



Tecnológico
de Monterrey

Proyecto Integrador (Gpo. 10)

Avance 0. Propuesta de proyecto y firma de convenios

Equipo #6

Julio César Pérez Zapata	A01793880
Christian Emilio Saldaña López	A00506509
Jorge Estivent Cruz Mahecha	A01793808

21 de Enero de 2024

Esta fase consiste en definir el alcance del proyecto integrador, considerando los objetivos comerciales, para evaluar con mayor precisión el impacto del proyecto en la empresa/organización. La propuesta de proyecto integrador debe contener los siguientes elementos, basándose en la metodología CRISP-ML:

1. Antecedentes

Se refieren a la información y contextos que proporcionan la base del proyecto. Incluyen la descripción de la empresa / organización y del proceso o procesos de negocios que serán impactados por el proyecto:

- Los modelos inteligentes basados en datos han evolucionado a tal grado que podemos pensar en un médico artificial que diagnostique alguna patología del habla de un paciente, estime su severidad y tenga la capacidad de monitorear su evolución en terapias, con solo abrir una página web o descargar una aplicación.

Los trastornos del habla son afecciones que dificultan crear o formar sonidos necesarios para comunicarse con otros, o relacionados con la calidad de la voz. Se caracterizan por una interrupción en el flujo o ritmo del habla, como el tartamudeo o falta de fluidez, de acuerdo con Medline Plus.

Pueden haber dificultades a la hora de formar sonidos, tono, volumen o calidad de la voz, incluso puede haber una combinación de los cuatro problemas.

Algunos de estos trastornos se deben a infecciones en los tejidos que generan la voz, cambios ambientales, distrofia muscular, dolor de cara, anormalidades genéticas, estrés, traumatismos en el cerebro, pérdida de la audición, o lesiones en las cuerdas vocales como los pólipos.

El director del Hub de Ciencia de Datos del Tec de Monterrey, Juan Arturo Nolasco Flores, compartió este proyecto con la maestría de MNA. El objetivo es crear un prototipo que fuera funcional y que se pudiera utilizar en el ramo médico (TecScience, 2022).

2. Entendimiento del negocio

- Formulación del problema: ¿Qué es lo que se intenta resolver?

El problema es que muchas personas padecen alguna patología vocal y no tienen acceso a servicios médicos por diversas situaciones o porque no tienen el tiempo de ir a una cita médica por falta de tiempo, se quiere brindar alternativas adicionales al personal médico que acerque al paciente a la identificación de su posible patología sin necesidad de acercarse a algún establecimiento a realizarse exámenes en una etapa temprana.

- Contexto: ¿Por qué es importante resolver este problema?

Dado que es muy diverso el origen y causas de estas enfermedades, este proyecto atendería padecimientos que tengan que ver con los músculos de la voz, cuerdas vocales y garganta.,el poder obtener un resultado positivo de este proyecto elevará la atención de pacientes con patologías vocales a un nivel más accesible y con una menor inversión tanto económica como de recursos médicos.

- Objetivos: ¿Cuál es la meta prevista?

En la parte inicial, se está trabajando en la detección de disfonías, ya que son los trastornos con mayor cantidad de muestras, lo cual es muy útil para generar un modelo inteligente basado en datos, el objetivo principal es obtener un modelo que permita mediante la identificación vocal el poder clasificar si un paciente tiene una patología vocal o no, y en pasos posteriores la correcta identificación de su patología específica.

En un futuro, un paciente podría abrir una página web o descargar una aplicación y, desde ahí, recibir un diagnóstico médico que le indique qué patología del habla tiene, le diga el grado de alteración —del uno al cinco, de menor a mayor grado de afectación—, le dé terapia y seguimiento a sus avances.

- Preguntas clave: ¿Cuáles son las preguntas clave que deben responderse?

¿Son suficientes los datos para generar un modelo preciso?, ¿qué tan factible es la implementación de un posible software para esta interacción con el paciente?, ¿a parte de la medicina específica este proyecto puede trascender a otros campos de acción?, ¿ Quien sería el primer ente en usar el prototipo de identificación?.

- Involucrados: ¿Quiénes participan: personal / equipos / departamentos?
¿En qué etapa y con qué nivel/tipo?

Los involucrados serían nosotros el equipo #6, la Tutora Asignada la Dra. Ludivina Facundo y el Dr Juan Arturo Nolasco Flores quienes validarán los avances periódicos que el equipo entregará..

3. Entendimiento de los datos

- Descripción de los datos: ¿Cuál es el contexto de estos datos? ¿Qué contienen?

Los datos están compuestos por audios en diferentes lenguajes y la información de la patología que representan. Los datos provienen de diferentes universidades y de sitios de datos públicos(Huckvale & Buciuleac, 2021).

- Técnica de ML: Supervisado (Regresión, Clasificación), No-supervisado, Profundo

Se considera utilizar un modelo de aprendizaje automatizado basado en supervisión. Posiblemente de clasificación.

El algoritmo supervisado a usar es CNN dado que se han utilizado para aprender representaciones de alto nivel a partir de espectrogramas de voz. Estas redes pueden identificar patrones complejos y características relevantes que son útiles para la detección de enfermedades.

- Identificación de las variables: ¿Cuáles son las entradas y la salida?

La variable de entrada directamente es el procesamiento de los audios del dataset los cuales se procesarán obteniendo los coeficientes cepstrales en las frecuencias de Mel, extraídos de los espectrogramas los cuales encapsulan de manera eficiente las características fundamentales de la voz de los pacientes.

La salida se definirá como la etiqueta donde se especifique si el paciente se encuentra con patología o se encuentra sano.

4. Convenios

Es importante asegurar el permiso y el acceso a los datos necesarios para realizar el proyecto integrador. Si la empresa/organización solicita un convenio de privacidad, indicarlo al profesor(a) titular para que les comparta el formato. En ese caso deberán incluir en el apartado de Anexos, de esta entrega, el convenio firmado por el representante de la empresa (quien les facilita el acceso a la información) y los miembros del equipo.

- No aplica dado que los sets de entrenamiento son públicos y directamente se está trabajando en el HUB de ciencia de datos del Tecnológico de Monterrey.

5. Bibliografía

- Huckvale, M., & Buciuleac, C. (2021). Automated Detection of Voice Disorder in the Saarbrücken Voice Database: Effects of Pathology Subset and Audio Materials. *Interspeech 2021*. <https://doi.org/10.21437/interspeech.2021-1507>
- TecScience. (2022, August 8). *Modelos inteligentes aplicados a terapia de lenguaje*. TecScience. <https://tecscience.tec.mx/es/tecnologia/software-para-patologias-del-habla/>