

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN. FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA**

**Proyecto IMA I**

“**Propuesta de mejora de sistema experto**”

**Supervisado por:**

Ing. Isaac Estrada García

HORA: N2 GRUPO: 016 SALÓN: 7101

|  |  |  |  |
| --- | --- | --- | --- |
| MATRICULA | NOMBRE | SEMESTRE | CARRERA |
| 1736243 | David Alejandro Jacobi García | 8° | IAS |

SEMESTRE AGOSTO – DICIEMBRE 2022

Pedro de Alba SN, Niños Héroes, Ciudad Universitaria, San Nicolás de los Garza, N.L.

Acerca de los autores

MATERIA: PROYECTO IMA I

EQUIPO: 6 PLAN: 401 DIA: L-M-V HORA: N2

|  |  |
| --- | --- |
| **Un hombre con traje y corbata  Descripción generada automáticamente** | **David Alejandro Jacobi García 1736243**  Estoy cursando el noveno semestre de Ingeniero administrador de sistemas en FIME, Actualmente soy practicante de Command Center en el corporativo de HEB México, también tengo experiencia ensamblando computadoras y programando en el lenguaje PYTHON, mis expectativas a largo plazo es comprender los procesos del retail y poner en practica mis conocimientos en el ámbito de control de procesos y ventas digitales para un mejor desempeño en mis cargos y tareas, teniendo certificaciones internacionales**.**  **Plan 401**  **Mail:** [**Davidalejandrojacobi@gmail.com**](mailto:Davidalejandrojacobi@gmail.com) |

Contenido

[1. Resumen 4](#_Toc121371990)

[2. Introducción 4](#_Toc121371991)

[3. Antecedentes y Estado del Arte 5](#_Toc121371992)

[4. Hipótesis 5](#_Toc121371993)

[5. Propuesta 5](#_Toc121371994)

[6. Objetivos 6](#_Toc121371995)

[7. Método 6](#_Toc121371996)

[7.1 Business Understanding. 7](#_Toc121371997)

[7.2 Data Understanding 8](#_Toc121371998)

[7.3 Prepara los datos 8](#_Toc121371999)

[7.4 Modeling 8](#_Toc121372000)

[7.5 Evaluation 9](#_Toc121372001)

[7.6 Deployment. 9](#_Toc121372002)

[8. Procedimiento 10](#_Toc121372003)

[9. Resultados 11](#_Toc121372004)

[9.1 interfaz de usuario 11](#_Toc121372005)

[Conclusiones 17](#_Toc121372006)

[Conclusiones individuales 17](#_Toc121372007)

[Bibliografía 18](#_Toc121372008)

# Resumen

El plan de mejora del sistema se basa en la implementación de una red neuronal, esta red neuronal puede identificar patrones faciales a través de datos biométricos utilizando una Red neuronal convolucional, MTCNN, el principal problema contemporáneo de este sistema es poco fiable a la hora de identificar patrones faciales con tapabocas, esta problemática tiene mucho rango de mejora, Se utiliza el lenguaje de programación Python, el sistema fue programado en Python 3.9 (se esta actualizando a Python 3.10), utilizaremos un IDE llamado pycharm, este entorno de desarrollo grafico nos ayudara a implementar correctamente las librerías especializadas de redes neuronales, la razón de estas herramientas es por que el sistema ya existe, vamos a mejorarlo y documentar a través del método científico los cambios vistos y esperados, Nuestra hipótesis es que la red neuronal sea capas de identificar patrones faciales, incluso utilizando tapabocas, esperamos que la eficacia de predicción de la inteligencia artificial sea superior al 80-85% donde es mas que aceptable.

Utilizaremos una metodología de entrenamiento para la red neuronal, esto comprende de 2 bases de datos para entrenarla, una donde tenga las imágenes del usuario con tapabocas y el otro no. Nuestro proyecto no implementará nada nuevo, pues ya existe tecnología muy avanzada capaz de un reconocimiento facial superior al 90%, nuestra aportación principal será la de implementar una metodología de plan de mejora para proyectos enfocados a la inteligencia artificial, actualizando viejos sistemas y implementándolos en problemas reales.

**Keywords:** Sistema experto, MTCNN, Python, redes neuronales

Keywords: production line improvement, production data, presser, standardization.

# Introducción

Mejorar la eficacia de reconocimientos de patrones faciales utilizando datos biométricos, adaptando un sistema viejo, “que una red neuronal pueda identificar caras incluso con tapabocas”. El proyecto no es nuevo, pero puede implementar bases para futuras aportaciones a la comunidad científica, desarrollando una metodología general para poder implementar en diferentes proyectos. Nuestro proyecto no implementará nada nuevo, pues ya existe tecnología muy avanzada capaz de un reconocimiento facial superior al 90%, nuestra aportación principal será la de implementar una metodología de plan de mejora para proyectos enfocados a la inteligencia artificial, actualizando viejos sistemas y implementándolos en problemas reales.

# Antecedentes y Estado del Arte

El sistema ya fue implementado por mi (David Jacobi), este sistema fue un proyecto de implementación de la redes neuronales para el aprendizaje de sistemas expertos, en mi investigación y desarrollo del sistema tuve la oportunidad de observar y identificar problemas relacionados con la eficacia del sistema, pues depende de muchos factores físicos y naturales para el optimo desempeño del sistema, por ejemplo me di cuenta que un cambio de iluminación puede ser la diferencia entre una identificación de patrones biométricos en el sistema experto, pude observar cambios del 30% aproximadamente.

La implementación de un sistema antiguo, puede conllevar de ciertas limitaciones de a cuerdo con el grado de actualización del mismo, por ejemplo el sistema cuenta con una versión de Python 3.9, esta versión corresponde un grado significativo si es que se quiere actualizar correctamente el sistema a una nueva generación, pero de poder lograr un grado optimo de desempeño del sistema podremos comprender la estructura de las redes neuronales y la implementación de sistemas expertos, dejándonos un alto valor de conocimientos para futuros proyectos

Se pudo observar áreas de mejora con la actual problemática de la pandemia actual por el covid-19, este problema conlleva a una implementación de seguridad para el virus, lo que se conoce comúnmente como tapabocas, esta sencilla prenda facial impone otra problemática en el área de desarrollo de sistemas expertos, pues si no esta correctamente entrenado una red neuronal puede fallar en cuanto se encuentre con un problema nuevo, como lo es un usuario con tapabocas, aquí es donde actualizamos el sistema y mejoramos su desempeño.

# Hipótesis

¿Se puede adaptar un sistema antiguo de red neuronal para poder reconocer rostros con datos biométricos?

# Propuesta

Implementando un nuevo entrenamiento a la red neuronal, este aprenderá a identificar rostros incluso con el 40-60% del rostro cubierto, solucionando el problema de baja eficacia a la hora de identificar rostros con tapabocas.

# Objetivos

Vamos a implementar en 5 meses y medio una mejora del sistema experto ya existente que identifica rostros a través del entrenamiento, este sistema implementa una red neuronal convolucional, podemos implementarlo en situaciones reales como en la entrada a una institución que obligue al usuario a ingresar con tapabocas.

1. Preparar una interfaz grafica simple y amigable para que el usuario no se complique con la entrada de datos.
2. Tenemos que evaluar la eficacia antes y después de la implementación de mejora del sistema.
3. Datos biométricos simples, estos comprenden de una simple foto del rostro identificando patrones y expresiones.
4. Podemos medir un umbral de aceptación de la red neuronal, teniendo en cuenta cual porcentaje es aceptable en un entorno real donde se necesite aplicar.
5. Comparamos los dos sistemas (viejo y nuevo), estos sistemas nos ayudan a comprender el alcance real de futuros proyectos implementando la misma metodología
6. El desarrollo del sistema implica una observación real de los resultados del sistema, comprendiendo la lógica computacional del mismo, aplicándolos de forma que se mejore en la nueva problemática.

# Método

Utilizaremos una metodología de entrenamiento para la red neuronal, esto comprende de 2 bases de datos para entrenarla, una donde tenga las imágenes del usuario con tapabocas y el otro no.

Metodología CRISP-DM

Fase I. Business Understanding. Definición de necesidades del cliente final (comprensión del negocio)

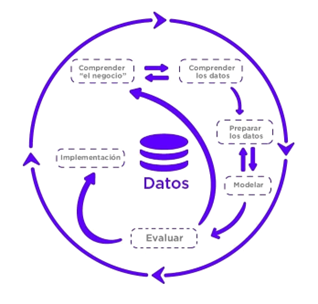
Fase II. Data Understanding. Estudio y comprensión de los datos

Fase III. Prepara los datos

Fase IV. Modeling. Modelado

Fase V. Evaluation. Evaluación (obtención de resultados)

Fase VI. Deployment. Despliegue (puesta en producción)



*Figura 1: ciclo de la metodología propuesta*

## 7.1 Business Understanding.

La fase de comprensión empresarial se centra en la comprensión de los objetivos y requisitos del proyecto. Aparte de la tercera tarea, las otras tres tareas en esta fase son actividades fundamentales de gestión de proyectos que son universales para la mayoría de los proyectos:

1. Determine los objetivos comerciales: primero debe "comprender a fondo, desde una perspectiva comercial, lo que el cliente realmente quiere lograr". (Guía CRISP-DM) y luego definir criterios de éxito empresarial.
2. Evaluar la situación: determinar la disponibilidad de recursos, los requisitos del proyecto, evaluar los riesgos y contingencias y realizar un análisis de costo-beneficio.
3. Determine los objetivos de la minería de datos: además de definir los objetivos comerciales, también debe definir cómo se ve el éxito desde una perspectiva técnica de minería de datos.
4. Producir plan de proyecto: seleccione tecnologías y herramientas y defina planes detallados para cada fase del proyecto.

Si bien muchos equipos se apresuran en esta fase, establecer un sólido entendimiento empresarial es como construir los cimientos de una casa: absolutamente esencial.

## Data Understanding

La siguiente es la fase de comprensión de datos. Agregando a la base de Business Understanding, impulsa el enfoque para identificar, recopilar y analizar los conjuntos de datos que pueden ayudarlo a lograr los objetivos del proyecto. Esta fase también tiene cuatro tareas:

1. Recopilar datos iniciales: Adquiera los datos necesarios y (si es necesario) cárguelos en su herramienta de análisis.
2. Describa los datos: examine los datos y documente sus propiedades superficiales, como el formato de los datos, la cantidad de registros o las identidades de los campos.
3. Explore los datos: profundice en los datos. Consúltelo, visualícelo e identifique las relaciones entre los datos.
4. Verifique la calidad de los datos: ¿Qué tan limpios/sucios están los datos? Documente cualquier problema de calidad.

## Prepara los datos

Esta fase, a la que a menudo se hace referencia como “recopilación de datos”, prepara los conjuntos de datos finales para el modelado. Tiene cinco tareas:

1. Seleccionar datos: determine qué conjuntos de datos se utilizarán y documente los motivos de inclusión/exclusión.
2. Limpiar datos: A menudo, esta es la tarea más larga. Sin él, es probable que seas víctima de la basura que entra y sale. Una práctica común durante esta tarea es corregir, imputar o eliminar valores erróneos.
3. Construir datos: derivar nuevos atributos que serán útiles. Por ejemplo, obtenga el índice de masa corporal de alguien a partir de los campos de altura y peso.
4. Integre datos: cree nuevos conjuntos de datos combinando datos de múltiples fuentes.
5. Formatear datos: Vuelva a formatear los datos según sea necesario. Por ejemplo, puede convertir valores de cadena que almacenan números en valores numéricos para poder realizar operaciones matemáticas.

## Modeling

Aquí probablemente construirá y evaluará varios modelos basados en varias técnicas de modelado diferentes. Esta fase tiene cuatro tareas:

1. Seleccione técnicas de modelado: determine qué algoritmos probar (p. ej., regresión, red neuronal).
2. Generar diseño de prueba: en espera de su enfoque de modelado, es posible que deba dividir los datos en conjuntos de entrenamiento, prueba y validación.
3. Modelo de compilación: por muy glamoroso que parezca, esto podría ser simplemente ejecutar unas pocas líneas de código como "reg = LinearRegression().fit(X, y)".
4. Evaluar el modelo: por lo general, varios modelos compiten entre sí y el científico de datos debe interpretar los resultados del modelo en función del conocimiento del dominio, los criterios de éxito predefinidos y el diseño de la prueba.

Aunque la Guía CRISP-DM sugiere "iterar la construcción y evaluación del modelo hasta que crea firmemente que ha encontrado los mejores modelos", en la práctica, los equipos deben continuar iterando hasta encontrar un modelo "suficientemente bueno", continuar con el CRISP. -Ciclo de vida de DM, luego mejore aún más el modelo en iteraciones futuras.

## Evaluation

Mientras que la tarea de evaluación del modelo de la fase de modelado se centra en la evaluación del modelo técnico, la fase de evaluación analiza de manera más amplia qué modelo se adapta mejor al negocio y qué hacer a continuación. Esta fase tiene tres tareas:

1. Evaluar resultados: ¿Los modelos cumplen con los criterios de éxito empresarial? ¿Cuál(es) debemos aprobar para el negocio?
2. Proceso de revisión: Revisar el trabajo realizado. ¿Se pasó algo por alto? ¿Se ejecutaron correctamente todos los pasos? Resuma los hallazgos y corrija cualquier cosa si es necesario.
3. Determine los próximos pasos: en función de las tres tareas anteriores, determine si debe continuar con la implementación, iterar más o iniciar nuevos proyectos.

## Deployment.

Un modelo no es particularmente útil a menos que el cliente pueda acceder a sus resultados. La complejidad de esta fase varía ampliamente. Esta fase final tiene cuatro tareas:

1. Planificar la implementación: desarrolle y documente un plan para implementar el modelo.
2. Supervisión y mantenimiento del plan: Desarrolle un plan completo de supervisión y mantenimiento para evitar problemas durante la fase operativa (o fase posterior al proyecto) de un modelo.
3. Producir informe final: el equipo del proyecto documenta un resumen del proyecto que puede incluir una presentación final de los resultados de la extracción de datos.
4. Revisar el proyecto: realice una retrospectiva del proyecto sobre lo que salió bien, lo que podría haber sido mejor y cómo mejorar en el futuro.

El trabajo de su organización podría no terminar ahí. Como marco de un proyecto, CRISP-DM no describe qué hacer después del proyecto (también conocido como "operaciones"). Pero si el modelo va a producción, asegúrese de mantener el modelo en producción. A menudo se requiere un monitoreo constante y un ajuste ocasional del modelo.

# Procedimiento

1. Preparar una interfaz grafica simple y amigable para que el usuario no se complique con la entrada de datos.
2. Comprender el objetivo del proyecto.
3. Estudio y comprensión de los Datos biométricos simples, estos comprenden de una simple foto del rostro identificando patrones y expresiones.
4. Preparar y ordenar los datos capturados y modelarlo.
5. Comparamos los dos sistemas (viejo y nuevo), estos sistemas nos ayudan a comprender el alcance real de futuros proyectos implementando la misma metodología y evaluamos.
6. Lanzamos el sistema

Diagrama

Descripción generada automáticamente

* *Figura 2: Diagrama de flujo del sistema*

# Resultados

El proyecto tuvo resultados prometedores y esperados de acuerdo a nuestra hipótesis inicial, también obtuvimos resultados concluyentes, con estos resultados podemos concluir el proyecto.

## 9.1 interfaz de usuario

* Esta interfaz tuvo como objetivo el de cumplir las necesidades de registrar e iniciar sesión para el usuario. “Rápida, intuitiva, confiable.”

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente

* *Figura 3: interfaz de usuario*

La interfaz en el apartado de registro

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente

* *Figura 4: interfaz de usuario - registro*

Se programo un registro para el usuario con inteligencia artificial.

Inicio de sesión.

* Este es el apartado mas importante del programa, aquí es donde demostramos la confiabilidad y eficacia del programa.

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente

* *Figura 5: interfaz de usuario – login*

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente

* *Figura 6: interfaz de usuario – registro demo*
* *Imagen que contiene persona, tabla, frente, joven

  Descripción generada automáticamente*
* *Figura 7: interfaz de usuario – registro demo*

Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente

* *Figura 8: Comprobación de registro exitoso.*

*Imagen de la pantalla de un video juego

Descripción generada automáticamente con confianza baja*

* *Figura 9: Comprobación de inicio exitoso – sin tapabocas.*

*Imagen que contiene Tabla

Descripción generada automáticamente*

*Interfaz de usuario gráfica, Sitio web

Descripción generada automáticamente*

* *Figura 10: Comprobación de inicio exitoso – con tapabocas.*

*Imagen que contiene Tabla

Descripción generada automáticamente*

# Conclusiones

Tuvimos una compatibilidad con la identificación de la imagen registrada 97.91% esto sin cubrebocas. Aparte de tener un índice alto de compatibilidad los datos biométricos se muestran estables cuando se requieran utilizar para otro tipo de usuario. Mientras que con cubrebocas tuvimos un índice de compatibilidad con la foto de registro del 99.39% esto demuestra que el proyecto de mejora tuvo un alto porcentaje de contabilidad, pero también hay que aceptar que los resultados fueron mediante el ambiente de desarrollo así que no es 100% fiable ya en producción.

## Conclusiones individuales

David Alejandro Jacobi García

Bien si es un punto clave la metodología de una mejora de algún sistema también es importante saber el contexto se tuvo para obtener dichos resultados, aparte de tener el objetivo propuesto para saber si van por buen camino o si es que vale la pena seguir con el proyecto, mi proyecto de mejora continua fue en parte posible al contexto del covid-19, tenia una idea y la pude implementar de forma que se adaptara a la problemática que se vivía en esos días, el programa en si no tuvo mas percances, a la hora de desarrollar una mejora tuvimos como propuesta la mejora de la metodología en base a la minería de datos, creando un modelo de datos para obtener los resultados esperados, también obtuvimos conclusiones certeras pero solo en nuestro ambiente de pruebas, para tener un ambiente de producción es complicado con las herramientas actuales que se tienen, ya sean personas o metadatos, así que como conclusión solo puedo decir que es muy interesante esta metodología para obtener datos resultante de un modelo de datos y partiendo de la minería de datos, esto claro teniendo un sistema viejo el cual se quiere mejorar, así que como conclusión tengo la experiencia para futuros proyectos en donde puedo implementar la misma metodología y considerar un cambio significativo para bien, aumentando los resultados satisfactorios obtenidos.

# Bibliografía

1. Shearer C., *el modelo CRISP-DM: el nuevo plan para la minería de datos,* almacenamiento de los datos J (2000); 5:13-22.
2. [↑](https://es.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining#cite_ref-2) [What IT Needs To Know About The Data Mining Process](https://www.forbes.com/sites/metabrown/2015/07/29/what-it-needs-to-know-about-the-data-mining-process/" \l "2065f3a3515f) Publicado por Forbes, el 29 de julio de 2015, recuperado el 24 de junio de 2018