



**Instituto Politécnico Nacional
Unidad Profesional Interdisciplinaria de
Ingeniería Campus Zacatecas**

**Área de ubicación para el desarrollo del
trabajo**

Ingeniería en Sistemas Computacionales.

Línea de investigación
Inteligencia Artificial.

Título del proyecto de Trabajo Terminal
Aplicación web para la clasificación de cáncer
de mama usando Deep Learning.

Presenta(n):

Lucia Iveth De La Vega Hernández.

Fernando Ruiz Correa.

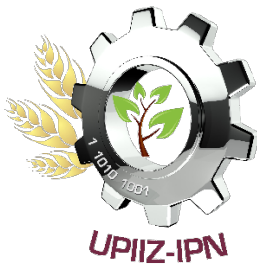
Jorge Ulises Zapata Arteaga.

Director:

Mayra Alejandra Torres Hernández.

Asesores:

Sergio Valadez Godínez.



Zacatecas, Zacatecas a 09 de Septiembre de 2025

Índices.

Índice de contenido

Descripción del proyecto.	1
Objetivo general del proyecto.....	1
Objetivos particulares del proyecto.	1
Marco metodológico.....	2
Cronograma de actividades.	8
Bibliografía.	8
Firmas.	10
Autorización.	10

Índice de figuras

Figura 1. Cronograma de actividades trabajo terminal I	8
Figura 2. Cronograma de actividades trabajo terminal II.....	8

Descripción del proyecto.

El presente proyecto consiste en el diseño e implementación de una aplicación web que emplea técnicas de Deep Learning utilizando redes neuronales convolucionales para la clasificación de posibles casos de cáncer de mama. Dicha aplicación procesará imágenes mamográficas con el objetivo de identificar patrones, anomalías u otro tipo de indicio asociados a la presencia de cáncer. Proporcionando los resultados en la interfaz web de la aplicación.

Objetivo general del proyecto.

Clasificar posibles casos de cáncer de mama utilizando técnicas de Deep Learning, desarrollando modelos multimodales, integrando diferentes perspectivas de una mamografía mediante el uso de una aplicación web.

Objetivos particulares del proyecto.

- Optimizar la calidad de las imágenes de rayos X mediante el pre-procesamiento de imágenes para aumentar la exactitud de la clasificación de las mamografías.
- Clasificar imágenes de rayos X mamarios usando técnicas de Deep Learning, con el fin de apoyar la detección temprana de posibles casos de cáncer de mama, mediante la implementación de una aplicación web.
- Entrenar redes neuronales convolucionales para la clasificación de casos de cáncer de mama.
- Analizar estadísticamente de los resultados de las redes neuronales convolucionales para evaluar el rendimiento del modelo.
- Visualizar los resultados de la clasificación para facilitar su interpretación por parte del personal medico mediante el desarrollo de una aplicación web.

Marco metodológico.

Actualmente, la ingeniería en software se ha visto afectada por numerosos problemas como la falta de confiabilidad, respuesta deficiente al cambio, agilidad limitada y costos excesivos. [2] Es por eso que surgen conceptos llamados "Metodologías ágiles", estas se caracterizan por su enfoque que pone como prioridad la colaboración, el lanzamiento rápido de software y la participación de los clientes, lo que da como resultado productos finales eficientes para los interesados. [1]

CRISP-DM.

La metodología CRISP-DM es una de las más empleadas actualmente para el desarrollo de proyectos de minería de datos. En 1997, se puso en marcha bajo el financiamiento del Programa de Investigación y Desarrollo en Tecnologías de información de la Unión Europea por 5 empresas. Actualmente IBM es la principal empresa que promueve el uso de esta metodología. [3]

La metodología CRISP-DM es un método probado para orientar trabajos de minería de datos, incluyendo descripciones de las fases de un proyecto, las tareas necesarias en cada fase y una explicación de las relaciones entre las tareas.

Esta metodología consiste en 6 fases que indican las dependencias más importantes y frecuentes entre fases. [4]

1. Entendimiento del negocio.

La primera etapa consiste en identificar el objetivo del proyecto y las necesidades del proyecto. Esta es la etapa más importante ya que, si no se tiene una correcta comprensión del problema, o negocio, de nada servirán las etapas siguientes. [4] Las principales actividades clave son:

- Definir los objetivos del proyecto.
- Identificar los criterios de éxito.
- Traducir las metas del negocio a metas de minería de datos.

- Obtener un plan de proyecto.[5]

2.Entendimiento de los datos.

Se hace la recolección de los datos, teniendo en cuenta que ninguna información importante quede fuera. Se identifican las variables y los datos estadísticos necesarios como la media, el error estándar, etc. [3] Las principales actividades clave de esta fase son:

- Recolección inicial de datos, obteniendo los datos a utilizar en el proyecto e identificando las fuentes y las técnicas empleadas en la recolección de los mismos.
- Gestionar la calidad de los datos, identificando problemas y proporcionando soluciones. [5]

3. Preparación de los datos.

Esta es la etapa que consume una mayor cantidad de tiempo. Con los datos ya recolectados se organizan y seleccionan de tal manera que sea posible observarlos de manera correcta.[3] Las principales actividades clave son:

- Eliminar inconsistencias, o errores en los datos. Aplicando diferentes técnicas como la normalización y el tratamiento de duplicados.
- Transformaciones en los datos, como normalización de los datos. Con el objetivo de cambiar el formato o estructura de ciertos datos sin modificar su significado.
- Construir un juego de datos apto para ser usado. [5]

4. Modelado.

En esta fase se construye un modelado que permita alcanzar los objetivos del proyecto. Las principales actividades de esta fase son:

- Seleccionar las técnicas de modelado más adecuadas. Se seleccionan las técnicas apropiadas de acuerdo con el problema a resolver y las herramientas disponibles como los algoritmos.
- Fijar una estrategia de verificación de la calidad del modelo.

- Construir el modelo a partir de las técnicas seleccionadas realizando un proceso iterativo de modificación de parámetros. [3]

5. Evaluación.

En esta fase se verifica si el modelo cumple con los objetivos del negocio o proyecto. Para esto, se realizan las siguientes actividades:

- Evaluar el modelo o modelos generados hasta el momento.
- Revisar todo el proceso de minería de datos que se ha realizado hasta el momento.
- Se establecen los siguientes pasos a tomar, si se trata de repetir fases o abrir nuevas líneas de investigación. [5]

6. Implementación.

Esta etapa explota, mediante acciones concretas, el conocimiento adquirido mediante el modelo. Aquí también es importante documentar los resultados de manera clara para el usuario final y asegurarse de que todas las etapas de la metodología se documenten debidamente para hacer una revisión del proyecto a fin de obtener lecciones aprendidas durante el proceso. [3] Además, se realiza la implementación del modelo en un entorno real de forma que sea propagado a los usuarios finales, así como el mantenimiento del software. En esta fase se realizan las siguientes actividades:

- Implementación del modelo o realización del proyecto.
- Realizar el seguimiento y mantenimiento del software con nuevos datos. [5]

KANBAN.

Dentro de este amplio abanico de metodologías encontramos Kanban, una metodología ágil que se implementa a través de tableros. Un tablero Kanban muestra el trabajo de un proyecto en forma de tablero organizado por columnas pudiendo ser tan simples como "Atraso", "En progreso" y "Listo", en estas columnas hay actividades que son representadas por tarjetas, promoviendo un flujo de trabajo eficiente, visualizando el progreso actual de las tareas y eliminando cuellos de botella. Kanban se ha convertido en una

metodología muy usada ya que ayuda al equipo a tener una comprensión clara de sus actividades actuales, aumentando la productividad al reducir el tiempo dedicado a tareas improductivas. [3]

Ahora bien, Kanban presenta algunos desafíos importantes a la hora de implementarse, ya que se necesita una coordinación y cohesión de equipo para obtener los beneficios completos de la metodología [3]. Sin embargo, dentro de sus beneficios encontramos: limitación de trabajo en proceso, monitoreo y control del proceso de producción, programación visual, mejora del flujo de trabajo, capacidad de respuestas, adaptabilidad y flexibilidad, facilita una alta producción, previene la sobreproducción, mejora la utilización de la capacidad y reduce el tiempo de producción. [4]

En cuanto a la implementación de Kanban en Proyectos de Desarrollo Software encontramos que: en un trabajo de aprendizaje de vocabulario básico de inglés realizado por la Universidad Nacional del Chimborazo, se concluyó que Kanban fue fundamental para mejorar la productividad y la transparencia del equipo en el desarrollo de software. Específicamente en este ámbito, Kanban demostró una gestión más eficiente y efectiva de las actividades, permitiendo concentrarse en las actividades prioritarias y completar el trabajo de manera más rápida y satisfactoria [6]. Así como también en diversos proyectos basados en la metodología Kanban se respaldan hechos como cumplimiento satisfactorio de requerimientos y la refinación del proyecto paso a paso. [6]

Así pues, para el desarrollo del proyecto se propone la integración de dos metodologías que nos serán de ayuda en la realización organizada de este: CRISP-DM para la orientación técnica en el ciclo de vida del modelo Deep Learning, y Kanban como metodología para la gestión del trabajo en equipo.

Aplicación de la metodología de UPIIZ.

En el presente proyecto se empleará una metodología estructurada que contempla la planificación, desarrollo, control y cierre del mismo. Para ello, se seguirán los siguientes pasos: la definición del plan metodológico y plan de proyecto, donde se establecerán los objetivos, alcance y cronograma; la elaboración de la matriz de trazabilidad, que permitirá verificar la cobertura de los requisitos en todas las etapas; la especificación de requisitos de

software (SRS), en la cual se documentarán de manera detallada los requerimientos funcionales y no funcionales; la gestión de riesgos, identificando posibles amenazas y estrategias de mitigación; la documentación del diseño del sistema, donde se representará la arquitectura y los modelos necesarios para su construcción; la planificación y diseño de pruebas, con los casos y criterios para validar el sistema; y finalmente, la presentación del reporte final, integrando los resultados obtenidos y conclusiones del proyecto [7].

Aplicación de la metodología CRISP-DM.

La metodología CRISP-DM será la guía para el desarrollo del modelo de clasificación, implementando sus 6 fases de la siguiente manera:

Entendimiento del negocio: Se definirán los objetivos generales, particulares y los resultados esperados que nos indicarán el índice de éxito de nuestro proyecto junto a los asesores. Además de identificar los requerimientos del proyecto, y posteriormente encontrar una manera de migrar la idea inicial de negocio hacia un enfoque sistemático que cumpla con los requerimientos previamente definidos.

Así mismo, se ideará un plan de acción para llevar a cabo las actividades necesarias de desarrollo que permitirán cumplir con los objetivos planteados; además de un plan de riesgo.

Entendimiento de los datos. En esta etapa se realizará la recolección y análisis de mamografías provenientes de una base de datos pública, identificando sus características, posibles sesgos y vistas.

Preparación de datos. Se aplicará preprocesamiento de imágenes como eliminación de ruido, normalización, rotación, aumento o reducción de escala, entre otras. Esto con la finalidad de mejorar la calidad del conjunto de entrenamiento.

Modelado. Elegir el modelo Deep Learning que se implementará la arquitectura de redes neuronales convolucionales (CNN). Así mismo, se definirá la estrategia de verificación de eficacia del modelo con parámetros variables iterativos.

Evaluación. Se validará el rendimiento del modelo mediante métricas dentro del ámbito médico de acuerdo con el diagnóstico de un radiólogo. En dado caso de fallas en el

modelo o que surja la necesidad de modificación, se seguirá el plan de contingencia o riesgo de acuerdo con la situación que se presente.

Implementación. Se integrará el modelo entrenado en la aplicación web, desplegándolo en entornos reales, acompañado de la documentación necesaria para su uso por parte de personal médico.

Aplicación de la metodología KANBAN.

Por otro lado, la metodología Kanban permitirá organizar y monitorear las actividades del equipo a lo largo de la metodología CRISP-DM. Para ello se implementará un tablero digital haciendo uso de la herramienta “Trello”, organizado en columnas que representarán el estado de las actividades, los cuales serán: Pendiente, En Progreso y Terminado.

Cada actividad será representada mediante una tarjeta que se moverá a lo largo del tablero de acuerdo con el estado de la actividad.

Tomando en cuenta que, para pasar de fase de la metodología CRISP, es necesario haber finalizado todas las actividades de la fase precedente, es decir, que todas las actividades planteadas estén en la columna Terminado.

La combinación entre CRISP-DM y Kanban nos puede ayudar a equilibrar la técnica necesaria para el procesamiento de datos y la organización visual que el proyecto requiere. Mientras CRISP-DM nos proporciona el rigor técnico de la metodología para el ciclo de clasificación, Kanban nos facilita la coordinación del equipo, optimizando tiempos y recursos para cumplir los objetivos del proyecto.

Cronograma de actividades.

A continuación, se presenta el plan de trabajo que será realizado mediante una gráfica de Gantt, la cual será utilizada para la administración de las acciones realizadas o por realizar durante el desarrollo del proyecto; haciendo uso de la herramienta “Instagantt”.

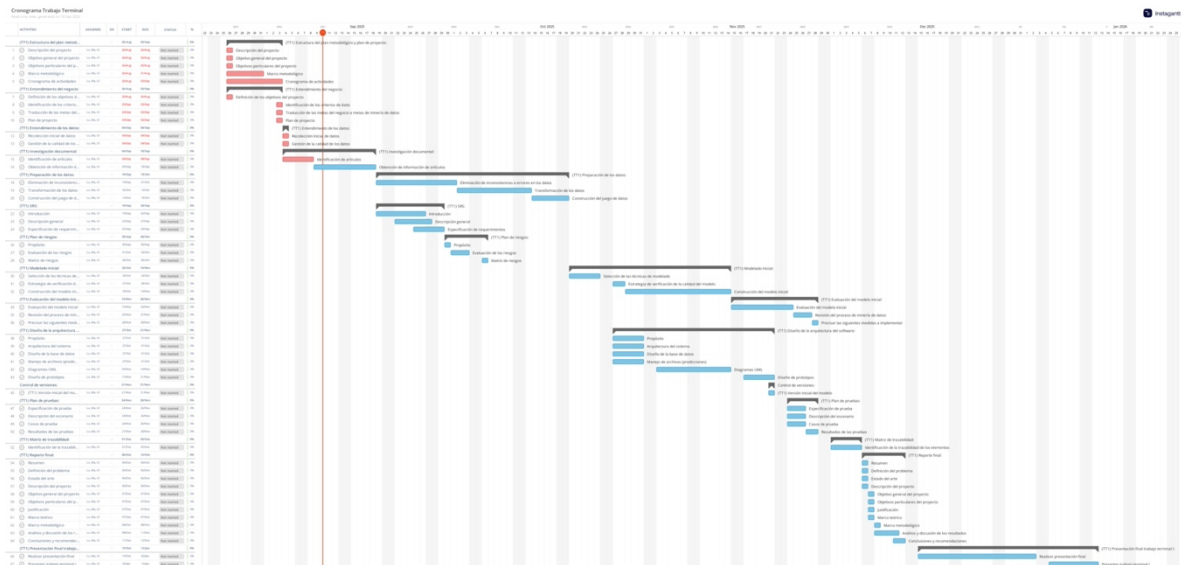


Figura 1. Cronograma de actividades trabajo terminal I

En la figura 1 se muestra el cronograma de actividades correspondiente a trabajo terminal I, el cual presenta una lista de actividades a realizar dividido por secciones, las personas responsables de realizar cada actividad y la fecha de inicio y fin de cada sección en conjunto de sus actividades.

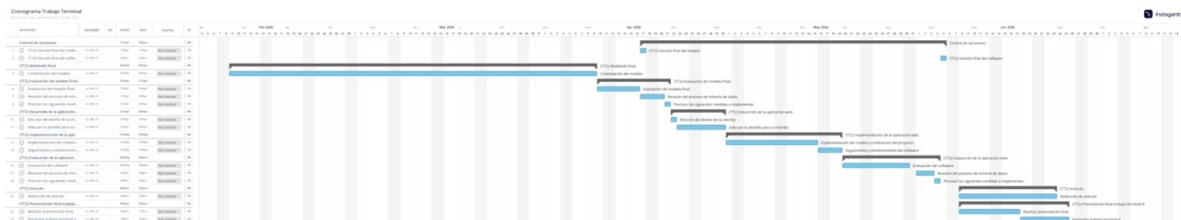


Figura 2. Cronograma de actividades trabajo terminal II

En la figura 2 se observa el cronograma con las actividades correspondientes a trabajo terminal II, detallando cada sección, persona responsables y fechas destinadas a cada actividad.

Bibliografía.

- [1] M. Castillo Anzules and E. J. Guaña Moya, "Kanban: Una metodología ágil para la gestión eficiente del flujo de trabajo en el desarrollo de software, una revisión sistemática," 2024.
- [2] M. O. Ahmad, D. Dennehy, K. Conboy, and M. Oivo, "Kanban in software engineering: A systematic mapping study," *Journal of Systems and Software*, vol. 137, pp. 96–113, 2018.
- [3] Zúñiga, J. J. E. (2020). Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. *Ingeniería Investigación y Tecnología*, 21(1), 1-13. <https://doi.org/10.22201/fi.25940732e.2020.21n1.008>.
- [4] Roberto, C. (2022, 31 mayo). _Crisp-DM: las 6 etapas de la metodología del futuro_. Blog MBA Esalq USP. <https://blog.mbauspesalq.com/es/2022/05/31/crisp-dm-las-6-etapas-de-la-metodologia-del-futuro/>
- [5] Rueda, J. F. V. (2019, 4 noviembre). _CRISP-DM: una metodología para minería de datos en salud - healthdataminer.com_. healthdataminer.com. <https://healthdataminer.com/data-mining/crisp-dm-una-metodologia-para-mineria-de-datos-en-salud/#:~:text=Como%20ejemplo%20de%20exploraci%C3%B3n%20visual,diferentes%20diagn%C3%B3sticos%20de%20los%20tumores.&text=Los%20histogramas%20sugieren%20que%20el,el%20%C3%A1rea%20y%20el%20per%C3%ADmetro.&text=de%20componentes%20principales-,La%20finalidad%20es%20observar%20c%C3%B3mo%20se%20distribuye%20cada%20uno%20de,diferentes%20diagn%C3%B3sticos%20de%20los%20tumores>
- [6] S. X. Zhuño Quizhpe, "Desarrollo de un software educativo aplicando la metodología Kanban, para el aprendizaje de vocabulario básico de inglés," Bachelor's thesis, Universidad Nacional de Chimborazo, Riobamba, 2023.
- [7] “Trabajo Terminal 1.” https://www.poluxis.com/trabajo_terminal/trabajo_terminal1.html

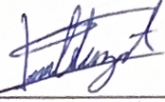
Firmas.



Lucia Iveth De La Vega Hernández.



Fernando Ruiz Correa.



Jorge Ulises Zapata Arteaga.

Autorización.

Por medio del presente autorizo la impresión y distribución del marco metodológico y cronograma de actividades, toda vez que lo he leído, comprendido en su totalidad, y estar de acuerdo con su desarrollo.

Atentamente;



Mayra Alejandra Torres Hernández.



Sergio Valadez Godínez.