

**UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO**  
**FACULTAD DE INGENIERÍA**  
**ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN**



**Aplicación móvil basado en machine learning para facilitar la comunicación de personas con discapacidad auditiva en el Perú**

**TESIS PARA OPTAR EL TÍTULO DE  
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

**AUTOR**

**Maricarmen Lopez Romero**

**ASESOR**

**Luis Augusto Zuñe Bispo**

<https://orcid.org/0000-0001-7838-8656>

**Chiclayo, 2026**

**Aplicación móvil basado en machine learning para facilitar la  
comunicación de personas con discapacidad auditiva en el Perú**

PRESENTADA POR  
**Maricarmen Lopez Romero**

A la Facultad de Ingeniería de la  
Universidad Católica Santo Toribio de Mogrovejo  
para optar el título de

**INGENIERO DE SISTEMAS Y COMPUTACIÓN**

APROBADA POR

Huiler Juanito Mera Montenegro  
PRESIDENTE

Jury Yesenia Aquino Trujillo  
SECRETARIO

Luis Augusto Zuñe Bispo  
VOCAL

## **Dedicatoria**

A mi padre, por su apoyo incondicional y por haberme acompañado con amor y fortaleza en cada etapa de mi vida.

## **Agradecimientos**

A toda mi familia, por su constante motivación y por enseñarme a enfrentar los desafíos con fortaleza y perseverancia. Su amor y apoyo incondicional han sido pilares fundamentales en mi formación personal y profesional.

# Aplicación móvil basado en machine learning para facilitar la comunicación de personas con discapacidad auditiva en el Perú

## INFORME DE ORIGINALIDAD



## FUENTES PRIMARIAS

<b>1</b>	<b>hdl.handle.net</b> Fuente de Internet	<b>4%</b>
<b>2</b>	<b>www.gob.pe</b> Fuente de Internet	<b>2%</b>
<b>3</b>	<b>tesis.usat.edu.pe</b> Fuente de Internet	<b>1%</b>
<b>4</b>	<b>revistatransregiones.com</b> Fuente de Internet	<b>1%</b>
<b>5</b>	<b>www.infobae.com</b> Fuente de Internet	<b>1%</b>
<b>6</b>	<b>revistamedica.com</b> Fuente de Internet	<b>&lt;1%</b>
<b>7</b>	<b>Submitted to Universidad Católica Santo Toribio de Mogrovejo</b> Trabajo del estudiante	<b>&lt;1%</b>
<b>8</b>	<b>Submitted to Universidad Carlos III de Madrid - EUR</b>	<b>&lt;1%</b>

## Índice

Resumen .....	6
Abstract .....	7
Revisión de literatura .....	11
Materiales y métodos .....	16
Resultados y discusión .....	17
Conclusiones .....	28
Recomendaciones.....	29
Referencias .....	30
Anexos.....	34

## Resumen

El proyecto tiene como objetivo desarrollar una aplicación móvil basado en machine learning para facilitar la comunicación de personas con discapacidad auditiva en el Perú. Esta investigación, de tipo aplicada, aborda como problema las limitaciones en la comunicación de las personas con discapacidad auditiva profunda. Para ello, se aplicaron métodos de investigación, entre ellos la revisión de literatura, el enfoque deductivo y técnicas de codificación. Además, entrevistas y encuestas a una muestra de 60 personas de la comunidad junto con evaluaciones realizadas por especialistas en Lenguaje de Señas Peruano. En el desarrollo del sistema, se implementaron dos metodologías: la metodología de Machine Learning, que se utilizó para la construcción del modelo computacional, y la metodología de Extreme Programming, para el desarrollo de la aplicación móvil. Se recopilaron 15,000 puntos claves correspondientes a 30 clases, pertenecientes a las señas estáticas y dinámicas que más se utilizan en una conversación, los cuales se utilizaron para entrenar una red neuronal densamente conectada. Posteriormente, se utilizó OpenNLP para estandarizar el texto de entrada durante el proceso de generación. Como resultado de este enfoque, se logró una precisión del 85% en entrenamiento y 97% en validación del modelo, lo que demuestra la eficacia del sistema propuesto. Por último, la aplicación móvil fue validada en entornos reales, permitiendo ajustes que mejoraron su accesibilidad y usabilidad para las personas con esta discapacidad, asegurando una respuesta más efectiva a sus necesidades.

**Palabras clave:** Discapacidad auditiva, lenguaje de señas, machine learning, señas estáticas, señas dinámicas.

## Abstract

The project aims to develop a mobile application based on machine learning to facilitate the communication of people with hearing disabilities in Peru. This research, of an applied type, addresses as a problem the limitations in the communication of people with profound hearing impairment. For this purpose, research methods were applied, including literature review, deductive approach, and coding techniques. In addition, interviews and surveys were conducted with a sample of 60 people from the community along with evaluations conducted by specialists in Peruvian Sign Language. In the development of the system, two methodologies were implemented: the Machine Learning methodology, which was used for the construction of the computational model, and the Extreme Programming methodology, for the development of the mobile application. A total of 15,000 key points corresponding to 30 classes, belonging to the static and dynamic signs most used in a conversation, were collected and used to train a densely connected neural network. Subsequently, OpenNLP was used to standardize the input text during the generation process. As a result of this approach, an accuracy of 85% in training and 97% in model validation was achieved, demonstrating the effectiveness of the proposed system. Finally, the mobile application was validated in real environments, allowing adjustments that improved its accessibility and usability for people with this disability, ensuring a more effective response to their needs.

**Keywords:** Hearing impairment, sign language, machine learning, static signing, dynamic signing.

## Introducción

La discapacidad auditiva (DA) se describe como la dificultad para detectar y comprender los estímulos sonoros del entorno [1]. Las personas que presentan esta discapacidad que varía desde niveles leves hasta graves, a menudo utilizan la comunicación oral y subtítulos. En contraste, quienes experimentan una discapacidad auditiva profunda, es decir que tienen dificultades significativas o nulas para escuchar, suelen comunicarse mediante el lenguaje de señas [2].

Según la Organización Mundial de la Salud al 2 de febrero de 2024, existen aproximadamente 430 millones de personas con DA en el mundo; y se proyecta que para el año 2050 esta cifra se elevará por encima de los 700 millones de personas [2]. Asimismo, de acuerdo al Instituto Nacional de Estadística al 2023, en España existen 86,315 personas que tienen DA, de las cuales 33,163 presentan discapacidad auditiva profunda [3]. En América Latina y el Caribe, las personas con DA representan el 21,52% de la población [4]. En relación con México, conforme a lo reportado por el Instituto de Salud para el Bienestar en 2023, se prevé que cerca de 2,3 millones de personas presentan DA. Por su parte, en Perú, según el registro nacional del Conadis (Consejo Nacional para la Integración de la Persona con Discapacidad) al 21 de enero de 2023, 39,048 personas con DA se encuentran inscritas; de las cuales, 17,500 son mujeres y 21,548 varones [5].

Según indica Edgar Espejo Muñoz, docente y especialista en Lenguaje de Señas Peruano (LSP) en Chiclayo, las personas con discapacidad auditiva (DA) atraviesan desafíos en diversos ambientes al intentar comunicarse con oyentes. Por ejemplo, en la educación, las tecnologías que emplearon las instituciones durante las clases virtuales en la pandemia de COVID-19 no resultaron efectivas para los estudiantes con DA, ya que solo el 10% de los profesores conocían y/o practicaban el LSP y utilizaban canales de mensajería para entregar las tareas y materiales educativos a estos estudiantes. A esto se suma que hay menos de 100 intérpretes acreditados, lo cual es una situación atribuida a la falta de acción del Estado y no al desinterés de las personas por aprender el LSP [6]. Estas limitaciones contribuyen a que las barreras sociales lleven a situaciones de exclusión de las personas con DA en reuniones remotas, conversaciones y actividades sociales y/o culturales. Asimismo, en las empresas, los escasos mecanismos adaptados en las entrevistas laborales remotas para estas personas limitan el acceso a información clave y oportunidades de participación (ver Anexo 02).

Ahora bien, la discapacidad auditiva (DA) puede manifestarse en diversos grados y tipos, y puede afectar a uno o ambos oídos. En particular, el tipo más grave de esta condición es la discapacidad auditiva profunda, que se define cuando el umbral de audición supera los 90 decibelios (dB). En consecuencia, las personas con esta condición pueden escuchar poco o nada

en absoluto y, con frecuencia, hacen uso del lenguaje de señas o de otros métodos de comunicación visual [7].

La discapacidad auditiva profunda puede desarrollarse en diferentes etapas de la vida debido a una variedad de condiciones. Estos incluyen factores genéticos, que pueden ser hereditarios o resultantes de mutaciones genéticas, infecciones que ocurren durante el embarazo, asfixia perinatal, hiperbilirrubinemia que puede causar daño auditivo severo en el recién nacido, entre otros [7].

Como consecuencia, la discapacidad auditiva profunda impacta significativamente en distintos ámbitos de la vida de una persona. Por ejemplo, puede causar problemas en la comunicación verbal y en la articulación del lenguaje, así como complicaciones en el procesamiento cognitivo y en la memoria debido a la ausencia de estimulación auditiva. Además, el aislamiento social y la soledad pueden ser frecuentes, ya que las dificultades para comunicarse pueden llevar a una mayor separación de los demás (oyentes) [7]. A esto se suma, la ausencia de adaptación en espacios recreativos y eventos de entretenimiento, junto con programas educativos especializados, contribuye a la exclusión y limita las posibilidades de desarrollo académico, profesional y social para estas personas [8].

Teniendo en cuenta lo anterior, el problema de investigación seleccionado son las limitaciones que tienen las personas con discapacidad auditiva profunda para comunicarse. Estas dificultades se deben a la variabilidad en las señas y la carencia de medios efectivos y no intrusivos, lo cual limita la comunicación de esta comunidad, contribuyendo a su aislamiento y exclusión en diversos contextos sociales. Además, los dispositivos de asistencia disponibles no siempre son accesibles o asequibles. Basándonos en lo mencionado anteriormente, se formula la siguiente pregunta: ¿De qué manera la tecnología de Inteligencia Artificial ayudará a las personas con discapacidad auditiva profunda a comunicarse con aquellos que tienen un conocimiento limitado del lenguaje de señas peruano?

La ejecución de esta aplicación móvil impulsada por Inteligencia Artificial tiene un impacto significativo en la dimensión social al abordar las limitaciones de comunicación que enfrentan las personas con discapacidad auditiva profunda. Al proporcionar una base inicial accesible para la interacción entre personas con esta discapacidad y aquellas con conocimiento limitado del Lenguaje de Señas Peruano (LSP), la aplicación funciona como un facilitador para el diálogo inclusivo. Además, esta herramienta tecnológica sirve como puente inicial, permitiendo establecer un primer contacto efectivo, lo que a su vez promueve la inclusión y la participación de esta comunidad en diversas esferas sociales. Esta iniciativa fomenta la igualdad de oportunidades al facilitar estos encuentros comunicativos iniciales, reduciendo así las barreras

de comunicación y sentando las bases para fortalecer los lazos sociales y las relaciones interpersonales a largo plazo.

Desde una perspectiva económica, la implementación de esta tecnología tiene múltiples beneficios. Por un lado, al mejorar la capacidad de comunicación de las personas con discapacidad auditiva profunda, se amplían sus posibilidades de participación en el mercado laboral y a una reducción de las disparidades económicas asociadas a la discapacidad. Además, el desarrollo y la comercialización de esta aplicación móvil puede generar oportunidades de empleo en el sector tecnológico y contribuir al crecimiento económico a través de la innovación y la creación de nuevos productos y servicios en el Perú.

Desde una perspectiva científica, representa un progreso en el campo de la Inteligencia Artificial aplicada a la accesibilidad y la inclusión. La combinación de técnicas de Machine Learning y Procesamiento de Lenguaje Natural para traducir gestos del LSP a texto representa un desafío técnico importante. La elección de modelos de aprendizaje automático apropiados, que sean ligeros y precisos, para el procesamiento y reconocimiento de gestos estáticos y dinámicos es crucial para crear sistemas efectivos en este ámbito. Además, abre nuevas oportunidades de investigación en tecnologías no intrusivas y la adaptación de sistemas de Inteligencia Artificial para atender las necesidades específicas de personas con discapacidad auditiva profunda.

En el planteamiento de la investigación se traza como objetivo general desarrollar una aplicación móvil basado en Machine Learning para facilitar la comunicación de personas con discapacidad auditiva en el Perú. Como objetivos específicos se identificaron y cuantificaron las características posturales y cinemáticas distintivas de los gestos del Lenguaje de Señas Peruano; posteriormente se evaluó comparativamente las herramientas de Machine Learning y Procesamiento de Lenguaje Natural para el reconocimiento, así como traducción en tiempo real del Lenguaje de Señas Peruano a texto español. Luego, se diseñó una interfaz de usuario amigable para la aplicación móvil basada en el estándar de calidad ISO/IEC 25010:2011, que permitió a las personas con discapacidad auditiva profunda utilizarla de manera efectiva. Asimismo, se implementó la integración de la funcionalidad de reconocimientos de gestos, asegurando su precisión y fiabilidad en el entorno de interacción del usuario. Adicionalmente, se verificó la funcionalidad de la aplicación móvil mediante evaluación de especialistas del Lenguaje de Señas Peruano que aseguró su correcto funcionamiento.

## Revisión de literatura

Amangeldy et al. [9] señalaron que el reconocimiento continuo de la lengua de signos (CSLR) enfrenta desafíos como la variabilidad de las señas, el contexto y la entonación, así como limitaciones tecnológicas. No obstante, su desarrollo es crucial para mejorar la comunicación y la autonomía de las personas con discapacidad auditiva. Para abordar estos desafíos, propusieron un método que combina CSLR y procesamiento de lenguaje natural (PNL). Para el CSLR, utilizaron MediaPipe para recopilar datos de gestos y una red LSTM de 1024 unidades para analizar 1495 muestras de gestos, cada una representada por 60 fotogramas, resultando en un corpus de entrenamiento de 89,700 fotogramas. En cuanto al PNL, desarrollaron un procesador específico para el kazajo con un analizador morfológico, sintáctico y un corrector morfológico, empleando una base de datos de más de 320,000 entradas y 30 características semánticas. Como resultado, el modelo alcanzó una precisión promedio de 0.97, lo que permitió la creación de 20 estructuras de oraciones kazajas y modelos de entonación.

Sreemathy et al. [10] indicaron que el lenguaje de signos indio utiliza ambas manos para la mayoría de las señas, lo que requiere un seguimiento preciso de la posición y orientación de las manos. Para facilitar la comunicación diaria de personas con discapacidades, desarrollaron un sistema experto utilizando dos modelos: YOLOv4 y Support Vector Machine (SVM) con media-pipe. Utilizaron un conjunto de datos de 80 clases con 13,000 imágenes para el reconocimiento de gestos estáticos con SVM y media-pipe, y pre procesaron las imágenes convirtiéndolas de BGR a RGB. Media-pipe se usó para detectar la palma y extraer puntos clave, dividiendo el conjunto de datos en 80% para entrenamiento y 20% para validación. Para el sistema con YOLOv4, emplearon Gramformer para corregir errores gramaticales en textos y etiquetaron imágenes en formato de texto mediante Roboflow, entrenándolas con DarkNet-53. El sistema experto combinó ambos modelos para comparar la precisión en tiempo real, obteniendo una precisión de 98.62% con SVM y 98.8% con YOLOv4.

Cataño et al. [11] abordaron la problemática de comunicación entre personas con discapacidad auditiva y oyentes, destacando la necesidad de herramientas que faciliten el entendimiento de la Lengua de Señas Mexicana (LSM). Para ello, desarrollaron un sistema de detección del abecedario LSM utilizando MediaPipe para extraer características de las imágenes de señas, generando vectores de coordenadas de 21 puntos clave de las manos. Además, crearon una base de datos propia con 5,000 imágenes, capturando 100 fotos por letra del abecedario en ambientes controlados, con versiones de una y dos manos. Para la clasificación, implementaron y compararon los algoritmos Support Vector Machine (SVM) y Random Forest. El sistema fue entrenado con el 80% de las imágenes y probado con el 20%

restante. Los resultados alcanzaron un 99.4% de efectividad con SVM para dos manos, mientras que Random Forest logró entre un 97.3% y 97.6%.

Bejarano et al. [12] señalaron que el desarrollo de modelos para el reconocimiento del Lenguaje de Señas Peruano se ve restringido por la escasez de datos en videos de señas continuas. Esto se debe tanto a la dificultad de grabar a usuarios nativos como al limitado número de intérpretes disponibles en el país. Para abordar esta limitación, diseñaron un conjunto de datos continuos compuesto por 2,000 instancias y más de 150 oraciones en señas, correspondientes a 500 señas únicas. Estos videos se obtuvieron de un programa educativo del gobierno llamado “Aprendo en Casa”, en el cual intérpretes traducían contenido audiovisual. Para el procesamiento, emplearon ELAN para la segmentación de los videos y posteriormente Mediapipe para extraer puntos clave en el rostro, cuerpo y manos. Además, adaptaron ChaLearn para optimizar las secuencias de video y utilizaron ResNet para la extracción de características. El modelo alcanzó una precisión del 80.3% y 52.4% en un conjunto de datos internos y externos respectivamente.

Portocarrero-Banda et al. [13] señalaron que las personas con discapacidad auditiva en el Perú enfrentan barreras de comunicación debido a la falta de herramientas tecnológicas que reconozcan el alfabeto dactilológico. Para abordar esta necesidad, recopilaron 3025 imágenes de señas estáticas del alfabeto con variaciones en iluminación y fondo. Seguidamente, en el pre procesamiento aplicaron técnicas como la conversión de color a YCbCr, detección de color de piel y la normalización. Posteriormente, implementaron una red neuronal convolucional (CNN) con cuatro capas de convolución y max pooling junto a una red completamente conectada, la cual fue integrada en un sistema web que permitió a los usuarios capturar imágenes en tiempo real para el reconocimiento de señas. Finalmente, el modelo alcanzó 99% de precisión en entrenamiento, 88% en validación y 84% en pruebas; lo que demostró que el sistema es robusto para reconocer los gestos en escenarios complejos.

Briones-Cerquín et al. [14] señalaron que las personas con discapacidad auditiva enfrentan obstáculos para acceder a servicios como la educación y la salud debido a la falta de herramientas tecnológicas que traduzcan el Lenguaje de Señas Peruano. Para abordar esta problemática, desarrollaron un traductor básico basado en una red neuronal convolucional (CNN). Empezaron por capturar 22,400 imágenes de señas estáticas, entre números (0 a 9) y letras del alfabeto (exceptuando J, Ñ, Z, O y W), utilizando Mediapipe y OpenCV. Posteriormente, emplearon técnicas de aumentación en el pre procesamiento para mejorar la precisión del modelo. Luego, construyeron la red de tres capas de convolución con funciones

de activación ReLu y MaxPooling para clasificar 32 clases. Los resultados que obtuvieron fueron 90% de precisión en entrenamiento, 86% en validación y 81% en pruebas.

#### Bases teóricas

Las formas de comunicación empleadas por las personas con discapacidad auditiva (DA) son fundamentales para su inclusión en los ámbitos social, educativo y laboral, promoviendo su independencia y permitiéndoles participar activamente en la sociedad.

- Audífonos: Los audífonos son aliados tecnológicos importantes para las personas con discapacidad auditiva (DA), puesto que les permiten seguir sus actividades diarias con independencia. Funcionan de la siguiente manera: utilizan un micrófono para captar los sonidos del entorno, posteriormente, estos sonidos son transmitidos a un procesador que los amplía y transforma en señales eléctricas. Por último, estas se transmiten al oído de la persona a través de un altavoz o receptor [15].
- Implantes cocleares: En Octubre del 2024, el Instituto Nacional de Salud del Niño lanzó el Programa de Tamizaje Neonatal Auditivo, con el cual se beneficiará a 30 niños mediante implantes cocleares, que serán realizados por especialistas del Servicio de Otorrinolaringología [16]. Estos implantes se colocan en las zonas afectadas del oído y activan el nervio auditivo, que, a su vez, transmite las señales generadas por el implante al cerebro, donde son interpretadas como sonidos [17].
- Lenguaje de señas: La lengua de señas se diseñó para ayudar a las personas con DA a expresar sus ideas y pensamientos mediante la gesticulación de dedos, manos y rostro [18].

De acuerdo con el Censo Nacional de Población realizado por el INEI en 2017, se reportó que 8,790 personas con DA registraron la Lengua de Señas Peruano (LSP) como su lengua materna [19]. A través de la Ley N° 29535, se otorga reconocimiento oficial a la LSP, exigiendo a las instituciones públicas y privadas garantizar el derecho de los estudiantes con DA a acceder a una educación bilingüe e intercultural que incluya la LSP, el castellano escrito, u otras lenguas originarias, según el contexto de cada estudiante [20].

El vocabulario del LSP se basa en la comunicación visual a través del cuerpo, manos y expresiones faciales; además, utiliza el espacio como un medio de comunicación (estructura fonológica) para indicar referencias temporales y transmitir significados de manera precisa y expresiva [21].

Para aprender el LSP de manera efectiva, es fundamental desarrollar la atención, percepción y memoria visual. Los ejercicios de aprestamiento permiten fortalecer diversos aspectos, tanto visuales como gestuales. En cuanto a los aspectos visuales, se trabajan la atención visual, la

discriminación y la memoria visuales. En los aspectos gestuales, se desarrollan la expresión corporal, la expresión facial y la motricidad digital y manual. Además, la Lengua de Señas posee una "fonología" compleja, analizada por la Querología, que estudia los queremos o parámetros de las señas, tales como la posición de las manos, la dirección del movimiento, el lugar de contacto, el plano y el componente manual [21].

Elementos del querema del LSP según el Ministerio de Educación del Perú:

- El lugar de articulación establece el marco espacial donde se ejecutan las señas, facilitando la percepción visual tanto para el emisor como para el receptor [21].
- La configuración de las manos varía en complejidad, desde gestos simples a combinaciones más elaboradas que añaden matices al lenguaje [21].
- El movimiento de la mano y la orientación complementan la información que se transmite, modificando el significado dependiendo de cómo se mueven o posicionan las manos [21].
- El punto de contacto y el plano señalan cómo se interactúa físicamente con el espacio corporal, añadiendo precisión a la seña [21].
- Los componentes no manuales refuerzan la intención y emoción detrás de las señas, brindando una comunicación más completa [21].

Con lo mencionado anteriormente, los elementos del querema presentan una complejidad que puede ser abordada mediante diversas herramientas de Inteligencia Artificial. En particular, el análisis de la configuración y el movimiento de las manos es crucial para comprender cómo se comunican significados y matices en este lenguaje.

Machine Learning o aprendizaje automático (ML) es una disciplina de la Inteligencia Artificial que se centra en crear algoritmos que pueden aprender de los datos y clasificar patrones. Esta tecnología resulta especialmente útil para estudiar las variaciones en la posición y el movimiento de las manos, permitiendo una interpretación más precisa y rica de las señas [22].

Existen tres enfoques principales en ML aplicables a este contexto. Por un lado, el aprendizaje supervisado, que utiliza conjuntos de datos etiquetados, permite que el sistema aprenda a predecir correctamente nuevas señas a partir de ejemplos conocidos, siendo útiles en tareas de clasificación de gestos. Por otro, el aprendizaje no supervisado trabaja con datos sin etiquetar, lo que permite al sistema identificar patrones ocultos en las señas, facilitando el reconocimiento de variantes desconocidas. Por último, el aprendizaje por refuerzo ajusta el comportamiento del sistema mediante recompensas o penalizaciones, optimizando su

capacidad para interpretar correctamente las señas a través de la interacción continua con su entorno [22].

Ahora bien, para asegurar la eficacia del algoritmo en el reconocimiento de gestos estáticos y dinámicos, es fundamental evaluar su desempeño mediante métricas de ML adecuadas. Este proceso implica analizar los errores y aplicar diversas métricas que permiten medir la efectividad del modelo. De esta manera, se pueden identificar áreas de mejora y garantizar que el algoritmo cumpla con los objetivos establecidos en la clasificación de los gestos [23].

Entre las métricas más importantes se incluyen la exactitud, que mide la proporción de gestos clasificados correctamente respecto al total; la sensibilidad, que es clave para identificar todos los casos positivos; y la puntuación F1, que combina exactitud y sensibilidad, indicando la eficacia del modelo en la identificación de gestos correctos. También se considera la especificidad, que mide la correcta clasificación de gestos negativos, y la precisión, que evalúa cuán confiables son las predicciones del modelo en relación con el total de gestos clasificados [23].

La implementación de un algoritmo de reconocimiento de gestos requiere una metodología bien definida, que proporcione un marco estructurado para guiar cada fase del proceso. Esta metodología es crucial para asegurar que se aborden de manera adecuada todos los aspectos del desarrollo, facilitando así la identificación de problemas y optimizando el rendimiento general del modelo [24].

La metodología general de Machine Learning proporciona este marco que facilita la implementación, evaluación y mejora continua del modelo. El proceso de implementación del algoritmo inicia con la definición del objetivo (Iteración 1), donde se identifica el problema a resolver. Luego, se procede a la recolección de datos (Iteración 2) para obtener la información adecuada para el entrenamiento. En la preparación de los datos (Iteración 3), se limpia y transforma la información para su análisis. Posteriormente, se selecciona el algoritmo de machine learning más adecuado (Iteración 4) y se entrena el modelo (Iteración 5) con los datos disponibles. Después, se realiza la validación del modelo (Iteración 6) para garantizar su precisión y robustez. Finalmente, se utiliza el modelo para hacer predicciones (Iteración 7), permitiendo realizar ajustes continuos que mejoren su rendimiento basado en los resultados obtenidos [24].

El procesamiento de lenguaje natural (PNL) permite a las máquinas interpretar correctamente el contexto de las oraciones. Utiliza componentes como la tokenización, el etiquetado de partes del discurso y el análisis sintáctico y semántico, siendo clave para aplicaciones como la traducción automática y la creación de resúmenes de texto. Además, el

PNL mejora la accesibilidad de la información para personas con discapacidad auditiva profunda, facilitando su acceso a contenidos a través de tecnologías de transcripción y subtítulos [25].

La metodología XP (Extreme Programming) es crucial en el desarrollo de aplicaciones móviles para personas con discapacidad auditiva profunda, ya que permite crear interfaces efectivas mediante un enfoque ágil y adaptativo. Inicia con la fase de exploración, que recopila datos sobre las necesidades de los usuarios y gestiona riesgos. Luego, en la fase de planificación, se definen historias de usuarios y se priorizan requerimientos. La fase de diseño se centra en la arquitectura del sistema y las interfaces accesibles. Posteriormente, la codificación aborda tareas específicas, y la fase de prueba asegura la calidad del sistema mediante diversas pruebas, garantizando la accesibilidad y funcionalidad de la aplicación [26].

### **Materiales y métodos**

Con lo mencionado por Macorra [27] sobre los métodos de investigación, se pudo relacionar que esta investigación es de tipo aplicada. Dicho método se basa en emplear todos los conocimientos adquiridos a través de la teoría para un propósito específico práctico.

El problema que se identificó fue las limitaciones que tienen las personas con discapacidad auditiva profunda para comunicarse, para lo cual se aplicó la teoría para desarrollar una aplicación móvil de apoyo.

De acuerdo con Zúñiga [28] en su investigación sobre los métodos en investigación científica, se logró analizar y elegir los siguientes:

*Tabla I. Métodos de Investigación*

Método	Descripción
Revisión de la literatura	Conocer que se realizó en investigaciones previas sobre la aplicación de Machine Learning para traducir el Lenguaje de Señas en texto en tiempo real
Deductivo	Establecer una solución a la dificultad que enfrentan las personas con discapacidad auditiva profunda para comunicarse
Implementación	Codificar la solución para cubrir las dificultades que enfrentan las personas con discapacidad auditiva profunda para comunicarse

Del mismo modo, se detalla las técnicas e instrumentos que se emplearon en la investigación.

*Tabla II. Técnicas de Investigación*

Técnicas	Instrumentos
----------	--------------

Entrevista y encuestas	Guía de entrevista y encuesta a un especialista en Lenguaje de Señas Peruano y la comunidad de personas con discapacidad auditiva en Chiclayo ( <b>ver Anexo 01 y Anexo 03</b> )
Evaluación de especialistas	Lista de cotejo que se aplicará a especialistas de Lenguaje de Señas Peruano para verificar la calidad del software al culminar su desarrollo ( <b>ver Anexo 02</b> )

### Población y muestra

El grupo objetivo de estudio se centró en personas con discapacidad auditiva profunda de la ciudad de Chiclayo, incluyendo tanto a jóvenes como adultos mayores que dominan el Lenguaje de Señas Peruano.

$$n = \frac{Z^2 N p q}{e^2 (N - 1) + Z^2 p q}$$

Donde:

- n= Tamaño de muestra buscado
- N= Tamaño de población (70)
- z= Nivel de confianza (95% o 1.96)
- e= Error de estimación aceptado (5% o 0.05)
- p= Probabilidad que ocurra el evento estudiado (0.5)
- q= Probabilidad de que no ocurra el evento estudiado (1-p o 0.5)

$$n = \frac{1.96^2 \times 70 \times 0.5 \times 0.5}{0.05^2 (70 - 1) + 1.96^2 \times 0.5 \times 0.5}$$

n= 59.34

### Resultados y discusión

#### Resultados

En base a los objetivos de la investigación:

#### **Identificar y cuantificar las características posturales y cinemáticas distintivas de los gestos del Lenguaje de Señas Peruano:**

De acuerdo a la guía proporcionada del Ministerio de Educación [21] y entrevistas realizadas a especialistas de Lenguaje de Señas Peruano, se detallaron los siguientes criterios por evaluar: claridad en la postura, ángulos críticos en los gestos, verificación de la simetría y estabilidad de los gestos, así como la captura precisa de los puntos clave.

Asimismo, se definió la herramienta de Machine Learning a utilizar, tomando en cuenta lo mencionado anteriormente, siendo esta Mediapipe. Esta librería permitió recolectar 21 keypoints de la mano, tanto de la derecha como de la izquierda, tomando como centro en un plano 2D la muñeca y extendiéndose hasta la punta de los dedos, lo que facilitó la captura de gestos estáticos y dinámicos.

La captura de los gestos se realizó de la siguiente manera:

- Preparación del entorno de captura: Se aseguró que la iluminación sea uniforme para una detección precisa, además se usó un fondo neutro y se posicionó a una distancia óptima de la cámara (40-60 cm).
- Activación del modo de recolección: Se activó un modo específico para la recolección de keypoints.
- Proceso de captura para cada clase: Se mantuvo el gesto estable durante la captura y posteriormente los datos se guardaron en un formato .csv.
- Verificación de la calidad de la captura: Para ello, se tomaron en cuenta lo siguiente: el dibujo claro de los landmarks (claridad en la postura), líneas claras entre landmarks para verificar ángulos (ángulos críticos), así como la posición natural y equilibrada de la mano (simetría).
- Captura de múltiples muestras: Para cada clase, se capturaron 500 muestras, la posición se alteró manteniendo el gesto base y se aplicó normalización para mantener consistencia entre estas.
- Verificación de estabilidad: Para cada gesto, se verificó que los puntos históricos sean estables y se evitó movimientos bruscos.
- Captura de variaciones: Para cada clase, se capturaron el gesto desde diferentes ángulos (+-15 grados), se alteró la escala (acercando/alejando la mano) y se mantuvo la esencia del gesto mientras se introducía variaciones.
- Validación de los datos recolectados y procesamiento final: Para cada clase, se revisó el archivo generado .csv y se preparó los datos para el entrenamiento.

```
0,0,0,0,0,-0.23270440251572327,-0.20125786163522014,-0.25157232704402516,-0.5408805031446541,-0.025157232704402517,-0.7295597484276729,0.132075
0,0,0,0,0,-0.215311004784689,-0.20574162679425836,-0.22009569377990432,-0.5406698564593302,0,0,-0.7272727272727273,0.15789473684210525,-0.89473
0,0,0,0,0,-0.22330097087378642,-0.22330097087378642,-0.21844660194174756,-0.558252427184466,0.009708737864077669,-0.7281553398058253,0.16019417
0,0,0,0,0,-0.22815533980582525,-0.19902912621359223,-0.22330097087378642,-0.5436893203883495,0.014563106796116505,-0.7233009708737864,0.1893203
0,0,0,0,0,-0.22488038277511962,-0.20574162679425836,-0.22009569377990432,-0.5454545454545454,0,0,-0.722488038277512,0.1626794258332056,-0.8851
0,0,0,0,0,-0.2222222222222222,-0.20772946859903382,-0.2222222222222222,-0.5458937198067633,0.00966183574879227,-0.7342995169082126,0.1835748792
0,0,0,0,0,-0.22596153846153846,-0.20673076923076922,-0.21634615384615385,-0.5480769230769231,0.014423076923076924,-0.7259615384615384,0.1730769
0,0,0,0,0,-0.22380952380952382,-0.20476190476190476,-0.21428571428571427,-0.5476190476190477,0.01904761904761905,-0.7333333333333333,0.19523809
```

Figura 1. Ejemplo de la captura de vectores de landmarks pertenecientes a la clase 0

**Evaluar comparativamente las herramientas de Machine Learning y Procesamiento de Lenguaje Natural para el reconocimiento y traducción en tiempo real del Lenguaje de Señas Peruano a texto español:**

Se realizó una revisión de la literatura sobre varios algoritmos de Machine Learning, considerando aspectos clave como si el método es supervisado o no supervisado y si la técnica corresponde a clasificación o regresión. Para esta investigación, se eligió el método supervisado, dado que se cuenta con un objetivo, y la técnica de clasificación, ya que el propósito es identificar 30 categorías de gestos [29].

La siguiente tabla ofrece una comparación detallada entre los algoritmos estudiados.

*Tabla III. Comparación de algoritmos [29]*

<b>Modelo</b>	<b>Data de procesamiento</b>	<b>Ejemplo de aplicación</b>
Naive Bayes	Datos categóricos	Clasificación de correos electrónicos como spam o no spam, análisis de sentimientos, diagnóstico médico
Logística	Valores numéricos	Predicción de eventos binarios
Árboles de clasificación	Datos numéricos y categóricos	Decisiones médicas, clasificación de clientes en marketing, diagnósticos en sistemas expertos
SVM	Datos numéricos	Clasificación de imágenes, reconocimiento de escritura, detección de patrones en texto
Red Neuronal Densamente Conectada	Datos numéricos	Reconocimiento de patrones, análisis de series temporales
Bagging	Datos numéricos y categóricos	Sistemas de voto en clasificación
Adaboost	Datos numéricos y categóricos	Detección de fraudes, reconocimiento facial, clasificación de textos

Posteriormente se llegó a la conclusión que el modelo secuencial de redes neuronales densamente conectadas cumple con los requisitos para el desarrollo de esta investigación. Su característica distintiva es la presencia de múltiples capas de neuronas interconectadas. Estas capas se disponen secuencialmente con cada capa conectándose a la siguiente.

En cuanto a la generación de texto, se analizaron bibliotecas pertenecientes a Machine Learning y Procesamiento de Lenguaje Natural, así como compatibles con java y Android Studio.

*Tabla IV. Bibliotecas de PNL [30]*

<b>Biblioteca</b>	<b>Objetivo</b>	<b>Técnicas</b>
-------------------	-----------------	-----------------

Apache OpenNLP	Corregir y estandarizar el texto de entrada	<ul style="list-style-type: none"> <li>• Tokens para modificar, analizar o corregir palabras específicas.</li> <li>• Reconocimiento de entidades nombradas, se encarga de detectar y categorizar elementos específicos en el texto, tales como nombres propios de individuos, ubicaciones geográficas, instituciones u organizaciones, expresiones temporales, entre otros.</li> <li>• Clasificación de sustantivo, verbo y adjetivo.</li> </ul>
Stanford CoreNLP	Transformar el texto en un formato adecuado para el análisis	<ul style="list-style-type: none"> <li>• Análisis básico hasta análisis profundo, como la comprensión semántica y la resolución de correferencia</li> </ul>

Se utilizó la herramienta de Apache OpenNLP para segmentar texto, identificar entidades (como nombres y fechas), y analizar la sintaxis, facilitando aplicaciones avanzadas de análisis de texto. Además, se destacó por ser más ligera que otras herramientas similares, lo que permite una implementación más eficiente en entornos con recursos limitados [30].

#### **Diseñar una interfaz de usuario amigable para la aplicación móvil basada en el estándar de calidad ISO/IEC 25010:2011:**

Se diseñó una interfaz basada en la usabilidad del estándar de calidad ISO/IEC 25010:2011.

Los criterios evaluados incluyeron la idoneidad, facilidad de aprendizaje, operatividad, prevención de errores de usuario, estética de la interfaz y accesibilidad. En cuanto a la estética de la interfaz, se revisó que los títulos de figuras junto con tablas fueran descriptivos además de únicos, la interfaz no estuviera sobrecargada de información, el uso de ventanas emergentes o notificaciones para destacar información relevante, así como la ubicación de elementos como botones y mensajes fuera apropiada para una apariencia visualmente atractiva. Para la facilidad de aprendizaje, se evaluó que el ingreso de datos y los resultados en el software fueran comprensibles, la navegación por las interfaces fuera intuitiva, el sistema permitiera regresar a pasos previos para ajustar datos y que se proporcionaran indicaciones claras para guiar al usuario en sus acciones [31].

A continuación, las interfaces más importantes de la aplicación móvil:

Interfaz #1: Login:

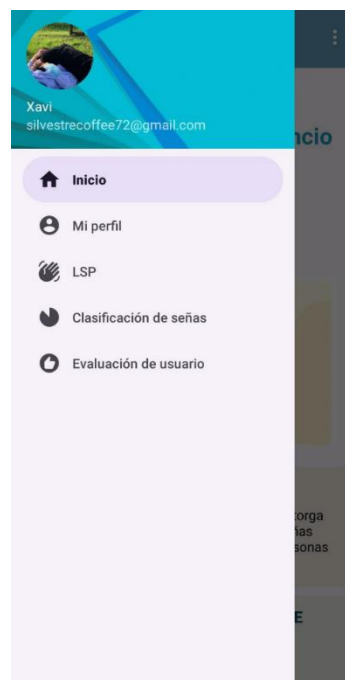
Esta interfaz consiste en el inicio de sesión de la aplicación móvil, y en la cual se validan los datos del usuario para poder acceder al menú y demás opciones.



*Figura 2. Login*

Interfaz #2 Menú:

Posteriormente del usuario iniciar sesión en la aplicación, se podrá visualizar el menú que tiene las siguientes opciones: Inicio, Mi perfil, LSP, Clasificación de señas y evaluación de usuario.



*Figura 3. Menú*

### Interfaz #3 LSP:

En esta interfaz, se tendrá que dar permisos de cámara, video y audio para poder realizar las señas en tiempo real; y consecutivamente su traducción a texto.

Además, la pantalla tiene otras opciones que son las siguientes: se puede utilizar la cámara frontal o trasera; grabar, detener y/o pausar el video; y enviar el video grabado a un contacto de WhatsApp, Instagram y/o Gmail.



*Figura 4. LSP*

### Interfaz #4 Reporte de clasificación de señas:

En este reporte se mostrará la cantidad de señas que puede traducir el modelo de Machine Learning en tiempo real en un gráfico circular.

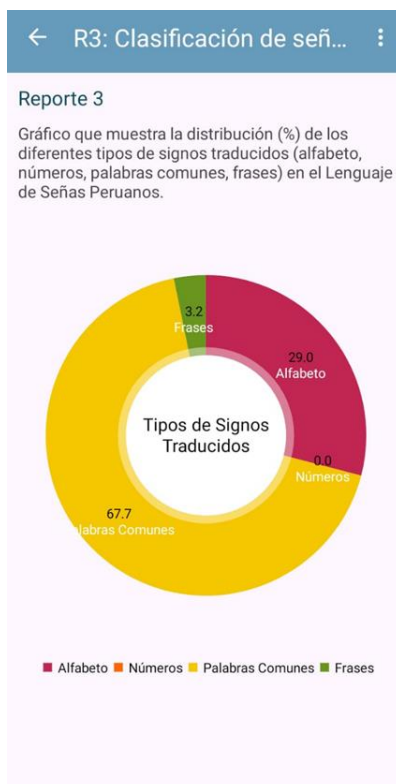


Figura 5. Formulario

**Implementar la integración de la funcionalidad de reconocimiento de gestos para asegurar su precisión y fiabilidad en el entorno de interacción del usuario:**

El modelo entrenado de reconocimiento de lenguaje de señas peruano se exportó en formato .tflite, permitiendo su ejecución eficiente en Android Studio. Además, se habilitó la configuración de la GPU/CPU, adaptándose a las características del dispositivo móvil en uso para mejorar la velocidad de ejecución en tiempo real. Por último, se integró Apache OpenNLP para convertir la salida en texto español de manera precisa.

```

Kotlin  Java
try {
    KeypointClassifier model = KeypointClassifier.newInstance(context);

    // Creates inputs for reference.
    TensorBuffer inputFeature0 = TensorBuffer.createFixedSize(new int[]{1, 42}, DataType.FLOAT32);
    inputFeature0.loadBuffer(byteBuffer);

    // Runs model inference and gets result.
    KeypointClassifier.Outputs outputs = model.process(inputFeature0);
    TensorBuffer outputFeature0 = outputs.getOutputFeature0AsTensorBuffer();

    // Releases model resources if no longer used.
    model.close();
} catch (IOException e) {
    // TODO Handle the exception
}

```

Figura 6. Modelo de clasificación de keypoints en Android Studio

Se puede observar el código para la detección de poses en imágenes. El modelo espera 42 valores de entrada (21 puntos x,y) y produce una clasificación como salida.

Asimismo, el modelo fue evaluado mediante diversas métricas de rendimiento: alcanzó una precisión del 85% aproximadamente durante el entrenamiento, mientras que en la fase de validación logró un 97%. Adicionalmente, se obtuvo una precisión específica de 0.97653 y una puntuación F1 de 0.9759, demostrando un alto nivel de efectividad en la clasificación.

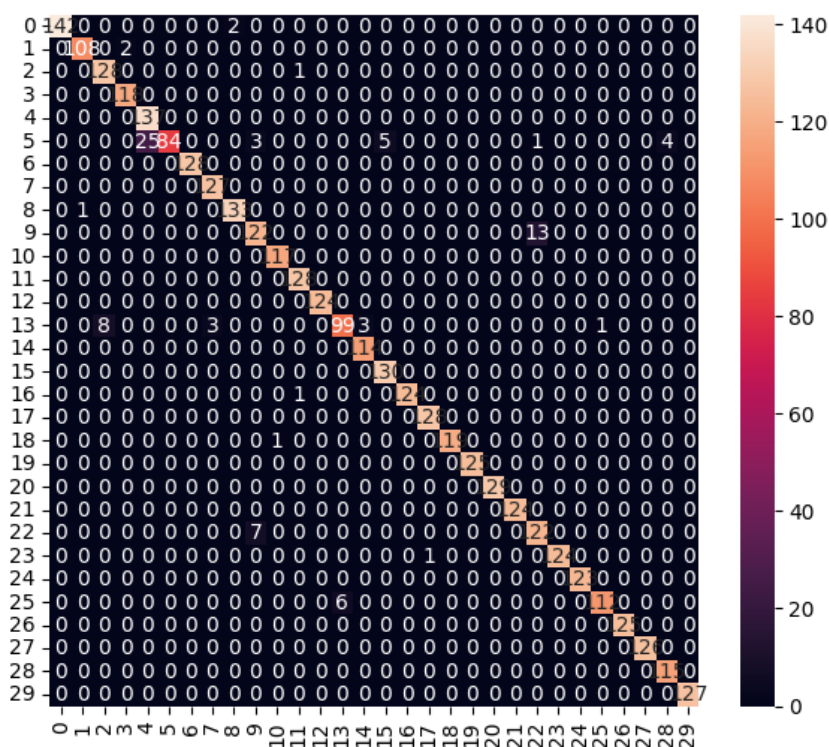


Figura 7. Matriz de confusión

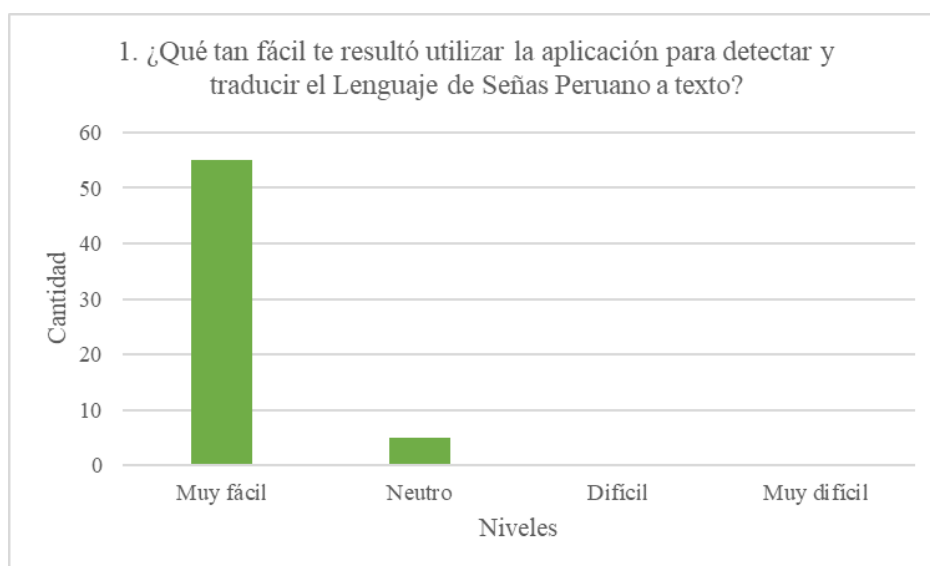
De igual manera se evaluó a través de la matriz de confusión obteniendo lo siguiente:

- De los 144 keypoints pertenecientes a la clase 0 enviadas, el algoritmo acertó 142. Se obtuvo 99% de recall y 1% de error.
- De los 110 keypoints pertenecientes a la clase 1 enviadas, el algoritmo acertó 108. Se obtuvo 98% de recall y 2% de error.
- De los 129 keypoints pertenecientes a la clase 2 enviadas, el algoritmo acertó 128. Se obtuvo 99% de recall y 1% de error.
- De los 118 keypoints pertenecientes a la clase 3 enviadas, el algoritmo acertó la mayoría de las señas. Se obtuvo un 98% de precisión y 2% de error.
- De los 122 keypoints pertenecientes a la clase 5 enviadas, el algoritmo acertó 84. Se obtuvo 82% de f1-score y 18% de error.
- De los 114 keypoints pertenecientes a la clase 13 enviadas, el algoritmo acertó 99. Se obtuvo 87% de recall y 13% de error.

### **Verificar la funcionalidad de la aplicación móvil mediante evaluación de especialistas del Lenguaje de Señas Peruano para asegurar su correcto funcionamiento:**

Se llevó a cabo una encuesta de validación de la aplicación para recopilar información valiosa sobre la experiencia de uso en esta comunidad, con el objetivo de mejorarla y adaptarla de manera más efectiva a sus necesidades. Este proceso incluyó la participación y comprometida de 60 personas con discapacidad auditiva profunda de las edades entre 25 y 50 años, reunidas en la “Iglesia Bautista de Chiclayo” y la “Casa Comunal de la Juventud” junto a la “Asociación de Sordos de Lambayeque”. La colaboración de especialistas de la I.E. Karl Weiss y la Universidad Nacional Pedro Ruiz Gallo en Lenguaje de Señas Peruano fue fundamental, ya que proporcionaron retroalimentación crucial que contribuyó significativamente a la optimización de la aplicación. La importancia de este proceso radica en su enfoque centrado en el usuario, que asegura que la aplicación no solo cumpla con las expectativas, sino que también atienda de manera efectiva las particularidades y desafíos que enfrenta esta comunidad. Al incorporar las perspectivas necesidades de los usuarios finales, se busca garantizar que la aplicación sea una herramienta accesible y útil que mejore su calidad de vida y facilite su comunicación (**ver Anexo 03**).

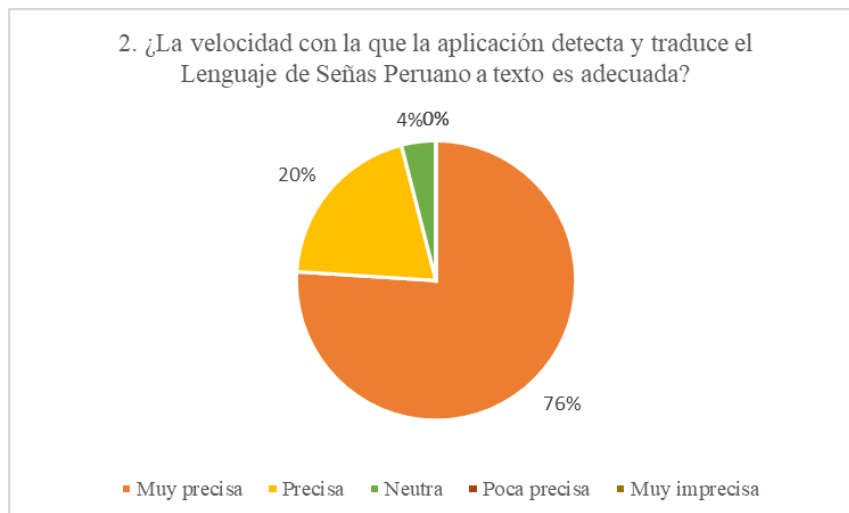
Los resultados más relevantes obtenidos a partir de la encuesta implementada fueron los siguientes:



*Figura 8. Uso de la aplicación*

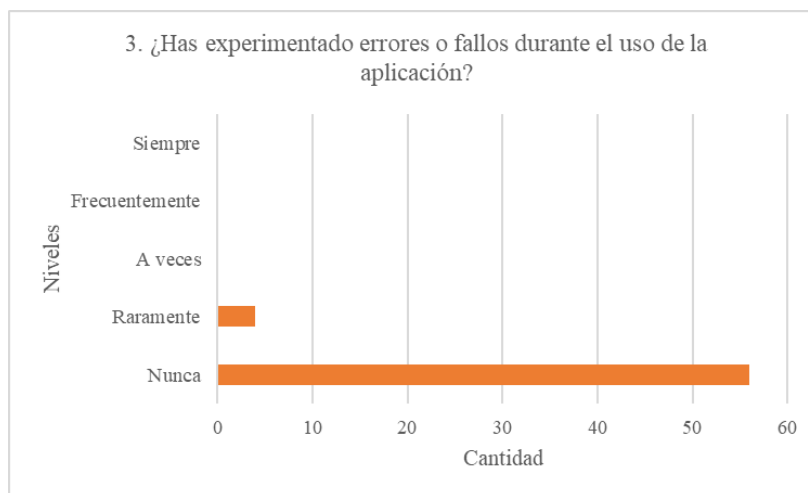
El gráfico muestra que 55 encuestados consideran que la aplicación móvil es "muy fácil" de usar, mientras que 5 la califican como "neutral". Esto sugiere que la aplicación cumple con los criterios de usabilidad de la norma ISO/IEC 25010:2011. Sin embargo, algunos usuarios

enfrentan dificultades para utilizarla, debido a la falta de familiaridad con ciertas funciones específicas.



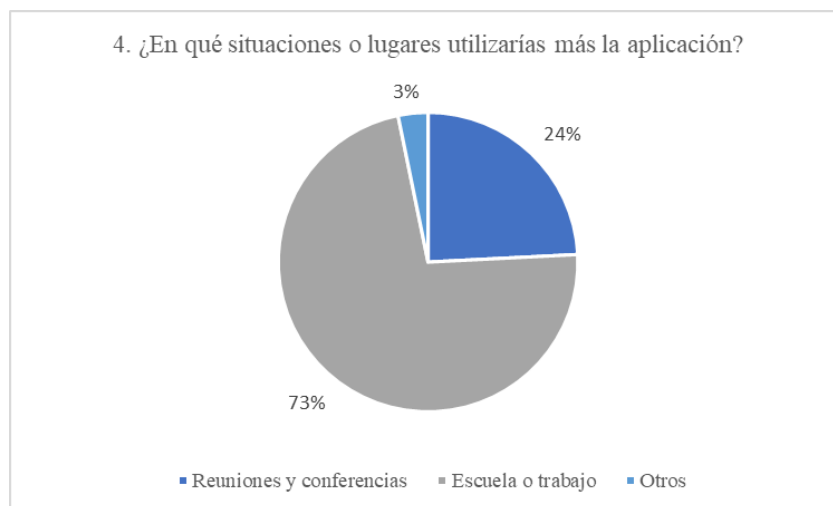
*Figura 9. Velocidad de detección y traducción*

El gráfico muestra que el 76% y el 20% de los encuestados consideran la velocidad de detección y traducción del modelo como “muy precisa” y “precisa”, respectivamente, mientras que el 4% la percibe como “neutral”, debido a que algunos dispositivos dependen del CPU en lugar de la GPU, lo que puede causar un leve retraso.



*Figura 10. Fallos de la aplicación*

El gráfico muestra que 56 de los encuestados indicaron “nunca” haber experimentado fallos, mientras que 4 indicaron “raramente”. Lo que sugiere que la aplicación responde correctamente a posibles fallos.



*Figura 11. Utilidad de la aplicación*

El gráfico indica que el 73% de los encuestados usaría la aplicación en “escuela o trabajo”, el 24% en “reuniones y conferencias” y el 3% en asuntos privados, lo cual muestra su adaptabilidad a diversos contextos.

Asimismo, se realizaron pruebas de caja negra y caja blanca. Las pruebas de caja negra se centran en evaluar cómo responde el software a diferentes entradas. En contraste, las pruebas de caja blanca examinan el código, permitiendo analizar la lógica del programa, y mejorar el rendimiento del sistema. Además, se llevaron a cabo pruebas exhaustivas de la aplicación móvil utilizando Firebase, las cuales permitieron evaluar aspectos clave como el rendimiento y la seguridad (**ver Anexo 04**).

### **Discusión**

En [9] se desarrolló un procesador específico para el idioma kazajo que abarca un analizador morfológico, sintáctico y un corrector morfológico, lo que permite un análisis profundo y personalizado del idioma. Además, implementaron un sistema de clasificación de oraciones que distingue tipos de palabras y analiza sus relaciones sintácticas. Mientras que, en esta investigación, si bien se trabajó igualmente en la corrección y estandarización de texto mediante el uso de tokens para analizar y corregir palabras, así como en la clasificación de sustantivos, verbos y adjetivos; el enfoque se realizó de manera más general y adaptable, con el fin de aplicarlo en diferentes contextos, particularmente en aquellos donde el idioma es el español.

En [10] utilizaron MediaPipe para capturar gestos complejos en coordenadas 3D de los landmarks de las manos, junto con un modelo Long Short-Term Memory (LSTM) para modelar secuencias largas y complejas de movimientos. Este sistema ofrece mayor precisión en gestos complejos y de larga duración, pero a costa de una mayor complejidad y requerimientos computacionales. En comparación con esta investigación, se empleó el reconocimiento de

gestos en 2D y se utilizó un historial de puntos para gestos dinámicos, junto con un modelo de redes neuronales densamente conectadas, las cuales son generalmente más ligeras y rápidas de ejecutar. Estas redes son eficaces para capturar patrones espaciales en los datos de entrada, lo que las hace ideales para aplicaciones en tiempo real con recursos limitados o cuando se dispone de menos datos de entrenamiento.

En [11], [13], [14] durante el pre procesamiento utilizaron Mediapipe para reconocer manos y landmarks, seguido de recorte de imágenes y pruebas en diversas condiciones de iluminación y tonos de piel. En contraste, esta investigación se centró exclusivamente en procesar las coordenadas de los landmarks, sin manipulación de imágenes. Ambos enfoques contribuyeron a lograr una precisión alta en el modelo de Machine Learning, pero el método basado en coordenadas normalizadas eliminó inherentemente variables como fondo, iluminación y tamaño, sin requerir pruebas adicionales en condiciones variadas. Esto sugiere que el modelo basado en coordenadas puede ser más eficiente en procesamiento de datos y más robusto frente a variaciones en las condiciones de captura, sin necesidad de pruebas exhaustivas en diferentes escenarios.

En [12] utilizaron una red neuronal convolucional ResNet pre entrenada para extraer características de alto nivel de imágenes, lo cual es común en el análisis de gestos continuos, incluyendo variaciones en la pose y la entonación del movimiento. Esta arquitectura tiene una notable capacidad para reconocer secuencias dinámicas y gestos complejos, aunque esto implica un mayor uso de recursos computacionales. En contraste, la investigación en cuestión opta por el uso de un sistema basado en 21 puntos clave junto a una red neuronal densamente conectada, lo que simplificó el modelo al centrarse en las características esenciales de los gestos en lugar de procesar imágenes completas. Esto facilitó el proceso de entrenamiento, reduciendo la carga computacional y disminuyendo el riesgo de sobreajuste, lo que hizo que el modelo fuera adecuado para la tarea de clasificación, y en última instancia, contribuyó a una mejora en el rendimiento del modelo.

## **Conclusiones**

Se concluye que, gracias al apoyo de especialistas en Lenguaje de Señas Peruano y al uso de un manual de guía del lenguaje, se logró identificar las características asociadas a los gestos estáticos y dinámicos, los cuales fueron: claridad en la postura, ángulos críticos en los gestos, verificación de la simetría y estabilidad de los gestos, así como captura precisa de los puntos claves. Estos criterios resultaron fundamentales para el entrenamiento del modelo y para evaluar su precisión.

Se estructuró un modelo computacional para reconocer los gestos estáticos y dinámicos del Lenguaje de Señas Peruano, compuesto por tres capas densas, lo que genera un modelo de aproximadamente 5,47 KB. Esta configuración optimiza la velocidad y la lectura de los datos, mejorando el tiempo de respuesta. Además, la generación de texto se apoyó en diversas técnicas, como el uso de tokens, reconocimiento de entidades nombradas, y la clasificación de sustantivos, verbos y adjetivos, con el fin de garantizar coherencia en el texto generado.

Se implementó el estándar de calidad ISO/IEC 25010:2011 para la interfaz de usuario, lo que permitió evaluar y mejorar atributos clave como el reconocimiento de la idoneidad, capacidad de aprendizaje, operabilidad, protección contra errores de usuario, estética de la interfaz de usuario y la accesibilidad, lo que garantizó una experiencia de usuario más intuitiva y eficiente, así como una mayor calidad en el rendimiento general de la aplicación, asegurando su adaptabilidad frente a posibles cambios o mejoras futuras.

Se integró la funcionalidad de reconocimiento de gestos a la aplicación móvil, logrando una precisión aproximada del 85% en entrenamiento y 97% en validación. Esta alta tasa de aciertos refuerza la eficacia del modelo, permitiendo una interacción más fluida y precisa con los usuarios.

Se validó la funcionalidad de la aplicación móvil en entornos reales, lo que permitió realizar ajustes para adecuarla a las necesidades de las personas con discapacidad auditiva profunda. Estos ajustes mejoraron la accesibilidad y la usabilidad, garantizando que la aplicación responda de manera más efectiva a las expectativas y requerimientos de este grupo de usuarios.

### **Recomendaciones**

Ampliar el repertorio de gestos estáticos y dinámicos que incluya números, letras del alfabeto y frases para facilitar una comunicación más amplia y fluida.

Implementar un sistema que incluya tanto el reconocimiento de gestos como el reconocimiento facial para capturar expresiones faciales, con el fin de complementar el significado del mensaje en el lenguaje de señas.

Realizar un modelo de traducción de texto a lenguaje de señas dentro de un sistema bidireccional, que permita a las personas oyentes interactuar de manera fluida y natural con usuarios del Lenguaje de Señas Peruano.

## Referencias

- [1] Gobierno del estado de México, «¿Qué es la discapacidad auditiva? | Instituto Mexiquense para la Discapacidad». Accedido: 15 de septiembre de 2024. [En línea]. Disponible en: <https://imedis.edomex.gob.mx/que-es-discapacidad-auditiva>
- [2] Organización Mundial de la Salud, «Sordera y pérdida de la audición». Accedido: 15 de septiembre de 2024. [En línea]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/deafness-and-hearing-loss>
- [3] Instituto Nacional de Estadística, «Encuesta de Discapacidad, Autonomía personal y Situaciones de Dependencia. Población residente en centros (EDAD centros)», INE. Accedido: 21 de septiembre de 2024. [En línea]. Disponible en: [https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736176782&menu=ultiDatos&idp=1254735573175](https://ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176782&menu=ultiDatos&idp=1254735573175)
- [4] Organización Panamericana de la Salud, «Salud auditiva - OPS/OMS». Accedido: 15 de septiembre de 2024. [En línea]. Disponible en: <https://www.paho.org/es/temas/salud-auditiva>
- [5] Consejo Nacional para la Integración de la Persona con Discapacidad, «14 jóvenes con discapacidad auditiva de diversas regiones del país se graduaron en el aprendizaje de la lengua de señas peruana». Accedido: 25 de mayo de 2023. [En línea]. Disponible en: <https://www.gob.pe/institucion/conadis/noticias/691684-14-jovenes-con-discapacidad-auditiva-de-diversas-regiones-del-pais-se-graduaron-en-el-aprendizaje-de-la-lengua-de-senas-peruana>
- [6] R. Nexos, «Manos que hablan: Los retos de la lengua de señas en el Perú», Nexos. Accedido: 28 de septiembre de 2024. [En línea]. Disponible en: <https://nexus.ulima.edu.pe/2024/09/23/manos-que-hablan-los-retos-del-lenguaje-de-senas-en-el-peru/>
- [7] E. Agudo Narvi3n, R. L3pez Huamani, y L. Delgado Guerrero, «D3ficit auditivo o p3rdida de audici3n», Ocronos - Editorial Cient3fico-T3cnica. Accedido: 15 de septiembre de 2024. [En l3nea]. Disponible en: <https://revistamedica.com/guia-deficit-auditivo-perdida-audicion/>
- [8] C. Cedeno Nadal, «Impacto de la inclusi3n social en la calidad de vida de adolescentes con discapacidad auditiva, una revisi3n sistem3tica», 2023. doi: 10.13140/RG.2.2.30355.04646.
- [9] N. Amangeldy et al., «Continuous Sign Language Recognition and Its Translation into Intonation-Colored Speech», *Sensors*, vol. 23, n.º 14, 2023, doi: 10.3390/s23146383.

- [10] R. Sreemathy et al., «Sign language recognition using artificial intelligence», *Educ. Inf. Technol.*, vol. 28, n.º 5, pp. 5259-5278, 2023, doi: 10.1007/s10639-022-11391-z.
- [11] E. J. M. Cataño et al., «Detección del abecedario de Lengua de Señas Mexicanas (LSM) usando MediaPipe, SVM y Random Forest», *Transregiones*, n.º 7, Art. n.º 7, feb. 2024.
- [12] G. Bejarano et al., «PeruSIL: A Framework to Build a Continuous Peruvian Sign Language Interpretation Dataset», *Proc. 10th Workshop Represent. Process. Sign Lang.*, pp. 1-8, jun. 2022.
- [13] G. Portocarrero-Banda et al., «Static Peruvian Sign Language Classifier Based on Manual Spelling Using a Convolutional Neural Network», *JINIS 2023 XXX Int. Conf. Syst. Eng.*, vol. 3693, pp. 69-76, oct. 2023.
- [14] A. Briones-Cerquín et al., «Peruvian Sign Language translator for people with hearing and/or communication disabilities using a convolutional neural network», *21st LACCEI Int. Multi-Conf. Eng. Educ. Technol. Leadersh. Educ. Innov. Eng. Framew. Glob. Transform. Integr. Alliances Integral Dev.*, jul. 2023, doi: <https://dx.doi.org/10.18687/LACCEI2023.1.1.1403>.
- [15] Consejo Nacional para la Integración de la Persona con Discapacidad, «Un gran aliado: ¿Sabes cómo funcionan los audífonos para personas con discapacidad auditiva?» Accedido: 1 de octubre de 2023. [En línea]. Disponible en: <https://www.gob.pe/institucion/conadis/noticias/719492-un-gran-aliado-sabes-como-funcionan-los-audifonos-para-personas-con-discapacidad-auditiva>
- [16] Instituto Nacional de Salud del Niño de Breña, «30 niños del INSN oirán por primera vez con la colocación de implantes cocleares». Accedido: 2 de noviembre de 2024. [En línea]. Disponible en: <https://www.gob.pe/institucion/insn/noticias/1037971-30-ninos-del-insn-oiran-por-primera-vez-con-la-colocacion-de-implantes-cocleares>
- [17] National Institute on Deafness and Other Communication Disorders, «Implantes cocleares», *Cent. Inf. NIDCD*, mar. 2021.
- [18] Consejo Nacional para la Integración de la Persona con Discapacidad, «Semana Internacional de la Persona Sorda: Mitos y verdades sobre la comunidad sorda que se debe conocer». Accedido: 1 de octubre de 2023. [En línea]. Disponible en: <https://www.gob.pe/institucion/conadis/noticias/837340-semana-internacional-de-la-persona-sorda-mitos-y-verdades-sobre-la-comunidad-sorda-que-se-debe-conocer>
- [19] Defensoría del Pueblo, «Defensoría del Pueblo: debe facilitarse el aprendizaje de la lengua de señas peruana y promover la identidad lingüística y cultural de las personas sordas». Accedido: 2 de noviembre de 2024. [En línea]. Disponible en:

<https://www.defensoria.gob.pe/defensoria-del-pueblo-debe-facilitarse-el-aprendizaje-de-la-lengua-de-senas-peruana-y-promover-la-identidad-linguistica-y-cultural-de-las-personas-sordas/>

- [20] Consejo Nacional para la Integración de la Persona con Discapacidad, «Conadis: Universidades e institutos deben contar con intérpretes de lengua de señas si brindan servicio a personas sordas». Accedido: 1 de octubre de 2023. [En línea]. Disponible en: <https://www.gob.pe/institucion/conadis/noticias/842142-conadis-universidades-e-institutos-deben-contar-con-interpretes-de-lengua-de-senas-si-brindan-servicio-a-personas-sordas>
- [21] Ministerio de Educación, «Lengua de señas peruana: guía para el aprendizaje de la lengua de señas peruana, vocabulario básico», *Minist. Educ.*, 2015, Accedido: 2 de noviembre de 2024. [En línea]. Disponible en: <https://repositorio.minedu.gob.pe/handle/20.500.12799/5545>
- [22] J. Sanchez Galan F, *Machine Learning y sus aplicaciones*. 2021, p. 90. doi: 10.47300/978-9962-5599-8-6-03.
- [23] S. kumar et al., «Malware Classification Using Machine Learning Models», *Procedia Comput. Sci.*, vol. 235, pp. 1419-1428, ene. 2024, doi: 10.1016/j.procs.2024.04.133.
- [24] M. Mandolini et al., «A cost modelling methodology based on machine learning for engineered-to-order products», *Eng. Appl. Artif. Intell.*, vol. 136, p. 108957, oct. 2024, doi: 10.1016/j.engappai.2024.108957.
- [25] J. Márquez Díaz, «Modelos de lenguaje natural en la investigación científica una descripción técnica», vol. 33, pp. 1-10, jul. 2023.
- [26] R. Mersita y D. Darwis, «Sistem Informasi Pembayaran SPP pada Sekolah di Kecamatan Gedung Tataan dengan Metode Extreme Programming», *J. Ilm. Sist. Inf. Akunt.*, vol. 2, pp. 45-53, nov. 2022, doi: 10.33365/jimasia.v2i2.1872.
- [27] J. C. D. L. Macorra García, *Manual de metodología de la Investigación. Curso 2023-24*. 2023. Accedido: 30 de septiembre de 2024. [En línea]. Disponible en: <https://hdl.handle.net/20.500.14352/91690>
- [28] P. Zúñiga, R. Cedeño, y I. Palacios, «Metodología de la investigación científica: guía práctica», *Cienc. Lat. Rev. Científica Multidiscip.*, vol. 7, pp. 9723-9762, sep. 2023, doi: 10.37811/cl\_rcm.v7i4.7658.
- [29] A. Singla et al., «Exploration of machine learning approaches for automated crop disease detection», *Curr. Plant Biol.*, vol. 40, p. 100382, dic. 2024, doi: 10.1016/j.cpb.2024.100382.

- [30] K. A. Shastry y A. Shastry, «An integrated deep learning and natural language processing approach for continuous remote monitoring in digital health», *Decis. Anal. J.*, vol. 8, p. 100301, sep. 2023, doi: 10.1016/j.dajour.2023.100301.
- [31] A. Moreira et al., «A social and technical sustainability requirements catalogue», *Data Knowl. Eng.*, vol. 143, p. 102107, ene. 2023, doi: 10.1016/j.datak.2022.102107.

**Anexos****ANEXO N° 01. GUÍA DE ENTREVISTA****Fecha:** 11/10/2023**Ciclo:** 2023-II**Responsable de las preguntas:** López Romero Maricarmen**Objetivo:**

- Identificar los problemas relacionados con la discapacidad auditiva y su influencia en el ámbito social, laboral y educativo.

**Persona entrevistada:**

- **Nombres y Apellidos:** Edgar Espejo Muñoz
- **Cargo:** Docente y especialista en Lenguaje de Señas Peruano en la I.E. Karl Weiss.
- **Tipo de Entrevista:** Presencial
- **Herramientas digitales por emplear:** Grabadora de audio

**Contenido:****1. ¿Cuál es el proceso de inclusión que deben seguir las instituciones y/o empresas para atender y apoyar a las personas con discapacidad auditiva?**

Las instituciones y empresas deben estar pendientes de la inclusión para que estas personas con diferentes habilidades no sean aisladas; asimismo, para que en un futuro puedan desenvolverse ya sea en el plano social, laboral, educativo y cultural.

**2. En el ámbito educativo, ¿qué necesidades identifica usted en la enseñanza para los estudiantes con discapacidad auditiva?**

Las necesidades que identifico es que algunos estudiantes requieren implantes, pero no pueden ser costeados debido a la situación económica de sus padres. Los estudiantes con discapacidad auditiva enfrentan dificultades en su proceso de aprendizaje y desarrollo de habilidades, por lo tanto, es importante el uso de audífonos para facilitar su acceso y comprensión. Además, es fundamental que el Gobierno brinde apoyo para que estos estudiantes puedan contar con dispositivos de asistencia.

**3. ¿Qué actividades realiza la institución para cubrir estas necesidades?**

Lo que llevo a cabo son adaptaciones curriculares basadas en las necesidades individuales de los estudiantes, con el objetivo de crear un entorno motivador que promueva la enseñanza del Lenguaje de Señas Peruano (LSP) en las aulas de cuarto año de secundaria A, B y C. Esto facilita la interacción entre los estudiantes oyentes y aquellos con

discapacidad auditiva.

**4. ¿Cómo se sienten los estudiantes con discapacidad auditiva con relación a la tecnología que utiliza la institución durante las clases virtuales?**

Durante la pandemia COVID-19 muchos de los estudiantes con discapacidad auditiva no entraban a clases virtuales, puesto que la institución no cuenta con un intérprete de Lenguaje de Señas Peruano (LSP) para asistirlos durante la enseñanza académica, a esto se suma que solo el 10% de los profesores conocen y practican algunas palabras básicas en LSP. Por ello, solo se comunicaban por los canales de mensajería para recibir las tareas y materiales educativos. En el caso de los estudiantes con discapacidad auditiva que asistieron a clases, necesitaban la compañía de un familiar para que estos pudieran ayudarles a comprender las sesiones.

**5. En el ámbito social, ¿de qué manera las barreras sociales afectan la participación y la integración de las personas con discapacidad auditiva en reuniones remotas?**

Tengo un familiar que tiene discapacidad auditiva y somos unos pocos que conocemos y practicamos el Lenguaje de Señas Peruano (LSP); esto es porque hay muy poca empatía. Las barreras sociales se deben a una falta de conciencia sobre las necesidades de comunicación de las personas con discapacidad auditiva, lo que resulta en una comunicación limitada. Esto puede llevar a situaciones de exclusión en conversaciones y actividades sociales.

**6. Usted tiene contacto diario con personas que tienen esta discapacidad, ¿cómo perciben la actitud de la sociedad hacia ellas en situaciones sociales?**

Pueden sentirse aisladas, mal comprendidas; a esto se suma que hay muy poco interés en aprender el Lenguaje de Señas Peruano y no es un curso obligatorio en muchos colegios y/o universidades, lo cuál sería indispensable.

**7. En el ámbito laboral, ¿cuáles son las barreras que afrontan las personas con discapacidad auditiva en las empresas?**

La falta de adaptaciones, como intérpretes de LSP, la necesidad de una asistencia auditiva en las entrevistas laborales remotas puede limitar el acceso a información clave y oportunidades de participación. Asimismo, el estigma que hay en nuestra sociedad asociado con la discapacidad auditiva afecta a la autoestima y la confianza en el entorno laboral, y esto puede influir en las oportunidades de empleo y el avance profesional.

**8. Según su opinión, ¿una aplicación móvil que permita la traducción del Lenguaje de Señas Peruano a texto facilitará la comunicación de personas con discapacidad**

**auditiva?**

Si, fuera esencial un aplicativo para que pueda, por ejemplo, usarse durante las clases virtuales de manera autónoma y, además, los estudiantes con discapacidad auditiva puedan interactuar con sus compañeros.

**9. ¿Cómo usted cree que debería funcionar esta aplicación móvil?**

Debe facilitar la comunicación, ya que necesitan las personas con discapacidad auditiva en el futuro desarrollar habilidades sociales y culturales, ya que no son incluidos en muchas ocasiones por las empresas, centros, entre otros.

**ANEXO N° 02. LISTA DE COTEJO PARA EVALUAR LA USABILIDAD DEL  
SOFTWARE BAJO LA NORMA ISO/IEC 25010:2011**

**ANEXO N° 04. LISTA DE COTEJO PARA EVALUAR LA USABILIDAD DEL  
SOFTWARE BAJO LA NORMA ISO/IEC 25010:2011**

Fecha: 24/06/24

Ciclo: 2024 - I

**1. Datos personales:**

- a. Nombre y Apellidos: *Edgar Robert Espejo Muñoz*  
 b. Cargo: *Coordinador Pedagógico*

**2. Introducción:**

El estándar ISO/IEC 25010:2011 aborda la usabilidad como un aspecto importante relacionado con el diseño, evaluación y experiencia de la interacción del usuario con el software. Asimismo, proporciona criterios que se deben tener en cuenta al momento de evaluar y medir el nivel de calidad del sistema en términos de usabilidad.

De acuerdo con lo anteriormente mencionado, se creó la siguiente lista de cotejo en la que se establecieron ciertos criterios de medición para evaluar el software.

**3. Recomendaciones previas:**

- Estudiar el software para adquirir conocimiento sobre su funcionamiento.
- Comprender y analizar los criterios de usabilidad utilizados en la prueba.

**4. Objetivo:**

- Evaluar el software en base a los criterios de usabilidad del estándar de calidad ISO/IEC 25010:2011.

**5. Evaluación:**

	No aplica (NA)	Muy en desacuerdo (MED)	En desacuerdo (ED)	De acuerdo (DA)	Muy de acuerdo (MDA)
Puntaje	0	1	2	3	4

CRITERIOS DE USABILIDAD	NA	MED	ED	DA	MDA
<b>RECONOCIMIENTO DE LA IDONEIDAD</b>					
El software es relevante para las necesidades y objetivos de los usuarios en el contexto de uso					X
El software permite a los usuarios realizar sus tareas de manera eficiente y sin obstáculos					X
El software se integra de manera efectiva con otros sistemas o servicios que los usuarios puedan necesitar utilizar					X
<b>CAPACIDAD DE APRENDIZAJE</b>					
El ingreso de datos y los resultados mostrados en el software son fáciles de entender					X
El usuario puede explorar con facilidad el software sin tener ningún inconveniente					X
El software permite volver a pasos anteriores para modificar los datos previamente ingresados					X
El software brinda indicaciones que permite al usuario completar acciones que se consideran complejas					X
El usuario puede identificar fácilmente figuras, tablas, elementos (clic) y el tipo de acción que debe ejecutar cada elemento del software					X
<b>OPERABILIDAD</b>					
El software responde de manera eficiente y no se vuelve lento o inmanejable con el tiempo					X
El software es estable y no se bloquea con frecuencia o experimenta errores críticos					X
El software se actualiza regularmente y es fácil de mantener					X
El software es escalable para manejar un aumento en la carga de trabajo o el número de usuarios					X
El software está disponible y accesible cuando los usuarios lo necesitan				X	
El software es compatible con diferentes dispositivos Android				X	
<b>PROTECCIÓN CONTRA ERRORES DE USUARIO</b>					
El software ayuda a los usuarios a evitar errores					X
El software informa adecuadamente sobre los errores que podrían presentarse en el proceso					X
<b>ESTÉTICA DE LA INTERFAZ DE USUARIO</b>					
Los títulos de las figuras y tablas son integralmente descriptivos y distintos					X
El software no tiene excesiva información dentro de su interfaz					X
El software utiliza ventanas emergente o notificaciones para mostrar información importante					X
La posición de los elementos como son: botones, mensajes, información, etc. son los adecuados					X
La apariencia del software es visualmente agradable y sencilla				X	
<b>ACCESIBILIDAD</b>					
Se ofrece contenido alternativo, como descripciones de imágenes, subtítulos y versiones accesibles, para permitir a usuarios con discapacidades acceder al contenido					X
La estructura y navegación del software son intuitivas y sencillas de entender, lo que beneficia a usuarios con discapacidades cognitivas					X
El software ha sido sometido a pruebas de accesibilidad con usuarios reales con discapacidades					X
Se proporciona soporte y recursos de accesibilidad, como tutoriales o documentación específica					X

EDGAR BARRERA MUÑOZ  
DNI: 27283017

24 Junio de 2024.

## ANEXO N° 03. ENCUESTA DE VALIDACIÓN DE LA APLICACIÓN MÓVIL

### Encuesta de validación de la aplicación basada en machine learning y procesamiento de lenguaje natural para facilitar la comunicación en personas con discapacidad auditiva en el Perú

#### Datos del encuestado

Nombre y apellidos:

Edad:

Ocupación:

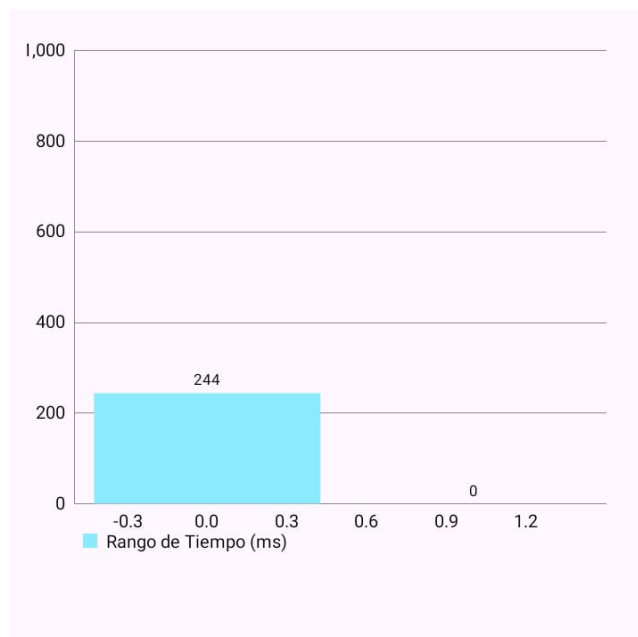
**OBJETIVO:** Esta encuesta busca recopilar información valiosa sobre la experiencia de las personas con discapacidad auditiva al utilizar la aplicación, con el fin de mejorarla y adaptarla mejor a sus necesidades.

**INSTRUCCIÓN:** Marca con "X" en las casillas según corresponda.

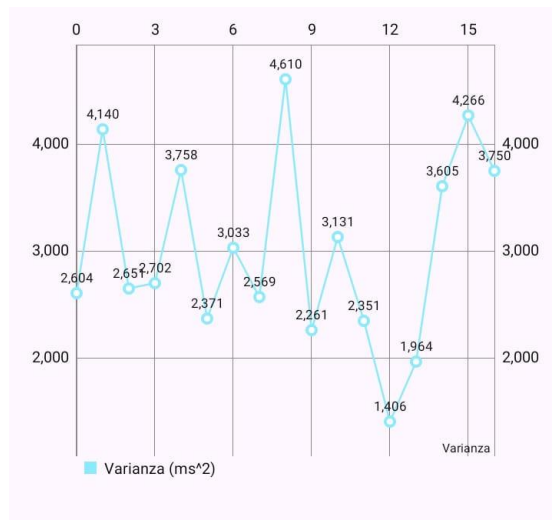
1. ¿Qué tan fácil te resultó utilizar la aplicación para detectar y traducir el Lenguaje de Señas Peruano a texto?
  - Muy fácil
  - Neutro
  - Difícil
  - Muy difícil
2. ¿La velocidad con la que la aplicación detecta y traduce el Lenguaje de Señas Peruano a texto es adecuada?
  - Muy precisa
  - Precisa
  - Neutra
  - Poca precisa
  - Muy imprecisa
3. ¿Has experimentado errores o fallos durante el uso de la aplicación?
  - Nunca
  - Raramente
  - A veces
  - Frecuentemente
  - Siempre
4. ¿En qué situaciones o lugares utilizarías más la aplicación? (Puedes seleccionar más de una opción)
  - Reuniones y conferencias
  - Escuela o trabajo
  - Otros (especifica)
5. En una escala del 1 al 10, siendo el 10 el más alto ¿cómo calificarías tu satisfacción general con la aplicación?
 

1    2    3    4    5    6    7    8    9    10
6. ¿Hay alguna función o mejora que te gustaría ver en futuras versiones de la aplicación? (Responde en texto abierto)

**ANEXO N° 04. PRUEBAS DEL PRODUCTO ACREDITABLE**



*Figura 12. Rango de tiempo (mínimo – máximo ms) del modelo al reconocer una seña en tiempo real*



*Figura 13. Varianza en el tiempo de predicción global (ms<sup>2</sup>)*



Figura 14. Pruebas de rendimiento

El acceso con redireccionamiento de origen cruzado en Google Chrome M115 y versiones posteriores ya no es compatible y dejará de funcionar el 24 de junio de 2024.

Identificador	Proveedores	Fecha de creación	Fecha de acceso	UID de usuario
lopez_romero_m@hotmail...		5 ago 2024	5 ago 2024	vLm0Vg81h7dQWnJlQmIbb9...
monicaltam21@gmail...		25 jul 2024	25 jul 2024	ACyW8hSicaVLIshrcRTf4RyyE...
silvestrecoffe72@gmail...		24 jul 2024	30 sept 2024	hRzmiz9LWygAh1pFah470KA...
lopezromeromarcarme...		24 jul 2024	30 sept 2024	PrVgAha8GmchYqKm49OSRz...

Filas por página: 50 1 - 4 of 4

Figura 15. Pruebas de seguridad