

Universidad Siglo 21



Trabajo final de grado: Ingeniería en software.

Sistema de Diagnóstico Predictivo de Células Anómalas mediante
Imágenes Médicas.

Autor: Matías Ezequiel Cativa Sonzini.
Legajo: SOF00624

Córdoba, noviembre 2021

Índice

<i>Índice de tablas</i> _____	- 2 -
<i>Índice de imágenes</i> _____	- 3 -
<i>Resumen</i> _____	- 4 -
<i>Abstract</i> _____	- 5 -
<i>Título</i> _____	- 6 -
<i>Introducción</i> _____	- 6 -
Antecedentes: _____	- 6 -
Descripción del área problemática: _____	- 6 -
<i>Justificación</i> _____	- 7 -
<i>Objetivo General Del Proyecto</i> _____	- 7 -
<i>Objetivo Específico Del Proyecto</i> _____	- 7 -
<i>Marco Teórico Referencial</i> _____	- 8 -
Dominio Del Problema _____	- 8 -
TIC _____	- 8 -
Competencia _____	- 9 -
<i>Diseño Metodológico</i> _____	- 9 -
Herramientas Metodológicas _____	- 9 -
Herramientas De Desarrollo _____	- 9 -
Recolección De Datos _____	- 10 -
Planificación _____	- 10 -
<i>Relevamiento</i> _____	- 11 -
Relevamiento estructural _____	- 11 -
Relevamiento funcional _____	- 13 -
<i>Diagnóstico y Propuesta</i> _____	- 14 -
Diagnóstico _____	- 14 -
<i>Nombre del proceso: Análisis y redacción de los resultados de la muestra del paciente.</i>	- 14 -
<i>Problemas</i> _____	- 14 -
Propuesta _____	- 15 -
<i>Objetivo, Límites y Alcance del Prototipo</i> _____	- 15 -
Objetivo Del Prototipo _____	- 15 -

Límites _____	- 15 -
Alcances _____	- 15 -
<i>Descripción del Sistema</i> _____	- 16 -
Product Backlog _____	- 16 -
Historia De Usuario _____	- 16 -
Sprint Backlog _____	- 19 -
Estructura De Datos _____	- 20 -
Prototipo De Interfaces De Pantallas _____	- 21 -
Diagrama De Arquitectura _____	- 25 -
<i>Seguridad</i> _____	- 26 -
<i>Análisis de Costos</i> _____	- 27 -
Análisis de costo de desarrollo _____	- 28 -
Análisis de costos operativos _____	- 28 -
Resumen de costos _____	- 29 -
<i>Análisis de Riesgos</i> _____	- 30 -
Estimación De Probabilidad: _____	- 30 -
Estimación De Impacto: _____	- 30 -
Exposición al riesgo: _____	- 31 -
Plan De Contingencia: _____	- 32 -
<i>Conclusiones</i> _____	- 32 -
<i>Demo</i> _____	- 33 -
<i>Referencias</i> _____	- 33 -
<i>Anexos</i> _____	- 34 -

Índice de tablas

<i>Tabla 1: relevamiento tecnológico.</i> _____	- 12 -
<i>Tabla 2: Diagnóstico</i> _____	- 14 -
<i>Tabla 3: Product Backlog</i> _____	- 16 -
<i>Tabla 4: Historia de usuario 001</i> _____	- 16 -
<i>Tabla 5: Historia de usuario 002.</i> _____	- 17 -
<i>Tabla 6: Historia de usuario 003.</i> _____	- 17 -
<i>Tabla 7: Historia de usuario 004.</i> _____	- 17 -
<i>Tabla 8: Historia de usuario 005.</i> _____	- 18 -
<i>Tabla 9: Historia de usuario 006</i> _____	- 18 -
<i>Tabla 10: Historia de usuario 007.</i> _____	- 18 -
<i>Tabla 11: Sprint Backlog 01</i> _____	- 19 -
<i>Tabla 12: Sprint Backlog 002.</i> _____	- 19 -

<i>Tabla 13: Análisis de costo de desarrollo.</i>	- 28 -
<i>Tabla 14: Análisis de costos operativos.</i>	- 28 -
<i>Tabla 15: Resumen de costos.</i>	- 29 -
<i>Tabla 16: Estimación de Probabilidad.</i>	- 30 -
<i>Tabla 17: Estimación de impacto.</i>	- 30 -
<i>Tabla 18: Exposición al riesgo.</i>	- 31 -
<i>Tabla 19: Plan de contingencia.</i>	- 32 -

Índice de imágenes

<i>Ilustración 1: tabla Gantt</i>	- 10 -
<i>Ilustración 2: diagrama de Gantt</i>	- 11 -
<i>Ilustración 3: ubicación del consultorio médico.</i>	- 12 -
<i>Ilustración 4: Organigrama.</i>	- 13 -
<i>Ilustración 5: DER</i>	- 21 -
<i>Ilustración 6: Prototipo completo.</i>	- 22 -
<i>Ilustración 7: Prototipo 01</i>	- 22 -
<i>Ilustración 8: Prototipo 02</i>	- 23 -
<i>Ilustración 9: Prototipo 03</i>	- 23 -
<i>Ilustración 10: Diagrama de arquitectura.</i>	- 26 -

Resumen

Este proyecto contribuyó en brindarle una herramienta a médicos especializados en anatomía patológica para que ellos puedan realizar sus tareas en el menor tiempo posible y de mejor calidad. Esto fue llevado a cabo por un sistema que, mediante inteligencia artificial, reconoce patrones de células malignas en imágenes de histologías médicas.

Actualmente, estos profesionales médicos no cuentan con herramientas que les facilite su trabajo y todo el proceso de toma de datos de pacientes, realización de informe y entrega del mismo es manual.

Por medio del proceso de desarrollo de software, se encontró una solución que, haciendo uso de nuevas tecnologías, se pueda obtener un impacto tecnológico en el proceso de brindar resultado, digitalizando por completo.

Para llevar adelante esa solución, se tuvo que obtener, analizar y limpiar muchos datos que fueran útiles a la hora de entrenar el algoritmo que realiza los informes, gracias a la buena predisposición de los profesionales de la salud se pudo determinar con rapidez la calidad de los datos.

Palabras clave: anatomía patológica, inteligencia artificial, histología.

Abstract

This project contributed to providing a tool to doctors specialized in pathological anatomy so that they can perform their tasks in the shortest possible time and of better quality. This was carried out by a system that, using artificial intelligence, recognizes patterns of malignant cells in images of medical histology. Currently, these medical professionals do not have tools that facilitate their work and the entire process of taking patient data, making a report and delivering it is manual. Through the software development process, a solution was found that, using new technologies, can obtain a technological impact in the process of delivering results, digitizing completely. To carry out this solution, it was necessary to obtain, analyze and clean up a lot of data that was useful when training the algorithm that makes the reports. Thanks to the good predisposition of the health professionals, the quality of the data could be quickly determined.

Keywords: pathological anatomy, artificial intelligence, histology.

Título

Sistema de diagnóstico predictivo de células anómalas por imágenes.

Introducción

Este proyecto fue llevado a cabo en la ciudad de Córdoba, provincia de Córdoba en Argentina.

Con el desarrollo del mismo se buscó generar una herramienta para médicos en la especialidad de anatomía patológica, que les facilite el trabajo a la hora de diagnosticar que se encontraron células anómalas en una muestra.

Antecedentes:

La inteligencia artificial (IA) se refiere a los sistemas o las máquinas que imitan la inteligencia humana para realizar tareas y que tienen la capacidad de mejorar iterativamente a partir de la información que recopilan. (Web de Oracle s.f)

La inteligencia artificial (IA) se está aplicando para muchos casos dentro de la medicina, como por ejemplo el software que desarrolló Google para la detección temprana de cáncer de mama, llamado Google Health. El estudio realizado arrojó los siguientes resultados "una reducción absoluta del 5,7% y 1,2% en falsos positivos y de un 9,4% y 2,7% en falsos negativos" (Xataka, 2020). Otro caso de Google Health en el que usó inteligencia artificial en la medicina es para ayudar a prevenir la ceguera ocasionada por diabetes. (Google Health, s.f).

Descripción del área problemática:

Los médicos especializados en anatomía patológica requieren de una herramienta que mejore la calidad de sus trabajos actualizando el proceso de detección de células cancerígenas de las muestras provenientes de los pacientes. Ya que, con el método actual de detección, se les dificulta poder devolver todos los informes a tiempo y esto conlleva a que tengan que trabajar horas extras, bajando el nivel de productividad.

Justificación

Se hizo necesario contar con una herramienta para automatizar el proceso de predicción de células anómalas en las muestras de diferentes pacientes. De esta manera, se pudo reducir el trabajo de los médicos especialistas y se ayudó a la detección temprana de enfermedades.

Este proyecto tiene un impacto en el campo tecnológico de forma que se usa el software para la medicina innovando en el proceso de la detección de enfermedades a nivel celular, pasando de un proceso manual a uno automatizado dándole valor al trabajo de los especialistas ayudando a la precisión de a la hora de detectar enfermedades y reduciendo las horas que este proceso conlleva.

Objetivo General Del Proyecto

Desarrollar un sistema de información que permita capturar, analizar y procesar imágenes de muestras de pacientes a nivel celular facilitando la obtención automatizada de resultados.

Objetivo Específico Del Proyecto

- Analizar los procesos de detección de enfermedades en pacientes.
- Estudiar la situación entre los médicos especialistas y las herramientas con la que cuentan para la detección de enfermedades, definiendo problemas.

- Recopilar información de las diferentes células que se van a analizar.
- Analizar las entradas para determinar la configuración óptima del modelo matemático para las predicciones.

Marco Teórico Referencial

Dominio Del Problema

Lo primero que se debe definir para analizar el dominio del problema es que es la inteligencia artificial. “La Inteligencia Artificial (IA) es la combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano.” (Iberdrola, s.f)

TIC

Es necesario mencionar las tecnologías que se usaron para el sistema y la explicación de cada una de ellas.

Python es un lenguaje de programación potente y fácil de aprender. Tiene estructuras de datos de alto nivel eficientes y un simple pero efectivo sistema de programación orientado a objetos. La elegante sintaxis de Python y su tipado dinámico, junto a su naturaleza interpretada lo convierten en un lenguaje ideal para scripting y desarrollo rápido de aplicaciones en muchas áreas, para la mayoría de plataformas. (Python, s.f)

Streamlit es una biblioteca de python que hace fácil crear aplicaciones web para mostrar resultados de tus análisis de datos. (Pybonacci, s.f)

SQL por sus siglas en inglés significa Lenguaje de Consulta Estructurada (Structured Query Language), es un lenguaje de programación diseñado para actualizar, obtener, y calcular información en bases de datos relacionales. (Developer Mozilla, s.f)

Competencia

No se encontraron competencias que realicen este trabajo.

Diseño Metodológico

Herramientas Metodológicas

El desarrollo del software se realizó bajo la metodología ágil Scrum.

La web para la difusión de la gestión ágil de proyectos proyectosagiles.org define Scrum como:

‘Scrum es un proceso en el que se aplican de manera regular un conjunto de buenas prácticas para trabajar colaborativamente, en equipo, y obtener el mejor resultado posible de un proyecto. Estas prácticas se apoyan unas a otras y su selección tiene origen en un estudio de la manera de trabajar de equipos altamente productivos.’

Herramientas De Desarrollo

Para el desarrollo de este proyecto se utilizaron diferentes tecnologías, tanto para la capa de presentación, Front End, como en la capa de acceso a los datos, Back End.

Para la primera, implica, Streamlit, una librería de Front End que ayuda a la creación de aplicaciones web para mostrar resultados de los análisis de datos.

Para la segunda capa (Back End), incluye Python que tiene diferentes frameworks nativos que facilitan el uso de Machine Learning.

Los datos van a ser administrados mediante SQL, tanto los datos necesarios para el Machine Learning como los datos necesarios para la aplicación.

Recolección De Datos

Desde el punto de vista del proyecto, la técnica de observación ayudó a percibir el estado actual del proceso de diagnóstico y permitió relevar su comportamiento, tanto desde el momento que el paciente llega al consultorio médico, hasta el momento de retirar el resultado.

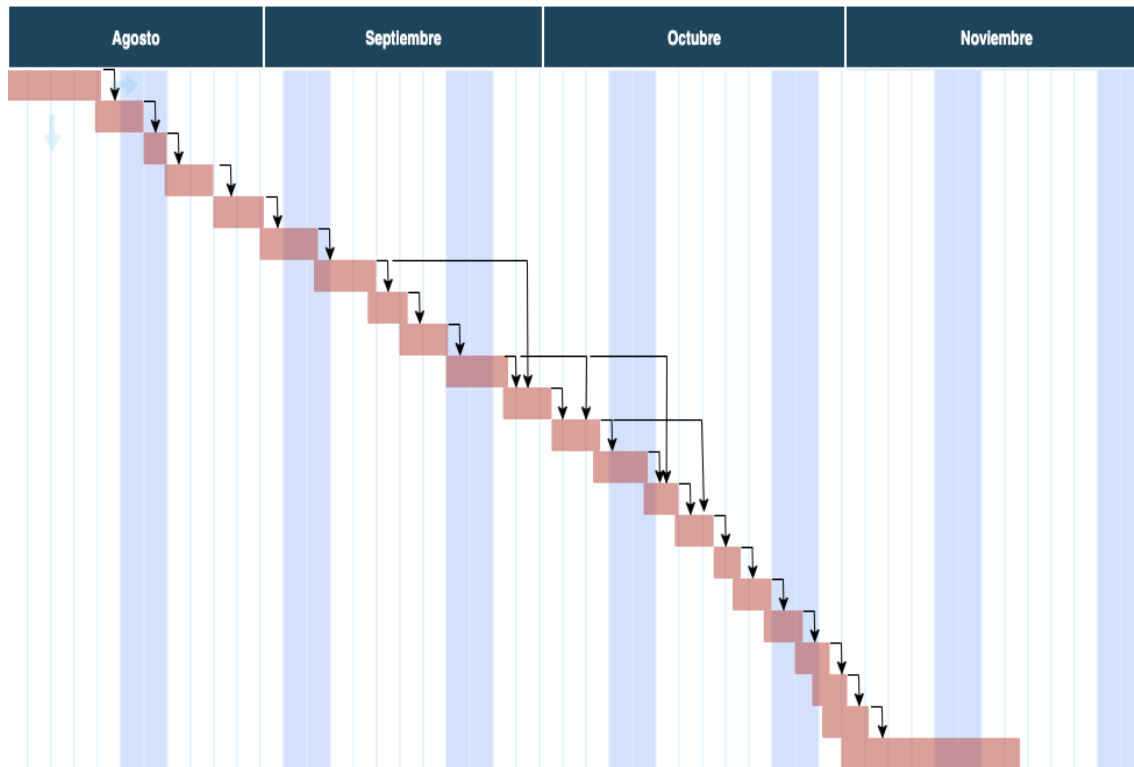
Todos los datos recolectados fueron tomados por observación, compartiendo días de trabajo con un médico especialista en su consultorio.

Planificación

Ilustración 1: tabla Gantt

	Nombre tarea	Duración	Comienzo	Fin	Predecesoras
1	Selección tematica	10 días	09/08/2021	19/08/2021	
2	Introducción	3 días	19/08/2021	21/08/2021	1
3	Justificación	2 días	21/08/2021	23/08/2021	2
4	Objetivo	3 días	23/08/2021	26/08/2021	3
5	Marco teórico referencial	5 días	26/08/2021	31/08/2021	4
6	Diseño Metodológico	4 días	31/08/2021	04/09/2021	5
7	Relevamiento	8 días	04/09/2021	12/09/2021	4
8	Diagnóstico	3 días	12/09/2021	16/09/2021	7
9	Propuesta	3 días	16/09/2021	19/09/2021	7,8
10	Objetivo	4 días	19/09/2021	23/09/2021	9
11	Límite y Alcance	3 días	23/09/2021	26/09/2021	10
12	Historias de usuario	3 días	26/09/2021	29/09/2021	10,11
13	Diagramas	5 días	29/09/2021	03/10/2021	10,11
14	Seguridad	2 días	03/10/2021	05/10/2021	13
15	Análisis de riesgos	4 días	05/10/2021	09/10/2021	13,12
16	Análisis de costos	3 días	09/10/2021	12/10/2021	15
17	Plan de contingencia	5 días	12/10/2021	17/10/2021	16
18	Conclusiones	4 días	17/10/2021	21/10/2021	18
19	Resumen	2 días	21/10/2021	23/10/2021	18
20	Indice	2 días	23/10/2021	25/10/2021	19
21	Referencias	4 días	25/10/2021	29/10/2021	21
22	Codificacion Prototipo	17 días	29/10/2021	14/11/2021	

Ilustración 2: diagrama de Gantt



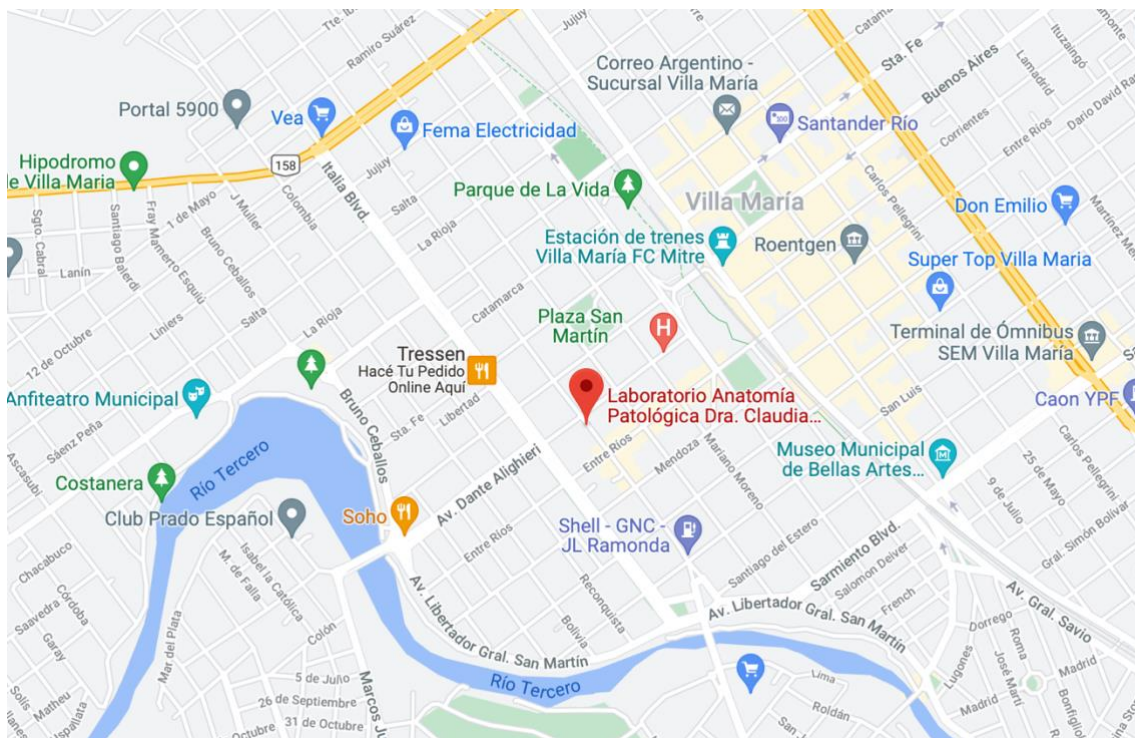
Fuente: elaboración propia.

Relevamiento

Relevamiento estructural

Laboratorio de anatomía patológica Dra. Claudia Sonzini, es un laboratorio médico especializado en la anatomía patológica ubicado en la ciudad de Villa María, Córdoba, Argentina. Donde se realizan extracción y video de colposcopia, análisis de biopsias y piezas quirúrgicas.

Ilustración 3: ubicación del consultorio médico.



Fuente: elaboración propia.

En cuanto a tecnología relevante que cuenta la organización se describe a continuación:

Tabla 1: relevamiento tecnológico.

	Cantidad/Descripción	Procesadores	Memoria RAM	Disco Duro	S.O
PC. Secretaria	1 pc de escritorio para el área de atención al cliente	Core i3	8GB	120 GB	Windows 10
PC. Depto. Medico	1 pc de escritorio para el departamento médico.	Core i7	16GB	500 GB	Windows 10
	Cantidad/	Marca	Modelo	Descripción	

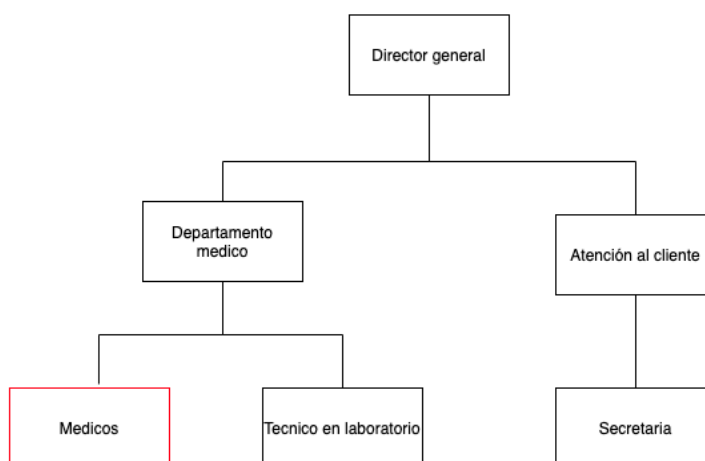
Microscopios	1 microscopio óptico Primo Star	Zeiss	SF20	El microscopio permite tomar imágenes de la muestra y almacenarla en la computadora.
	1 microscopio óptico Olympus	Olympuse	CH30	

Fuente: elaboración propia.

Relevamiento funcional

A continuación, se presenta el organigrama de la organización relevada:

Ilustración 4: Organigrama.



Fuente: elaboración propia.

En la actualidad cuenta con una secretaria, un técnico en laboratorio, tres médicas, una de ella se encarga de las extracciones de muestras para los Papanicolaou y dos que son especialistas en anatomía patológica.

Como puede observarse en la ilustración anterior, el área involucrada en el proceso de definición de los objetivos cumple la siguiente función:

Médicos: Son los encargados de ver las muestras mediante el microscopio y realizar el informe donde se redacta el resultado de las muestras.

A continuación, se describe el proceso relevante para el objetivo definido:

- Nombre del proceso: Análisis y redacción de los resultados de la muestra del paciente.
- Roles: Medico.
- Ejecución: El médico encargado del análisis de la muestra, toma la biopsia y la introduce en el microscopio. La biopsia cuenta con un id numérico y una orden del médico que deriva al paciente donde tiene datos relevantes. El médico especialista en anatomía patológica en una hoja en blanco escribe el id y datos relevantes del paciente mientras empieza a observar la biopsia. En esta misma hoja empieza a anotar datos que sean útiles para luego redactar el informe. Luego de ver la muestra y tomar todos los datos necesarios el médico especialista empieza a confeccionar el informe en la computadora.

Diagnóstico y Propuesta

Diagnóstico

Tabla 2: Diagnóstico

Nombre del proceso: Análisis y redacción de los resultados de la muestra del paciente.	
Problemas	Causas
1. El médico escribe en diferentes oportunidades el informe dando oportunidad al error.	<ol style="list-style-type: none"> 1. Falta de tecnología en el proceso de desarrollo del informe. 2. Falta de hábito de uso tecnológico.
3. Mucho tiempo de redacción de informes.	<ol style="list-style-type: none"> 1. Se redacta un informe a mano y luego se pasa a computadora. 2. Puede que a veces los datos del informe a mano no coincidan con los datos del paciente.
3. Margen de error humano.	1. A pesar de ser médicos especialistas y experimentados siempre hay un margen de error humano impredecible a la hora de diagnosticar.
4. No puede medir el error.	1. El informe al ser algo que puede cambiar dependiendo del punto de

	vista del médico, no se puede saber el margen de error.
--	---

Fuente: elaboración propia.

Propuesta

El sistema que se desarrolló ayudó a la realización del diagnóstico y del informe de manera automatizada. Esto se logró mediante tecnología de inteligencia artificial que reconoce las imágenes de las biopsias y con esto encuentra patrones en las células que son anómalas. El software al contar con información proveniente de la base de datos y de su entrenamiento, puede determinar qué células son cancerígenas, brindar información relevante para el médico, redactar un informe y mostrar el margen de error que obtuvo al realizar todo el proceso automatizado para que luego el médico especialista decida si los resultados obtenidos son confiables.

El software, es una aplicación web para que pueda ser accedido desde cualquier dispositivo con conexión a internet y pueda ser ejecutado por cualquier especialista que cuente con la tecnología para hacerlo.

Objetivo, Límites y Alcance del Prototipo

Objetivo Del Prototipo

Desarrollar un prototipo de sistema que ayude al desarrollo seguro, confiable y preciso de informes médicos mediante inteligencia artificial.

Límites

Desde que el médico ingresa la imagen de la biopsia hasta que se entrega el informe.

Alcances

- Registro de los datos de la biopsia.
- Carga de la imagen de la biopsia.
- Verificación de los resultados.
- Control del informe creado.
- Envío del informe vía email.

Descripción del Sistema

Product Backlog

Tabla 3: Product Backlog

ID	Historia de usuario	Prioridad	Puntos de historia	Dependencias
HU-001	Carga de datos del paciente	Alta	8	
HU-002	Cargar la imagen	Alta	8	HU-001
HU-003	Analizar la imagen	Alta	13	HU-002
HU-004	Realizar el informe	Alta	13	HU-003
HU-005	Visualización del informe	Media	5	HU-004
HU-006	Confirmación del informe por parte del profesional.	Media	5	HU-005
HU-007	Envío del informe vía email	Baja	3	HU-006

Fuente: elaboración propia.

Historia De Usuario

Tabla 4: Historia de usuario 001

ID	HU-001	Nombre	Carga de datos del paciente
Descripción	Como usuario quiero cargar los datos del paciente para poder analizar las imágenes.		

Criterios de aceptación	<ol style="list-style-type: none"> 1. Dado un campo incompleto cuando el usuario intenta guardar los datos, entonces, el sistema dará aviso de que todos los campos son requeridos. 2. Dado un campo incorrecto cuando el usuario intente guardar los datos, entonces, el sistema dará aviso que cierto campo es incorrecto. 		
Prioridad	Alta	<i>Puntos estimados de historias.</i>	8

Fuente: elaboración propia.

Tabla 5: Historia de usuario 002.

ID	HU-002	Nombre	Cargar la imagen
Descripción	Como usuario quiero cargar la imagen tomada por el microscopio de la biopsia para poder ser analizada por el sistema.		
Criterios de aceptación	<ol style="list-style-type: none"> 1. Dada una imagen corrompida cuando el usuario intenta analizarla, entonces el sistema dará aviso de que la imagen tiene un error. 2. Dada una imagen con demasiada calidad cuando el usuario intenta analizarla, entonces el sistema dará aviso de que la imagen es demasiado pesada. 3. Dada una imagen que no sea la que el sistema espera cuando el usuario intenta analizarla, entonces, se le preguntará si desea analizar de igual manera. 		
Prioridad	Alta	<i>Puntos estimados de historias.</i>	8

Fuente: elaboración propia.

Tabla 6: Historia de usuario 003.

ID	HU-003	Nombre	Analizar la imagen
Descripción	Como usuario quiero que el sistema analice la imagen de manera automatizada para poder ver el resultado.		
Criterios de aceptación	<ol style="list-style-type: none"> 1. Dada una imagen previamente cargada, cuando ocurra un error al analizarla, entonces, se le preguntará si desea re intentarlo. 		
Prioridad	Alta	<i>Puntos estimados de historias.</i>	13

Fuente: elaboración propia.

Tabla 7: Historia de usuario 004.

ID	HU-004	Nombre	Realizar el informe
Descripción	Como usuario quiero que el sistema realice el informe automáticamente para tener el resultado.		
Criterios de aceptación	<ol style="list-style-type: none"> 1. Dada una imagen analizada, cuando ocurra un error al analizarla, entonces, se le preguntará si desea re intentarlo. 		

	2. Dado una imagen analizada, cuando ocurra un error inesperado, entonces, mostrará un aviso de error.		
Prioridad	Alta	Puntos estimados de historias.	13

Fuente: elaboración propia.

Tabla 8: Historia de usuario 005.

ID	HU-005	Nombre	Visualización del informe
Descripción	Como usuario quiero ver el informe realizado de manera autónoma por el sistema para poder validarlo.		
Criterios de aceptación	<ol style="list-style-type: none"> 1. Dado un informe realizado previamente de manera autónoma, cuando ocurra un error, entonces se le preguntará si desea reintentarlo. 2. Dado un informe realizado previamente de manera autónoma, cuando el sistema lo muestre, entonces también visualizará el margen de error. 		
Prioridad	Media	Puntos estimados de historias.	5

Fuente: elaboración propia.

Tabla 9: Historia de usuario 006

ID	HU-006	Nombre	Confirmación del informe por parte del profesional.
Descripción	Como usuario quiero validar el informe realizado de manera autónoma por el sistema para poder darle al paciente un informe confiable.		
Criterios de aceptación	<ol style="list-style-type: none"> 1. Dado un informe realizado previamente de manera autónoma, cuando el usuario encuentre un error, entonces se le preguntará si desea editarlo. 2. Dado un informe realizado previamente de manera autónoma, cuando el usuario apreté “aceptar”, entonces el sistema guardará el informe. 		
Prioridad	Media	Puntos estimados de historias.	5

Fuente: elaboración propia.

Tabla 10: Historia de usuario 007.

ID	HU-007	Nombre	Envío del informe vía email
Descripción	Como usuario quiero mandar el informe realizado vía email al paciente para que el paciente lo tenga de manera más rápida.		
Criterios de aceptación	<ol style="list-style-type: none"> 1. Dado un email que no exista o que no esté cargado correctamente, cuando el usuario apreté “enviar vía email”, entonces, el sistema mostrará un mensaje “el email ingresado es incorrecto”. 		

	2. Dado un email que no exista o que no esté cargado correctamente, cuando el usuario apreté “enviar vía email” entonces, el sistema mostrara la opción de imprimir el informe.		
Prioridad	Baja	<i>Puntos estimados de historias.</i>	3

Fuente: elaboración propia.

Sprint Backlog

Tabla 11: Sprint Backlog 01

Sprint	Historia de usuario	ID	Tareas	Prioridad	Estimado	Estado
1	HU-001 Carga de datos del paciente	01	Diseñar diagramas correspondientes al módulo.	Alta	2 días	Hecho
		02	Codificar módulo correspondiente a la historia de usuario.	Alta	4 días	Hecho
		03	Codificar la base de datos correspondiente a la historia de usuario	Alta	3 días	Hecho
		04	Diseñar e implementar la interfaz gráfica.	Alta	4 días	Por hacer
		05	Integrar módulo al sistema.	Media	2 días	Por hacer
		06	Realizar testing sobre el módulo integrado.	Alta	1 día	Por hacer

Fuente: elaboración propia.

Tabla 12: Sprint Backlog 002.

Sprint	Historia de usuario	ID	Tareas	Prioridad	Estimado	Estado
		01	Adquisición de datos de aprendizaje	Alta	4 días	Por hacer

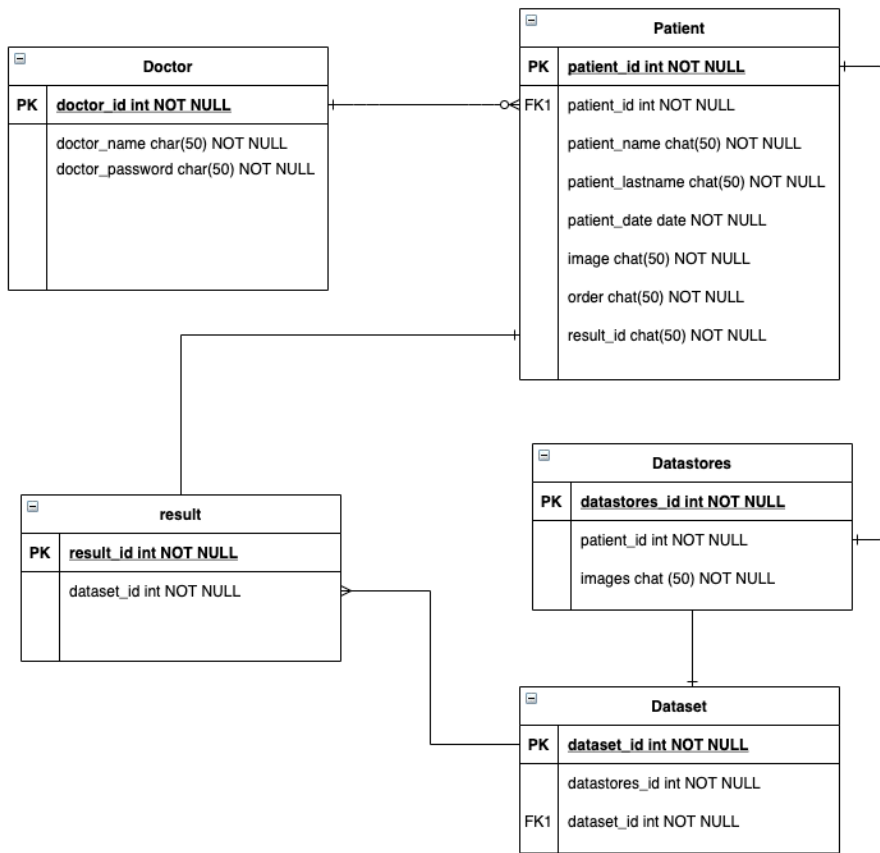
3	HU-004 Realizar informe	02	Preparar y limpiar los datos.	Alta	5 días	Por hacer
		03	Analizar, explorar los datos	Alta	3 días	Por hacer
		04	Elegir un modelo de aprendizaje.	Alta	4 días	Por hacer
		05	Construir el modelo.	Alta	5 días	Por hacer
		06	Realizar testing sobre el módulo integrado.	Alta	1 día	Por hacer

Fuente: elaboración propia.

Estructura De Datos

Diagrama de entidad relación:

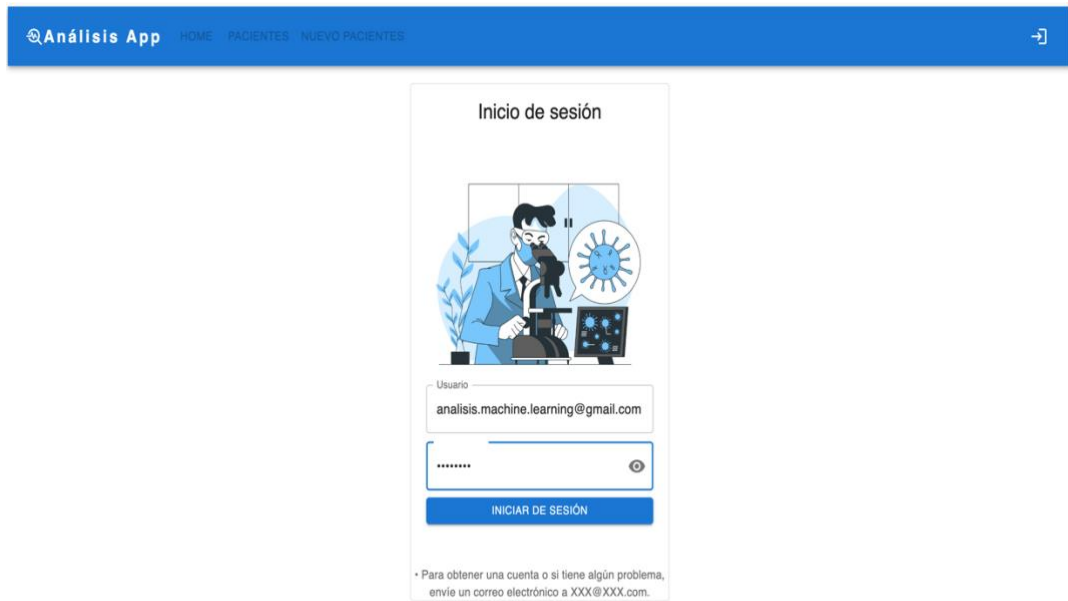
Ilustración 5: DER



Fuente: elaboración propia.

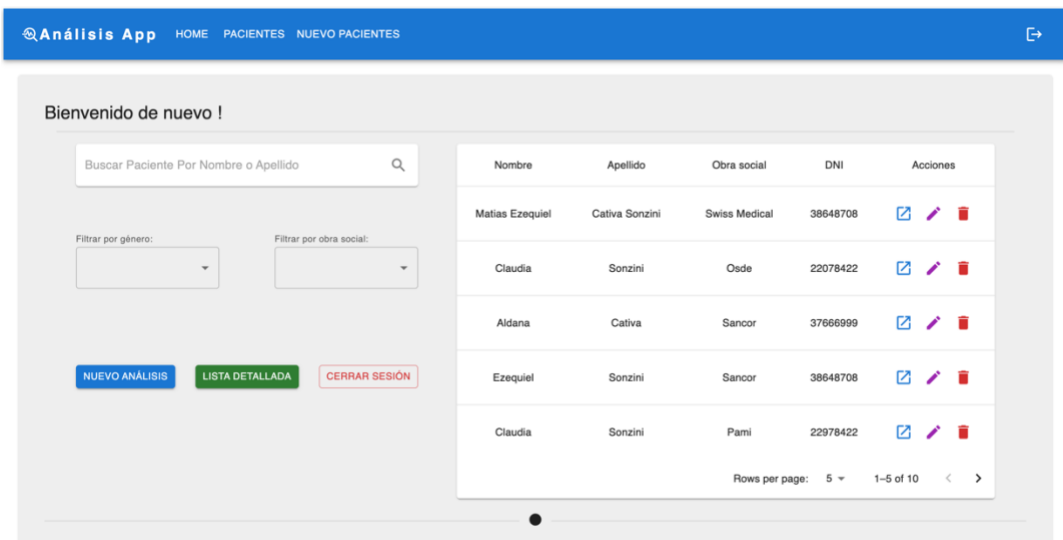
Prototipo De Interfaces De Pantallas

Ilustración 6: Prototipo 01 - Inicio de sesión.



Fuente: elaboración propia.

Ilustración 7: Prototipo 02 - Home



Fuente: elaboración propia.

Ilustración 8: Prototipo 03 – Cargar paciente.

The screenshot shows the 'Nuevo Paciente' form in the 'Análisis App'. The form is titled 'Nuevo Paciente' and contains several input fields for patient information. The fields are: Nombre (Ezequiel), Apellido (Sonzini), DNI (38648708), Obra Social (Swiss Medical), Género (Masculino), Fecha de nacimiento (13/02/1995), Edad (28), Email (matii.catiiva@gmail.com), Telefono (353 427-5471), Dirección (Bv illia 50), and Ciudad (Cordoba). There is a 'Descripción médica' section with a text area containing the text: '- Este es un caso fuera de lo común en el que se consultó con diferentes medicos que dieron un su visión sobre la enfermedad y se cree que es positivo.' Below this is a 'Cargar Imagen' section with a 'Seleccionar archivo' button and a 'Sin archivos seleccionados' message. At the bottom right, there are two buttons: 'NUEVO PACIENTE' (green) and 'CANCELAR' (red).

Fuente: elaboración propia.

Ilustración 9: Prototipo 04 – Paciente.

The screenshot shows the patient profile page in the 'Análisis App'. The page is titled 'Paciente Sonzini' and displays the patient's information: Género: masculino, DNI: 38648708, Edad: 28, Obra Social: Swiss Medical, Email: matii.catiiva@gmail.com, Teléfono: 353 427-5471, Dirección: Bv illia 50, Localidad: Cordoba. There is a section for 'Informe del 10 julio 2023' which contains a histological image of a tissue sample. Below the image is the text 'La imagen es cancerígena'. To the right of the image are four buttons: 'DESCARGAR ANÁLISIS' (blue), 'ENVIAR POR EMAIL' (blue), 'EDITAR' (green), and 'ELIMINAR' (red). At the bottom, there is a 'Descripción del caso' section with the text: '- Este es un caso fuera de lo común en el que se consultó con diferentes medicos que dieron un su visión sobre la enfermedad y se cree que es positivo.'

Fuente: elaboración propia.

Ilustración 10: Prototipo 05 – PDF generado de manera automática.

Informe creado por Inteligencia Artificial

Especialista en anatomía patológica martes, 11 de julio de 2023

Información del paciente:

- Nombre: Ezequiel
- Apellido: undefined
- Género: masculino
- Fecha de nacimiento: 12/02/1995
- Edad: 31/12/1969 años
- DNI: Fecha de nacimiento inválida
- Obra Social: Fecha de nacimiento inválida

Descripción médica del caso:

- Este es un caso fuera de lo común en el que se consultó con diferentes médicos que dieron su visión sobre la enfermedad y se cree que es positivo.

Reporte generado por Inteligencia Artificial

La imagen es cancerígena

Documento creado por un modelo de inteligencia artificial.
Respaldo por el doctor NOMBRE-APELLIDO.
Firma y sello del médico

Fuente: elaboración propia.

Ilustración 11: Prototipo 06 – Editar paciente.

Análisis App HOME PACIENTES NUEVO PACIENTES

Editar Paciente Ezequiel Sonzini

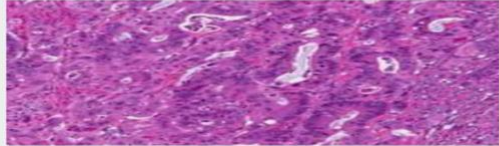
Nombre: Ezequiel | Apellido: Sonzini | DNI: 38648708 | Obra Social: Swiss Medical

Género: Masculino | Fecha de nacimiento: 13/02/1995 | Edad: 28

Email: mati.cativa@gmail.com | Teléfono: 353 427-5471 | Dirección: Bv illa 50 | Ciudad: Cordoba

Descripción médica: Este es un caso fuera de lo común en el que se consultó con diferentes médicos que dieron un su visión sobre la enfermedad y se cree que es positivo.

La imagen es cancerígena



ACTUALIZAR PACIENTE | **CANCELAR**

Fuente: elaboración propia.









Ilustración 12: Prototipo 04 – Lista de pacientes.

Análisis App HOME PACIENTES NUEVO PACIENTES

Lista de Pacientes

Nombre, Apellido o DNI

Filtrar por género: | Filtrar por obra social: | Filtrar por ciudad: | Filtrar por edad:

Nombre	Apellido	Fecha Nacimiento	DNI	Obra social	Dirección	Ciudad	Genero	Reporte	Fecha Reporte	Acciones
Matias Ezequiel	Cativa Sonzini	12/02/1995	38648708	Swiss Medical	Bv illa 50	Cordoba	masculino	✓	09/07/2023	 
Claudia	Sonzini	08/03/1971	22078422	Osde	Rivadavia 8	Villa Maria	femenino	✓	07/07/2023	 
Aldana	Cativa	13/04/1992	37666999	Sancor	Rivadavia 8	Villa Maria	femenino	✓	07/07/2023	 
Ezequiel	Sonzini	12/02/1995	38648708	Sancor	Bv illa 50	Cordoba	masculino	✓	07/07/2023	 

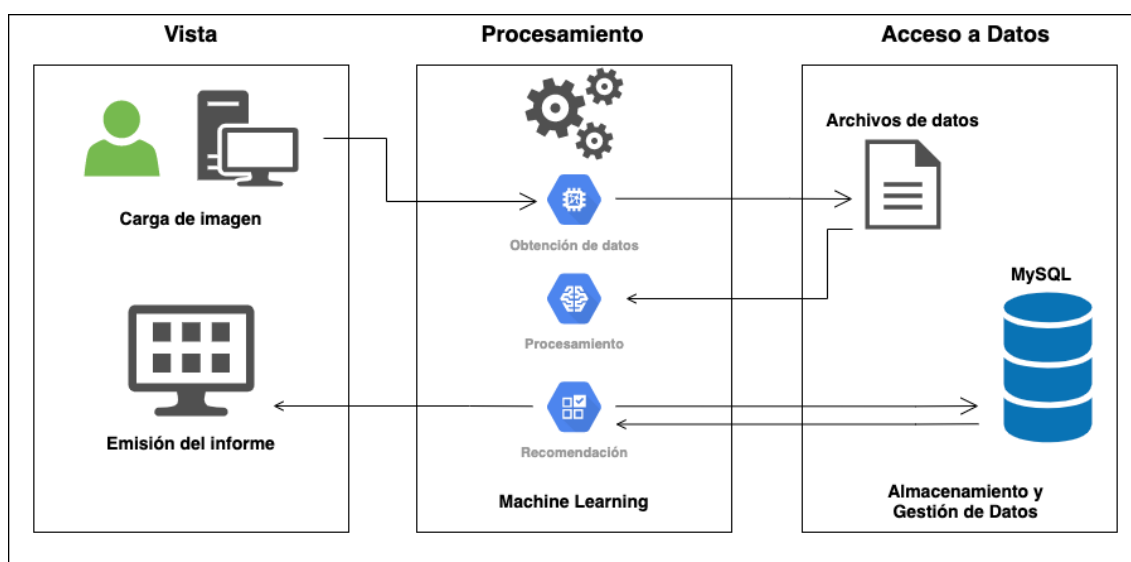
Rows per page: 4 | 1-4 of 11

Fuente: elaboración propia.

Diagrama De Arquitectura

En la siguiente ilustración se puede visualizar la arquitectura de la aplicación web. Partiendo desde que el usuario carga la imagen, luego esta es procesada y guarda sus resultados, para luego devolver el informe en la pantalla del usuario. Médicos especialistas previamente han examinado varias veces y catalogado manualmente las imágenes de biopsias para poder clasificar cada una de ellas según el resultado que creen que este arrojaría. Las imágenes clasificadas se introducen en un algoritmo de reconocimiento de imágenes, tras introducir en el algoritmo las imágenes que fueron clasificadas por los profesionales, este (el algoritmo), empezará a reconocer los patrones que son de células malignas.

Ilustración 13: Diagrama de arquitectura.



Fuente: elaboración propia.

Seguridad

1. Al usuario se le brindará un nombre de usuario y una contraseña para acceder a la aplicación.
2. La contraseña puede ser modificada bajo los siguientes requisitos:
 - a. Tener, como mínimo, 16 caracteres.
 - b. Contener, al menos, una letra mayúscula.

- c. Poseer, al menos, una letra minúscula.
 - d. Tener, al menos, un carácter especial.
3. Al quinto intento de inicio de sesión fallido, la contraseña caducará.
 4. La contraseña se encripta mediante la función bcrypt, que permite almacenarla de forma segura en la base de datos.
 5. Los datos sensibles de los pacientes van ser encriptados mediante la función cryptography, que permite encriptar y desencriptar datos por medio de una contraseña.

Para resguardar la información de la aplicación, ya sea la de los usuarios como la del código fuente, se contará con dos copias de cada una, con el objetivo de asegurar la redundancia de las mismas en caso de pérdidas, a saber:

1. Nube: la primera copia estará en la nube de AWS Cloud Hosting Service, donde el servidor y la base de datos están instanciadas. Aquí se tendrán los datos originales con los cuales trabaja la aplicación además de la copia de seguridad.
2. Almacenamiento externo: la segunda copia del código fuente y la base de datos estarán alojados en discos duros externos, los se encontrarán en las oficinas del equipo.

Para mantener los datos actualizados, en la nube se realizan actualizaciones diarias a la hora de terminar el día. En el almacenamiento externo las actualizaciones de base de datos de los usuarios son a diario, pero la del código fuente se realiza cada vez que se hace un cambio importante.

Análisis de Costos

Análisis de costo de desarrollo

A continuación, se va a exhibir la tabla donde se describe el análisis de costo de desarrollo para el proyecto.

Tabla 13: Análisis de costo de desarrollo.

Rol	Honorarios mensuales ar\$	Meses Totales	Subtotal ar\$
Programador IA	169.435	6	1016610
Analista funcional	96.615	1	96.615
Tester	120.167	6	721002
Administrador de base de datos	123.885	6	743310
Project Manager	187.775	6	1126650
Diseñador Jr	60.162	1	60.162
Total, Desarrollo ar\$			3.764349

Fuente: elaboración propia.

Los datos salariales fueron obtenidos del consejo profesional de ciencias informáticas de la Provincia de Córdoba. (CPCIPC, 2021)

Análisis de costos operativos

La siguiente tabla muestra los análisis de costos operativos del proyecto.

Tabla 14: Análisis de costos operativos.

Recurso	Cantidad	Fuente	Subtotal inicial ar\$	Subtotal mensual ar\$
Lenovo ThinkPad E15 Core I7	1	https://www.lenovo.com/ar/es/laptops/thinkpad/serie/E15-	173.999	n/a

		G2/p/20TESO X200		
Disco externo 5tb seagate	1	https://www.mexx.com.ar/productos-rubro/almacenamiento/37177-disco-externo-5tb-seagate-expansion.html	15.999	n/a
AWS Aurora Mysql*	1	https://calculator.aws/#/estimate?id=675e5c6de5c24cf685a85f1f41524ef5b98671f1	n/a	11.672,7
AWS EC2*	1	https://calculator.aws/#/estimate?id=675e5c6de5c24cf685a85f1f41524ef5b98671f1	n/a	12.585,3
Total, operativo \$			189.998	24.258

Fuente: elaboración propia.

Los precios están a los pesos argentinos al tipo de cambio de dólar oficial extraídos de US\$ 1 = AR\$ 105,23 (Banco Nación, 2021).

Resumen de costos

A continuación, se describe un resumen de los costos en una tabla.

Tabla 15: Resumen de costos.

Descripción	Inicial \$	Recurrencia mensual \$
Total, de costos de hardware y servicios.	189.998	24.258
Total, de costos de costos de desarrollo	n/a	3.764349
Totales	189.998	3.788607

Fuente: elaboración propia.

Análisis de Riesgos

Estimación De Probabilidad:

En el siguiente cuadro se presentan los riesgos identificados en el proyecto y su estimación de probabilidad:

Tabla 16: Estimación de Probabilidad.

ID	Riesgo	Expresión en lenguaje natural	Probabilidad
001	Subestimar tiempo de entrega	Bajo	15%
002	Crecimiento en el alcance del proyecto	Bajo	15%
003	Personal incapacitado con la tecnología	Moderado	45%
004	Datos incorrectos para el machine learning	Moderado	45%
005	Tecnología obsoleta	Bajo	15%

Fuente: elaboración propia.

Estimación De Impacto:

En el siguiente cuadro se muestra el impacto que producirían las ocurrencias de los riesgos del cuadro anterior:

Tabla 17: Estimación de impacto.

ID	Riesgo	Expresión en lenguaje natural	Impacto
001	Subestimar tiempo de entrega	Medio	3

002	Crecimiento en el alcance del proyecto	Medio	3
003	Personal incapacitado con la tecnología	Alto	4
004	Datos incorrectos para el machine learning	Muy alto	5
005	Tecnología obsoleta	Medio	3

Fuente: elaboración propia.

Exposición al riesgo:

Es el resultado de la multiplicación de la probabilidad de ocurrencia por el impacto que tendría el mismo en el proyecto.

ER (exposición al riesgo):

- Riesgo bajo: $ER \geq 0,50$.
- Riesgo medio: $0,51 \leq ER \leq 0,99$.
- Riesgo alto: $ER \geq 1$.

Tabla 18: Exposición al riesgo.

ID	Riesgo	Probabilidad	Impacto	Exposición
001	Subestimar tiempo de entrega	15%	3	0,45
002	Crecimiento en el alcance del proyecto	15%	3	0,45
003	Personal incapacitado con la tecnología	45%	4	1,8
004	Datos incorrectos para el machine learning	45%	5	2,25
005	Tecnología obsoleta	15%	3	0,45

Fuente: elaboración propia.

Plan De Contingencia:

Tabla 19: Plan de contingencia.

Plan de contingencia para los ER ≥ 1	
Personal incapacitado con la tecnología	Se asignará la capacitación necesaria al personal que lo necesite y se le asignará la tarea a otro personal que pueda realizarla.
Datos incorrectos para el machine learning	Verificar que los datos para el aprendizaje sean de calidad y útiles.

Fuente: elaboración propia.

Conclusiones

El proyecto sistema de diagnóstico predictivo de células anómalas por imágenes se llevó a cabo por una motivación personal de querer entrelazar dos mundos muy interesantes para mí, el software y la medicina. Como hijo de médico siempre tuve un interés particular por el área. Los objetivos planteados fueron alcanzados con éxito, ya que el software permite determinar que celular son malignas y cuáles no por medio de inteligencia artificial y creando una base de datos que permite la creación de informes de manera autónoma. El sistema puede mostrar el nivel de error y la calidad de los datos, para que el profesional de la salud sea el último eslabón a determinar el resultado del informe. De esta manera puede haber un impacto tecnológico en el área, ya que lo digitaliza en su totalidad, uno de los problemas planteados en el análisis.

El desarrollo del proyecto en todo su conjunto, significó un desafío muy grande, que me permitió utilizar los recursos adquiridos en la carrera de ingeniería en software a lo largo de todas las etapas del mismo. En cuanto al desarrollo del sistema como tal, a pesar de dedicarme al desarrollo me llevó a tener que investigar y aprender mucho, sobre la inteligencia artificial, el análisis de datos e interiorizarse en el mundo hermoso de la ciencia de datos. Es por ello que me resulta grato y satisfactorio el haber cumplido los objetivos, los cuales me

permitieron aprender sobre herramientas y sobre todo de cómo trabajarlas a nivel personal y profesional.

Demo

En el siguiente enlace se podrá descargar el prototipo desarrollado para su posterior ejecución junto con su código fuente y su documento “Readme” donde explica su ejecución.

Link: <https://github.com/MatiCativa/TFG-Analysis>

En el siguiente enlace se podrán encontrar los diferentes modelos de machine learning con las bases de datos de las imágenes para entrenar los modelos.

Link:

<https://drive.google.com/drive/folders/1WpLQb7CLdRz4WnC3w7PfaE0CmXp7xWX-?usp=sharing>

Referencias

Xataka. (2 de enero 2020) Obtenido de <https://www.xataka.com/inteligencia-artificial/google-desarrolla-inteligencia-artificial-para-detectar-cancer-mama-capaz-superar-a-expertos-humanos>

Google Health (s.f) Obtenido de <https://about.google/stories/seeingpotential/>

Oracle. (s.f) Obtenido de <https://www.oracle.com/mx/artificial-intelligence/what-is-ai/>

Python (s.f) Obtenido de <https://docs.python.org/es/3/tutorial/index.html>

Developer Mozilla (s.f) Obtenido de <https://developer.mozilla.org/es/docs/Glossary/SQL>

Iberdrola (s.f) Obtenido de <https://www.iberdrola.com/innovacion/que-es-inteligencia-artificial>

Pybonacci (s.f) Obtenido de <https://pybonacci.org/2020/02/20/haciendo-un-dashboard-con-streamlit-en-python/>

Ccipc, (31 de agosto 2021) Obtenido de https://www.ccpic.org.ar/sites/default/files/honorarios_recomendados_agosto_2021.pdf

Banco Nación (31 de agosto 2021) Obtenido de <https://www.bna.com.ar/Personas>

Anexos

Cuestionario del día de observación:

¿Cómo toma los datos de los pacientes?

¿Cómo almacena la información de los pacientes?

¿Dónde almacena la información de los pacientes?

¿Anota las características de lo que ve en el microscopio en la computadora o lo pasa luego?

¿Se le asigna a cada paciente un ID?

¿Primero realiza el informe a mano y luego lo pasa a la computadora?

¿Qué pasa si algún dato escrito a mano esta mal escrito o no se entiende la letra?

¿Cuánto tiempo le lleva aproximadamente todo el proceso?

¿También lo realiza los fines de semana?

¿Les manda el resultado a los pacientes por email?

¿Qué pasa si un paciente se olvida de retirar un estudio o lo necesita con urgencia?