

UNIVERSIDAD PRIVADA FRANZ TAMAYO
FACULTAD DE INGENIERÍA EN SISTEMAS
CARRERA DE INGENIERIA EN SISTEMAS



“Evaluación Procesual”

AUTORES:

Carlos Andrew Luvi Aguilar

TUTOR:

Enrique Alejandro Laurel Cossio

LAPAZ - BOLIVIA

2025

Título

Plataforma Inteligente para la Mejora de la Fonética y la Dicción en Personas con Problemas Orofaciales.

Introducción

El presente proyecto busca desarrollar un software de logopedia enfocado en mejorar la pronunciación y fonética de niños de 3 a 8 años con dificultades del habla. Mediante el uso de tecnologías de Big Data e Inteligencia Artificial, el sistema analizará grabaciones de voz para identificar errores de pronunciación y generar recomendaciones personalizadas.

La etapa de recolección de datos es esencial, ya que permitirá construir una base de audios reales sobre la que se entrenarán los modelos de análisis fonético. Cada grabación será realizada con consentimiento de los padres o tutores y bajo condiciones controladas, garantizando calidad, privacidad y precisión en los resultados.

Este proyecto combina el aprendizaje tecnológico con una aplicación social importante: apoyar el desarrollo comunicativo infantil a través de una herramienta educativa accesible y moderna.

Planteamiento del Problema

En Bolivia y en muchos países de Latinoamérica, miles de personas especialmente niños con trastornos orofaciales, adultos en rehabilitación neurológica y pacientes con dificultades del habla carecen de acceso a terapias de logopedia de calidad. Las terapias tradicionales suelen ser presenciales, costosas y limitadas en su alcance, lo que genera brechas de acceso para personas en áreas rurales o con bajos recursos. Además, los métodos convencionales no aprovechan las grandes cantidades de datos generados durante las prácticas de rehabilitación (grabaciones de voz, métricas de fonética, progresos por sesión). Esto provoca que no exista un análisis profundo ni una personalización adecuada de los tratamientos, reduciendo la eficacia en la rehabilitación del habla.

Objetivos

Objetivo general

Desarrollar una plataforma inteligente de logopedia que, mediante técnicas de Big Data e inteligencia artificial, mejore la fonética y la dicción de personas con

problemas orofaciales, ofreciendo evaluaciones automáticas, seguimiento de progreso y recomendaciones personalizadas para terapeutas y pacientes.

Objetivos específicos

- Diseñar e implementar el flujo de captura, preprocesado y almacenamiento de grabaciones de voz y métricas acústicas.
- Desarrollar modelos de Machine Learning que evalúen la calidad fonética y estimen la evolución del paciente.
- Construir un dashboard web para visualización de métricas, informes y recomendaciones.
- Validar la plataforma mediante pruebas con casos controlados y métricas de mejora.

Alcance

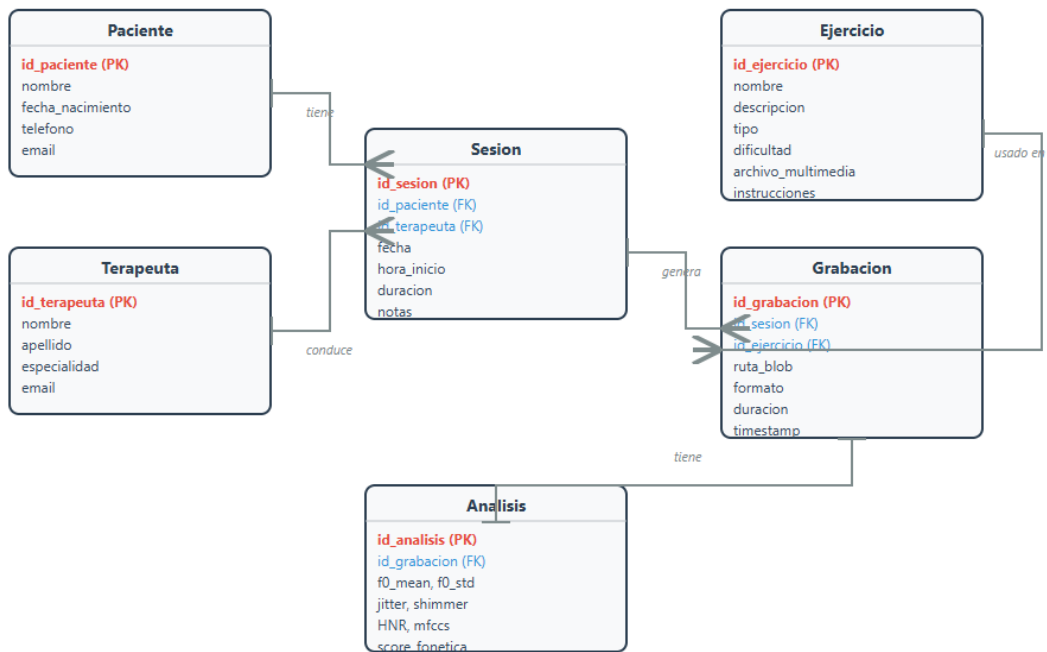
El proyecto incluirá:

- Módulo de captura (web) para grabar ejercicios y enviar las grabaciones al servidor.
- Pipeline de preprocesamiento: normalización, extracción de features (F0, jitter, shimmer, HNR, MFCCs).
- Almacenamiento centralizado en la nube (bases de datos orientadas a big data y almacenamiento de blobs para audios).
- Modelos de clasificación/regresión y dashboard para terapeutas.
- Biblioteca de ejercicios y motor de recomendaciones personalizadas.

No incluye:

- Integración completa con historia clínica de terceros.
- Certificaciones médicas oficiales.
- Hardware propietario (se usará micrófono estándar del dispositivo).

Fuente de datos



Esquema de tablas

- pacientes (id_paciente PK, nombre, fecha_nac, genero, telefono, email, observaciones)
- terapeutas (id_terapeuta PK, nombre, credenciales, email)
- sesiones (id_sesion PK, id_paciente FK, id_terapeuta FK, fecha, nota)
- grabaciones (id_grabacion PK, id_sesion FK, id_ejercicio FK, ruta_blob, duracion, formato, timestamp)
- analisis (id_analisis PK, id_grabacion FK, f0, jitter, shimmer, hnr, mfcc_json, score_fonetica, created_at)
- ejercicios (id_ejercicio PK, nombre, descripcion, tipo, dificultad, media_url)
- planes (id_plan PK, id_paciente FK, id_terapeuta FK, fecha_inicio, fecha_fin, objetivos)

Plan de recolección de datos (3–8 años)

Objetivo:

Obtener grabaciones de voz de niños de 3 a 8 años para analizar su pronunciación y entrenar modelos que detecten errores de dicción.

Población y muestra:

30 niños (15 con dificultades de habla y 15 sin ellas), reclutados en colegios y centros de terapia, con consentimiento de los padres.

Datos a recolectar:

- Audios WAV
- Edad, sexo, tipo de ejercicio, fecha
- Observaciones del terapeuta

Procedimiento:

1. Consentimiento del tutor
2. Grabación de vocales, sílabas y palabras simples
3. Subida del audio al sistema con metadatos
4. Revisión y etiquetado por terapeuta

Criterios de calidad:

Audio limpio ($\text{SNR} \geq 20 \text{ dB}$), duración adecuada, datos anónimos y cifrados.

Resultado esperado:

Dataset inicial de 150 audios infantiles listos para análisis fonético

Importancia del Proyecto

Este proyecto busca impacto social y tecnológico, respondiendo a necesidades reales:

- Contribuir a la inclusión educativa y social de personas con problemas de comunicación.
- Aprovechar el Big Data y la inteligencia artificial para generar modelos predictivos y adaptativos en terapias del habla.
- Reducir costos y ampliar el acceso a terapias mediante una plataforma web y móvil conectada a un sistema centralizado de datos.
- Brindar a profesionales de la salud un dashboard avanzado con métricas, patrones de progreso y predicciones sobre la evolución del paciente.

Posible Solución

Se plantea el desarrollo de una plataforma inteligente de logopedia asistida por Big Data, que:

1. Capture la voz del usuario a través de un micrófono de PC o dispositivo móvil.
2. Procese la señal con algoritmos de reconocimiento de patrones y análisis acústico (frecuencia fundamental, jitter, shimmer, HNR).
3. Genere un dataset estructurado y lo almacene en la nube para análisis masivo.
4. Aplique Big Data Analytics y Machine Learning para:
 - Identificar patrones de mejora o retroceso en la fonética.
 - Predecir tiempos estimados de rehabilitación.
 - Recomendar ejercicios personalizados.
5. Muestre resultados en un panel interactivo para logopedas, familias y pacientes.

Cómo se enfrentará técnicamente la solución

- Recolección de datos: Grabaciones de voz, evaluaciones clínicas y métricas en tiempo real.
- Preprocesamiento: Normalización de audio, extracción de características acústicas y limpieza de datos.
- Almacenamiento: Base de datos en la nube.
- Análisis: Algoritmos de Machine Learning (clasificación, regresión y clustering).
- Visualización: Dashboard web con gráficos dinámicos para médicos y terapeutas (Power BI, Tableau o librerías como D3.js).
- Interacción: Aplicación móvil/web para los usuarios, con ejercicios y feedback en tiempo real.

Plan de trabajo (resumen — 6 meses estimados)

Fase	Duración	Entregables
Diseño y setup	2 semanas	Documentación, arquitectura, dataset plan.
Medio de captura + storage	4 semanas	Cliente básico + API + almacenamiento.
Pipeline y extracción	4 semanas	Workers y extracción de features.
Dashboard básico	4 semanas	Visualizaciones y reportes básicos.
Modelado ML inicial	6 semanas	Modelos de baseline y pruebas.
Integración y mejoras	4 semanas	Recomendaciones automáticas y mejoras UI.
Pruebas con usuarios	4 semanas	Resultados de prueba y ajustes.
Documentación final	2 semanas	Informe, manual de uso, resultados.

Diagrama de Explicación del Proyecto

