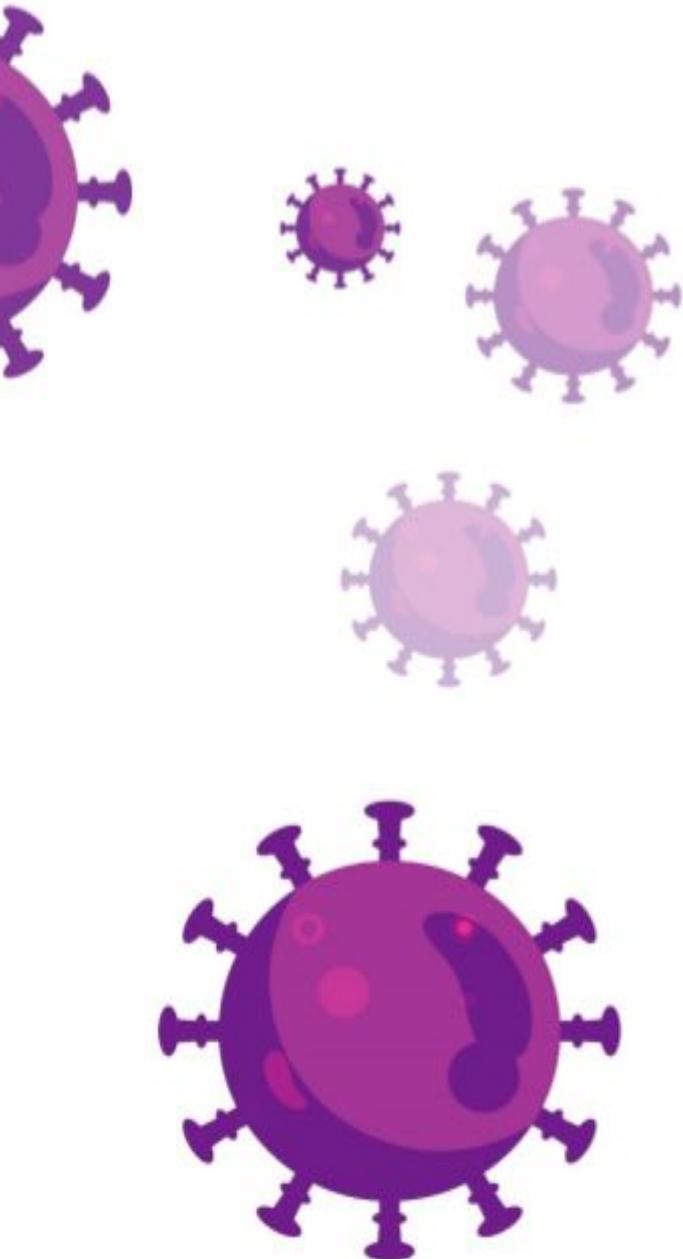


PROYECTO PRIORIZACIÓN DE CASOS

COVID -19

Máster en Business Intelligence y Big Data (2019 - 2020)

Escuela De Organización Industrial (EOI)



Santo Domingo, República Dominicana.

1. Definición del problema

A finales del 2019 surge en Wuhan, ciudad de China, un brote de una enfermedad altamente contagiosa. Esta enfermedad fue identificada como un nuevo tipo de coronavirus y fue nombrada SARS-CoV-2 (eventualmente la WHO la denominó Covi-19). Rápidamente esta enfermedad se expande por casi todos los países del mundo y es declarada una pandemia global por la WHO. Aunque la mayoría de los contagiados presentan síntomas leves, debido a la alta tasa de contagio que tiene esta enfermedad existe una gran cantidad de pacientes que presentan un síndrome respiratorio agudo severo y se está sufriendo un enorme déficit de camas y respiradores en los Hospitales.

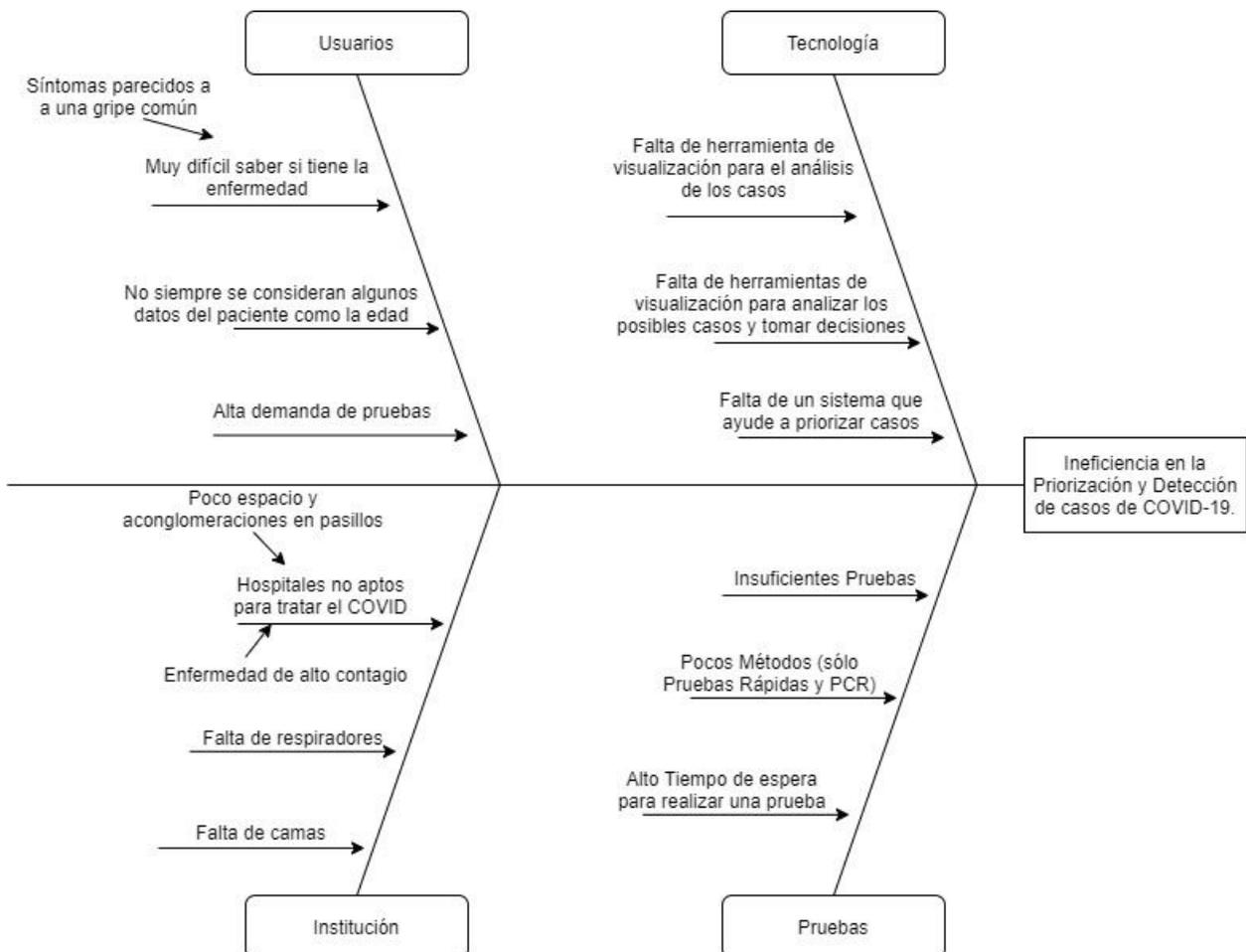
Esto supone un enorme reto sanitario en todos los países y aunque algunos han logrado controlar el avance de la enfermedad, la gran mayoría aún está en medio de esta batalla. En República Dominicana a la fecha se tiene cerca de 65,000 casos y ya los hospitales y clínicas se encuentran ya colapsados debido a la gran cantidad de pacientes. Al mismo tiempo el país no puede cubrir la gigantesca demanda de pruebas para detectar la enfermedad, siendo este uno de los factores más importantes para el control de la misma, ya que mientras más casos se identifiquen es posible tomar medidas de aislamiento oportunas y así prevenir futuros contagiados a través de estas personas.

Por esto, resulta imperativo identificar un método inteligente de realizar estas pruebas que apunte con un poco más de precisión a los casos que tengan mayor probabilidad de ser positivos y así sacar más provecho de las pruebas que se poseen, eficientizando el proceso y disminuyendo efectivamente la propagación de la pandemia.

1.1 Causas del problema

En República Dominicana, la detección del COVID-19 en muchos casos no es eficiente, existen casos en que pacientes han tardado más de una semana en recibir una prueba. Una de las causas principales de este problema es la poca cantidad de pruebas diarias que se realizan, ya que no son las suficientes para satisfacer su demanda. Lo mismo ocurre al tratar y priorizar a los pacientes con la enfermedad ya que, en la mayoría de casos, casi todas las camas están ocupadas.

En este diagrama, recolectamos las principales causas de ineficiencia en la priorización y detección de casos de COVID-19 en el país.



1.2 Consecuencias

Preselección de problemas

- ❖ **Colapso de hospitales:** debido a la gran demanda de pacientes con probable caso de covid-19, los hospitales han estado saturados por la cantidad de personas que presentan síntomas y desean ser atendidos.
- ❖ **Larga lista de espera para poder realizarse la prueba de covid-19:** Cuando una persona presenta ciertos síntomas desea hacerse una prueba para validar si es la enfermedad, pero al haber tantos posibles casos, se debe de realizar por turnos por lo que para obtener la prueba se puede llevar incluso semanas.
- ❖ **Escasos medicamentos y recursos para dar atención a todos los pacientes**
- ❖ **Gran necesidad de gestionar mejor a las pacientes que llegan con síntomas,** ya que al llegar muchos a la vez, no hay espacio suficiente para atenderlos.
- ❖ **Fallecimiento de personas que no recibieron el tratamiento adecuado a tiempo debido a la tardanza para realizarse la prueba.**

Priorización de problemas

Este es el orden de importancia que tienen los problemas que se estarían solucionando con este proyecto, en base al impacto que está causando actualmente en el país.

- 1) Colapso de hospitales
- 2) Escasos medicamentos y recursos para dar atención a todos los pacientes
- 3) Fallecimiento de personas que no recibieron tratamiento a tiempo, debido a la tardanza para realizarse la prueba.
- 4) Larga lista de espera
- 5) Gran necesidad de gestionar mejor a los pacientes que llegan con síntomas

1.3 ¿Por qué se está produciendo el problema?

Identificación de las causas y sus relaciones

El coronavirus proviene de una familia de virus que pueden causar enfermedades como el resfriado común, el síndrome respiratorio agudo grave y el síndrome respiratorio de Oriente Medio. El virus que se enfrenta actualmente a nivel mundial, con especial mención en este estudio en la República Dominicana, corresponde al síndrome respiratorio agudo grave antes mencionado y se identificó en el 2019 por causa de un brote de enfermedades de fuente desconocida que se originó en China según especuladores por murciélagos o pangolines, declarado en marzo de 2020 por la Organización Mundial de la Salud (OMS) como la pandemia del COVID-19.

Desde entonces, no se podido determinar la fuente oficial causante del problema ni una vacuna para contrarrestar la misma y la propagación continua aumentando, siendo el

riesgo mayor para cualquiera que tenga contacto cercano con personas que tienen el virus, como es el caso de los trabajadores de la salud o incluso ciudades altamente pobladas, provocando estos motivos el colapso de hospitales debido a la gran demanda de pacientes con probables casos de covid-19 que presentan síntomas y desean ser atendidos.

Los primeros indicios de síntomas del COVID-19 pueden incluir pérdida del gusto o del olfato, pero los signos y síntomas que frecuenta la enfermedad pueden aparecer entre dos y 14 días después de la exposición al virus, destacándose los siguientes más comunes que serán los más importantes para el estudio de la solución que se plantea:

- Fiebre
- Tos
- Cansancio

Otros síntomas a incluir para ser considerados en los pacientes debido a que pueden estar relacionados serían:

- Falta de aire o dificultad para respirar
- Dolores en los músculos
- Escalofríos
- Dolor de garganta
- Goteo de la nariz
- Dolor de cabeza
- Dolor en el pecho

Bajo este contexto, los síntomas en las diferentes personas pueden variar, no ocurrir o ser muy leves o extremos, no obstante, tomando en cuenta los síntomas anteriores así como la edad y condición de las personas, se pueden llegar a priorizar casos donde el riesgo es mayor y por ende requieren más atención. Algunas de estas condiciones que pueden tener un riesgo más alto de enfermarse gravemente incluyen:

- Enfermedades cardíacas graves, como insuficiencia cardíaca.
- Enfermedades de las arterias coronarias, o miocardiopatía.
- Cáncer.
- Enfermedad pulmonar obstructiva crónica (EPOC).
- Diabetes tipo 2.
- Obesidad grave.
- Enfermedad renal crónica.
- Enfermedad de células falciformes.
- Sistema inmunitario debilitado por trasplante de órganos sólidos.
- Asma
- Enfermedad hepática

- Enfermedades pulmonares crónicas, como fibrosis quística
- Afecciones del cerebro y del sistema nervioso
- Sistema inmunitario debilitado por trasplante de médula ósea, VIH, o algunos medicamentos
- Diabetes tipo 1
- Presión arterial alta.

Priorización de las causas a intervenir

- Pacientes bajo los signos y síntomas descritos.
- Pacientes con algunas de las condiciones indicadas.
- Pacientes adultos mayores o envejecientes.

1.4 Importancia del problema

Desde que el COVID-19 empezó a provocar miles de muertes y el colapso de diferentes áreas como el comercio, esta enfermedad se ha alzado como uno de los mayores problemas en los tiempos actuales. Para evitar que este virus contagie a otras personas es necesario realizar pruebas a aquellos que cumplan con los síntomas, de esta forma poder aislar y tratar de manera adecuada a los que presenten positivo en los resultados.

Si los hospitales están saturados de personas que intentan realizarse la prueba, se dificulta el poder evitar que la enfermedad se siga expandiendo, lo que a su vez desencadenaría en más muertes y el colapso de diferentes sistemas de nuestra sociedad.

1.5 Para quiénes es importante resolver el problema

En términos generales, este es un problema que involucra a todas las personas que vivan en una área en donde la enfermedad esté en auge. Aunque los que más beneficios les otorga el resolver este problema son:

- Los hospitales y centros clínicos.
- Potenciales portadores del virus COVID-19.
- Para aquellos cuyos comercios o empresas se vean afectados por las contra medidas de la enfermedad.
- Personas que se encuentren en los hospitales y centros clínicos por razones ajenas al COVID-19.

Supuestos

- Los hospitales no tienen cantidad suficiente de pruebas para COVID-19.
- La priorización de los casos con mayor probabilidad ayudará a detectar los casos positivos de forma más rápida.
- Las autoridades necesitan un sistema más efectivo de realización de pruebas.

- No se cuenta con sistema entrenados para brindar la atención eficaz a los pacientes.
- Los hospitales están saturados por la cantidad de casos de COVID-19.
- Los pacientes que desean realizarse la prueba de COVID-19 presentan muchos problemas.
- Los más interesados en este sistema serían los hospitales.
- La mayoría de la población no ha podido hacerse la prueba del COVID-19.
- Aplicando un método predictivo se contribuirá al control de la pandemia.
- Los pacientes que desarrollan síntomas graves tienen mayor probabilidad de fallecer.
- Aumentando la velocidad de detectar el covid se reducirían las personas esperando por hacerse pruebas.
- Las 5 funcionalidades principales de mi producto deben ser:
 - Agilización en el proceso de realizar las pruebas.
 - Disminución de la propagación de los casos, ya que las personas tendrán acceso a los resultados de una forma más rápida, según sea su situación.
 - Reducir el congestionamiento de hospitales.
 - Atención priorizada de los casos de la enfermedad.
 - Detección de la posible enfermedad de forma automatizada.

Hipótesis

❖ **Hipótesis Descriptiva**

1. La mayoría de los tratamientos utilizados para tratar el COVID-19 son efectivos.

❖ **Hipótesis Exploratoria**

2. Una gran parte de los hospitales tienen una demanda de pruebas mayor de la cantidad de pruebas que poseen.
3. El 50% de las autoridades estarían dispuestas a usar un 20% del presupuesto en pruebas COVID-19 para conseguir un método más efectivo de realizar las mismas.
4. Hay más pacientes que son atendidos en su casa que en el hospital.
5. Más del 50% del total de pruebas realizadas han dado negativo.
6. La mayoría de los pacientes que solicitan hacerse una prueba de COVID-19 tardan más de 1 semana en obtener respuesta de la misma.
7. La cantidad de casos activos es mayor a la de recuperados.
8. El 90% del total de la población no han tenido acceso a las pruebas de COVID-19.
9. La mayoría de las personas que mueren por la enfermedad son mayores de 60 años.

❖ **Hipótesis Correlacional**

10. Si aumenta la velocidad en la que se detecta el virus, se podrá controlar más su propagación.

2. Investigación y toma de datos

Para poder realizar las validaciones de las hipótesis, se obtuvieron informaciones a través de encuestas realizadas a personas que habían pasado por el proceso de solicitar realizarse pruebas de Covid-19, así como personal médico involucrado en el día a día para dar apoyo a este proceso. De las encuestas mencionadas, 22 personas que realizaron el proceso para poder hacerse la prueba, respondieron a nuestras preguntas y de parte del personal médico, 12 participaron en las encuestas, por lo que obtuvimos una muestra considerable para tener una noción de cómo aproximadamente se encuentra la situación actualmente.

Las hipótesis validadas con hechos públicos, fueron soportadas con investigaciones de noticias en revistas, periódicos, datos e instituciones confiables que proveían información que ayuda a corroborar cada una de las hipótesis, donde cada una posee su referencia a la fuente que la soporta.

2.1 Validación de hipótesis

2.1.1 Identificación de hipótesis a validar

❖ Validadas con entrevistas y encuestas

- ***Una gran parte de los hospitales tienen una demanda de pruebas mayor de la cantidad de pruebas que poseen.***

Según los resultados de las encuestas, el 75% de los médicos encuestados han indicado que la cantidad de personas que solicitan diariamente las pruebas de covid-19 es de 11 a 50 personas, sin embargo, el 50% ha indicado que tiene una capacidad de realizar de 1 a 20 pruebas diarias, por lo que en una gran cantidad de hospitales se supera la cantidad de las pruebas que tienen disponibles por día.

- ***La mayoría de los pacientes creen que es necesario un nuevo método y recursos para poder apoyar la detección más temprana del virus.***

La mayoría de las personas que participaron en la encuesta indicaron que el proceso actual para detectar el virus no es suficiente para cubrir la demanda y se necesita más recursos para poder agilizar la detección del virus.

- ***La mayoría de los pacientes que solicitan hacerse una prueba de COVID-19 tardan más de 1 semana en obtener respuesta de la misma.***

El 63.6% indicó que tardan de 1 día a una semana para realizarse la prueba, aunque el 18.2% indicó que tarda más de un mes en realizarse esa prueba, ya

que debían de obtener la confirmación de la cita y esperar a que les tocara su turno, y esto podría tardar hasta más de un mes en algunos casos, por lo que depende mucho del centro médico y los recursos con los que se encuentre.

❖ **Validadas con investigación de hechos públicos**

- ***La mayoría de los tratamientos utilizados para tratar el COVID-19 son efectivos.***

Actualmente se han ido desarrollando una gran cantidad de medicamentos para poder combatir este virus, pero lamentablemente en muchos de ellos no está comprobada su efectividad ya que funcionan para ciertos tipos de pacientes y necesitan más investigación, aunque hay algunos que han sido autorizados para su uso en caso de emergencias por la FDA¹.

- ***Hay más pacientes que son atendidos en su casa que en el hospital.***

La autoridad sanitaria de República Dominicana indicó: “Hay 1,173 personas se encuentran en camas de hospitalizaciones en todo el país y 322 están en Unidades de Cuidados Intensivos, para un total de 1,484 pacientes. Mientras que 166 personas se encuentran bajo ventilación.”². También indican que hubo una reducción al 70% de ocupación hospitalaria, lo cual sigue representando un gran porcentaje para los recursos que se poseen, ya que no dan abasto.

Se debe tener en cuenta que de la cantidad de infectados que serían unos 33,110 los que están en hospitales, lo cual representa alrededor de un 18% y los hospitales se encuentran saturados con esta cantidad de personas.

- ***Más del 50% del total de pruebas realizadas han dado negativo.***

Según lo publicado en el boletín #145 del Ministerio de Salud de la República Dominicana, correspondiente al 11 de agosto de 2020, se han realizado 300,134 pruebas en total, de las cuales 81,094 han dado positivo, lo que representa un 27.01% y el restante ha dado negativo, lo que es un 72.99%, lo que hace notar

¹ "Tratamientos y medicamentos para el coronavirus: monitoreo" 24 ago.. 2020, <https://www.nytimes.com/es/interactive/2020/science/coronavirus-tratamientos-curas.html>. Se consultó el 26 ago.. 2020.

² "Se reduce a 70 % la ocupación hospitalaria de ... - Listín Diario." 3 ago.. 2020, <https://listindiario.com/la-republica/2020/08/03/629113/se-reduce-a-70--la-ocupacion-hospitalaria-de-pacientes-covid-en-el-gran-santo-domingo>. Se consultó el 26 ago.. 2020.

que la mayoría han dado negativo y esas pruebas pudieron haberse realizado a personas que tenían mayor posibilidad de dar positivo.³

- ***La cantidad de casos activos es mayor a la de recuperados.***

Según el boletín #145, hay 81,094 casos confirmados, 45,666 casos recuperados y 34,082 activos, por lo que el 56.3% de los casos confirmados se han recuperado el 11 de agosto de 2020.⁴

- ***El 90% del total de la población no han tenido acceso a las pruebas de COVID-19.***

Según el reporte de Worldometer, aproximadamente el menos del 2.77% de la población ha tenido acceso a las pruebas de COVID-19 en República Dominicana, dividiendo la cantidad de personas en el país y las que se han hecho pruebas.⁵

- ***La mayoría de las personas que mueren por la enfermedad son mayores de 60 años.***

Según lo expresado por el médico e inmunobiólogo Janko Nikolich-Zugich de la Facultad de Medicina de la Universidad de Arizona afirma que «Las personas mayores no son tan buenas para reaccionar a los microorganismos que no han encontrado antes». Llama a la vejez «el crepúsculo de la inmunidad».⁶

También se indica que el 78 % de los pacientes que salieron positivos en EE. UU. tienen menos de 65 años pero, 78 % de los fallecidos tienen 65 años o más.

Se debe tomar en cuenta que la gran mayoría de los ancianos fallecidos también ha sido porque presentaban algún tipo de enfermedad previa.

³ "República Dominicana reporta 293 personas en UCI, 595"
<https://coronavirusrd.gob.do/2020/08/11/republica-dominicana-reporta-293-personas-en-uci-595-nuevos-casos-de-covid-19-y-18-fallecimientos/>. Se consultó el 26 ago.. 2020.

⁴ "República Dominicana reporta 293 personas en UCI, 595"
<https://coronavirusrd.gob.do/2020/08/11/republica-dominicana-reporta-293-personas-en-uci-595-nuevos-casos-de-covid-19-y-18-fallecimientos/>. Se consultó el 26 ago.. 2020.

⁵ "Coronavirus Update (Live): 23,472,214 Cases and 810,262"
<https://www.worldometers.info/coronavirus/>. Se consultó el 26 ago.. 2020.

⁶ "Más del 95 % de los muertos por COVID son ancianos." 20 abr.. 2020,
<https://es.panampost.com/mamela-fiallo/2020/04/20/95-muertos-covid-ancianos/>. Se consultó el 26 ago.. 2020.

- ***Si aumenta la velocidad en la que se detecta el virus, se podrá controlar más su propagación.***

De acuerdo a lo publicado en la revista RSNA, sobre el sistema predictivo utilizado en China, ellos afirman que “El diagnóstico temprano de la enfermedad por coronavirus 2019 (COVID-19) es crucial para el tratamiento y control de la enfermedad”.

También indican que la “Tomografía Computarizada de tórax tiene una alta sensibilidad para el diagnóstico de COVID-19. Nuestros datos y análisis sugieren que la TC de tórax debe considerarse para la detección de COVID-19, la evaluación integral y el seguimiento, especialmente en áreas epidémicas con alta probabilidad de enfermedad antes de la prueba. La detección rápida con alta sensibilidad de la infección viral puede permitir un mejor control de la propagación viral.”⁷

Como hemos visto, China, es uno de los países que tiene un mayor control de la propagación del virus hasta el momento y estos han indicado que utilizar métodos de detección y predicción con las tomografías de tórax, es una forma confiable, práctica y rápida de diagnosticar y evaluar el virus, lo que ayudaría a controlar la propagación, ya que las personas serían ingresadas a los hospitales dependiendo de la gravedad de su caso o mantendrían un aislamiento domiciliario para evitar contagiar a otros.

⁷ "Correlation of Chest CT and RT-PCR" <https://pubs.rsna.org/doi/10.1148/radiol.2020200642>. Se consultó el 27 ago.. 2020.

Resumen de las validaciones de las hipótesis

Hipótesis	Resultado
1. Una gran parte de hospitales tienen una demanda de pruebas mayor de la cantidad de pruebas que poseen.	+
2. La mayoría de los pacientes creen que es necesario un nuevo método y recursos para poder apoyar la detección más temprana del virus.	+
3. La mayoría de los pacientes que solicitan hacerse una prueba de COVID-19 tardan más de 1 semana en obtener respuesta de la misma.	-
4. La mayoría de las personas que mueren por la enfermedad son mayores de 60 años.	+
5. La mayoría de los tratamientos utilizados para tratar el COVID-19 son efectivos.	-
6. Hay más pacientes que son atendidos en su casa que en el hospital.	+
7. Más del 50% del total de pruebas realizadas han dado negativo.	+
8. La cantidad de casos activos es mayor a la de recuperados.	-
9. El 90% del total de la población no ha tenido acceso a las pruebas de COVID-19.	+
10. Si aumenta la velocidad en la que se detecta el virus, se podrá controlar más su propagación.	+

2.2 Análisis y conclusiones derivadas de las entrevistas y encuestas

Encuesta a profesionales de los centros de salud.

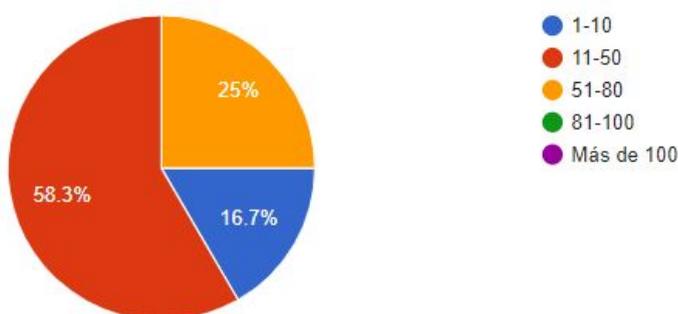
Conforme a los resultados de las encuestas y entrevistas realizadas a los profesionales de salud, se ha podido identificar que la mayoría de las personas que completaron la misma corresponden a doctores del área de medicina, perfil que buscábamos para poder focalizar los datos a obtener.

Dichos encuestados se encuentran ejerciendo su profesión en instituciones tales como el Hospital Docente Universitario Dr. Darío Contreras, la Clínica Gómez Patiño, Marcelino Vélez y la Clínica C J, lo cual los coloca en buena posición de aportar con relación a la situación que se está viviendo en cada uno de los centros citados.

Determinar qué porcentaje de la población realiza diagnósticos diariamente es uno de los factores más importantes a la hora de establecer medidas y acciones en beneficio de la reducción de la propagación de la pandemia en todo el mundo y es debido a estos datos que pudimos observar que en estos centros de República Dominicana se realizan en su mayoría entre 11 a 80 solicitudes diarias, incluyendo el porcentaje representativo que posee el rango de 51 a 80.

¿Alrededor de cuántas personas solicitan la pruebas diariamente de covid-19 en esa institución?

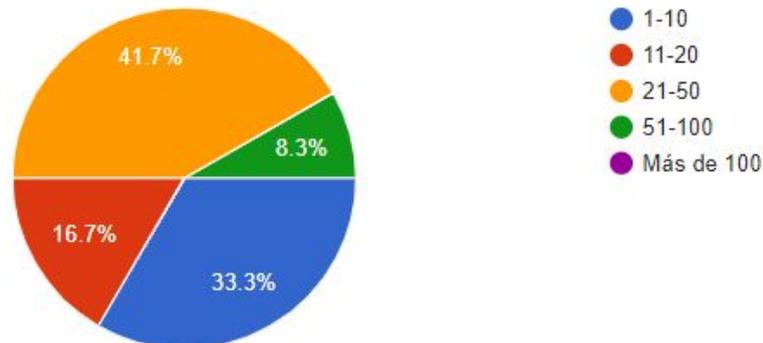
12 respuestas



Así mismo, pudimos comprobar que la capacidad de pruebas que se pueden realizar en los diferentes centros no cubre en su totalidad la demanda de pacientes que acuden a los hospitales y/o centros en búsqueda de diagnósticos ya sea porque cuentan con síntomas o por prevención ante condiciones, ya que el 50% indica que es mayor la cantidad de pruebas solicitadas que las realizadas y en el porcentaje de 41.7% que indica una capacidad de realizar de 21 a 50 pruebas, muchas de las respuesta fueron del mismo hospital, por lo que este porcentaje sería menor.

¿Cuál es la cantidad aproximada de pruebas diarias que pueden realizar?

12 respuestas



Con respecto a los métodos actuales para la realización de pruebas, los datos muestran que la satisfacción de los pacientes con relación a la efectividad es pasivamente complaciente ya que sólo un 75% de estos indicaron que no están de acuerdo en su totalidad con el método actual, otro 8.3% de los encuestados reflejaron que si es efectivo y el 16.7% restante se siente seguro con la PCR, pero no con las pruebas rápidas ya que sus resultados no son muy precisos y en ocasiones hay que realizar hasta una segunda prueba para confirmar su efectividad.

¿Crees que el método actual para realizar las pruebas del covid-19 es efectivo?

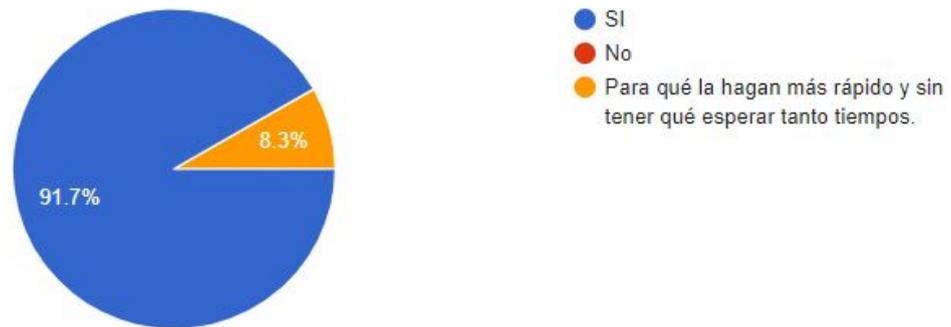
12 respuestas



Debido a lo antes visto con respecto a la cantidad de diagnósticos por día, la capacidad de los centros hospitalarios para atender la cantidad de pacientes y adicionando que el 100% de los encuestados están a favor, entendemos propicio enfatizar en mejorar la eficiencia de las pruebas, siendo necesario para ello contar con mayores recursos o financiamiento para responder a ante esta pandemia, además de que si se adicionan mejores prácticas como la automatización de los metodos de diagnosticos, se pudieran reducir los tiempos de espera de los pacientes que dado los resultados en su mayoría indicaron que rondan entre de 1 a 20 minutos y el restante dependiendo la prueba entre 21 a 40 minutos de ser la rápida o en el caso de la PCR de 12 a 48 horas.

¿Cree usted necesario conseguir más recursos para mejorar la eficiencia de las pruebas?

12 respuestas



¿Cuánto tiempo suele tardar el proceso para realizar una prueba de covid-19?

12 respuestas



Entre las medidas que actualmente están realizando en sus procedimientos estos centros para priorizar sus pruebas se encuentran:

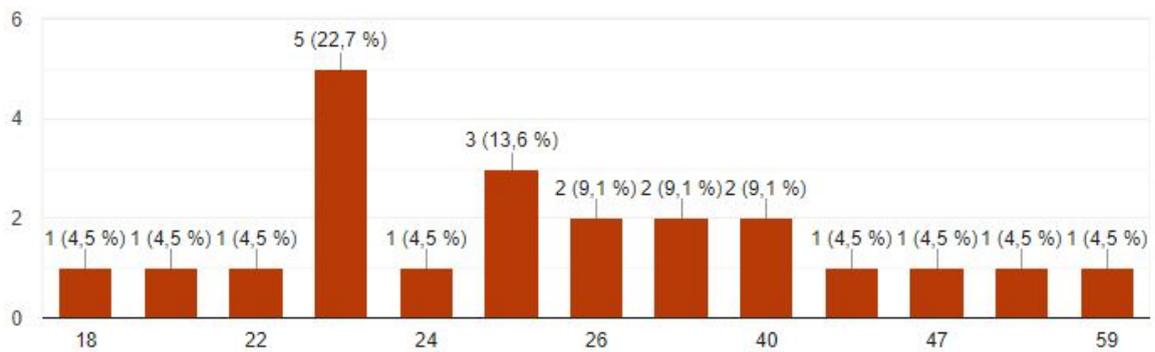
- Colocar etiquetas a los tubos de ensayo con la indicación de urgencia.
- Prueba rápida a los que presentan síntomas de tos, fiebre, dificultad respiratoria y radiografía de tórax positiva.
- Realizar la prueba a los pacientes que serán ingresados, ya sea en sala, UCI o cirugía.
- Se les realiza la prueba rápida de COVID 19 a todos los pacientes que se van a ingresar, ya sea en UCI, planta o cirugía.
- A los que necesitan prueba con urgencia, se les realizan pruebas rápidas y se entiende que no es el mejor método para detectar el virus, puesto que esta prueba arroja muchos falsos positivos.
- Dependiendo de la gravedad del paciente se priorizan los casos.

Encuesta a pacientes del proceso de pruebas para detectar el virus en los centros de salud.

Conforme a los resultados de las encuestas y entrevistas realizadas a los pacientes, se ha podido identificar que la mayoría de estos cuentan con seguro médico y que las edades de los encuestados rondan entre los 18 a 59 años de edad, factor que nos permite contar con datos muy variantes en relación a las posibles condiciones de estos y su satisfacción con los métodos de diagnósticos actuales.

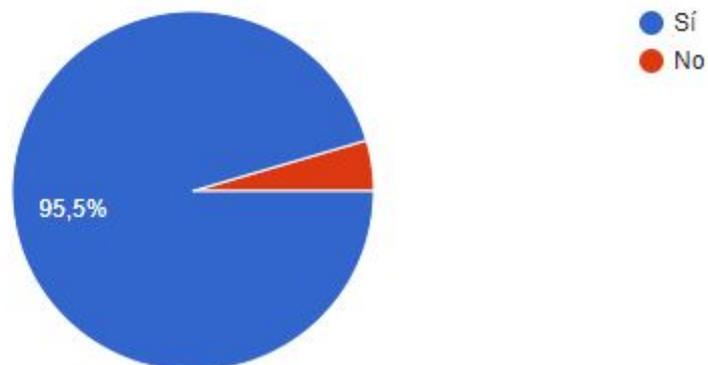
¿Cuál es su edad?

22 respuestas



¿Posee seguro médico?

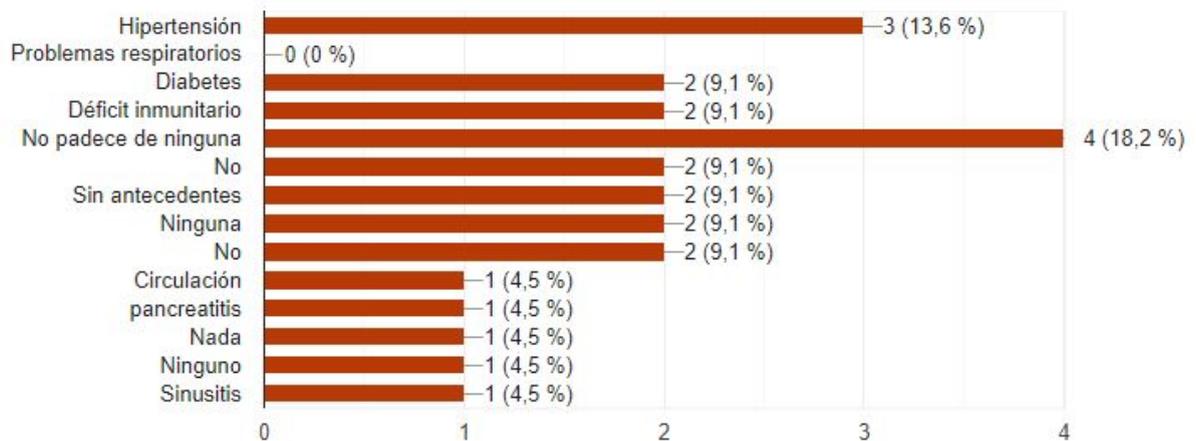
22 respuestas



Según los datos, la mayoría de los encuestados no padecen de alguna condición de las citadas en la misma. Sin embargo, un 13,6% de los mismos indicaron ser hipertensos, otro 9,1% de estos contar con diversas de las enfermedades citadas como diabetes y déficit inmunitario y un 4,5% contar con problemas de circulación, pancreatitis y sinusitis.

¿Padece de alguna de las enfermedades de alto riesgo para covid-19?

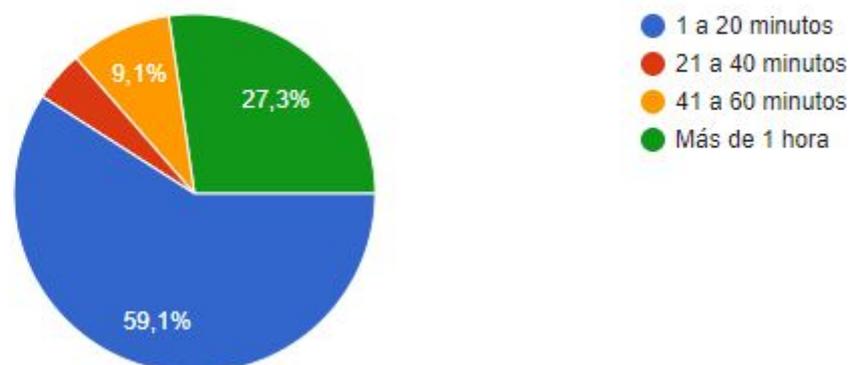
22 respuestas



De igual forma, un 59.1% de los pacientes reflejaron igual que los profesionales encuestados con relación a que el tiempo de realizarse la prueba tarda de 1 a 20 minutos o más de una hora como indicó un 27,3%, pero adicionalmente en menor porcentaje 9,1% indicaron que puede demorar entre 21 a 60 minutos, sin contar el tiempo previo para llegar a obtener una cita que puede durar entre 1 día a 1 semana.

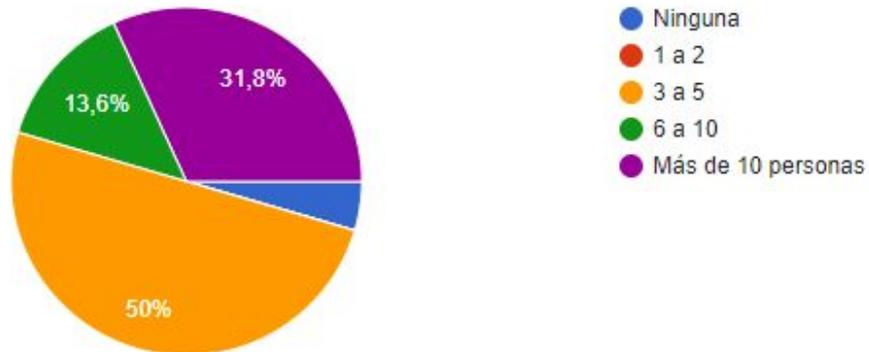
¿Cuánto tiempo ha tardado el proceso para hacerse una prueba de covid-19?

22 respuestas



¿Alrededor de cuántas personas conocidas sabe usted que han solicitado realizarse la prueba?

22 respuestas



De esas personas conocidas, ¿Cuánto tiempo aproximado han tardado en obtener una cita?

22 respuestas



Existen muchos motivos por los cuales los pacientes pueden acudir a realizarse una prueba de COVID-19, y esto provoca que la capacidad de los centros se vean desbordados, hecho que se podría mitigar si se priorizan los casos de una forma más ágil, considerando como prioritario aquellos cuyas condiciones lo ameriten. Conforme a los datos, el 50% de estos acudieron por diversos síntomas relacionados a los que provoca el virus, otro 18,2% por brotes en la institución donde labora y el resto por prevención ante casos cercanos, curiosidad o por ser una condición para otros análisis que requieren realizarse.

¿Qué le hizo tomar la decisión de realizarse la prueba?

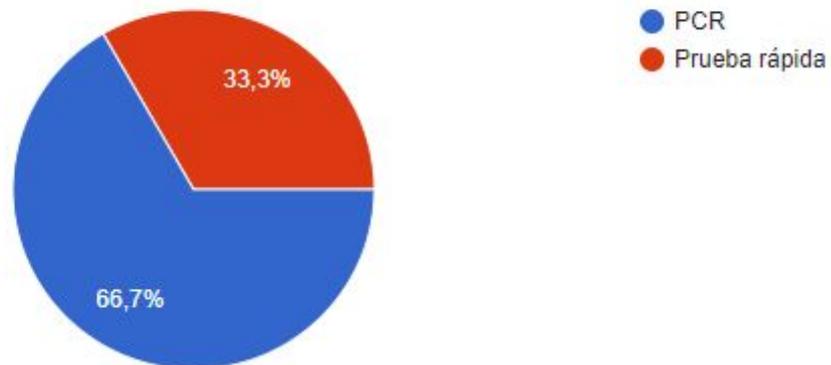
22 respuestas



Hemos podido comprobar que la mayoría de los pacientes acuden a los centros a realizarse las pruebas proteína C reactiva (PCR), porque prefieren la misma ante las pruebas rápidas conforme a la fiabilidad que les ofrece, un 66,7% de los encuestados soportaron estas informaciones, mientras que el resto con un porcentaje de 33,3% se estuvo realizando una prueba rápida.

¿Qué tipo de prueba se realizó?

21 respuestas



2.3 Análisis preliminar de datos

Dataset de radiografías y tomografías de pacientes sospechosos de COVID-19

Contamos con dos datasets. El primero recopila radiografías y tomografías de pacientes sospechosos de COVID-19 de múltiples repositorios públicos un aproximado de 904 imágenes de radiografías y tomografías de 454 distintos pacientes que han sido sospechosos principalmente de COVID-19 u otra neumonía viral (SARS, ARDS y MERS), a nivel internacional, y comprende otras informaciones acerca de la imagen tomada y el paciente.⁸ (Ver anexo A)

Diagrama Entidad-Relación

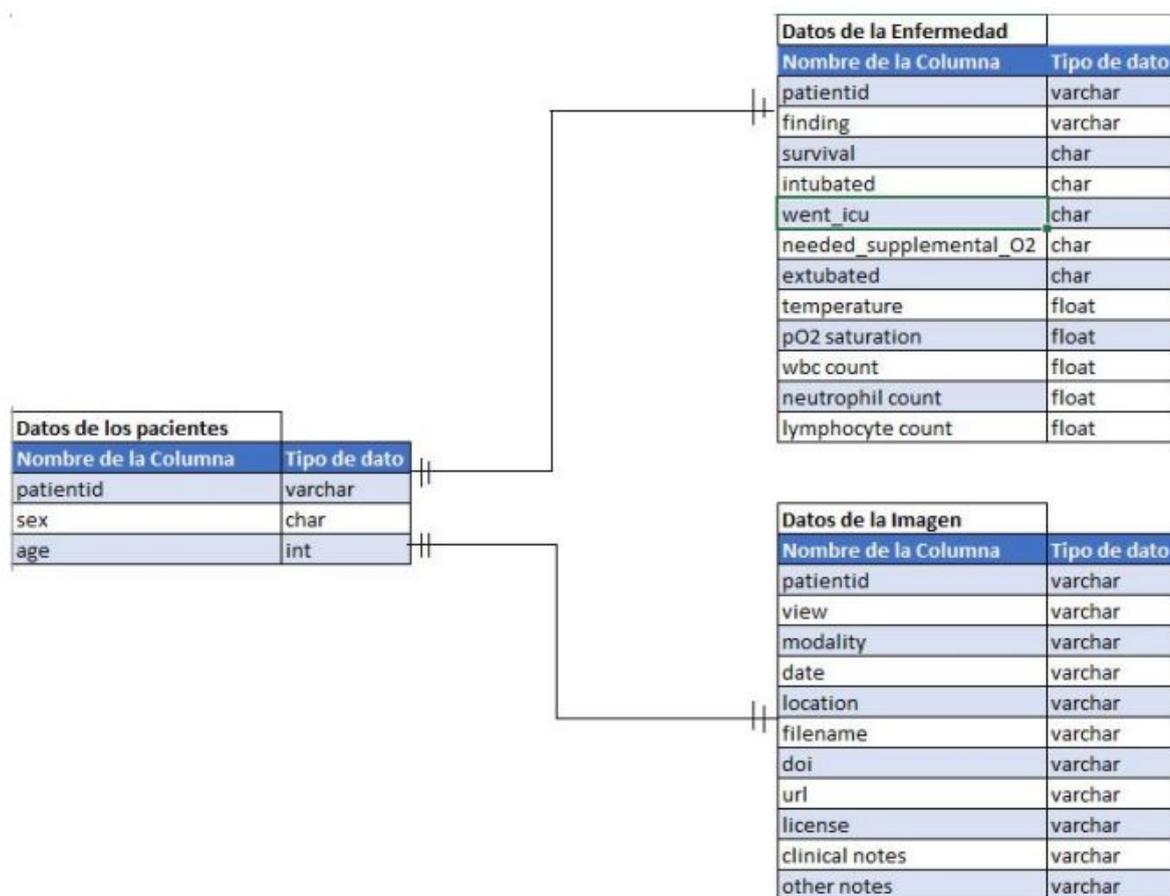


Figura 4. Diagrama Entidad-Relación de la fuente de datos de pacientes con radiografías y/o tomografías, sospechosos a COVID-19 y otras neumonías.

⁸ "ieee8023/covid-chestxray-dataset: We are" <https://github.com/ieee8023/covid-chestxray-dataset>. Se consultó el 26 ago.. 2020.

El segundo repositorio de datos a usar, cuenta con aproximadamente 265,000 pruebas que se han realizado a dominicanos al 31 de julio del presente año. Cuenta con datos de los pacientes, si sobrevivieron o no, enfermedades según su historial clínico, donde fue hospitalizado, entre otros. (Ver anexo B)

Datos Generados

Entre los datos que se pueden generar a partir del resultado del análisis los datos e imágenes recolectados, son:

- Probabilidad de tener un tipo de neumonía, principalmente COVID-19.
- Indicar el nivel de riesgo de un paciente.
- Priorización de pacientes.

Con el objetivo de aumentar la precisión de nuestro modelo, le daremos la opción a los médicos de los hospitales y clínicas que usen nuestros servicios, de poder enviarnos datos del paciente, así como rayos X o tomografías si es que tiene, si fue diagnosticado positivo o no de COVID-19, etc.

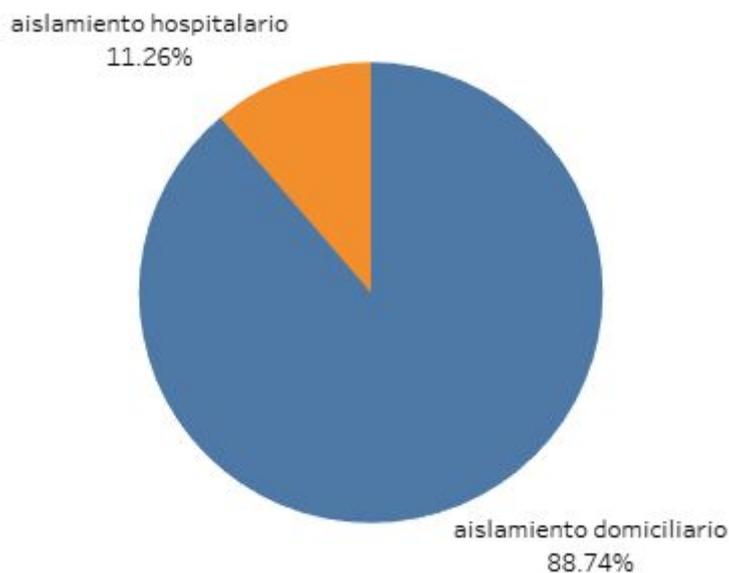


Figura 1. Porcentaje de pacientes aislados en hospitales y en domicilios en Rep. Dom. Capturado el 31 de Julio 2020.

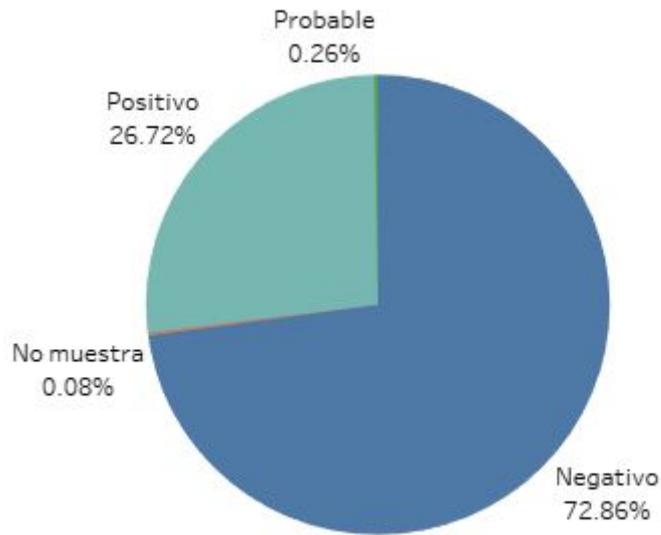


Figura 2. Casi el 73% de las pruebas realizadas a posibles sospechosos de COVID-19 en Rep. Dom. han resultado negativas. Capturado el 31 de Julio 2020.

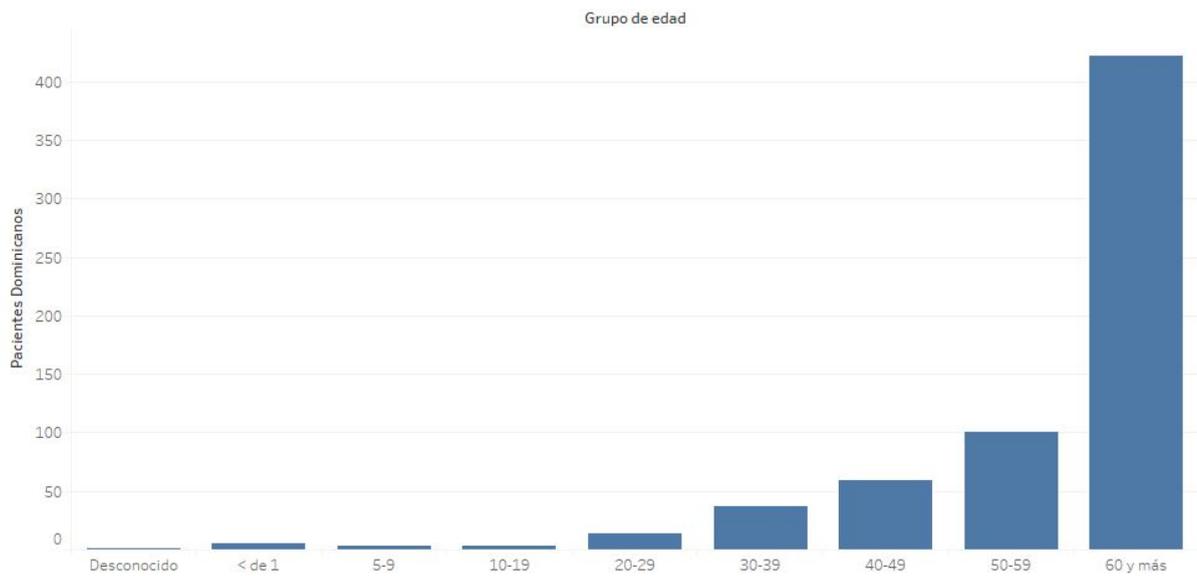


Figura 3. Cantidad de pacientes dominicanos por edad que resultaron positivos a la prueba de COVID-19 y han fallecido. Capturado el 31 de Julio 2020.

2.4 Análisis entorno

Por la naturaleza de nuestro proyecto, hay una gran cantidad de factores del entorno que tienen gran influencia en el mismo. Los principales son:

- **Factores político-jurídico:** Actualmente en República Dominicana, se presenta un cambio de gobierno que puede afectar la rapidez y las negociaciones que se pueden realizar con el sector de salud pública al estar todavía en plena transición.
- **Factores económicos:** Es incierto la factibilidad económica que podamos tener con el proyecto ya que actualmente nos encontramos en una profunda crisis económica provocada por el COVID-19. En el área de la salud, se están alojando todos los recursos para contener la epidemia y puede ser difícil una inversión en métodos tecnológicos aún no probados.
- **Factores tecnológicos:** Uno de los factores más importantes es la creación de la vacuna contra el COVID-19 que actualmente lidera Rusia, la cual de probar satisfactoria y de contar con las cantidades suficientes para vacunar un buen porcentaje de la población, pero en caso de que la vacuna fuese efectiva, el proyecto se podría reutilizar para la detección y priorización de otros tipos de enfermedades pulmonares.

2.5 Análisis competitivo

El continuo crecimiento de la tecnología Big Data ha permitido superar problemas en varios ámbitos, tanto profesionales como de la vida cotidiana. Uno de estos problemas ha sido la duración a la hora de detectar ciertas patologías en los pacientes.

Por esto, en la última década han aparecido algoritmos cuyo propósito es el detectar diferentes enfermedades como puede ser la pulmonía utilizando el Big Data para analizar las imágenes de las radiografías, aunque en República Dominicana se están tomando en cuenta este tipo de estudios, no se está utilizando un proceso automatizado que ayude a identificar la severidad en los pacientes. Con respecto a la detección del COVID-19 utilizando estos algoritmos, ya han aparecido algunas propuestas interesantes de entre las que se encuentran:

Xi'an Jiaotong University: Es una institución de educación superior de China, en la cual se ha propuesto el análisis y detección de COVID-19 utilizando Redes neuronales convolucionales y Machine learning, el cual ya habían utilizado anteriormente para la detección de neumonía con imágenes radiográficas, con la que alcanzaron un acierto de detección del 96%. Ahora planean utilizar el mismo sistema para poder clasificar los resultados entre COVID-19, Normal o neumonía.⁹

⁹ "Classification of COVID-19 from Chest X-ray ... - medRxiv." 18 jun.. 2020, <https://www.medrxiv.org/content/10.1101/2020.05.01.20088211v2>. Se consultó el 27 ago.. 2020.

Bluedocai: Blue Eye Soft es una compañía enfocada en proveer servicios de gestión y soluciones IT con software personalizados a grandes y medianas empresas en norte América. Están desarrollando un software llamado Bluedocai, el cual tiene la capacidad de analizar rápidamente imágenes radiográficas para detectar la presencia del COVID-19 con una precisión mayor del 90%, esto lo hace utilizando Machine Learning e Inteligencia artificial.¹⁰

DarkCovidNet: Es un modelo propuesto para la detección del COVID-19 utilizando imágenes de rayos x. Tiene un acierto del 98% en caso de clasificación binaria (COVID-19 o normal) y un acierto del 87% en caso de clasificación múltiple (COVID-19, neumonía o normal). Este utiliza mapas de calor para localizar las regiones afectadas en la radiografía.¹¹

COVID-19 Sounds App: La Universidad de Cambridge y el European Research Council están apoyando el desarrollo de una APP que tiene como objetivo el de detectar el COVID-19 con el sonido de la voz, respiración y tos, utilizando algoritmos de aprendizaje automático.¹²

ACPSEM: Es una organización profesional de ingenieros biomédicos los cuales están realizando un modelo en el que se utilizan imágenes de Rayos X y transfer learning con Redes neuronales convolucionales para poder detectar el COVID-19, dentro de este modelo se utilizan diferentes redes neuronales para obtener diferentes resultados, como VGG19, MobileNet v2, Inception, Xception e Inception ResNet v2; Obteniendo resultados de aciertos de hasta 98%, 97%, 86% para clasificación binaria.¹³

¹⁰ "Blue Eye Soft Announces Commodity Designation and FDA" 20 jul.. 2020, <https://www.globenewswire.com/news-release/2020/07/20/2064630/0/en/Blue-Eye-Soft-Announces-Commodity-Designation-and-FDA-Acknowledgement-of-BLUEDOCAITM-an-AI-driven-Medical-Imaging-Clinical-Decision-Tool.html>. Se consultó el 27 ago.. 2020.

¹¹ "Automated detection of COVID-19 cases using deep neural" 28 abr.. 2020, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7187882/>. Se consultó el 27 ago.. 2020.

¹² "University of Cambridge - COVID-19 Sounds App." <https://www.covid-19-sounds.org/es/>. Se consultó el 27 ago.. 2020.

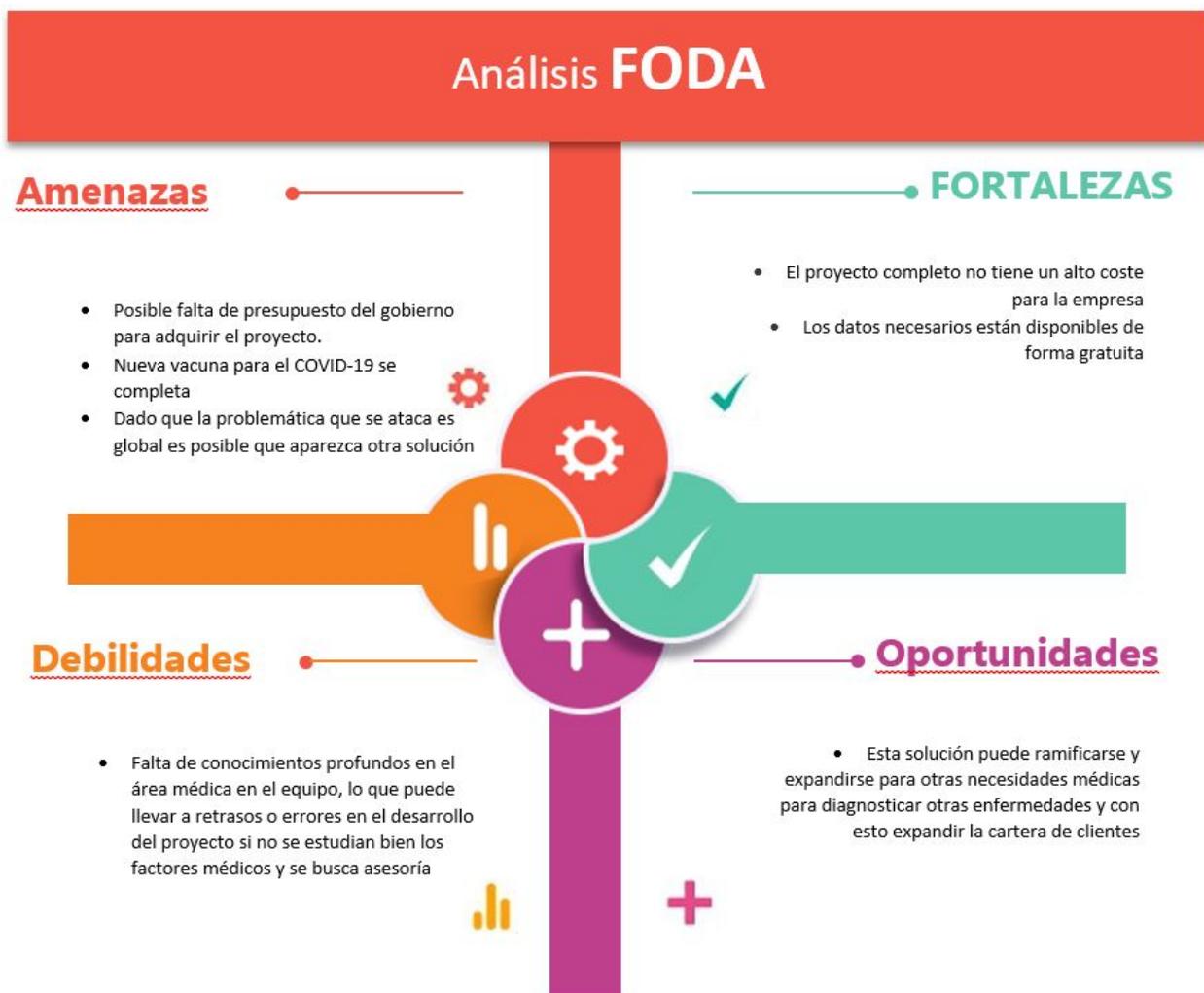
¹³ "Covid-19: automatic detection from X-ray images utilizing" 3 abr.. 2020, <https://link.springer.com/article/10.1007/s13246-020-00865-4>. Se consultó el 27 ago.. 2020.

4. Análisis y diagnóstico/Plan estratégico-acción

4.1 Análisis y diagnóstico

4.1.1 Análisis DAFO

Mediante el análisis DAFO podemos reconocer y actuar sobre todos los factores internos y externos que de alguna manera puedan afectar o beneficiar nuestro proyecto. Esto nos permitirá catapultar la probabilidad de éxito al enfocarnos en mejorar las debilidades, ser cautos con las amenazas, priorizar nuestras fortalezas y aprovechar las oportunidades.



Debilidades

Falta de conocimientos profundos en el área médica en el equipo, lo que puede llevar a retrasos o errores en el desarrollo del proyecto si no se estudian bien los factores médicos y se busca asesoría.

Amenazas

Nuestros clientes principales serían los hospitales públicos, por lo tanto, es posible que no logremos capitalizar la venta ya que este proyecto puede afectar intereses de funcionarios, así como también depende del presupuesto actual del gobierno que debido a la misma crisis sanitaria está muy recortado.

Al momento, tenemos 3 países desarrollando vacunas contra el COVID-19 y de alguno de ellos lograr completar todas las pruebas y poder comercializar en República Dominicana antes de nosotros concluir el proyecto, el mismo se podría tornar innecesario contra esta enfermedad.

Debido a que la problemática que estamos tratando es un tema global, es posible que aparezca una solución parecida que ataque esta necesidad y por lo tanto nos veamos en la necesidad de competir con la misma.

En caso de presentarse mutaciones, mientras el virus siga afectando a los pulmones, el modelo seguirá funcionando correctamente. En caso de presentarse una nueva característica, el modelo debe ser adaptado y entrenado para identificar la nueva anomalía.

Fortalezas

El proyecto completo no tiene un alto coste para la empresa. Los softwares a utilizar son gratuitos y los costos tecnológicos se pueden mantener en un mínimo hasta el momento de la venta ya que sólo cubrirán servidores y personal. Se cuenta con el personal necesario para el desarrollo del proyecto, el cual es una necesidad a nivel global y causaría un gran impacto positivo para el control de la enfermedad y ayudará a la carencia de pruebas COVID-19 que existe en la actualidad en el país.

Los datos que necesitamos en un inicio para hacer funcionar el proyecto están disponibles de forma gratuita y abierta.

Oportunidades

Esta solución puede ramificarse y expandirse para otras necesidades médicas las cuales con el conocimiento adquirido de terminar esta fase de proyecto es posible adaptar el algoritmo para diagnosticar otras enfermedades y con esto expandir la cartera de clientes.

4.2 Definición modelo de negocio

Para el análisis del modelo de negocio estaremos utilizando la herramienta de Canvas con el propósito de identificar cómo nuestro producto va a conseguir generar ingresos, así como señalar las áreas más críticas de trabajo y enfocar nuestros esfuerzos en ellas. Con esto estaremos en camino a asegurar una buena relación con los clientes a la vez que obtenemos los ingresos deseados.

Asociados claves

Nuestros principales socios claves se componen de los siguientes:

- Centros de salud
- Laboratorios clínicos
- Inversionistas
- Ministerio de salud pública

Actividades claves

- Desarrollar la solución.
- Promocionar nuestro sistema a los diferentes hospitales y contactarnos con ellos para dar a conocer el proyecto.
- Establecer un vínculo de confianza con los clientes sobre el sistema para desarrollar una red de usuarios.
- Instalar nuestros sistemas en los hospitales y alimentar el banco de datos lo suficiente para mantener un sistema robusto.
- Cuando ya se tenga una base establecida y todo esté ejecutándose como es debido, es clave la mejora continua del sistema a través de un sistema de feedback para seguir brindando un servicio de calidad.

Recursos claves

- Recursos físicos: Equipos informáticos.
- Recursos económicos: Apoyo económico por parte del gobierno o entidades privadas que requieran de la solución.
- Recursos intelectuales: Datos, licencias, código fuente, etc..
- Recursos humanos: Personal encargado de la parte administrativa, financiera, tecnológica y de marketing.

Propuesta de valor

Sistema computarizado y entrenado para responder ante probabilidades de personas tener el virus conforme a lo detectado en las radiografías, como punto de referencia para poder discernir los casos que podrían ser positivos, de los negativos.

Entre los beneficios que proporciona se encuentran:

- Agilizar el proceso para realizar las pruebas.
- Reducir la propagación de los casos, ya que las personas tendrán acceso a los resultados de una forma más rápida, según sea su situación.
- Reducir el congestionamiento de hospitales.
- Atención priorizada de los casos de la enfermedad.
- Detección de la posible enfermedad de forma automatizada.
- Brindar atención a las personas que puedan desarrollar mayores complicaciones al contraer el virus.

Canales

Se realizará promoción de la solución a través de redes sociales, página web, así como ventas directas ofreciendo el sistema a los potenciales clientes.

Relación con los clientes

Al inicio del proyecto, se pretende fidelizar a los clientes (hospitales) mostrando los resultados de sus pacientes utilizando nuestro sistema, que podrán ser corroborados con pruebas si así lo desean. Al demostrar que nuestro sistema es confiable y tiene un aporte real optimizando el proceso de las pruebas, podrán ayudarnos alimentando al sistema con nuevos datos.

Una vez ya establecido en varios hospitales, laboratorios clínicos y otros sitios con autoridades competentes a realizar las pruebas, los usuarios podrán reportar problemas respecto al sistema a través de la página para poder solucionar los inconvenientes, así como poder contactarnos vía telefónica para solicitar apoyo remoto.

Segmento de Clientes

- Nuestros cliente objetivo son los hospitales, los cuales adquirirán el sistema entrenado para brindar la atención eficaz a los beneficiarios quienes serían el público en general cuyos síntomas lo llevan a estar interesados en detectar si poseen el virus.
- Los laboratorios clínicos y demás lugares habilitados por las autoridades competentes para realizar este tipo de pruebas.

- Como potenciales consumidores de los datos finales, se encuentran las autoridades de salud y presidencia quienes a través de estos pudieran mantenerse informados del estado actual del país y tomar decisiones oportunas.

Costes

Nuestra estructura de costos es principalmente:

- ◆ Servicio cloud y storage
- ◆ Desarrollo frontend y backend de la web de acceso
- ◆ Salario del equipo desarrollador
- ◆ Publicidad y estrategias de introducción al mercado

Ingresos

- Ingresos por pago a suscripción por servicio para acceder a la plataforma de parte de los centros médicos e instituciones gubernamentales como el ministerio de salud pública para habilitar su uso en hospitales.
- Soporte Post-Venta

4.3 Plan de acción

4.3.1 Definición del alcance del proyecto: objetivos y métricas

4.3.1.1 Objetivos

En nuestro plan de acción contamos con objetivos que son nuestra guía para el avance del proyecto.

- Alcanzar un nivel de productividad mayor al porcentaje de confianza de las pruebas rápidas de modo que nuestra solución provea una alternativa más eficiente a las pruebas rápidas para detección de COVID-19. Esta meta se debe lograr en un periodo de 3- 5 semanas.
- Marketing del producto en las clínicas privadas y oficinas del gobierno correspondientes para ir promoviendo el producto y asegurar clientes. Esto se debe hacer de 4 - 8 semanas.
- Preparar la data e integrar con el servicio de 3-4 semanas.
- Satisfacción del cliente mediante la actualización y soporte continuo, además recibiendo quejas y sugerencias para mejorar nuestro producto.

- Ambiente laboral relajado y efectivo. Para mantener la productividad alta se debe crear un ambiente de confianza y trabajo en equipo que fomente la innovación.

4.3.1.2 Métricas (KPIs), Cuadro de mando.

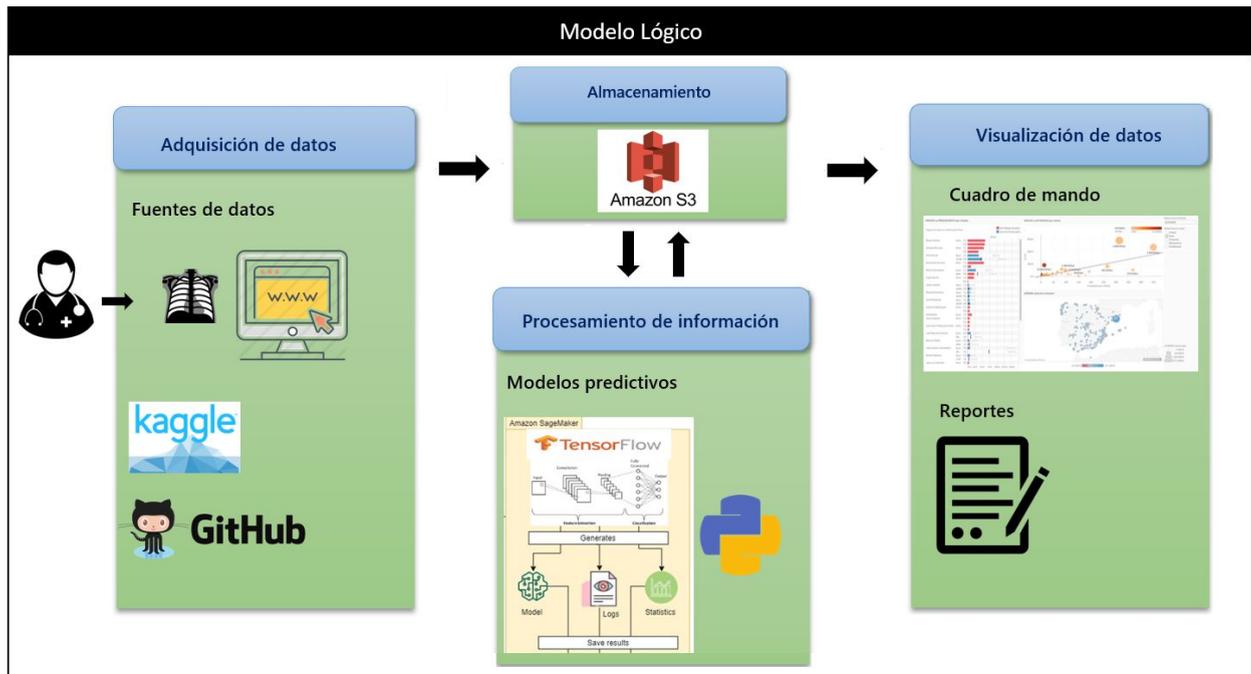
Los indicadores claves de rendimiento de nuestro proyecto se centran en la efectividad del producto, la satisfacción del cliente y de nuestros empleados. Mediante la continua monitorización de estos se nos permitirá mantener nuestros objetivos.

KPI	Target
Precision modelo predictivo	>95%
Numero de quejas por cliente	<10
Quejas reportadas vs quejas solucionadas	100%
Satisfaccion del empleado (poll)	>8.5 pts
Capacitacion especializada del personal	>95%

4.3.2 Análisis de actividades: Modelo lógico y Arquitectura técnica

4.3.2.1 Modelo lógico

El modelo lógico está compuesto por la parte de adquisición de datos , procesamiento de la información y la visualización de los datos.



1. Identificación y captura de los datos:

Los datos serán obtenidos de repositorios públicos que han sido identificados con datasets de imágenes de rayos X y tomografías, pero en un primer momento, sólo estaremos utilizando radiografías, la cual posee casos positivos como negativos del COVID-19, pacientes con neumonía o en condiciones normales.

Las imágenes que el laboratorio o clínica necesite clasificar y priorizar deberán ser cargadas en una página web y también podrán aportar con imágenes ya clasificadas por sus instituciones, para hacer más robusto el modelo.

Actualmente contamos con esta cantidad de radiografías:

Categoría	Total de Radiografías
Normal	1582
Neumonía	4272
COVID-19 Positivo	504
COVID-19 Positivo con Gravedad	114

2. Transformación y validación de los datos

Se realizará la depuración debida de los datos para asegurar calidad e integridad de los mismos, para que el modelo pueda realizar las predicciones de la forma más correcta posible y posteriormente pasar a la unificación de los mismos para poder ser almacenados y analizados.

En la base de pruebas de COVID-19 en República Dominicana, detectamos que hay muchos campos relacionados con el historial médico de los pacientes que la mayoría de valores se desconocen (Ver Anexo C). Se realizaron los siguientes cambios antes de visualizar los datos:

- Aquellos valores que sean inválidos o desconocidos, se reemplaza por NULL. Algunas de estas palabras son: “Desconocido”, “.”, “No muestra”, etc.
- Los valores de las columnas de *Muestra*, *Manejo* e *Ingreso en UCI* fueron llevados a minúscula.
- Se asigna NULL a aquellas edades que tengan un valor menor a cero.
- Se realizaron pequeñas correcciones manuales detectadas al verificar los valores únicos por cada columna. Por ejemplo: Un grupo de edad aparece como **19-Oct**, pero se entiende que el rango debe ser **10-19**.

Con respecto al dataset de radiografías, para aumentar la cantidad de imágenes de nuestro dataset, se hizo uso de una técnica llamada **Data Augmentation**, la cual que consiste en generar, de forma aleatoria, copias a partir de imágenes ya existentes con ligeros cambios. Los criterios utilizados fueron los siguientes:

- Reescalado del pixel: [0, 1]. Debido a que cada píxel contiene tres valores entre 0 y 255, dividimos estos números entre 255 para reducir el rango significativamente.
- Rango de rotación: 20°
- Rango de desplazamiento (anchura y altura): 20%
- Rango de zoom: 20%

Su objetivo es disminuir el overfitting de un modelo de aprendizaje automatizado, es decir, impedir que dicho modelo pueda predecir significativamente mejor los datos con los que ha

sido entrenado en comparación con los datos que aún no ha visto, no generalizando como debería.

La resolución de las imágenes son cambiadas a 224x224x3 antes de ser entrenados para que puedan ser usadas al entrenar el modelo.

De las 30 imágenes que indican la gravedad de pacientes con COVID-19 obtenidas del sitio web Radiology Class¹⁴, como la mostrada debajo, se extrajo la primera mitad que representa la imagen original:



Los criterios que se usaron para clasificar la gravedad de la neumonía por COVID-19 son:

- Leve: Se categoriza de esta forma cuando la radiografía presenta un porcentaje mayor de color negro en comparación al color blanco.
- Moderado: La radiografía presenta un porcentaje similar de ambos colores.
- Severo: La radiografía muestra un porcentaje mayor de color blanco en comparación al color negro.

Las imágenes restantes fueron clasificadas según el criterio anteriormente mencionado y la valoración en opacidad de 3 médicos sobre dichas imágenes¹⁵. Para estos datos, se seleccionó como criterio el nivel de opacidad de la radiografía, cuya puntuación está en escala del 0 al 6:

- Leve: El nivel de opacidad es menor a 2.3.
- Moderado: El nivel de opacidad está entre 2.3 y 3.7.
- Severo: El nivel de opacidad es mayor o igual que 3.7.

3. Almacenamiento de los datos:

Se estará utilizando Amazon S3 para la parte de almacenamiento, donde serán unificados los datos obtenidos de las diferentes fuentes, luego de haberse asegurado la calidad de los mismos.

¹⁴ "COVID-19 Image Gallery - Classification of severity."
<https://www.radiologymasterclass.co.uk/gallery/covid-19/covid-19-image-gallery-2>. Se consultó el 14 oct.. 2020.

¹⁵ "covid-chestxray-dataset/covid-severity-scores.csv at master"
<https://github.com/ieee8023/covid-chestxray-dataset/blob/master/annotations/covid-severity-scores.csv>.
Se consultó el 14 oct.. 2020.

Cada imagen descargada se coloca en una carpeta que coincida con la categoría a la que pertenece. Los datos estructurados estarán guardados en archivos .csv con el formato indicado (Ver Figura 4).

4. Descubrimiento y modelado:

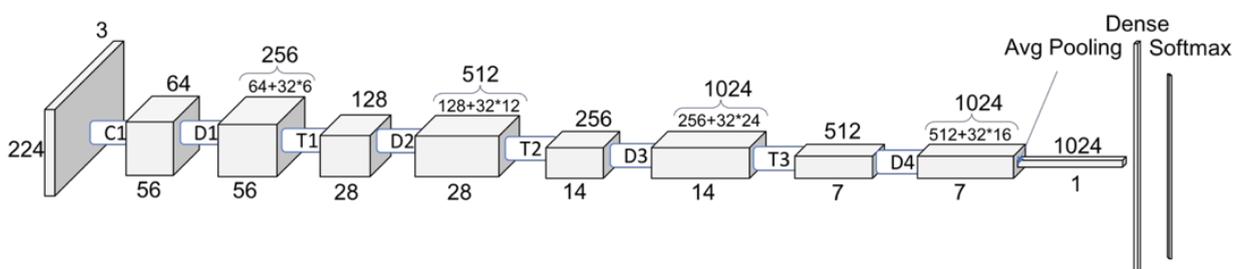
En esta fase, se utilizarán los distintos modelos predictivos para realizar los análisis y hacer la detección de los casos con neumonía, los positivos en covid-19, así como la priorización de los mismos.

Para la selección de modelos de aprendizaje automatizado, se hizo uso de un paper que muestra el proceso de clasificar radiografías con redes neuronales, incluyendo pacientes con COVID-19¹⁶. Estos modelos fueron entrenados previamente con millones de imágenes para clasificar una imagen en 1000 categorías diferentes, sin embargo, fueron entrenadas adicionalmente con las radiografías ya mencionadas y fue sobrescrita la capa más alta de estos para predecir sólo entre 2 y 3 clases.

Hacemos uso de tres modelos:

- A. **VGG-16:** Clasificación binaria entre Neumonía o Normal.
- B. **Densenet-169:** Clasificación binaria entre COVID-19 Positivo o Negativo. Se hace uso de este modelo sólo si se detecta que el paciente tiene neumonía o el médico así lo indica en la página web.
- C. **Densenet-169:** Clasificación múltiple de la gravedad del COVID-19, si es Leve, Moderado o Severo. Se hace uso de este modelo sólo si se detecta que el paciente es positivo al COVID-19 o el médico así lo indica en la página web.

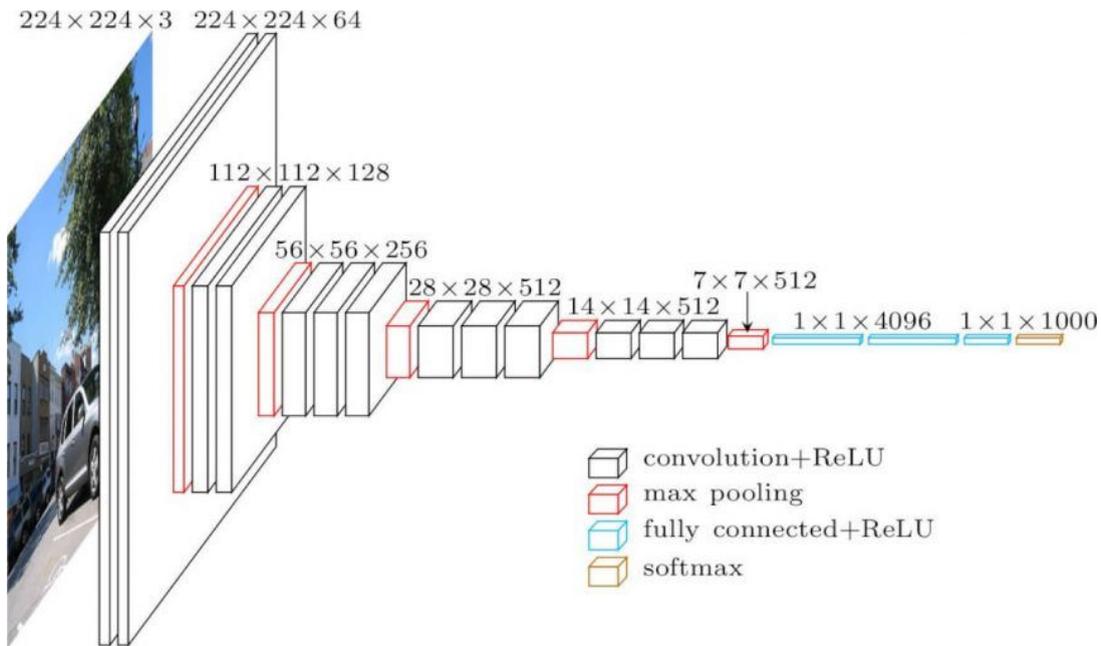
A continuación, se muestran ambas arquitecturas seleccionadas:



Arquitectura Densenet-169 ¹⁷

¹⁶ "Chest X-ray classification using Deep learning for automated" 23 jun.. 2020, <https://www.medrxiv.org/content/10.1101/2020.06.21.20136598v1>. Se consultó el 14 oct.. 2020.

¹⁷ "Understanding and visualizing DenseNets | by Pablo Ruiz" 10 oct.. 2018, <https://towardsdatascience.com/understanding-and-visualizing-densenets-7f688092391a>. Se consultó el 14 oct.. 2020.



VGG-16 ¹⁸

Las redes neuronales tienen algunas diferencias con respecto al documento basado originalmente, las cuales son:

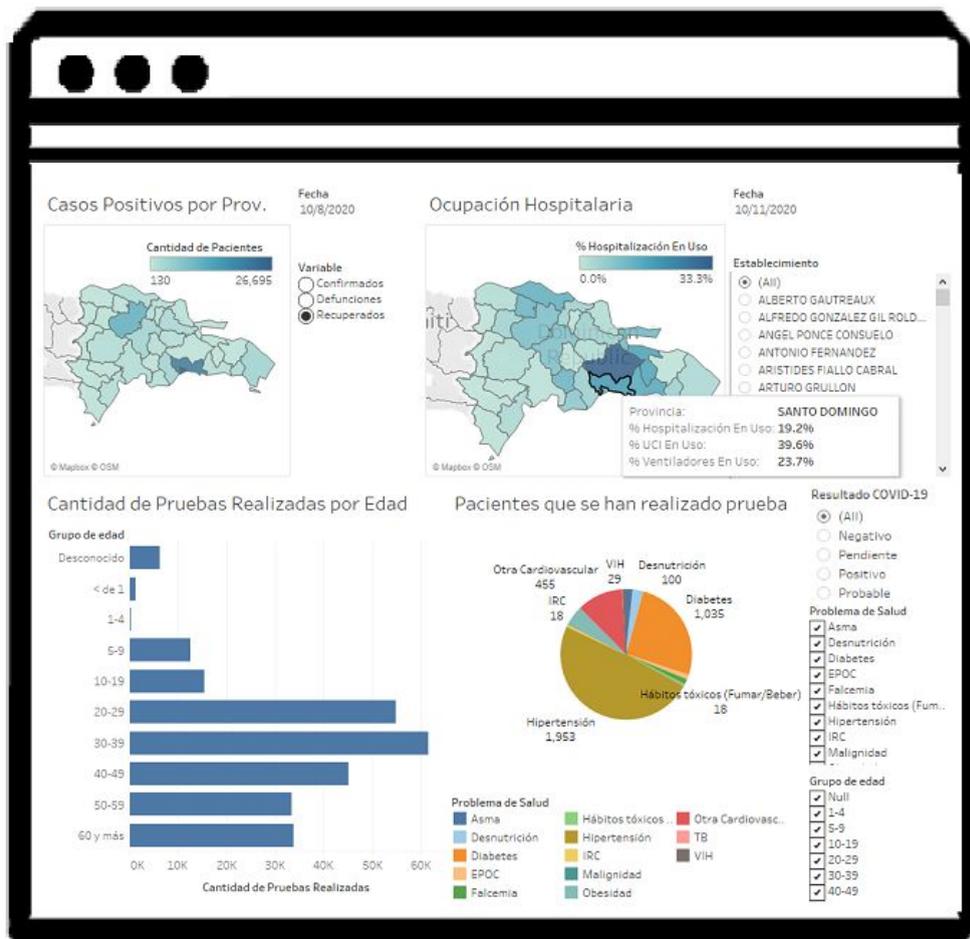
- A. En el caso de los dos modelos B y C, en vez de usar el mejor modelo, seleccionamos el segundo con mayores resultados según el paper, debido a que no se encuentran ya entrenados en Tensorflow, librería que seleccionamos.
- B. Las imágenes que fueron usadas para entrenar constan de 3 canales, a diferencia del paper que sólo usa 1 canal, es decir, radiografías en escala de grises. También fueron inicialmente reescaladas a 64 x 64 píxeles.
- C. Algunos de los hiper parámetros de los modelos, como el ratio de aprendizaje (learning rate), el número de ciclos (epochs), entre otros, fueron modificados ligeramente.

Debido a esto, es claro que no se obtendrán los mismos resultados que en el trabajo de investigación que se basó el presente trabajo, sin embargo, los resultados se acercan aunque siguen estando por debajo.

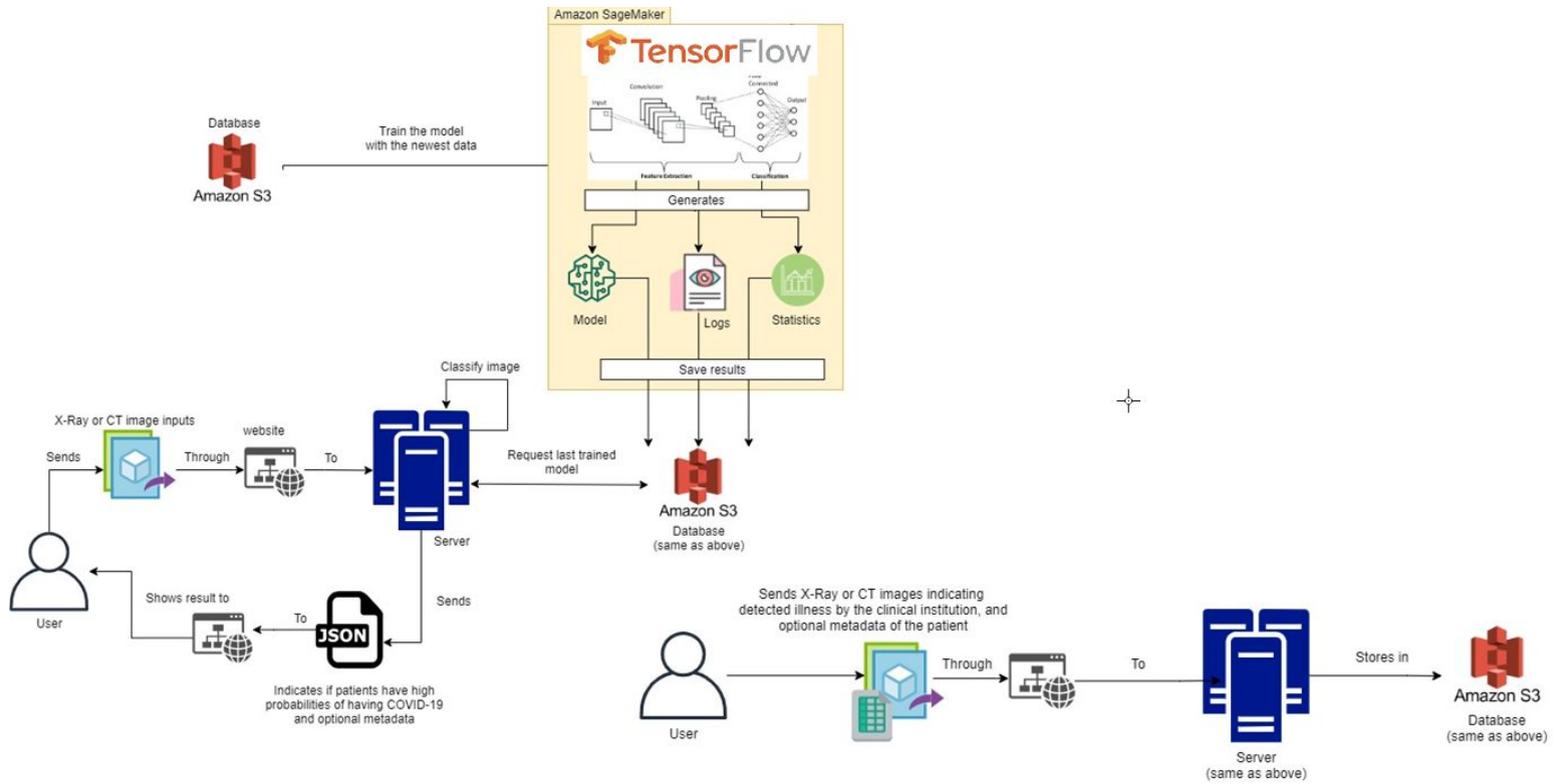
¹⁸ "VGG16 - Convolutional Network for" 20 nov.. 2018, <https://neurohive.io/en/popular-networks/vgg16/>. Se consultó el 14 oct.. 2020.

5. Visualización de datos:

Será realizado un cuadro de mando para la presentación gráfica de los casos identificados, gravedad y por los aspectos que el usuario considere convenientes, ya sea la edad, que el paciente necesite urgentemente realizarse una cirugía, presente síntomas graves etc..



4.3.2.2 Solución tecnológica: arquitectura técnica



Flujo Básico de Procesos Técnicos

El sistema permite subir una o más imágenes para predecir si el paciente tiene COVID-19 o no.

El usuario sube una o más radiografías en una página web, además, puede indicar si ya han sido previamente clasificadas por alguna entidad hospitalaria, de forma tal que el modelo se enfoque en predecir las categorías desconocidas.

Una vez el servidor recibe la imagen, lo clasifica de acuerdo con el último modelo entrenado disponible en las siguientes categorías especificando la probabilidad de que haya acertado. Este es el listado de clasificaciones predecidas:

- a. NORMAL
- b. NEUMONÍA
 - i. COVID-19 POSITIVO
 - 1. Leve
 - 2. Moderado
 - 3. Severo
 - ii. COVID-19 NEGATIVO

Los datos se compartirán en formato JSON, como se muestra en el siguiente ejemplo:

```
{
  "data": [{
    "id": 1,
    "nombre_del_archivo": "x-ray-01.png",
    "Categoría": "COVID-19",
    "Gravedad": "Leve"
  },
  {
    "id": 2,
    "nombre_del_archivo": "x-ray-02.jpg",
    "Categoría": "PNEUMONÍA",
    "Gravedad": null
  },
  // ...
]
```

Finalmente, la página web (cliente) recibe el resultado del análisis al usuario.

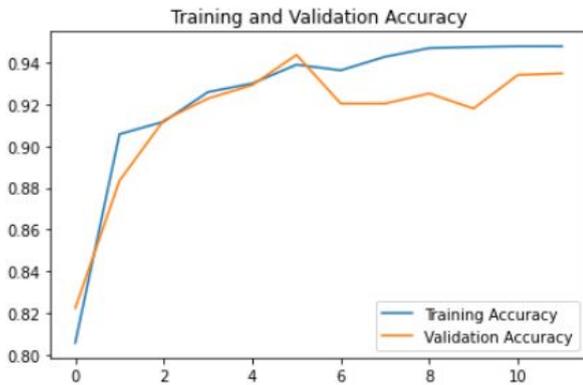
El sistema también les permitirá a los médicos subir radiografías que ya han sido clasificadas por la clínica u hospital en que laboran, y opcionalmente pueden añadir otros datos sobre los pacientes, que serían utilizados para entrenar el modelo con un mayor número de imágenes y, por tanto, esperando que aumente la probabilidad de acertar.

Los modelos se entrenarán, inicialmente una vez al día, en un servidor a parte del servidor web y con una alta capacidad de GPU.

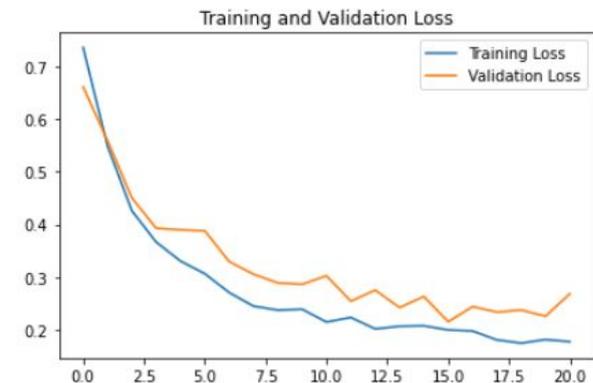
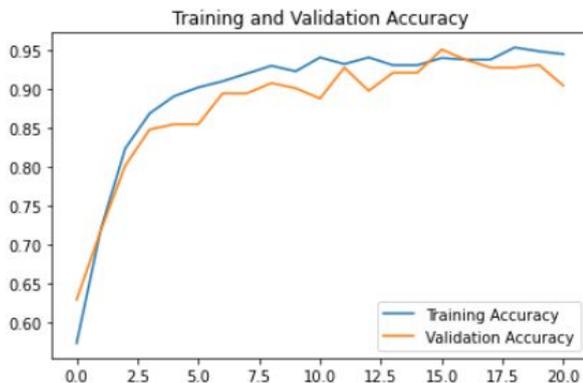
A continuación, se muestran detalles y métricas de los modelos entrenados:

No.	Modelo	Epochs	Learning Rate	Métrica	Training Set	Validation Set	Test Set
1	VGG-16	12	2e-05	Loss	0.1343	0.1576	0.2758
1	VGG-16	12	2e-05	Accuracy	0.9478	0.9349	0.9022
1	VGG-16	12	2e-05	Precision	0.9475	0.9356	0.9048
1	VGG-16	12	2e-05	Recall	0.9482	0.9341	0.8990
1	VGG-16	12	2e-05	AUC	0.9895	0.9855	0.9545
2	DenseNet-169	20	1e-06	Loss	0.1776	0.226	0.1589
2	DenseNet-169	20	1e-06	Accuracy	0.9444	0.9305	0.9567
2	DenseNet-169	20	1e-06	Precision	0.9407	0.9333	0.9536
2	DenseNet-169	20	1e-06	Recall	0.9487	0.9272	0.96
2	DenseNet-169	20	1e-06	AUC	0.978	0.9666	0.9845
3	DenseNet-169	23	1e-05	Loss	0.2472	1.31	1.2967
3	DenseNet-169	23	1e-05	Accuracy	0.8904	0.68	0.72

Cantidad de imágenes usadas en cada modelo por categoría y uso del dataset.



Resultados Modelo #1



Resultados Modelo #2

Debido a la poca cantidad de imágenes para entrenar, podemos observar que hay un overfitting en el modelo, es decir, no está generalizando.



Resultados Modelo #3

4.3.3 Análisis de recursos: talento humano y recursos físicos

Para el objetivo y/o funciones del proyecto se requiere de la participación de los siguientes profesionales que de una forma u otra representan y aportan valor al mismo:

4.3.3.1 Talento humano

1. Estructura Organizativa



2. Denominación del recurso humano y tareas asignadas

- **CEO:** director ejecutivo, quien será la cara del proyecto y asumirá la responsabilidad de liderar el mismo en términos de que fluya y funcione de la manera correcta.
- **Director administrativo y financiero:** encargado de coordinar todas las tareas contables, administrativas y financieras, así como lograr financiación con bancos y proveedores.
- **Director comercial:** encargado de la estrategia comercial del proyecto en colaboración con la Dirección ejecutiva, de manera que pueda ser contratado el sistema una vez ejecutado.
- **Analista:** el analista constituye el perfil clave y requiere experiencia en adaptación de modelos, conocimientos diversos de las tecnologías que estaremos utilizando y buena presencia y relaciones interpersonales para que sirva de apoyo a los implementadores y en la capacitación de los potenciales clientes.
- **Implementadores:** dentro de los implementadores se encuentra un desarrollador, quien será el responsable de llevar a cabo el desarrollo del sistema conforme a los requisitos y modelos descritos por el/la analista y un ingeniero en infraestructura que prestaría servicios de soporte al mismo cuando se presenten temas que internamente no se puedan resolver.

Conforme a lo antes descrito, el equipo estará compuesto por 5 miembros principales y un prestador de servicios, quienes de manera conjunta trabajarán en colaboración constante para lograr el objetivo, adicional a estos, el proyecto no contará con más personal fijo debido a que no es sustentable de manera inicial y quedará como un punto dependiente de las necesidades que se presenten u otros futuros proyectos.

Dos de estos miembros y el prestador de servicio, se dedicarán a la parte técnica del proyecto, lo cual contempla el desarrollo de la solución tecnológica, la implementación de los modelos que permitan la evaluación de los pacientes y respondan ante los mismos en función a los síntomas y condiciones, y los otros miembros restantes se dedicarán a la parte del negocio que contempla crear la estrategia a emplear, el financiamiento, búsqueda de prospectos y demás factores asociados.

En primera instancia, las cargas de trabajo de nuestra estructura organizativa será horizontal y enfocada a resultados, esto presidido a que puedan responder a los requerimientos con precisión y calidad.

4.3.3.2 Recursos físicos

Para la realización del proyecto se han utilizado los recursos de cada uno de los integrantes del grupo, ya sea ordenadores o dispositivos móviles. Debido a la situación actual de la pandemia global, no han sido necesarios gastos de transporte para realizar las reuniones, tampoco para realizar las encuestas. Los detalles de los recursos en cuanto a la infraestructura física y servicios externos son:

- 5 ordenadores, uno por cada integrante del grupo.
- Softwares de uso libre: Python, Excel, Google docs y Trello.
- Para la coordinación y reunión del grupo: WhatsApp, Google Hangouts y Google Meeting.
- Servicio de conexión de internet.
- Servicio de servidores en línea.

4.3.4 Gestión del tiempo (cronograma)

4.3.4.1 Diagrama de Gantt

En el diagrama de Gantt se representarán las diferentes actividades a realizar durante el proyecto de Priorización de Casos Covid-19, para facilitar la visualización del mismo, se ha dividido en dos fases, Planificación e implementación.

Planificación: En esta fase se representan las actividades de planificación del proyecto, como el planteamiento del problema, la investigación y la toma de datos, la creación del modelo de negocios y los beneficios del proyecto.

Implementación: En esta fase se muestran las actividades respecto al desarrollo, pruebas e implementación del proyecto.

5. Optimización de resultados

5.1 Análisis financiero

En este apartado vamos a estar desglosando el detalle del análisis financiero realizado para el proyecto y así validar su viabilidad, en donde mostraremos la inversión inicial que se debe de realizar para poder iniciar con la implementación del mismo, así como los ingresos proyectados según el resultado de la captación de clientes. También se estará incluyendo el cálculo para indicar la rentabilidad que tendría el desarrollo del mismo.

Esto nos ayudará a poder tener una visión clara de lo que se necesita para poner en marcha el proyecto. Esta solución beneficiaría a nuestros clientes en la reducción de gastos, debido a que las pruebas rápidas de Covid-19 que se realizan actualmente salen más costosas que teniendo nuestra solución, y estas tampoco ofrecen una gran precisión.

El proyecto estará enfocado en la detección y priorización de casos de covid-19, pero también tiene la capacidad de detectar neumonías, por lo que se estarán realizando mejoras para que en el futuro se utilice para detectar otras enfermedades pulmonares, como la tuberculosis.

5.1.1. Cuenta de resultados

Ingresos

Hemos realizado un análisis sobre la percepción de los ingresos que se estaría estimando, así como los costos y gastos que se deben de asumir para la implementación del proyecto.

Estos son los ingresos estimados que tenemos proyectado en un periodo comprendido en 3 años:

	Año 1 (semestres)			
Ingresos	Semestre 1	Semestre 2	Año 2	Año 3
Cantidad de Clientes Nuevos	5	20	10	5
Cantidad de Clientes Viejos (Año anterior)	0	0	20	25
Cantidad de Pacientes / día	50	200	200	300
Precio de Instalación (Sólo clientes nuevos)	\$ 1,000.00	\$ 1,000.00	\$ 1,050.00	\$ 1,100.00
Precio x Pacientes (x Radiografía)	\$ 3.50	\$ 3.50	\$ 2.80	\$ 2.50
Precio de mantenimiento	\$ 400.00	\$ 400.00	\$ 400.00	\$ 400.00
Ingresos x Clientes	\$ 7,000.00	\$ 28,000.00	\$ 22,500.00	\$ 17,500.00
Ingresos x Usuarios	\$ 31,937.50	\$127,750.00	\$204,400.00	\$273,750.00
Ingresos Totales	\$ 38,937.50	\$155,750.00	\$226,900.00	\$291,250.00

Los precios por cliente para la instalación fueron aumentados en un 5% anual para contrarrestar el impacto de la inflación y este solo se estará cobrando al adquirir el sistema. La cantidad de pacientes promedio diarios determina el mayor porcentaje del ingreso del proyecto, desde un 80% hasta un 94%, por lo que si esta cantidad aumenta en el tiempo, el precio se adaptará a esta demanda.

El precio de mantenimiento se cobrará en conjunto con el precio de la instalación en los clientes nuevos, para los clientes viejos se seguirá cobrando anualmente.

5.1.2 Inversión, costes y gastos

Se hará uso de servidores en la nube ofrecidos por Amazon EC2 y para el almacenamiento Amazon AWS.

Costos de Implementación	Año 1	Año 2	Año 3
Servidor p2.xlarge (4 horas diarias) - Precio/hora = \$0	\$ -	\$ 1,314.90	\$ 1,314.90
Servidor p2.xlarge (4 horas diarias) - Precio/hora = \$0	\$ 1,314.90	\$ 1,314.90	\$ 1,314.90
Servidor m6gd.2xlarge - Precio/hora = \$0.3616	\$ -	\$ 3,181.79	\$ 3,181.79
Servidor m6gd.2xlarge - Precio/hora = \$0.3616	\$ 3,181.79	\$ 3,181.79	\$ 3,181.79
AWS S3 - 200GB	\$ 4.60	\$ 55.20	\$ 55.20
Costos Totales	\$ 4,501.29	\$ 7,733.68	\$ 7,733.68

Debido a que para el año 2 y 3 se proyecta que se duplicará la cantidad de pacientes con relación al primer año del proyecto, se añadirán dos servidores adicionales a partir del segundo año para poder suplir la demanda de los nuevos usuarios del sistema. A continuación se detalla la funcionalidad principal de estos servidores:

- **Servidor p2.xlarge:** Estos servidores se especializan para trabajos en general que hagan uso significativo del GPU, por lo que será usado para entrenar los modelos de aprendizaje automatizado. Entre sus especificaciones técnicas se encuentran:
 - **GPU:** NVIDIA K80, hasta 24 GB.
 - **CPU:** 4 núcleos
 - **RAM:** 61 GB

Fue estimado que para entrenar los modelos diariamente es suficiente hacer uso de un máximo de 4 horas de este servidor, al menos con las radiografías recolectadas durante el primer año del proyecto.

- **Servidor m6gd.2xlarge:** Se hará uso de estos servidores para procesar todos los requests de nuestro sistema, incluyendo la predicción de la radiografía.
 - **CPU:** 8 núcleos
 - **RAM:** 32 GB

- **Almacenamiento:** 474 GB NVMe SSD
- **AWS S3:** Servirá como repositorio principal para almacenar el resultado de los modelos ya entrenados, backups, logs, radiografías y cualquier otro dato obtenido de los pacientes.

Otros costos incluyen el pago del dominio y certificado del sitio web del sistema, y un servicio de correos para mantener comunicación con nuestros usuarios, ya sea para enviar correos de propaganda, sobre actualizaciones o cualquier otra información. A continuación, se detallan estos servicios:

Otros servicios	Costo anual
Dominio y certificado del sitio web	\$ 30.00
Sendgrid (Mensajería)	\$ 180.00
Total	\$ 210.00

Costos de Personal

Estos son los costos de los recursos que serán necesarios para el desarrollo e implementación del proyecto. Actualmente se cuenta con el personal adecuado para la implementación de la solución, por lo que no habría que hacer una búsqueda del personal para el mismo.

RH	Coste mensual	Coste anual
CEO	\$ 1,400.00	\$ 16,800.00
Director administrativo y financiero	\$ 1,100.00	\$ 13,200.00
Director comercial	\$ 1,100.00	\$ 13,200.00
Analista	\$ 1,000.00	\$ 12,000.00
Implementadores (2)	\$ 1,800.00	\$ 21,600.00
Costos Totales	\$ 6,400.00	\$ 76,800.00

Gastos

Dentro de los gastos de publicidad, se tiene proyectado realizar una inversión en marketing de US \$60,000 el primer año, para poder captar la mayor cantidad de clientes posibles, luego en los siguientes años se estarían disminuyendo los costes, ya que se invertiría menos en publicidad. Esto siempre estará sujeto a la necesidad del negocio, ya que si la cantidad de clientes disminuye drásticamente habrá que hacer una inversión mayor, según lo necesario.

Gastos de Publicidad	Año 1	Año 2	Año 3
Marketing online y % representativo	\$ 50,000.00	\$ 30,000.00	\$ 20,000.00
Gastos Totales	\$ 50,000.00	\$ 30,000.00	\$ 20,000.00

Tenemos estimado que en el primer año estaríamos obteniendo una ganancia alrededor del 31% con respecto a los costos y gastos, luego en los siguientes años aumentarán en un 50% y 64% respectivamente, ya que los gastos de publicidad estarían disminuyendo, y se estarán captando más clientes y pacientes.

	Año 1	Año 2	Año 3
Ingresos	\$194,687.50	\$226,900.00	\$ 291,250.00
Costos y gastos	\$134,533.68	\$114,533.68	\$ 104,533.68
Ganancias (%)	31%	50%	64%
Neto	\$ 60,153.82	\$112,366.32	\$ 186,716.32

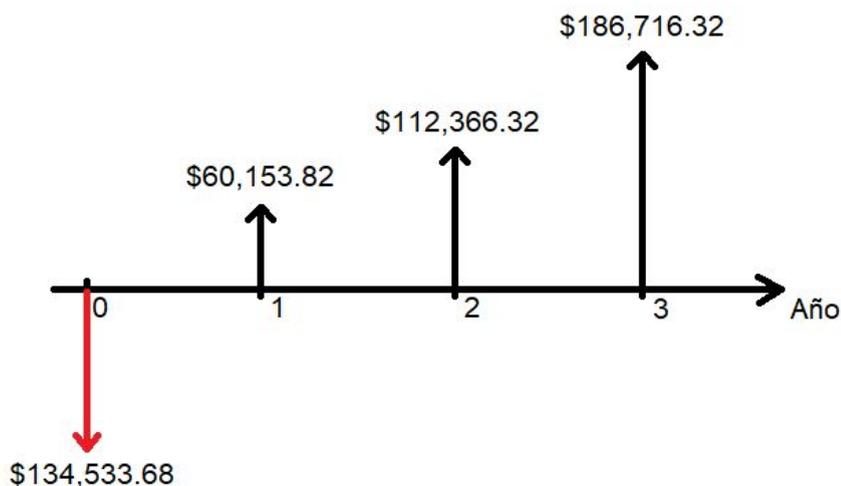
5.3.3 Rentabilidad económica

Flujo de Caja

En el flujo de caja representamos la inversión necesaria para iniciar con el proyecto, donde se incluyen los costos y gastos necesarios para implementar el proyecto. También se encuentran las ganancias netas que se estarán obteniendo a través de los años con los beneficios de los clientes que utilicen nuestros servicios.

A continuación se presenta el flujo con una proyección de 3 años.

Descripción	Año	Flujo de Caja
Inversión inicial	0	\$(144,533.68)
	1	\$ 27,341.32
	2	\$ 112,366.32
	3	\$ 184,716.32



Observamos de forma gráfica el flujo que se mantiene constantemente en aumento y de forma positiva, por lo que indica que hay ganancias a través del tiempo.

Teniendo esto en cuenta, calculamos la rentabilidad del proyecto, así como el tiempo de retorno de la inversión inicial realizada.

VAN	TIR	Pay-Back
\$287,832.75	56%	Al final del 2do año

Según se visualiza, la variable de VAN es mayor a 0, TIR es mayor a la tasa de interés colocada, la cual fue de un 10% y la inversión se estará recuperando al final del segundo año, por lo que el proyecto es financieramente viable.

Anexos

A. Diccionario de datos del dataset de radiografías y tomografías de pacientes sospechosos de COVID-19

- **Paciente ID:** Identificador único.
- **Offset:** Número de días desde el comienzo de los síntomas o la hospitalización para cada imagen. Si un informe indica después de unos días, entonces se supone que son 5 días. Esto es muy importante tenerlo cuando hay múltiples imágenes para que el mismo paciente pueda seguir la progresión..
- **Sexo:** Masculino (M), Femenino (F), o en blanco.
- **Edad:** Edad del paciente en años.
- **Enfermedad Encontrada:** Tipo de neumonía.
- **Supervivencia:** Sí (S) o no (N) o en blanco si se desconoce.
- **Intubado:** Sí (S) si el paciente fue intubado y ventilado en algún momento de esta enfermedad o No (N) o en blanco si se desconoce..
- **Estuvo en UCI:** Sí (S) si el paciente estuvo en la UCI (unidad de cuidados intensivos) o en la UCC (unidad de cuidados críticos) en algún momento de esta enfermedad o No (N) o en blanco si se desconoce..
- **Necesitó oxígeno suplementario:** Sí (S) si el paciente requirió oxígeno suplementario en cualquier momento durante esta enfermedad o No (N) o en blanco si se desconoce.
- **Extubado:** Sí (S) si el paciente fue extubado con éxito o No (N) o en blanco si se desconoce.
- **Temperatura:** La temperatura del paciente en Celsius en el momento de la imagen.
- **Saturación de pO2:** Porcentaje de la presión parcial de saturación de oxígeno en el momento de la imagen.
- **Recuento de wbc:** Recuento de glóbulos blancos en unidades de $10^3/uL$ en el momento de la imagen.
- **Recuento de neutrófilos:** El recuento de células de neutrófilos en unidades de $10^3/uL$ en el momento de la imagen.
- **Recuento de linfocitos:** Recuento de células linfocitarias en unidades de $10^3/uL$ en el momento de la imagen.
- **Vista:** Posteroanterior (PA), Anteroposterior (AP), AP Supino (APS), o Lateral (L) para los rayos X; Axial o Coronal para las tomografías computarizadas.
- **Modalidad:** Medio por el cual fue capturada la imagen. Puede ser CT, Rayos X, o algo más.
- **Fecha:** Fecha en la que se adquirió la imagen.
- **Ubicación:** Nombre del hospital, ciudad, estado, país.
- **Nombre de archivo:** Nombre con extensión.
- **DOI:** El identificador digital de objetos (DOI) del artículo de investigación.
- **URL:** URL del periódico o sitio web de donde vino la imagen.
- **Licencia:** Licencia de imagen como CC BY-NC-SA. En blanco si se desconoce.
- **Notas clínicas:** Sobre la imagen y/o el paciente.

- **Otras notas:** Por ejemplo, el crédito.

B. Diccionario de datos del dataset de radiografías y tomografías de pacientes sospechosos de COVID-19

- **No.:** Indicador único de las pruebas realizadas o a realizar en pacientes.
- **Edad:** Edad del paciente.
- **Grupo de edad:** El grupo al que pertenece según la edad del paciente.
- **Sexo:** Toma los valores masculino o femenino.
- **Fecha Resultado:** Fecha en que se obtuvo el resultado de la prueba.
- **Resultado COVID-19:** Resultado de la prueba. Los valores que toma son: positivo, negativo, pendiente, no muestra, probable.
- **Provincia de residencia en RD:** Provincia en que reside el paciente.
- **Municipio de residencia en RD:** Municipio en que reside el paciente.
- **Barrio:** Barrio en el que reside el paciente.
- **Manejo:** Si se aisló en el hospital o en su casa.
- **Fecha notificación:** Fecha en que se notificó el resultado de la prueba.
- **Muestra:** Si se obtuvo o no una muestra del paciente.
- **Fecha toma de muestra:** Fecha en que se tomó la muestra al paciente.
- **Laboratorio que confirma COVID-19:** Laboratorio que indica el resultado de la prueba del paciente.
- **Fecha de confirmación:** Fecha en que se confirma el resultado de la prueba.
- **Ingreso en UCI:** Si el paciente ingresó a cuidados intensivos (UCI).
- **Condición (Vivo/Muerto):** Indica si el paciente falleció o no.
- **Fecha defunción:** Fecha en que fallece el paciente, si aplica.
- **Recuperado:** Indica si el paciente da negativo al COVID-19, luego de haber dado positivo.
- **Observaciones:** Notas acerca del paciente.
- **Problema de salud:** Indica si el paciente sufre de hipertensión, diabetes, asma, malignidad, otra cardiovascular, EPOC, IRC, TB, VIH, obesidad, desnutrición, falcemia y/o hábitos tóxicos (fumar o beber).
- **Fecha de egreso:** Fecha en que el paciente sale del hospital, en caso de estar aislado en dicho lugar.

C. Cantidad de valores nulos por columna en el dataset de pruebas de COVID-19 realizadas a pacientes dominicanos al 31 de julio de 2020, el cual cuenta con 264.706 registros

Columna	Cantidad de Nulos
No.	0
Edad	45955
Grupo de edad	6258
Sexo	390
Fecha Resultado	7289
Resultado COVID-19	0
Provincia de residencia en RD	0
Municipio de residencia en RD	27731
Barrio	255984
Manejo	1
Fecha notificación	257752
Muestra	2170
Fecha toma de muestra	13423
Laboratorio que confirma COVID-19	1494
Fecha de confirmación	3517
Ingreso en UCI	228659
Condición (Vivo/Muerto)	0
Fecha defunción	263290
Recuperado	0
Observaciones	258475
Hipertensión	101069
Diabetes	101128
Asma	254316
Malignidad	254288
Otra Cardiovascular	254260
EPOC	254286
IRC	254313
TB	101208
VIH	101207
Obesidad	101201
Desnutrición	101208
Falcemia	101208
Hábitos tóxicos (Fumar/Beber)	254772
Fecha de egreso	261425

D. Código de Parámetros de Data Augmentation para los modelos en Tensorflow

```
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator

INPUT_SHAPE = 224
OUTPUT_SHAPE = len(categories)
BATCH_SIZE = 32

train_image_gen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=20,
    width_shift_range=.2,
    height_shift_range=.2,
    zoom_range=.2,
    fill_mode="nearest"
)

val_image_gen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=20,
    width_shift_range=.2,
    height_shift_range=.2,
    zoom_range=.2,
    fill_mode="nearest"
)

train_data_gen = train_image_gen.flow_from_directory(
    batch_size=BATCH_SIZE,
    directory=train_dir,
    shuffle=True,
    target_size=(INPUT_SHAPE, INPUT_SHAPE)
)

val_data_gen = val_image_gen.flow_from_directory(
    batch_size=BATCH_SIZE,
    directory=val_dir,
    shuffle=True,
    target_size=(INPUT_SHAPE, INPUT_SHAPE)
)
```

E. Código de Definición, Entrenamiento y Resultados de los Modelos Predictivos

```
import matplotlib.pyplot as plt

metrics = [
    tf.keras.metrics.Accuracy(),
    tf.keras.metrics.Precision(),
    tf.keras.metrics.Recall(),
    tf.keras.metrics.AUC(name='auc')
]

base_model_1 = tf.keras.applications.VGG16(
    False,
    input_shape=(INPUT_SHAPE, INPUT_SHAPE, 3))

for layer in base_model_1.layers:
    layer.trainable=False

output_1 = tf.keras.layers.Flatten()(base_model_1.output)
output_1 = tf.keras.layers.Dense(4096, 'relu')(output_1)
output_1 = tf.keras.layers.Dense(4096, 'relu')(output_1)
output_1 = tf.keras.layers.Dense(OUTPUT_SHAPE, 'sigmoid')(output_1)

model_1 = tf.keras.Model(inputs=base_model_1.input, outputs=output_1)
model_1.compile(
    loss=tf.keras.losses.BinaryCrossentropy(),
    optimizer=tf.keras.optimizers.Adam(2e-06),
    metrics=metrics)

base_model_2 = tf.keras.applications.DenseNet169(
    False,
    input_shape=(INPUT_SHAPE, INPUT_SHAPE, 3))

output_2 = tf.keras.layers.Flatten()(base_model_2.output)
output_2 = tf.keras.layers.Dense(1000, 'relu')(output_2)
output_2 = tf.keras.layers.Dense(2, 'sigmoid')(output_2)

base_model_2 = tf.keras.Model(inputs=base_model_2.input, outputs=output_2)

base_model_2.compile(
    loss=tf.keras.losses.BinaryCrossentropy(),
```

```

optimizer=tf.keras.optimizers.Adam(1e-06),
metrics=metrics)

base_model_3 = tf.keras.applications.DenseNet169(
    False,
    input_shape=(INPUT_SHAPE, INPUT_SHAPE, 3))

output_3 = tf.keras.layers.Flatten()(base_model_3.output)
output_3 = tf.keras.layers.Dense(1000, 'relu')(output_3)
output_3 = tf.keras.layers.Dense(3, 'softmax')(output_3)

model_3 = tf.keras.Model(inputs=base_model_3.input, outputs=output_3)

model_3.compile(
    loss=tf.keras.losses.CategoricalCrossentropy(),
    optimizer=tf.keras.optimizers.Adam(2e-05),
    metrics=metrics[0]
)

epochs=100
early = tf.keras.callbacks.EarlyStopping(
    monitor='val_accuracy',
    patience=6)

# Apply to all three models the next blocks of code
history_1 = model_1.fit(
    train_data_gen,
    steps_per_epoch=int(np.ceil(train_data_gen.n / float(BATCH_SIZE))),
    epochs=epochs,
    validation_data=val_data_gen,
    validation_steps=int(np.ceil(val_data_gen.n / float(BATCH_SIZE))),
    callbacks = [early]
)

def plot_metric(result, val_result, epochs_range, metric='', loc=''):
    plt.plot(epochs_range, auc, label=f'Training {metric}')
    plt.plot(epochs_range, val_auc, label=f'Validation {metric}')
    if loc != '':
        plt.legend(loc='lower right')
    plt.title(f'Training and Validation {metric}')
    plt.show()

```

```
for metric in ['Loss', 'Accuracy', 'Precision', 'Recall', 'AUC']:
    plot_metric(result=history_1.history[metric.lower()],
                val_result=history_1.history[f'val_{metric.lower()}'],
                epochs_range=range(early.stopped_epoch+1)
                metric=metric)

test_image_gen = ImageDataGenerator(rescale=1./255)
test_data_gen = test_image_gen.flow_from_directory(
    batch_size=BATCH_SIZE,
    directory=test_dir,
    shuffle=False,
    target_size=(INPUT_SHAPE, INPUT_SHAPE)
)
model.evaluate(test_data_gen, verbose=2)
```

Bibliografía

Diario, L. (2020, 3 agosto). *Se reduce a 70 % la ocupación hospitalaria de pacientes Covid en el Gran Santo Domingo*. listindiario.com.
<https://listindiario.com/la-republica/2020/08/03/629113/se-reduce-a-70--la-ocupacion-hospitalaria-de-pacientes-covid-en-el-gran-santo-domingo>

Corum, J. (2020, 22 julio). *Tratamientos y medicamentos para el coronavirus: monitoreo de efectividad*. <https://www.nytimes.com/#publisher>.
<https://www.nytimes.com/es/interactive/2020/science/coronavirus-tratamientos-curas.html>

X. (2020, 11 agosto). *República Dominicana reporta 293 personas en UCI, 595 nuevos casos de COVID-19 y 18 fallecimientos*. Coronavirus República Dominicana.
<https://coronavirusrd.gob.do/2020/08/11/republica-dominicana-reporta-293-personas-en-uci-595-nuevos-casos-de-covid-19-y-18-fallecimientos/>

Coronavirus Update (Live): 20,521,644 Cases and 745,918 Deaths from COVID-19 Virus Pandemic - Worldometer. (2020). worldometers.
<https://www.worldometers.info/coronavirus/>

Flor, M. F. (2020, 20 abril). *Más del 95 % de los muertos por COVID son ancianos*. PanAm Post.
<https://es.panampost.com/mamela-fiallo/2020/04/20/95-muertos-covid-ancianos/>

Hirald, M. I. (2020, 29 marzo). *Patologías más afectadas por COVID-19*. www.diariolibre.com.
<https://www.diariolibre.com/actualidad/salud/patologias-mas-afectadas-por-covid-19-0117978457>

Asif, S. (2020, 1 enero). *Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Networks*. medRxiv.
<https://www.medrxiv.org/content/10.1101/2020.05.01.20088211v2>

Soft, B. E. (2020, 20 julio). *Blue Eye Soft Announces Commodity Designation and FDA Acknowledgement of BLUEDOCAITM — an AI-driven Medical Imaging Clinical Decision Tool*. Blue Eye Soft. <https://blueeyesoft.com/>
<https://www.globenewswire.com/news-release/2020/07/20/2064630/0/en/Blue-Eye-Soft-Announces-Commodity-Designation-and-FDA-Acknowledgement-of-BLUEDOCAITM-a-n-AI-driven-Medical-Imaging-Clinical-Decision-Tool.html>

Automated detection of COVID-19 cases using deep neural networks with X-ray images. (2020, 1 junio). PubMed Central (PMC).

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7187882/>

Covid-19 Sounds app. Capturado en agosto 27, 2020.

<https://www.covid-19-sounds.org/es/>

Apostolopoulos, I. D. (2020, 3 abril). Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*.

<https://link.springer.com/article/10.1007/s13246-020-00865-4>

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., & Xia, L. (2020). Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology*, 296(2), E32-E40.

<https://pubs.rsna.org/doi/10.1148/radiol.2020200642>

Doris Pantaleón. (2020). Sistema hospitalario a punto del colapso por el Covid-19. *Listin Diario*. Recuperado de:

<https://listindiario.com/la-republica/2020/06/20/622736/sistema-hospitalario-a-punto-del-colapso-por-el-covid-19>

Alexa Olivero. (2020). Colapsa disponibilidad hospitalaria en el Gran Santo Domingo con un 100% de ocupación. julio 27, 2020, de *Diario Libre*. Sitio web:

<https://www.diariolibre.com/actualidad/salud/colapsa-disponibilidad-hospitalaria-en-el-gran-santo-domingo-con-un-100-de-ocupacion-PJ20384830>

Yulissa Álvarez. (2020). Para hacerse una prueba de COVID-19 en RD hay que pasar una odisea. junio 24, 2020, de *Diario Libre*. Sitio web:

<https://www.diariolibre.com/actualidad/salud/para-hacerse-una-prueba-de-covid-19-en-rd-hay-que-pasar-una-odisea-OJ19671906>

Yulissa Álvarez. (2020). Continúan las dificultades para la entrega de resultados de pruebas PCR en el país. julio 10, 2020, de *Diario Libre*. Sitio web:

<https://www.diariolibre.com/actualidad/salud/continuan-las-dificultades-para-la-entrega-de-resultados-de-pruebas-pcr-en-el-pais-PB20034649>

Anónimo. (2020). Hijo de paciente de Hemodiálisis en Hospital Regional Taiwán de Azua denuncia frente a puerta de Unidad de Hemodiálisis pusieron puerta de entrada y salida del personal médico al área del COVID-19. julio 20, 2020, de *Somos Pueblo RD*. Sitio web:

https://www.instagram.com/p/CC3M_CdD8ZN/

Radiologist -, C. F. M. M. L. G. B. A.-. (2020). COVID-19 Image Gallery - Classification of severity. *COVID-19 Image Gallery - Classification of severity* « Previous.

<https://www.radiologymasterclass.co.uk/gallery/covid-19/covid-19-image-gallery-2>

Ankita Shelke, Madhura Inamdar, Vruddhi Shah, Amanshu Tiwari, Aafiya Hussain, Talha Chafekar, Ninad Mehendale. (2019). Chest X-ray classification using Deep learning for automated COVID-19 screening. China: medRxiv.

<https://www.medrxiv.org/content/10.1101/2020.06.21.20136598v1.full.pdf>