

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



Agente de reconocimiento de lengua de señas con archivos multimedia y plataforma para su aprobación y distribución

Trabajo de graduación presentado por Brandon Josué Hernández Marroquín para optar al grado académico de Licenciado en Ingeniería en Ciencias de la Computación y Tecnologías de la Información

Guatemala,

2023

UNIVERSIDAD DEL VALLE DE GUATEMALA
Facultad de Ingeniería



Agente de reconocimiento de lengua de señas con archivos multimedia y plataforma para su aprobación y distribución

Trabajo de graduación presentado por Brandon Josué Hernández Marroquín para optar al grado académico de Licenciado en Ingeniería en Ciencias de la Computación y Tecnologías de la Información

Guatemala,

2023

Vo.Bo.:



(f) _____
M.Ed. Carmen Lucía del Pilar Guerrero Abascal de Prado

Tribunal Examinador:



(f) _____
M.Ed. Carmen Lucía del Pilar Guerrero Abascal de Prado



(f) _____
MSc. Douglas Barrios



(f) _____
Ing. Gabriel Brolo Tobar

Fecha de aprobación: Guatemala, 05 de diciembre de 2023.

Lista de figuras	VII
Resumen	IX
Abstract	XI
1. Introducción	1
2. Antecedentes	3
3. Justificación	5
4. Objetivos	7
4.1. Objetivo general	7
4.2. Objetivos específicos	7
5. Alcance	9
6. Marco teórico	11
6.1. Lengua de señas	11
6.1.1. Tipos de lengua de señas	11
6.1.2. Lengua de señas en Guatemala	12
6.2. Machine Learning	12
6.2.1. Tipos de modelos	12
6.3. Computación en la nube	12
6.3.1. Modelos de servicio	14
6.3.2. Google MediaPipe sobre Hand Landmarker	14
6.4. Red de Entrega de Contenido (CDN)	15
6.4.1. ¿Es una CDN igual a un servidor de sitios web?	15
6.4.2. ¿Cuáles son los beneficios de utilizar una CDN?	15
6.4.3. ¿Cómo funciona una CDN?	16
6.5. Aplicaciones	16
6.5.1. Frontend	16

6.5.2.	Backend	17
6.5.3.	Interfaz de Programación de Aplicaciones (API)	17
6.6.	Servicios de la nube	18
6.6.1.	Firebase Authentication	18
6.6.2.	Compute Engine (GCE)	19
6.6.3.	Cloud Run (GCR)	19
6.6.4.	Cloud Functions (GCF)	19
6.6.5.	Cloud Storage (GCS)	20
6.6.6.	MongoDB	20
7.	Metodología	21
7.1.	Modelo de Machine Learning	21
7.2.	Desarrollo de aplicación	22
7.2.1.	Validación de conocimientos	26
7.2.2.	Flujos de la aplicación	26
7.3.	Recolección de resultados	28
8.	Resultados	29
8.1.	Predicción de recursos en LENSEGUA	29
8.2.	Predicciones de recursos en ASL	30
8.3.	Aumentar el rango	31
8.4.	Validación de contenido	33
8.5.	Observaciones	34
9.	Análisis de resultados	35
10.	Conclusiones	37
11.	Recomendaciones	39
12.	Bibliografía	41

Lista de figuras

1.	Comparación de servicios en la nube	13
2.	Puntos de referencia de la mano	15
3.	Flujo de un API	18
4.	Flujo 1 de Figma	22
5.	Flujo 2 de Figma	22
6.	Flujo 3 de Figma	23
7.	Flujo 4 de Figma	23
8.	Pantalla de inicio de sesión	23
9.	Pantalla de creación de cuenta	24
10.	Pantalla principal	24
11.	Pantalla de perfil	25
12.	Pantalla de deletreo	25
13.	Flujo predicción de recursos	27
14.	Predicción de la palabra hola en deletreo	29
15.	Predicción de la palabra pescado en video	30
16.	Predicción de la palabra inhalador en video	30
17.	Predicción de la palabra see you later en video	31
18.	Aumento de rango de los usuarios que hicieron la prueba	31
19.	Métricas para avanzar en el rango	32
20.	Aumentar rango con base en el tiempo	32
21.	Aumentar rango con base con cantidad de recursos aceptados	33
22.	Ver recursos en la pantalla de inicio y dar feedback	33
23.	Ver recursos propios y observar estado	34

En Guatemala aproximadamente viven 250 mil personas que utilizan la lengua de señas para comunicarse, a pesar de ello la comunidad no cuenta con un reglamento oficial. Dificultando la comunicación entre la comunidad sorda por todas las variaciones que existen de la lengua. Se creó una aplicación para centralizar el conocimiento de una lengua de señas de una variación. Utilizando modelos de Machine Learning se creó un sistema inteligente que predice palabras, la aplicación móvil se trabajó en React Native, mientras que el backend en NestJS alojado en Cloud Run que es un servicio de Google Cloud Platform. Luego de la implementación quedó claro que se desarrolló una aplicación móvil para la distribución de contenido multimedia, predecir el mensaje con base en la lengua de señas seleccionada por el usuario y verificar la calidad con base a los votos de la misma comunidad.

In Guatemala there are approximately 250 thousand people who use sign language to communicate, despite this the community does not have official regulations. Making communication difficult among the deaf community due to all the variations that exist in the language. An application was created to centralize knowledge of a sign language of one variation. Using Machine Learning models, an intelligent system was created that predicts words, the mobile application was built in React Native, while the backend was built in NestJS hosted on Cloud Run, which is a Google Cloud Platform service. After the implementation, it was clear that a mobile application was developed for the distribution of multimedia content, predicting the message based on the sign language selected by the user and verifying the quality based on the votes of the same community.

CAPÍTULO 1

Introducción

En Guatemala viven unas 250 mil personas con algún grado de discapacidad auditiva y no se cuenta con un reglamento oficial. [1] La lengua de señas no es una lengua universal; teniendo alrededor de 300 lenguas diferentes. Lo que resulta difícil para una persona aprender sobre un mismo idioma.

Con esto en mente, se realizó una plataforma en donde las personas puedan grabarse haciendo alguna oración o palabra, compartiéndolas en la comunidad. En dónde la misma aprobara o rechazará este archivo, como resultado la persona podrá leer porque sus señas no fueron efectivas al tratar de dar el mensaje. Todo esto clasificado dentro de la misma lengua, dando por resultado que, aunque la expresión este correcta, esta sea rechazada si no es del mismo idioma. Para finalizar, se trabajó con un modelo de Machine Learning para identificar lo que está queriendo decir el usuario con la lengua seleccionada.

La lengua de señas es un medio de comunicación utilizado por personas sordas. Al compararla con el lenguaje hablado, es innegable que la lengua de señas es empleada por una minoría, lo que crea una brecha entre aquellos que no dominan ambas formas de comunicación. Debido a esto, se han desarrollado diferentes proyectos para reducir esta distancia.

Lenovo ha creado una inteligencia artificial que traduce en tiempo real la lengua de señas brasileña a voz, con el propósito de facilitar las interacciones en situaciones donde no esté disponible un intérprete. Esta tecnología ha estado en desarrollo durante cuatro años y ha sido entrenado utilizando más de mil horas de videos. Considera diversos gestos faciales y corporales, como los hombros, codos y muñecas, así como un mapeo detallado de la mano. [2]

Por otro lado, existen diversas aplicaciones dedicadas a la enseñanza de lenguas de señas, una de ellas (y la más popular) es 'ASL Pocket Sign', que funciona de manera similar a Duolingo. Esta aplicación ofrece lecciones divididas en diferentes categorías, permitiendo a los usuarios aprender la lengua de señas americana (ASL) y contar con una fuente confiable sobre el idioma. [3] Además, está 'Kitsord', una aplicación móvil con una dinámica similar a la de Duolingo, diseñada para reducir la brecha entre las comunidades sordas y oyentes. A diferencia de 'ASL Pocket Sign', 'Kitsord' ofrece varias lenguas de señas, incluyendo Lensegua, Lengua de signos americana (ASL) y Lengua de signos española (LSE). [4]

En Guatemala aproximadamente viven 250 mil personas que utilizan la lengua de señas para comunicarse, a pesar de ello la comunidad sorda en Guatemala no cuenta con un reglamento oficial. [1] Como resultado, se cuenta con diferentes señas para una misma idea o palabra. El Congreso de la República aprobó el Decreto 3-2020 reconociendo a la lengua de señas de Guatemala y aprobó oficialmente la lengua de señas como el medio de comunicación integrado por el conjunto de gestos, formas y movimientos corporales característicos con gramática propia de las personas sordas. Con el fin de unificar y estandarizar la lengua de señas que se utiliza en Guatemala y aprovechando el decreto 03-2020, se ha creado una mesa técnica de Lensegua conformada por distintas organizaciones, principalmente organizaciones de sociedad civil de la comunidad de sordos.

La lengua de señas no es una lengua universal, se puede relacionar como los idiomas que cuenta con el español e inglés. La Federación Mundial de Personas Sordas señala que en el mundo, existen más de 300 lenguas de señas. Por ende, una persona tendrá dificultades en aprender lengua de señas, dado que puede encontrar traducciones en otro idioma y por lo tanto, su comunicación con el resto no sería efectiva.

Con base en lo anteriormente expuesto, al finalizar el proyecto se contará con una aplicación donde centralice la lengua de señas. Una de las finalidades de la aplicación es brindar la posibilidad que los propios usuarios puedan seguir alimentando esta base de datos y de esta manera el conocimiento aumente. Así mismo brinda la oportunidad que dos personas sordas y oyentes puedan comunicarse al consultar en la aplicación lo que desean expresar de tal forma que interactúen mutuamente.

4.1. Objetivo general

Desarrollar una aplicación móvil para la distribución de contenido multimedia, predecir el mensaje con base a la lengua de señas seleccionada por el usuario y verificar la calidad.

4.2. Objetivos específicos

- Implementar dentro de la aplicación dos lenguas de señas diferentes.
- Diseñar un proceso de verificación de recursos en el que los usuarios puedan emitir un voto para aceptar o rechazar las señas registradas.

Este proyecto incluirá la creación del Frontend, Backend y la arquitectura necesaria. En un primer paso, se configurará toda la infraestructura requerida en Google Cloud Platform (GCP) para alojar el backend y el modelo, y se desarrollará el backend para respaldar las funciones de la aplicación móvil.

Es importante señalar que, aunque se empleará un modelo de predicción de palabras o frases en archivos multimedia, no se incluirá ningún modelo específico. La precisión y cantidad de predicciones de palabras o frases dependerá de la calidad del modelo utilizado. No se planea publicar la aplicación en tiendas de aplicaciones.

Al concluir el proyecto, se entregará una aplicación móvil para el sistema Android. Los usuarios podrán crear cuentas, cargar archivos multimedia y la aplicación intentará predecir su contenido. Además, podrán interactuar con los recursos compartidos por otros usuarios, agregar comentarios y verificar la corrección en su lengua de señas respectiva. El sistema incluirá una función de votación en la que los usuarios obtendrán un rango que aumentará mediante un examen que evaluará sus conocimientos, y un rango más alto dará mayor peso a sus votos. Los usuarios podrán ver los comentarios de otros usuarios y seleccionar la lengua de señas que prefieran para la aplicación, con la opción de cambiar entre ASL y LENSEGUA

6.1. Lengua de señas

La lengua de señas se define como un sistema lingüístico de carácter visual, gestual y manual, utilizando tradicionalmente por personas sordas, con discapacidad auditiva y/o sordociegas. [5] Se basa en expresiones faciales y en diversos movimientos de las manos, los brazos y el cuerpo.

6.1.1. Tipos de lengua de señas

La lengua de señas, al igual que el lenguaje hablado, cuenta con diferentes idiomas alrededor del mundo. Haciendo que se tenga registrado alrededor de 300 lenguas diferentes, que puede variar de región en región. Debido a que las comunidades suelen desarrollar sus propias formas de comunicarse con base en su forma de vivir, creando de esta manera las variaciones. Las lenguas más conocidas son el estadounidense (ASL), chino (CSL o ZGS) y brasileño (BSL). Siendo el chino el más grande con 20 millones de usuarios, pero el más popular del lado americano es el ASL. [6]

Dentro del lenguaje de hablado existe el acento y el dialecto, estos suelen estar influenciados por la geografía o las variaciones regionales. El acento se refiere a cómo se pronuncian las palabras mientras que el dialecto está relacionado al vocabulario, la gramática y la pronunciación. ¿Cómo es que se relacionan estas definiciones con la lengua de señas? Si bien no es estrictamente cierto que existan los acentos, los dialectos sí lo hacen. Muchos signatarios han desarrollado su propio dialecto según cómo se les enseñó, pero también influyen factores como su edad, etnia e incluso si son sordos u oyentes. En algunos casos, los diferentes signos para comunicar la misma palabra o idea permiten que se puedan diferenciar las variaciones regionales. [7]

6.1.2. Lengua de señas en Guatemala

En Guatemala, el 10.2% de la población total tiene alguna discapacidad y dentro de este porcentaje aproximadamente 250 mil habitantes cuentan con algún grado de discapacidad auditiva, según datos obtenidos en el último censo de 2018. Por esta razón el Congreso de la República aprobó el Decreto 3-2020, reconociendo oficialmente la Lengua de señas de Guatemala (Lensegua) como el medio de comunicación. Con este decreto se exhorta a las instituciones a llevar a cabo actividades de socialización y divulgación, charlas e intercambios de experiencias con la comunidad de lengua de señas. Asimismo, el Ministerio de Educación (Mineduc) tiene el compromiso de implementar un reglamento que dicte las directrices de la enseñanza y aprendizaje de esta modalidad. Con este Decreto, Lensegua está trabajando en un documento oficial en donde se establecen las lenguas oficiales de este país. [1]

6.2. Machine Learning

Machine Learning es una rama de la inteligencia artificial, ciencias de la computación y matemáticas que se enfoca en utilizar algoritmos y datos para imitar el aprendizaje humano. Al tener la velocidad de una computadora es perfecto para el análisis de millones de datos, dando como resultado que sea fuertemente utilizado en el área de ciencia de los datos para generar modelos que clasifican o predicen.

6.2.1. Tipos de modelos

Los modelos están separados en tres categorías:

- **Supervisado:** Utiliza un conjunto de datos al cual se le agrega la respuesta para una mayor precisión al momento de clasificar o predecir. Es un método bastante utilizado en el mundo real, como la clasificación de spam, identificación de intrusos entre otros. Algunos algoritmos muy utilizados son las redes neuronales, las regresiones, random forest y support vector machine (SVM).
- **No Supervisado:** Este analiza y agrupa el conjunto de datos para encontrar información o patrones sin la intervención de un ser humano. Normalmente es utilizado para la exploración de datos, segmentación y reconocimiento de imágenes. Algunos algoritmos utilizados son: K-Means, redes neuronales y agrupación probabilística.
- **Semi-Supervisado:** Este combina diferentes aspectos del no supervisado y el supervisado. Utiliza un pequeño conjunto de datos que cuenta con la respuesta y luego se le introduce otro conjunto más grande, pero sin la respuesta. [8]

6.3. Computación en la nube

La computación en la nube es una tecnología que permite a los usuarios acceder y utilizar una variedad de recursos informáticos a través del internet. Los proveedores más conocidos

son Amazon Web Services (AWS), Google Cloud Platform (GCP) y Microsoft Azure. Las principales características de la computación en la nube son:

- **Servicio Bajo Demanda:** Los usuarios pueden rentar y gestionar los recursos que se proveen, haciendo que la inversión sea conforme el uso de estos servicios.
- **Elasticidad:** Los recursos en la nube pueden escalarse rápidamente hacia arriba o hacia abajo según las necesidades de carga de trabajo, lo que permite flexibilidad y optimización de costos.

En la siguiente imagen se puede observar una comparación de servicios entre los diferentes proveedores más populares:



Figura 1: Comparación de servicios en la nube

Como se pudo observar en la imagen, los proveedores cuentan con servicios equivalentes entre sí, hablando en términos de AWS, lo necesario para este proyecto son los servicios de: EC2, Lambdas y S3. Que se puede decir en términos de Azure como Virtual Machine, Azure Functions y Blob Storage. De la misma manera, en GCP sería Compute Engine, Cloud Functions y Cloud Storage. Todos estos proveedores cuentan con una prueba gratis.

Para este proyecto se utilizará GPC, los factores decisivos para elegir al proveedor es la experiencia que se cuenta en esta plataforma y un servicio en específico que utiliza el modelo para predecir la frase o palabra, el servicio es Google MediaPipe sobre Hand Landmarker. Debido a que debe haber comunicación entre el modelo y este servicio, es más barato que el modelo este en una maquina virtual en GCP debido a que no se cobra la transferencia de datos dentro de los mismos servicios de GCP. Estas son las razones por la cual se utiliza GCP en vez de AWS o Azure.

6.3.1. Modelos de servicio

La computación en la nube ofrece varios modelos de servicio:

- **Infraestructura como servicio (IaaS):** Ofrece recursos informáticos virtualizados a través de Internet, como máquinas virtuales, almacenamiento y redes.
- **Plataforma como servicio (PaaS):** Proporciona una plataforma y un entorno para que los desarrolladores construyan, implementen y gestionen aplicaciones sin preocuparse por la infraestructura subyacente.
- **Software como servicio (SaaS):** Ofrece aplicaciones de software completamente funcionales a través de Internet, eliminando la necesidad de instalación y mantenimiento local. [9]

6.3.2. Google MediaPipe sobre Hand Landmarker

Este servicio detecta los puntos de referencia de una imagen. Opera sobre datos de imágenes con un modelo de aprendizaje automático que produce puntos de referencia de las manos en coordenadas de imagen, puntos de referencia de las manos en coordenadas del mundo y también se indica de que mano son las coordenadas.

El modelo es compatible con los siguientes formatos: imágenes, videos y transmisiones en vivo. También se le puede configurar el número máximo de manos detectadas, la puntuación mínima de confianza para la detección y el seguimiento de las manos.

Este utiliza dos modelos empaquetados, uno para la detección de la palma y otro para los puntos de referencia de los dedos. El segundo, hace referencia a 21 puntos de las articulaciones de la mano dentro de las regiones de manos detectadas [Figura 2]. Fue entrenado con aproximadamente 30 000 imágenes reales y con otros modelos. [10]

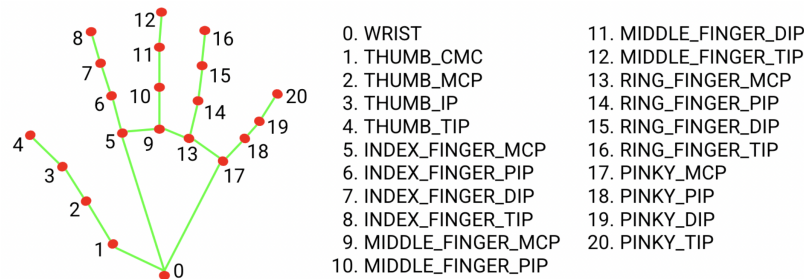


Figura 2: Puntos de referencia de la mano

6.4. Red de Entrega de Contenido (CDN)

Una red de entrega de contenido (CDN) es un conjunto de servidores ubicados estratégicamente que almacenan y proporcionan contenido de Internet, como páginas web, imágenes y videos, de manera rápida y eficiente. La CDN permite acelerar la entrega de contenido a los usuarios finales, mejorando así la experiencia de navegación en línea. Grandes plataformas, como Facebook, Netflix y Amazon, dependen en gran medida de las CDN para ofrecer un servicio rápido y confiable.

Además de mejorar la velocidad de entrega, una CDN bien configurada también puede desempeñar un papel en la seguridad en línea, protegiendo los sitios web contra ataques maliciosos, como los ataques de denegación de servicio distribuido (DDOS).

6.4.1. ¿Es una CDN igual a un servidor de sitios web?

Aunque una CDN no reemplaza a un servidor web tradicional, juega un papel fundamental al almacenar en caché contenido en servidores distribuidos en todo el mundo. Esta estrategia mejora el rendimiento del sitio web, reduce la carga del servidor de origen y, en muchos casos, resuelve problemas de ancho de banda. Los beneficios de una CDN la convierten en una opción preferida para mitigar los desafíos asociados con el alojamiento web tradicional.

6.4.2. ¿Cuáles son los beneficios de utilizar una CDN?

Los beneficios de una CDN varían según las necesidades y el tamaño de un sitio web, pero generalmente se pueden dividir en cuatro categorías:

- **Mejora de los tiempos de carga del sitio web:** Distribuir el contenido de un sitio web a través de servidores CDN cercanos a los visitantes acelera significativamente los tiempos de carga. Esto reduce la tasa de rebote y aumenta la retención de visitantes, ya que los usuarios tienden a abandonar sitios web lentos.
- **Reducción de los costes de ancho de banda:** El almacenamiento en caché y otras optimizaciones de la CDN reducen la cantidad de datos que los servidores de origen

deben proporcionar, disminuyendo los costos de ancho de banda.

- **Aumento de la disponibilidad de contenido y redundancia:** Las CDN son altamente tolerantes a fallos y pueden manejar grandes volúmenes de tráfico. Esto asegura que el contenido esté disponible incluso en situaciones adversas.
- **Mayor seguridad del sitio web:** Las CDN refuerzan la seguridad con mitigación de DDoS, mejoras en certificados de seguridad y otras optimizaciones.

6.4.3. ¿Cómo funciona una CDN?

En esencia, una CDN es una red de servidores interconectados que buscan ofrecer contenido de manera rápida y segura. Para mejorar la velocidad y la conectividad, una CDN coloca servidores en puntos estratégicos de intercambio de Internet. A través de estos puntos, diferentes proveedores de Internet se conectan para acceder al tráfico de sus redes.

Al tener servidores en estos puntos de intercambio, una CDN reduce los costos y el tiempo de entrega de datos, lo que se traduce en una experiencia de usuario más rápida y eficiente. [11]

6.5. Aplicaciones

6.5.1. Frontend

Este se encarga del diseño de una interfaz gráfica para el usuario, desde su estructura hasta sus estilos, colores, fondos, tamaños, animaciones y efectos. Estos pueden trabajar para aplicaciones web, móviles o escritorio.

Para el desarrollo del frontend se utilizó framework de React Native, el cual es un marco de desarrollo de aplicaciones móviles de código abierto desarrollado por Meta. Permite crear aplicaciones para varias plataformas, como Android, iOS, Web y Windows, al proporcionar controles de interfaz de usuario nativa y acceso a las funciones nativas de cada plataforma. Para trabajar con React Native, es necesario tener conocimientos básicos de JavaScript. [12]

Para el diseño inicial se utilizan plataformas para el desarrollo de las pantallas, aquí se enfoca en el UI (User Interface) y una de estas plataformas es **Figma**. Es una plataforma de diseño colaborativo basada en la nube que permite a los equipos de diseño trabajar de manera conjunta en proyectos en tiempo real. Algunas de las características destacadas de Figma incluyen:

- **Colaboración en Tiempo Real:** Permite a los miembros del equipo trabajar simultáneamente en un proyecto, facilitando la comunicación y la toma de decisiones en tiempo real.
- **Diseño Responsivo:** Figma es conocido por su capacidad para adaptarse a diferentes tamaños de pantalla, lo que facilita el diseño de interfaces de usuario responsivas para diversas plataformas y dispositivos.

- **Prototipado Interactivo:** Las CDN son altamente tolerantes a fallos y pueden manejar grandes volúmenes de tráfico. Esto asegura que el contenido esté disponible incluso en situaciones adversas.
- **Versionado e Historial:** Ofrece funciones de versionado y un historial detallado, lo que facilita el seguimiento de cambios y la posibilidad de retroceder a versiones anteriores si es necesario. [13]

6.5.2. Backend

Se encarga de procesar la información que alimenta al frontend de datos. Es el conjunto de sistemas de software que computan y procesan la lógica tecnológica que permite que un sitio web funcione. Es responsable de crear funciones que simplifican el proceso de desarrollo, conectarse a la base de datos, utilizar bibliotecas web del servidor y garantizar la correcta manipulación de la información en la aplicación. [14]

Para el desarrollo del backend se utilizó el framework de NestJS. Este es un framework para construir aplicaciones eficientes y escalables del lado del servidor en Node.js. Utiliza JavaScript progresivo, está construido y admite completamente TypeScript (aunque aún permite a los desarrolladores codificar en JavaScript puro) y combina elementos de la programación orientada a objetos (OOP), la programación funcional (FP) y la programación reactiva funcional (FRP). [15]

Bajo la superficie, Nest utiliza frameworks robustos de servidor HTTP como Express (el predeterminado) y opcionalmente se puede configurar para usar Fastify también.

Nest proporciona un nivel de abstracción por encima de estos frameworks comunes de Node.js (Express/Fastify), pero también expone sus API directamente al desarrollador. Esto brinda a los desarrolladores la libertad de utilizar la gran cantidad de módulos de terceros disponibles para la plataforma subyacente.

6.5.3. Interfaz de Programación de Aplicaciones (API)

Las APIs trabajan en un modelo cliente-servidor. La aplicación que envía la solicitud es el cliente y la aplicación que envía la respuesta es el servidor. Por ejemplo, en un restaurante el frontend son los comensales, el Backend son los cocineros y el API es el mesero que hace la comunicación posible entre frontend y Backend.

Existen cuatro tipos principales de APIs, cada uno con sus propias características y casos de uso:

- **APIs SOAP:** Estas APIs utilizan el Protocolo Simple de Acceso a Objetos (SOAP). El cliente y el servidor intercambian mensajes usando XML. Las APIs SOAP son menos flexibles y fueron más populares en el pasado.
- **APIs RPC:** Conocidas como Llamadas a Procedimientos Remotos, estas APIs permiten que el cliente complete una función (o procedimiento) en el servidor, y el servidor envía el resultado de vuelta al cliente.

- **APIs WebSocket:** Las APIs WebSocket utilizan objetos JSON para enviar datos y admiten comunicación bidireccional entre las aplicaciones del cliente y el servidor. El servidor puede enviar mensajes de devolución de llamada a los clientes conectados, lo que la hace más eficiente que las APIs REST.
- **APIs REST:** Estas son las APIs más populares y flexibles que se encuentran en la web hoy en día. El cliente envía solicitudes al servidor como datos. El servidor utiliza esta entrada del cliente para iniciar funciones internas y devuelve los datos de salida al cliente. [16]

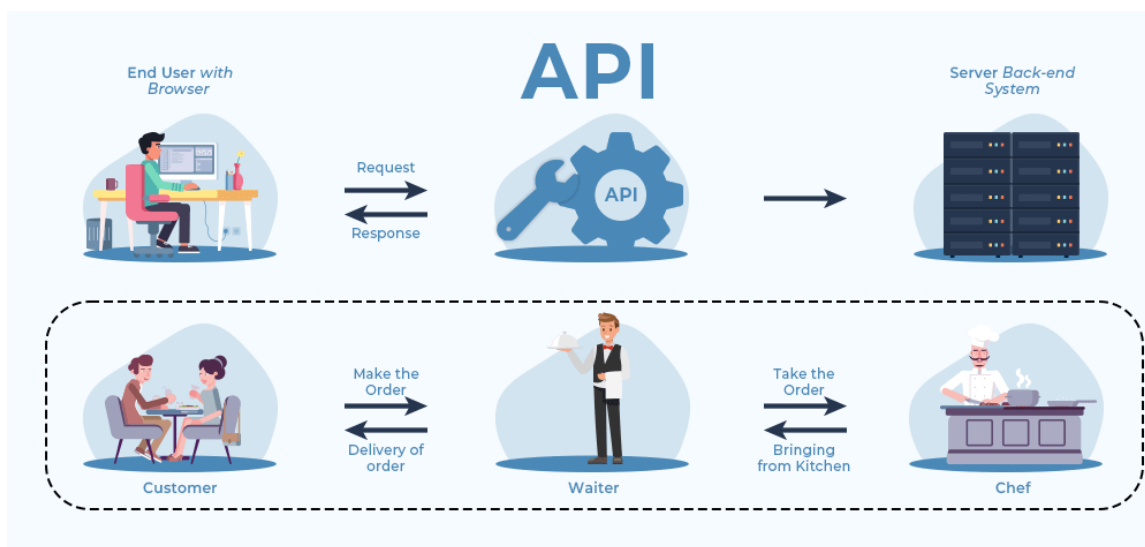


Figura 3: Flujo de un API

6.6. Servicios de la nube

6.6.1. Firebase Authentication

La identificación de usuarios es esencial para la mayoría de las aplicaciones, ya que permite almacenar los datos de manera segura en la nube y brindar una experiencia personalizada en todos los dispositivos del usuario.

Firebase Authentication ofrece una solución completa para autenticar a los usuarios en la aplicación. Proporciona servicios de backend, SDK fáciles de usar y bibliotecas de IU preconstruidas. Admite varios métodos de autenticación, como contraseñas, números de teléfono y proveedores de identidad federada populares como Google, Facebook y X.

Firebase Authentication se integra de manera fluida con otros servicios de Firebase y cumple con estándares de la industria como OAuth 2.0 y OpenID Connect, lo que facilita su integración en tu backend personalizado.

Al actualizar a Firebase Authentication con Identity Platform, se tiene acceso a características adicionales como la autenticación de múltiples factores, bloqueo de funciones,

seguimiento de actividad del usuario, registro de auditoría, compatibilidad con SAML y proveedores genéricos de OpenID Connect, funcionalidad multiusuario y soporte empresarial. [17]

6.6.2. Compute Engine (GCE)

Google Compute Engine es una plataforma que proporciona máquinas virtuales alojadas en los centros de datos de Google. Estas máquinas virtuales son altamente eficientes y están respaldadas por la red de fibra global de Google. Destacan por su inicio rápido, ofrecen almacenamiento en discos persistentes y garantizan un rendimiento constante.

Esta plataforma brinda una amplia flexibilidad, permitiendo escalar desde una sola instancia hasta una red global de cómputo en la nube con balanceo de cargas. Los usuarios pueden elegir entre varias configuraciones predefinidas o crear tipos de máquinas personalizados según sus necesidades específicas. Además, Google Compute Engine se distingue por su alto rendimiento en SSD local, lo que lo posiciona como líder en comparación con otros servicios de infraestructura como servicio (IaaS).

Entre sus ventajas se encuentra la facturación por minuto, opciones de configuración de máquinas virtuales flexibles y descuentos automáticos por uso continuo, convirtiéndolo en la elección principal en términos de precio y rendimiento para diversas cargas de trabajo.[18]

6.6.3. Cloud Run (GCR)

Cloud Run es una plataforma de procesamiento administrada que permite ejecutar contenedores directamente sobre la infraestructura escalable de Google. Permite la implementación de código en varios lenguajes de programación, gestionando la compilación de imágenes de contenedor de manera opcional. En resumen, Cloud Run simplifica el proceso para los desarrolladores, permitiéndoles centrarse en la escritura de código sin preocuparse por la operación, configuración o escalabilidad.

En Cloud Run, el código puede ejecutarse de dos formas: como un servicio que responde a solicitudes web o eventos, o como un trabajo que realiza una tarea específica y se detiene cuando se completa. Ambos tipos de ejecución comparten el mismo entorno y tienen acceso a las integraciones con otros servicios de Google Cloud. [18]

6.6.4. Cloud Functions (GCF)

Google Cloud Functions es un entorno sin servidores que permite a los usuarios desarrollar y conectar servicios en la nube de manera eficiente. Este servicio permite la programación de funciones simples y específicas que se vinculan a eventos generados por la infraestructura y los servicios en la nube. La función se activa automáticamente en respuesta a eventos observados, y su código se ejecuta en un entorno administrado, sin necesidad de gestionar servidores ni infraestructura.

Una ventaja de Cloud Functions es que admite varios lenguajes de programación, lo que

proporciona flexibilidad a los desarrolladores. Además, las funciones se pueden ejecutar en entornos estándar para los lenguajes compatibles, lo que facilita la portabilidad y las pruebas locales. Existen dos versiones de producto: Cloud Functions (1st gen) y Cloud Functions (2nd gen). [19]

6.6.5. Cloud Storage (GCS)

El Cloud Storage es un servicio de almacenamiento de datos en el que la información digital se guarda en servidores ubicados externamente. Estos servidores son administrados y protegidos por un proveedor externo, asegurando la accesibilidad de los datos a través de conexiones de Internet públicas o privadas.

Esta solución brinda a las organizaciones la capacidad de almacenar, acceder y gestionar datos sin la necesidad de mantener sus propios centros de datos, lo que conlleva a un cambio de los costos de capital a costos operativos. Además, el Cloud Storage es altamente escalable, permitiendo a las organizaciones ajustar su capacidad de almacenamiento según sus necesidades cambiantes. [20]

Para poder subir videos/imágenes a la Bucket, es necesario utilizar una URL firmada, es un enlace que otorga permisos limitados en tiempo para realizar una solicitud, en este caso es para subir archivos. Estas URLs contienen información de autenticación en la cadena de consulta, permitiendo a usuarios sin credenciales llevar a cabo acciones específicas en un recurso. [21]

6.6.6. MongoDB

MongoDB es una solución de base de datos NoSQL de código abierto que utiliza documentos flexibles en lugar de tablas y filas. Al no requerir un sistema de gestión de bases de datos relacionales, ofrece un modelo de almacenamiento de datos versátil que simplifica la gestión de bases de datos y proporciona un entorno escalable para aplicaciones multiplataforma. Los documentos de MongoDB, con formato BSON, pueden almacenar varios tipos de datos y distribuirse en diferentes sistemas. Su diseño de esquema dinámico brinda flexibilidad a los usuarios para crear registros, consultar colecciones y analizar datos de manera eficiente. [22]

Este proyecto se divide en tres grandes áreas. Primero, el Modelo de Machine Learning, que tiene como objetivo predecir palabras, frases y deletreos. Seguido del Frontend, que en este caso corresponde a la aplicación móvil y abarca toda la experiencia del usuario y el diseño de la aplicación. Por último, el Backend, encargado de gestionar la base de datos, los usuarios, los recursos y el flujo del modelo.

7.1. Modelo de Machine Learning

Para esta sección se dividieron las predicciones en dos partes:

- **Videos:** Se utilizaron tres modelos para la predicción de la frase o palabra en cada lengua. En total, se desarrollaron seis modelos, tres para ASL y tres para LENSEGUA. El video se procesa con todos los modelos, lo que permite elegir una de las tres opciones y aumenta la probabilidad de que la predicción conjunta sea correcta.
- **Deletreo o ráfagas de imágenes:** Para esta parte, se utilizó un único modelo de redes convolucionales para cada imagen. Cada imagen se procesa de manera individual y luego se unen los resultados en función del orden en que se tomaron las imágenes.

Estos agentes inteligentes se tomaron del trabajo de graduación **Plataforma de Aprendizaje de Lengua de Señas y Accesibilidad de Textos para Personas con Discapacidad Auditiva en Guatemala**.

7.2. Desarrollo de aplicación

Para poderle dar uso al modelo de Machine Learning fue necesaria una aplicación con la que interactuaria el usuario. Debido a esto se diseño e implemento la aplicación llamada Deafleans Studio. De la cual se tienen los siguientes módulos:

Para los servicios de backend se utilizó Google Cloud Platform, de este se utilizaron los siguientes servicios:

- **Firestore:** Para la autenticación y autorización de los usuarios.
- **Compute Engine:** Este servicio es necesario debido a que aquí es en dónde se estará prediciendo.
- **Cloud Run:** Este se encargará de hostear el API que atenderá las peticiones de la aplicación.
- **Cloud Functions:** La cual se encargará de generar las vistas previas de los recursos y avisar a la instancia del modelo que se debe de predecir para el recurso recién subido.
- **Cloud Storage:** Este tendrá cómo objetivo ser una CDN.
- **MongoDB:** Esta es la base de datos que se utilizó para el proyecto y que contiene toda la información de los recursos.

Para el desarrollo de la aplicación, fue necesario contar con una referencia, por lo que el primer paso consistió en crear un proyecto en Figma que incluyera los flujos principales de la aplicación.

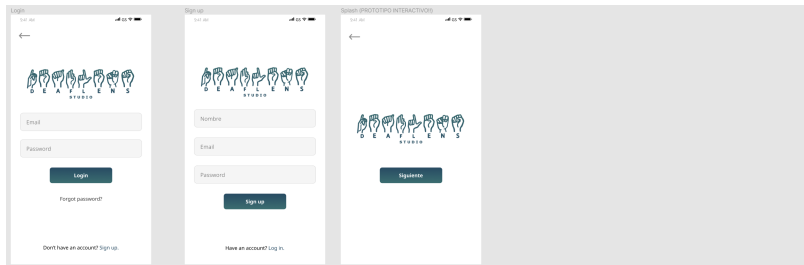


Figura 4: Flujo 1 de Figma

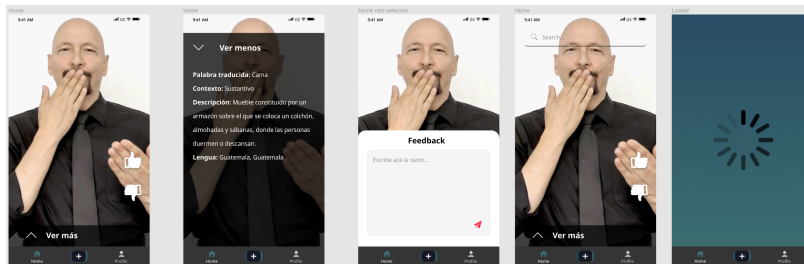


Figura 5: Flujo 2 de Figma

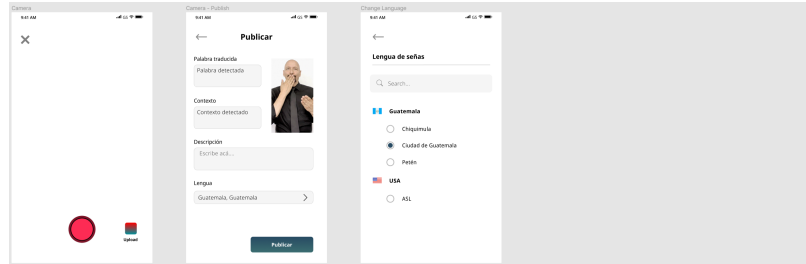


Figura 6: Flujo 3 de Figma

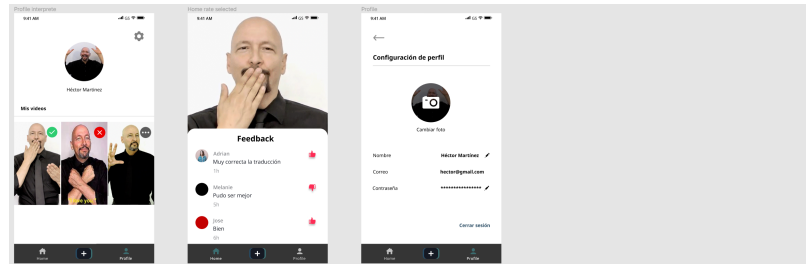


Figura 7: Flujo 4 de Figma

A continuación, se presentan las pantallas que se desarrollaron siguiendo esta guía:

- Pantalla de inicio:** Se creó la pantalla de inicio de sesión para que los usuarios con cuentas puedan disfrutar de todas las funcionalidades de la aplicación. Los usuarios se autentican con el API y, junto con **Firestore Authentication**, se les otorga acceso o no a sus cuentas.

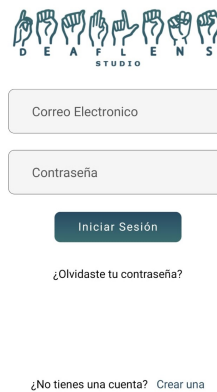


Figura 8: Pantalla de inicio de sesión

- Pantalla de creación de cuentas:** El propósito de esta pantalla es identificar a las personas y proporcionarles una cuenta. Esto les permitirá crear contenido, ya sea a

través de un video o deletreando palabras. Además, les dará la posibilidad de participar en la valoración del contenido multimedia creado por la comunidad.

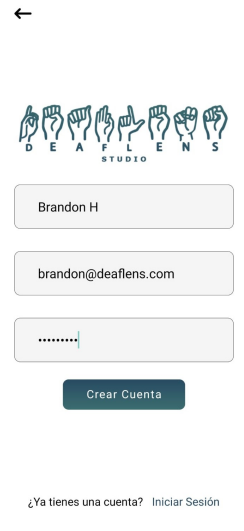


Figura 9: Pantalla de creación de cuenta

- **Pantalla de términos y condiciones:** Esta pantalla busca informar al usuario sobre qué deberá de aceptar para poder utilizar la aplicación.
- **Pantalla de selección de idioma:** El propósito es poder cambiar el idioma con el que se configurará la aplicación.
- **Pantalla principal:** Se creó la pantalla de inicio, en dónde los usuarios podrán ver el contenido de otros usuarios y agregar su voto para aceptar o rechazar el contenido, junto con un comentario.



Figura 10: Pantalla principal

- **Pantalla de perfil:** El objetivo de esta pantalla es permitir a los usuarios visualizar los recursos que han subido a la aplicación y observar la retroalimentación que han recibido.

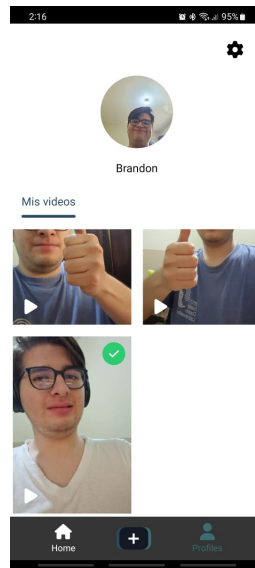


Figura 11: Pantalla de perfil

- **Pantalla de cambio de rango:** Esta pantalla es muy importante ya que le permite al usuario subir de rango para poder tener un mayor peso dentro de la aplicación. Aquí se le pasa un examen que aumenta la dificultad conforme aumenta el rango.
- **Pantalla de crear recurso:** El propósito de esta pantalla la de grabar un video o tomarse fotos para el deletreo. Esta pantalla estandariza el formato que debe de tener el recurso para poder ser procesado correctamente por el modelo.

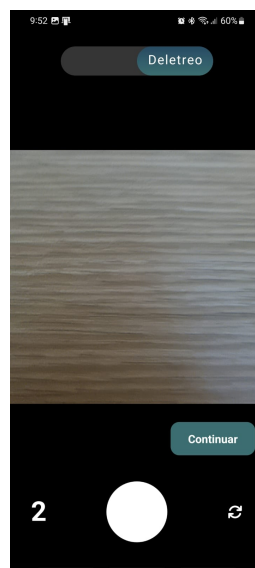


Figura 12: Pantalla de deletreo

- **Pantalla de vista previa:** Esta tiene por objetivo poder visualizar el contenido creado para que el usuario este seguro de lo que quiere subir a la plataforma.
- **Pantalla de subir recurso:** Como lo indica su nombre, esta pantalla sube el o los recursos y espera a que se procesen del lado del modelo.
- **Pantalla de publicar:** Por último se creó la pantalla de publicar el recurso, el cual le muestra al usuario la predicción de su recurso y le da la opción al usuario de modificarla y agregar cierta información, como lo es el contexto y la descripción.

7.2.1. Validación de conocimientos

En la aplicación hay diferentes tipos de rangos, que demuestran el conocimiento del usuario sobre una lengua de señas en específico, estos rangos son:

- Sin rango
- Poco conocimiento
- Básico
- Intermedio
- Avanzado
- Experto

Conforme aumente el rango, también aumenta el peso del voto para rechazar o aceptar el recurso. De esta manera se necesitarán menos votos para aceptar un recurso en el que un usuario **Experto** votó; que uno con usuarios con **Pocos conocimientos**. Cabe resaltar que el flujo está diseñado para que los usuarios recién registrados que no cuenten con un rango, no tengan un peso real en la votación de aceptación o rechazo del recurso. Esto se debe a que no se sabe realmente si conoce o no sobre la lengua.

7.2.2. Flujos de la aplicación

Flujo de predicción de recursos

1. La aplicación solicita al API las URLs autenticadas para cargar los archivos multimedia. Para cada imagen del deletreo se debe crear su URL autenticada, lo que permite a los usuarios de Deafflens cargar archivos de manera exclusiva. La aplicación también informa al servidor sobre el idioma utilizado para mantener un orden dentro de la CDN.
2. El API devuelve la cantidad de URLs autenticadas necesarias para que el usuario pueda cargar su video o deletreo en la CDN.
3. El usuario carga los archivos hacia la CDN a través de la aplicación.

4. La CDN genera un evento que indica la creación de uno o varios archivos.
5. Un Cloud Function atiende el evento generado, creando un avance del video con el primer fotograma o copiando la primera imagen del deletreo como avance.
6. Además, el servidor responsable de mantener el modelo es notificado sobre este nuevo recurso que debe predecir, indicando la cantidad de archivos, si es deletreo o no y el idioma a predecir.
7. Los modelos realizan las predicciones y envían los resultados al API, que los introduce en la base de datos.
8. La aplicación solicita las predicciones al API y, cuando estas están disponibles, se las devuelve.

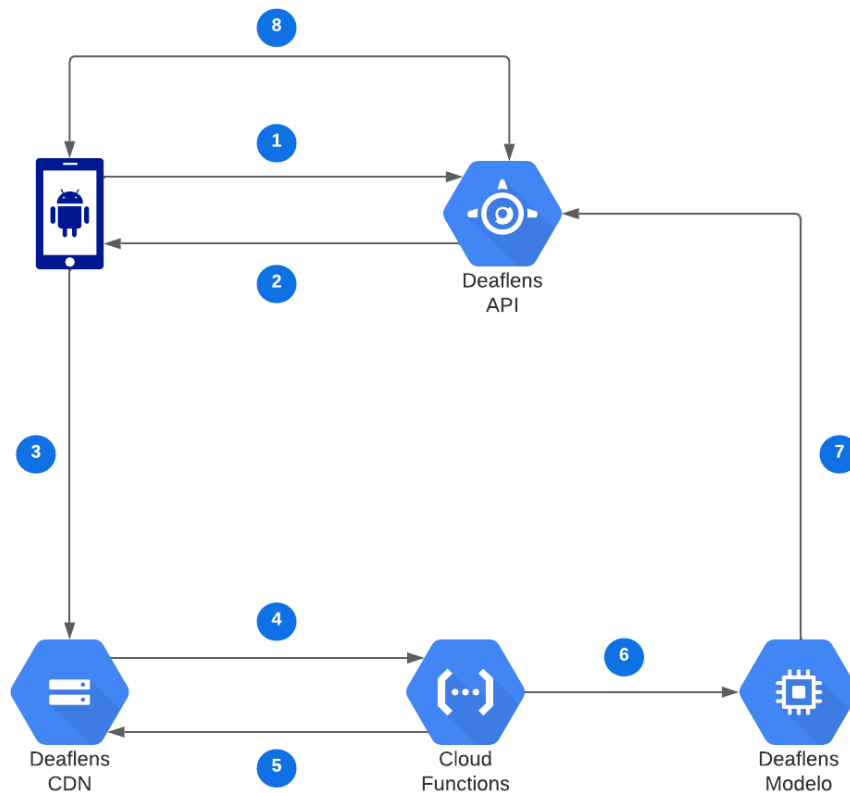


Figura 13: Flujo predicción de recursos

Flujo validación de recursos

1. El usuario publica su recurso, ya sea un video o una predicción.
2. Otro usuario visualiza el recurso, verifica el contexto, la descripción y lo que se intenta comunicar en el video. Luego emite su voto de calidad mediante "me gusta." "no me gustaz proporciona un comentario explicando su decisión.

3. El API registra el voto del usuario, verifica su rango y decide si aprueba o no el recurso. De esta manera, se le asigna una puntuación de calidad al recurso, y los votos de los usuarios con un mayor conocimiento de la lengua de señas tienen un peso más significativo.
4. Cuando el recurso alcanza un cierto valor de votos positivos o negativos, el API cambia su estado a "aceptado" o "no aceptado".

7.3. Recolección de resultados

Para poder validar la aplicación fue necesario la creación de una encuesta. La cual guiaba al usuario alrededor de la aplicación dándole pequeñas tareas y les hace varias preguntas con respecto a los flujos, tomando en cuenta su opinión.

Se seleccionaron diferentes tipos de usuarios con diferentes rangos de conocimientos. De esta manera se obtendrá el punto de vista desde alguien que no tiene conocimientos de las lenguas de señas hasta alguien que conoce perfectamente una o más lengua de señas y lo que involucra. Luego de recolectar los resultados se tabularon y se generaron gráficos para su interpretación.

En esta sección se mostrarán todos los resultados de las preguntas de la encuesta. Estos resultados se agruparon según su categoría, por lo que no necesariamente están en el mismo orden en el que aparecieron en la encuesta.

8.1. Predicción de recursos en LENSEGUA

El objetivo de esta sección es verificar la calidad del agente inteligente al momento de predecir los recursos que el usuario ha subido a la aplicación. La siguiente figura muestra el resultado del deletreo de una palabra, mientras que las otras dos figuras representan la predicción de dos palabras en un video.

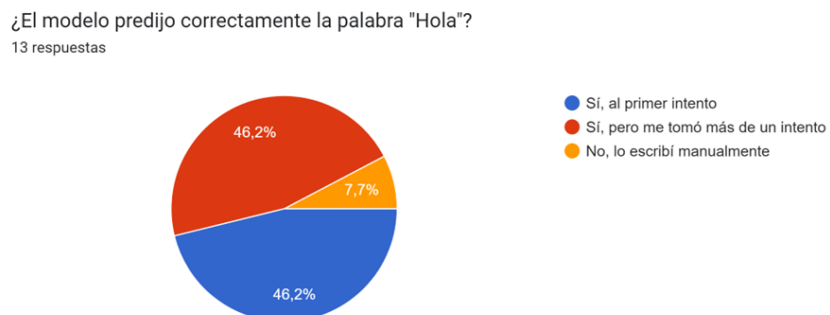


Figura 14: Predicción de la palabra **hola** en deletreo

Como se observa en la figura, la predicción de la palabra 'hola' tuvo una tasa de aciertos

del **92.3%**, mientras que solo el **7.7%** no pudo obtener una predicción correcta, incluso después de varias iteraciones.



Figura 15: Predicción de la palabra **pescado** en video

La predicción de la palabra 'pescado' tuvo una tasa de aciertos del **100%**. A pesar de esto, el **7.7%** tuvo que iterar varias veces para obtener la predicción correcta. Esto significa que el **92.3%** pudo ser predicho correctamente en la primera iteración.



Figura 16: Predicción de la palabra **inhalador** en video

Por otro lado, la predicción de la palabra 'inhalador' tuvo una tasa de aciertos del **100%**, al igual que la palabra anterior. Sin embargo, en el caso de esta palabra, el **15.4%** tuvo que iterar varias veces para obtener la predicción correcta. Esto significa que el **84.6%** pudo ser predicho correctamente en la primera iteración.

8.2. Predicciones de recursos en ASL

En esta sección se muestra la predicción de la lengua de señas ASL con un video.

¿El modelo predijo correctamente la expresión "see you later"?
13 respuestas

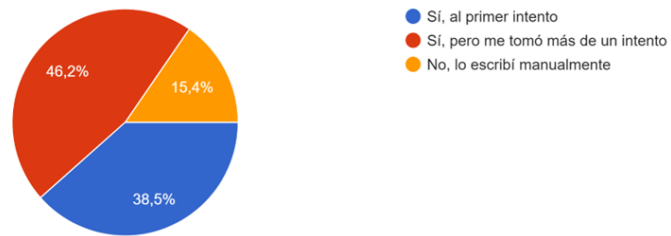


Figura 17: Predicción de la palabra **see you later** en video

Por otro lado, en el caso de la lengua de señas ASL, se logró predecir correctamente el **84.6%** de las veces, ya sea en una o varias iteraciones. Sin embargo, el **15.4%** no obtuvo la predicción correcta incluso después de varias iteraciones.

8.3. Aumentar el rango

En la presente sección se muestra como se comportaron los usuarios al momento de pasar al rango de **Poco conocimiento**

Luego de terminar la prueba, cuéntanos, ¿Conseguiste avanzar de nivel?
13 respuestas

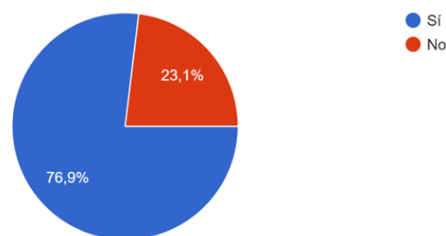


Figura 18: Aumento de rango de los usuarios que hicieron la prueba

En el examen para aumentar el rango en la aplicación LENSEGUA, el **76.9%** de los usuarios lograron subir de rango con éxito. Sin embargo, el **23.1%** no pudo hacerlo, por lo que estos usuarios deberán estudiar más para poder avanzar de rango.

Como mencionabamos con anterioridad, el nivel de la cuenta es predeterminado, para la siguiente iteración, ¿cómo consideras que sería la mejor forma para avanzar de nivel?

13 respuestas

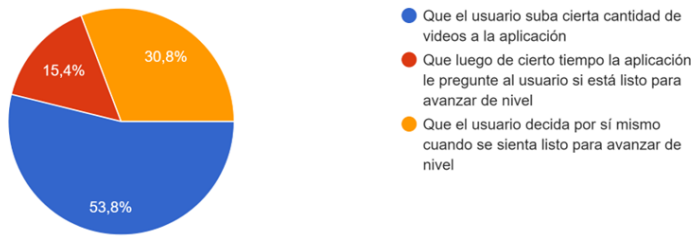


Figura 19: Métricas para avanzar en el rango

Tras realizar el examen, se observó que el **53.8%** de los usuarios consideraron que sería una buena idea que los usuarios tengan varios recursos subidos para optar a subir de rango. El **30.8%** indicó que sería adecuado que los usuarios puedan intentar subir de rango cuando se sientan preparados. Por último, el **15.4%** de los usuarios encuestados opinaron que sería mejor que los usuarios avancen de rango después de haber pasado cierto tiempo en la aplicación.

¿Qué tiempo consideras correcto para preguntarle al usuario si está listo para avanzar de nivel?

2 respuestas

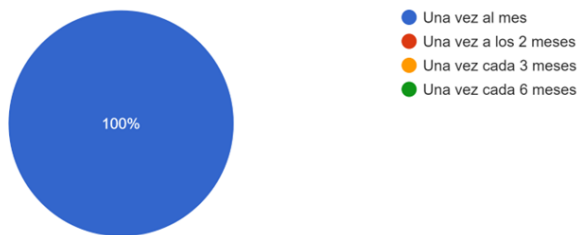


Figura 20: Aumentar rango con base en el tiempo

De las personas que opinaron que es mejor que el usuario avance de rango después de un cierto tiempo, el **100%** sugiere que hacerlo cada mes sería una buena idea.

¿Qué cantidad de videos consideras correcta a subir para avanzar de nivel?
7 respuestas

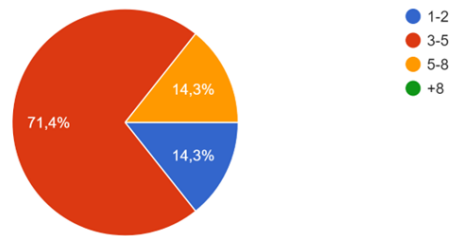


Figura 21: Aumentar rango con base con cantidad de recursos aceptados

De los encuestados que opinaron que es mejor que el usuario avance de rango luego de subir cierta cantidad de recursos, el **71.4%** mencionó que sería una buena idea hacerlo después de subir entre 3 y 5 recursos. El **14.3%** indicó que sería adecuado entre 5 y 8 recursos, con el mismo porcentaje (**14.3%**) sugiriendo que entre 1 y 2 recursos sería una buena idea, y ningún encuestado consideró un número mayor a 8 recursos como adecuado.

8.4. Validación de contenido

Esta sección tiene como propósito guiar al usuario sobre cómo es que se valida la calidad del contenido del resto de compañeros y verificar cómo es la calidad de los mismos.

¿En tu pantalla de **home** te aparecen los videos de otras personas? Recuerda que en el flujo 2 publicaste los videos, deberías de poder ver los de otros compañeros. Revisa el contenido y agregale un voto de aceptación o rechazo al de tus compañeros. ¿Lograste darle feedback a tus compañeros?

13 respuestas

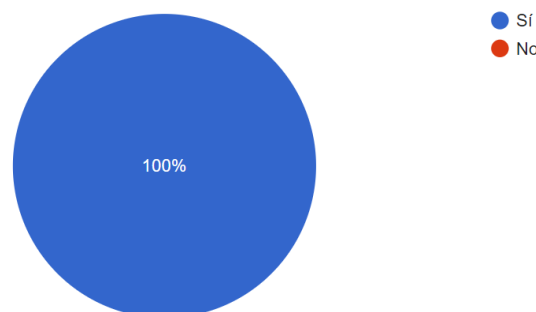


Figura 22: Ver recursos en la pantalla de inicio y dar feedback

Se observa que el **100%** de los usuarios lograron brindar algún tipo de comentario a los recursos de otros usuarios dentro de la plataforma.

¿En tu pantalla de **profile** te aparecen tus propios videos? Recuerda que en el flujo 2 publicaste los videos, deberías de poder ver tus propios videos. ¿Alguno de los recursos que publicaste, tuvo un cambio de estado a rechazado (representado con una X) o aceptado (representado con un cheque)?

13 respuestas

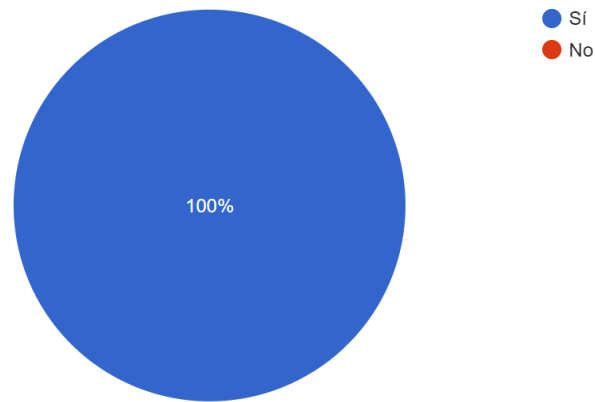


Figura 23: Ver recursos propios y observar estado

Se observa que el **100 %** de los usuarios lograron observar cómo su recurso fue aceptado o rechazado dentro de la plataforma.

8.5. Observaciones

Al momento de pasar las encuestas se observó lo siguiente:

Los usuarios tuvieron problemas al momento de instalar la aplicación. Debido a que para descargar la aplicación necesitan hacerlo de una manera diferente a como están acostumbrados los usuarios, que sería descargarlo directamente de la Play Store. En general no se tuvieron dificultades al realizar las tareas. Pero una en específico sí causó bastantes problemas, el cual fue el examen para subir rango. A las personas que no contaban con conocimientos se les tuvo que ayudar dándoles pistas, pero no dándoles las respuestas. De esta manera lograron subir de rango en un 75%. Además los usuarios se desesperaron debido a que el modelo tardaba en predecir la palabra o frase por lo que fue un buen punto que se podía mejorar de la aplicación.

Análisis de resultados

Los objetivos planteados para este trabajo son: desarrollar una aplicación móvil para la distribución de contenido multimedia, predecir el mensaje con base en la lengua de señas seleccionada por el usuario y verificar la calidad. Implementar dentro de la aplicación dos lenguas de señas diferentes dentro de la aplicación. Diseñar un proceso de verificación de recursos en el que los usuarios puedan emitir un voto para aceptar o rechazar las señas registradas. En dónde se cumplieron satisfactoriamente los objetivos planteados.

Dentro de la aplicación, se implementó la función de utilizar diferentes modelos para cada una de las lenguas de señas. En el caso de LENSEGUA, se puede observar en las figuras **14**, **15**, y **16**: cómo la comunidad logró una predicción positiva, el resultado es que la mayoría de los casos se obtuvo la predicción correcta, sin embargo, en la **Figura 14**, un usuario no fue capaz de obtener la predicción correcta del deletreo de 'hola'. Al revisar el resultado de la predicción en la base de datos, se observa que la palabra que se predijo fue 'holg'.

También se observa que las **figuras 15** y **16** cuentan con una tasa de éxito en el primer intento superior al **80%**, con un **0%** de errores en las predicciones. Por otro lado, en la **Figura 17**, que intenta predecir la palabra 'See you later' en ASL, tuvo una exactitud y precisión menor en comparación con LENSEGUA. Esto se debe a que solo el **38.5%** de los usuarios pudo obtener la predicción correcta en el primer intento. No obstante, el **84.6%** de los usuarios obtuvo la respuesta correcta después de varios intentos.

Cabe resaltar que la calidad del modelo está fuera del alcance del proyecto. Sin embargo, es importante que el modelo tenga una exactitud suficiente para que el usuario tenga una buena experiencia al usar la aplicación. Por esta razón, es necesario validar el flujo desde la creación del contenido, pasando por su predicción por parte del modelo y la verificación del resultado del modelo por parte del usuario. Basándonos en los resultados, es evidente que se pudo implementar un agente inteligente capaz de reconocer la lengua de señas de un recurso dentro de la aplicación.

Por otro lado, se cuentan con las figuras **18**, **19**, **21** y **20**, donde se observa que aproximadamente un **80 %** de los usuarios lograron ascender de rango. Esto permite que los usuarios comiencen a validar los recursos de otros usuarios y que sus votos tengan un peso significativo. Al avanzar de rango y participar en la validación de recursos, estos pueden ser rechazados o aceptados, proporcionando una crítica confiable y otorgando un grado de calidad a los recursos con numerosos votos positivos.

Por último, como se observa en las figuras **22** y **23**, se analiza la interacción de los usuarios en la aplicación. En la **Figura 22**, el **100 %** de los usuarios visualizó el contenido publicado por otros usuarios. También se puede notar en la **Figura 23** que el **100 %** pudo observar cómo al menos uno de sus recursos cambió de estado a aceptado o rechazado. Con esto en mente, queda claro que la creación de un flujo para la validación de recursos por parte de la comunidad fue exitosa.

CAPÍTULO 10

Conclusiones

- Al identificar el porcentaje de usuarios que lograron observar el cambio de estado de sus recursos subidos, es evidente que se logró establecer un flujo de verificación en el que los usuarios proporcionan un voto para aceptar o rechazarlos.
- Al observar la tasa de éxito de la predicción de los recursos, que queda claro que se pudo implementar con éxito las lenguas de señas ASL y LENSEGUA.
- Es indiscutible que se desarrolló una aplicación móvil para la distribución de contenido multimedia, predecir el mensaje con base en la lengua de señas seleccionada por el usuario y verificar la calidad con base a los votos de la misma comunidad.

- Se propone la creación de un flujo automatizado basado en la nube que facilite el entrenamiento de palabras, frases y deletreos, ya sean conocidos o nuevos.
- Implementar anuncios dentro de la aplicación para establecer una fuente constante de ingresos que respalde el mantenimiento de los servidores.
- Ampliar la compatibilidad de la aplicación para incluir el sistema operativo iOS.
- Implementar técnicas de 'lazy loading' en la página principal con el fin de reducir el tiempo de carga inicial de la aplicación.
- Optimizar la configuración de la CDN (Red de Distribución de Contenido) para garantizar que los recursos se almacenen geográficamente cerca de los usuarios. Esta estrategia tiene como objetivo reducir significativamente los tiempos de carga de videos e imágenes en la aplicación.

- [1] M. Palma. “LENSEGUA, LEY QUE FOMENTA LA INCLUSIÓN SOCIAL.” (2022), dirección: https://www.congreso.gob.gt/noticias_congreso/9131/2022/4#gsc.tab=0.
- [2] Lenovo. “Lenovo’s AI-powered sign language translation solution empowers signers in Brazil.” (2023), dirección: <https://news.lenovo.com/ai-powered-sign-language-translation-solution-hearing-brazil/>.
- [3] MobiReactor. “Sign Language ASL Pocket Sign.” (2023), dirección: <https://play.google.com/store/apps/details?id=com.mobireactor.signlanguage&hl=en&pli=1>.
- [4] Kitsord. “Kitsord - Lengua de señas.” (2023), dirección: https://play.google.com/store/apps/details?id=com.kitsord&hl=es_GT&gl=US.
- [5] RAE. “lengua de signos.” (N.A), dirección: <https://dpej.rae.es/lema/lengua-de-signos>.
- [6] SignSolutions. “What are the different types of sign language?” (2021), dirección: <https://www.signsolutions.uk.com/what-are-the-different-types-of-sign-language/>.
- [7] SignSolutions. “Are there different accents in sign language?” (2021), dirección: <https://www.signsolutions.uk.com/are-there-different-accents-in-sign-language/>.
- [8] IBM. “What is machine learning?” (N.A), dirección: <https://www.ibm.com/topics/machine-learning>.
- [9] AWS. “What is cloud computing?” (N.A), dirección: <https://aws.amazon.com/what-is-cloud-computing/>.
- [10] Google. “Hand landmarks detection guide.” (2023), dirección: https://developers.google.com/mediapipe/solutions/vision/hand_landmarker.

- [11] Cloudflare. “¿Qué es una Red de entrega de contenido (CDN)?” (N.A), dirección: <https://www.cloudflare.com/es-es/learning/cdn/what-is-a-cdn/#:~:text=En%20esencia%2C%20una%20CDN%20es,de%20intercambio%20entre%20diferentes%20redes..>
- [12] Microsoft. “Introducción a la creación de una aplicación de escritorio con React Native para Windows.” (2023), dirección: <https://learn.microsoft.com/es-es/windows/dev-environment/javascript/react-native-for-windows> (visitado 2023).
- [13] Figma. “Figma.” (2024), dirección: <https://www.figma.com/> (visitado 2024).
- [14] N. Chapaval. “Qué es Frontend y Backend: características, diferencias y ejemplos.” (2018), dirección: <https://platzi.com/blog/que-es-frontend-y-backend/> (visitado 2018).
- [15] NestJS. “NestJS.” (2023), dirección: <https://docs.nestjs.com/> (visitado 2023).
- [16] AWS. “What Is An API (Application Programming Interface)?” (N.A), dirección: <https://aws.amazon.com/what-is/api/>.
- [17] Google. “Firebase.” (2023), dirección: <https://firebase.google.com/docs/auth?hl=es-419> (visitado 2023).
- [18] Google. “Compute Engine.” (2023), dirección: https://cloud.google.com/compute/?hl=es_419 (visitado 2023).
- [19] Google. “Descripción general de Cloud Functions.” (2023), dirección: <https://cloud.google.com/functions/docs/concepts/overview?hl=es-419> (visitado 2023).
- [20] Google. “¿Qué es Cloud Storage?” (2023), dirección: <https://cloud.google.com/learn/what-is-cloud-storage?hl=es-419> (visitado 2023).
- [21] Google. “URL firmadas.” (2023), dirección: <https://cloud.google.com/storage/docs/access-control/signed-urls?hl=es-419#overview> (visitado 13-12-2023).
- [22] IBM. “¿Qué es MongoDB?” (2023), dirección: <https://www.ibm.com/mx-es/topics/mongodb> (visitado 2023).