

**PONTIFICIA UNIVERSIDAD
CATÓLICA DEL PERÚ**

Escuela de Posgrado



Metodología Basada en Analítica Espacio-temporal y Machine Learning para Evaluar la Calidad del Servicio de Red de Telecomunicaciones y Determinar la Factibilidad de Conectividad para Aplicaciones e-Health

Trabajo de investigación para obtener el grado académico de Maestro en Informática con mención en Ciencias de la Computación que presenta:

John Kevin Barrera Contreras

Asesores:

Edwin Rafael Villanueva Talavera

Miguel Nuñez del Prado Cortez

Lima, 2024

Informe de Similitud

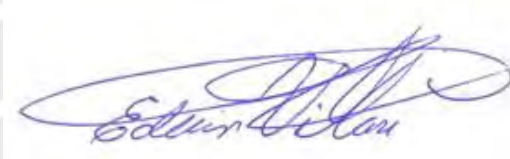
Yo, Edwin Rafael Villanueva Talavera, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis/el trabajo de investigación titulado

Metodología basada en Analítica Espacio-temporal y Machine Learning para evaluar la calidad del servicio de red de telecomunicaciones y determinar la factibilidad de conectividad para aplicaciones e-Health, del autor John Kevin Barrera Contreras, dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 10%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 18/09/2024.
- He revisado con detalle dicho reporte y la Tesis, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:

Lima, 18 de septiembre del 2024.

Apellidos y nombres del asesor / de la asesora: <u>Villanueva Talavera, Edwin Rafael</u>	
DNI: 29714308	Firma 
ORCID: 0000-0002-6540-1230	

Resumen

El presente trabajo propone una metodología que integra el aspecto espacio-temporal de los datos y el machine learning para evaluar la calidad de la infraestructura digital frente al despliegue de aplicaciones e-health en Perú. Haciendo énfasis en la importancia del sistema de salud en el bienestar social, y que fue evidenciada por la creciente relevancia de la telemedicina después de la pandemia de COVID-19.

El estudio tiene como objetivo evaluar la infraestructura digital en el Perú en relación a los servicios de telemedicina, utilizando datos de conectividad del servicio de Internet de Speedtest, densidad poblacional modelada en hexagonos H3, y la lista de servicios de telemedicina. El uso de técnicas de machine learning recrean un mapa de la infraestructura digital del Perú. Luego, estos resultados se extienden a evaluar la calidad en zonas adyacentes a los centros de salud. La metodología contempla la preparación de datos, incluyendo la indexación e imputación de datos faltantes de calidad del servicio (QoS) utilizando el regresor K-Neighbors, y la agrupación de áreas utilizando K-means para categorizar la calidad del servicio. Los hallazgos revelan que más del 90% de la población reside en áreas donde la infraestructura soporta más del 70% de los servicios de telemedicina catalogados.

Los resultados revelan desafíos y oportunidades clave sobre la infraestructura digital de Perú. La identificación de áreas con limitaciones y la categorización de hospitales en diferentes niveles de capacidad ofrecen una visión integral de la situación actual. Estos indicadores pueden orientar políticas públicas y decisiones de salud pública, enfocándose en mejorar la adopción y uso efectivo de las aplicaciones e-health, especialmente en el ámbito de la telemedicina.

Abstract

This work proposes a methodology that integrates the spatio-temporal aspect of data and machine learning to evaluate the quality of digital infrastructure for the deployment of e-health applications in Peru. Emphasizing the importance of the health system in social well-being, which was highlighted by the growing relevance of telemedicine following the COVID-19 pandemic.

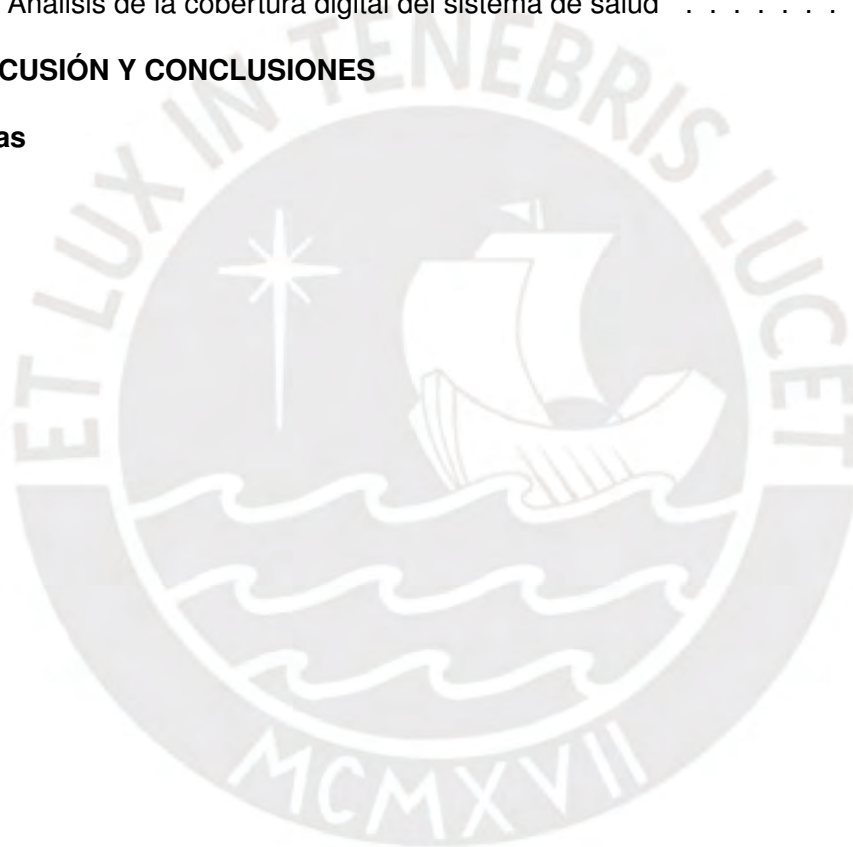
The study aims to evaluate the digital infrastructure in Peru in relation to telemedicine services, using internet connectivity data from Speedtest, population density modeled in H3 hexagons, and the list of telemedicine services. Machine learning techniques are used to recreate a map of Peru's digital infrastructure. These results are then extended to evaluate the quality in areas adjacent to health centers. The methodology includes data preparation, indexing and imputing missing quality of service (QoS) data using the K-Neighbors regressor, and clustering areas using K-means to categorize service quality. The findings reveal that more than 90% of the population resides in areas where the infrastructure supports more than 70% of the cataloged telemedicine services.

The results reveal key challenges and opportunities regarding Peru's digital infrastructure. Identifying areas with limitations and categorizing hospitals into different levels of capacity offer a comprehensive view of the current situation. These indicators can guide public policies and health decisions, focusing on improving the adoption and effective use of e-health applications, especially in the field of telemedicine.

Índice general

Resumen	I
Abstract	II
Índice general	III
Índice de figuras	V
Índice de cuadros	VII
Introducción	1
I GENERALIDADES	3
1.1 Problema de Investigación	3
1.2 Objetivos del Estudio	4
1.2.1 Objetivo General	4
1.2.2 Objetivos Específicos	4
1.3 Justificación	5
II LITERATURA REVISADA	6
2.1 Accesibilidad a los Servicios de Salud	6
2.2 Accesibilidad y Crítica de la Infraestructura en Perú	7
2.3 Impacto de la Infraestructura Digital Durante la Pandemia	9
2.4 Definición y Estado de la Telemedicina en Perú	10
III METODOLOGÍA	13
3.1 Descripción de los datos utilizados	13
3.1.1 Mapas de regiones administrativas del Perú	13
3.1.2 Datos de Población	13
3.1.3 Servicios de telemedicina y requisitos técnicos de conexión	16
3.1.4 Speedtest - Proxy de calidad de infraestructura digital	16
3.2 Esquema metodológico	18
3.3 Etapas	18

3.3.1	Preparación de datos	18
3.3.2	Análisis de cobertura general - Nacional	21
3.3.3	Análisis de la cobertura digital del sistema de salud	22
IV	RESULTADOS	26
4.1	Preparación de datos	26
4.2	Análisis de cobertura general - Nacional	29
4.2.1	Visualización del Análisis de la Calidad del Servicio de e-Health	32
4.3	Análisis de la cobertura digital del sistema de salud	34
V	DISCUSIÓN Y CONCLUSIONES	42
	Referencias	45
	Anexo	48



Índice de figuras

2.1	<i>Mapa global del tiempo de viaje a un centro de salud, sin acceso a transporte motorizado.</i>	7
2.2	<i>El mapa global del tiempo de viaje a un centro de salud, con acceso a transporte motorizado.</i>	7
2.3	<i>Mapa de accesibilidad, a escala de unidades censales.</i>	8
2.4	<i>Asociación entre la reducción de la movilidad y la calidad de la infraestructura digital.</i>	10
2.5	<i>COVID-19: Stringency Index for Peru.</i>	11
3.1	<i>Kontur: Densidad poblacional en el Perú</i>	14
3.2	<i>Establecimientos de Salud por categoría (nivel) de atención</i>	15
3.3	<i>Distribución de Indicadores de Calidad de Infraestructura Digital en Perú.</i>	17
3.4	<i>Esquema metodológico para llevar a cabo el análisis de accesibilidad digital.</i>	18
4.1	<i>Mapas temáticos de indicadores de calidad de red en Perú.</i>	26
4.2	<i>Mapas de Lima y Callao antes de completar los datos ausentes.</i>	27
4.3	<i>Mapas de indicadores de calidad de red en Perú despues de completar los datos ausentes.</i>	28
4.4	<i>Mapas de Lima y Callao despues de completar los datos ausentes.</i>	28
4.5	<i>Scores para Elbow analysis.</i>	30
4.6	<i>Gráfico de siluetas con número de clusters igual a 4.</i>	31
4.7	<i>Proxy de calidad de infraestructura digital.</i>	33
4.8	<i>Gráfico de la capacidad de Infraestructura Digital para Telemedicina en Lima y Callao.</i>	38
4.9	<i>Cantidad de Hospitales por Servicio y Categoría.</i>	40
4.10	<i>Cantidad de Centros de atención primaria por Servicio y Categoría.</i>	41

0.1	<i>Mapas de indicadores de calidad de red en Perú antes de completar los datos ausentes.</i>	48
0.2	<i>Gráfico de la capacidad de Infraestructura Digital para Telemedicina en el Perú. .</i>	49



Índice de cuadros

3.1	Casos de uso de TI de salud y requisitos de infraestructura asociados	23
4.1	Scores para <i>Silhouette analysis</i>	29
4.2	Centroides de los Clústeres	31
4.3	Ranking de centroides basados en el índice de Similitud Coseno	32
4.4	Capacidad de Infraestructura Digital para Telemedicina en el Perú	35
4.5	Capacidad de Infraestructura Digital para Telemedicina en Lima y Callao	36
4.6	Hospitales sin capacidad o infraestructura que soporte servicios de telemedicina	39



Introducción

Los sistemas de salud son uno de los pilares de la sociedad que garantizan el bienestar de la población. Sin embargo, el acceso a la atención médica es un requisito para el bienestar humano que está limitado, en parte, por la asignación de recursos de atención médica en relación con la población humana dispersa geográficamente (Weiss y cols., 2020). Por esta razón, el tiempo de accesibilidad al sistemas de salud ha sido estudiado con el fin de implementar políticas en el sector salud (Tariverdi, Nunez-del Prado, Leonova, y Rentschler, 2023). Esto debido a que las poblaciones que deben realizar largos desplazamientos para llegar al sistema de salud tienen menos probabilidades de buscar asistencia cuando la necesitan (Battle y cols., 2016; Alegana, Wright, Pezzulo, Tatem, y Atkinson, 2017), y las consecuencias de no buscar asistencia incluyen un aumento de la mortalidad y la morbilidad por afecciones tratables (Manongi y cols., 2014).

Durante la pandemia del COVID-19, se ha observado la creciente relevancia de las aplicaciones e-health, especialmente en el ámbito de la telemedicina. También se destaca la complejidad de la adopción de aplicaciones e-health en la lucha contra enfermedades infecciosas. Factores sociodemográficos, socioeconómicos y epidemiológicos influyen en la percepción del riesgo y la adopción de medidas, resaltando la importancia del acceso a herramientas digitales en el seguimiento de pautas de distanciamiento social durante emergencias sanitarias (Gozzi, Comini, y Perra, 2022; Bank, Año de la publicación 2022). Estas herramientas han demostrado ser vitales para mantener la continuidad de la atención médica y reducir el impacto de la crisis sanitaria. Sin embargo, para aprovechar al máximo el potencial de estas tecnologías, es esencial contar con una infraestructura digital sólida y accesible que permita su implementación efectiva (Gozzi y cols., 2022; Bank, Año de la publicación 2022).

La revisión de la literatura nos ha permitido identificar áreas de interés y brechas, proporcionando un contexto para entender cómo la telemedicina puede abordar desafíos como la accesibilidad a la atención médica. A través de la metodología propuesta, que incluye la identificación de zonas con cobertura limitada, la evaluación de la calidad del servicio de Internet y la

medición de la accesibilidad, buscamos generar resultados que informen sobre la capacidad de la infraestructura digital en Lima Metropolitana y su relación con la implementación efectiva de aplicaciones de e-health.

La utilización de estos diversos conjuntos de datos nos permitio realizar un análisis integral de la infraestructura digital, la población, enfermedades y la disponibilidad de servicios de salud en Perú. Esto nos proporciona una base sólida para comprender los desafíos y las oportunidades relacionadas con la adopción de aplicaciones e-health y que servicios pueden ser desplegados tomando en cuenta la calidad de la infraestructura digital en el país.



Capítulo I

GENERALIDADES

1.1. Problema de Investigación

La carencia de una infraestructura digital desarrollada puede restringir la capacidad de implementar políticas y aplicaciones de e-health, especialmente en regiones con déficit de personal médico capacitado (Weiss y cols., 2020). En este escenario, surge la necesidad imperante de proporcionar métricas que evidencien el estado de la infraestructura y comprender qué servicios y aplicaciones de telemedicina pueden ser ofrecidos.

Este planteamiento nos conduce a la siguiente interrogante: ¿En qué medida la infraestructura digital actual en Perú es suficiente y facilita la adopción eficaz de aplicaciones de e-health, abarcando aspectos como el acceso a Internet, la cobertura geográfica y la calidad del servicio, con el objetivo de mejorar la prestación de servicios de salud a nivel nacional?

En este contexto, buscamos alcanzar el siguiente objetivo del estudio: Evaluar de manera integral la infraestructura digital en todo el territorio peruano, focalizándonos en aspectos clave como el acceso a Internet, la cobertura geográfica y la calidad del servicio, con el propósito de identificar oportunidades y desafíos para mejorar la prestación de servicios de salud a nivel nacional. Este objetivo se llevará a cabo mediante una metodología basada en analítica espacio-temporal y machine learning, para evaluar la calidad del servicio de la red de telecomunicaciones y determinar la factibilidad de conectividad para servicios de e-Health en Perú.

1.2. Objetivos del Estudio

1.2.1. Objetivo General

Evaluar de manera integral la infraestructura digital en todo el territorio peruano, focalizándose en aspectos clave como el acceso a Internet, la cobertura geográfica y la calidad del servicio, con el propósito de identificar oportunidades y desafíos para mejorar la prestación de servicios de salud a nivel nacional.

1.2.2. Objetivos Específicos

1. Caracterizar la calidad de la infraestructura digital del Perú utilizando datos de conectividad a internet.
 - Se busca entender las variaciones en la calidad del servicio de internet y cómo estas afectan la implementación de servicios de telemedicina.
 - Identificar las zonas geográficas en todo el Perú que cuentan con cobertura de internet.
2. Identificar un vector ideal de calidad de servicio para diferentes aplicaciones de telemedicina, considerando los requisitos técnicos específicos de cada servicio.
 - Establecer los parámetros técnicos necesarios para garantizar el funcionamiento eficiente de diversas aplicaciones de telemedicina.
3. Clasificar las áreas geográficas en diferentes categorías de calidad de servicio utilizando machine learning.
 - La clasificación de las áreas geográficas según la calidad de infraestructura permitirá identificar áreas que requieren mejoras para optimizar los servicios de telemedicina.

4. Evaluar la capacidad de los centros de salud para proporcionar servicios de telemedicina en función de la calidad de la conectividad a internet.
 - Este objetivo se centra en identificar los centros de salud (hospitales y PHC) que están ubicados en zonas geográficas con cobertura para ofrecer servicios de telemedicina y cuáles necesitan mejoras en su infraestructura digital.

1.3. Justificación

Al abordar estos objetivos, buscamos comprender el estado actual de la calidad de servicio de la infraestructura digital existente en el Perú, y también en Lima Metropolitana y su capacidad para respaldar la adopción y uso efectivo de aplicaciones de salud electrónica. Esto proporcionará información relevante para la toma de decisiones en los sectores de la salud y telecomunicaciones, contribuyendo al fortalecimiento de la infraestructura digital en beneficio de la población.

Capítulo II

LITERATURA REVISADA

La investigación se sitúa en el contexto de trabajos previos que han abordado la adopción de aplicaciones e-health, la accesibilidad y la calidad de la infraestructura digital. Estos estudios han destacado la importancia de una infraestructura sólida y accesible para garantizar la efectividad de las aplicaciones e-health y mejorar la atención médica, y como la telemedicina se destaca como un componente crucial en la evolución de la prestación de servicios de salud. La literatura existente, nos ayudo a identificar áreas de interés y brechas que motivaron el presente estudio.

2.1. Accesibilidad a los Servicios de Salud

En específico, en la esfera de la accesibilidad a los servicios de salud, la literatura destaca la relevancia crítica de los tiempos de acceso como factor determinante en la eficiencia y calidad del sistema de salud a nivel global. Donde, cuantificar el acceso a la atención médica a nivel global a través de la caracterización detallada del tiempo de viaje a las instalaciones sanitarias, nos ayuda comprender las inequidades en el acceso a los servicios de salud (Weiss y cols., 2020).

La heterogeneidad espacial del acceso a la atención médica a nivel global, no solo esta definida por la geografía Figura 2.1, sino que el hecho de contar con acceso a transporte motorizado modifica el mapa drásticamente Figura 2.2, pero es no necesariamente significa que sea accesible para todo el mundo.

Los mapa de tiempo de acceso a un centro de salud fueron elaborados utilizando una metodología implica la generación de indicadores de accesibilidad global utilizando información de tiempos de viaje a ciudades (Nelson y cols., 2019), y los datos utilizados fueron una recopilación

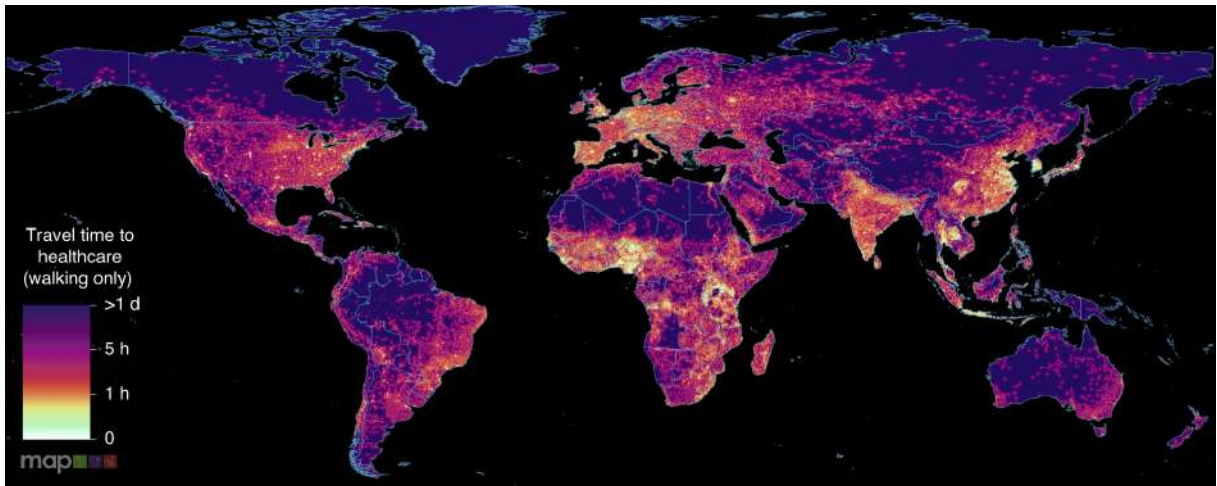


Figura 2.1: *Mapa global del tiempo de viaje a un centro de salud, sin acceso a transporte motorizado.*

Fuente: (Weiss y cols., 2020)

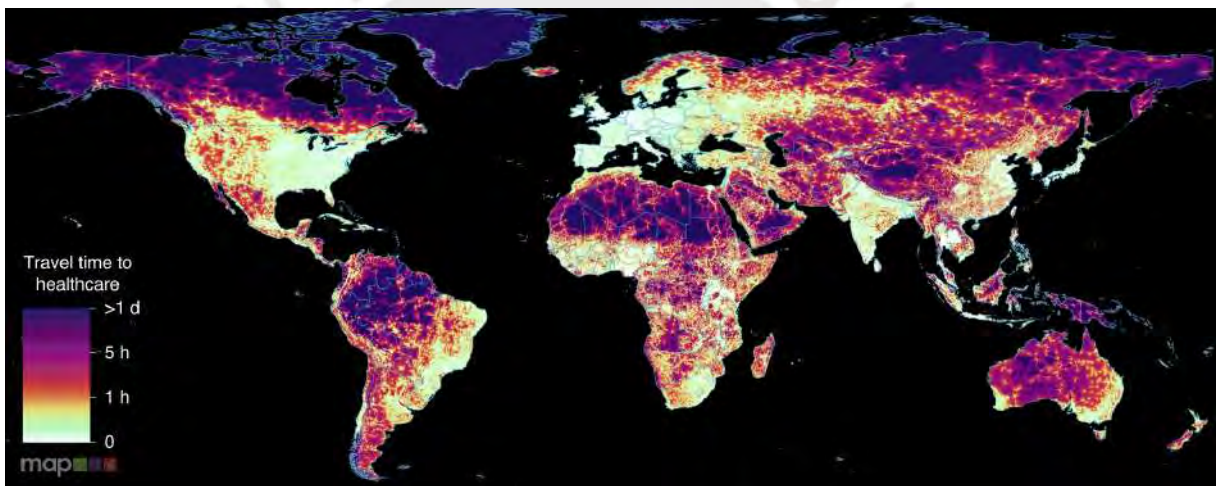


Figura 2.2: *El mapa global del tiempo de viaje a un centro de salud, con acceso a transporte motorizado.*

Fuente: (Weiss y cols., 2020)

ron de ubicaciones de instalaciones sanitarias provenientes de OpenStreetMap, Google Maps y otras fuentes académicas.

2.2. Accesibilidad y Crítica de la Infraestructura en Perú

Acotando la investigación y enfocándonos en la geografía del Perú, otro estudio propone una nueva metodología para evaluar la accesibilidad de los servicios públicos esenciales y la

criticidad de la infraestructura, centrándose en las necesidades de los usuarios. Cuyo enfoque esta aplicado particularmente a las instalaciones de atención médica, pero puede extenderse a otros sectores y servicios públicos.

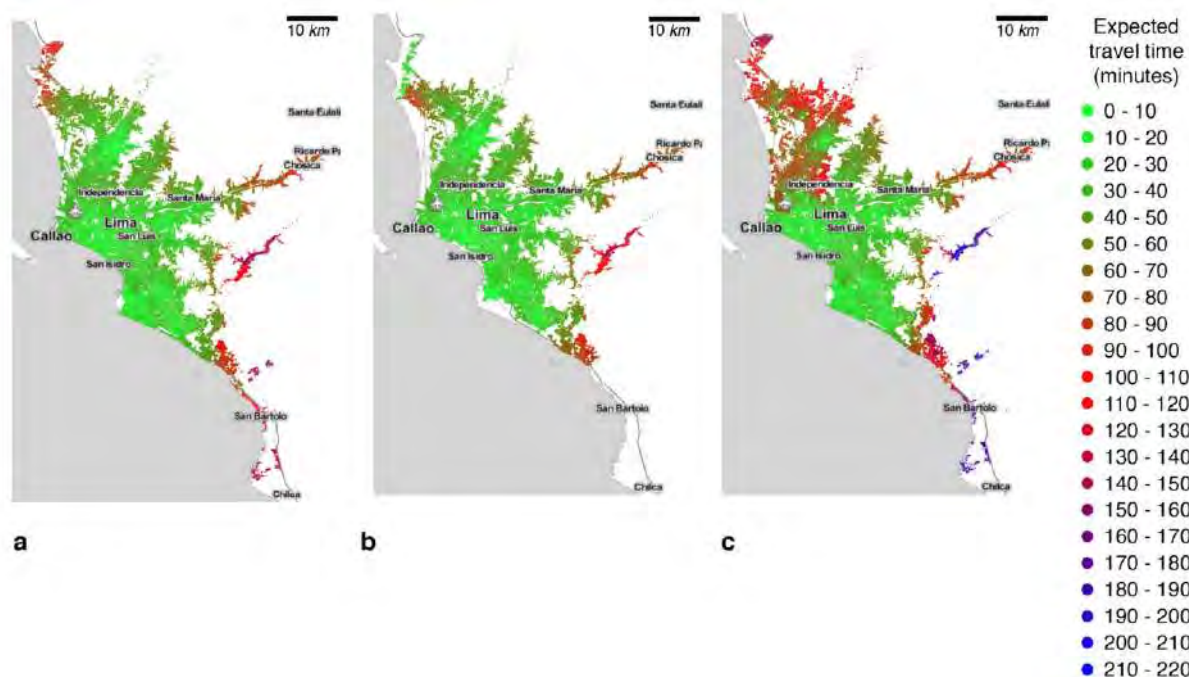


Figura 2.3: Mapa de accesibilidad, a escala de unidades censales.

Para instalaciones de atención médica en la ciudad de Lima. Donde (a) es el escenario sin interrupciones, (b) es el escenario de terremoto con un área de influencia de 500 m y un período de retorno de 50 y 100 años, (c) es el escenario de inundación basado en el mapa de peligros de inundación con un período de retorno histórico.

Fuente: (Tariverdi y cols., 2023)

Se utiliza información de la red vial, densidad de población, mapas de riesgo, ubicaciones de las instalaciones y velocidades de viaje promedio para calcular tiempos de viaje esperados y probabilidades de acceso a las instalaciones. Además, se consideran los efectos de los desastres naturales, como inundaciones y eventos sísmicos, en la infraestructura vial y se cuantifica el impacto en la accesibilidad y criticidad, Figura 2.3. Los resultados proporcionan información valiosa para la planificación de emergencias y la toma de decisiones en la mejora de la resiliencia de los servicios públicos críticos. (Tariverdi y cols., 2023)

Los trabajos vistos presentan los tiempo de acceso a los sistemas de salud presencialmente,

y evidenciamos las deficiencias para poder acceder a este, es en este escenario donde surgen las propuestas de TI para mitigar el problema de acceso al sistema de salud.

2.3. Impacto de la Infraestructura Digital Durante la Pandemia

Un estudio reciente realizado durante la primera ola de la pandemia de COVID-19 en Colombia, Ecuador y El Salvador, se analizaron los determinantes del cumplimiento de intervenciones no farmacéuticas (INF) a nivel municipal, los indicadores socioeconómicos, sociodemográficos y epidemiológicos, y se investigó el papel de la infraestructura digital como posible barrera para la adopción de las INFs. Donde se encontró una correlación significativa entre la calidad de la infraestructura digital y la reducción de la movilidad (Gozzi y cols., 2022), Figura 2.4. Este hallazgo sugiere que los municipios con una mejor conectividad tuvieron mayores tasas de adhesión a las medidas de distanciamiento social. Además, se observó que las reducciones de movilidad fueron más pronunciadas en municipios de mayor tamaño, densidad y riqueza. (Gozzi y cols., 2022; Bank, Año de la publicación 2022)

Para el caso peruano tenemos el *COVID-19 Stringency Index* Figura 2.5 (Mathieu y cols., 2020) , es una medida compuesta de nueve métricas de respuesta (cierre de escuelas; cierre de lugares de trabajo; cancelación de eventos públicos; restricciones en reuniones públicas; cierre de transporte público; requisitos de quedarse en casa; campañas de información pública; restricciones en movimientos internos; y controles de viaje internacional) que se calcula como el promedio de las puntuaciones de las nueve métricas, cada una con un valor entre 0 y 100, donde 100 es la respuesta más estricta (Mathieu y cols., 2020; Hale y cols., 2021).

Otra metodología propuesta por *Data for Good at Meta* (for Good at Meta, 2022) nos proporciona información de la reducción de movilidad durante la pandemia del COVID-19 mediante el *Movement Range Maps* para Perú.

Ambos mapas toman como línea base fechas anteriores a la pandemia del COVID-19, y cuan-

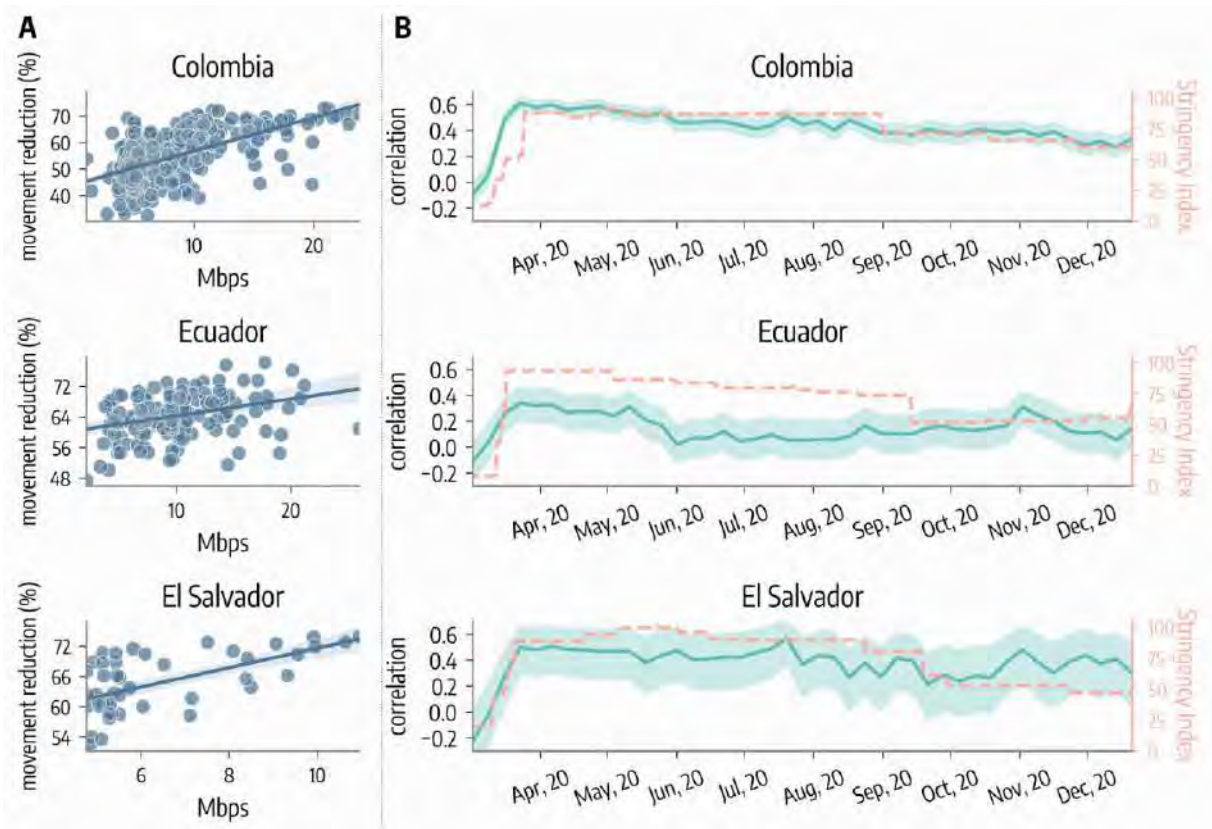


Figura 2.4: Asociación entre la reducción de la movilidad y la calidad de la infraestructura digital.

Para diferentes municipios de Colombia, Ecuador y El Salvador. (A) Reducción de la movilidad en función de la velocidad promedio de descarga en diferentes municipios. (B) Coeficiente de correlación de Pearson (mediana e intervalo de confianza del 95 %) entre la reducción promedio de movimiento semanal y la velocidad promedio de descarga en diferentes municipios.

Fuente: (Gozzi y cols., 2022)

tifican la reducción de la movilidad. Sin embargo a pesar de la políticas para mitigar el impacto del COVID-19, el sistema de salud peruano se vio colapsado, y las aplicaciones de telemedicina no cumplieron las expectativas para minimizar el daño, además el concepto y aplicaciones de telemedicina no tenían una definición robusta (Cuba García, 2021).

2.4. Definición y Estado de la Telemedicina en Perú

Entender que servicios y aplicaciones comprende la telemedicina, nos llevo a revisar la literatura sobre el tema y hacer una consulta a la paginas oficiales del pais para entender en estado y madurez de las aplicaciones de telemedicina.

COVID-19: Stringency Index, Peru



The stringency index is a composite measure based on nine response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest).



Data source: Hale, T., Angrist, N., Goldszmidt, R. et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat Hum Behav* 5, 529–538 (2021). <https://doi.org/10.1038/s41562-021-01079-8>
CC BY

Figura 2.5: COVID-19: Stringency Index for Peru.
Fuente: (Mathieu y cols., 2020)

La *telemedicina* es el uso de tecnologías de telecomunicaciones para respaldar la prestación de todo tipo de servicios médicos, diagnósticos y relacionados con el tratamiento, generalmente por médicos. En contraste, *telehealth* o *teleasistencia* son similares a la *telemedicina*, pero incluyen una variedad más amplia de servicios de atención médica remota más allá de la relación médico-paciente. Además, tenemos el término *telecare* o *teleasistencia*, que se refiere a la tecnología que permite a los consumidores permanecer seguros e independientes en sus propios hogares (*Telehealth, Telemedicine, and Telecare: What's What*, s.f.).

Pero los términos *telehealth* y *telemedicine* a menudo son utilizados de manera indistinguible por el personal de salud. (Balestra, 2018)

En el contexto de la telemedicina es importante y esencial considerar algunos aspectos técnicos. Algunas aplicaciones implican desde enviar un mensaje de texto hasta la telecirugía

y telepresencia (Steele y Lo, 2013), por eso rescatamos algunos servicios y aplicaciones. Tabla 3.1, el ancho de banda recomendado para subida y descarga (Federal Communications Commission, 2010; Steele y Lo, 2013), y la latencia por tipo de servicio (Federal Communications Commission, 2010).

En Perú, el Ministerio de Salud (MINSA) opera la Dirección de Telemedicina, una entidad dedicada a proporcionar servicios de teleorientación, telemonitoreo y teleconsulta, disponible las 24 horas del día, los 7 días de la semana (*Dirección de Telemedicina del Ministerio de Salud (MINSA)*, s.f.). Esta dirección destaca por ser un componente específico enfocado en la implementación y gestión de servicios de telemedicina en el país.



Capítulo III

METODOLOGÍA

En esta sección, presentamos una descripción detallada de los datos utilizados en el presente trabajo de investigación y el proceso metodológico.

3.1. Descripción de los datos utilizados

Los datos se dividen en diferentes categorías para abordar aspectos geográficos, calidad de la infraestructura digital, población, casos epidemiológicos, y la ubicación de hospitales. A continuación, se detalla cada una de estas categorías:

3.1.1. Mapas de regiones administrativas del Perú

Se utiliza un mapa de la división geopolítica del Perú y sus subdivisiones (Instituto Nacional de Estadística e Informática, 2022). Este mapa proporciona información sobre las divisiones administrativas del país, como regiones, provincias y distritos, que son relevantes para el análisis geográfico.

3.1.2. Datos de Población

La información sobre la población de Perú se obtiene a partir de los datos de Kontur Figura 3.1, donde tenemos 259,609 registros, que buscan proporcionar una representación más precisa de la distribución de la población al integrar diversas fuentes y aplicar ajustes para reflejar de manera coherente la presencia humana en áreas geográficas específicas. Los datos de Kontur sobre la población se basan en el índice espacial jerárquico hexagonal (H3) de 400 metros y utilizan una combinación de la Capa Global de Asentamientos Humanos (*GHSL*) y los datos de población de la Capa de Asentamientos de Alta Resolución de Facebook (*HRSL*),

corrigiendo posibles artefactos con información de OpenStreetMap (Kontur, Año de la última actualización 2023).

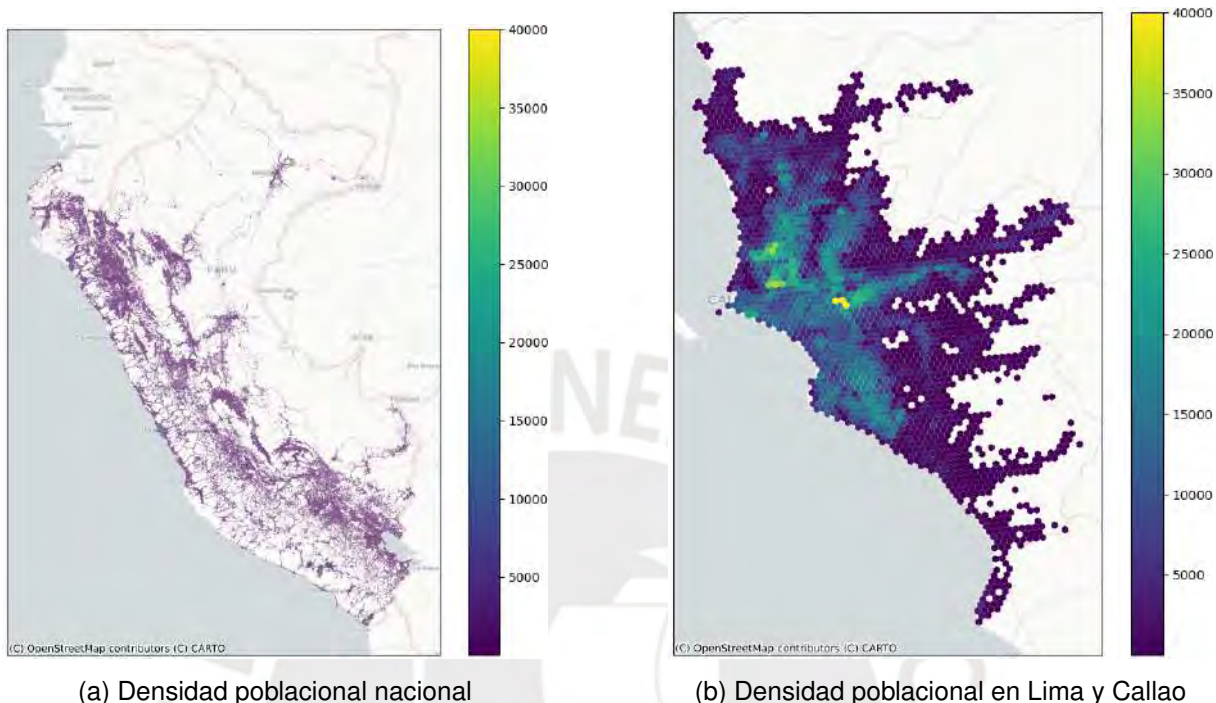


Figura 3.1: Kontur: Densidad poblacional en el Perú
Fuente: Elaboración propia

Hospitales

Los datos sobre la ubicación de los hospitales se obtienen de fuentes oficiales del Ministerio de Salud (MINSA) disponibles en Open Data Perú y RENIPRESS (Ministerio de Salud, 2021; Superintendencia Nacional de Salud (SUSALUD), 2023). Estos datos proporcionan información sobre la ubicación y distribución de los hospitales en diferentes áreas geográficas.

El Nivel de Atención en el sistema de salud se organiza según la relación entre las necesidades de salud de la población y la capacidad de resolución. Hay tres niveles (Ministerio de Salud, 2011):

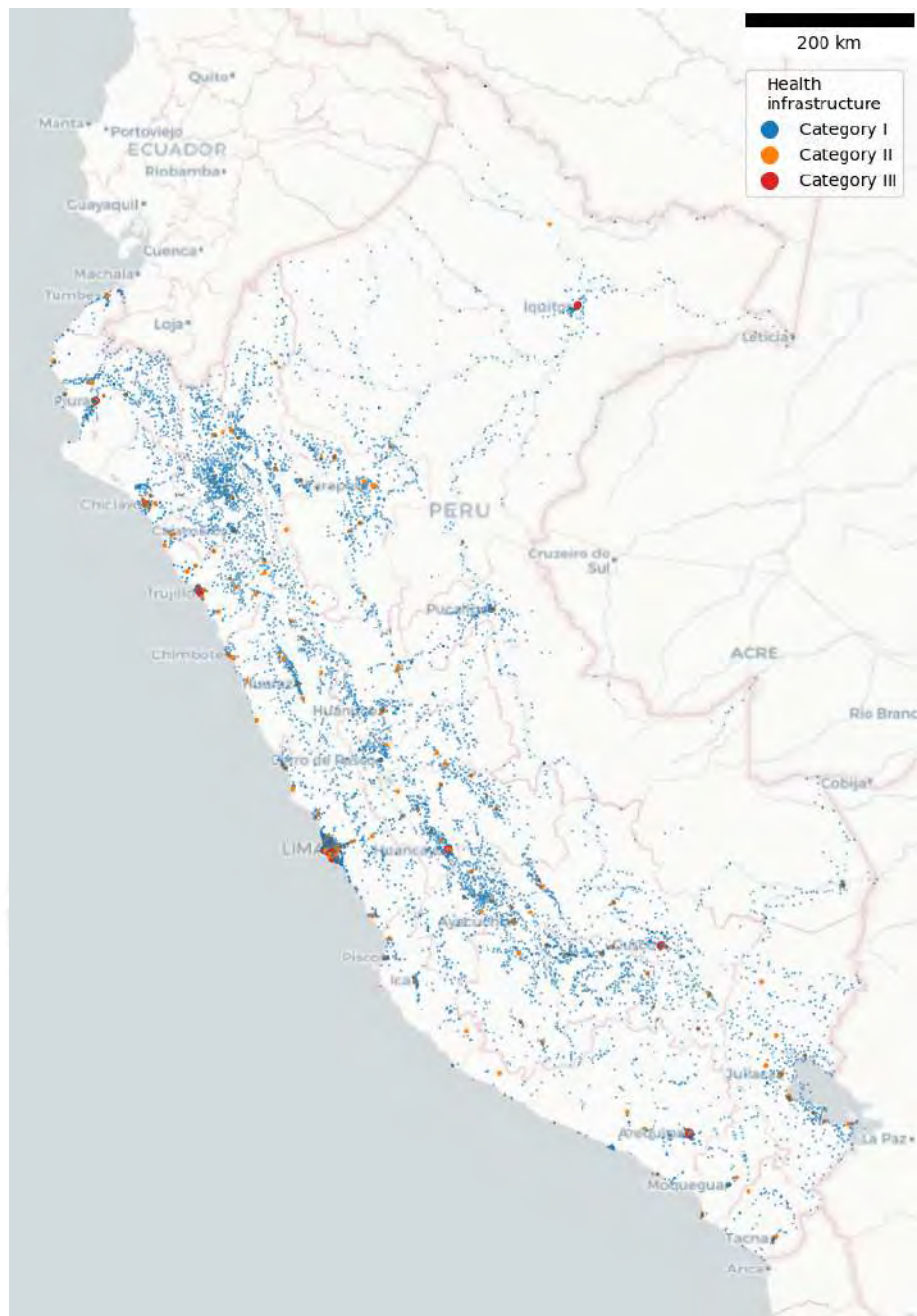


Figura 3.2: Establecimientos de Salud por categoría (nivel) de atención

Fuente: Elaboración propia

1. **Primer Nivel de Atención:** Es la entrada principal al sistema de salud y se centra en actividades como la promoción de la salud, prevención de riesgos, diagnóstico temprano y tratamiento oportuno. Aborda las necesidades de salud más comunes con una oferta de atención de gran tamaño y baja complejidad. Actúa como coordinador del flujo de usuarios

en el sistema.

2. **Segundo Nivel de Atención:** Complementa la atención integral del primer nivel, agregando mayor especialización en recursos humanos y tecnológicos. Ofrece servicios de salud dirigidos a resolver problemas de personas referidas del primer nivel o que acuden por urgencias o emergencias. También realiza actividades preventivas y promocionales.
3. **Tercer Nivel de Atención:** Es el nivel más especializado y con mayor capacidad resolutive en recursos humanos y tecnológicos. Se enfoca en resolver necesidades de salud referidas de niveles anteriores y atiende casos de urgencia o emergencia. También lleva a cabo actividades preventivas y promocionales según sea necesario.

3.1.3. Servicios de telemedicina y requisitos técnicos de conexión

El informe de la Comisión Federal de Comunicaciones (FCC) proporciona una lista detallada de servicios de salud y sus respectivos requisitos de ancho de banda y latencia (Federal Communications Commission, 2010). Por otro lado, la Telehealth Technology Taxonomy del CITL proporciona una referencia para los requisitos mínimos de ancho de banda de diversas funcionalidades implementables en una aplicación de telemedicina (Steele y Lo, 2013). Más adelante, la metodología nos permitirá contrastar y retroalimentar los aspectos técnicos, los datos referentes al ancho de banda y latencia, ya que son vitales para este trabajo Tabla 3.1.

3.1.4. Speedtest - Proxy de calidad de infraestructura digital

Buscar una fuente de datos que proporcione información sobre la calidad de la infraestructura digital en Perú resultó ser un desafío. Por esta razón, optamos por utilizar los resultados de *Speedtest*[®] proporcionados por *Ookla*[®] (Ookla, 2023), que además, pueden ser utilizados como un indicador proxy de la calidad de Internet (Gozzi y cols., 2022).

Speedtest es una herramienta popular que permite a los usuarios medir la velocidad de su conexión a Internet (Feamster y Livingood, 2020). La herramienta evalúa varias métricas clave

que describen el rendimiento de la conexión, proporcionando una visión general del estado de la red.

La solución *Speedtest Intelligence*[®] de *Ookla*[®] ofrece análisis de métricas de rendimiento del Internet para redes fijas y móviles a nivel nacional:

- **iOS:** Dispositivos Apple como iPhone y iPad.
- **Android:** Dispositivos que ejecutan el sistema operativo Android.
- **Desktop:** Computadoras de escritorio u ordenadores portátiles.

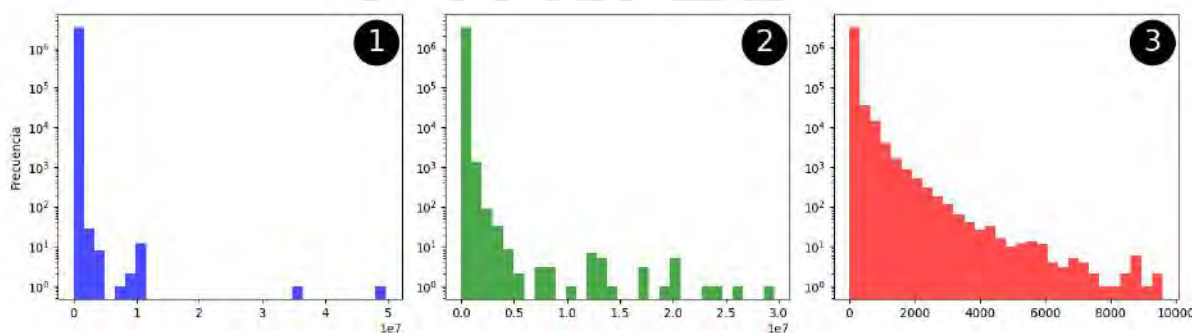


Figura 3.3: *Distribución de Indicadores de Calidad de Infraestructura Digital en Perú.*

(1) Histograma de Velocidad de Descarga (kbps): Revela una distribución zipfiana, resaltando disparidades notables en la calidad de la infraestructura digital. (2) Histograma de Velocidad de Carga (kbps): Muestra variaciones significativas en capacidades de carga entre ubicaciones, también siguiendo una distribución zipfiana. (3) Histograma de Latencia (ms): Presenta la variabilidad en los tiempos de respuesta, indicando que algunas regiones experimentan latencias más bajas que otras.

Fuente: Elaboración propia

Además, para el trabajo realizado fue necesario estandarizar los parámetros de medición. La observación de que los gráficos Figura 3.3 muestran una distribución zipfiana, este patron destaca la heterogeneidad en la calidad de la infraestructura de Internet en Perú, y nos da a entender cuan crucial pueden ser las estrategias de mejora y equidad en la conectividad digital.

- **download (kbps):** Velocidad de descarga en kilobits por segundo.
- **upload (kbps):** Velocidad de carga en kilobits por segundo.
- **latency (ms):** Latencia en milisegundos.

3.2. Esquema metodológico

A continuación se presenta el esquema metodológico del presente trabajo:

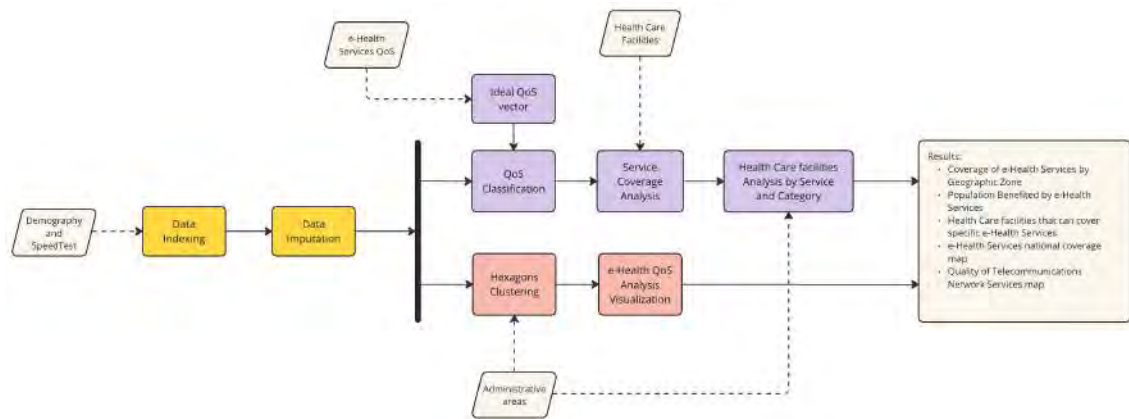


Figura 3.4: Esquema metodológico para llevar a cabo el análisis de accesibilidad digital.

Fuente: Elaboración propia

3.3. Etapas

3.3.1. Preparación de datos

Indexación de datos

En la etapa inicial, nuestro objetivo es organizar y extraer un indicador (promedio) las métricas de los registros de datos de Speedtest, como un *proxy* de calidad del servicio de internet (QoS) dentro de áreas geográficas específicas para obtener una visión clara y comprensible de la infraestructura digital. Dado que los puntos de datos de internet se recogen de manera casi aleatoria y en diferentes momentos del día, es necesario delimitar estos puntos a áreas geográficas bien definidas. Para esto, utilizamos los datos poblacionales aproximados en hexágonos H3 de *Kontur Data*, que nos permiten agrupar los datos de manera coherente con la densidad poblacional y la cobertura geográfica.

El proceso consiste en:

- **Asignación geográfica**

Cada registro geolocalizado de velocidad de internet (tanto de descarga como de carga) y latencia se asigna al hexágono H3 correspondiente. Esto garantiza que cada punto de datos se ubique correctamente dentro del área poblacional correspondiente, reflejando así la calidad del servicio en esa región específica.

- **Cálculo de indicadores**

Se calculan los promedios de las velocidades de descarga, velocidad de carga y latencia para cada hexágono H3. Este paso es esencial porque los datos de velocidad de internet pueden variar significativamente dependiendo de la hora del día y otros factores temporales, en un análisis más profundo se puede analizar en ventanas de tiempo. Al promediar estos valores, nos ayuda a superar la dispersión y temporalidad de los datos de internet, proporcionando una estructura coherente y representación más estable y confiable de la calidad del servicio en cada área.

- **Integración de métricas y densidad poblacional**

Cada hexágono H3 no solo contiene las métricas de QoS (calidad del servicio) promediadas, sino también la densidad poblacional. Esta integración nos permite correlacionar la calidad del servicio de internet con la densidad de la población, proporcionando una visión más detallada y completa del estado de la infraestructura digital en diversas áreas.

Imputación de datos

En la etapa, enfrentamos el desafío de que los datos de calidad del servicio (QoS) no están presentes en todas las áreas pobladas. Esto significa que no siempre contamos con información completa sobre la velocidad de descarga, velocidad de carga y latencia en cada región poblacional. Para abordar esta falta de datos, necesitamos una estrategia efectiva de imputación que nos permita estimar y completar estos valores faltantes.

- **Desafío:** Los datos de QoS no están presentes en todas las áreas pobladas, por lo que

necesitamos una estrategia para imputar los valores faltantes.

Para resolver este desafío optamos por imputar los valores faltantes mediante técnicas de aprendizaje automático. Este enfoque se basa en la idea de que los valores de QoS en una región pueden ser estimados a partir de los valores en regiones cercanas y similares. Para abordar las brechas de datos, utilizamos el regresor de vecinos más cercanos (**K-Neighbors Regressor**) con distancia euclidiana. Esto permite al modelo estimar la calidad del internet en todas las áreas pobladas.

El proceso consiste en:

- **Entrenamiento del modelo de imputación**

Primero, tomamos los datos existentes de las áreas donde tenemos información completa sobre QoS. Estos datos incluyen las coordenadas geográficas (x, y) y las métricas de QoS (velocidad de descarga, velocidad de carga y latencia). Utilizamos estos datos para entrenar un modelo de imputación adecuado.

- **Predicción de valores faltantes**

Luego, aplicamos el modelo entrenado para predecir los valores de QoS en las áreas donde esta información falta. El modelo utiliza las coordenadas geográficas para encontrar las áreas más cercanas y similares, y a partir de estos vecinos, estima los valores faltantes de QoS.

- **Asignación de datos imputados**

Finalmente, asignamos los valores de QoS predicho por el modelo a las áreas correspondientes. Esto nos permite tener un conjunto de datos más completo y coherente para el análisis posterior.

3.3.2. Análisis de cobertura general - Nacional

Agrupación de Hexágonos

En la etapa de *Hexagons Clustering*, utilizamos los resultados de la imputación de datos para clasificar las áreas geográficas en diferentes categorías de calidad de servicio (QoS) mediante el algoritmo K-means. Este enfoque nos permite generar un mapa detallado que muestra la distribución de la calidad de la infraestructura digital en todo el país.

- Primero, tomamos los datos imputados y ajustamos las geometrías para que coincidan con las áreas administrativas correspondientes, donde las áreas administrativas están definidas por el país estudiado.
- Luego, aplicamos el algoritmo K-means para agrupar los hexágonos en diferentes categorías basadas en la calidad del servicio de internet. Esta clasificación se realiza utilizando los valores de velocidad de descarga, velocidad de carga y latencia.
- Finalmente, agregamos y etiquetamos los resultados para tener una distribución más clara de la calidad de la infraestructura digital, lo que nos permite identificar áreas que requieren mejoras y planificar intervenciones más efectivas para optimizar los servicios de telemedicina en todo el país.

Visualización del Análisis de la Calidad del Servicio de e-Health

El objetivo en esta etapa es identificar y destacar las áreas con mejor y peor infraestructura digital. Para ello, evaluamos y comparamos la calidad de cada centroide, empleamos la distancia coseno respecto al **vector ideal (I)**. (Ecuación 3.1) Además, realizamos agregaciones y utilizamos tablas, mapas y gráficos de los resultados anteriores, enfocándonos especialmente en las ciudades más importantes. En el caso de Perú, nos centramos en Lima, ya que representa un tercio de la población nacional, sin descuidar el resto del país

3.3.3. Análisis de la cobertura digital del sistema de salud

Construcción de vector ideal de calidad de servicio

En esta etapa, nos enfocamos en establecer los requisitos mínimos de calidad de servicio (QoS) para una variedad de aplicaciones de telemedicina. La construcción de estos vectores ideales implica un análisis exhaustivo de la bibliografía existente, sobre aplicaciones de telemedicinas y requerimiento de red (Federal Communications Commission, 2010; Steele y Lo, 2013); cotejando y organizando la información para crear una tabla que detalle los requisitos específicos de velocidad de descarga, velocidad de carga y latencia para cada tipo de servicio.

Con la información recopilada, construimos la Tabla 3.1 que resume los requisitos técnicos específicos para cada aplicación de telemedicina. La tabla incluye:

- Velocidad de Carga (kbits/s): Indica la velocidad mínima de datos necesaria para que el servicio funcione correctamente. Este valor es bidireccional, lo que significa que la velocidad de descarga y carga son las mismas.
- Latencia Máxima (ms): Define el tiempo máximo de retraso aceptable en la transmisión de datos.

Esta tabla proporciona una visión clara de los requisitos técnicos necesarios para cada servicio de telemedicina, permitiendo una planificación más precisa y efectiva de la infraestructura digital. Una vez construida la tabla con los requisitos mínimos para cada servicio, el siguiente paso es generar un vector ideal general para la calidad del servicio (QoS).

Este **vector ideal** se derivará de los valores máximos requeridos para la velocidad de descarga y carga, y los valores mínimos, para la latencia.

- Velocidad de Descarga y Carga: Para estos dos parámetros, buscamos el ancho de banda mínimo necesario para cubrir el servicio más demandante. Con esto nos aseguramos que el vector ideal represente las condiciones más exigentes y, por lo tanto, pueda servir

Complejidad de Comunicaciones	Ancho de banda recomendado (kbits/s)	Latencia Máxima (ms)
Telepresencia	1000000	20
Telecirugía	1000000	20
Aplicaciones de Telemedicina en Tiempo Real basadas en Streaming de Video de Alta Tasa de Bits	1000000	20
Sistemas de Referencia Médica para Médicos	1000	20
Actualización en el Sitio del Historial Electrónico de Salud	100000	20
Aplicaciones de Telemedicina basadas en Imágenes de Alta Resolución (Teleradiología y Teledermatología)	100000	20
Aplicaciones de Telemedicina en Tiempo Real basadas en Streaming de Video de Baja Resolución	10000	50
Monitoreo de Salud Móvil	100000	50
Monitoreo de Salud en el Hogar (en Tiempo Real)	100000	50
Gestión de Enfermedades Crónicas	1540	50
Monitoreo de Salud en el Hogar (almacenamiento y reenvío)	10000	50
Historial Electrónico de Salud o Historial de Salud Personal, (incluyendo imágenes y videos de alta resolución)	100000	50
Educación Médica Continua (EMC) a través de Dispositivos Móviles	10000	50
Educación Médica Continua (EMC)	1540	50
Notificación de Emergencia basada en Smartphone	1540	120
Historial Electrónico de Salud o Historial de Salud Personal basado en Texto	364	120
Aplicaciones de Telemedicina de Almacenamiento y Reenvío (Telepsiquiatría, Teleprescripción, ...)	1540	120
Comunicación e Informes a través de Correos Electrónicos (Incluyendo correos electrónicos con calidad médica-alta)	364	120
Notificación de Emergencia basada en Texto, Teléfono o SMS	364	120
Comunicación de Salud a través de Medios Sociales Relacionados con la Salud	364	120

Cuadro 3.1: Casos de uso de TI de salud y requisitos de infraestructura asociados
Fuente: Elaboración propia

como una guía robusta para la implementación de infraestructuras que soporten a los otros servicios.

- Latencia: En cuanto a la latencia, un valor cercano a cero es el escenario ideal. Se busca que la latencia sea lo más bajo posible para garantizar una comunicación eficiente y sin interrupciones o retrasos.

El **vector ideal** I se define como:

$$I = \{1000000, 1000000, 0,000001\} \quad (3.1)$$

Nota: Este vector ideal es específico para el presente estudio. En caso de actualizaciones de parámetros, métricas o servicios, será necesario actualizar la tabla y el vector ideal correspondiente.

Clasificación de la Calidad del Servicio

En la etapa, se procederá a evaluar la calidad de los servicios proporcionados. Para ello, utilizaremos la tabla de servicios junto con sus valores correspondientes. Paralelamente, analizaremos los resultados obtenidos de la imputación de datos. Con esta información, calificaremos cada hexágono en función de los servicios que pueden cubrir, considerando las características específicas de la calidad del internet.

Análisis de Cobertura de Servicios

En esta etapa, tomamos el mapa de cobertura, resultado del proceso de imputación de datos, y analizamos en qué áreas se encuentran ubicados los centros de salud. Este análisis nos permite identificar las zonas de influencia de cada centro de salud y evaluar qué servicios pueden ofrecer, basándonos en la calidad de la conectividad a internet y otros factores relevantes.

La superposición del mapa de cobertura con la localización de los centros de salud nos permite determinar qué servicios de telemedicina son factibles en cada ubicación. Por ejemplo, en áreas con excelente cobertura de internet, los centros de salud pueden ofrecer una gama más amplia de servicios, como consultas virtuales en tiempo real y monitoreo remoto de pacientes. En contraste, las zonas con cobertura limitada pueden estar restringidas a funciones básicas como teleconsultas asincrónicas.

Este análisis nos permite conocer qué centro de salud puede operar de manera efectiva

dentro de las limitaciones de su infraestructura tecnológica.

Análisis de Establecimientos de Salud por Servicio y Categoría

En esta etapa, utilizamos los resultados previos para generar mapas, tablas y gráficos, y otros recursos visuales que nos permitirán analizar el acceso potencial de la población a los servicios de salud electrónica (e-health) y evaluar la calidad del servicio (QoS) en cada zona geográfica poblada. Mediante este análisis, también podemos determinar la capacidad de los centros de salud para ofrecer servicios de e-health, identificar áreas que necesitan mejoras y optimizar la distribución de recursos.



Capítulo IV

RESULTADOS

4.1. Preparación de datos

Indexación de datos

Inicialmente contamos con los datos de los puntos de registro de Speedtest ploteados sobre el mapa nacional. Dado que tenemos tres características distintas, la velocidad de descarga/subida (expresada en megabites por segundo) y la latencia (en milisegundos), hemos generado tres mapas correspondientes, además los datos Geoespaciales fueron retroalimentado con los mapas hexagonales de la población.(Figura 4.1)

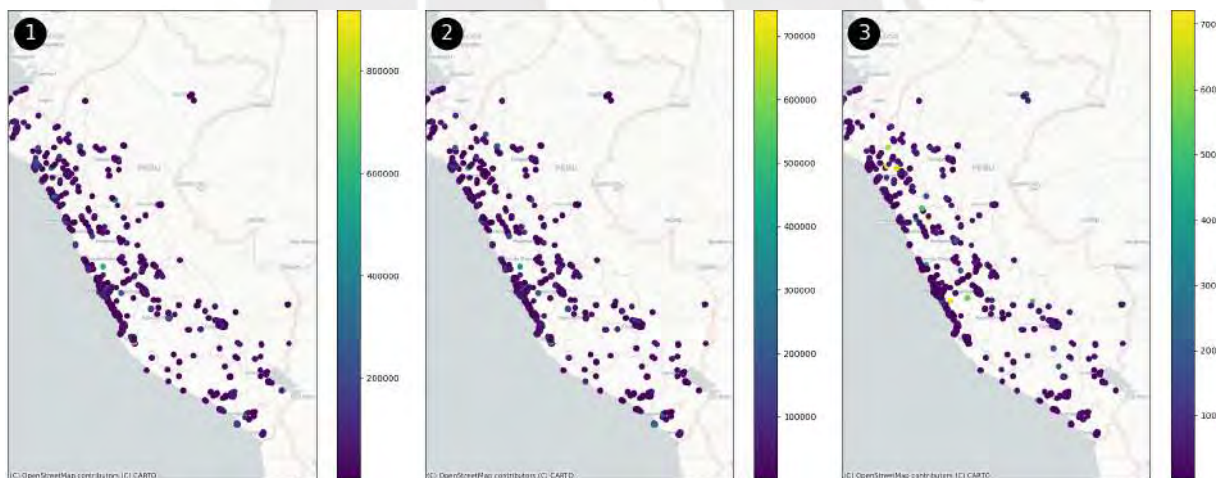


Figura 4.1: Mapas temáticos de indicadores de calidad de red en Perú.

(1) muestra la velocidad de descarga (kbps) en diferentes ubicaciones. (2) visualiza la velocidad de carga (kbps),(3) muestra la latencia.

Fuente: Elaboración propia

A los mapas de puntos iniciales o datos iniciales, procedemos a asignarles un hexágono H3 poblacional. Luego, calculamos el promedio de las métricas en cada hexágono, minimizando el impacto y la variabilidad temporal de los datos. (Figura 0.1 - Anexo V)

La visualización de resultados a nivel nacional es compleja debido a la amplitud del territorio. Por este motivo, se ha optado por una aproximación más detallada, centrándonos en Lima y Callao. (Figura 4.2)

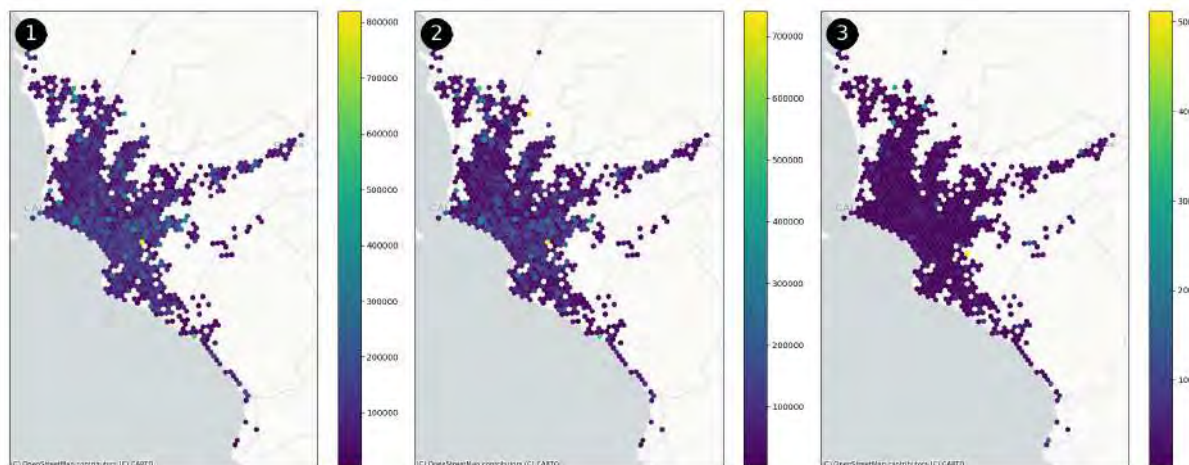


Figura 4.2: Mapas de Lima y Callao antes de completar los datos ausentes.

(1) velocidad de descarga (kbps), (2) velocidad de carga (kbps), y (3) latencia; resultados después de completar las lagunas de datos.

Fuente: Elaboración propia

Los resultados en esta sección proporcionan una visión detallada de la accesibilidad a Internet en las áreas poblacionales consideradas. Tras calcular la media de los parámetros de medición, incluyendo download (kbps), upload (kbps), y latency (ms) para cada hexágono poblacional, se ha generado un mapa de calor que actúa como un proxy representativo de la calidad de la conexión a Internet en dichas áreas.

Imputación de datos

En los resultados de la sección anterior, podemos observar algunas lagunas de datos. Por ello, es necesario realizar una imputación como se indica en la metodología y tomar en cuenta todas las áreas pobladas. Para ello, empleamos el algoritmo K-Nearest Neighbors (KNN), con el cual completamos los hexágonos con datos faltantes de manera exitosa (Figura 4.3), ofreciendo una cobertura más completa y precisa. Gracias a este proceso, logramos imputar exitosamente 257,140 valores a partir de los 2,469 valores iniciales a nivel nacional.

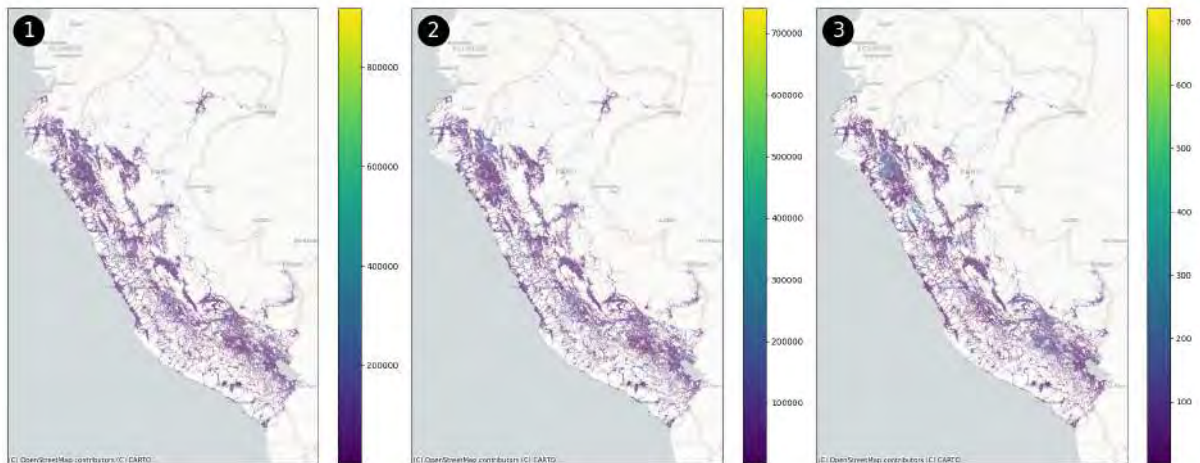


Figura 4.3: Mapas de indicadores de calidad de red en Perú después de completar los datos ausentes.

(1) velocidad de descarga (kbps), (2) velocidad de carga (kbps), y (3) latencia; resultados después de completar las lagunas de datos.

Fuente: Elaboración propia

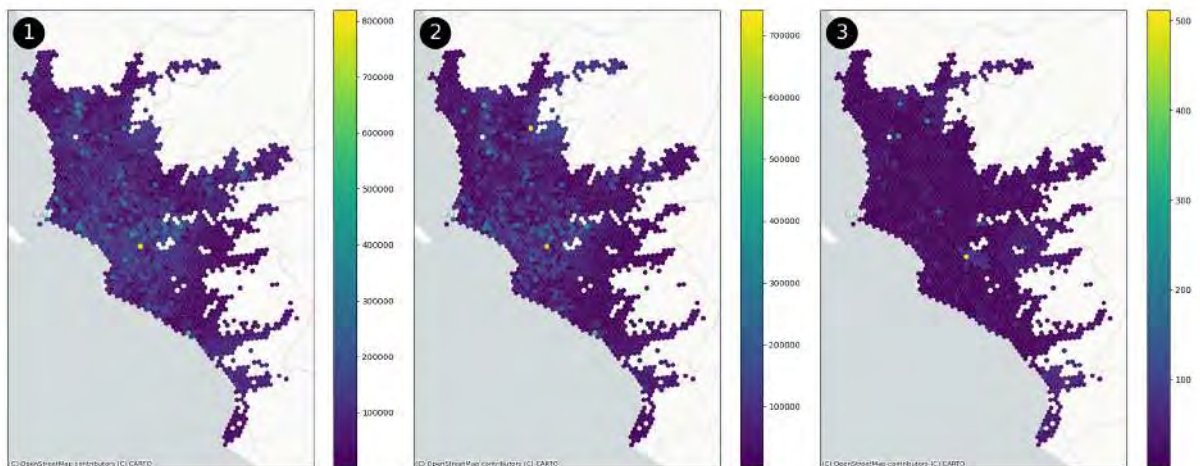


Figura 4.4: Mapas de Lima y Callao después de completar los datos ausentes.

(1) velocidad de descarga (kbps), (2) velocidad de carga (kbps), y (3) latencia; resultados después de completar las lagunas de datos.

Fuente: Elaboración propia

Para mejorar la visualización de resultados hicimos una aproximación más detallada en la región Lima y Callao. (Figura 4.4) Esta focalización ayuda a tener una apreciación más precisa de las variaciones en la accesibilidad a Internet, brindando una perspectiva detallada y visualmente accesible sobre cómo la aplicación de KNN ha mejorado la representación de los datos de

nuestros datasets.

Estos resultados constituyen la base sólida sobre la cual avanzaremos en las dos siguientes fases del estudio, permitiéndonos entender la variabilidad en la accesibilidad a Internet en función de las características poblacionales y geoespaciales.

4.2. Análisis de cobertura general - Nacional

Agrupación de Hexágonos

En esta sección, presentamos los resultados derivados de la fase de agrupación utilizando el algoritmo K-Means para generar categorías representativas de la infraestructura de Internet en áreas poblacionales. Tras completar la categorización, se obtuvieron resultados significativos, comparando 2 métodos: *Silhouette analysis* (Tabla 4.1) y *Elbow analysis*.(Figura 4.5)

Número de Clústeres	Puntuación Promedio de Silueta
2	0.5629
3	0.5019
4	0.4821
5	0.4838
6	0.479

Cuadro 4.1: Scores para *Silhouette analysis*
Fuente: Elaboración propia

Estos resultados proporcionan una indicación de la cohesión y separación de los clústeres para diferentes configuraciones de agrupación y en ambos resultados observamos que la puntuación más alta se observa cuando el número de clústeres es 2 ($K = 2$), lo que sugiere que dos clústeres proporcionan la mejor cohesión y separación entre los datos. Sin embargo, la elección del número óptimo de clústeres en este estudio considera otros factores y validaciones.

En la elección del número de clústeres para este estudio, se optó por considerar 4 clústeres debido a que esta configuración se alinea con el estándar utilizado en categorizaciones similares observadas en otros estudios afines. Esta decisión se basa en la coherencia con prácticas

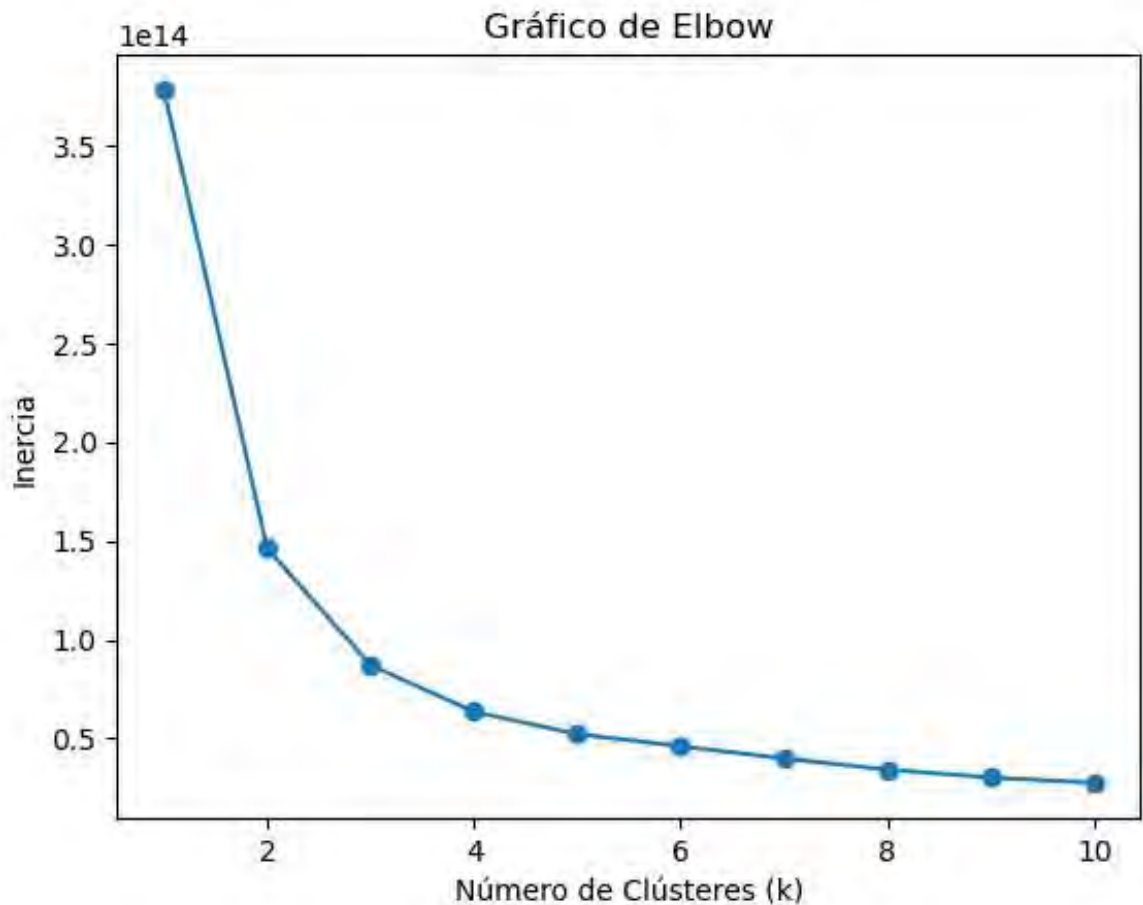


Figura 4.5: Scores para Elbow analysis.

Fuente: Elaboración propia

previas en la literatura y permite establecer comparaciones y correlaciones más efectivas. Además, al evaluar los resultados obtenidos a través de métricas como la puntuación de silueta y el método del codo (elbow), se observó que la configuración con 4 clústeres presenta valores razonables y no indica un decrecimiento abrupto en la varianza explicada. (Figura 4.6) Es importante destacar que al visualizar el gráfico de siluetas para $k=4$, se aprecia que todos los grupos se sitúan por encima de la línea roja vertical, indicando una buena cohesión y separación entre los clústeres. Esta convergencia de evidencia respalda la elección de 4 clústeres como una configuración apropiada para el análisis de la infraestructura de Internet en las áreas poblacionales consideradas en este estudio.

Los centroides, representando los valores medios de las métricas clave para cada clúster,

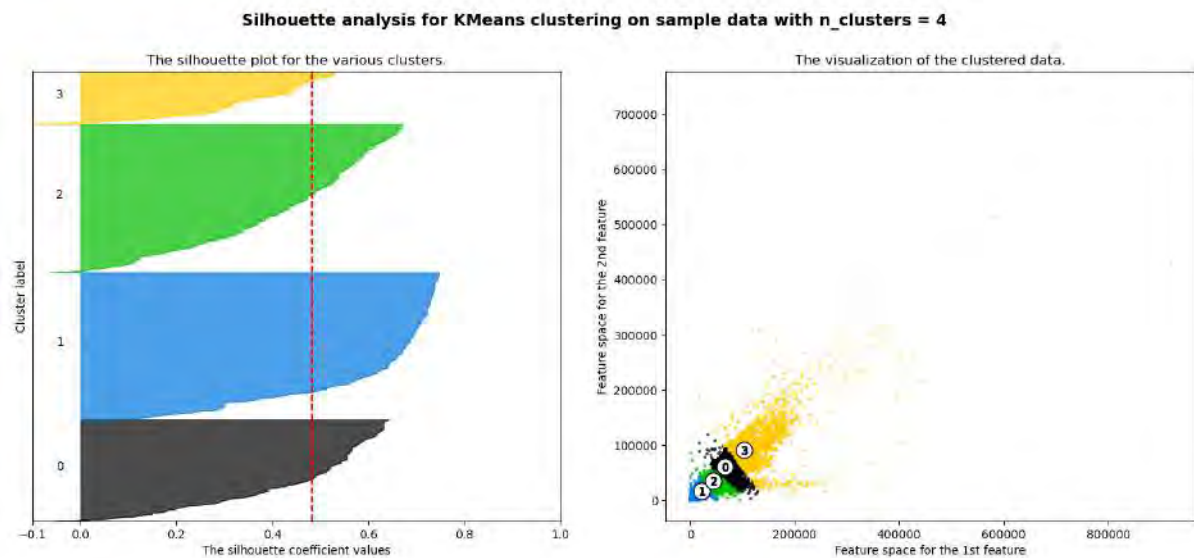


Figura 4.6: Gráfico de siluetas con número de clusters igual a 4.

La puntuación promedio es de 0.0.4821.

Fuente: Elaboración propia

brindan una instantánea de las características únicas de cada grupo identificado por el algoritmo de K-Means. En la Tabla 4.2, se detallan estos centroides para los cuatro clústeres, destacando diferencias significativas en las velocidades de descarga, carga y tiempos de latencia. El clúster 1 muestra un rendimiento particularmente alto, mientras que el clúster 3 presenta valores más bajos en comparación. Estos centroides proporcionan una comprensión concisa de las variaciones en la infraestructura de Internet entre las áreas poblacionales evaluadas.

download (kbps)	upload (kbps)	latency (ms)
44917.97	35818.08	44917.97
102950.31	90445.71	102950.31
65848.27	60390.68	65848.27
21676.64	17734.82	21676.64

Cuadro 4.2: Centroides de los Clústeres

Fuente: Elaboración propia

Sin embargo, que un centroide sea mejor que otro en algún aspecto no implica necesariamente que sea mejor en general, por lo que es necesario rankearlos objetivamente.

4.2.1. Visualización del Análisis de la Calidad del Servicio de e-Health

Para evaluar y comparar la calidad de cada centroide, se emplea la distancia coseno con respecto a un **vector ideal (I)** (3.1) derivado de los valores máximos de descarga y subida de datos identificados en la sección **Servicios de telemedicina y requerimientos técnicos de conexión**. Estos valores extremos sirven como referencia para definir un rendimiento ideal, donde se busca maximizar las velocidades de descarga y subida, y minimizar la latencia. Los límites máximo y mínimo de descarga se extraen de los Casos de uso de TI de salud y requisitos de infraestructura asociados Tabla 3.1, proporcionando un contexto realista para la evaluación de los centroides. Este enfoque permitió determinar qué centroides se acercan más a los estándares ideales establecidos por los requisitos técnicos de infraestructura asociados con servicios de telemedicina.

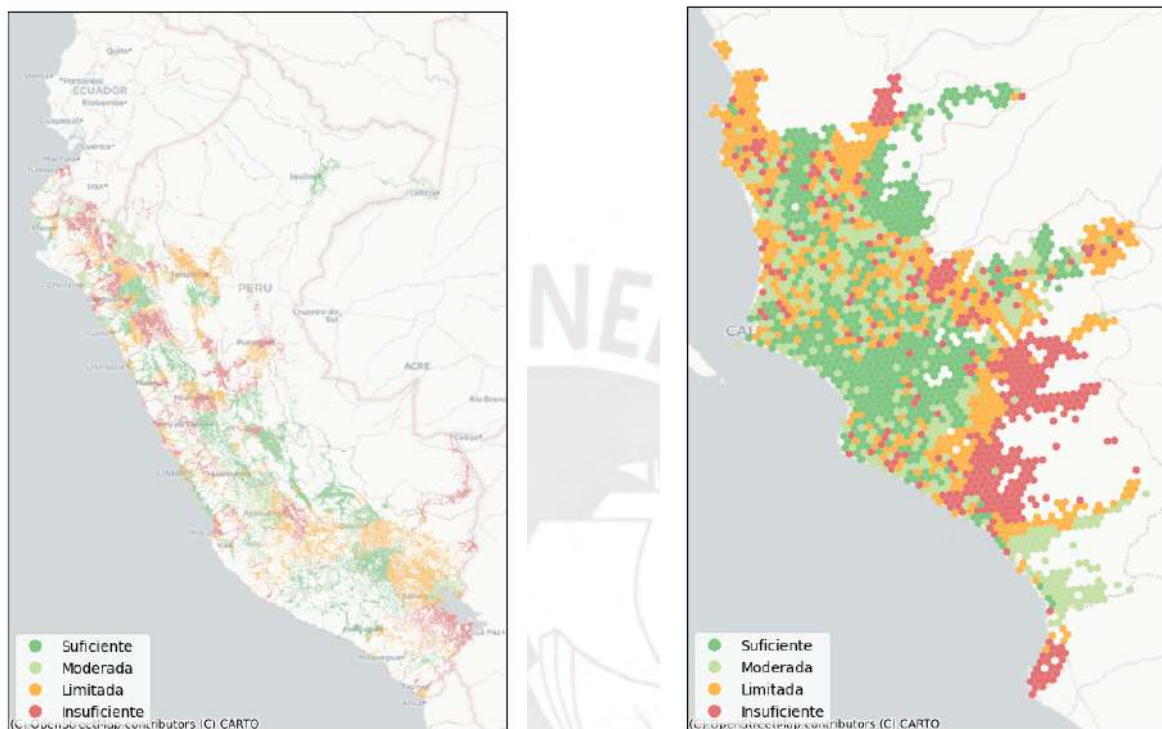
Los resultados Tabla 4.3 revelan la coherencia y similitud pronunciada entre los clusters, destacando el Cluster 3 como el más homogéneo con un índice de similitud coseno excepcionalmente alto de 0.9991. Este patrón se refleja en su primer puesto en el ranking de categorías, seguido por el Cluster 2, Cluster 1 y el Cluster 4 en orden descendente. El índice de similitud coseno, que cuantifica la similitud entre los vectores en cada cluster, sugiere una agrupación eficaz en términos de características compartidas. Los vectores característicos de cada cluster exhiben tendencias distintivas, como altos valores en todas las dimensiones para el Cluster 2 y valores más moderados para el Cluster 4.

Cluster	Vector	Índice de Similitud Coseno	Ranking de categorías
Cluster 3	(65848.27,60390.68,39.68)	0.9991	1
Cluster 2	(102950.31,90445.71,31.76)	0.9979	2
Cluster 1	(44917.97,35818.08,39.76)	0.9937	3
Cluster 4	(21676.64,17734.82,58.36)	0.995	4

Cuadro 4.3: Ranking de centroides basados en el índice de Similitud Coseno
Fuente: Elaboración propia

Estos resultados proporcionan una visión estructurada de la organización de datos en clus-

ters, las 4 categorías propuestas reflejan de manera precisa la capacidad de la infraestructura para satisfacer las necesidades específicas de la telemedicina en diferentes contextos. Para ello, generamos mapas geograficos de la calidad de la infraestructura de internet en el país Figura 4.7.



(a) *Ranking Nacional.*

(b) *Ranking en Lima y Callao.*

Figura 4.7: *Proxy de calidad de infraestructura digital.*

Fuente: Elaboración propia

Para una evaluación clara y ajustada a las visualización, proporcionando las siguientes etiquetas: Infraestructura de Internet Suficiente para Telemedicina, indicando un rendimiento sólido y consistente; Infraestructura de Internet Moderada para Telemedicina, destacando una opción aceptable con algunas limitaciones; Infraestructura de Internet Limitada para Telemedicina, señalando áreas con notables restricciones; y Infraestructura de Internet Insuficiente para Telemedicina, reservado para aquellas regiones donde la infraestructura no cumple con las expectativas mínimas.

4.3. Análisis de la cobertura digital del sistema de salud

Clasificación de la Calidad del Servicio

En esta sección, compartimos los resultados derivados del algoritmo implementado para evaluar la capacidad de cada hexágono de soportar cada servicios específicos de telemedicina Tabla 3.1, para esto fue necesario generar un código único para cada servicio.

- **Servicio 1:** Telepresencia
- **Servicio 2:** Telecirugía
- **Servicio 3:** Aplicaciones de Telemedicina en Tiempo Real (Streaming de Videos de Alta Tasa de Bits)
- **Servicio 4:** Sistemas de Referencia Médica para Médicos
- **Servicio 5:** Actualización en el Sitio del Historial Electrónico de Salud
- **Servicio 6:** Aplicaciones de Telemedicina basadas en Imágenes de Alta Resolución
- **Servicio 7:** Aplicaciones de Telemedicina en Tiempo Real (Streaming de Videos de Baja Resolución)
- **Servicio 8:** Monitoreo de Salud Móvil
- **Servicio 9:** Monitoreo de Salud en el Hogar (en Tiempo Real)
- **Servicio 10:** Gestión de Enfermedades Crónicas
- **Servicio 11:** Monitoreo de Salud en el Hogar (almacenamiento y reenvío)
- **Servicio 12:** Historial Electrónico de Salud o Historial de Salud Personal Rico en Contenido
- **Servicio 13:** Educación Médica Continua (EMC) a través de Dispositivos Móviles

- **Servicio 14:** Educación Médica Continua (EMC)
- **Servicio 15:** Notificación de Emergencia basada en Smartphone
- **Servicio 16:** Historial Electrónico de Salud o Historial de Salud Personal basado en Texto
- **Servicio 17:** Aplicaciones de Telemedicina de Almacenamiento y Reenvío
- **Servicio 18:** Comunicación e Informes a través de Correos Electrónicos
- **Servicio 19:** Notificación de Emergencia basada en Texto, Teléfono o SMS
- **Servicio 20:** Comunicación de Salud a través de Medios Sociales Relacionados con la Salud

Servicio	Población Habilitada	Porcentaje	Hexágonos Habilitadas
1	0	0.00 %	0.00 %
2	0	0.00 %	0.00 %
3	0	0.00 %	0.00 %
4	14773817	45.47 %	14.36 %
5	2604145	8.02 %	0.77 %
6	2604145	8.02 %	0.77 %
7	27445134	84.47 %	72.28 %
8	2900348	8.93 %	1.48 %
9	2900348	8.93 %	1.48 %
10	28666075	88.23 %	73.64 %
11	27445134	84.47 %	72.28 %
12	2900348	8.93 %	1.48 %
13	27445134	84.47 %	72.28 %
14	28666075	88.23 %	73.64 %
15	31798549	97.87 %	96.98 %
16	31926440	98.27 %	97.01 %
17	31798549	97.87 %	96.98 %
18	31926440	98.27 %	97.01 %
19	31926440	98.27 %	97.01 %
20	31926440	98.27 %	97.01 %

Cuadro 4.4: Capacidad de Infraestructura Digital para Telemedicina en el Perú
Fuente: Elaboración propia

En el análisis a nivel nacional Tabla 4.4, se evidencia que varios servicios de telemedicina enfrentan desafíos significativos en términos de accesibilidad. Por ejemplo, de los 20 servicios

evaluados, solo 3, como Monitoreo de Salud en el Hogar, Educación Médica Continua a través de Dispositivos Móviles y Historial Electrónico de Salud basado en Texto, tienen un potencial más amplio, llegando a una cobertura de más del 80% de la población.

Servicio	Población Habilitada	Porcentaje	Hexágonos Habilitados
1	0	0.00%	0.00%
2	0	0.00%	0.00%
3	0	0.00%	0.00%
4	8913411	82.79%	71.41%
5	2018140	18.75%	10.40%
6	2018140	18.75%	10.40%
7	10123348	94.03%	90.98%
8	2098733	19.49%	11.38%
9	2098733	19.49%	11.38%
10	10445150	97.02%	95.07%
11	10123348	94.03%	90.98%
12	2098733	19.49%	11.38%
13	10123348	94.03%	90.98%
14	10445150	97.02%	95.07%
15	10693113	99.32%	99.11%
16	10715780	99.53%	99.41%
17	10693113	99.32%	99.11%
18	10715780	99.53%	99.41%
19	10715780	99.53%	99.41%
20	10715780	99.53%	99.41%

Cuadro 4.5: Capacidad de Infraestructura Digital para Telemedicina en Lima y Callao
Fuente: Elaboración propia

En contraste, servicios como Telepresencia, Telecirugía y Aplicaciones de Telemedicina en Tiempo Real basadas en Streaming de Video de Alta Tasa de Bits presentan altas limitaciones, con un 0% de cobertura, sin beneficiar a algunos de los hexágonos con infraestructura habilitada.

Por otro lado, resultados para la región de Lima y Callao Tabla 4.5 muestran que, en términos de población que podría ser beneficiada, el servicio con mayor impacto es el Servicio 15, con un 99.32% de la población cubierta, seguido del Servicio 16, que tiene una cobertura del 99.53% de la población. Estos servicios demuestran una alta viabilidad y representan soluciones efectivas en el contexto de la infraestructura digital disponible en la región. Por otro lado,

servicios como el Servicio 1, Servicio 2 y Servicio 3 muestran una baja viabilidad, con un 0% de población beneficiada y sin cobertura de hexágonos con infraestructura habilitada.

Análisis de Cobertura de Servicios

Para complementar los resultados previos, realizamos una comparación gráfica entre el mapa de tiempos de accesibilidad (Figura 2.3) y los mapas de cobertura generados para cada servicio. Se analizó la capacidad de la infraestructura digital para soportar servicios de telemedicina tanto a nivel nacional, (Figura 0.2 - Anexo V). como a nivel Lima y Callao. (Figura 4.8) Ambos análisis destacan las limitaciones previamente mencionadas en la capacidad de la infraestructura digital para respaldar ciertos tipos de servicios de telemedicina en esta región específica.

Estos graficos junto a las métricas cuantitativas anteriores, proporcionan una comprensión detallada de cómo la infraestructura digital disponible impacta la accesibilidad a servicios de telemedicina en Lima y Callao, facilitando la identificación de áreas de mejora y la toma de decisiones informadas en la implementación de políticas de salud digital.

Análisis de Establecimientos de Salud por Servicio y Categoría

En esta sección, presentamos los resultados derivados de la evaluación de la capacidad de hospitales y población para proporcionar y recibir servicios de telemedicina.

En esta primera parte del análisis nos centraremos exclusivamente en los centros de salud de categoría 2 y 3, al enfocarnos en estas 2 categorías, buscamos obtener una imagen más precisa y aplicable de la capacidad de telemedicina en los centros de salud que están potencialmente equipados para brindar estos servicios.

En los resultados previos Tabla 4.6, identificamos 3 hospitales que no se encuentran dentro de la geometría de nuestra población kountur, lo que impide estimar sus indicadores de infraestructura de internet. Además, se identificaron 11 hospitales que, según la información



Figura 4.8: Gráfico de la capacidad de Infraestructura Digital para Telemedicina en Lima y Callao.

Fuente: Elaboración propia

Descripción del hallazgo	Hospitales
Hospitales que no caen en el shape de población Kountur	Hospital de Especialidades Medicas San Ignacio de Loyola, Clínica San Pedro, Hospital Distrital de Pacasmayo
Hospitales que no tienen capacidad de brindar telemedicina	Hospital de Apoyo de Pomabamba Antonio Caldas D., Santo Tomas, Hospital de Apoyo Sihuas, Hospital de Apoyo Otuzco Elpidio Beróvides Pérez, Santa Maria de Cutervo, Hospital de Apoyo Caraz, Clínica San Juan, Servicios Medicos Clinica Santa Anita SRL, Hospital Provincial de Tayabamba, Hospital de Apoyo Yungay, Hospital II Tarapoto

Cuadro 4.6: Hospitales sin capacidad o infraestructura que soporte servicios de telemedicina
Fuente: Elaboración propia

disponible, no tienen la capacidad de brindar ningún servicio de telemedicina. Este resultado proporcionan una visión importante sobre la disponibilidad de servicios de telemedicina en hospitales específicos y destacan la necesidad de considerar la ubicación geográfica al planificar la implementación de servicios de salud digital.

Continuando con el análisis y después de pasar por los filtros anteriores quedan 566 entre los de categoría 2 y 3, de los cuales hicimos una evaluación de capacidad de hospitales y población, y confirmamos que la infraestructura de Internet en ciertas áreas facilita la provisión de servicios de telemedicina, y podemos visualizar la distribución de hospitales en categorías 2 y 3, (Figura 4.9) según la capacidad de brindar diferentes servicios.

Por ejemplo, el servicio 5 - Actualización en el Sitio del Historial Electrónico de Salud cuenta con 192 hospitales de categoría 2 y 27 de categoría 3, lo que da un total de 219 hospitales. Asimismo, el servicio 6 - Aplicaciones de Telemedicina basadas en Imágenes de Alta Resolución tiene 158 hospitales de categoría 2 y 13 de categoría 3, sumando un total de 171 hospitales. Este análisis permite identificar las áreas donde la infraestructura digital respalda eficientemente la prestación de servicios de telemedicina, brindando información valiosa para la planificación y mejora de la atención médica remota.

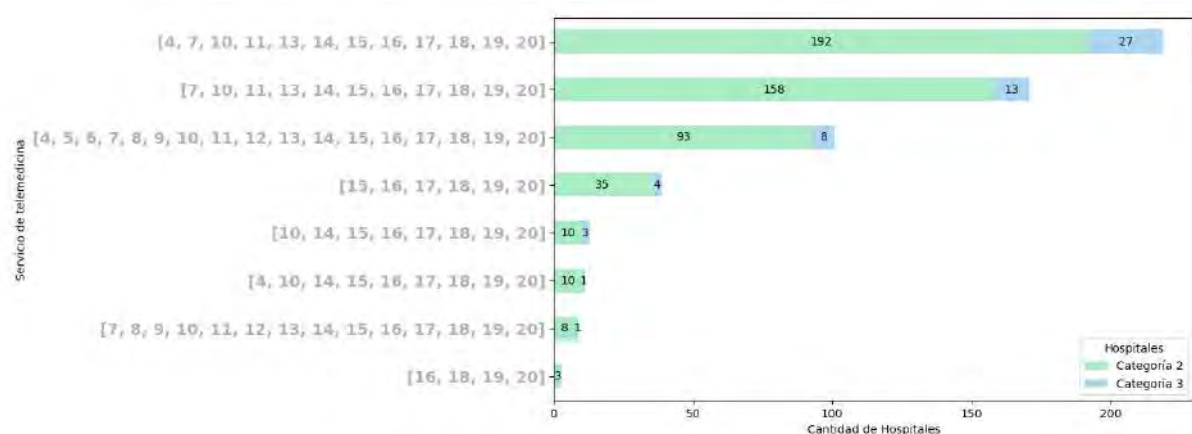


Figura 4.9: Cantidad de Hospitales por Servicio y Categoría.

Fuente: Elaboración propia

En esta segunda parte del análisis, nos enfocamos en los centros de categoría 1. Estos centros, por lo general, no ofrecen servicios de telemedicina o, en el peor de los casos, carecen de registros de datos relevantes para nuestro estudio. Sin embargo, es crucial incluirlos por varias razones:

- **Cobertura geográfica:** Los centros de categoría 1 tienen una distribución más densa, proporcionando una cobertura geográfica tanto a zonas rurales como urbanas.
- **Impacto social:** Mejorar la infraestructura de los centros de categoría 1 puede tener un impacto significativo en la calidad de vida de las poblaciones rurales.
- **Evaluación de infraestructura:** Aunque estos centros no atienden casos complejos, evaluar su infraestructura digital es esencial para identificar brechas y planificar mejoras en la red de telemedicina. Además, estos centros requieren menos inversión comparado con los de categoría 2 y 3 para poder brindar este servicio.

Al analizar los centros de salud primaria (PHC) (Figura 4.10), podemos observar que están ubicados en zonas con una alta capacidad para cubrir servicios de telemedicina relacionados con comunicaciones y notificaciones, con cerca de 10,000 PHC, representando aproximadamente el 90% de los centros evaluados. En contraste, la capacidad para ofrecer servicios más avanzados y especializados, como aplicaciones de telemedicina basadas en imágenes de alta

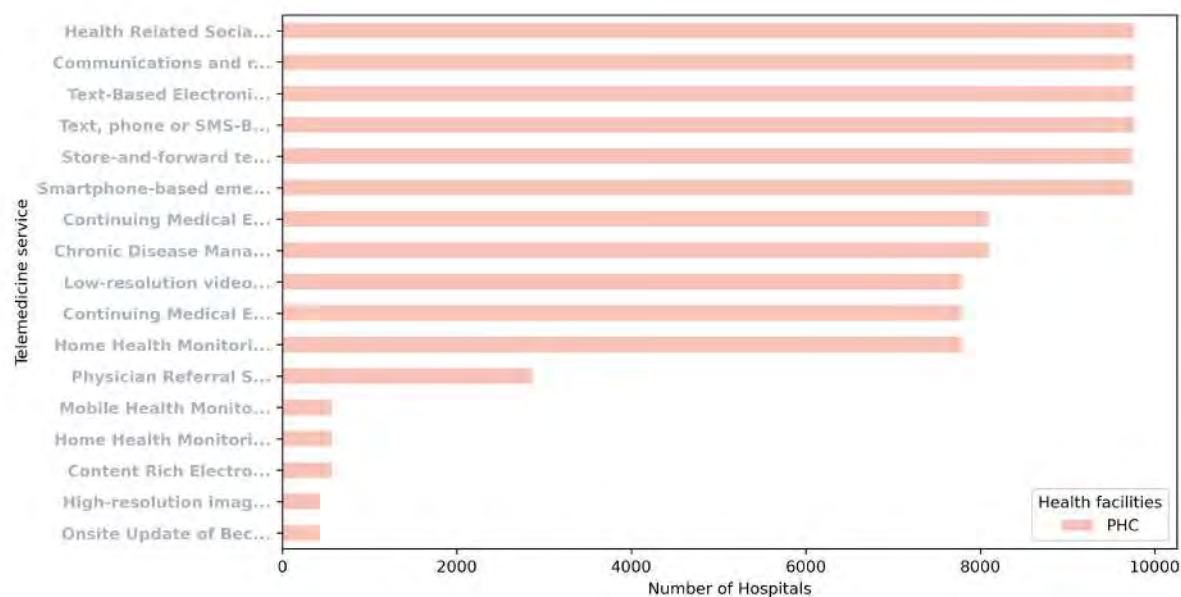


Figura 4.10: Cantidad de Centros de atención primaria por Servicio y Categoría.

Fuente: Elaboración propia

resolución y la actualización en sitio de registros electrónicos de salud, es mucho menor, con menos del 10%. Esto indica que estos centros de salud solo necesitarían una infraestructura básica adecuada para ciertos servicios, lo que les permitiría cubrir una gran cantidad de la población. (Tabla 4.4)

Capítulo V

DISCUSIÓN Y CONCLUSIONES

La metodología aplicada en este estudio ha sido esencial para proporcionar una base sólida para entender la robustez de la infraestructura digital a nivel nacional y evaluar su calidad. Asimismo, la catalogación de servicios y sus requerimientos mínimos es uno de los puntos más importantes de este estudio. Esto nos permitió establecer un ranking de la calidad de la infraestructura digital en relación con estos servicios y visualizar geográficamente las zonas que podrían cubrir algún servicio.

En este trabajo, se han revelado valiosos *insights* sobre la infraestructura digital y la capacidad de los centros de salud para brindar servicios de telemedicina en diversas regiones del país. La categorización de la población en hexágonos H3 de Kontur Data permitió una evaluación precisa de la conectividad, evidenciando áreas con limitaciones y otras con una infraestructura robusta. La aplicación de técnicas de clustering mediante K-Means proporcionó una segmentación clara, permitiendo categorizar la conectividad en grupos homogéneos.

Los resultados indican que más del 90% de la población se encuentra en áreas donde la calidad de la infraestructura soporta más del 70% de los servicios de telemedicina catalogados. Sin embargo, esta capacidad no se refleja en la adopción y uso efectivo de estos servicios en el Perú. Esta brecha entre la infraestructura disponible y su utilización práctica en aplicaciones de e-health debe ser abordada con rigor y podría deberse a factores sociodemográficos, como la pobreza de internet y la pobreza monetaria, o a factores técnicos, como la falta de capacitación, la resistencia al cambio tecnológico o problemas regulatorios. Algo a resaltar de los resultados es que ninguna zona geográfica está en capacidad de implementar servicios con requerimientos tecnológicos más elevados, como telepresencia, telecirugía o streaming de videos de alta tasa de bits. En contraste, en casi toda el área nacional, 11 de los servicios están cubiertos en más de un 70% y los 6 restantes están cubiertos en algunas zonas, lo cual incluye a los centros de

salud.

En particular, la categorización de hospitales en las categorías 2 y 3, y la posterior evaluación de su capacidad para servicios específicos, sugiere la necesidad de abordar las brechas tecnológicas de manera focalizada para cada servicio. Además, la identificación de centros de salud que no se encuentran en una zona de cobertura subraya la importancia de considerar la expansión de la infraestructura digital a nivel nacional.

Se podría incentivar que los centros de atención primaria (PHC) brinden los servicios básicos de telemedicina, ya que los requerimientos especializados no son altos y la calidad de la infraestructura digital ya está en capacidad de soportarlos.

Por otro lado, trabajar con la distribución de población ajustada como hexágonos H3 Kountur, plantea la posibilidad de agregación a niveles más amplios, como distritales, provinciales, regionales y nacionales. Esta escalabilidad permite una contextualización más amplia de los resultados y facilita la identificación de patrones a nivel macro, proporcionando información valiosa para formuladores de políticas y planificadores de salud a diferentes escalas administrativas.

Este trabajo estuvo limitado por varios factores importantes. En primer lugar, la evaluación de la infraestructura digital se basó exclusivamente en datos de pruebas de velocidad en dispositivos de escritorio mediante Speedtest. La omisión de datos móviles podría generar una visión sesgada, ya que la conectividad móvil juega un papel vital, especialmente en áreas donde las conexiones fijas son limitadas. Este enfoque centrado en datos de escritorio proporciona una instantánea inicial, pero futuros trabajos deben integrar datos móviles para obtener una comprensión más completa del panorama digital en salud.

Además, la manera en que abordamos la temporalidad de los datos, utilizando un único indicador por polígono, podría ser mejorada. En trabajos futuro planeamos considerar los picos de demanda y generar múltiples escenarios para reflejar mejor las variaciones en la conectividad. Sin embargo, la limitación en este aspecto es computacional, ya que requeriría un procesamiento de datos significativamente mayor para modelar estos escenarios de manera efectiva.

En síntesis, la primera parte de la metodología propone cómo preparar los datos y llenar las lagunas de datos, limitándonos a la distribución poblacional. La segunda parte de la metodología ofrece una evaluación nacional de la calidad de la infraestructura, sin embargo, la presentación elegida está limitada al formato utilizado por el MINSA, contemplando cuatro categorías. La tercera parte, la introducción al estudio de los centros de salud, proporciona una visión holística de la situación, brindando una base sólida para la toma de decisiones en materia de salud pública y resaltando la necesidad de inversiones y estrategias específicas para mejorar la infraestructura digital en ciertas áreas.

Finalmente, el enfoque metodológico proporciona una herramienta escalable y adaptable para futuros análisis de salud pública, contribuyendo significativamente al panorama general de la telemedicina en el país.

En trabajos futuros, se considerará la inclusión de los horarios de atención de los centros de salud. Además, se categorizará la calidad de la infraestructura en las áreas que ya cuentan con cobertura dentro del análisis de la cobertura de servicios. Por otro lado, este mismo estudio integrará factores sociodemográficos de la población para realizar un análisis más profundo, permitiendo así entender la factibilidad de que un habitante pueda acceder a servicios que le permitan conectarse a la infraestructura digital.

Referencias

- Alegana, V. A., Wright, J., Pezzulo, C., Tatem, A. J., y Atkinson, P. M. (2017, 20 de Apr). Treatment-seeking behaviour in low- and middle-income countries estimated using a bayesian model. *BMC Medical Research Methodology*, 17(1), 67. Descargado de <https://doi.org/10.1186/s12874-017-0346-0> doi: 10.1186/s12874-017-0346-0
- Balestra, M. (2018). Telehealth and legal implications for nurse practitioners. *The Journal for Nurse Practitioners*, 14(1), 33-39. Descargado de <https://www.sciencedirect.com/science/article/pii/S1555415517308085> doi: <https://doi.org/10.1016/j.nurpra.2017.10.003>
- Bank, W. (Año de la publicación 2022). *Is el salvador's health care system ready for telemedicine?* Descargado de <https://blogs.worldbank.org/digital-development/el-salvadors-health-care-system-ready-telemedicine> (Consultado en julio de 2023)
- Battle, K. E., Bisanzio, D., Gibson, H. S., Bhatt, S., Cameron, E., Weiss, D. J., ... Gething, P. W. (2016, 11 de Jan). Treatment-seeking rates in malaria endemic countries. *Malaria Journal*, 15(1), 20. Descargado de <https://doi.org/10.1186/s12936-015-1048-x> doi: 10.1186/s12936-015-1048-x
- Cuba García, H. (2021). *La pandemia en el Perú*. Colegio Médico del Perú. (Ciencias médicas Medicina)
- Dirección de telemedicina del ministerio de salud (minsa). (s.f.). <https://www.gob.pe/23164-ministerio-de-salud-direccion-de-telemedicina>. (Consulta: diciembre de 2023)
- Feamster, N., y Livingood, J. (2020, nov). Measuring internet speed: Current challenges and future recommendations. *Commun. ACM*, 63(12), 72–80. Descargado de <https://doi.org/10.1145/3372135> doi: 10.1145/3372135
- Federal Communications Commission. (2010). *Health care broadband in america*. <https://transition.fcc.gov/national-broadband-plan/health-care>

- broadband-in-america-paper.pdf. (Consulta: [Noviembre 2023])
- for Good at Meta, D. (2022). *Movement range maps*. Descargado de <https://data.humdata.org/m/dataset/movement-range-maps> (Consultado en abril de 2023)
- Gozzi, N., Comini, N., y Perra, N. (2022). *The adoption of non-pharmaceutical interventions and the role of digital infrastructure during the covid-19 pandemic in colombia, ecuador, and el salvador*.
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., ... Tatlow, H. (2021, 01 de Apr). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 5(4), 529-538. Descargado de <https://doi.org/10.1038/s41562-021-01079-8> doi: 10.1038/s41562-021-01079-8
- Instituto Nacional de Estadística e Informática. (2022, diciembre). *Infraestructura de datos espaciales inei*. Descargado de <https://ide.inei.gob.pe/> (Consultado en diciembre de 2022)
- Kontur. (Año de la última actualización 2023). *Peru: Population density for 400m h3 hexagons*. Descargado de <https://data.humdata.org/m/dataset/kontur-population-peru> (Consultado en agosto de 2023)
- Manongi, R., Mtei, F., Mtove, G., Nadjm, B., Muro, F., Alegana, V., ... Reyburn, H. (2014, marzo). Inpatient child mortality by travel time to hospital in a rural area of tanzania. *Trop Med Int Health*, 19(5), 555–562.
- Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., ... Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*. (<https://ourworldindata.org/coronavirus>)
- Ministerio de Salud. (2011). *Resolución ministerial n°546-2011-minsa*. (Publicado en el Boletín de Normas Legales del Diario Oficial .El Peruano)
- Ministerio de Salud. (2021). *Establecimientos de salud*. Datos Abiertos. Ministerio de Salud. Descargado de <https://www.datosabiertos.gob.pe/dataset/establecimientos-de-salud> (Consultado en 2023)

- Nelson, A., Weiss, D. J., van Etten, J., Cattaneo, A., McMenemy, T. S., y Koo, J. (2019, 07 de Nov). A suite of global accessibility indicators. *Scientific Data*, 6(1), 266. Descargado de <https://doi.org/10.1038/s41597-019-0265-5> doi: 10.1038/s41597-019-0265-5
- Ookla. (2023). *Ookla for good*. Descargado de <https://www.speedtest.net/insights/blog/tag/ookla-for-good/> (Datos proporcionados en 2023)
- Steele, R., y Lo, A. (2013, 01 de Mar). Telehealth and ubiquitous computing for bandwidth-constrained rural and remote areas. *Personal and Ubiquitous Computing*, 17(3), 533-543. Descargado de <https://doi.org/10.1007/s00779-012-0506-5> doi: 10.1007/s00779-012-0506-5
- Superintendencia Nacional de Salud (SUSALUD). (2023). *Registro nacional de instituciones prestadoras de servicios de salud (renipress)*. Descargado de <http://app20.susalud.gob.pe:8080/registro-renipress-webapp/listadoEstablecimientosRegistrados.htm?action=mostrarBuscar#no-back-button> (Consultado en 2023)
- Tariverdi, M., Nunez-del Prado, M., Leonova, N., y Rentschler, J. (2023, 28 de Jan). Measuring accessibility to public services and infrastructure criticality for disasters risk management. *Scientific Reports*, 13(1), 1569. Descargado de <https://doi.org/10.1038/s41598-023-28460-z> doi: 10.1038/s41598-023-28460-z
- Telehealth, telemedicine, and telecare: What's what*. (s.f.). Descargado de <https://www.fcc.gov/general/telehealth-telemedicine-and-telecare-whats-what> (Consulta: [fecha de tu consulta])
- Weiss, D. J., Nelson, A., Vargas-Ruiz, C. A., Gligorić, K., Bavadekar, S., Gabrilovich, E., ... Gething, P. W. (2020, 01 de Dec). Global maps of travel time to healthcare facilities. *Nature Medicine*, 26(12), 1835-1838. Descargado de <https://doi.org/10.1038/s41591-020-1059-1> doi: 10.1038/s41591-020-1059-1

Anexo

Mapas de hexagonos indexados

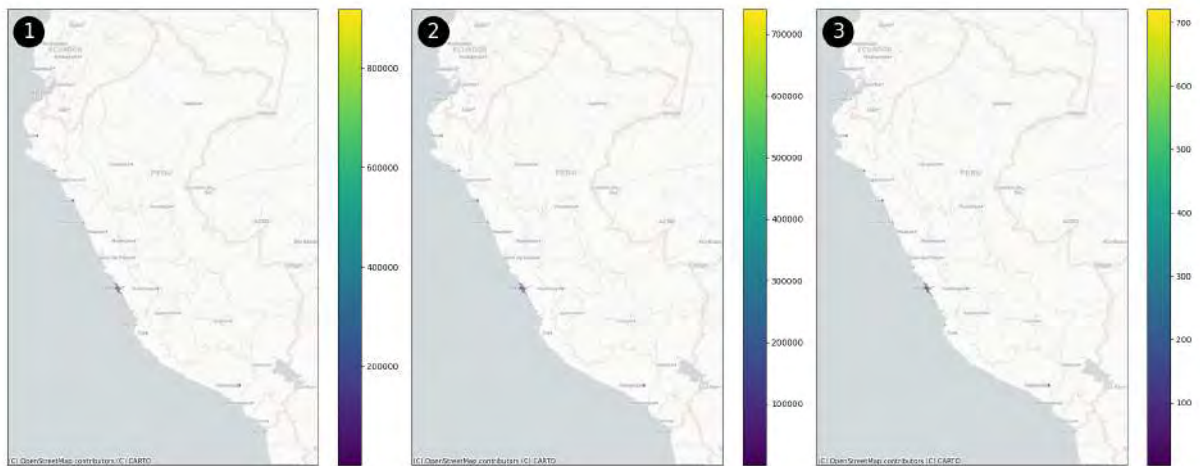


Figura 0.1: Mapas de indicadores de calidad de red en Perú antes de completar los datos ausentes.

(1) velocidad de descarga (kbps), (2) velocidad de carga (kbps), y (3) latencia; resultados después de completar las lagunas de datos.

Fuente: Elaboración propia

Capacidad de Infraestructura Digital para Soportar Servicios de Telemedicina: Un Análisis Nacional

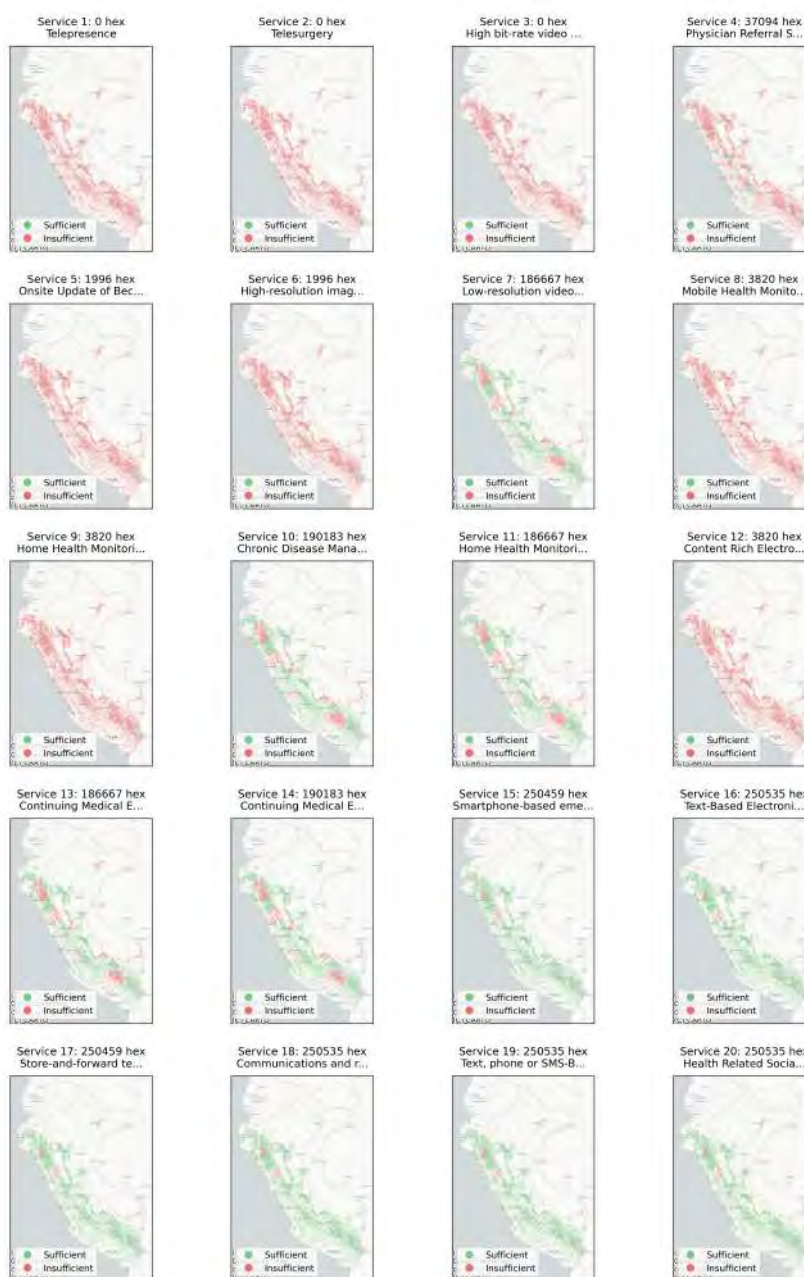


Figura 0.2: Gráfico de la capacidad de Infraestructura Digital para Telemedicina en el Perú.
Fuente: Elaboración propia