo en la nube permitiendo ejecutar y administrar servidores de manera escalable y flexible, aquí es donde se alojará la plataforma 5C Online.

Amazon Auto Scaling: Este es un servicio integrado con Amazon EC2, el cual proporciona una de las características más importantes y diferenciadoras de la nube, es decir, el escalamiento horizontal, lo que aumentará la disponibilidad y el rendimiento de la plataforma 5C Online, traduciéndose así, en una mejor experiencia para el usuario final.

Amazon RDS: Amazon Relational Database Service, es un servicio que proporciona servidores de bases de datos completamente administrados por AWS. Este servicio será usado para ejecutar la capa de bases de datos para la plataforma 5C Online, se integra de manera nativa con el servicio Amazon EC2, lo que permite trabajar la capa de aplicación y de base de datos en diferentes servidores.

Amazon ElastiCache: Es un servicio de almacenamiento en cache que se integra de manera directa con servicios como Amazon RDS, garantizando así una mejora en el rendimiento y escalabilidad de la plataforma 5C Online, ya que es aquí donde se almacenarán los datos de cache de factores importantes como lo son la sesión y la aplicación, de manera tal que sea fácil recuperar los datos importantes de la plataforma logrando así reducir la carga de la base de datos, ya que se almacenarán en cache los datos que con frecuencia son utilizados por la plataforma.

Amazon S3: Es un servicio de almacenamiento escalable, para el almacenamiento de objetos que puede hacer las funciones de un Data WareHouse. En este servicio se almacenará la data que se usará posteriormente para el entrenamiento del algoritmo de Machine Learning, se integra directamente con servicios como Amazon RDS, ya que es de las bases de datos, de donde se extraerá dicha información y será almacenada en Buckets de S3 como archivos separados por comas CSV.

Amazon SageMaker: Es un servicio completamente administrado por AWS que se usa para desarrollar, entrenar e implementar modelos de Machine Learning, este servicio permite la preparación de datos, el entrenamiento de modelos de Machine Learning y la implementación de los modelos en entornos de producción. Este servicio se integra directamente con Amazon S3, ya que es de este servicio de almacenamiento de donde se obtendrá la data que se usará para el entrenamiento del modelo de Machine Learning que usará la plataforma 5C Online para poder realizar las respectivas recomendaciones.

AWS Backup: Es un servicio que proporciona la posibilidad de realizar copias de seguridad y configuración de políticas de ciclo de vida de las copias de distintos servicios de AWS como Amazon RDS y Amazon EC2, se integra directamente con Amazon S3, ya que se usan los Buckets para almacenar esta información.

AWS CloudWatch: Es un servicio de monitoreo que permite la recolección de métricas, registros y eventos al interior de la cuenta de AWS, se usará para realizar el monitoreo y detección de anomalías en los diferentes servicios desplegados en la arquitectura que soportará la plataforma 5C Online, así como la creación de alarmas cuando un evento sea detectado.

Mockups de la plataforma

Ilustración 23

Mockup de la UI del Login

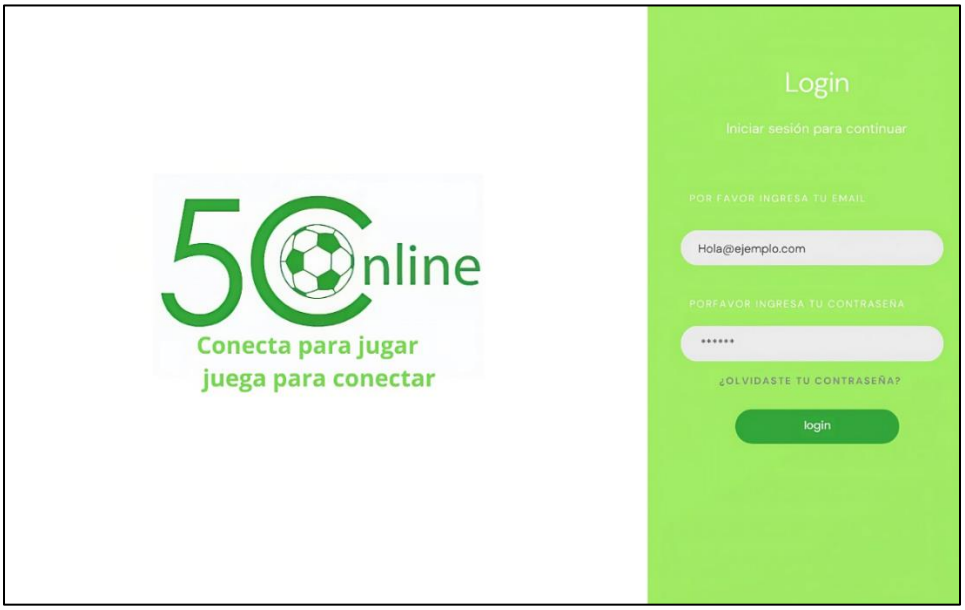


Ilustración 24

Mockup de la UI para la Recuperación de la Cuenta

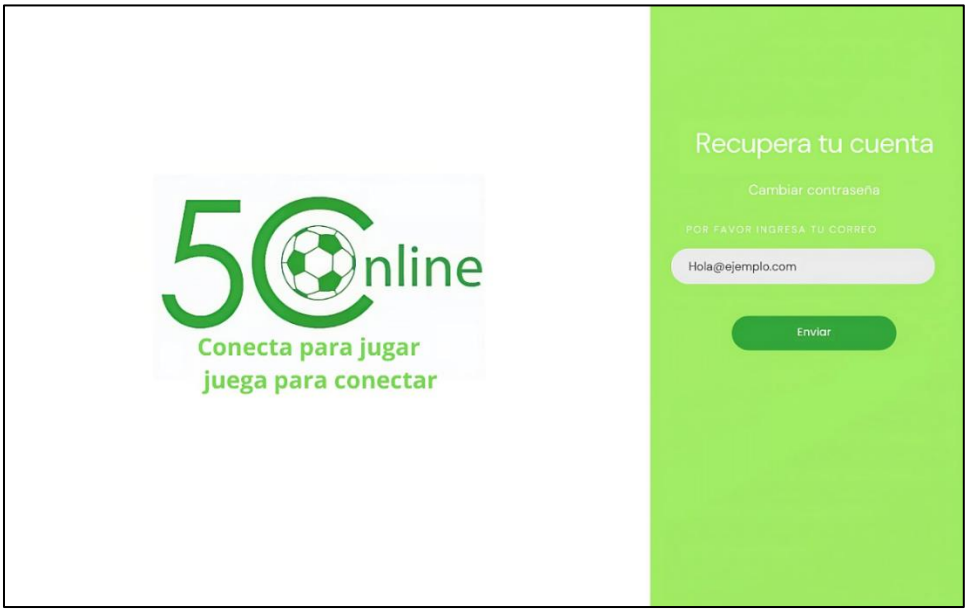


Ilustración 25

Mockup de la UI para el Cambio de Contraseña



Ilustración 26

Mockup de la UI de Confirmación de Correo Enviado

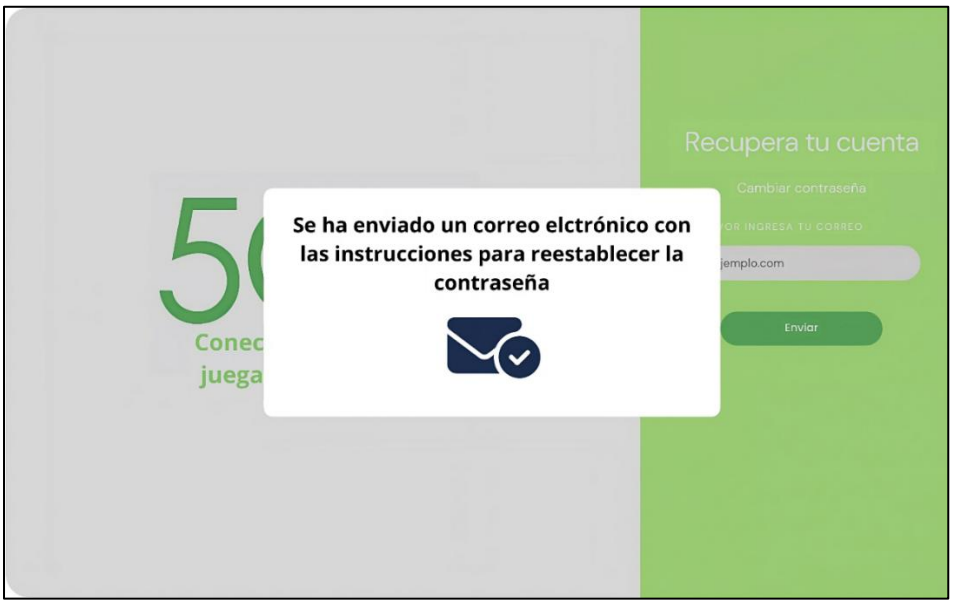


Ilustración 27

Mockup de la UI para la Creación de Usuarios



Ilustración 28

Mockup de la UI de Elección de Intereses 1



Ilustración 29

Mockup de la UI de Elección de Intereses 2

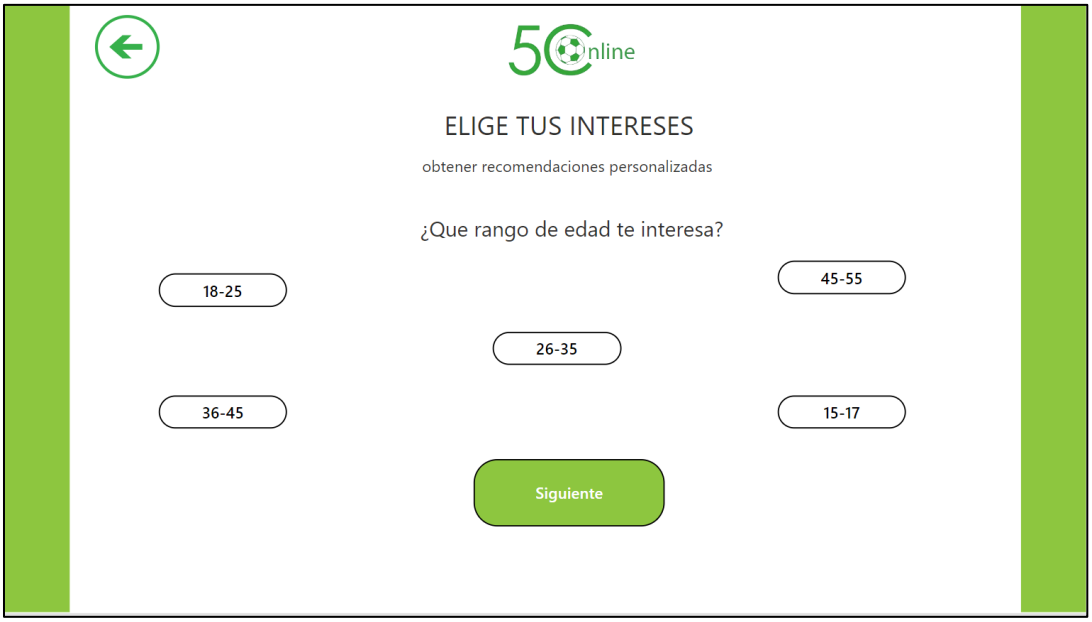


Ilustración 30

Mockup de la UI de Elección de Intereses 3



Ilustración 31

Mockup de la UI de Elección de Intereses 4



Ilustración 32

Mockup de la UI de Elección de Intereses 5



Ilustración 33

Mockup de la UI de Elección de Intereses 6



Ilustración 34

Mockup de la UI de Elección de Intereses 7



Ilustración 35

Mockup de la UI de la Página Principal

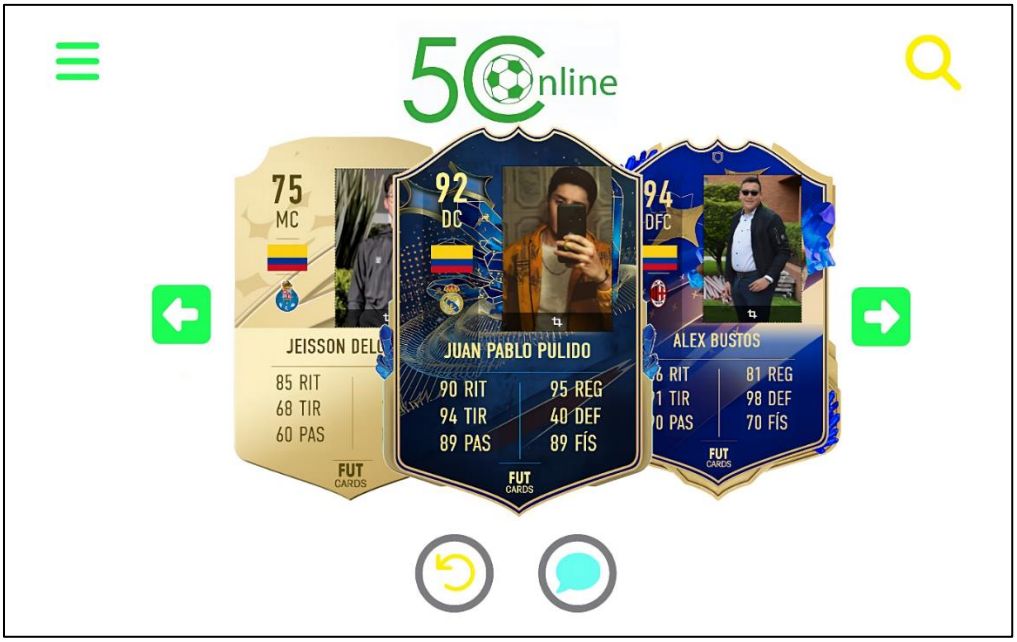


Ilustración 36

Mockup de la UI de la Página Principal 2

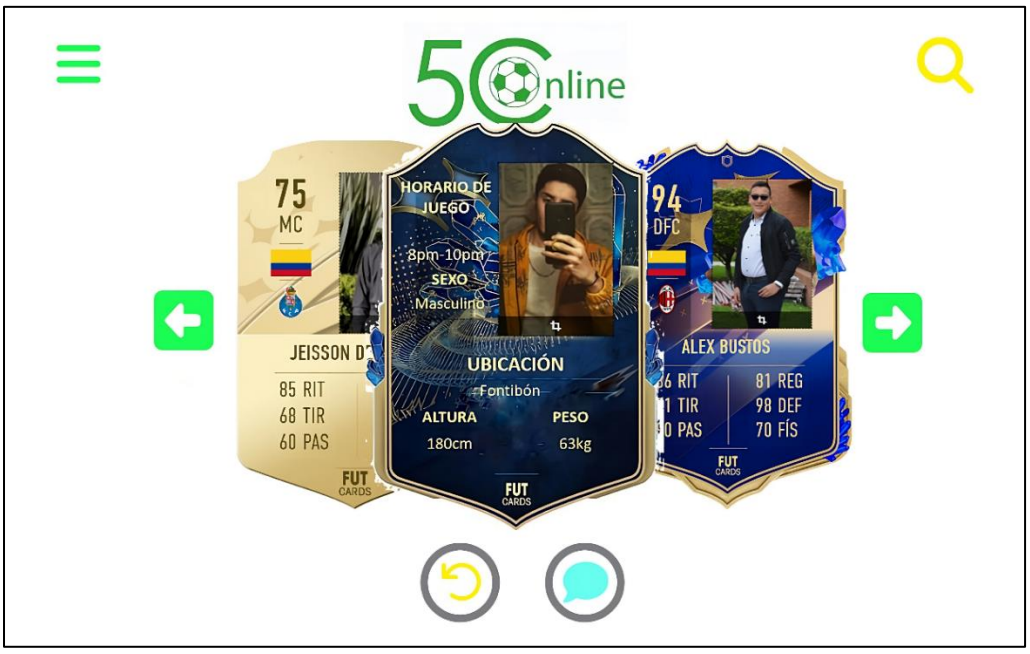


Ilustración 37

Mockup de la UI del Menú Contextual

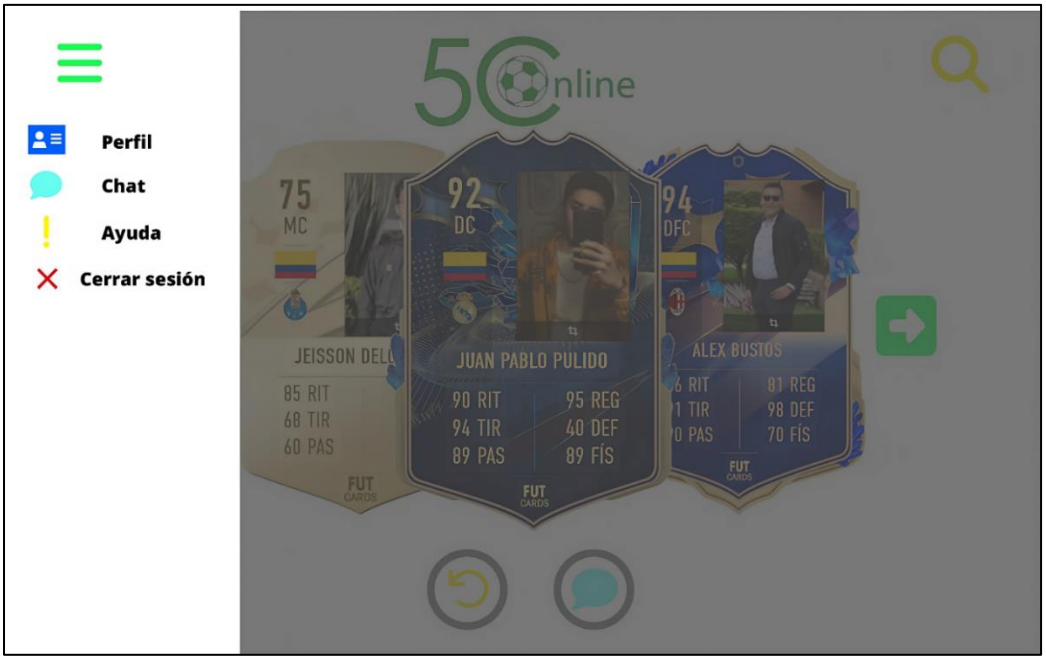


Ilustración 38

Mockup de la UI para la Edición de Perfil



Ilustración 39

Mockup de la UI de la Búsqueda por Filtros

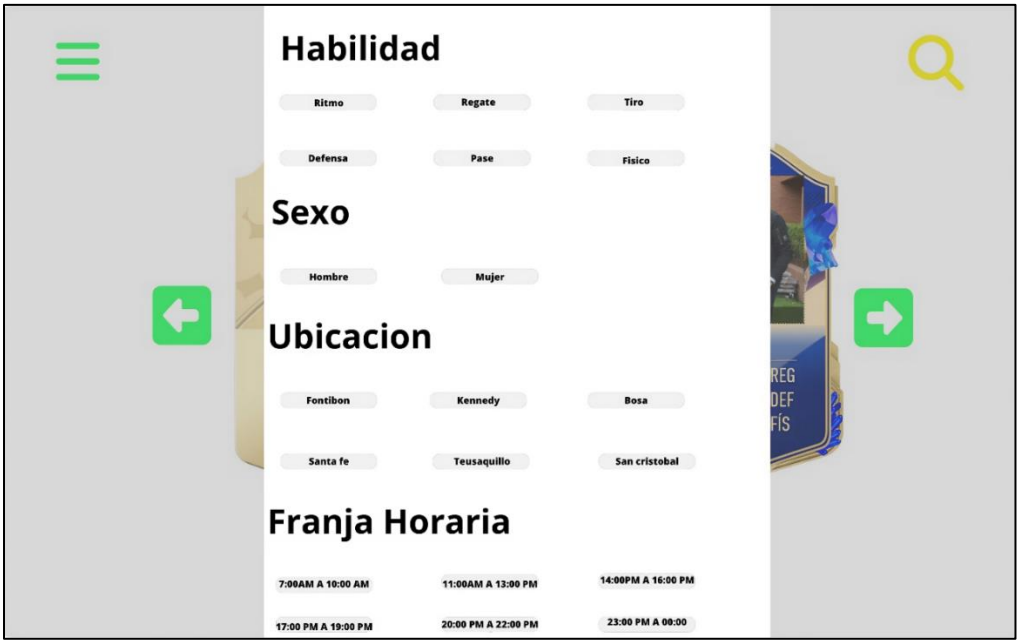


Ilustración 40

Mockup de la UI del InBox



Ilustración 41

Mockup de la UI del Chat

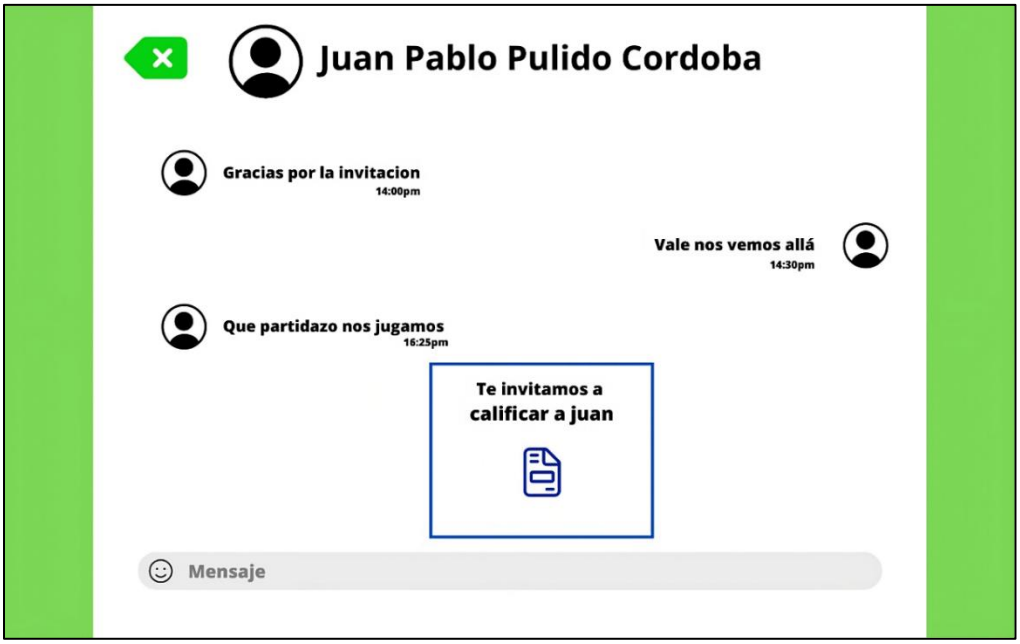
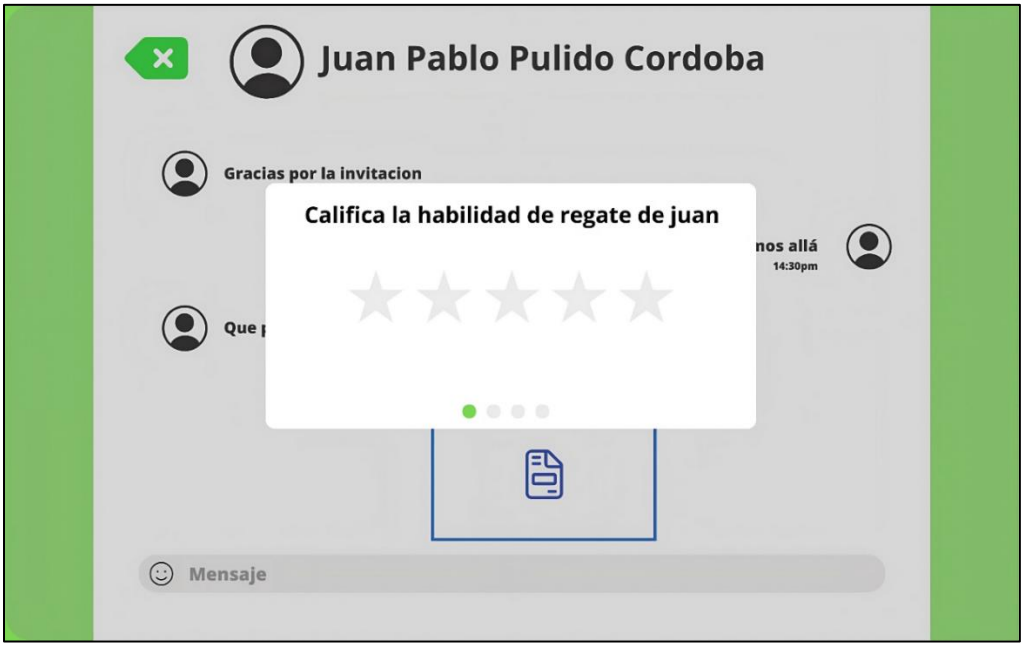


Ilustración 42

Mockup de la UI para la Calificación de Usuarios



Funcionamiento del Algoritmo

Filtrado Colaborativo

Un filtrado colaborativo es una técnica utilizada en el machine learning que permite recomendar nuevos usuario o artículos dependiendo de los características, interacciones o patrones que haya tenido dicho usuario o articulo con los elementos dentro del algoritmo o sistema.

En este caso, se evidencia el filtrado colaborativo en las posibles recomendaciones que recibirá de acuerdo con las interacciones con los usuarios que ha cliqueado, chateado y calificado.

La ponderación de puntos es de 1 a 7 de acuerdo con un rango definido de puntuación, donde se pondera según el nivel de interacción, como se observa en la Tabla 22:

Tabla 22

Ponderación por cada tipo de interacción

Tipo de Interacción	Puntuación
Sin Interacción	0
Clic[a]	1
Chat[b]	2
Calificación 1[c]	3
Calificación 2[d]	4
Calificación 3[e]	5
Calificación 4[f]	6
Calificación 5[g]	7

- a Hace referencia a cuando el usuario ha dado Clic en determinado perfil de usuario.
- b Hace referencia a cuando el usuario inicia un chat con determinado usuario de la plataforma.
- c Hace referencia a cuando un usuario califica a otro usuario con 1 estrella calificadora.
- d Hace referencia a cuando un usuario califica a otro usuario con 2 estrellas calificadora.
- e Hace referencia a cuando un usuario califica a otro usuario con 3 estrellas calificadora.

- f Hace referencia a cuando un usuario califica a otro usuario con 4 estrellas calificadora.
- g Hace referencia a cuando un usuario califica a otro usuario con 5 estrellas calificadora.

Existe un proceso de búsqueda de jugador, el cual se divide en dos partes, la barra de búsqueda del sistema que solo permitirá buscar basada en filtros en un orden jerárquico con el nivel de importancia que se la ha otorgado, en este caso:

1. Habilidad (60%)

2. Sexo (20%)

3. Disponibilidad horaria (10%)

4. Localidad preferida (10%)

Se predefinió de esta forma con forme a las necesidades del usuario final; en las canchas de futbol 5 predominan las habilidades que tiene un jugador sobre la posición en la que juega, porque no existe gran variedad de estas últimas debido al tipo de cancha en la que se juega, la mayoría de las personas están de acuerdo en que es posible cambiar de posición dentro de la cancha, lo que no es posible cambiar, es la habilidad que se tiene con el balón. Siguiendo esta misma línea, luego de conocer que habilidades, interesa saber en donde le gusta jugar y sobre que franja horaria.

Si un usuario (a) consulta el perfil de un usuario (b) dándole clic al perfil, se le asignará un punto (1), que se almacenará para trabajar más adelante si no hace un tipo de interacción adicional, supongamos que va a realizar los 3 tipos de interacciones, luego de dar clic, mensajea

con el usuario (b), el hecho de que haya una respuesta por el usuario (b) le otorga 2 puntos, se sobrescribe sobre el punto anterior y este es ahora el que se tendría en cuenta, por último, luego de que los dos usuarios hayan pactado y jugado un partido, se habilitará la opción de calificar al otro usuario, la calificación de 1 a 5 se otorgará de acuerdo al criterio propio, cómo se explicó anteriormente, si se otorga 1 estrella, equivaldrá a 3 puntos que se reemplazarán sobre los dos puntos que ya tenía por haber conversado con el otro usuario, de esta manera con las posibles puntuaciones restantes.

Los puntos que se le han otorgado al usuario servirán para recomendarle un futuro usuario, entre más interacciones con otros usuarios haya tenido, más certera será la recomendación, el algoritmo funciona de la siguiente manera utilizando la codificación One-Hot, según Lekhana, 2019 "…varios modelos de aprendizaje automático no funcionan con datos categóricos y, para ajustar estos datos al modelo de aprendizaje automático, es necesario convertirlos en datos numéricos" Es decir, que One-Hot permite adaptar cualquier tipo de tabla-matriz categórica en una que el modelo de aprendizaje profundo pueda entender. Y es lo que se realiza. Se cuenta con una matriz de interacción, que es donde se almacenan las puntuaciones que previamente se otorgaron, el usuario (a) puntuó al usuario (b) con 4 estrellas, por ende, tiene una puntuación general de (6), con la información proporcionada hasta el momento, la matriz de interacciones quedaría como se evidencia en la Ilustración 43.

Ilustración 43

Matriz Interacción

		Usuario calificado					
		User a	User b	User c	User d	User e	User f
Usuario calificador	User a	-	6	4	0	0	2

Ahora bien, para codificar esta matriz en una que el algoritmo pueda utilizar, se requiere entender los siguientes puntos.

Habilidad: Cuando un usuario se registra dentro de la plataforma, se solicitan unos datos iniciales, algunos de ellos, y que son los que importan para el ejemplo, altura, peso, sexo y edad, teniendo en cuenta estos atributos se realiza un cálculo que permite identificar que promedio tendría en una habilidad específica, de esta manera cada persona que se registre tendrá una media diferente en cada habilidad. Además de los otros atributos adicionales ya mencionados que tenemos en cuenta para hacer recomendaciones. La matriz de atributos iniciales quedaría de la siguiente manera:

Sexo: Se categorizó como segundo atributo importante para la recomendación de futuros jugadores, ¿cómo sabemos si el próximo jugador que va a recomendar es un hombre o una mujer? Cuando se interactúa con un jugador se asignará (1) al campo de sexo si es hombre y (-1) si es mujer. De esta manera se colocará el número que más se repita dentro de la cantidad de usuarios que haya calificado. Si son 10 registros y hay 6 (-1) sobre 4 (1) se realizará el cálculo sobre -1.

Si la siguiente persona que se va a recomendar es hombre, tendrá un 20% de probabilidad adicional de que sea recomendado, si es mujer, no se le va a asignar ese 20% pero si es cero se adicionará ese 20% a cualquier sexo.

Disponibilidad: Todos los jugadores cuando se registran deberán indicar cuál es la franja horaria más concurrente en la que suelen jugar, dentro de un rango máximo de 4 horas y mínimo de 1 hora, por ejemplo: 11 a 15. El algoritmo lo que hará es revisar el histórico de jugadores con los que ha interactuado y ver cuál es la hora inicial que más se repite, al igual que con la hora

final, esta va a ser la hora que será tomada en cuenta para asignar el 10% adicional de recomendación. Por ejemplo, supongamos que hay nuevamente 10 registros de interacciones con jugadores, de los 10 jugadores, 6 prefieren 12 como hora inicial y 14 como hora final, los otros 4 manejan rangos de hora diferente. Se creará un nuevo rango de hora de 12 a 14, si el jugador que voy a recomendar está dentro de esta franja, por ejemplo: de 11 a 15 se le asignará el 10% adicional. En este caso 11 está por debajo de la hora inicial y 15 por arriba de la hora final, se plantea una fórmula que verifique si está dentro del rango, está dada de la siguiente manera:

Las dos condiciones de "SI" deben ser verdaderas; Si la hora inicial es mayor o igual a la hora inicial del usuario a recomendar y menor o igual a la hora final del usuario a recomendar O si la hora final es mayor o igual a la hora inicial del usuario a recomendar y la hora final es menor o igual a la hora final a recomendar, asigne el 10%

OR((Hi1>=Hi2 and Hi1<=Hf2) and (Hf1>=Hi2 and Hf1 <=Hf2)):

Hi1 Corresponde a la hora inicial del rango que se ha creado de acuerdo con pasadas interacciones

Hi2 Corresponde a la hora inicial del rango del posible usuario a recomendar

Hf1 Corresponde a la hora final del rango que se ha creado de acuerdo con pasadas interacciones

Hf2 Corresponde a la hora final del rango del posible usuario a recomendar

Si la formula retorna verdadero, se asigna 10% de lo contrario 0%

Localidad: Con este atributo sucede algo similar que con la disponibilidad, funciona de la siguiente manera; se evalúa la cantidad de localidades que se repiten y se asignan porcentajes

conforme a la cantidad de repetidos que se encuentren, por ejemplo si hay 10 registros y de esos 10 hay 7 que digan “bosa”, se le asignará un 7% adicional para recomendar y el 3% que sobre se asignarán a las otras localidades según la cantidad que haya, si ese 3% corresponde a “suba”, y el posible usuario a recomendar es de suba se le agregará ese 3%.

De esta manera, se tendría una matriz resultante que se vería de como se muestra en la Ilustración 44:

Ilustración 44

Matriz Inicial Resultante

		Atributos iniciales							
		Regate	Pase	Tiro	Atajada	Sexo	H:inicial	H:Final	Localidad
Usuarios	User b	90	53	26	38	1	11	13	Bosa
	User g	41	8	72	39	-1	12	15	Bosa

Con Esto se soluciona un problema común en el diseño de los algoritmos, el arranque en frío.

Según Silvi (2016) “La falta de información ante un usuario o ítem nuevo dificulta la tarea de encontrar resultados relevantes, pues no hay historial anterior con el cual comparar.”

Al asignar un promedio inicial a cada habilidad se está omitiendo el error de asignar valores iniciales iguales a todos los jugadores que se registren, ya la media de dichos jugadores irá cambiando de acuerdo con el tipo de interacciones que tengan dentro de la plataforma

Hasta ahora se endiente que cuando un usuario se registra, se le asigna una media por habilidad y que cuando se puntúa a un jugador, se asigna una ponderación que se guarda en la matriz de recomendación. Con estos puntos claros se puede proceder con la codificación one-hot,

la matriz perfilamiento sería el resultante de la multiplicación entre la matriz interacciones y la matriz atributos iniciales, quedando como se muestra en la Ilustración 45:

Ilustración 44

Matriz Perfilamiento

		Atributos iniciales							
		Regate	Pase	Tiro	Atajada	Sexo	H:inicial	H:Final	Localidad
Usuarios	User b	540	318	156	228	1	11	13	Bosa

Se multiplica cada habilidad por la puntuación que se le asignó en la matriz de interacciones (6). Para iniciar con la recomendación se procede de la siguiente manera: se suman los totales de las habilidades, se multiplica dicha sumatoria por 60, que es el porcentaje que se le asignó al apartado de habilidades, esto es hasta ahora lo que se ve en la imagen.

Si sumamos el total de las habilidades se obtiene un total de 1242, lo vamos a tener en cuenta para obtener el porcentaje del total de cada habilidad si se hace la operación correspondiente para obtener un porcentaje equivalente al 60% inicial de habilidad se vería como se muestra en la Ilustración 46:

Ilustración 46

Matriz Perfilamiento 2

		Atributos iniciales							
		Regate	Pase	Tiro	Atajada	Sexo	H:inicial	H:Final	Localidad
Usuarios	User b	540	318	156	228	1	11	13	Bosa
		26%	15%	8%	11%				

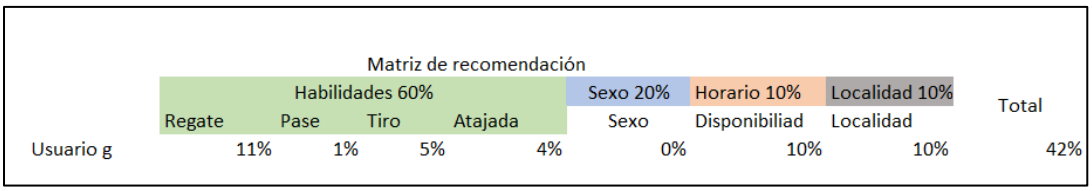
La sumatoria de los porcentajes en amarillo no pueden superar el 60%, para los demás atributos, como no hay usuarios adicionales, no es necesario revisar que tanto se repiten los

datos, sabemos que la franja horaria a buscar será de 11 a 13 y si la localidad es bosa se le otorgará el adicional del 10%.

Ahora se puede pasar a la matriz final de recomendación, la cual indicará que jugador se debe recomendar como se muestra en la Ilustración 47:

Ilustración 47

Matriz Recomendación



Matriz de recomendación								
	Habilidades 60%				Sexo 20%	Horario 10%	Localidad 10%	Total
	Regate	Pase	Tiro	Atajada	Sexo	Disponibiliad	Localidad	
Usuario g	11%	1%	5%	4%	0%	10%	10%	42%

Para llegar a los porcentajes que se muestran en habilidad, se multiplica cada habilidad del usuario (g) que es el único en la lista aparte del (b), por el ponderado de la habilidad del usuario b, esto determinará que tal se encuentran las habilidades y/o atributos del usuario (g) frente al acumulativo de los usuarios con los que se interactuó, para al final concluir que al usuario (g) se le asignaría un 42% de probabilidad de ser recomendado frente a otros posibles jugadores que se encuentren en la plataforma.

Ahora se explicará con una matriz un poco más grande para evidenciar el funcionamiento del algoritmo:

Como se muestra en la Ilustración 48, el usuario 1 ha calificado a 8 usuarios:

Ilustración 48

Matriz Interacción 2

Usuario calificador	Usuario calificado							
	User 1	User 2	User 3	User 4	User 5	User 6	User 7	User 8
User 1	-	5	4	0	0	2	5	5

Estos 8 usuarios tienen ya definidas las medias que se han otorgado en función de sus atributos físicos como se muestra en la Ilustración 49

Ilustración 49

Matriz Inicial 2

	Habilidad				Disponibilidad		Localidad
	regate	pase	Tiro	atajada	H_Inicial	H_Final	Localidad
usuario2	62	41	1	26	11	13	Bosa
usuario3	35	85	24	72	12	20	Fontibon
usuario4	48	28	9	49	15	17	Bosa
usuario5	50	62	23	95	12	18	Fontibon
usuario6	52	56	94	14	10	15	Bosa
usuario7	97	63	94	48	20	22	Suba
usuario8	71	93	58	44	13	18	Bosa
Usuario 9	84	4	10	47	21	22	Bosa
Usuario10	36	4	18	10	10	20	Fontibon
Usuario 11	46	54	45	18	11	15	Madrid

Estas medias se multiplican por el nivel de puntuación que hayan recibido producto de las interacciones que hayan tenido, se obtiene un porcentaje equivalente a la sumatoria de las habilidades y que corresponde al 60% que se le había otorgado en el inicio. El resultado de esta multiplicación se muestra en la Ilustración 50.

Ilustración 50

Matriz Perfilamiento 3

Matriz de perfilamiento								
	regate	pase	Tiro	atajada	Genero	H:inicial	H:Final	Localidad
usuario2	310	205	5	130	1	11	13	Bosa
usuario3	140	340	96	288	1	11	20	Fontibon
usuario4	0	0	0	0	1	11	17	Bosa
usuario5	0	0	0	0	-1	11	18	Fontibon
usuario6	104	112	188	28	-1	11	15	Bosa
usuario7	485	315	470	240	-1	11	22	Suba
usuario8	355	465	290	220	-1	24	18	Bosa
Total perfil buscad	1394	1437	1049	906	-1	11	18	7
Ponderado	17%	18%	13%	11%				
4786								

Se utilizan los porcentajes obtenidos para verificar cuáles de los usuarios que aún no han sido calificados son más probables de recomendar para pactar un encuentro, en este caso el usuario 9 como se evidencia en la Ilustración 51.

Ilustración 51

Matriz Recomendación 2

	Habilidades 60%				Sexo 20%	Horario 10%	Localidad 10%		
	regate	pase	tiro	atajada	Sexo	Disponibiliad	Localidad	total habilidad	total
Usuario 9	15%	1%	1%	5%	20%	0%	6%	22%	48%
Usuario10	6%	1%	2%	1%	20%	10%	3%	11%	43%
Usuario 11	8%	10%	6%	2%	0	10%	0%	26%	36%

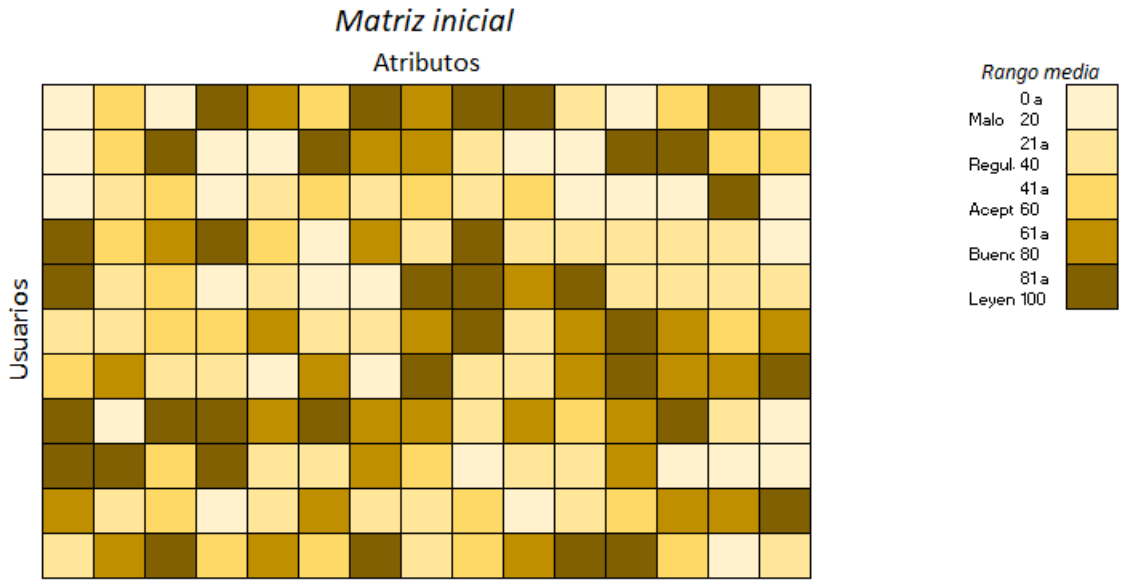
Factorización de Matrices

Factoriza una matriz en diferentes matrices que reciben información de alta calidad, es decir, información explicita del usuario o artículo. Que luego multiplica estas matrices entre sí para acercarse a la matriz inicial. En el presente caso se emplea una matriz completa de interacción de la siguiente manera.

Existe una matriz inicial que se representa en la Ilustración 52 y que por sí sola no brinda información suficiente para determinar qué próximo usuario se podría recomendar, por ende, esta matriz será llamada como matriz inicial, esta matriz contiene las medias de los jugadores ya definidas en función de sus atributos físicos.

Ilustración 53

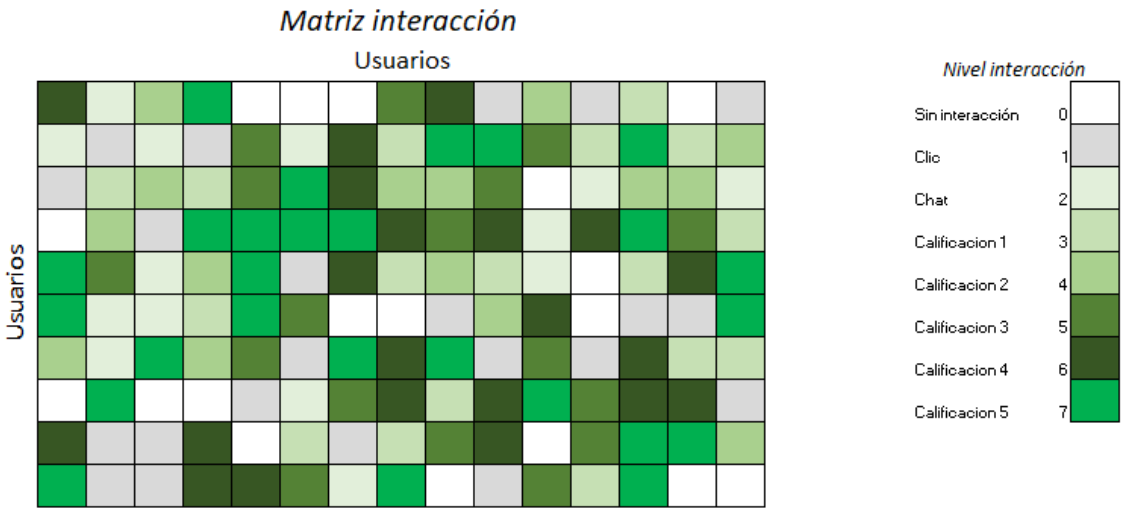
Matriz Inicial 3



Por otro lado, existe la matriz interacción representada en la Ilustración 54, que es la que almacena las puntuaciones de las interacciones que se realizan.

Ilustración 54

Matriz Interacción 3



El resultado del producto de estas dos matrices es la matriz perfilamiento representada en la Ilustración 54, que ayuda a entender los porcentajes que se asignan en función de los porcentajes iniciales mencionados

1. Habilidad (60%)

2. Sexo (20%)

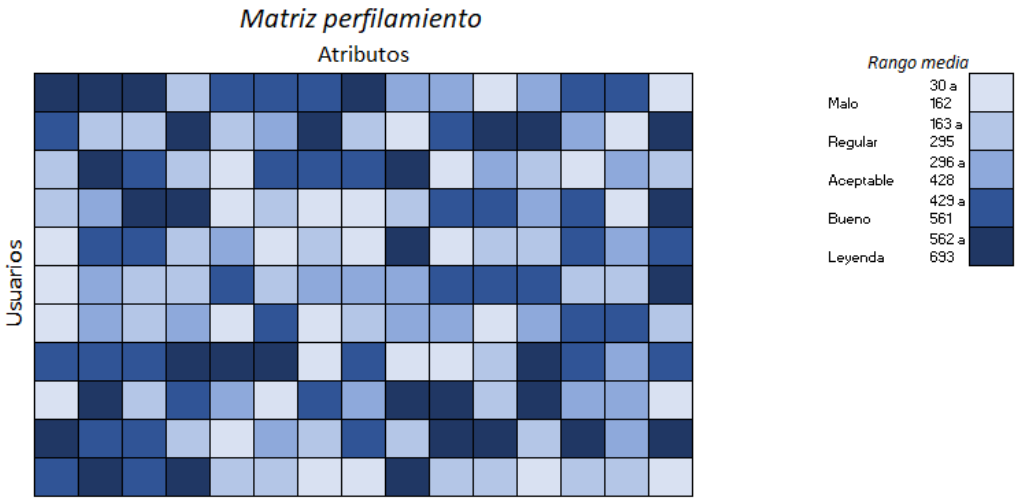
3. Disponibilidad horaria (10%)

4. Localidad preferida (10%)

En esta matriz se asignarán porcentajes con estas especificaciones:

Ilustración 54

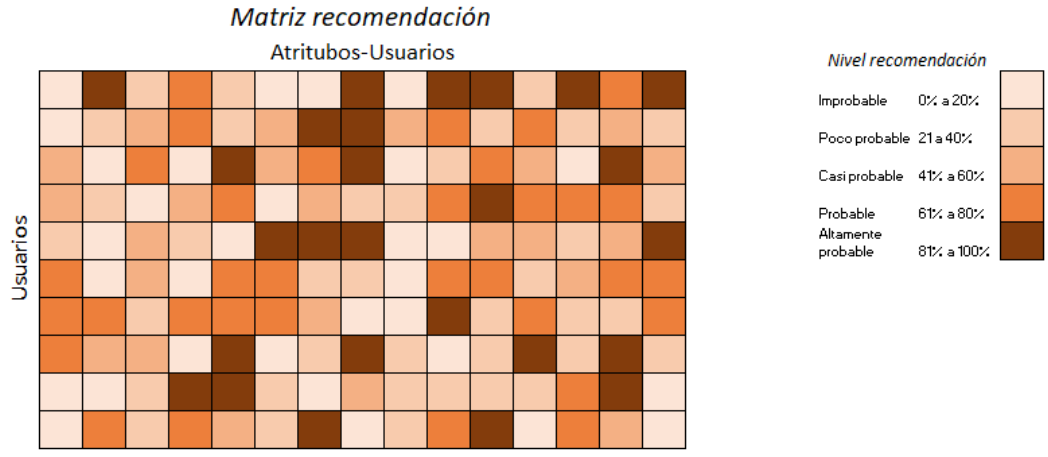
Matriz Perfilamiento 54



Ahora, el producto de estas matrices (inicial x perfilamiento) representada en la Ilustración 55, explicará que jugador es el que se va a recomendar de la lista de jugadores que se tienen en la base de datos:

Ilustración 55

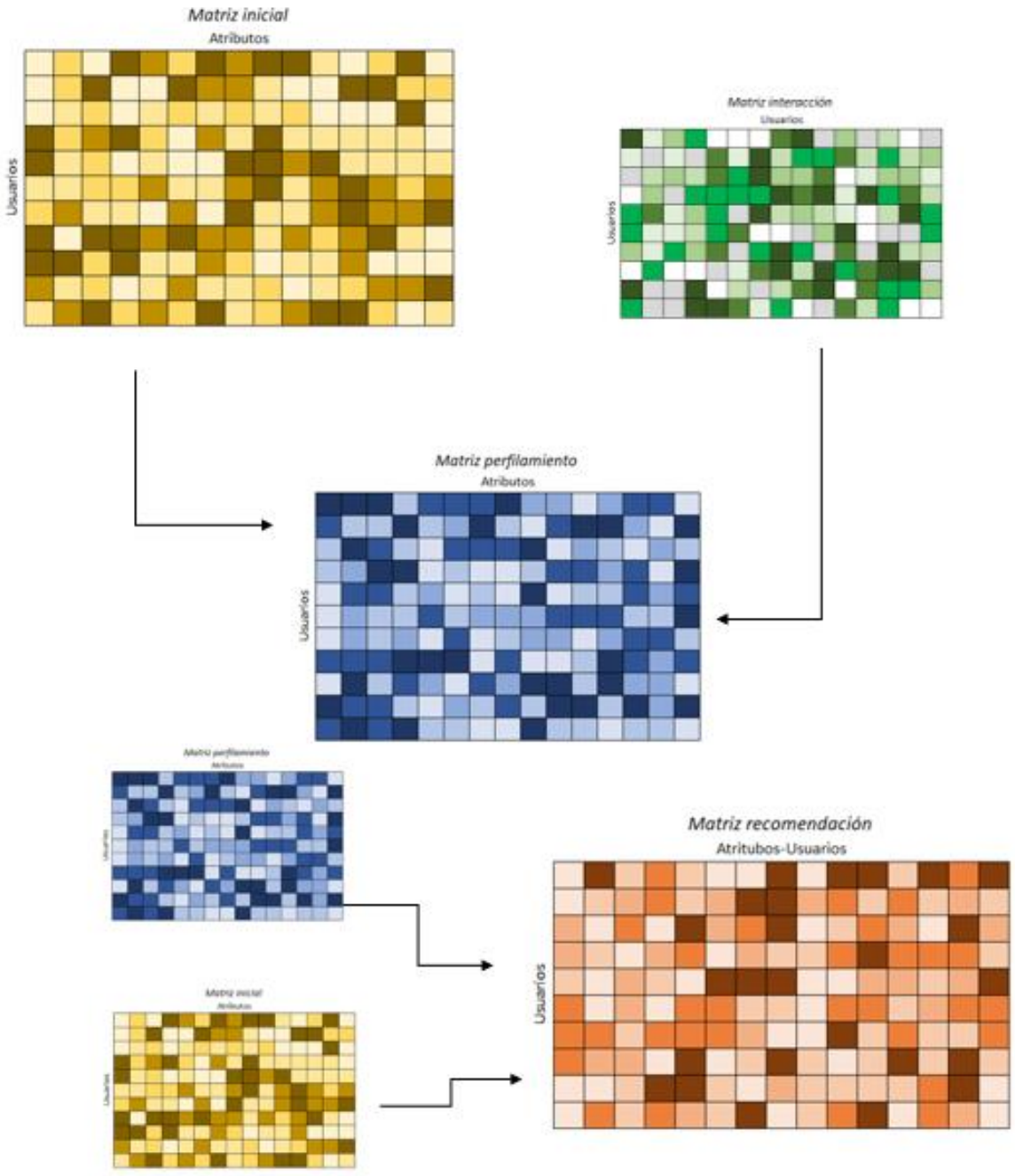
Matriz Recomendación 3



A continuación, se muestra la Ilustración 56 que proporcionará un mejor entendimiento a nivel conceptual del proceso de factorización de matrices.

Ilustración 56

Interacción de Matrices Factorizadas



La forma en que se dividen las matrices y como se realizan los cálculos están basados en la siguiente formula:

$$recomendacion = \sum_{l=1} usuario * perfilamiento$$

En donde:

- Interacciones corresponde a cualquier tipo de interacción definida.
- i corresponde a un tipo de interacción en específico: Habilidad
- j corresponde a un tipo de interacción en específico: Chat
- k corresponde a un tipo de interacción en específico: Clic
- l corresponde desde usuario 1
- usuario corresponde al usuario que califica y es evaluado.

Codificación del Algoritmo en Amazon SageMaker

A continuación, se explica cómo se realizó la codificación del algoritmo de recomendación basado en filtrado colaborativo en el servicio de Amazon SageMaker, haciendo uso de las instancias proporcionadas por este servicio, así como los Notebooks de Jupyter que se emplearán como IDE de desarrollo.

La sección del notebook de Jupyter llamada importación de librerías, como su nombre lo sugiere, importa diferentes librerías de Python que permitirán trabajar con datos y machine learning en Python, elementos que son cruciales para la codificación del algoritmo de recomendación.

Por una parte, se importa módulos de MXnet, que se puede definir como un Framework de Deep learning, el cual es usado para construir, entrenar y desplegar redes neuronales; entre estos módulos importados se encuentran las librerías gluon, nd y ndarray, los cuales son parte clave para construir y entrenar modelos de Deep Learning.

Por otro lado, se importa la librería Pandas, que es una librería de análisis de datos y permite leer, escribir y manipular datos en formatos de tablas.

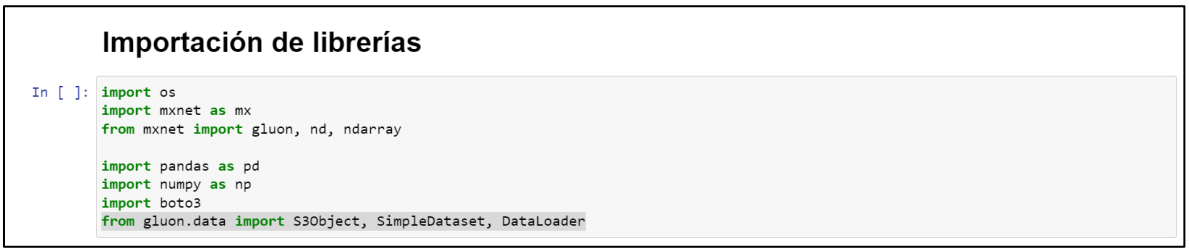
También se hace necesario importar la librería Numpy, que es una librería de procesamiento numérico y que proporciona estructuras de datos para almacenar matrices, así como una gran variedad de funciones para realizar operaciones entre matrices.

Por otra parte, la librería boto3, es una API para interactuar con servicios de AWS, lo cual se hará necesario para trabajar con los objetos almacenados en Amazon S3.

Finalmente, se importan los módulos S3Object, SimpleDataset y DataLoader del paquete gluon.data perteneciente a la librería Apache MXnet como se muestra en la Ilustración 57.

Ilustración 57

Importación de Librerías



En la siguiente sección del código es en donde se definen e inicializan las variables o hiperparamentros que se usarán para el entrenamiento del algoritmo, en este caso la variable data_path corresponde a la ruta en donde se encuentra almacenada la data que se va a utilizar en modelo de machine learning, que en este caso corresponde a una carpeta en la instancia local de Amazon SageMaker.

El siguiente hiperparametro definido es num_emb que corresponde a la capa de embedding y define al número de dimensiones que se usará para representar los datos en un formato vectorial; en la ilustración 58 se definió en 64, lo que quiere decir que cada usuario se representará dentro del modelo como un vector de 64 dimensiones.

La variable opt hace referencia al optimizador que se utilizará en el entrenamiento del modelo de ML, en este caso se definió usar Adam, el cual es un algoritmo de optimización para ajustar los parámetros de un modelo durante el entrenamiento.

El siguiente hiperparametro lr establece en qué medida se irán ajustando los parámetros del modelo de ML en la fase del entrenamiento hasta lograr una posible exactitud por parte del algoritmo, como se videncia en la ilustración 58 se estableció una tasa de 0.002 inicialmente, este valor se puede interpretar como la velocidad a la que el modelo aprenderá.

Se define también el hiperparametro mmnt que hace referencia al momento utilizado por el modelo de aprendizaje y que determina la manera en que este se ajusta a los datos para que la pérdida o error del entrenamiento se minimice y se logre una precisión y rendimiento aceptable.

El siguiente hiperparametro es batch_size el cual determina el tamaño del lote de datos que alimentará al modelo, es decir, las muestras que se utilizarán para entrenar el algoritmo se hacen de esta manera para evitar alimentar el modelo con todos los datos de una sola vez, por lo

que se divide en lotes más pequeños de manera tal que el algoritmo se pueda entrenar lote por lote, lo que garantizará un mejor rendimiento en el procesamiento del entrenamiento.

Finalmente se establece la variable ctx, que hace referencia al contexto en donde se ejecutará el procesamiento del modelo, es decir, el sistema de hardware que se utilizará, el cual, como se evidencia en la ilustración 58 definió el uso de un sistema mx.gpu(4) de la instancia de Amazon SageMaker donde se está desarrollando el algoritmo, lo que significa que se usara la GPU número 4 del sistema, esto se logra a haciendo uso del framework MXnet.

Ilustración 58

Inicialización de Variables

Inicialización de variables

```
In [ ]: data_path = 'ml-100k/'
        num_emb = 64
        opt = 'Adam'
        lr = 0.002
        mmntm = 0.
        batch_size = 50
        ctx = mx.gpu(4)
```

La siguiente celda del notebook de Jupyter (Ilustración 59), establece una función llamada download_ml_data_from_s3, la cual descarga los datos desde el Bucket de Amazon S3, sin embargo, antes de realizar la descarga, valida si esta data ya existe en el sistema de archivos local haciendo uso del método os.path.exists de la librería os, de ser así, omite la descarga, de lo contrario, procede con la descarga.

En la función se hace uso de los parámetros bucket_name que corresponde al nombre del Bucket de Amazon S3 donde se encuentran almacenados los archivos csv y del cual se obtendrá la data.

El parámetro prefix corresponde a la ruta dentro del Bucket donde se almacena la información, y local_path corresponde a la ruta local dentro de la instancia de Amazon SageMaker donde se almacenará la data.

Esto se consigue interactuando con los servicios de AWS, en concreto el servicio de almacenamiento de objetos Amazon S3 mediante el uso de la librería boto3.

Ilustración 59

Descarga de Datos



En la sección del Notebook mostrada en la Ilustración 60, se define una función llamada max_id, la cual toma como parámetro el nombre del archivo fname, luego de esto, se abre el archivo haciendo uso de la instrucción with open(fname), para después iniciar un bucle for que iterará cada línea del archivo con el objetivo de dividir cada línea en campos separados por tabulación para almacenarlo en una lista llamada tks; posteriormente se valida si esta lista contiene exactamente 9 elementos, de no ser así no la procesará y continuará con la siguiente iteración del bucle esto con el fin de tener como resultado el número total de usuarios en el archivo más uno.

Ilustración 60

Obtener Valores Máximos de Columnas en un Archivo de Datos

Obtener valores máximos de columnas en un archivo de datos

```
In [ ]: def max_id(fname):
            mu = 0
            with open(fname) as file:
                for line in file:
                    tks = line.strip().split('\t')
                    if len(tks) != 9:
                        continue
                    nu = max(mu, int(tks[0]))
                    mu = nu
            return nu + 1

        max_users = max_id(data_path + 'u.data')
```

La siguiente sección permite cargar la data de entrenamiento y pruebas desde archivos CSV que se encuentran en la ruta data_path de la instancia de Amazon SageMaker, esto se consigue haciendo uso de la librería pandas. Los primeros datos de entrenamiento se cargan desde el archivo ul.base en un DataFrame de pandas, se hace uso del parámetro header = None que establece que el DataFrame no tiene una fila de encabezado en el archivo CSV y el parámetro sep = \t indica que las columnas del archivo están separadas por tabulaciones.

De igual forma se hace el mismo proceso para la data de pruebas la cual es cargada desde el archivo ul.test.

Una vez se cargan los datos de entrenamiento y pruebas, se deben convertir en matices, y eso se consigue haciendo uso de la librería Numpy utilizando el método .values de cada uno de los objetos almacenados en el DataFrame de pandas; el índice de la columna 0 se usa para representar un identificador del usuario y a su vez, la columna 1 se selecciona como etiqueta del conjunto de entrenamiento y pruebas.

También se hace uso de la función **nd.array**, para convertir los matices de **NumPy** en arreglos de la librería Gluon con el tipo de datos np.float32 (). Las matrices de datos se asignan a

las variables **train_data** y **test_data** mientras que las matrices de etiquetas se asignan a las variables **tain_label** y **test_label** como se evidencia en la Ilustración 61.

Ilustración 61

Carga de Datos de Dos Archivos de Texto Utilizando Pandas

Carga de datos de dos archivos de texto utilizando Pandas

```
In [ ]: train_df = pd.read_csv(data_path + 'ul.base', header = None, sep = '\t')
        test_de = pd.read_csv(data_path + 'ul.test', header = None, sep = '\t')

        train_data = nd.array(train_df[[0,1]].values, dtype = np.float32)
        train_label = nd.array(train_df[2].values, dtype = np.float32)

        test_data = nd.array(test_df[[0,1]].values, dtype = np.float32)
        test_label = nd.array(test_df[2].values, dtype = np.float32)
```

El siguiente bloque de código se muestra en la Ilustración 62, es donde se define una clase llamada SparseMatrixDataset en la que se tienen tres métodos, init() que es el método constructor, getitem() el cual es usado para obtener elementos específicos de los objetos tomando idx como argumento y devolviendo tanto el primer como el segundo elemento de las filas correspondientes en self.data y self.label las cuales son variables definidas en el método constructor de la clase, en las que se almacenan los datos y etiquetas luego de una validación realizada en la que se comprueba que el número de datos y etiquetas sen el mismo; finalmente se tiene el método len() que devuelve la longitud de los datos en el objeto, es decir, la cantidad de muestras en el conjunto de datos.

Ilustración 62

Definición de la clase SparceMatrixDataset

Definición de una clase de conjunto de datos para matrices dispersas con etiquetas

```
In [ ]: class SparseMatrixDataset(guon.data.Dataset):
            def_init_(self, data, label):
                assert data.shape[0] == len(label)
                self.data = data
                self.label = label
                if isintance(label, ndarray.NDArray) and len(label.shape) ==1:
                    self._label = label.asnumpy()
                else:
                    self.label = label

            def getitem(self, idx);
                return self.data[idx, 0], self.data[idx, 1], self.label[idx]

            def _ len_(self):
                return self.data.shape[0]
```

Finalmente en la Ilustración 64, se muestra la sección del Notebook de Jupyter que define una clase llamada MFBlock que hereda de la clase gluon.block perteneciente a la librería Gluon del framework Apache MXnet y que permite establecer un bloque de construcción de un modelo de Machine Learning.

El método constructor de la clase MFBloock define e inicializa todos los atributos de la clase, así como el alcance de los nombres que se usarán para agrupar y operar las capas definidas en el bloque.

También se hace uso del método forward el cual es heredado de la clase gluon.Block, en la cual se obtienen las representaciones vectoriales de los usuarios a través de la capa de incrustación, las cuales capturan las características relevantes de los usuarios en un espacio de dimensiones reducidas, luego de esto, se aplican capas conectadas a esas incrustaciones obtenidas y que ayudarán a transformar los datos y extraer de ellos las características más significativas para finalmente, realizar la predicción final, y esto se logra multiplicando elemento a elemento los vectores de la matiz en la que están las incrustaciones de usuarios para luego

sumar los resultados y así obtener el tensor que tiene la predicción global, esta predicción representa la posible interacción entre los usuarios en el modelo de factorización matricial.

Ilustración 63

Implementación del Modelo MF

Implementación de una capa de redes neuronales para el modelo de factorización de matriz (MF)

In [ ]:
```
class MFBlock(gluon.Block):
    def __init__(self, max_users_1, max_users_2, num_emb, dropout_p=0.5):
        super(MFBlock, self).__init__()

        self.max_users_1 = max_users_1
        self.max_users_2 = max_users_2
        self.dropout_p = dropout_p
        self.num_emb = num_emb

        with self.name_scope():
            self.user_1_embeddings = gluon.nn.Embedding(max_users_1, num_emb)
            self.user_2_embeddings = gluon.nn.Embedding(max_users_2, num_emb)
            self.dropout = gluon.nn.Dropout(dropout_p)

    def forward(self, users1, users2):
        a = self.user_1_embeddings(users1)
        b = self.user_2_embeddings(users2)
        predictions = self.dropout(a) * self.dropout(b)
        predictions = nd.sum(predictions, axis=1)
        return predictions
```

Resultados

Recordando la pregunta inicial, la cual plateaba la idea de mejorar la experiencia de usuario por medio de un algoritmo de recomendación, a lo largo del proyecto se identificaron los intereses de los usuarios frente a esta idea inicial por medio de los requerimientos funcionales, junto con los antecedentes de proyectos similares, que tienen el mismo objetivo pero sin el plus agregado de conectar dichos usuarios basándose en los intereses y preferencias compartidas; que en este caso son llamadas interacciones, como el chatear, hacer clic y puntuar a un jugador, con estas interacciones ya es posible iniciar un proceso de aprendizaje en el algoritmo, puesto que se almacena la información en matrices que posteriormente se factorizan y multiplican para obtener cierto porcentaje de recomendación frente a otros jugadores. El algoritmo fue pensado desde un

inicio para que funcionará por medio de un filtrado colaborativo, el cual ayuda a obtener mejores resultados, este filtrado cuenta con algunos problemas, como lo son el arranque en frío, lo que significa que al ser un algoritmo de aprendizaje, no cuenta con la suficiente información para basarse y hacer recomendaciones acertadas, para solucionar este problema, se tuvieron en cuenta como se asignaba la media a los jugadores que se registraban; cuando un jugador se registra, se realiza un cálculo en función de sus atributos físicos y se asigna una media a cada habilidad, estas medias(60%) son almacenadas en una matriz inicial junto con otros atributos como lo son, el sexo (20%), la disponibilidad (10%) y la localidad preferida para jugar(10%) y se multiplica por la matriz de interacciones que almacena la puntuación que se le otorga a los jugadores, el resultado de esto es una nueva matriz llamada matriz de perfilamiento, hasta ahora existen 3 matrices, para obtener la última matriz se multiplica la matriz inicial por la de perfilamiento. Esta última matriz explica qué jugadores que se encuentran registrados puede ser más apto para recomendar en función de la interacción del usuario 1 que ha calificado a diferentes jugadores, no solo se recomendará a un jugador sino también a aquellos por debajo del porcentaje de recomendación más apto. Teniendo todo esto en cuenta se puede concluir que el algoritmo es viable para una posible futura implementación.

Referencias

- Ballara, Y., & Martín, C. (02 de Octubre de 2020). Aplicación para la suscripción a eventos deportivos. Valencia, España.
- Cáceres, J. G. (2008). Comunicología y fútbol. La vida social, el deporte y el espectáculo desde una perspectiva constructiva. *Revista Educación Física y Ciencia, 10*(2314-2561), págs. 15-16.
- Essenwanger, L., Humberto, V., Rojas, S., & Moisés, G. (2014). *REPOSITORIO ACADÉMICO USMP*. Recuperado el Marzo de 2023, de Sistema de gestión de eventos deportivos basado en redes sociales, aplicaciones web y móviles: https://hdl.handle.net/20.500.12727/1625
- Fubles.com. (s.f.). *Fubles.com*. Obtenido de https://www.fubles.com/es/pages/about
- Gómez, P., Guarda, T., Cedeño, J., Benavides, A., Alejandro, C., Mosquera, G., . . . Benavides, V. (Febrero de 2019). Sistemas de Recomendación: un enfoque a las técnicas de filtrado. *Revista Ibérica de Sistemas y Tecnologias información*(E18), págs. 286-293. Obtenido de https://www.proquest.com/openview/a2be2fdf640fbedd3baea0ff7e57eb9f/1?pq-origsite=gscholar&cbl=1006393
- León Bedoya, D., Macpherson Espinosa, R. M., Nuñez Solano, Y. F., & Olea Córdova, K. R. (2019). *Repositorio Académico de la UPC.* Obtenido de Aplicación móvil Gol App organizadora de partidos de fútbol (Trabajo de investigación).: https://repositorioacademico.upc.
- López, M. E. (Noviembre de 2009). Valores para una buena convivencia. *Temas para la educación revista digital para profesionales de la enseñanza, 5*(109), págs. 1-7. Obtenido de https://www.feandalucia.ccoo.es/docu/
- Orjuela, M., Villar, J. C., & Herrera, J. P. (Septiembre de 2017). *Reporte sobre la información en materia de Propiedad Intelectual en Colombia.* Obtenido de sic.gov.co: https://www.sic.gov.co/sites/default/files/files/Proteccion_Competencia/Estudios_Economicos/Documentos_elaborados_Grupo_Estudios_Economicos/Reporte-informacion-en-materia-de-Propiedad-Intelectual-en-Colombia.pdf
- Pérez Muñoz, I. (Enero de 2020). *O2, repositorio institucional de la UOC.* Recuperado el Marzo de 2023, de Desarrollo de la APP Android: Footman: http://hdl.handle.net/10609/105026

- Super Intendencia de Industria y Comercio (SIC). (Noviembre de 2022). Cartilla Formatos modelo para el cumpliemiento de obligaciones establecidas en la ley 1581 de 2012 y sus decretos reglamentarios. Colombia.
- Trasciatti, M. (12 de Octubre de 2011). Una red social para futbolistas aficionados. (B. Mundo, Entrevistador)
- Vanguardia, L. (10 de Octubre de 2011). La red social para practicar fútbol más activa de Europa. *La Vanguardia*.
- *Ley 1341 de 2009 - Gestor Normativo*. (s/f). Gov.co. Recuperado el 31 de marzo de 2023, de https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=36913
- *Ley 1356 de 2009 - Gestor Normativo*. (s/f). Gov.co. Recuperado el 31 de marzo de 2023, de https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=37645
- *Ley 1581 de 2012 - Gestor Normativo*. (s/f). Gov.co. Recuperado el 31 de marzo de 2023, de https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981
- Silvi, L. (2016). Aprendizaje Activo para mejorar el Arranque en Fr´ıo de Sistemas de Recomendaci´on. https://rdu.unc.edu.ar/bitstream/handle/11086/4125/Silvi.pdf?sequence=1