I.A Course

Proyecto Final Módulo I.A

Formulario de presentación de ideas

ⓒ2024 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of this document.

This document is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this document other than the curriculum of Samsung innovation Campus, you must receive written consent from copyright holder.

Formulario de Presentación de Ideas

**Plasme su idea de prueba rellenando este sencillo formulario.**

**POR FAVOR TOMAR EN CUENTA LAS SIGUIENTES INDICACIONES**

**CRONOGRAMA DE ENTREGAS Y CIERRE SIC 2025**

* ***Del 17 de febrero al 22 de febrero– Se deben entregar actividades faltantes de la semana 12 a la 16.***
* ***Del 24 de febrero al 01 de marzo– Se deben entregar actividades faltantes de la semana 17 a la 22.***

|  |  |  |
| --- | --- | --- |
| ***Acción*** | ***Hackathon*** | ***Cierre*** |
| ***Cierre de Venezuela*** | ***11 – 14 de marzo*** | ***14 de marzo*** |
| ***Cierre de El Salvador*** | ***24 – 26 de marzo*** | ***26 de marzo*** |
| ***Cierre de Guatemala*** | ***25 – 27 de marzo*** | ***27 de marzo*** |

**Los participantes del programa deben estar atentos por los grupos de comunicación ya que estaremos dando u ofreciendo más detalles relacionados con el cierre del programa.**

**DESARROLLO DEL PROYECTO FINAL DEL MÓDULO IA**

**PROGRAMA SAMSUNG INNOVATION CAMPUS. SIC – 2025.**

1. **Nombre y función de los miembros del equipo**

|  |  |
| --- | --- |
| Nombre de los Integrantes | Papel a desempeñar en el equipo |
| Anderson Perdomo | Como líder de este equipo me encargo de que el trabajo se haga se manera ordenada, concisa y este bien estructurado para disminuir el porcentaje de error. |
| Diego Alviarez | Soy el segundo al mando y me encargare de apoyar al líder y a los integrantes del grupo en las tareas que realicemos, daré mis opiniones e ideas para desarrollar un buen Proyecto. Soy el encargado de realizar la página web |
| Jeremy Vicent | Con todo mi esfuerzo y desempeño apoyare a mis compañeros en este proyecto. Soy el encargado de realizar el código para la limpieza del dataset. |
| Kevin Rodriguez | Con todo mi esfuerzo y desempeño apoyare a mis compañeros en este proyecto. Soy el encargado de realizar el código que muestre todas las gráficas basadas en estudios realizados. |
| Greymel Moreno | Como integrante de este grupo, daré todo mi esfuerzo y desempeño, apoyare a mis compañeros en este proyecto. Seré el encargado de finalizar todo código que requiera información extra, para darle un toque visual más agradable al proyecto. |

1. **Nombre del proyecto**

Escribe el nombre del proyecto seleccionado del banco de proyectos o propuesta que construirá tu equipo.

WildPassPro: Generador y Validador de Contraseñas Seguras con IA

1. **Nombre del equipo**

Escribe el nombre del equipo con el que se van a identificar en la realización de la propuesta

The Wild Project

1. **Descripción del proyecto**

Describa en qué consistirá el proyecto o propuesta de su equipo. Describa el problema que le gustaría resolver (incluya cualquier dato que pueda tener para apoyar que este problema existe) y el concepto general para resolver este problema).

Este proyecto no solo generará y validará contraseñas seguras, sino que también utilizará un modelo de red neuronal entrenado con TensorFlow para predecir la fortaleza de una contraseña basándose en patrones de contraseñas comunes y vulnerabilidades conocidas. El modelo aprenderá a identificar contraseñas débiles, medianas, fuertes, predecibles o comprometidas, y proporcionará recomendaciones inteligentes para mejorarlas.

1. ¿Qué valor social genera su idea (medioambiental, social, financiero, etc.)?

Nuestro proyecto WildPassPro genera un valor social significativo y multidimensional al abordar uno de los problemas más críticos de la era digital: la seguridad de las credenciales.

**1. Prevención de Ciberdelitos y Protección de Datos**

* **Proteger a individuos y organizaciones** de:
  + **Robo de identidad**: Evita que hackers accedan a cuentas personales (correos, redes sociales, banca online).
  + **Pérdidas económicas**: Mitiga fraudes financieros (acceso a cuentas bancarias).
  + **Extorsiones**: Reduce el riesgo de secuestro de cuentas (*ransomware*).
* **Datos clave**:
  + El **81% de las brechas de datos** se deben a contraseñas débiles.
  + El costo promedio de una filtración de datos es $4.45 millones.

**2. Inclusión Digital Segura**

* **Empodera a usuarios no técnicos**:
  + Genera contraseñas robustas sin necesidad de conocimientos avanzados.
  + Educa mediante análisis de IA sobre buenas prácticas de seguridad.
* **Reduce la brecha digital** al proteger a grupos vulnerables:
  + Adultos mayores (víctimas frecuentes de phishing).
  + Personas con bajo nivel de alfabetización digital.

**3. Defensa de la Privacidad**

* **Protege información sensible**:
  + Historiales médicos, comunicaciones personales, datos laborales.
  + Previene el uso malicioso de información para acoso o discriminación.

**4. Resiliencia para PyMEs y Empresas**

* **Protege infraestructuras críticas**:
  + Evita ataques a sistemas empresariales (ej.: robo de datos de clientes).
  + Reduce el riesgo de paralización operativa por ciberataques.
* **Cumplimiento regulatorio**:
  + Ayuda a cumplir normas como GDPR (UE) o CCPA (California) que exigen protección de datos.

**5. Concientización y Cultura de Seguridad**

* **Educa mediante retroalimentación de IA**:
  + Explica en lenguaje simple por qué una contraseña es débil.
  + Enseña patrones de riesgo (ej.: uso de nombres comunes, secuencias numéricas).
* **Promueve hábitos sostenibles**:
  + Fomenta la actualización periódica de contraseñas.
  + Incentiva el uso de autenticación de dos factores (2FA).

**6. Impacto Global y Ético**

* **Democratiza la seguridad**:
  + Funciona como herramienta accesible (gratuita o de bajo costo).
  + Puede adaptarse a idiomas locales para comunidades no angloparlantes.
* **Contribuye a los ODS (Objetivos de Desarrollo Sostenible)**:
  + **ODS 9**: Infraestructuras resilientes.
  + **ODS 16**: Sociedades pacíficas e inclusivas.

WildPassPro no es solo una herramienta técnica; es un escudo social que protege derechos fundamentales (privacidad, propiedad, seguridad) en el mundo digital. Al fortalecer el eslabón más débil de la ciberseguridad (las contraseñas), contribuyes a crear una sociedad digital más justa, resiliente y consciente.

1. ¿Hay alguna consideración que deba tenerse en cuenta para la comunidad (¿cómo crees que le beneficiaría a la comunidad esta idea, afecta negativamente a alguna persona?)

Nuestro proyecto WildPassPro tiene un potencial transformador para la comunidad.

**Beneficios Clave para la Comunidad**

1. **Empoderamiento Digital Masivo**
   * **Para usuarios comunes**: Les brinda herramientas profesionales de seguridad sin necesidad de conocimientos técnicos.
   * **Ejemplo**: Una persona mayor podrá generar contraseñas seguras sin depender de familiares o expertos.
2. **Reducción de la Brecha de Seguridad**
   * **Democratiza la ciberseguridad**: Ofrece protección de nivel empresarial a individuos y pequeñas empresas.
   * **Impacto**: Una PyME podría evitar un ataque de ransomware que le costaría $200,000 en promedio (Sophos, 2023).
3. **Educación en Tiempo Real**
   * **Feedback educativo**: La IA no solo bloquea contraseñas débiles, sino que explica *por qué* son peligrosas.
   * **Ejemplo**: Un adolescente aprenderá a evitar contraseñas como "amordepelícula123" tras recibir un análisis detallado.
4. **Protección de Grupos Vulnerables**
   * **Adultos mayores**: Menos riesgo de caer en estafas por contraseñas predecibles.
   * **Niños/adolescentes**: Evita el ciber acceso a sus cuentas por contraseñas como "123456".
5. **Ahorro Económico Colectivo**
   * **Reducción de costos por filtraciones**: Si el 30% de usuarios mejora sus contraseñas, se evitarían pérdidas millonarias a nivel nacional.

**Consideraciones y Riesgos Potenciales**

1. **Dependencia Excesiva en la Herramienta**
   * **Riesgo**: Usuarios podrían confiar ciegamente en la IA y descuidar otros métodos (ej.: autenticación en dos pasos).
   * **Mitigación**: Incluir recordatorios para usar múltiples capas de seguridad.
2. **Falsa Sensación de Seguridad**
   * **Riesgo**: Si la IA no detecta ciertos patrones complejos (ej.: contraseñas largas pero predecibles), usuarios podrían pensar que están protegidos cuando no es así.
   * **Mitigación**: Auditorías periódicas del modelo de IA y actualizaciones constantes.
3. **Privacidad de Datos**
   * **Riesgo**: Al analizar contraseñas, el sistema podría almacenar patrones sensibles.
   * **Mitigación**: **Cifrado extremo a extremo** y política de **"zero-knowledge"** (nunca almacenar contraseñas completas).
4. **Exclusión Digital**
   * **Riesgo**: Personas sin acceso a internet o con dispositivos obsoletos no podrían usar la herramienta.
   * **Mitigación**: Versión offline funcional y soporte para dispositivos de gama baja.
5. **Sobre restricción**
   * **Riesgo**: Si la IA es demasiado estricta, podría rechazar contraseñas válidas, pero poco convencionales (ej.: "CaféConLeche$2024").
   * **Mitigación**: Permitir ajustes manuales de sensibilidad en la configuración
6. ¿Cómo afectará a la comunidad (positiva o negativamente)?

Nuestro proyecto WildPassPro tiene el potencial de afectar positivamente a la comunidad en múltiples dimensiones, pero también es importante considerar algunos riesgos o impactos negativos que podrían surgir si no lo gestionamos adecuadamente.

WildPassPro tiene el potencial de transformar positivamente la comunidad al mejorar la seguridad digital, educar a los usuarios y proteger a los grupos más vulnerables. Sin embargo, es crucial abordar los riesgos potenciales con medidas proactivas para garantizar que el impacto sea netamente positivo. Si se implementa con responsabilidad, WildPassPro no solo prevendrá ciberataques, sino que también construirá una sociedad digital más consciente y resiliente.

1. ¿Alguna pregunta pendiente y/o suposición a la que pueda responder sobre su idea?

1. Pregunta sobre Privacidad

"Si usan contraseñas de la dark web para entrenar la IA, ¿cómo garantizan que no están exponiendo datos sensibles de usuarios reales?"

Respuesta: Todos los datos son anonimizados y provienen de brechas públicas con consentimiento legal. Eliminamos cualquier información personal (nombres, emails) antes del entrenamiento, usando solo patrones generales de contraseñas, no datos identificables.

2. Pregunta Técnica Clave

¿Cómo evitan que la IA genere contraseñas que parezcan seguras, pero sigan patrones hackeados?

Respuesta: Nuestro modelo tiene una capa de adversarial testing: se entrena contra contraseñas generadas por GANs (Redes Generativas Antagónicas) que simulan ataques de hackers. Así, aprende a evitar incluso patrones que los humanos no detectaríamos.

3. Suposición Legal

¿No es riesgoso usar datos de la dark web? ¿Cómo se protegen legalmente?

Respuesta: Trabajamos con datasets de acceso público como Have I Been Pwned, que son legales y éticos.

1. ¿Cuáles son las preguntas más importantes que quedan por responder antes de seguir desarrollando esta idea?’

1. ¿Cómo garantizan que el modelo de IA no replique patrones inseguros presentes en el dataset de la dark web?

2. ¿Qué métricas usan para medir la efectividad real de las contraseñas generadas?

3. ¿Cómo protegen los datos de usuarios almacenados en la bóveda (passwords.json.encrypted) ante un ataque directo al sistema?

4. ¿Cómo obtuvieron el dataset de contraseñas de la dark web y qué medidas toman para evitar violar leyes de protección de datos?

1. ¿Cuál es el objetivo principal o la métrica sobre la que intenta influir con esta prueba (por ejemplo, compras, valor medio de los pedidos, envío de formularios, etc.)?

Nuestro objetivo es reducir el riesgo de brechas de seguridad causadas por contraseñas débiles o predecibles, y la métrica clave sobre la que buscamos influir es el porcentaje de contraseñas vulnerables en uso (ej. débiles, reutilizadas o comprometidas).

Métricas Específicas para Medir el Impacto

1. Reducción en el uso de contraseñas débiles.

2. Adopción de contraseñas únicas.

3. Disminución de incidentes de seguridad.

1. Indica, por qué consideras que tu proyecto es novedoso

**1. Integración de IA Generativa y Ciberseguridad**

* **Novedad**: Combina modelos de lenguaje avanzados (como Llama3) con análisis técnico de contraseñas, algo que pocas herramientas hacen.
* **Ejemplo**:
  + **Solución tradicional**: Genera contraseñas aleatorias.
  + **WildPassPro**:
    - Genera contraseñas.
    - Explica en lenguaje natural por qué son seguras.
    - Recomienda mejoras contextuales (ej.: "Esta contraseña es fuerte para redes sociales, pero débil para banca online").

**2. Educación Proactiva Mediante Chatbots**

* **Innovación**: Usas IA no solo para validar contraseñas, sino para enseñar seguridad digital interactivamente.
* **Caso de uso**:
  + Un usuario escribe: "¿Por qué no debo usar mi fecha de nacimiento?"
  + El chatbot responde con ejemplos de ataques de fuerza bruta, estadísticas de filtraciones y alternativas creativas.

**3. Personalización Contextual Única**

* **Diferencial**: Adapta las contraseñas al contexto específico del usuario:
  + **Para empresas**: Longitud y complejidad según el nivel de acceso (empleado vs. administrador).
  + **Para juegos online**: Sugiere contraseñas fáciles de recordar, pero seguras.
  + **Para niños**: Genera contraseñas divertidas pero robustas (ej.: "Dino$aurio123!").

**4. Protección Activa de Grupos Vulnerables**

* **Enfoque Único**: Diseñas funcionalidades específicas para:
  + **Adultos mayores**: Interfaz simplificada con explicaciones auditivas.
  + **Personas con discapacidad visual**: Integración con lectores de pantalla y contraseñas en formato audio.
  + **Niños**: Sistema de "contraseñas familiares" para cuentas compartidas con padres.

**5. Sistema de Riesgo Predictivo**

* **Tecnología Innovadora**: Tu IA no solo analiza la contraseña actual, sino que predice futuros riesgos:
  + Detecta si una contraseña podría ser vulnerable a nuevos tipos de ataques (ej.: IA de descifrado cuántico).
  + Sugiere actualizaciones preventivas basadas en tendencias de ciberataques globales.

**6. Combate la "Fatiga de Contraseñas"**

* **Solución Creativa**: Abordas un problema psicológico poco tratado:
  + **Función "Contraseñas Memorables"**: Genera frases seguras tipo "CaféCon3Leches$Verano" en lugar de strings aleatorios.
  + **Recordatorios Inteligentes**: La IA detecta cuándo el usuario reutiliza contraseñas y ofrece alternativas sin saturarlo.

**INSTRUCCIONES PARA EL DESARROLLO DEL PROYECTO FINAL DEL MÓDULO I.A. PROGRAMA SAMSUNG INNOVATION CAMPUS. SIC – 2025.**

* Leer e identificar uno de los proyectos propuestos en el documento. Si Ud. y su equipo conformado por 5 personas como máximo. Si desean proponer un proyecto, deben presentar la propuesta a su respectivo docente/tutor del aula para hacer seguimiento del desarrollo y la entrega.
* Luego de identificar el proyecto a trabajar, desarrollarlo con las técnicas aprendidas en todo el programa Python e I.A
* Responder las preguntas de las páginas 3 y 4 de este documento.
* Todos los proyectos tienen un tema, sin embargo, los participantes deben darle una orientación, por ejemplo. Si el proyecto es de un programa de apoyo a estudiantes, se pudiera ejemplificar: "Machine, el aliado estudiantil", entre otros nombres.
* Se invita a todos los participantes a generar ideas novedosas inspiradas en las mentorías brindadas, productos e innovación del equipo de SAMSUNG.
* Todos los proyectos deben tener un enfoque tecnológico o una personalización.
* En el PPT que se encuentra cargado en la plataforma, semana 24, deben desarrollar de forma resumida su propuesta o desarrollo mostrando**: planteamiento, objetivos, impacto en el interés público, herramientas utilizadas y una breve explicación del resultado del proyecto donde se muestre la funcionabilidad del mismo.**
* Es **obligatorio que el PPT donde cargarán el desarrollo resumido del proyecto, se entreguen con la línea gráfica de SAMSUNG.**
* **Todos los grupos obligatoriamente deben grabar un video de 60 segundos como máximo y cargar en la plataforma, indicando:** 
  + Nombre del Grupo
  + Nombre del Proyecto
  + Impacto en el interés público de su proyecto: por qué creen que el proyecto es relevante e importante.
  + Debe hablar el líder del proyecto

**NOTA IMPORTANTE:**

SI UD Y SU EQUIPO DESEAN PROPONER EL DESARROLLO DE OTRO PROYECTO DEL MÓDULO PARA CUBRIR EL CONTENIDO DEL PROGRAMA, DEBEN TOMAR EN CONSIDERACIÓN LO SIGUIENTE:

* **QUE EL CONTENIDO SE ENMARQUE EN LO VISTO DURANTE EL DESARROLLO DEL PROGRAMA SIC 2025.**
* DEBEN ATENDER Y ENTENDER LAS PAUTAS DEL PROYECTO**,** PARA LUEGO CONTINUAR CON EL DESARROLLO DEL MISMO DE FORMA ORDENADA.
* NOTIFICAR AL DOCENTE/TUTOR DE AULA LA FORMA DEL DESARROLLO DE SU PROYECTO, PARA QUE EL TENGA CONOCIMIENTO Y PUEDA HACER CORRECCIONES Y EL RESPECTIVO SEGUIMIENTO DE ENTREGA.
* SI TIENE ALGUNA CONSULTA O DUDA, PUEDE DIRIGIRSE A SU DOCENTE, TUTOR O CON LA COORDINADORA DEL PROGRAMA.

**Propuestas de Proyectos de Cierre del Samsung Innovation Campus**

**SIC – 2024**

ⓒ2024 SAMSUNG. All rights reserved.

Samsung Electronics Corporate Citizenship Office holds the copyright of this document.

This document is a literary property protected by copyright law so reprint and reproduction without permission are prohibited.

To use this document other than the curriculum of Samsung innovation Campus, you must receive written consent from copyright holder.

Propuestas de Proyectos

**Escoja un proyecto de los que se muestran a continuación y personalice su formulación o en su defecto, proponga uno y evalúe su propuesta con su docente de aula.**

**PROPUESTAS DE PROYECTOS**

**Proyecto 1:** **Análisis y Visualización de Datos de Salud Pública**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 2: Detector de Emociones en Textos**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 3: Clasificación de Imágenes para Diagnóstico Médico**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 4: Sistema de Recomendación de Películas o Libros**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 5: Generador de Texto Automático**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 6: Análisis de Sentimientos en Opiniones de Productos**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 7: Simulador de Predicción Meteorológica**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 8: Traductor Automático de texto**

* Los equipos tienen total libertad de diseñar la propuesta y utilizar datos de la red.

**Proyecto 9:** **Automatización del servicio a través de chatbots**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que estén alineadas al desarrollo del programa SIC 2024

**Objetivo**: Desarrollar un **chatbot** que contenga técnicas que permitan aprender del contexto de la conversación, el cual sea capaz de ofrecer respuestas informativas en un contexto determinado ante la consulta de un usuario.

Debe implementar algún algoritmo de Machine Learning o Deep Learning y técnicas de procesamiento de lenguaje natural para generar un modelo de probabilidad que decida el conjunto de acciones a realizar en función del conjunto anterior de entradas del usuario.

**Actividades**:

* Desarrollar los flujos de conversación; con el fin de delinear la estructura de la conversación establecer a través de un diagrama qué tipo de conversación se desea iniciar, cuántas preguntas debe responder el chatbot y cuántos contextos se deben mantener entre una pregunta y otra.
* Seleccionar la herramienta de desarrollo; para la creación de un chatbot existen cientos de herramientas, en la mayoría de casos son pagas, pero también se pueden encontrar versiones gratuitas con limitaciones que podrán funcionar. A continuación se mencionan algunas herramientas.
  + Rasa
  + Python
  + Tenserflow
  + Chatterbot
* *Nota*: En caso de trabajar con una plataforma asegúrese de que pueda desarrollar parte del código y realizar el entrenamiento del modelo, consultar con su profesor esta opción.
* Aplicar técnicas de NLU.
  + Búsqueda de información para la creación de la base de conocimiento.
  + Preparar los datos de entrenamiento.
    - En caso de ser necesario, creación de:
* Respuestas
* Intenciones
* Entidades
* Implementación de un algoritmo de Machine Learning o deep learning y técnicas de procesamiento de lenguaje natural.
  + Implementar *algoritmo de* Machine Learning.
  + Implementar técnicas de procesamiento de lenguaje natural (PLN).
  + Desarrollo de acciones personalizadas que detecten y almacenen: *el nombre y número telefónico del usuario* implementando expresiones regulares.
* Desarrollo del manejo de excepciones.
* Implementación en un canal.
  + Discord
  + Telegram
  + Slack
  + Entre otros
* Evaluación del modelo.
  + Matriz de confusión.
  + Entre otros.
* Desarrollo de interfaz

# **Proyecto 10****: Análisis de Redes Sociales**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que estén alineadas al desarrollo del programa SIC 2024

**Objetivo**: Realizar el análisis de redes sociales con el fin de obtener una visión más profunda de lo que motiva la interacción, encontrar conexiones, patrones y tendencias; hacer uso de algoritmos de PLN para detectar y clasificar las emociones expresadas por los usuarios, analizar el rendimiento de las publicaciones en las redes sociales, entre otros.

**Actividades**:

* Recopilar datos de redes sociales como Twitter, Facebook o Reddit.
* Analizar tendencias de conversación, sentimientos del usuario, influenciadores clave, etc.
* Utilizar técnicas de procesamiento de lenguaje natural (NLP) para analizar el contenido de los mensajes y detectar temas relevantes.
* Implementación de un algoritmo de Machine Learning o deep learning y técnicas de procesamiento de lenguaje natural.
  + Implementar *algoritmo de* Machine Learning.
  + Implementar técnicas de procesamiento de lenguaje natural (PLN).
  + Desarrollo de acciones personalizadas que detecten y almacenen: *el nombre y número telefónico del usuario* implementando expresiones regulares.
* Desarrollo del manejo de excepciones.
* Implementación en un canal.
  + Discord
  + Telegram
  + Slack
  + Entre otros
* Evaluación del modelo.
  + Matriz de confusión.
  + Entre otros.
* Desarrollo de interfaz

# **Proyecto 11****: Análisis de Resultados de Equipos Deportivos basados en su historial de fichajes**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que estén alineadas al desarrollo del programa SIC 2024

**Objetivo**: desarrollar un modelo que permita predecir los posibles resultados del equipo basados en su historial de fichajes.

**Actividades**:

* Obtener datos de fichajes y resultados de partidos de ligas deportivas como la Premier League o la NBA.
* Análisis de rendimiento histórico: Examina el desempeño de los equipos en temporadas anteriores, incluyendo estadísticas como victorias, derrotas, empates, goles a favor, goles en contra, y posición en la tabla de la liga.
* Análisis de fichajes: Evalúa el impacto de los fichajes en el rendimiento del equipo. Esto incluye analizar el historial de transferencias de jugadores, como el costo de los fichajes, la posición en el campo, la edad y el rendimiento previo.
* Análisis de correlación: Busca relaciones entre los fichajes de jugadores y el rendimiento del equipo. Examina si ciertas características de los fichajes (por ejemplo, la posición del jugador, la edad, el rendimiento previo) están correlacionadas con el éxito del equipo.
* Modelado predictivo: Utiliza técnicas de aprendizaje automático para predecir el rendimiento futuro del equipo en función de sus fichajes históricos. Esto podría incluir la construcción de modelos de regresión para predecir la posición final en la tabla de la liga o modelos de clasificación para predecir si el equipo alcanzará ciertos hitos (por ejemplo, clasificación para torneos europeos).
* Análisis comparativo: Compara el rendimiento de diferentes equipos en función de su historial de fichajes. Examina qué equipos han tenido más éxito en el campo en relación con sus inversiones en fichajes.
* Análisis de tendencias a largo plazo: Identifica tendencias a lo largo del tiempo en el rendimiento del equipo y sus fichajes. ¿Ha habido cambios significativos en el rendimiento del equipo después de ciertos fichajes o cambios en la política de fichajes?
* Análisis de impacto económico: Evalúa el impacto económico de los fichajes en el equipo, considerando el retorno de la inversión en términos de ingresos por taquilla, patrocinios, ventas de camisetas, etc.

Algunos links que podrían ayudar:

<https://www.kaggle.com/code/ekrembayar/an-introduction-to-football-analytics>

<https://www.kaggle.com/code/ekrembayar/an-introduction-to-football-analytics#tm7>

# **Proyecto 12: Sistema de reconocimiento facial para garantizar acceso.**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que esten alineadas al desarrollo del programa SIC 2024

**Objetivo**: Desarrollar un sistema de reconocimiento facial utilizando técnicas de inteligencia artificial y aprendizaje profundo para garantizar el acceso seguro a una instalación mediante la identificación precisa de individuos autorizados.

**Actividades**:

* Extracción de características o features que permitan la identificación del personal en una instalación.
* Se debe generar un registro automático del registro ingresado a la instalación.
* Se debe optimizar la búsqueda en la base de datos para que los tiempos de reconocimiento y sean lo más cortos posibles.
* Explorar técnicas de preprocesamiento de imágenes para mejorar la calidad y consistencia de las imágenes faciales.
* Investigar y seleccionar algoritmos de reconocimiento facial basados en inteligencia artificial y aprendizaje profundo, como redes neuronales convolucionales (CNN).
* Entrenar el modelo de reconocimiento facial utilizando el conjunto de datos preparado y las técnicas seleccionadas.
* Evaluar el rendimiento del modelo mediante pruebas de validación cruzada y métricas de precisión, sensibilidad y especificidad.
* Implementar el sistema de reconocimiento facial en una aplicación o plataforma utilizando bibliotecas de visión por computadora y aprendizaje automático en Python, como OpenCV y TensorFlow.
* Integrar el sistema de reconocimiento facial con cámaras de seguridad y dispositivos de control de acceso en la instalación.
* Realizar pruebas exhaustivas del sistema en condiciones reales para verificar su precisión, velocidad y robustez.

# **Proyecto 13****: Entrenamiento de modelo predictivo de mantenimiento basado en señales de maquinaria.**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que esten alineadas al desarrollo del programa SIC 2024

**Objetivo**: Desarrollar un modelo que indique mantenimientos programados para evitar fallas inesperadas en los equipos.

**Actividades**:

* Obtener información sobre señales vibratorias y de sonido de los equipos.
* Detección de anomalías en las señales provenientes de los equipos.
* Extracción de feactures o características de las señales de medición provenientes de los equipos, algunos de los analisis para extraer características de las señales son las siguientes:
  + Transformada de Fourier.
  + Análisis en el dominio del tiempo: media, desviación estándar, varianza, curtosis etc.
  + Análisis de la envolvente.
  + Análisis de la frecuencia cepstral.
  + Análisis de tiempo-frecuencia

# **Proyecto 14**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que esten alineadas al desarrollo del programa SIC 2024

## **Fijación de precios dinámica, basada en la demanda y la oferta de un producto**

**Objetivo**: Entrenar un modelo de Machine learning que permita reconocer patrones, pronosticar situaciones futuras y actualizar precios en tiempo real.

**Actividades**:

* Implementación de un algoritmo de Machine Learning o deep learning que permita fijar precios de forma dinámica, precisa y rápida.
  + La tarifa dinámica con IA considera la demanda de un producto en el momento y el comportamiento de los usuarios, así como datos externos como noticias, clima, eventos locales, tiempo, tráfico, etc.
* Desarrollo del manejo de excepciones.
* Evaluación del modelo.
  + Matriz de confusión.
  + Entre otros.
* Implementación en un canal.
  + Discord
  + Telegram
  + Slack
  + Entre otros
* Desarrollo de interfaz

# **Proyecto 15****: Personalización del product**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que estén alineadas al desarrollo del programa SIC 2024

**Objetivos**: Recomendar el producto perfecto para cada persona. En base a datos de entrada (imagen de cuerpo, pies, entre otros) realizar un procesamiento de estos datos, para hacer referencias cruzadas con la información del producto y generar una recomendación personalizada. Un ejemplo de esta aplicación está en Nike Fit cuyo objetivo es recomendar el zapato perfecto para cada persona, a través de realidad aumentada toma como datos la medición, morfología y anatomía del píe del usuario.

**Actividades**:

* Implementación de un algoritmo de Machine Learning o deep learning que permita generar una recomendación personalizada, a partir de datos de entrada cómo lo pueden ser imagen de cuerpo, pies, entre otros.
* Desarrollo del manejo de excepciones.
* Evaluación del modelo.
  + Matriz de confusión.
  + Entre otros.
* Implementación en un canal.
  + Discord
  + Telegram
  + Slack
  + Entre otros
* Desarrollo de interfaz

# **Proyecto 16****: Curación de contenidos**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que esten alineadas al desarrollo del programa SIC 2024

**Objetivos**: En base a una imagen de referencia o palabra clave, realizar una búsqueda de contenido que recomiende resultados en tendencia o resultados que estén alineados a los gustos particulares del usuario.

**Actividades**:

* Implementación de un algoritmo de Machine Learning o deep learning que permita identificar patrones en los datos de entrada para hacer recomendaciones alineadas con la investigación a los gustos e intereses del usuario.
* Desarrollo del manejo de excepciones.
* Evaluación del modelo.
  + Matriz de confusión.
  + Entre otros.
* Implementación en un canal.
  + Discord
  + Telegram
  + Slack
  + Entre otros
* Desarrollo de interfaz

# **Proyecto 17****: Búsqueda personalizada**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que esten alineadas al desarrollo del programa SIC 2024

**Objetivos**: Clasificar los resultados de una búsqueda en orden de relevancia para cada usuario; a través del entrenamiento de un modelo con deep learning que tenga la capacidad de aprender cada vez más sobre los intereses del usuario y para comprender lo que quieren encontrar implementar un algoritmo de PLN como BERT.

**Actividades**:

* Implementación de un algoritmo de Machine Learning o deep learning y técnicas de procesamiento de lenguaje natural.
  + Implementar un *algoritmo de* Machine Learning o deep learning.
  + Implementar técnicas de procesamiento de lenguaje natural (PLN), con el fin de comprender el significado, cómo se relacionan y qué intenciones hay detrás de una búsqueda.
* Desarrollo del manejo de excepciones.
* Evaluación del modelo.
  + Matriz de confusión.
  + Entre otros.
* Implementación en un canal.
  + Discord
  + Telegram
  + Slack
  + Entre otros
* Desarrollo de interfaz

# **Proyecto 18**

* Los equipos tienen total libertad de rediseñar la propuesta o proponer otras funciones a las indicadas en el desarrollo, siempre que estén alineadas al desarrollo del programa SIC 2024

## Análisis del Caminar para la Detección de Enfermedades: Implementación de Modelos de Inteligencia Artificial

Objetivos: Desarrollar un modelo que permita la detección temprana de algunas enfermedades.

* Recopilación de datos: Obtén un conjunto de datos de videos de personas caminando, preferiblemente con etiquetas que indiquen la presencia o ausencia de enfermedades relacionadas con el caminar.
* Preprocesamiento de datos: Prepara los datos de video para su análisis, incluyendo la extracción de fotogramas clave y la segmentación de las regiones relevantes (piernas, pies, etc.).
* Extracción de características: Utiliza técnicas de procesamiento de imágenes para extraer características relevantes del caminar, como la longitud de zancada, la velocidad de paso, el ángulo de la pierna, entre otros.
* Entrenamiento del modelo: Utiliza algoritmos de aprendizaje automático, como redes neuronales convolucionales (CNN) o modelos basados en características, para entrenar un modelo que pueda predecir la presencia o gravedad de enfermedades a partir de las características del caminar.
* Validación del modelo: Evalúa el rendimiento del modelo utilizando técnicas de validación cruzada y métricas de evaluación apropiadas, como precisión, sensibilidad, especificidad, entre otras.
* Generación de resultados preliminares: Genera resultados preliminares que muestren la capacidad del modelo para detectar correlaciones entre el caminar y las enfermedades.
* Documentación y presentación: Documenta el proceso de desarrollo del modelo, incluyendo los datos utilizados, las técnicas de preprocesamiento y extracción de características, el diseño del modelo, los resultados obtenidos y las limitaciones del modelo inicial. Prepara una presentación para compartir los resultados con el equipo o los interesados.